Combining single molecule and bulk phase approaches to investigate damage detection by nucleotide excision repair proteins

Jamie Barnett

A thesis submitted to the School of Biosciences for the degree of Doctor of Philosophy in Biochemistry

University of Kent, September 2019

No part of this thesis has been submitted in support of an application for any degree or other qualification of the University of Kent, or any other University or Institution of learning

Acknowledgements

I would like to acknowledge the people who have contributed materials or data used in this thesis. The National Bioresource Project in Japan for providing the *uvrA* gene. Quentin Smith for pUC18 preparation. Dr. Alex Moores for removing the *uvrA* gene from the mScarlet vector. Dr. Nicola Don for her work on T4 Ligase. Robert Charman for his contributions to some calculations in chapter 3. Laurens Heling for helping with the final formatting of this thesis. Luke Springall for his work on the p44 protein and single molecule assays referred to in this thesis. I would especially like to thank our collaborators Dr. Jochen Kuper and Professor Caroline Kisker for their purification of TFIIH subunits and contribution of data to chapter 4, as well as many discussions about the data.

On a more personal note I would like to thank people whom without the last four years would have been a struggle. My lab mates Luke, Alessio, Jingyu, Laurens and Rob. My good friends Joe, Karl, Claudia, and Nathan for the frequent non-scholastic pub visits. But most of all to my supervisor Neil, besides it “being his job”, he has provided constant motivation and support throughout my PhD. Without him I doubt I would have been inspired to remain in academia.

Table of contents

Chapter 1 Introduction

Chapter 2 Experimental methods

Chapter 3 Understanding the coupling between DNA damage detection and UvrA's ATPase using bulk biochemical and single molecule kinetics

Chapter 4 The TFIIH components p44/p62 act as a damage sensor during nucleotide excision repair

Chapter 5 Overall discussion

List of figures & tables

List of Tables

Table of abbreviations

6-4PPs	6-4 pyrimidine-pyrimidone photoproducts
ATP	Adenosine triphosphate
ATPγS	Adenosine 5′-[γ-thio]triphosphate
CPD	Cyclobutane pyrimidine dimers
Cryo-EM	Cryogenic electron microscopy
CS	Cockayne syndrome
Fe-S	Iron Sulphur cluster
GGR	Global-genome repair
Mant ATP	(2'-(or-3')-*O*-(*N*-Methylanthraniloyl) Adenosine 5'-Triphosphate
MSD	Mean squared displacement
NER	Nucleotide excision repair
OAF	Oblique angle fluorescence
ROS	Reactive oxygen species
TCR	Transcription-coupled repair
TFIIH	Transcription factor II H
TTD	Trichothiodystrophy
UV	Ultraviolet
vWA	von Willebrand factor-A like domain
XP	Xeroderma pigmentosum

Abstract

DNA repair is crucial for the survival of all organisms. However, cellular DNA repair pathways are complex and many of the molecular details remain elusive. Understanding how DNA is repaired is paramount, as aberrant repair is implicated in numerous diseases that lead to premature ageing and increased cancer risk. Nucleotide excision repair (NER) is found across all kingdoms of life and protects the genome following exposure to diverse types of DNA damage, including UV light and chemotherapeutics. In this thesis, a combination of single molecule fluorescence and bulk biochemical techniques are used to investigate how NER proteins find damage in the genome. Individual molecules of DNA repair proteins are visualised interacting with the DNA using fluorescent tags. We show both DNA damage and nucleotide cofactors affect the attached lifetime of a prokaryotic DNA damage detecting protein (UvrA) to DNA. Based on these data and supporting biochemical ATPase experiments we devise a hypothesis where UvrA sequentially hydrolyses nucleotide in each ATPase site and reconcile this with UvrA's role in DNA repair. We also investigate two recent cryo-EM structures of the major transcription factor TFIIH. Two previously identified 'structural' subunits (p44/p62) are shown to make extensive contacts with a damage verifying helicase, XPD. We show these subunits stimulate XPD's ATPase on damaged DNA. In addition, we show p44/p62 form an independent complex that is capable of binding to DNA and discriminating damage. This suggests p44/p62 is a novel DNA-binding entity in the TFIIH complex that may have a role in the detection or verification of DNA damage. The results in this thesis have relevance for understanding cancer and ageing, as well as answering controversial questions in the field and contributing to our understanding of how DNA is repaired.

Chapter 1

Introduction

1.1 Genome damage and the importance of maintenance

A spectre haunts our cells, a multitude of agents that can impair vital cellular processes by damaging DNA. Human survival relies on accurate replication and transcription of DNA for cell division. When DNA is damaged, rapid repair is necessary to ensure genome integrity, even a single double-stranded DNA break can be lethal to a cell (1). Fortunately, complex networks of DNA repair pathways have evolved to deal with the variety of genomic damage that can lead to cell death. The human genome acts as a cellular instruction manual that defines who we are and differentiates us from plants and other animals. The information stored in our genome is an architectural blueprint of the cell and the internal milieu within it. Maintaining DNA is paramount for successful cell division that produces healthy daughter cells needed to replace old and damaged cells. Damage to our genome can prevent cell division, cause daughter cells to become diseased, or cause cell death. DNA damage is caused by chemicals interacting with DNA and modifying the native structure. These modifications can disrupt the base pairing of DNA, affect the structure of the helix, or cause breaks in the DNA. All of these modifications can lead to mutations that over time accumulate and contribute to aging and diseases such as cancers (2). However, not all mutations are detrimental to the cell, with some providing an enhanced ability to adapt to a new environment or otherwise being advantageous to the organism. There is a certain level of intrinsic genetic instability caused by errors in the DNA replication machinery, metabolism, and exposure to mutagens that is necessary for evolution. This intrinsic genetic instability is increased by chemicals that damage the DNA and cause cellular damage that can result in disease (3).

1.2 Types of DNA damage and repair

Although the structure of DNA is remarkably stable, both exogenous and endogenous compounds can interact with and damage the structure of DNA (**figure 1.1**). If DNA damage is not repaired it contributes to mechanisms of aging and disease, as well as the development and progression of cancers (2). It has been imperative for our survival to evolve redundant mechanisms to cope with a wide array of DNA damage. Key cellular processes such as DNA replication and transcription require protein-DNA interactions, DNA damage can disrupt these processes, or lead to mutations that alter the genetic code and affect the function of the gene product. In response to encountering DNA damage, proteins such as ATM + ATR (in eukaryotic cells) rapidly inhibit cell cycle progression and DNA replication by phosphorylating enzymes until the DNA has been repaired (4, 5).

1.2.1 DNA damage

In general, DNA repair is carried out using three mechanisms; direct reversal without disrupting the helix, excision of damage from DNA, or through recombination with the sister chromatid (6-8). The bulk of this thesis is focused on the nucleotide excision repair pathway and its role in repairing ultraviolet (UV) radiation induced DNA damage. UV radiation is particularly important as the majority of life forms are exposed to sunlight during their life cycle. The specifics of both prokaryotic and eukaryotic NER pathways are discussed in depth later in this chapter (section 1.5 and 1.6 respectively). The specifics of other types of DNA damage and repair pathways are beyond the scope of this thesis, but the broad picture of DNA damage and repair will be discussed, with several major types of DNA damage shown in **figure 1.1** and described below;

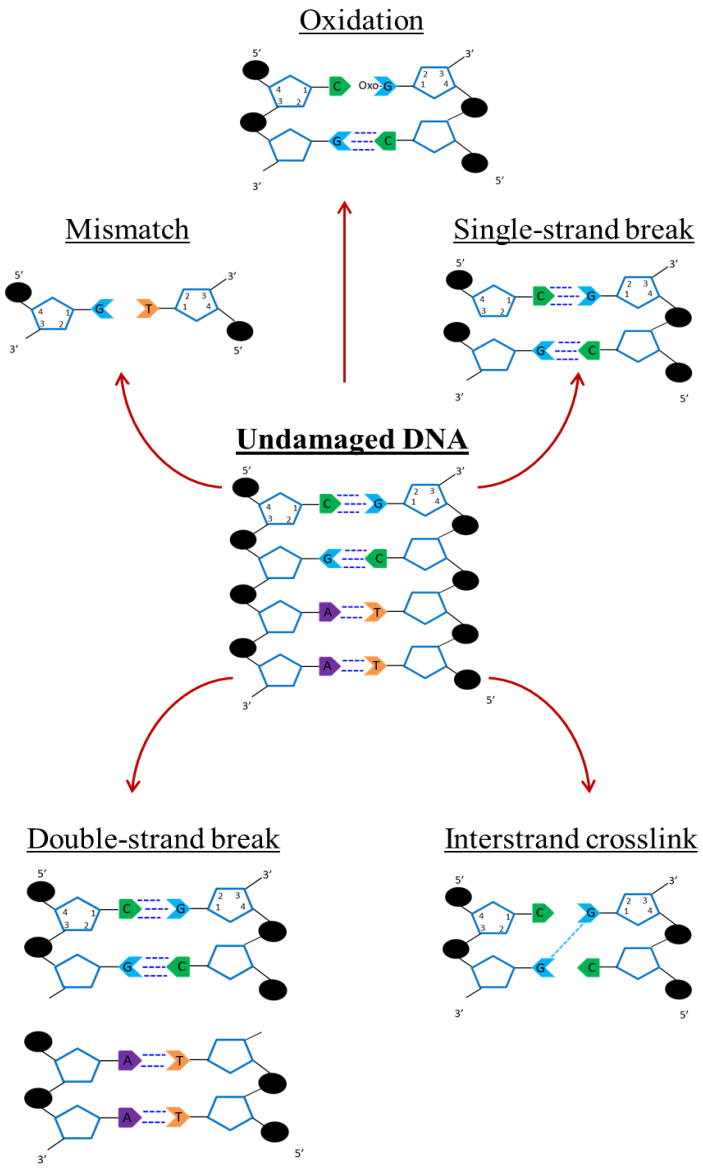


Figure 1.1 Major structural changes caused by various types of DNA damage. The normal base pairing and phosphate backbone of DNA are shown, along with the disrupted structures that occur with certain types of damage.

1.2.2 Base modifications

DNA bases can be modified by many cellular compounds or exposure to chemicals such as those found in cigarette smoke (9). A major source of endogenous damage results from reactive oxygen species (ROS) attacking the DNA backbone to release the nucleotide base from DNA and generate an abasic site (10). Abasic sites also occur spontaneously, and are often produced as a by-product of other DNA repair pathways (2). Up to 200,000 abasic sites are estimated to occur per cell, per day in mammals (10). Base modifications cause cell death and mutagenesis by causing DNA polymerases to make mistakes when replicating DNA (7, 10, 11). An example of a frequent base modification is 7,8-dihydro-8-oxo-guanine (8-OxoG), resulting from ROS oxidising a guanine base (12). These lesions are repaired by base-excision repair (BER) (13, 14), however, if left unrepaired 8-OxoG can base pair with either Cytosine or Adenine and cause a mutation in the DNA (15).

1.2.3 Insertions & deletions / mismatches

During DNA replication, DNA polymerase can insert an incorrect base to form a mismatched base pair. If mismatches are not identified and repaired they can be incorporated into the DNA (16, 17). DNA polymerase can also slip on DNA, resulting in either extra bases being inserted, or skipping several bases leading to a deletion (18). Insertions and deletions are recognised and repaired similarly because they both lead to extra bases on one strand. If these damages occur in a coding region of the genome, the impact on a polypeptide could be catastrophic, for example, 70% of cystic fibrosis cases are caused by a three base pair deletion in the CFTR gene that causes the deletion of a single amino acid (19).

1.2.4 Single-stranded breaks

A single-stranded break (SSB) occurs when the phosphodiester backbone is broken in one of the helical chains of DNA. Tens of thousands of SSBs arise every day caused by ROS, radiation, and as a by-product of DNA repair. A single-strand break is also able to develop into a more toxic double-strand break if not repaired (20).

1.2.5 Double-stranded breaks

Double-stranded breaks (DSBs) occur when there is a complete break in the DNA. Up to 9 spontaneous DSBs occur per cell, per day (7) and are the most toxic form of DNA damage (21). DSBs can result from SSBs that are not repaired properly, replication fork breakage, or by exposure to radiation. The toxicity of DSBs results from their ability to lose genetic information encoded around the break, and the ability of entire chromosomes to break and fuse with another segment – wreaking havoc with gene expression.

1.2.6 Crosslinks

DNA-protein and DNA-DNA crosslinks can be caused by carcinogens as well as chemotherapeutic agents such as cisplatin and mitomycin C (7, 22). DNA-protein crosslinks are able to block enzymes from interacting and moving along the DNA, inhibiting both replication and transcription. DNA-DNA crosslinks can occur either within the same DNA strand (intra-strand crosslink) or between the two strands (inter-strand crosslinks). These DNA-DNA crosslinks covalently link the strands together, preventing unwinding by helicases that is required for transcription and DNA replication. Even a single DNA-DNA crosslink can kill a cell (23, 24).

1.2.7 DNA repair

These damages generally have one major route for repair, but frequently cross-talk occurs between various DNA repair pathways, and this may be a redundancy mechanism to provide a dynamic, adaptable response. Below, damage repair pathways are described with examples from both prokaryotic and eukaryotic systems, where the mechanism is often conserved, but the eukaryotic pathway generally involves more proteins.

1.2.8 Direct reversal

Some types of DNA damage can be directly reversed, an example of this in eukaryotes is O^6-alkylguanine DNA alkyltransferase (AGT). AGT removes the methyl group from the O^6 position on methylation-damaged guanine bases. The addition of the aberrant methyl group forms a base pair with thymine rather than cytosine, causing mutations. Cells with lower levels of AGT are more sensitive to methylating agents and experience a higher risk of carcinogenesis (25).

1.2.9 Base excision repair

Modified DNA bases are the most common form of damage and are repaired by BER; BER needs to be able to respond to a large number of damage sites and rapidly repair them to prevent mutations in the genome (26). In eukaryotes, damaged bases such as 8-OxoG are cleaved by a glycosylase enzyme to produce an abasic site (27), an enzyme called APE then cuts the DNA backbone at the abasic site to generate a SSB with a free 3'OH group. Polymerase β loads at the hydroxyl group and resynthesizes the correct nucleotides (28). DNA ligase then seals the nick at 5' phosphate, restoring the DNA structure (14, 29). Additionally, BER can remove longer patches of damaged bases using an exonuclease to digest the damaged strand several nucleotides in before allowing polymerase to resynthesize the DNA (30).

1.2.10 Mismatch repair

Mismatch repair (MMR) recognises insertions, deletions, and substitutions that escape DNA polymerases proofreading capability (31). In *Escherichia coli (E. coli),* MutS can recognise mispaired bases up to 4 nucleotides long with high affinity (32, 33). The mispaired bases are then excised from the DNA backbone to generate a SSB. If the mismatch is longer, an exonuclease is recruited to resect the DNA further before allowing polymerases to load and resynthesize the correct sequence (34, 35). Because mismatches can lead to mutations if incorporated into the genome, MMR inactivation is highly mutagenic and has been associated with several cancers (36).

1.2.11 Single-stranded break repair

SSBs in eukaryotic cells are immediately detected by RPA and PAR. RPA binds to the exposed single-stranded DNA and protects it from further degradation and prevents the formation of a double-strand break (37). PAR ribosylates itself and other DNA repair proteins with chains of poly ADP-ribose and recruits XRCC1 - a scaffolding protein that interacts with other DNA repair proteins, to the ends of the break (38). Frequently, the ends of the DNA are either damaged or modified and therefore are not able to be ligated until restored to a 3' OH and 5' phosphate. Several nucleases are involved in removing these structures and restoring the correct chemical groups (39). The 3'OH enables DNA polymerases to load on the DNA and resynthesize the removed bases, and then ligation

follows to seal the break and restore DNA (40). SSBs in prokaryotes are repaired via homologous recombination by the RecBCD proteins (41).

1.2.12 Double-strand breaks: NHEJ & HR

As DSBs are the most toxic form of damage, several repair methods exist in both prokaryotic and eukaryotic cells;

Non-homologous end joining (NHEJ) - In eukaryotes, a ring shaped protein, Ku binds each end of the DSB and recruits DNA-PK (42). Interactions between Ku:DNA:DNA-PK bridge the broken ends and allow ligation through homology between each end of the DSB (43). However, NHEJ can cause a loss of genetic information by resecting both DNA ends before re-joining (44).

Homologous recombination (HR) - HR uses sister chromatid homology to repair the break and is therefore error-free (44-46). In HR different exonuclease complexes digest around the break, the single-stranded DNA is then bound by RAD51 and RPA (RecA and single-strand DNA-binding protein in *E. coli*) to form a nucleoprotein filament (44, 45, 47-49). Fascinatingly, this filament then uses a combination of three-dimensional and one-dimensional searching to find the homologous undamaged sister chromatid and uses the complementary DNA sequence to repair the damaged DNA (50, 51).

1.2.13 ICL repair

Intra-strand crosslinks are repaired by nucleotide excision repair (NER) to cause two single-strand breaks either side of the crosslinked in a similar fashion to BER and MMR (52). Inter-strand crosslinks (ICLs hereafter) are more complex; in eukaryotes, quiescent cells repair ICLs using sequential NER incisions, the first creates a substrate for loading translesion polymerases that bypass the crosslink (53, 54). A second round of NER then removes the crosslinked DNA, and allows resynthesis of the damaged strand (24, 53, 55). However, during S phase DSBs are observed in response to ICL causing agents, suggesting ICLs are repaired via a DSB intermediate (56, 57). Many proteins in the Fanconi anaemia pathway have been implicated in ICL repair in S phase, but this is still not well understood, and is not discussed here (58-60). Interestingly, in prokaryotes, NER is the primary method for removing crosslinks from DNA, but may also involve proteins from other repair pathways (61). For an excellent review see Hashimoto et al 2016 (62).

1.3 UV damage products depend on the wavelength of light

UV radiation has long been linked with cell death by causing chemical changes in DNA (63, 64). Since then it has been shown UV radiation causes DNA damage in two main forms; cyclobutane pyrimidine dimers (CPDs) and 6-4 pyrimidine-pyrimidone photoproducts ((6-4) PPs) (shown in **figure 1.2**) (2, 65). The type of DNA damage formed depends on the wavelength of light. UV light is conventionally separated into UVA (320-400 nm), UVB (280-320 nm), and UVC (200-280 nm). Atmospheric ozone blocks out most of the UVB and UVC radiation, with ~95% of UVA and ~5-10% of UVB radiation reaching the surface of earth (64, 66). UVB and UVC light are absorbed by DNA and react to cause CPDs and (6-4) PPs. It is estimated around 70-80% of UV-lesions are CPDs, and 20-30% are (6-4) PPs following exposure to sunlight (67, 68), however, (6-4) PPs are produced more efficiently at lower wavelengths of light that are mostly blocked by ozone (69). Although UVA light is not efficiently absorbed by DNA, it reacts with other cellular macromolecules that can indirectly damage DNA by generating ROS (66, 69, 70). (6-4) PPs are produced from a covalent bond between adjacent pyrimidine bases (**figure 1.2**), these lesions are more mutagenic than CPDs as they can cause transition mutations in the DNA (67). CPDs are formed by covalent bonding between two adjacent pyrimidine bases to form a cyclic butane ring (**figure 1.2**) (71). Both (6-4) PPs and CPDs are lethal to cells unless repaired. This is highlighted by the fact that when the NER pathway that repairs these lesions is disrupted it causes diseases in humans (discussed in section 1.7) (72).

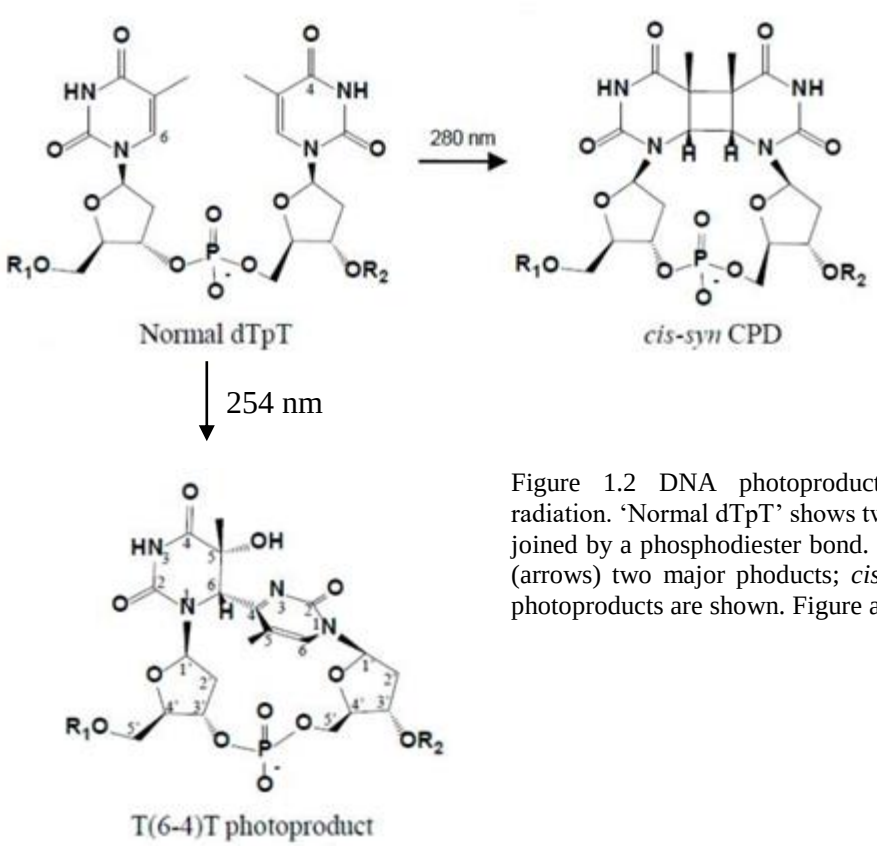


Figure 1.2 DNA photoproducts formed by UV radiation. 'Normal dTpT' shows two adjacent thymines joined by a phosphodiester bond. Upon UV irradiation (arrows) two major phoducts; *cis-syn* CPD and (6-4) photoproducts are shown. Figure adapted from (73).

1.4 Nucleotide excision repair

NER is evolutionarily conserved across all domains of life. Simultaneous observations in 1964 showed the removal of UV damage from DNA, and repair of the DNA backbone (74-77). Since then it has been found that NER primarily repairs UV damage such as CPDs and (6-4) PPs that form bulky adducts on DNA (78, 79). NER is a promiscuous pathway that responds to almost all types of damage, this can be seen as a safe option to the cell. Instead of utilising many enzymes to deal with different lesions, excision of damage containing DNA provides a system able to react and remove a vast array of damage in a relatively non-specific manner. NER involves recognition of DNA damage through either transcription-coupled repair (TCR) or through global genome scanning. The large variety of damage types are thought to be identified by disrupted base pairing around the lesion. The more a lesion disrupts the DNA duplex, the faster it is repaired (80). (6-4) PPs are more distorting than CPD lesions, enhancing their recognition by NER and leading to a nine-fold increased rate of removal compared to CPDs (81-83). Various chemical DNA damage can also distort the structure of the DNA helix, especially bulky chemical groups that covalently bind to DNA and disrupt base pairing such as those caused by polyaromatic hydrocarbons, and benzo[a]pyrene (8).

1.4.1 Transcription-coupled nucleotide excision repair

TCR was first proposed when groups observed CPDs in actively transcribed genes were repaired faster than the rest of the genome (82, 84). In *E. coli,* lesions in the transcribed strand stall RNA polymerase as it transcribes DNA, Mfd is then recruited to displace RNA polymerase from DNA, and UvrA is subsequently recruited to initiate DNA repair (85-87). Interestingly, damage on the non-transcribed strand of a gene does not trigger TCR and is repaired at the same rate as the rest of the genome, showing the strand specificity of TCR (82, 84, 85). TCR is conserved and occurs in eukaryotes where the functional equivalent of Mfd is CSA/B (88).

1.5 Global-Genome repair in bacteria

Lesions are detected through TCR or global-genome repair (GGR) and then the pathways merge together. GGR begins with UvrA detecting a lesion and is shown in **figure 1.3**. Once damage is detected, DNA unwinding and verification takes place to ensure the lesion is a substrate for NER – an important step to prevent incision on undamaged DNA (89, 90). The DNA is then excised with sequential endonuclease reactions to liberate ~12 nucleotides of DNA containing the lesion (25-30 nucleotides in eukaryotes (91, 92)) (78). After incision has taken place, the incision complex and oligonucleotide are removed from DNA (93). Polymerases are then able to load on the DNA to carry out repair resynthesis with ligation of the backbone completing repair (94). *E. coli* proteins are mechanistically homologous to eukaryotic NER proteins, however the bacterial counterparts are easier to culture, purify, and require only 4 proteins to successfully excise DNA (NER in mammalian cells requires ~ 30 proteins) (78, 95-100); providing a good model system to understand individual protein function in the context of a more complex system. In chapter 3 of this thesis purified bacterial proteins are used to understand the ATPase activity of UvrA and how it contributes to damage detection in NER.

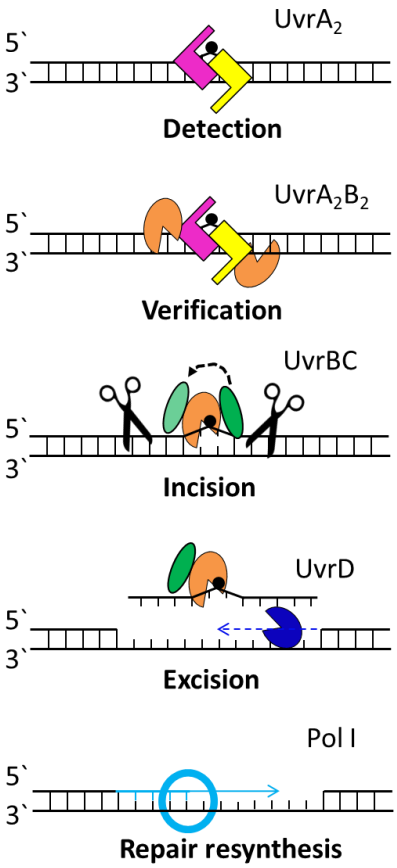


Figure 1.3 Nucleotide excision repair in bacteria.

Initial damage detection is carried out by UvrA (GGR) or by the transcriptional machinery (TCR)

UvrB is then recruited to verify the damage and initiate incision of the DNA

UvrA is removed from the UvrB-DNA complex to allow UvrC to incise the DNA either side of the lesion

The UvrD helicase then removes the damage containing oligonucleotide and proteins from the DNA

Pol I and DNA ligase then complete repair resynthesis of the DNA

1.5.1 Damage recognition in nucleotide excision repair

UvrA is responsible for searching the genome for DNA damage in *E. coli* (98, 101), it does this as a $UvrA_2$ homodimer (hereafter UvrA for clarity) (90, 101, 102). UvrA has an increased affinity for UV damaged DNA (~2-4 fold) (101, 103-105), and an even higher affinity for single-stranded DNA (97, 106), suggesting damage recognition is based on recognising disrupted base pairing. This explains the diverse array of lesions recognised and repaired by NER. UvrA's affinity for damage substrates is dependent on the nucleotide bound to each ATPase site in UvrA (discussed in section 3.3). Upon binding to DNA, UvrA bends the helix to cause local unwinding (104, 105). Previous work has shown that DNA bending is important for protein:DNA interactions (107), likely affecting the thermodynamics of complex stability, and enhancing the recognition of damaged DNA among structurally similar undamaged DNA (104, 105). Once UvrA recognises damage, it is able to recruit and load UvrB onto the DNA to verify the presence of damage (96). While UvrA binds DNA and searches for damage as a homodimer, it can also search for damage in complex with UvrB (102-104, 108, 109). There are discrepancies over the complex stoichiometry; either $UvrA_2B_2$ (98, 109-112) or $UvrA_2B_1$ (90, 110). Single molecule fluorescence experiments showed UvrAB forms in solution and binds to DNA where it slides along the backbone, whereas UvrA alone uses a three-dimensional DNA search (109). The collapse in dimensionality from three- to one- dimension caused by the addition of UvrB to UvrA increases the number of base pairs able to be interrogated, and therefore increases the speed of finding damage (109).

Alongside UvrA and UvrAB, the UvrBC complex has been shown both *in vitro* and *in vivo* to be capable of excising damaged DNA (90, 113, 114). This recent observation provides a UvrA-independent mechanism to detect and repair DNA damage. This pathway may be important when a large number of lesions exist in a cell, where additional damage recognition proteins such as Mfd are able to recruit the UvrBC complex to carry out incision. However, the UvrBC complex is also capable of binding and diffusing on DNA to recognise damage (113, 115). UvrC is an endonuclease that incises the damaged DNA strand either side of the lesion (78, 116-118), however, as few copies of UvrC exist per cell (~10 (117)) it is likely found in complex with UvrB *in vivo* (~ 150 copies (95)) which may prevent unwarranted DNA incision (113, 117). Springall et al visualised a mixture of UvrA, UvrB, and UvrC binding to DNA as a heterogenous population (**table 1.1**). This diversity is likely

more controlled *in vivo*, where regulation of protein concentration defines the equilibrium of different complexes. However, it is likely that UvrA reflects the major damage sensor, distinct from being in complex with UvrB. UvrC is capable of binding DNA alone, but considering estimated copy numbers, probably exists mainly bound to UvrB. Both UvrAB and UvrBC have enhanced affinity for damaged DNA and likely reflect damage verification complexes but may also search the genome for damage (103, 108).

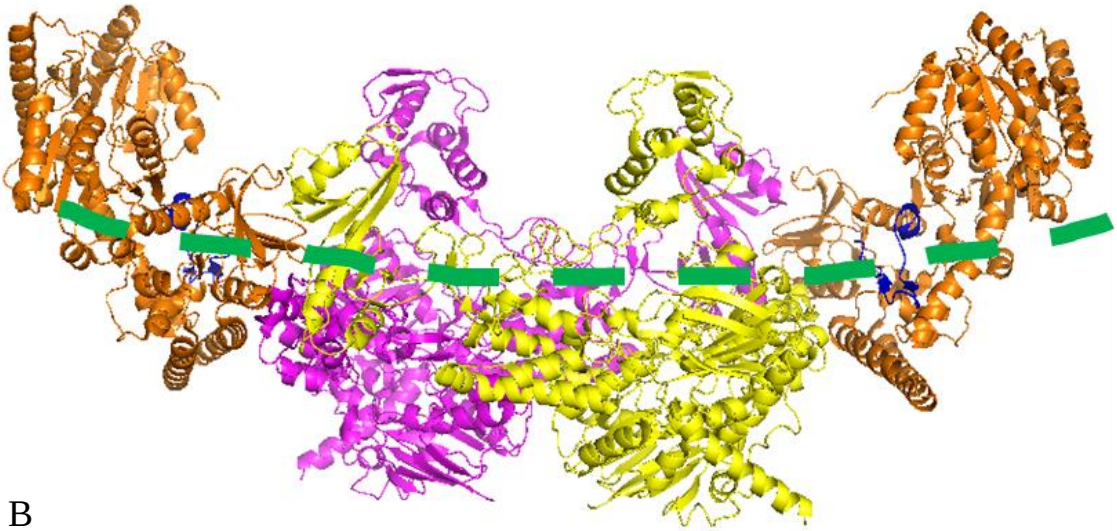
Protein complex	% Observed
UvrA	61.5 %
UvrC	15.8 %
UvrBC or UvrAB *	19 %
UvrABC***	1.6 %
UvrAC**	2.1 %

Table 1.1 Heterogenous populations of UvrABC repair proteins. Springall et al imaged the various complexes that form on DNA using different coloured fluorescent QDOTs (108). *UvrB alone does not bind DNA so was counted as UvrBC or UvrAB. UvrAB and UvrBC complexes were imaged as dual colour complexes and these data are pooled. **UvrAC complexes were only observed with UvrB present. ***Triple-coloured complexes of all three Uvr proteins.

1.5.2 UvrA's interaction with UvrB

In a crystal structure of the UvrAB complex, UvrB binds each end of the UvrA homodimer, forming a BAAB sandwich (**figure 1.4**) (112). This UvrAB complex has been suggested to form in solution and find damage (102, 104, 109), or form on DNA after UvrA finds damage and loads UvrB (90, 119, 120). The structure in **figure 1.4** shows a DNA binding channel down the centre of the protein that is estimated to be capable of binding around 45 base pairs of DNA. This location of UvrB molecules is some distance from the lesion (assuming it is bound in the centre of UvrA). This means that once UvrA is displaced from the UvrAB-DNA complex in an ATP-dependent reaction (90), UvrB has to translocate towards the lesion on the damaged strand of DNA. It is known that only one UvrB molecule exists in the incision complex when UvrC cuts the DNA, suggesting the higher affinity for DNA damage and disrupted base pairing around a lesion encourages formation of the UvrB-damage complex rather than an undamaged bound complex, causing one UvrB molecule to leave the DNA (102, 110, 112).

A



B

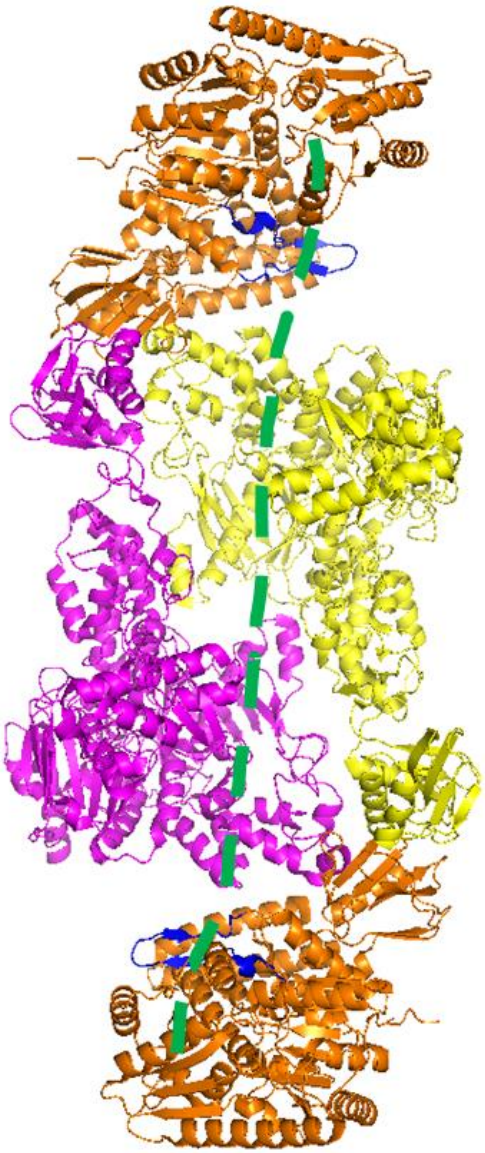


Figure 1.4 $UvrA_2B_2$ complex crystal structure. 3UWX crystal structure of dimeric UvrA (coloured magenta and yellow, centre) sandwiched by two UvrB monomers (orange) (112). The beta-hairpin of UvrB that interacts with the DNA is shown in blue. A) The side and B) the front on view show the DNA binding site forms a deep groove that arcs across the full structure of the protein. A rough guide of where the DNA would bind is shown with a dashed green line. Structure created in PyMol (121)

1.5.3 UvrB's beta hairpin verifies the presence of DNA damage

UvrB acts as the hub of prokaryotic NER, interacting with every protein in the pathway; UvrA, UvrC, UvrD, and Polymerase I (122, 123). Whilst these interactions are key for NER, UvrB's main role is to verify the presence of DNA damage. UvrB requires ATP hydrolysis and UvrA to form the pre-incision complex with damaged DNA and carry out NER (96, 105, 124). Helicase motifs were identified in UvrB during a systematic analysis of nucleotide-binding proteins (125), these motifs bear strong similarity to the eukaryotic counterparts RAD3 (XPD) and RAD25 (XPB) (examined in chapter 4) that are involved in unwinding the DNA and verifying the lesion in eukaryotic NER. This discovery was closely followed by the crystal structure of *Bacillus caldotenax* UvrB (**figure 1.5**). In this structure the ATP binding site and helicase motifs were observed, as well as residues involved in coupling these activities together (126). However, UvrB is not a true helicase. The UvrAB complex is capable of displacing oligonucleotides, but does not processively unwind the DNA (89, 127-129); instead it induces local conformational changes that lead to unwinding (128).

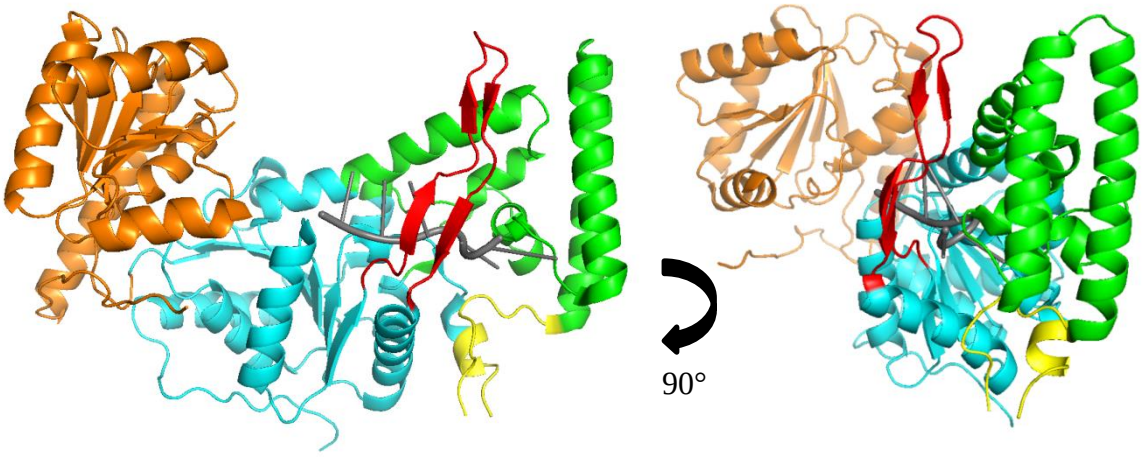


Figure 1.5 The structure of UvrB bound to DNA. *Bacillus caldotenax* UvrB structure colour coded by domain; Domain 1a cyan, domain 1b green, domain 2 yellow, domain 3 orange, and the beta hairpin in red. This structure shows the beta hairpin pinning the single-stranded DNA (grey) against domain 1b as suggested in the padlock model (130). Structure 2FDC from Truglio et al (100) with domains assigned as in Theis et al (130). Structure created in PyMol.

UvrB is able to reach a lesion by diffusing in complex with UvrA, or alternatively, UvrB can be directly loaded onto a lesion by UvrA (109, 113). Upon binding a lesion, conformational changes lead to the release of UvrA and formation of a UvrB-damage complex. Once UvrB is bound to a lesion it must then verify the presence of genuine damage; this involves further bending and unwinding of the DNA (89, 131, 132). Comparing UvrB's crystal structure to closely related counterparts PcrA, NS3, and Rep showed that if UvrB were to bind DNA in the same orientation, the translocating strand would be obscured by a beta-hairpin structure found in UvrB (133). This beta hairpin is critical for UvrB's ability to discriminate DNA damage, with extensive mutational studies identifying a stretch of residues at the tip of this hairpin (**figure 1.5**) being vital for DNA incision, damage binding and 'helicase' activity (127). Furthermore, the beta hairpin is key for translocation, and formation of the pre-incision complex. The padlock model proposed by Van Houten and colleagues describes how the beta hairpin holds the non-translocating strand against domain 1b (**figure 1.5**) (126, 134) while the translocating strand passes in front of the beta hairpin where the aromatic tyrosine residues flip the scanned nucleotides out of the helix, probing for helix distortions and detecting DNA damage (135, 136). Upon binding to a lesion UvrB's ATPase activity is stimulated, causing conformational change that cause UvrA to release from the UvrAB complex (116, 130). This interaction is critical, deletion of the UvrA binding domain from UvrB (domain 2) prevents formation of the pre-incision complex, 'helicase activity', and ATPase activity leading to DNA incision (137, 138). Once UvrA is released, the UvrB-damage complex forms a substrate for UvrC to bind (130).

1.5.4 UvrC performs both incisions on DNA

UvrC is a dual endonuclease that excises ~12 nucleotides of DNA around the UvrB-DNA complex with sequential incisions 5' and 3' of the lesion on the damaged strand (78, 116, 131). Interactions between UvrC and UvrB likely trigger both incision reactions (139, 140), although this is still unclear. It has also been suggested that the 3' incision causes further unwinding of the DNA, that then triggers the 5' incision on DNA (89). Furthermore, it has been observed that UvrBC can diffuse on DNA to find damage (108, 113). It is thought that spurious DNA incision by UvrBC is prevented by UvrB's ability to verify damage, and this likely represents the population of UvrC *in vivo* (108).

1.5.5 Repair resynthesis and ligation

When cells are irradiated, potentially thousands of damage sites are formed on DNA. Accurate estimates of cellular copy numbers of UvrA, UvrB, and UvrC are lacking, with early biochemical reports estimating ~10 molecules of UvrC (117), ~20 molecules of UvrA (95), and ~150 molecules of UvrB (95), while more recent *in vivo* observations suggest higher concentrations of UvrA (~85 molecules per cell) (119). However, both sets of numbers suggest that efficient turnover of repair machinery is required to deal with damage in a timely manner. After excision, UvrBC form a stable complex with damage containing DNA (78, 123, 141). The combination of DNA Polymerase I (Pol I) and UvrD together are able to displace UvrB and UvrC as well as the excised oligonucleotide from DNA (141-144). The removal of UvrBC from the DNA produces a substrate for Pol I to bind, allowing resynthesis of the DNA (145). DNA ligase can then ligate the remaining nick to reseal the DNA and complete repair (141, 142).

1.6 Eukaryotic NER

NER is evolutionarily conserved across all domains of life. All NER pathways involve DNA damage detection, verification of the lesion and then excision by structure specific endonucleases (146). However, the eukaryotic pathway is more complex, and requires over 30 proteins (147, 148). A simplistic overview of eukaryotic NER is shown in **figure 1.6**. Eukaryotic NER, like prokaryotic can be stimulated by either transcriptional stalling (transcription-coupled repair (TCR)), or by global-genome scanning repair (GGR) – the latter is discussed here.

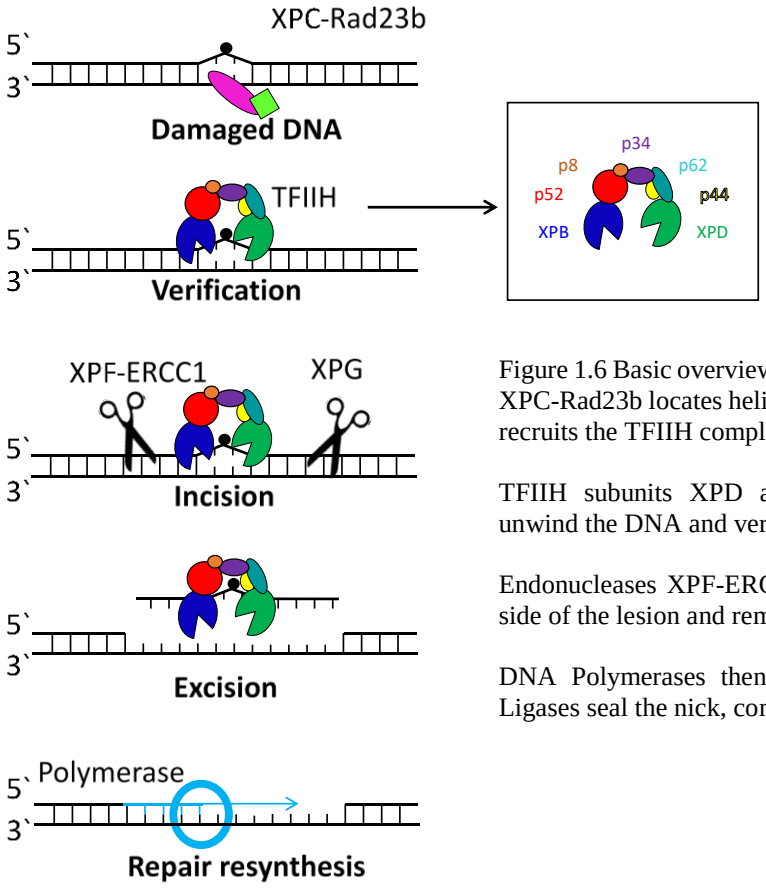


Figure 1.6 Basic overview of events in eukaryotic NER. XPC-Rad23b locates helix distorting DNA damage and recruits the TFIIH complex.

TFIIH subunits XPD and XPB are helicases that unwind the DNA and verify the presence of damage.

Endonucleases XPF-ERCC1 and XPG then cut either side of the lesion and remove the damaged DNA.

DNA Polymerases then resynthesize the DNA and Ligases seal the nick, completing repair.

1.6.1 In eukaryotes TFIIH carries out NER

The components involved in eukaryotic NER were initially discovered by observations in yeast showing that mutating certain genes affected cell viability and increased sensitivity to UV irradiation. RAD1 (XPF), RAD2 (XPG), RAD3 (XPD), RAD4 (XPC), and RAD25 (XPB) among others were found to be epistatically linked and confer sensitivity to UV irradiation (149-153). Then, in 1991 TFIIH (Transcription factor II H) was discovered as a major transcription factor key for cell survival (154). Additional studies then showed the UV sensitive genes in yeast encoded subunits of TFIIH and established a link between TFIIH and DNA repair (150, 155, 156). It is now known that TFIIH has a major role in NER, and defects in the XP genes such as XPB and XPD disrupt DNA repair and lead to diseases discussed later (section 1.7) (157-159). Here TFIIH's role in DNA repair is discussed, for an excellent review of TFIIH's role in transcription see Zurita and Merino (160).

1.6.2 XPC detects DNA damage

Damage detection in GGR is carried out by the XPC-hHR23B complex (XPC hereafter (**figure 1.6**)) (83, 161-163). NER responds to a broad array of DNA lesions that requires a dynamic and broad detection capability from XPC (78, 164). *In vivo*, XPC is assisted by UV-DDB (162, 165-167). UV-DDB is able to bind lesions located in chromatin compacted DNA, whereas XPC is not (168). UV-DDB binding to damage is thought to alter the structure of the chromatin to allow the NER machinery to access the lesion, and repair to take place (169, 170). Interestingly, a crystal structure of XPC bound to DNA damage (PDB ID: 2QSG) shows it does not directly interact with the lesion but inserts beta-hairpin structures into the DNA to detect disruption to the duplex caused by damage; similar to UvrB (115, 171, 172). When XPC interacts with the DNA it bends the duplex ~45° regardless of damage presence. This bending of the DNA may aid damage detection as it is thought damaged DNA forms a more thermodynamically stable complex with XPC than undamaged DNA, providing a better NER substrate (173, 174). Studies using the yeast homologue of XPC (Rad4-Rad23) demonstrated that to find damage, Rad4-Rad23 slides along the DNA (175). However, recently it has was shown human XPC-RAD23B locates lesions by hopping along the DNA, therefore avoiding obstacles (176). However, both studies show once disrupted base pairing is located, XPC switches to constrained DNA diffusion – restricted movement around the destabilised DNA that may be important for recruiting TFIIH (163, 175-178). XPC can also recognise many DNA sites that do not contain bona fide damage;

therefore, additional verification of damage by XPD and XPB helicases within TFIIH is key in preventing unwarranted DNA incision (171, 172, 175).

1.6.3 TFIIH is recruited to the lesion, and unwinds the DNA

In the XPC:damage complex, the N-terminal region of XPC interacts with XPB and p62, recruiting the TFIIH complex to the DNA (161, 178, 179). Once TFIIH is loaded onto the DNA the two helicases, XPD and XPB unwind around ~10-25 base pairs of DNA (178, 180, 181). XPB's ATPase is essential for this, and for remodelling of the TFIIH-DNA complex. Mutations that disrupt XPB's ATPase prevent TFIIH's recruitment to DNA and cause disease (178, 182-184). Surprisingly, the 3'-5' helicase activity of XPB is unnecessary for TFIIH's of unwinding of DNA (184). However, the 5' – 3' helicase activity of XPD is absolutely required for NER (150, 185-187). These observations have led to speculation that XPB may act as a molecular switch where ATP hydrolysis allows conformational changes in the DNA that aid XPD in unwinding the DNA from the 5' side of a lesion (188-190).

1.6.4 Damage verification by XPD

XPD is part of the superfamily 2 helicases (SF2) and is responsible for verifying the presence of DNA damage on the DNA; however, it also contains an iron sulphur cluster (4Fe-4S, Fe-S hereafter) (190, 191). Fe-S clusters can be structural and aid in protein folding as well as thermal stability but are emerging as being important for DNA binding and enzymatic activity (192-196). In the yeast homologue of XPD, mutations to the Fe-S cluster coordinating residues leads to a loss of helicase activity (191, 197-199), and damage verification (200). XPD translocates along DNA using an inchworm mechanism, where the hydrolysis of ATP provides free energy for conformational changes in the helicase domains, allowing a ratcheting movement along the DNA in distinct steps (shown in **figure 1.7)** (190, 201-204). The current model for XPD's translocation on DNA involves the arch domain and Fe-S cluster breaking the DNA duplex into single-stranded DNA, explaining the loss of helicase activity with Fe-S mutants (191, 197). The 5' ssDNA then passes through a pore formed between the arch and Fe-S domain to the two helicase domains (199, 205, 206). **Figure 1.8** shows the four domains of XPD with a 5 nucleotide stretch of single-stranded DNA bound to the HD2 domain of XPD, supporting the path of DNA described here (207), and has been further supported by a recent cryo-EM (Cryogenic electron microscopy) structure (discussed in chapter 4) (208). A DNA binding site between helicase domain 1 (HD1) and the Fe-S cluster is key for verifying damaged bases on the translocating strand (206, 209) and explains the loss of damage recognition in Fe-S mutants (200). Surprisingly, XPD can also detect damage on the non-translocating strand (200). Recognition of a lesion stalls XPD in an ATP-dependent manner (189, 200, 210), this signals the presence and location of DNA damage and is required for incision to take place (189, 200, 210).

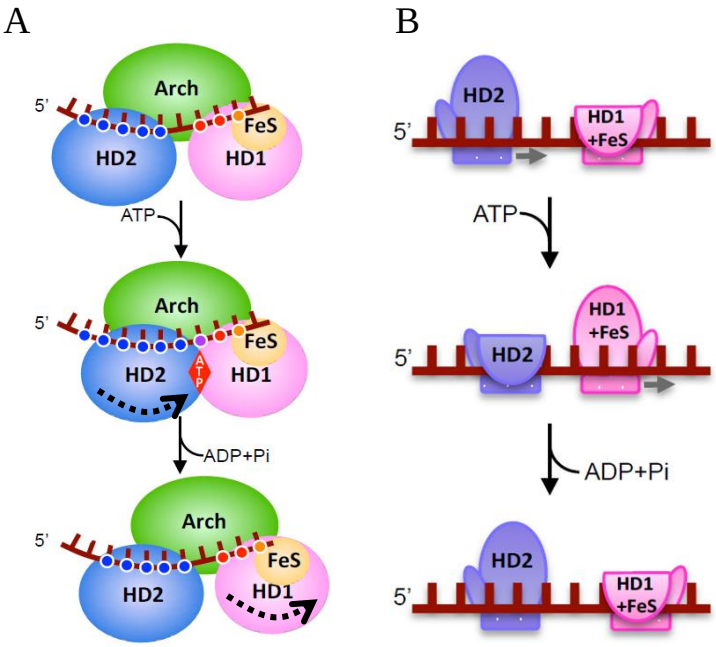


Figure 1.7 Translocation by SF2 family helicases. ATP binding induces conformational changes pushing HD2 along the DNA (towards the 3' end) whilst HD1 is tightly bound to ssDNA. Upon ATP hydrolysis the HD2 domain then becomes tightly bound to ssDNA, while the HD1 domain slides along DNA. A) Protein domain contacts with DNA during translocation for SF2 helicases. Colour coded dots represent contacts with respectively coloured domains. B) Model of HD1, HD2, and Fe-S domains moving along ssDNA. Taken from Cheng & Wigley 2018 (201).

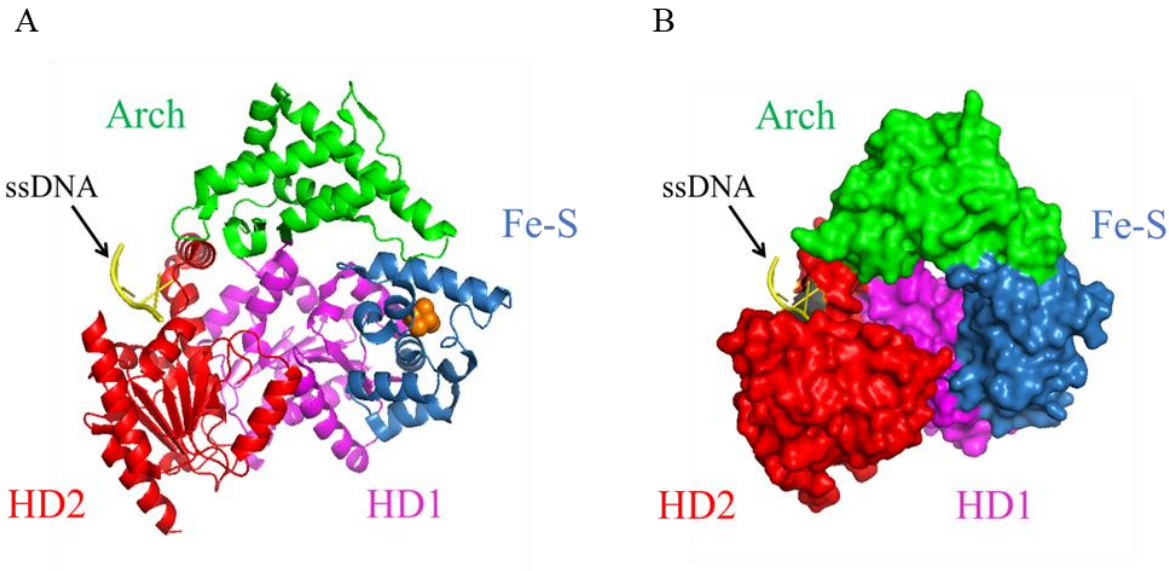


Figure 1.8 Crystal structure of *Thermoplasma acidophilum* XPD. The four domains of XPD are colour coded and labelled in A) as cartoon B) as space fill. A 5 nucleotide single-stranded DNA (ssDNA) yellow) substrate can be seen bound to the HD2 domain (red). The iron atom can also be seen (orange) in the Fe-S domain (blue). Domains were defined using the data from Kuper et al 2012. Figure was created in PyMol from PDB structure 4A15 (211).

1.6.5 Dual incisions remove the DNA damage

To remove DNA damage, concerted dual endonuclease incisions release an oligonucleotide of 24-32 bases containing the damage (212). The endonuclease reactions themselves are independent of ATP, but ATP is crucial to form the substrate for incision (213). The 3' and 5' incisions are carried out by XPG and the XPF-ERCC1 complex (XPF hereafter) respectively (164, 214, 215). These structure-specific, oppositely polar endonucleases require the correct orientation of TFIIH on DNA in order to successfully remove the damage (213, 215). This is achieved by XPD unwinding the DNA to form a bubble, RPA and XPA are then recruited to hold the DNA open - allowing access for both endonucleases to the damaged strand (164, 178, 180, 181, 216). Historically, it was thought XPG is the first endonuclease to cut and then XPF (215, 217, 218). However, catalytically dead mutants showed only the presence of XPG is necessary for XPF to cut, but not the 3' incision itself (219, 220).

1.6.6 Repair synthesis

After incision, the damaged oligonucleotide remains bound to TFIIH and is then handed off to RPA and degraded (221, 222). The 5' incision by XPF generates a free 3' hydroxyl group on the backbone which is a substrate for DNA polymerase (220). XPF also recruits proliferating cell nuclear antigen (PCNA) which acts as a processivity factor for DNA polymerases (Pol δ/ε) to carry out successful repair synthesis (37, 220, 223-225). Pol δ/ε move from the 3'OH loading site at the 5' incision site, in a 3'-5' direction, leaving a 5' phosphate that is ligated by DNA ligase I to complete the NER reaction (37, 223, 226).

1.6.7 Mutations in TFIIH disrupt protein-protein interactions

Mutations that disrupt the NER pathway can lead to disease in humans. Most disease-causing mutations in TFIIH are found in XPB, XPD, and p8 (157). However, it is likely that mutations occur throughout TFIIH but are either lethal or have little effect and therefore go undetected. Reported mutations have been shown to not actually affect the enzyme activity itself but disrupt critical protein-protein interactions that lead to a disease phenotype (157, 199, 227). Xeroderma pigmentosum (XP), Cockayne's syndrome and trichothiodystrophy are all caused by mutations in TFIIH and are discussed extensively elsewhere (227). In chapter 4 of this thesis, the helicase critical for DNA damage verification, XPD, is studied. XPD is stimulated by p44 (159, 207), and the disruption of this interaction leads to XP.

Around 87% of XP group D patients have mutations on the C-terminus of XPD where p44 binds (159, 184). These C-terminal mutations prevent p44 interacting with XPD, decreasing helicase activity, and reducing the overall viability of DNA repair (159, 207). Two clinical mutants, R616P and G675R in XPD have been shown to disrupt XPD's co-purification with TFIIH. These mutants disrupt interactions with the anchor region of p62, a key protein-protein interaction domain and have led to speculation that p62 may be involved in regulating XPD (177, 228).

1.6.8 The overall structure of TFIIH

TFIIH is highly conserved and forms a horseshoe shaped structure (**figure 1.9**). The core of human TFIIH comprises two helicases, XPD and XPB that form the base of the horseshoe, and five other subunits p44, p62, p34, p8, p52 that form the arch – shown in **figure 1.9** (177, 186, 187, 205, 226, 229-232). The CDK activating kinase complex (CAK) is attached to TFIIH via XPD, the CAK module is formed of Cdk7, MAT1, and CCNH. MAT1 has been shown to inhibit XPD's helicase activity, possibly by clamping down on the arch domain to which it is associated, preventing DNA passing through (203, 206, 233-235). Cdk7 phosphorylates the CTD domain of RNA polymerase II which is key for transcription (236-238). The CAK complex is not visible in **figure 1.9** (besides MAT1) because it is flexibly tethered to TFIIH (231, 239). However, for NER to occur, MAT1 and the CAK complex must be released (233). Recently, a high resolution cryo-EM structure of human TFIIH has been obtained, bearing strong similarities to previous work (205, 239, 240). The 3.7 Å resolution structure captures the transcriptional state of TFIIH and resolves all subunits, including the previously elusive p62 protein (239). These cryo-EM structures reveal the damage verifying helicase XPD is contacted extensively by p62 and p44; while p44 has previously been shown to stimulate XPD's ATPase activity, p62's role is unknown (159, 207). In chapter 4 the interactions between p44, p62, and XPD are investigated for their role in DNA repair.

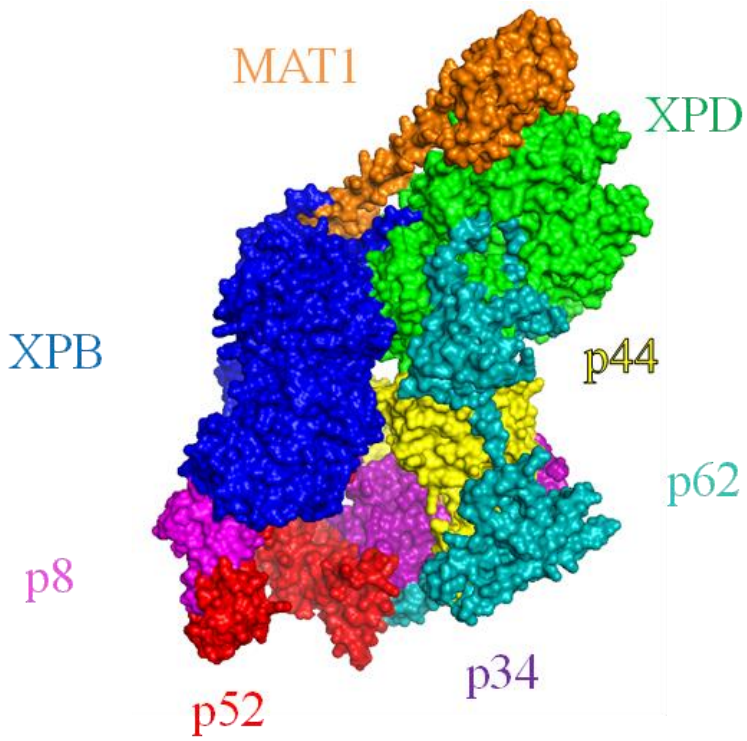


Figure 1.9 Recent Cryo-EM structure of transcriptional TFIIH. This high-resolution structure shows the core TFIIH subunits and the architecture of the complex. For the first time the majority of p62 can be assigned unambiguously. The extensive contacts between p44, p62, and XPD can be seen here. This interaction is explored further in chapter 4. Structure created in PyMol using PDB: 6NMI (239).

1.7 Disruption of DNA repair leads to disease

Mutations in genes encoding DNA repair proteins can impair the proteins ability to repair DNA. This is highlighted by the tumour suppressor protein p53, found mutated in ~ 50% of cancers (241). Mutations in the gene coding for p53 show a "UV mutation signature", a specific marker that UV radiation disrupts p53's function. This disruption compromises the ability of cells to repair DNA (242). These defects in cellular DNA repair machinery can lead to a number of diseases, three main diseases that arise from defects in NER are described below:

Xeroderma pigmentosum (XP) is phenotypically characterized by UV hyper-sensitivity and increased incidence of skin cancers. XP patients have a 1000-10,000 fold increased frequency of skin cancer and usually succumb to metastatic carcinoma early in life (243-246). Mutations that cause XP can be found in various NER proteins and are classified into 7 complementation groups depending on which protein is disrupted. Interestingly, some XP complementation group patients develop progressive neurological degeneration, linked to TFIIH's role in transcription and other cellular roles beyond that of DNA repair (247). Most disease causing mutations found in XPD disrupt protein-protein interactions with the core of TFIIH but not XPD's own activity (199, 207, 228, 248).

Trichothiodystrophy (TTD) clinically presents with short stature, mental retardation, ataxia, spasticity, and brittle hair (249, 250). Mutations in three TFIIH genes; *XPB*, *XPD*, and *GTF2H5* underlie the pathophysiology (251). In around 50% of cases photosensitivity is reported, similar to XP (249). However, there is a large variance in clinical severity in TTD, linked with a large number of mutations in three different genes (252). Additional physiological symptoms such as short stature and sterility are likely caused by the crossover of DNA repair and transcriptional roles of TFIIH. It has been shown that TTD mutations reduce the cellular concentration of TFIIH (253, 254) and this reduction may affect the expression of genes in certain tissues at critical periods (157). For more information on TTD see Stefanini et al (252).

Cockayne syndrome (CS) symptoms include dwarfism, mental retardation, microcephaly as well as skeletal and retinal abnormalities (255, 256). CS is caused by disrupted TC-NER (256). The CSB protein (eukaryotic functional equivalent of Mfd) aids RNA polymerase

stalled on DNA, and helps recruit NER factors to DNA (88). As well as this, CSB itself acts as a transcription factor (257). For a great review see Karikkineth et al 2017 (258).

These diseases can be very severe, but are rare, with XP having an incidence of around one in a million people (Europe and the USA (259)). The diverse features caused by mutated DNA repair proteins are not yet fully understood. Furthermore, crossover between DNA repair and transcriptional roles of TFIIH make deciphering the pathophysiology difficult, especially when diseases such as XP and CS share clinical features. However, these diseases do provide an insight into how critical DNA repair is for organism survival, and may help us understand the bigger picture of multifunctional proteins such as TFIIH.

1.8 How do DNA repair enzymes locate DNA damage?

Biological processes such as replication, transcription, and DNA repair require protein movement along DNA (260). In the case of a DNA repair enzyme finding damage, this must occur on a temporal scale that enables cell survival. With such a vast number of undamaged base pairs, and a relatively miniscule number of damage sites, how are proteins able to detect DNA damage and carry out repair before cells are subject to potentially dangerous mutations? In a cellular environment, the cytosol is packed with macromolecules. A protein using a diffusive search for a damaged base among millions of structurally similar undamaged bases is analogous to finding a needle in a haystack. Constant collisions with macromolecules and incorrect substrates take up time and affect the proteins trajectory within a cell. Biological complexities also have to be considered when thinking about diffusion in a cell; DNA is littered with other proteins that roadblock protein movement, complex formation may have to occur either on or off of the DNA, and electrostatic interactions with DNA are affected by the local salt concentration (260, 261). A study using molecular dynamic simulations provides a fascinating insight into how proteins might interact with a target in cells. Kanada et al found that crowding in a cell enhances the on rate of proteins associating with DNA, but that the diffusion constant is slowed when the DNA contains many nucleosomes. Strikingly, despite having a slower diffusion constant, the simulations revealed that as the nucleosome density increases, a transcription factor was actually able to find its target faster due to the higher local concentration of DNA (262).

Using physical models, it is possible to calculate the rate a protein can find its target. A proteins molecular mass and radius can be determined and then using the Stokes-Einstein equation the velocity of diffusion can be determined, accounting for drag and viscosity of

the solution. The Smoluchowski limit describes the maximum probability of the protein and DNA damage site encountering each other using a random three-dimensional search on DNA (263). The probability of an enzyme binding directly to a lesion is very low and would require a large number of binding and releasing events in order to find damage. In 1970, Riggs et al experimentally measured the association rate of the Lac repressor for the lac operator DNA sequence. They found the value was 1000x faster than the calculated limit for a three-dimensional diffusive search on DNA (264). The association rate was dependent on ionic strength, suggesting that DNA binding proteins roll or hop along the DNA to find their target (264). Since then several mechanisms for proteins locating their targets have been modelled and demonstrated experimentally, these are shown in **figure 1.10**, and discussed below.

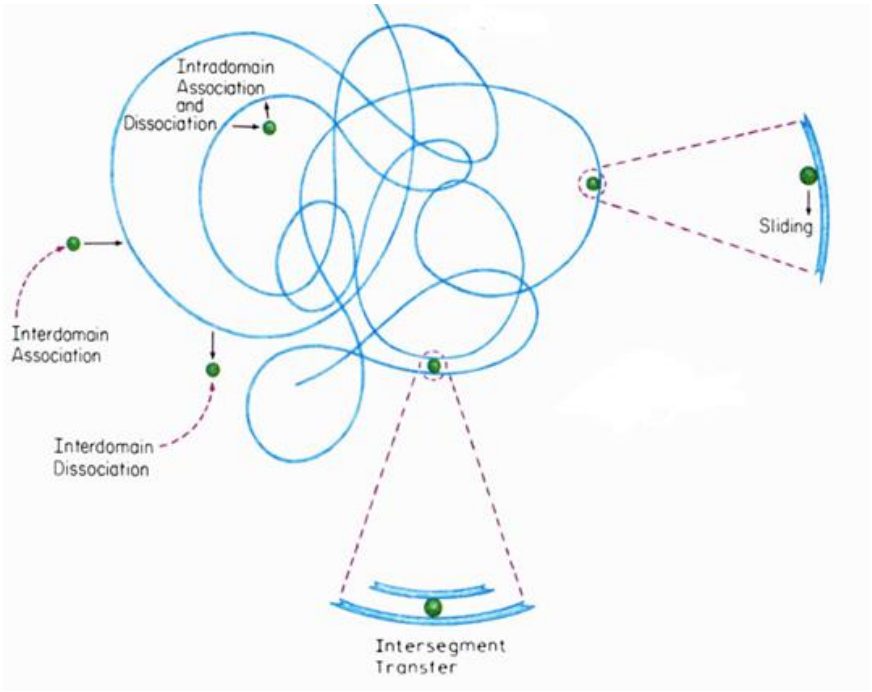


Figure 1.10 Different protein-DNA associations. Various proteins interact with the DNA in order to find their target site. Figure adapted from (263)

1.8.1 Sliding

Sliding was first described by Adam and Delbruck as a one-dimensional search along the contour length of DNA, scanning adjacent sites (265). Riggs and others then provided experimental data for this, showing that a collapse in dimensionality, from three dimensions to one dimension greatly enhances the rate at which a protein can find its target on DNA (260, 263, 264, 266, 267). Sliding is not an active process but occurs through the relatively weak electrostatic interactions between the positively charged DNA binding site of the protein and the negatively charged DNA. This binding energy keeps the protein in close proximity to the DNA backbone and facilitates sliding. As proteins slide they must displace positively charged ions from the DNA. This displacement of ions is driven by the thermal energy of the system. As the protein displaces ions in the direction of travel, ions rebind to the previously occupied site. This results in no overall displacement of ions and means sliding is not dependent on the ionic strength (263). If the salt concentration is reduced, proteins spend more time sliding on the DNA before releasing into solution. Whereas at high salt concentration the ionic strength weakens the protein-DNA interactions and favours dissociation into solution (268). As proteins slide along DNA, they follow the helical backbone. In doing this, proteins experience rotation-coupled drag proportional to their hydrodynamic radius, Schurr proposed a model to calculate the theoretical upper limit for protein sliding along the DNA backbone by using the Stokes-Einstein equation and accounting for rotation-coupled drag (263, 269).

1.8.2 Hopping

The condensed structure of DNA in the nucleus provides an ideal environment for hopping, where full dissociation of the protein from DNA is accompanied by a high probability of rapid rebinding, resulting in a fast but local search (260, 270). Unlike sliding, hopping is affected by the ionic strength, as the complete dissociation and re-association of a protein to DNA is shielded by increasing salt concentrations (270). In higher salt solutions, the attachment time of a protein to DNA is reduced, while there is the potential for more rapid diffusion (260). This sensitivity to salt can be tested experimentally to determine how a protein moves on DNA (270).

1.8.3 Intersegmental transfer

In cells, where DNA is densely compact, proteins are able to jump across to other DNA molecules, if the protein possesses multiple DNA binding sites, it is possible to bind multiple DNA molecules and translocate between them – known as intersegmental transfer (260, 263). An example of intersegmental transfer is the $UvrA_2B_2$ complex (discussed in section 1.5), which can jump between two DNA molecules, without full dissociation from the DNA (263, 270).

1.8.4 How to reconcile these different mechanisms of translocation

In vivo DNA is bound by many proteins such as histones and transcription factors. Collisions between proteins and these obstacles on DNA affect the rates observed, furthermore molecular crowding in solution and local salt concentrations all contribute, but are difficult to recapitulate *in vitro*. For a protein to find its target site, a combination of these search mechanisms such as one-dimensional and three-dimensional diffusion are required for an efficient search. Repeated binding and releasing from DNA, interspersed with sliding along the DNA in order to find a lesion is likely how proteins are able to find their targets in a time frame that allows for cell survival (263, 270, 271).

1.9 Using single molecule approaches to study DNA repair

Single molecule techniques have revolutionised our ability to investigate and understand how biological processes occur. The use of bulk phase studies prevents the individual behaviour of proteins being distinguished, with only an average being represented in the final measurements. Given the complexity of DNA repair pathways and frequent cross talk, this makes it hard to define a cause & effect relationship. However, using single molecule approaches it is possible to isolate individual functions of proteins and then reconstruct the system *in vitro* (226, 272). Super resolution microscopy now allow sensitive observations of individual protein molecules interacting with the DNA. This is a particularly powerful technique for studying protein:DNA interactions, where proteins can be observed binding to DNA in real time, and the rates at which this occur can be quantified. The use of single molecule techniques has been excellently demonstrated by seminal findings showing the mechanical properties of DNA, how a protein finds its target, and how helicases unwind DNA (273-279). Classic fluorescent proteins such as GFP photobleach quickly, and prevent imaging for long times, or at high speeds. However, progression in fluorophore design has

yielded several new, bright, and stable fluorophores such as mScarlet that is used in this thesis (280). Additionally, fluorescent semiconductor nanocrystal particles, quantum dots (QDOTs) are resistant to photobleaching, and have sharp, extremely bright emission peaks, but a very broad excitation wavelength making them ideal for single molecule imaging of proteins interacting with DNA (273).

1.10 Aims of this thesis

The aim of the work in this thesis is to understand the mechanistic details of how nucleotide excision repair proteins can find and verify the presence of DNA damage. Discovering the minute details of DNA repair is key for understanding the cellular processes that maintain genomic stability but when disrupted, lead to aging and diseases such as cancer. Using a combination of biochemical and modern single molecule imaging techniques I aim to break down the complexity of the DNA repair pathways and understand individual processes that can be put into a broader context with recent crystal and cryo-EM structures (239, 281).

Chapter 2

Experimental methods

2.1 Common buffers and solutions

XPD buffer;
20 mM Tris pH 8.0, 1 mM TCEP, 10 mM KCl, 1 mM EDTA, 5 mM $MgCl_2$

ABC buffer;
50 mM Tris pH 7.5, 50 mM KCl, 10 mM $MgCl_2$

UvrA storage buffer;
50 mM Tris pH 7.5, 500 mM KCl, 0.1 mM EDTA, 5 mM DTT

p44/p62, p62 storage buffer;
20 mM Tris pH 8.0, 200 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP

XPD, N-p44 storage buffer;
20 mM Tris pH 8.0, 10 mM KCl, 5 mM $MgCl_2$, 1 mM TCEP

All stock solutions and protocols were based on Molecular cloning - A laboratory manual, or the manufacturers guidelines unless stated otherwise (282). Reagents were purchased from Sigma-Aldrich UK or Thermo Fisher Scientific UK unless stated otherwise.

2.2 Cloning of UvrA gene

To create a UvrA-mScarlet fusion protein the UvrA gene was cloned into a vector upstream of the fluorescent protein gene (pET21a). The pET21a vector contained an unrelated gene (RLC) upstream of mScarlet that was removed using primers that amplified the vector outside of the RLC gene region. These primers also contained a tail region that was complementary to the start and end sequences of the UvrA gene to enable Gibson assembly to take place (**figure 2.1**). The amplified, linearised plasmid with RLC removed was analysed on an agarose gel and extracted. Primers used to remove RLC and incorporate complementary sequences to the UvrA gene:

Reverse primer;
5'CGCCCCGAACTTCGATCTTATCCAT**ATGTATATCTCCTTCTTAAAGTT** 3'

Forward primer;

5' CTTCCTTAAGCCGATGCTG**TCAGGTAGC TCTGGCACTTCA** 3'

The bold regions are complementary to the pET21a vector and remove the RLC gene from the plasmid. The underlined regions are complementary to the start (reverse primer) and end (forward primer) of the UvrA gene to be inserted into the plasmid.

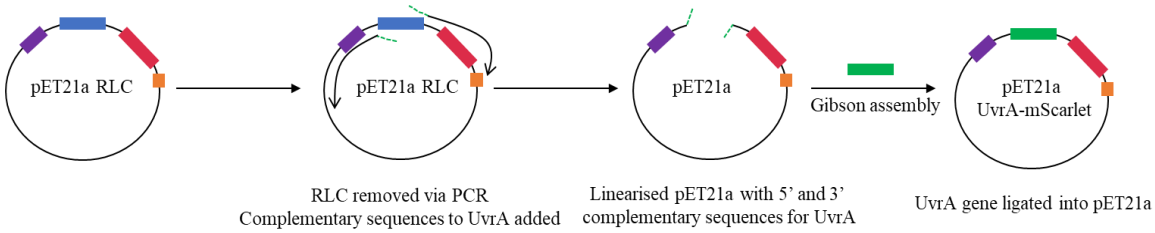
The gene for *E. coli* UvrA was obtained from the pCA24N vector (Kindly provided by the National Bioresource Project (NIG, Japan)) and was amplified using the following primers:

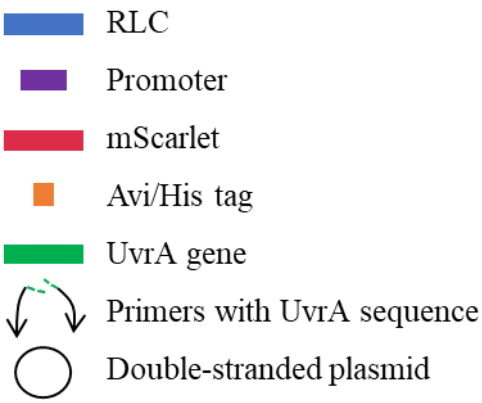
Forward primer; 5' ATGGATAAGATCGAAGTTCGGGGCG 3'

Reverse primer; 5' CAGCATCGGCTTAAGGAAGCGTGC 3'

The UvrA gene was then isolated from the plasmid on an agarose gel and extracted. Both the purified insert (UvrA gene) and the vector (pET21a) were then ligated together using the Gibson assembly kit (New England Biolabs (NEB)). The final construct contained the UvrA gene with a C-terminal mScarlet fluorescent protein separated by a flexible linker. Further C-terminal to mScarlet was a hexa-histidine (His) purification tag and an AviTag. The correct gene product was confirmed by site-specific restriction digests and sequencing using generic T7 primers and is shown in **figure 2.1**.

A





B

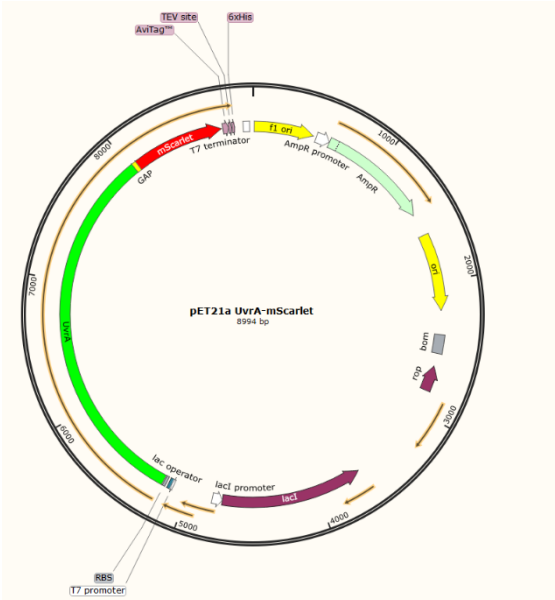


Figure 2.1 Cloning of UvrA into pET21a. A) Schematic of the removal of RLC from the pET21a vector via PCR and the insertion of the UvrA gene using the Gibson assembly kit (NEB). B) A plasmid map of the sequenced pET21a construct that was used to purify UvrA-mScarlet.

2.3 Protein purification - UvrA

The UvrA-mScarlet vector was transformed into BL21-DE3 cells (NEB) using standard protocols. A single colony was grown in lysogeny broth (LB) with 100 µg/ml ampicillin overnight. In the morning cultures were diluted 100-fold and inoculated into fresh media with antibiotic. After reaching OD600 0.4-0.6, 20 mL of the cells were inoculated into 1 L LB flasks (prewarmed, with fresh antibiotic). Flasks were grown at 37°C, shaking at 260 rpm until hitting mid exponential phase (OD600 0.4-0.6). 1 mM IPTG was added to induce the culture. Inclusion bodies formed with a 27°C degrees 3 hour induction, therefore an overnight induction at 18°C was used. Induction of UvrA-mScarlet production was confirmed by SDS-PAGE of cell samples pre- and post- addition of IPTG. Induced cells were pelleted at 5000 RPM for 10 mins at 4°C in a JA10 rotor (Beckman). The scarlet coloured pellet was resuspended in 10 mM Tris pH 8 and 200 mM NaCl and frozen at -20°C until purification.

UvrA proved difficult to purify; it precipitated in low salt conditions and bound to spin concentrator membranes (Amicon). There were also problems with protein degradation and non-specific binding of proteins to both the HiTrap and Heparin columns (GE healthcare). To prevent precipitation, a buffer precipitation test was used to determine UvrA's minimum salt requirements. ABC buffer solutions containing KCl concentrations from 0.1 M to 0.3 M in 50 mM steps were created. A scarlet coloured His-column elution fraction was diluted into 2 mL of different salt solutions and left on ice with gentle shaking for 2 hours, then centrifuged. No pellet was observed in any sample. Samples were then left overnight at 4°C. After brief centrifugation, precipitation was noted up to 200 mM KCl; buffers for purification after this contained ≥ 200 mM KCl / NaCl. Cell pellets were thawed on ice with halt protease inhibitor cocktail ((no EDTA) Thermo Fisher scientific)) and kept on ice for sonication; 30 seconds on, 59 off, ~60% amp, for 5 minutes of 'on'. The cell lysate was spun at 20k RPM for 20 mins at 4°C. The pellet was discarded, and the supernatant was spun in an ultracentrifuge at 50k RPM for 1 hour at 4°C. Sodium phosphate and imidazole were added to the clarified supernatant to final concentrations of 50 mM. The supernatant was then loaded on a HiTrap Nickel column (GE Healthcare) equilibrated with buffer A (50 mM imidazole 500 mM NaCl, 50 mM sodium phosphate pH 7.5) and washed with 30 CV (column volumes, 1 mL) of buffer A. Proteins were eluted with a linear gradient from 50-200 mM imidazole run over 40 CV. Fractions containing UvrA-mScarlet were identified by

their scarlet colour, and were then analysed by SDS-PAGE with coomassie staining (**figure 2.2**). The purest fractions were pooled together and diluted in low salt buffer B (50 mM Tris pH 7.5, 10 mM $MgCl_2$, and 250 mM NaCl) and loaded on a HiTrap Heparin column (GE healthcare). Proteins were eluted with a gradient from 250 mM to 1 M NaCl. Fractions were again identified by their colour, and analysed by SDS-PAGE (**figure 2.2)**.

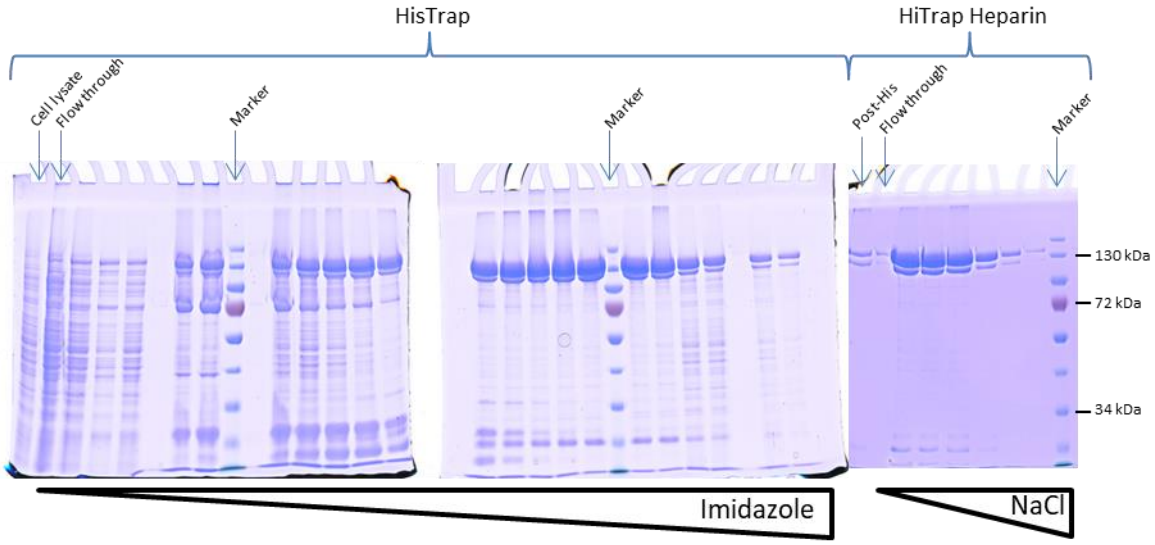


Figure 2.2 Purification of UvrA-mScarlet. SDS-PAGE of HisTrap (left two gels) and HiTrap Heparin column fractions (right gel). Purification of overexpressed UvrA-mScarlet (~130 kDa) can be seen from *E. coli* cell lysate through to almost pure protein fractions from the heparin column. Heparin fractions containing a large proportion of UvrA-mScarlet also contained a slightly lower molecular weight product, most likely a degraded form of UvrA that was unable to be separated further.

Pure fractions were pooled and spin concentrated in a Vivaspin 100 kDa filter into UvrA storage buffer and made 50% w/v with glycerol, stored at -20°C. Concentration of UvrA-mScarlet was determined spectrophotometrically. The absorbance of mScarlet at OD569 (ε 100,300 M^{-1} cm^{-1}) across a dilution series gave a concentration of 3 µM, giving a protein yield of 73 µg/L of culture.

2.4 Protein purification - TFIIH subunits

Purified N-p44, p62, p44/p62 and XPD were a kind gift from Professor Caroline Kisker & Dr Jochen Kuper (University of Würzburg, Germany) as part of a collaboration. The purification procedure for these proteins is described below. The p44/p62 complex contained full length p44, and p62 protein, whereas p44 alone (referred to as N-p44) contains only the first 285 residues of the protein.

Cloning and purification was carried out by Dr Jochen Kuper using the following method. The genes encoding full length p44 and p62 were cloned from *Chaetomium thermophilum* cDNA. p62 was cloned into the pETM-11 vector (EMBL) without a tag. p44 was cloned into the pBADM-11 vector (EMBL) containing an N-terminal hexa-Histidine tag. p62 and p44 were co-expressed in *E. coli* BL21 CodonPlus (DE3) RIL cells (Agilent) and were co-purified via immobilized metal affinity chromatography (Ni TED, Machery-Nagel), followed by size exclusion chromatography (SEC), and anion exchange chromatography (AEC). SEC was conducted with a HiLoad 16/600 Superdex 200 prep grade column (GE Healthcare) in 20 mM Hepes pH 7.5, 250 mM NaCl, and 1 mM TCEP. AEC was conducted with a MonoQ 5/50 GL column (GE Healthcare). The proteins were eluted via a salt gradient ranging from 50 to 1000 mM NaCl. AEC buffers were composed of 20 mM HEPES pH 7.5, 50/1000 mM NaCl, and 1 mM TCEP. The p44/62 protein complex was concentrated to approximately 20 mg/ml and flash frozen in liquid nitrogen for storage.

XPD and N-p44 (residues 1-285) were also from *C. thermophilum.* XPD was N-terminally His-tagged and expressed in *E. coli* ArcticExpress (DE3)-RIL cells (Agilent). Cells were grown in TB medium at 37°C until they reached an OD_{600} of 0.6. Expression was induced with 0.05% L-arabinose at 11°C for 20 h. N-p44 (His tagged) was expressed in *E. coli* BL21-CodonPlus (DE3)-RIL cells (Stratagene). Cells were grown as described for ctXPD and expression was started by adding 0.1 mM IPTG at 14°C for 18 h. XPD and N-p44 were purified to homogeneity by metal affinity chromatography (Ni-IDA, Macherey&Nagel) followed by SEC (20 mM HEPES pH 7.5, 200 mM NaCl) and an additional AEC step in the case of XPD. AEC was performed with 20 mM HEPES pH 7.5, 50 mM NaCl, and 1 mM TCEP as loading buffer and the same buffer containing 1 M NaCl was used for elution. The final buffer after AEC was 20 mM HEPES pH 7.5, 250 mM NaCl, and 1 mM TCEP. Proteins were concentrated and flash frozen for storage at -80°C.

2.5 Complementation assays

To investigate whether our UvrA-mScarlet construct can rescue $UvrA^{WT}$ knockout cells (ΔUvrA), we performed *in vivo* UV complementation assays. ΔUvrA (Keio) cells were obtained from National Bioresource Project (NIG, Japan) and transformed with UvrA-mScarlet or empty mScarlet vector. Transformed cells were grown in LB with 25 µg/ml chloramphenicol to OD600 = 0.5. 5 uL of undiluted, and three, ten-fold serial dilutions were plated on LB agar and subjected to either no UV or 5 J/m^2 UV (254 nm) irradiation and incubated overnight at 37°C in the dark.

2.6 NADH-linked ATPase assay

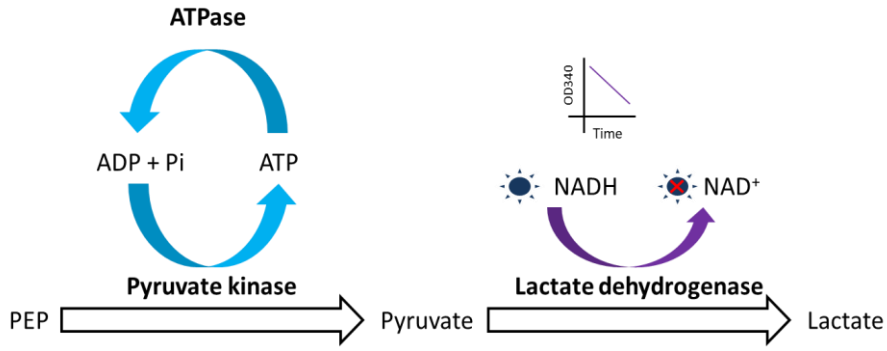


Figure 2.3 The NADH-linked ATPase assay. A schematic of how ATP turnover is coupled to a decrease in absorbance at OD340 as NADH is converted to NAD+.

Two variations of the ATPase assay are described in this thesis, the cuvette format (UvrA) and the plate reader format (p44/p62). In this assay the turnover of ATP to ADP is coupled to the production of pyruvate from phosphoenol pyruvate (PEP) by pyruvate kinase (PK). This recycles ADP and P_i back to ATP, preventing ADP accumulation. Pyruvate is then converted into lactate by lactate dehydrogenase (LDH) oxidising NADH to NAD^+. This process is shown in **figure 2.3**. Absorbance of NADH (340 nm) is monitored and decreases as ATP is turned over and NADH is oxidised.

The NADH-reaction components are assembled as follows;

- PEP – dissolved into the reaction buffer (either ABC buffer or XPD buffer) to a final concentration of 0.5 mM and frozen on dry ice and stored at -20°C
- NADH – dissolved in 0.1 M Tris pH 8.0 to a final concentration of 30 mg/ml and aliquots flash frozen in liquid nitrogen to avoid freeze-thaw cycles. Stored at -20°C.
- PK and LDH are a premixed stock solution consisting of (PK 600-1000 U/mL) and LDH (900-1400 U/mL) from Sigma-Aldrich (P0294). 10 µL of this solution is added per 500 µL of ATPase reaction

2.6.1 UvrA

For the cuvette format ABC buffer containing 0.5 mM PEP was thawed on ice and 1 mM DTT added. This solution was used to blank a cuvette at 340 nm in a spectrophotometer. NADH was added to a final concentration of 210 µM along with the PK/LDH premixed stock. Purified UvrA was added to a final concentration of 25 nM and the turnover of ATPase rate measured after starting the reaction with the addition of ATP (1 mM final) and various DNA substrates (all 50 ng).

2.6.2 XPD, N-p44 and p44/p62

For the plate reader format the NADH-reaction components were assembled as a master mix with 1 mM TCEP. This solution was dispensed into separate tubes where the individual reaction components were added before being plated into a 96 well plate (100 µL per well). Proteins were added to a final concentration of 100 nM (equimolar concentrations for XPD & N-p44 and XPD & p44/p62) and DNA substrates to a concentration of 50 nM. The reaction was started with the addition of 1 mM ATP to each well. The change in OD340 was monitored every 8 seconds/well over 30 minutes at room temperature in a plate reader (Clariostar, BMG Labtech).

For both the cuvette and plate reader format the change in OD340 was fitted linearly to calculate loss of NADH (6220 M^{-1} cm^{-1} at 340 nm), enabling calculation of k_{cat}. Reactions were repeated three times and either standard deviation (UvrA) or S.E.M (XPD) used as error. An example of the absorbance versus time plots for both UvrA and XPD are shown in **figure 2.4**.

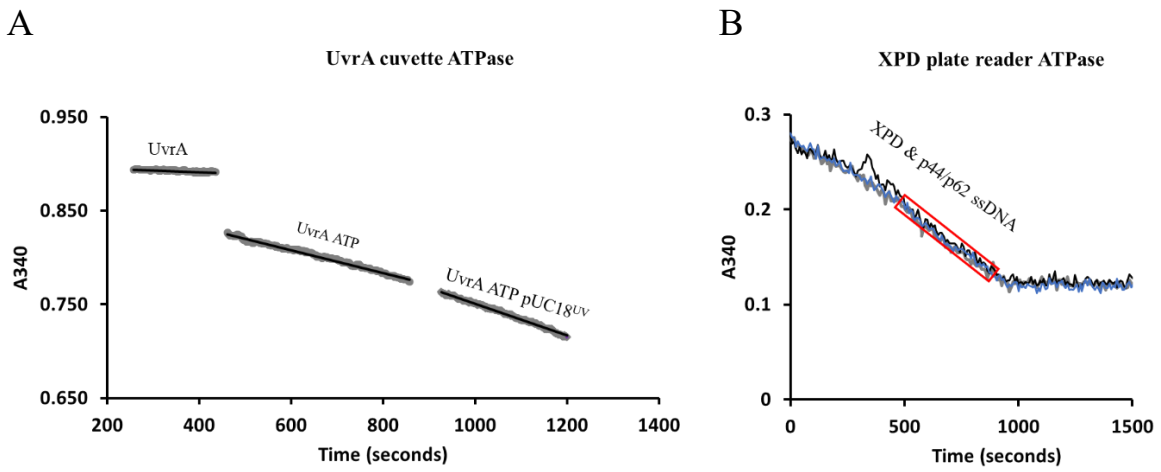


Figure 2.4 Real-time NADH-coupled ATPase assay raw data. A) Rates of NAD^+ production by UvrA alone, UvrA with 1 mM ATP added and after the addition of 50 ng 1000 J/m^2 irradiated pUC18. Data traces are fit to straight lines (black lines) to determine the rate of ATP hydrolysis per UvrA monomer, k_{cat} as described throughout this thesis. Gaps between the data traces are due to the spectrophotometer being opened and are easily removable from the data. B) Raw data for the plate reader ATPase assay. Three wells data are overlaid and show XPD's ATPase in the presence of p44/p62 and undamaged ssDNA. The reaction reaches the end point where the absorbance plateaus. As fluctuations in the linearity occur within the data, only the linear portion of the rate (shown by a red box in B) is used to determine the rate constant.

2.7 ATPase DNA substrates

A variety of different substrates were used in the ATPase assay; a short oligonucleotide (F26,50), lambda DNA, and pUC18 plasmid. These substrates were either undamaged or damaged. All oligonucleotides were purchased from Sigma-Aldrich. Lambda DNA and pUC18 were obtained from NEB.

pUC18; Double stranded plasmid cloning vector pUC18 was prepared by Quentin Smith (undergraduate student) using the QIAGEN Midi prep kit.

Double-stranded oligonucleotides (F26,50 undamaged and damaged) were produced by mixing equimolar concentrations of the reverse complementary non-damaged oligonucleotide, and either the damaged or undamaged forward oligonucleotide in 1x TE buffer (10 mM Tris, 1 mM EDTA, pH 8) heated to 95°C for 5 minutes and left to cool slowly to room temperature.

Undamaged F26,50;

5`GACTACGTACTGTTACGGCTCCATCTCTACCGCAATCAGGCCAGATCTGC 3`

Damaged F26,50 contains a fluorescein moiety covalently attached to thymine (*), previously demonstrated as a target for NER (134);

5`GACTACGTACTGTTACGGCTCCATCT*CTACCGCAATCAGGCCAGATCTGC 3`

F26,50 reverse complementary oligonucleotide;

5`GCAGATCTGGCCTGATTGCGGTAGCGATGGAGCCGTAACAGTACGTAGTC 3`

To introduce UV damage into pUC18 and lambda DNA a calibrated 254 nm UV lamp was used (ENF-240C/FE, Spectronics, New York). DNA substrates were irradiated to 1000 J/m^2 (unless stated otherwise) immediately before being used. The DNA solution was pipetted onto parafilm over a beaker of ice to prevent evaporation of the solution. DNA was irradiated at a concentration of 25 ng/µL to prevent excessive absorbance, permitting uniform damage throughout the solution. The lamp was then fixed at a distance above the substrate that was calibrated to induce a known amount of radiation to the sample in a time frame.

2.8 Confirming the presence of DNA damage

To practically demonstrate the presence of damage in the UV irradiated substrates we used gel based assays. T4 endonuclease V (T4EV (NEB)) is a DNA glycosylase that recognises CPD lesions and nicks the DNA 5' of the lesion. For plasmids, irradiation with UV induces CPD lesions that when nicked by T4EV cause loss of supercoiling, visible on an agarose gel. A high concentration of ethidium bromide (10 µg/mL) induces plasmid supercoiling, allowing differentiation of supercoiled and nicked (open circle) bands by molecular weight. For long, linear lambda DNA molecules this assay cannot be used. Instead, an assay using T4EV and S1 nuclease (Thermo Fisher scientific) was devised. S1 nuclease cuts at nicks in the DNA, causing double-strand DNA breaks. If CPDs are present, T4EV will nick the DNA 5' of the lesion, which S1 Nuclease will digest to cause a double-strand break. As CPDs occur randomly on DNA, the loss of the lambda band to a smear of DNA products is observed on a 1% agarose gel in the presence of both enzymes and UV damage. For the gels and more explanation see section 3.8.

UV damaged DNA substrates were irradiated with 254 nm light to a final dose of 300 J/m^2 (pUC18) or 1000 J/m^2 (lambda DNA), with 150 ng of DNA being used per reaction. Undamaged control substrates used the same mass of DNA without UV irradiation. T4EV reactions for both pUC18 and lambda DNA were carried out in T4 PDG buffer (NEB) with 100 µg/ml BSA and 10 units of T4EV enzyme for 30 minutes at 37°C. S1 nuclease reactions were carried out in S1 nuclease buffer with 10 units of S1 enzyme for 15 minutes at 37°C. For reactions that were digested with T4EV, before S1 the volume of the reaction was increased 2-3 fold when diluting into S1 nuclease buffer, as S1 is sensitive to EDTA and high salt concentrations. All reactions that contained S1 (either alone or with T4EV) were heat inactivated at 70°C for 10 minutes and immediately loaded onto gels. Reactions with only T4EV cannot be heat inactivated but were kept on ice after the incubation period and then immediately loaded on a gel.

2.9 Helicase assay

Helicase assays were performed by the Kisker lab using a fluorescence-based assay. A dabcyl fluorescence quencher and Cy3 fluorophore are in close proximity at the ends of a double-stranded, open fork substrate. As a helicase unwinds the substrate the quenching is relieved, giving an increase in fluorescence. The substrate is shown below, with the underlined section not base pairing to form the fork before the duplex DNA.

5`AGCTACCATGCCTGCACGAATTAAGCAATTCGTAATCATGGTCATAGC-Cy3 3`
3`GATGTCAAGCAGTCCTAAGGTTCGTTAAGCATTAGTACCAGTATCG-Dabcyl 5`

Assays were carried out in 20 mM HEPES pH 7.5, 50 mM KCl, 5 mM $MgCl_2$, and 1 mM TCEP. DNA was used at a concentration of 250 nM, with equimolar concentrations of XPD, N-p44, and/or p62. The mix of reagents, were preincubated at 37°C and the reaction was subsequently started with the addition of 5 mM ATP. Kinetics were recorded with a Fluostar Optima plate reader (BMG Labtech). Fluorescence was detected at an excitation wavelength of 550 nm and an emission wavelength of 570 nm. Initial velocities were fit using the MARS software package and represent averages of at least three different reactions.

2.10 Flow cells

A custom microfluidic flow chamber was used to create DNA tightropes. A standard glass slide (Thermo Fisher scientific) forms the base; two holes are then created using a diamond drill tip (Dremel). Cleaned (section 2.10) glass slides have two inlet tubes inserted into the holes, and pulled tight (**figure 2.5).** The inlet tubes are then glued to the slide using UV curable glue (NOA58, Thor labs). It is important that the tubes are tight and well-sealed to the glass to prevent leaking during tightrope preparation. Once set, a double-sided sticky tape gasket, cut using a printer (Explore air, Cricut) to 15 mm x 10 mm is placed over the slide, as shown in **figure 2.5**. The top layer is peeled off, leaving a 180 µm thick adhesive layer. A coverslip (24 x 40 mm, #1.5 thickness) that has been silanised (described in section 2.10) is placed onto, and firmly compressed against the adhesive using the back of a scalpel. This forms the basic flow cell used for all DNA tightrope experiments described in this thesis.

N.b. The exact dimensions of the inlet tubing are important as they affect the flow rate; the flow cell tubing has an inner diameter of 1.27 mm and a 0.82 mm wall (GE 0127-082). To connect a syringe pump to this, a 1.15 mm inner diameter tubing is used, with a 0.2 mm wall (GE-0115-020).

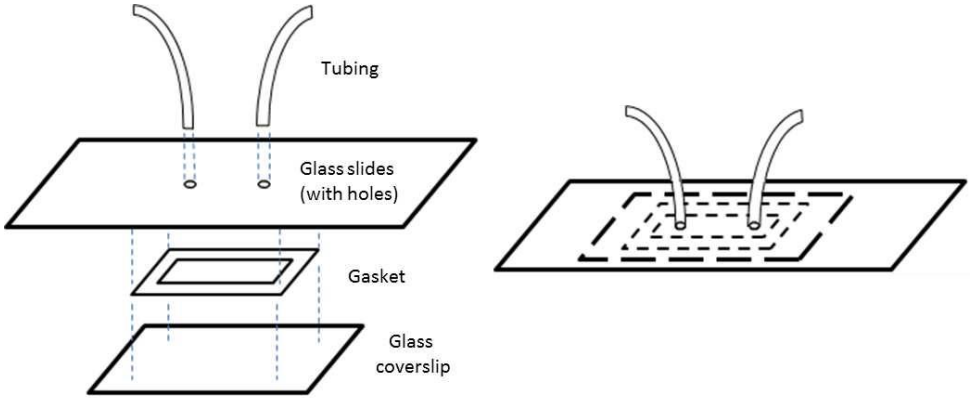


Figure 2.5 Flow chamber construction. A microfluidic flow cell is constructed from a glass slide, gasket, and coverslip. Drilled holes in the glass slide allow tubing to be inserted and glued in place on the chamber to allow solutions to be introduced. Figure taken from Springall et al (283).

2.11 DNA tightropes

DNA tightropes are a powerful single molecule technique that can be used to study DNA:protein interactions in real time. Individual DNA molecules are elevated above a surface, enabling fluorescence imaging of the focal plane with reduced background from the surface. The microfluidic chamber used to create DNA tightropes gives precise control of buffer conditions and sequential introduction of reagents. Before creating DNA tightropes, flow chambers are blocked to prevent non-specific surface binding of proteins, fluorophores, or other objects. Flow cells are incubated overnight in mPEG buffer (25 mg/ml $mPEG_{5000}$ (Sigma-Aldrich #10859) in 250 mM $NaHCO_3$, pH 8.2). The mPEG is functionalised with *N*-succinimidyl ester that bonds to the aminopropyl groups deposited by silanisation (described later). They are then washed with dH_2O and incubated overnight in ABT buffer (10 mg/ml BSA, 0.1 % tween & 0.1% sodium azide). Before creating tightropes the flow cell is washed with dH_2O. Silicon beads (5 µm (MicroSil Microspheres, Whitehouse Scientific)) are coated with 350 µg/ml poly-L-lysine (P5899, Sigma-Aldrich) to enable them to stick to the coated glass coverslip. However, to reduce aggregation of the beads, they are washed with dH_2O to remove excess poly-L-lysine and recovered by centrifugation immediately before preparing tightropes. Beads are sonicated for ~ 1 second, in 3 separate bursts to disperse them and are injected into the flow chamber, taking care not to introduce any air. A peristaltic pump (AL1000-220, world precision instruments) is then attached via a syringe and tubing and used to control the movement of reagents through the flow chamber. Tubing is attached to the other side, and flowed through to an Eppendorf reservoir containing a drilled hole to introduce fluids (**figure 2.6**).

It is imperative not to introduce any air at this stage, as it will disrupt the beads and prevent tightrope formation. DNA substrates are introduced into the reservoir and withdrawn into the flow chamber where a loop is set, flowing the DNA back and forth over the beads in the flow chamber to create DNA tightropes (**figure 2.6**). A constant flow rate of 300 µL/min is used at all steps in this protocol. For normal lambda DNA tightropes ~ 500 ng of DNA is used. After approximately 20 minutes, the loop is cancelled and buffer flowed through the flow chamber to wash out any excess DNA that has not formed DNA tightropes. The chamber is then ready for the introduction of proteins etc. An intercalating fluorescent DNA dye such as YOYO-1 (Invitrogen) can be used to show the presence of DNA tightropes, shown in **figure 2.6**.

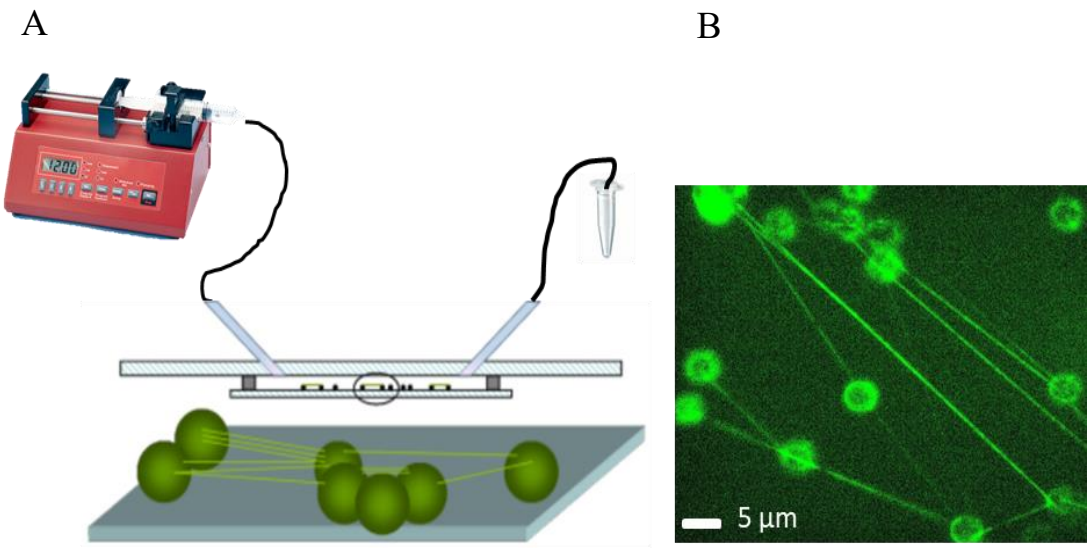


Figure 2.6 An assembled flow cell containing DNA tightropes.
A) A diagram showing the setup of a peristaltic pump connected to a flow chamber and a reservoir on the other side. The DNA tightrope structures can be seen between immobilised silica beads on the coverslip surface.
B) YOYO-1 stained DNA tightropes as viewed on the microscope system. Individual molecules of DNA can be seen between the beads.

After the experiment the tubing, flow cell chamber, and syringes are washed with 100% methanol and rinsed with dH_2O before being used again. Flow chambers are recycled by shaking in acetone overnight to dissolve the sticky tape and plastic. The coverslip and inlet tubes are not recycled. The glass slides are cleaned using acetone, ethanol, and vigorous scrubbing. Once cleaned, the glass slides and new coverslips are silanised in 2% (3-Aminopropyl) triethoxysilane (in acetone) for 10 minutes, and dried under a stream of nitrogen gas. Silanisation deposits aminopropyl groups on the glass surface that are then PEGylated by the $mPEG_{5000}$ 3-(Acetylthio) propionic acid *N*-succinimidyl ester solution. At this stage the coverslips and glass slides are ready to be made into flow chambers.

2.12 UV damaged DNA tightropes

For experiments with damaged tightropes, the procedure described in section 2.7 was applied to lambda DNA, immediately before being strung up to make tightropes.

2.13 Single-stranded gap tightropes

Single-stranded gap tightropes were created using oligonucleotides complementary to the cohesive ends of lambda DNA. Separately, 5 µg of lambda DNA is incubated with a 10 fold molar excess of oligonucleotide. The concentration of the lambda ends able to base pair with the oligonucleotides is 16 nM. Lambda and 'ssDNA oligo 1' or 'ssDNA oligo 2 bio' are incubated at 62°C for 5 minutes, and allowed to cool slowly to room temperature. The solution is made to 1 x T4 Ligase buffer, and 1 unit of T4 ligase is added. The reaction is incubated at room temperature overnight, and stored at 4°C in the morning. At this stage, lambda should be ligated to its complementary oligonucleotide at one end. Equal volumes of 'ssDNA oligo 1: lambda' and 'ssDNA oligo 2 bio: lambda' are then combined and heated to 62°C for 5 minutes. After slowly cooling to room temperature, 1 mM ATP, 1 unit of T4 Ligase are added, along with PEG_{6000} (10% final concentration). This reaction is incubated at room temperature for 1 hour. The lambda molecules should now be linked together with a 38 nucleotide single-stranded patch separating the molecules (**figure 2.7**). NaCl is added to dissociate T4 ligase from the DNA, and prevent the DNA becoming entangled (Dr Nicola Don, appendix I). This substrate is then used to form DNA tightropes; due to the lower efficiency of ligations 0.5-1.5 µg of DNA is used per flow cell. The inclusion of a biotin tag on one oligonucleotide allows it to be labelled with a QDOT, enabling spatial localisation using a fluorescence microscope.

ssDNA oligo 1;

5'GGGCGGCGACCTGCGTGATCTTTGCCTTGCGACAGACTTCCTTGGCTGGGCG GGCTGGC 3'

ssDNA oligo 2 bio;

5'AGGTCGCCGCCCGCCAGCCCG (TEG-bio) 3'

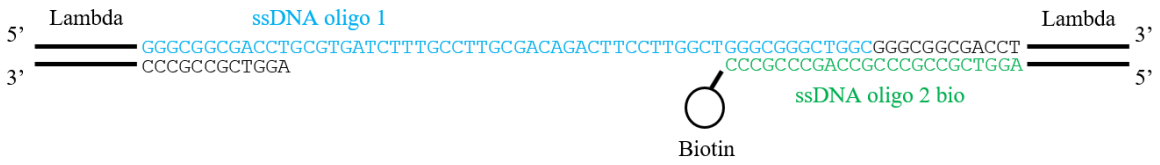


Figure 2.7 The single-stranded lambda DNA substrate. Oligonucleotides (blue and green) complementary to the overhanging ends of lambda (black sequence) base pair. Ligating these oligonucleotides bridges the two lambda molecules together to create a 38 nucleotide single-stranded patch. The green oligo here contains a biotin that can be fluorescently labelled with a QDOT to allow visualisation and colocalization with proteins.

2.14 Nucleotides in the DNA tightrope assay

Nucleotides (ATP, ADP, ATPγS (Adenosine 5′-[γ-thio]triphosphate)) were used at 1 mM final concentration. For experiments with P_i, 1 mM ADP with 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM DTT was supplemented with 19 mM sodium phosphate (freshly autoclaved to remove pyrophosphate) and the ionic strength balanced by the removal of 50 mM KCl (284, 285).

2.15 Protein labelling for imaging

Purified UvrA-mScarlet contained a fluorescent protein and no further labelling was necessary for visualisation. XPD, p44/p62 and p62 or N-p44 alone were visualised using fluorescent QDOTs (286). XPD, N-p44, and p62 contained N-terminal hexa-histidine tags (as did full length p44 in the p44/p62 complex), aliquots of 1 µM protein were thawed in the hand rapidly before being placed on ice. An equimolar concentration of 'primary' Penta-His antibody (Mouse, ID 34660 QIAGEN) was added to the protein and incubated on ice for ~ 30 minutes. The 'secondary' QDOTs ((F(ab')2-Goat anti-Mouse IgG (H+L) Secondary Antibody, ((Q11022MP 655 nm, Q11002MP 605 nm, or Q11032MP 565 nm) Thermo Fisher scientific)) was then incubated at three-fold higher concentration than the protein:primary solution for ~ 30 minutes on ice.

When labelling the p44/p62 complex only when labelled at a high initial concentration did p44/p62 molecules show abundant binding to DNA. The p44/p62:primary antibody complex was labelled with a 1.5 molar excess of QDOTs (1.5 fold concentration of IgG). This lower excess of QDOTs was used as imaging single molecules was very challenging with high background fluorescence in solution, and on the flow cell surface. This means there is unlabelled p44/p62 in our assay. The final concentration of p44/p62 used was 5 nM labelled, 15 nM total. An excess of QDOTs ensures only a single fluorophore is labelling each protein (287). Protein:primary:secondary were then diluted in imaging buffer ready for introduction into the flow chamber.

2.16 Single molecule fluorescence imaging

Imaging was performed using a custom-built oblique angle fluorescence (OAF) microscope (288) capable of imaging multiple different colour fluorophores simultaneously. OAF uses an inclined beam to penetrate deeper into solution than epi-illumination; increasing the signal/background ratio and enabling clear imaging of fluorescently labelled single molecules in the imaging plane. The microscope is built on an Olympus IX50 frame. For QDOT imaging, a continuous-wave 20 mW 488 nm DPSS laser (JDSU) was focused through a 100x 1.45 N.A. objective to achieve OAF at the sample. UvrA-mScarlet was excited using a 561 nm diode OBIS LS laser (Coherent, USA) optically combined with the 488 nm laser (**figure 2.8**). Images were recorded onto an OrcaFlash 4.2 camera (Hamamatsu) using 1x1 binning at a frame rate of 10 frames per second (fps) unless stated otherwise. Pixel size was determined using a calibration slide (Graticules ltd, UK) as 63.2 nm.

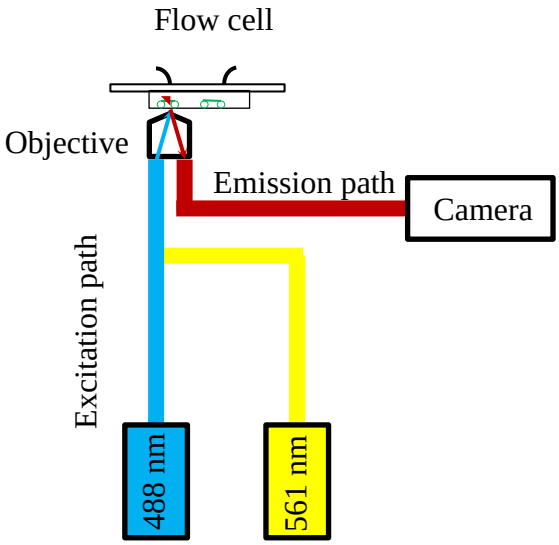


Figure 2.8 Simplified fluorescence microscope setup. Two lasers, 488 and 561 nm are combined into an excitation path that is focused through an objective at an oblique angle through the flow cell. The light excites fluorophores on the DNA (red triangle) that emit light at a different wavelength, collected through the objective and into an emission path to a camera.

2.17 Data analysis

2.17.1 Colocalizing multiple fluorophores

To determine if two QDOT labelled objects colocalize, a fiducial marker is used to correlate spatial position between the slightly offset fluorescence channels. Fiducial markers were collected for every flow cell imaged that requires dual colour analysis; frequently consisting of a coverslip surface heavily decorated with QDOTs and unidentifiable contaminants (dust, particles etc) that fluoresce in every channel. Using these points of alignment a custom written ImageJ macro aligns the channels to make a composite image that can be judged as colocalizing or not by eye. Colocalizing diffusing molecules were easy to identify as their kymographs displayed the same trajectories. For static molecules the fiducial marker is important to ensure accurate spatial accuracy. To prevent bleed-through light affecting experiments, band pass filters were used on the 565 and 655 nm channels to reduce the risk of stray or bright light from another channel being counted as a different molecule.

2.17.2 Kymographs

To analyse the images, videos were transformed into kymographs using ImageJ (289). Videos were projected as a Z-stack by maximum intensity to identify the tightrope position between beads. A line was then copied from the Z-stack to the original video and a kymograph generated. A kymograph in the context of DNA tightropes is a transformation of the position along the DNA tightrope to the Y axis, through every frame (time) – the X axis (shown graphically in **figure 2.9**). The intensity value of fluorescence in each frame is plotted on the X vs Y coordinates. A 5 pixel rolling ball background subtraction was applied to kymographs to average the intensity of bright and dim binders; the contrast was increased for easier visualisation of binders. Any molecules that looked anomalous, aberrantly bright, touched the edge of the kymograph, or that stayed bound for longer than the video were excluded from analysis.

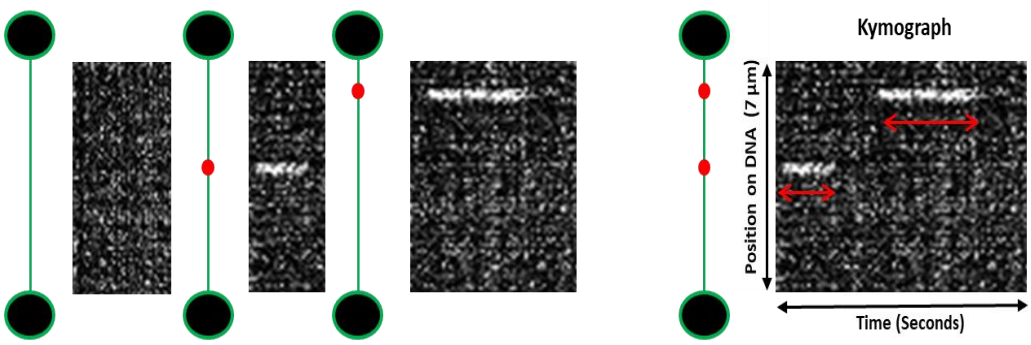


Figure 2.9 Kymograph transformation of protein binding. A diagram of DNA tightropes (black and green structures) with fluorescently labelled proteins (red) binding. The binding of the protein correlates to a streak in time (X) at a position (Y) in the video. The end result shows two proteins (streaks) binding and releasing at different positions on the DNA.

2.17.3 Data fitting

All data were fit in Microsoft Excel using sum of square difference fitting with the statistics add in package and built in solver function. Additional analysis was performed in Sigma-plot to determine fitting errors using the parameters from Excel fits. Equations were used based on previous studies.

2.18 Single molecule imaging and analysis

2.18.1 Data analysis for UvrA mScarlet

For DNA tightrope experiments, UvrA-mScarlet was diluted to 2 nM final concentration in ABC buffer supplemented with the nucleotide of choice, and 10 mM fresh DTT. UvrA-mScarlet was then imaged using a 561 nm laser between 1-5 mW power. Once a suitable tightrope was identified by visualising binders along a one-dimensional axis, the area was photobleached with 50 mW laser power for approximately 30 seconds. Photobleaching reduces the background signal from molecules attached to DNA and the surface. Immediately after photobleaching a video is collected, and therefore only capture ‘fresh’ binders from solution. This enables us to identify molecules that bind and release within one video. Videos were collected with a 300 ms exposure time for around 300 frames with 2x2 binning to avoid over- or under sampling the camera.

2.18.2 Lifetimes of individual UvrA molecules

Videos were transformed into kymographs as described. The bright patches on the kymograph indicate proteins binding to, residing on, and then releasing from DNA. The streaks intensity is that of the fluorescently labelled UvrA, while the length of the streak is proportional to the lifetime of the molecule on DNA. By knowing the exposure time, and measuring the length of streaks, we are able to determine the lifetime of individual UvrA molecules on DNA. The occurrence of proteins releasing from DNA follows a Poisson distribution, as the events are independent of each other, occur randomly, and have a constant rate (109). A Poisson process can be described by an exponential fit. Lifetimes of individual molecules in each condition were collected and plotted as cumulative frequency graphs. Plotting data as cumulative frequency removes any bin size dependency seen in histogram distributions. Data were fit to a single or double exponential equation where the constant *k* is the detachment rate constant for the population of molecules.

2.18.3 Single versus double exponentials

The requirement for single versus double exponential fits was determined using several methods:

- Residual plots show the difference between the experimental data and the fit described by the equation. A plot of the modelled data versus the experimental data shows how similar they are. The model that is closer to zero on the plot describes the data best.
- As the attached lifetimes are exponentially distributed in normal space, taking the log of the Y axis (frequency) should linearise the data if a single exponent is involved. Any non-linearity suggests an additional component is required to fit the data.
- An F-test compares the sum of squared differences between two models fitting the data. If a simpler model with fewer degrees of freedom, i.e. single exponential describes the data, the F-test ratio is around 1, as only the additional freedom parameters contribute to an increased sum of square differences. Conversely, if a more complex model, e.g. a double exponential describe the data, the F-test ratio will be significantly larger than 1, and suggests this model is correct. The F-test ratio is determined using the following equation:

$$\mathrm{F-test\ ratio} = \frac{(SS1 - SS2)/(DF1 - DF2)}{(SS2/DF\ 2)}$$

Where SS1 and SS2 are the sum of square differences of the simpler and more complex model respectively. DF1 and DF2 are the degrees of freedom of both the simpler and more complex model respectively.

2.18.4 Photobleaching experiments

To determine if photobleaching of the fluorescent protein mScarlet affected our observed attached lifetimes we measured durations of attachment for the longest associating nucleotide condition, UvrA-ATPγS at different laser powers. Videos were collected with a 300 ms exposure time for 90 seconds using 2x2 binning.

2.19 Data analysis for N-p44, p62, p44/p62, and XPD

2.19.1 Imaging conditions

Proteins were diluted into imaging buffer prior to being introduced into the flow chamber. Imaging buffer was supplemented with 1 mM fresh TCEP. For experiments with ATP, 1 mM final concentration was added to imaging buffer. For the salt experiments, the KCl concentration was altered by addition of a concentrated stock to the noted final concentration. Videos of N-p44, p62, p44/p62, XPD were obtained using the OAF system. The majority of videos were captured with 100 ms exposure (10 frames per second) for various time periods from 30 seconds to 5 minutes. Videos of up to 30 minutes were collected using 1 second exposures. Videos were transformed into kymographs as described previously. Diffusing molecules were clear from the kymographs, but an arbitrary rule of determining diffusers is used in our lab; proteins must move 3 pixels over three consecutive frames to be determined as diffusing (108, 113)

2.19.2 Rotation-coupled diffusion

The method and equations used to determine the diffusion limit of a protein tracking the helical backbone of DNA are described during chapter 4, and are derived from Schurr 1979 (269).

2.19.3 Diffusion constant

Diffusion constants of p44/p62 were determined using mean-squared displacement analysis (MSD) (109). Protein diffusion on DNA is modelled using a random walk. On DNA tightropes if a protein binds at position $x = 0$ and takes 10 random direction 'steps' of 1 base pair forwards (towards $+x$) or backwards (towards $-x$), the average position of the protein is zero; as the probability of moving in either direction is equal from every position. Therefore, a better measure of diffusion is how far the protein moves from the start point over time, described by the mean-squared displacement. Squared numbers are always positive, eliminating negative values that average to zero movement. As the number of steps increases (proportional to time (t)), the displacement of a molecule increases with the square-root of time. These observations lead to the MSD equation, where MSD is related to the diffusion coefficient, D, that describes the motion of the protein in one-dimension, and a time component:

$MSD = 2Dt^{\alpha}$

Diffusion can be obstructed by physical or chemical barriers, therefore a diffusive exponent (α) is added to the equation to describe diffusion (290). A diffusive exponent of 1 shows unobstructed random walking, whereas $\alpha < 1$ suggests obstructed diffusion, and $\alpha > 1$ shows directed motion.

Kymographs showing diffusion were isolated and fit to a Gaussian distribution, super-resolving molecule position through time. Each frame position of every molecule was filtered for an R^2 value of > 0.7. Using a custom-written macro the average displacement per frame was plotted against time for incremental time windows. The diffusion constant can be determined from the MSD equation (above). Diffusion constants are distributed normally in logarithmic space, so taking the log of this equation gives $\text{Log}(MSD) = \text{Log}(2D) + \alpha\,\text{Log}(t)$ which fits to a straight line. Plotting MSD vs time gives a straight line where the gradient = 2D (**figure 2.10**). The entire MSD cannot be fit as eventually the plot deviates from linearity (109, 113). In the analysis here at least the first 10% of the data are fit. The number of data to include was determined using R^2 values of fits, R^2 values < 0.7 were discarded. To obtain the diffusive exponent, α, Log(MSD) vs Log(time) is plotted where the gradient of a linear fit gives α (**figure 2.10**). This value describes the diffusive behaviour of a molecule (109, 113, 290). Around 30 individual kymographs were analysed for D and α. A plot of diffusion constants versus α shows a normal distribution of the values, and allows identification of

any sub-populations that may require further fitting. The diffusion constants and errors were averaged in log space and propagated to normal space. Average α values and S.E.M were determined in linear space.

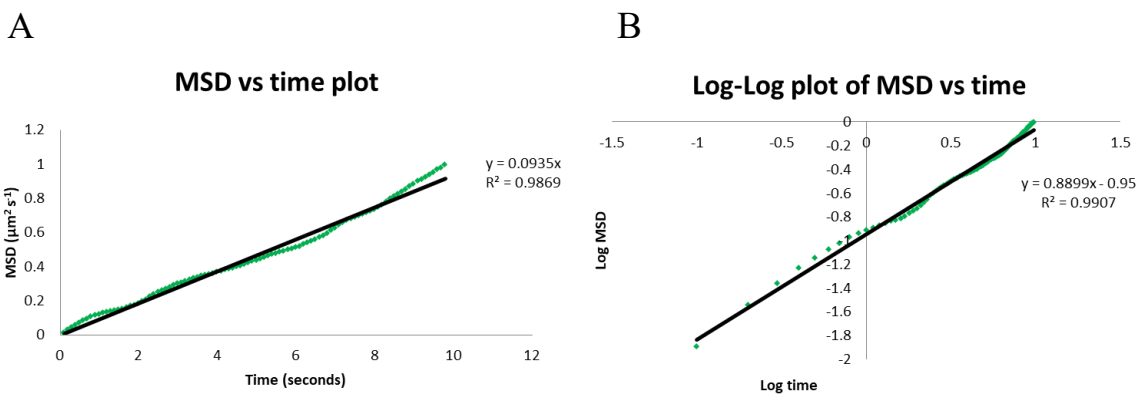


Figure 2.10 Diffusion constant and alpha plots.
A) MSD vs time where the gradient of the linear fit (black) is 2D. The R^2 value must be above 0.7 and at least 10% of the data included for analysis.
B) Taking the data in the left window and plotting the logs (Log MSD vs Log time). The gradient of a linear fit gives the α value.

Chapter 3

Understanding the coupling between DNA damage detection and UvrA's ATPase using bulk biochemical and single molecule kinetics

Introduction

3.1 The role of ATP in NER

Almost every step of the prokaryotic NER pathway requires nucleotide binding or hydrolysis in order to successfully repair DNA (78, 98, 291). While the current model of NER shows these steps are dependent on nucleotide, the exact mechanism behind each process is still largely unknown.

3.2 Structure and function of UvrA

UvrA binds to DNA as a homodimer that occurs through head to head interactions in the N-terminal domain (**figure 3.1**) (292), dimerization is crucial for UvrA to bind DNA, and is heavily dependent on the nucleotide state of UvrA as well as the concentration of protein (101, 102, 106, 292, 293). UvrA then attaches to double-stranded DNA with a dissociation constant of ~ 10 nM (293). As discussed previously, UvrB and UvrC alone have no significant affinity for DNA damage (96, 97) whereas UvrA has a high affinity for both UV-irradiated DNA and single-stranded DNA (97, 106, 112, 281). Sequence analysis of the *uvrA* gene product showed two 'Walker A' motif ATP binding cassette ATPase sites in each UvrA monomer, termed the N-terminal (residues 31-45) and C-terminal (residues 640-654) site respective to the amino acid sequence (294-296). These sites and surrounding regions are structurally similar, and are thought to have arisen from gene duplication (295). In the three-dimensional structure of dimeric UvrA, the N-terminal ATPase domain of one UvrA is in close proximity with the C-terminal ATPase site. Furthermore, all four sites in the dimer are positioned beneath the DNA binding interface that runs across the surface of the protein (**figure 3.1**) (281, 297). Each of the two ATPase motifs is interrupted by a zinc finger ((two in the N-terminal, one in the C-terminal) **figure 3.1 & figure 3.2**) that sits in close proximity to the DNA (295). Previously, the zinc fingers of UvrA have been implicated in damage discrimination and DNA binding (298-300). Each zinc finger in UvrA is coordinated by four cysteine residues (298) and it is known that mutations in the N-terminal zinc fingers do not affect NER (100, 299, 300). However, disruption of the C-terminal zinc finger compromises bacterial survival when challenged with UV and leads to insoluble protein being produced (295, 299).

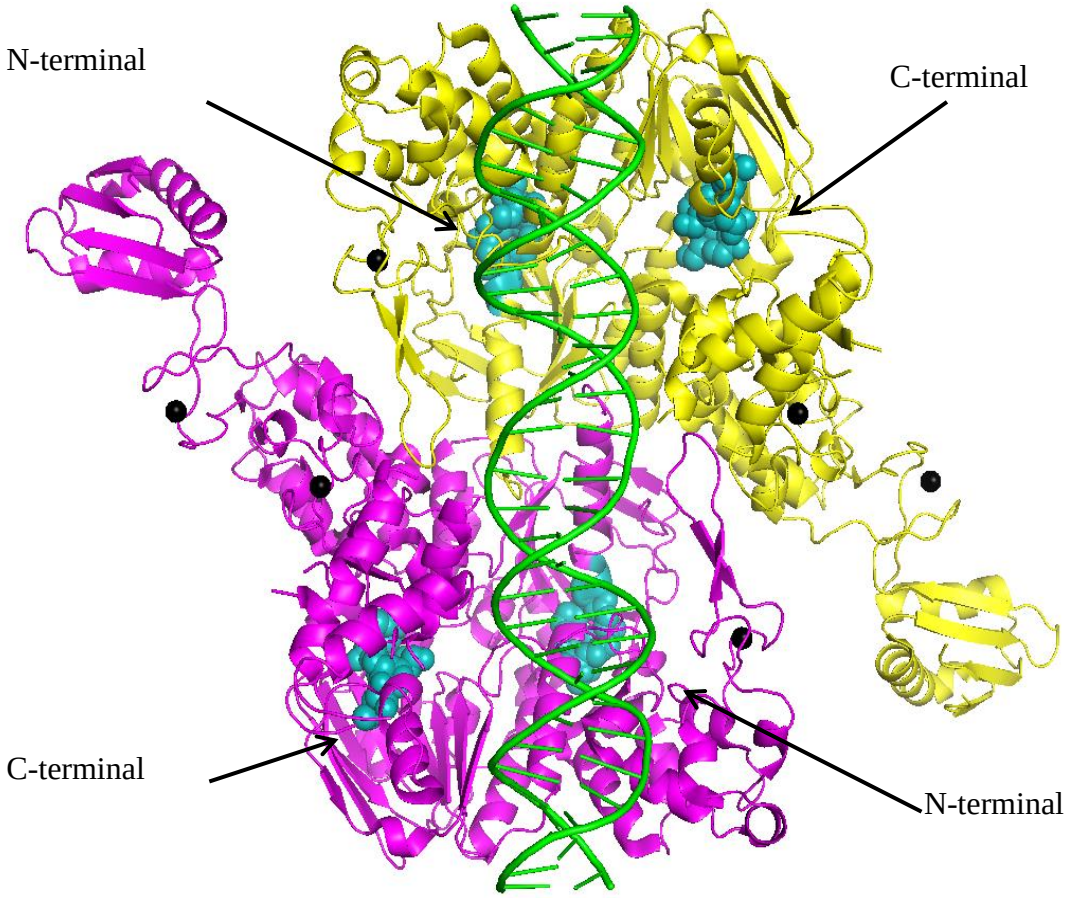


Figure 3.1 The crystal structure of UvrA. One monomer of the homodimer in yellow, one in magenta. DNA (green) can be seen bound across the centre of the dimeric protein, with the ATPase site residues (teal spheres, labelled) located beneath the DNA. Zinc atoms are shown as black spheres and indicate the position of the zinc fingers in UvrA. Structure generated in PyMol using PDB:3PIH (281).

Crystal structures of UvrA bound to ADP (PDB:2R6F), or nucleotide free (PDB:3PIH) are relatively similar; however the C-terminal zinc finger shows a drastic displacement between these structures (281). In **figure 3.2** the C-terminal zinc finger is located away from the DNA, however in an ADP bound structure (297) the zinc finger is superimposed to clash with the DNA. Mutating the C-terminal zinc finger leads to a loss of specificity for DNA damage, but an increased affinity for DNA damage, most likely related to a conformational change in the protein during damage verification process (281, 300, 301). This was further supported by a recent study that utilised crosslinking of UvrA in the presence of different nucleotides and DNA substrates to look at zinc finger movement (discussed in section 3.3) (302).

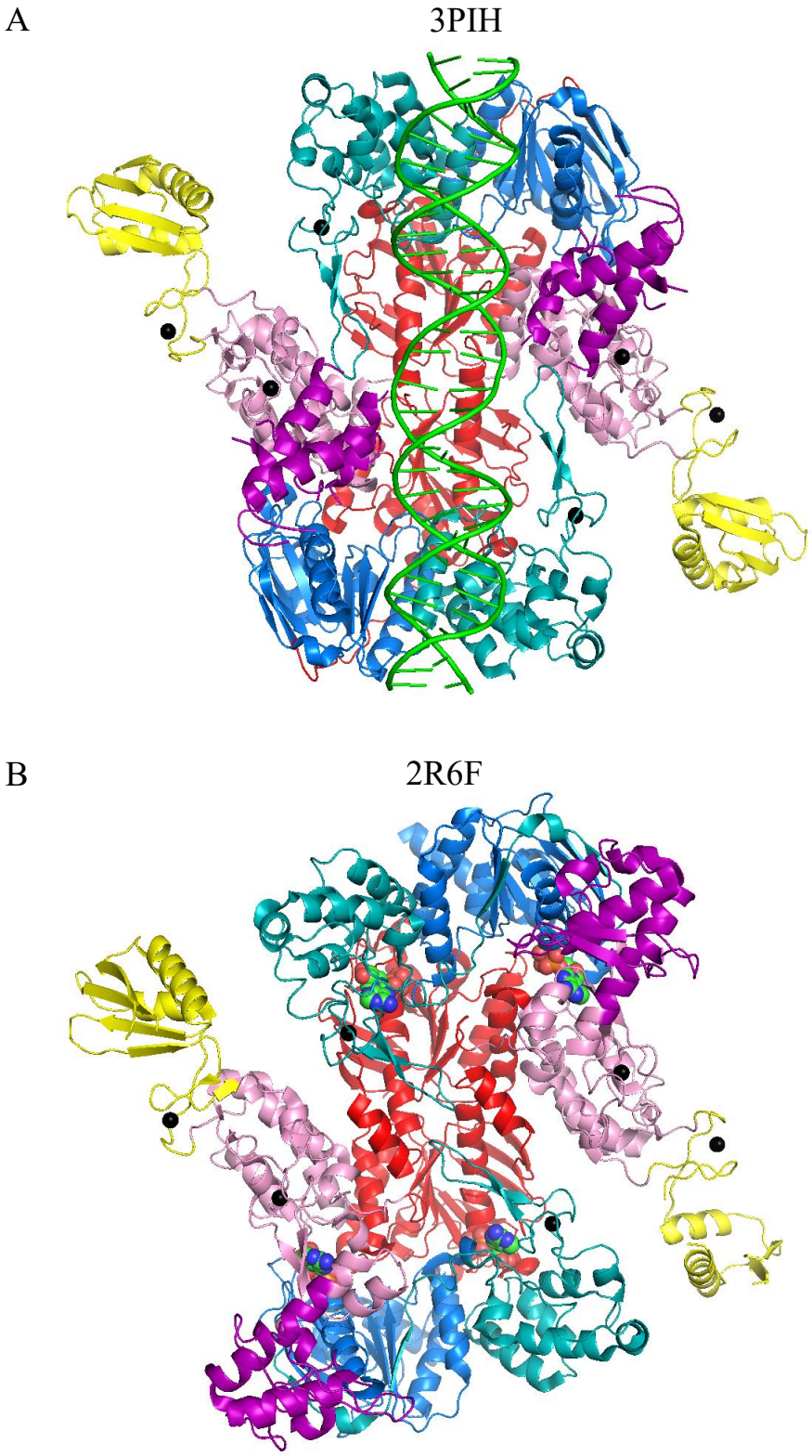


Figure 3.2 Zinc finger positioning between structures. Using the Uniprot gene sequence for the *Thermatoga maritima* sequence (Q9WYV0), two crystal structures of UvrA are colour coded by domain to show movement between A) 3PIH from *Thermatoga maritima* that contains no nucleotide, but is bound to double-stranded DNA containing a fluorescein lesion that acts as damage for NER (281) and B) 2R6F from *Bacillus stearothermophilus* that is bound to ADP (shown as spheres that are bound in all four ATPase sites) and not DNA (297). There is a clear difference in the position of the C-terminal zinc finger (contained in signature domain II (teal)) between 3PIH where the zinc finger supports the bound DNA, and 2R6F where the zinc finger clashes with the position of DNA, suggesting the ADP bound state is unable to bind DNA. Figure created in PyMol.

3.3 Conflicting evidence on UvrA's ATP hydrolysis

3.3.1 Mutating each ATPase site affects the other, and UvrA's function

The ATPase of UvrA has a huge impact on the efficiency of NER both *in vivo* and *in vitro* (99, 120, 293). Numerous studies have shown that binding or hydrolysis of nucleotide affects UvrA's ability to bind DNA (90, 101, 102, 106, 291). It is generally agreed that both ATPase sites in a UvrA monomer are cooperative (120, 292, 293, 297), but their ability to function independently is disputed. The N-terminal and C-terminal ATPase domains of UvrA have been isolated and purified, with both showing ATPase activity, but when recombined they have a reduced V_{max} compared to wild-type UvrA, highlighting this cooperativity (292).

Myles and Sancar first showed that mutating the catalytic lysine residue in each Walker A site inhibits ATP hydrolysis (293). Since then, many publications have been based on these mutations (99, 119, 120, 302, 303), despite evidence that mutating the C-terminal ATPase site causes protein precipitation and inclusion body formation in cells (293). Clearly this mutation disrupts protein structure, however, *in vitro* catalytic rate constants have been determined (99), and both ATPase UvrA mutants have been imaged *in vivo* (119). Furthermore, mutating the N-terminal ATPase domain destroys cooperativity between the two sites (292, 293), supported by Wagner et al where they observed a complete loss of all ATPase activity upon mutating either site (120). Both K37A (N-terminal) and K646A (C-terminal) mutants display UV sensitivity when complementing wild-type UvrA *in vivo*, and are defective in repair *in vitro* (99, 119, 120, 293). Furthermore, it is unclear how both mutants fold or affect protein dimerization and the zinc fingers that are associated with UvrA's ATPase and are critical for DNA repair (100, 297)

3.3.2 K37A (N-terminal ATPase site, proximal site)

Studies using the K37A mutant attempted to reconcile UvrA's activities with the active C-terminal ATPase site. K37A mutants have been shown to either abolish all of UvrA's ATPase activity (120, 303), or retain ATPase activity (293), even when digested into just the C-terminal domain (292). Although hydrolysis in the N-terminal ATPase site was abolished by the K37A mutant, this site possibly still bound either ATP or ADP (303). Hydrolysis in the C-terminal ATPase site is critical for full activity of UvrA (120, 293), with this site having a higher affinity for ATP than the N-terminal ATPase site (292, 303), which is thought to be responsible for UvrA releasing from undamaged DNA (99) as well as being important for protein dimerization (106, 119, 292, 293). UvrA appeared to be relatively unaffected by K37A when compared to the K646A mutant, with K37A mutants demonstrating similar purification and UV survival to wild-type UvrA (293). Although this mutant demonstrated similar UV survival to wild-type UvrA, *in vitro* complementation of incision was poor in the hands of Myles and Sancar (293). Conversely, Wagner et al found K37A supported incision on short DNA substrates, but was inhibited on longer DNA substrates (120) this can be speculated to be caused by inhibition of UvrA's ability to release and then rebind DNA. However, K37A has been shown to release from DNA, but is unable to discriminate damaged DNA (99), suggesting that it is not binding and releasing, but increased competition from undamaged DNA that explains this observation. Inconsistent with these data, *in vivo* studies showed K37A increases the number of immobile UvrA molecules bound to DNA compared to wild-type UvrA (suggesting impaired release), they also showed K37A can discriminate UV damage, further increasing the immobile fraction of protein and directly disagreeing with Thaligalingam and Grossman (99, 119).

3.3.3 K646A (C-terminal ATPase site, distal site)

Similar to the N-terminal ATPase site, the activities of the C-terminal ATPase site are disputed, with conflicting evidence complicating interpretation of the data. The main discrepancy with data concerning the mutation of the C-terminal ATPase site are reports of insoluble protein, aggregation, and inclusion body formation (120, 293), however not all studies note difficulties with this mutant (99, 119). In addition to the clear perturbation of normal protein folding, K646A does not complement UV survival *in vitro* or *in vivo* (293). Furthermore, k_{cat} values show K646A mutants have no overall ATPase (303), but some reports dispute this and show K646A has ~33% of wild-type UvrA ATPase activity, and

maintains cooperativity between the ATPase sites (99, 293). The N-terminal ATPase site is thought to have a high affinity for ADP, after being observed co-purifying with nucleotide (120, 292). Conversely, K646A leaves the protein with no affinity for ATP or ADP, showing that the C-terminal site has the higher affinity for nucleotide (303). Experimentally, K646A loses the ability to discriminate DNA damage, instead forming a long-lived stable complex on DNA that does not release (99, 119). In addition to this, a fast diffusing species is also observed with this mutant, ascribed to a reduction in DNA binding caused by reduced dimerization (119). Together these data suggest the N-terminal ATPase site is linked with damage recognition, but does not possess the ability to dissociate from DNA. In addition, the N-terminal site is required for loading UvrB onto the DNA, in a process that requires nucleotide hydrolysis at the N-terminal site, but while nucleotide is bound to the C-terminal ATPase site (119). A dual mutant of both catalytic lysine residues has no ATPase activity and demonstrates the highest DNA binding, as it is unable to release from DNA (99, 293).

3.3.4 Alteration of UvrA's ATPase upon addition of DNA

A peculiar observation in these studies is the various responses of UvrA's ATPase upon the addition of DNA. Some groups report a drop in ATPase activity upon addition of undamaged or UV damaged DNA (99, 120). However, there are reports of ATP turnover being accelerated by the addition of DNA, but no further stimulation in the presence of damage (97, 102). Furthermore, a recent study using thermophilic UvrA observed DNA stimulated ATPase activity, and further acceleration upon the addition of UV damaged DNA (303). The discrepancies between these data are unclear, but are most likely caused by differences in protein quality and organism, assays, and analysis.

3.3.5 Source of conflicts within these studies

The evidence described here clearly shows direct conflicts between studies where the same gene product is purified in a similar manner, sometimes in identical assays. However, it is apparent that experimental differences, temperature, contaminating ATPases, and the quality of purified protein are all potential sources of error. There are also differences between the damage used; cholesterol, fluorescein, psoralen, and CPDs have all been used in these assays, and could affect UvrA's ATPase differently. Clearly, the K646A mutant disrupts UvrA's structure, leading to aggregation and inclusion body formation in *E. coli*, therefore experimental evidence from this mutant must be treated with care. In addition, disruption of the cooperativity between the two ATPase sites has an unknown effect on the protein and its activity. Some groups work with *E. coli* UvrA, whereas others use thermophilic organisms for the protein stability. However, these different gene products require significantly different temperatures for optimal activity, and this will clearly influence rate constants and assay. The approach taken to study the individual ATPase domains of UvrA carries with it some concerns. For example, Myles and Sancar cleave UvrA into the N- and C-terminal domains with a collagenase site in the middle of the protein, although they observe ATPase activity in both purified domains, the additional effects upon UvrA are likely significant (292). Likewise, the ATPase is known to be linked with movement of the zinc fingers (302) that regulate DNA binding and damage discrimination, and therefore a specific observation is not necessarily linked to the ATPase site, but could be affecting the ability of UvrA to interact with DNA.

3.3.6 Two-site hypothesis

A recent study used crosslinking to determine the structural conformation of UvrA (302). When in a closed conformation a higher percent of UvrA was crosslinked, whereas if UvrA was mostly in an open conformation the protein is less efficiently crosslinked. They found that wild-type UvrA is efficiently crosslinked in the presence of DNA and ATP, where it forms a closed conformation on the DNA. When they mutated the N-terminal ATPase site they found crosslinking was enhanced, suggesting that this conformation is closed more often on the DNA. Conversely when they mutated the C-terminal ATPase site they found there was less crosslinking, suggesting the protein is locked in an open conformation – they then relate this to the fact that the zinc fingers in the protein are unable to move when this ATPase site is mutated (302). This study relates the direct effect of the ATPase site mutants

on the zinc fingers, and conformational state of UvrA when bound to DNA, reconciling some of the observations described here. A model proposed by Myles and Sancar (293) and more recently built on (119, 120) suggest a two-step mechanism for damage searching and recruiting UvrB to DNA. This hypothesis suggests that the N-terminal and C-terminal ATPase sites turnover and have separate activities within UvrA, but are cooperative (293). An *in vivo* study using K37A and K646 investigated the two-site hypothesis and suggested that the C-terminal ATPase site is involved in binding and releasing from DNA while the N-terminal ATPase site is necessary for recruiting UvrB onto the DNA from solution (119). The perpetual use of both K37A and K646A mutants to understand UvrA's ATPase is not fully justified. Although some studies using these mutants are important for discerning the need for binding, or hydrolysis of nucleotide in certain activities, the conflicting outcomes complicate accurate interpretation. Therefore, the role of ATP in UvrA's activities has not yet been clearly defined. Questions still remain as to how other nucleotide cofactors affect UvrA's behaviour, the mechanism of how nucleotides affects both damage discrimination and dimerization, and what this all means for the whole mechanism of NER in *E. coli*.

3.4 Using single molecule and biochemical techniques to examine UvrA's ATPase

In this chapter we avoid the study of K37A and K646A mutants and instead use fully functional wild-type *E. coli* UvrA labelled with a monomeric scarlet fluorescent protein (280). We investigate how UvrA's ATPase activity is controlled and how it contributes to DNA repair. Using a combination of biochemical and single molecule techniques the effect of different nucleotide cofactors and DNA substrates are examined for their impact on UvrA's ATPase.

Results

3.5 UvrA-mScarlet complements UvrA knockout cells *in vivo*

It has previously been shown that cells lacking UvrA are sensitive to UV irradiation, but ectopic expression rescues their UV sensitivity (293). To investigate whether UvrA-mScarlet is active in DNA repair, and can complement $UvrA^{WT}$ knockout cells (ΔUvrA), *in vivo* UV complementation assays were performed. ΔUvrA cells were transformed with UvrA-mScarlet, or mScarlet vector with the UvrA gene removed (kindly provided by Dr Alex Moores). Transformed cells were subjected to 5 J/m^2 UV (254 nm) irradiation or no UV and plated. Plates were incubated overnight in the dark to prevent photolyase repair interfering with the experiment. Images of the plates are shown in **figure 3.3**. As expected, ΔUvrA cells and ΔUvrA cells transformed with empty mScarlet vector demonstrated diminished survival when challenged with UV irradiation. ΔUvrA cells transformed with UvrA-mScarlet fully rescued UV survival, suggesting that UvrA-mScarlet is active in DNA repair *in vivo*.

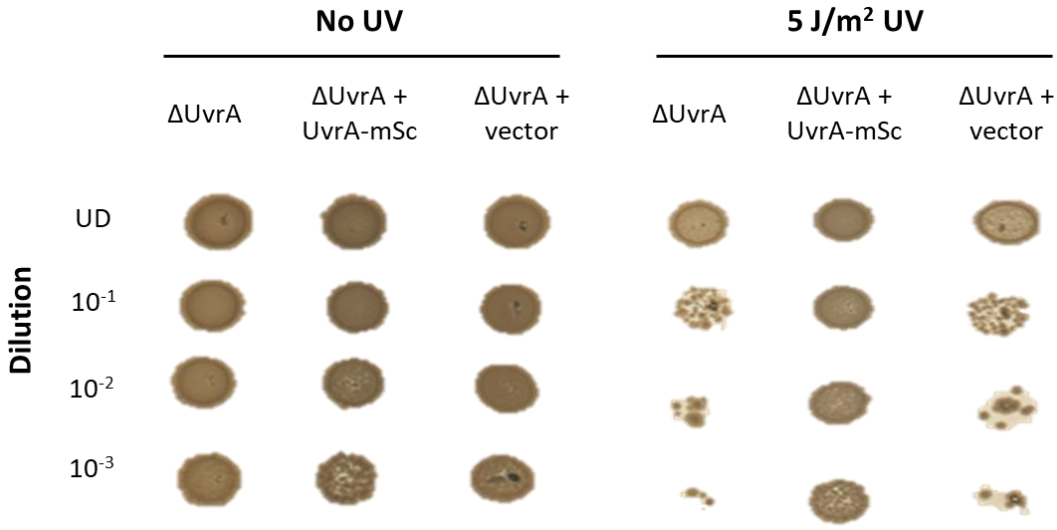


Figure 3.3 UvrA-mScarlet complements UvrA knockout cells. Serial dilutions of wild type UvrA knockout cells (ΔUvrA), ΔUvrA cells transformed with UvrA-mScarlet (ΔUvrA + UvrA-mSc), or empty mScarlet vector without UvrA (ΔUvrA + vector). In the absence of UV irradiation there is no inhibition of growth after overnight incubation. When the plate is subjected to 5 J/m^2 irradiation there is inhibition of growth in ΔUvrA and ΔUvrA + vector cells, but not in the ΔUvrA cells transformed with UvrA-mScarlet. UD = Undiluted solution of 50 uL 0.6 OD600 cells, with serial dilutions of this shown.

3.6 Purified UvrA has a DNA-stimulated ATPase

Each of UvrA's two Walker motif ATPase sites has been linked with specific activities. A particularly intriguing result reported a 50% reduction in UvrA's ATPase activity upon the addition of undamaged DNA, however there was no increase in ATPase in the presence of damaged DNA. That is, DNA damage does not affect the ATPase of UvrA alone, but undamaged DNA reduces the ATPase rate by 50% (120). In order to try and understand these results, the steady state ATPase of purified UvrA-mScarlet was measured using a NADH-linked assay to calculate k_{cat} per UvrA monomer (See section 2.5).

The addition of plasmid DNA (pUC18) to UvrA stimulated the ATPase activity from a k_{cat} of 0.71 ± 0.05 ATP·UvrA^{-1} s^{-1} in the absence of DNA to 1.08 ± 0.09 ATP·UvrA^{-1} s^{-1} ($P < 0.0001$). Surprisingly, this is the opposite of the reduction in ATPase activity seen in Wagner et al (120), despite the use of the same DNA substrate. To investigate if the plasmid substrate was the problem, several different substrates were examined. A short 50 bp double-stranded DNA oligo (F26,50) and the lambda phage genome were substituted in the same assay, both demonstrated identical stimulation of UvrA's ATPase (**figure 3.4**). Combined averages for the three DNA substrates show the steady state ATPase k_{cat} increases significantly ($P<0.0001$) from 0.71 ± 0.05 ATP·UvrA^{-1} s^{-1} to 1.07 ± 0.08 ATP·UvrA^{-1} s^{-1} upon addition of DNA (**figure 3.5**). The faster ATP turnover in the presence of DNA suggests a step in UvrA's kinetic cycle is accelerated upon binding DNA.

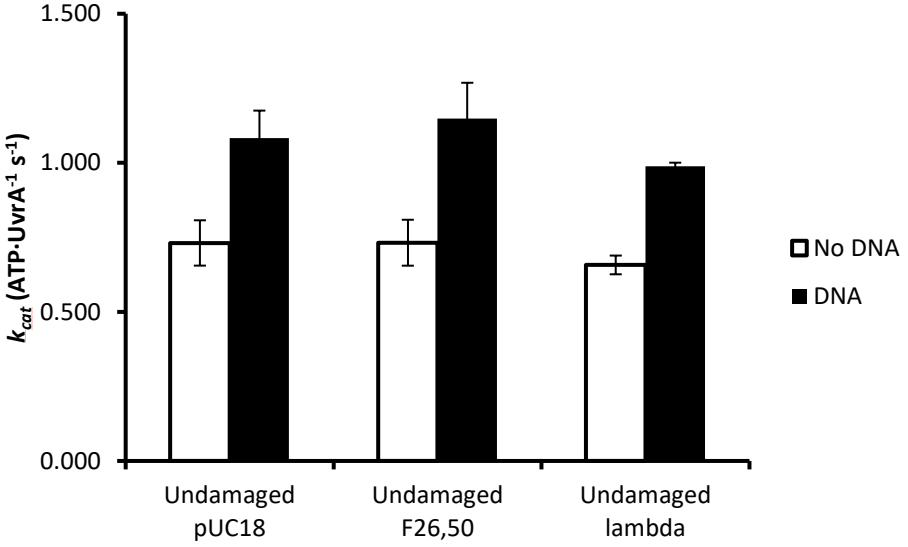


Figure 3.4 Steady state ATPase of UvrA on various undamaged DNA substrates. Data show k_{cat} value per UvrA monomer, per second in the absence of DNA (white) or after DNA is added (black) measured using a NADH-linked assay. The data show UvrA has a DNA-stimulated ATPase, but there is no difference between plasmids, short oligonucleotides and lambda DNA in the enhancement of UvrA's ATPase. Error bars show the standard deviation of three repeats.

3.7 DNA damage does not further stimulate UvrA's ATPase

To study the effect of DNA damage on UvrA's ATPase the NADH-linked ATPase assay was repeated in the presence of various damaged DNA substrates. Both pUC18 and lambda DNA were irradiated with 1000 J/m^2 of 254 nm UV light, randomly introducing CPDs and 6-4PP's into the DNA (estimated that CPDs form ~ 7 fold more than 6-4PP's at this wavelength (304)). Undamaged F26,50 was replaced with a 50 bp double-stranded oligonucleotide containing a fluorescein labelled thymine at the 26^{th} nucleotide, a substrate previously shown to be a target for NER (137). All three substrates stimulated UvrA's ATPase activity compared to the absence of DNA (**figure 3.5**). F26,50 (1.06 ± 0.142 ATP·$UvrA^{-1}$ s^{-1}), UV damaged pUC18 (1.13 ± 0.05 ATP·$UvrA^{-1}$ s^{-1}) and lambda phage DNA (1.03 ± 0.12 ATP·$UvrA^{-1}$ s^{-1}) compared to UvrA & ATP alone (0.71 ± 0.05 ATP·$UvrA^{-1}$ s^{-1} ($P < 0.0001$)). However, no damaged substrate increased the ATPase activity significantly more than its undamaged equivalent ($P = 0.9594$ for combined data) (**figure 3.5**). Both linear and plasmid substrates stimulated the ATPase of UvrA equally, indicating the DNA ends play no part in the observed activation.

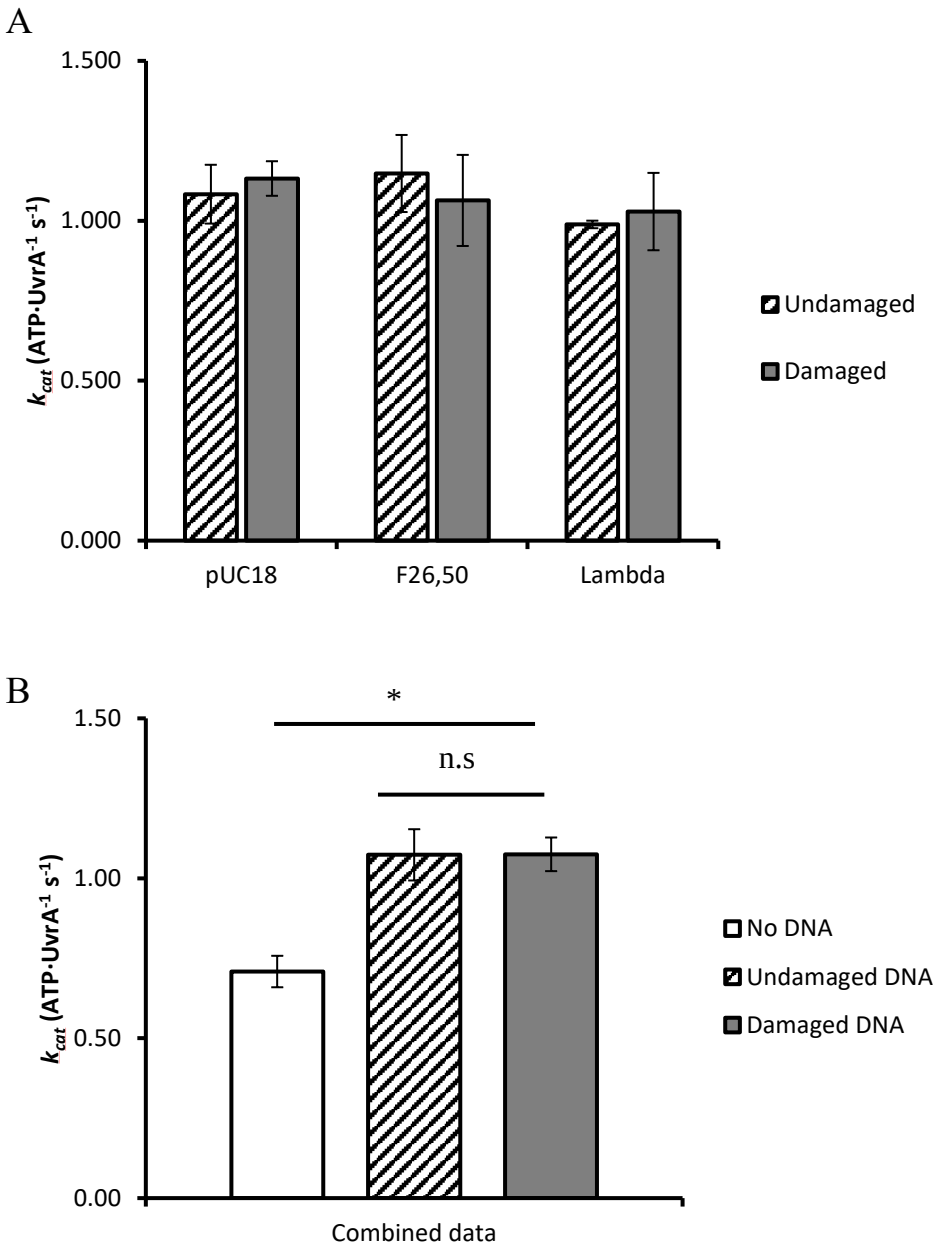


Figure 3.5 DNA damage does not further accelerate UvrA's ATPase.
A) k_{cat} values for UvrA monomers on undamaged (grey stripes) or damaged DNA substrates (solid grey) measured using a NADH-linked assay. There is no difference between the acceleration of UvrA's ATPase between undamaged or damaged DNA. Additionally there is no difference in the stimulation between a plasmid, short oligo, or lambda DNA substrate. Error bars show the standard deviation of three repeats.

B) Average values for the stimulation of UvrA's ATPase upon addition of undamaged DNA (white to grey stripes) or damaged DNA (white to solid grey). These data show DNA stimulates UvrA's ATPase, but DNA damage does not further increase ATPase activity. *P = 0.0001, n.s = not significant (P = 0.9594) using a Student t-test. Error bars represent the standard deviation between pooled data.

3.8 DNA damage is present in the steady state ATPase assay

The lack of further stimulation of UvrA's ATPase by damaged substrates is surprising considering UvrA's role as a damage sensor in NER. To ensure DNA damage was present in our assays we employed gel based assays. Using a DNA glycosylase that recognises CPD lesions and nicks the DNA 5' of the lesion (T4 endonuclease V (T4EV)), it is possible to confirm the presence of UV lesions on DNA substrates. For plasmids, irradiation with 254 nm UV induces CPD lesions. When lesions are nicked by T4EV there is a loss of plasmid supercoiling, visible on an agarose gel when run with a high concentration of ethidium bromide (**figure 3.6**).

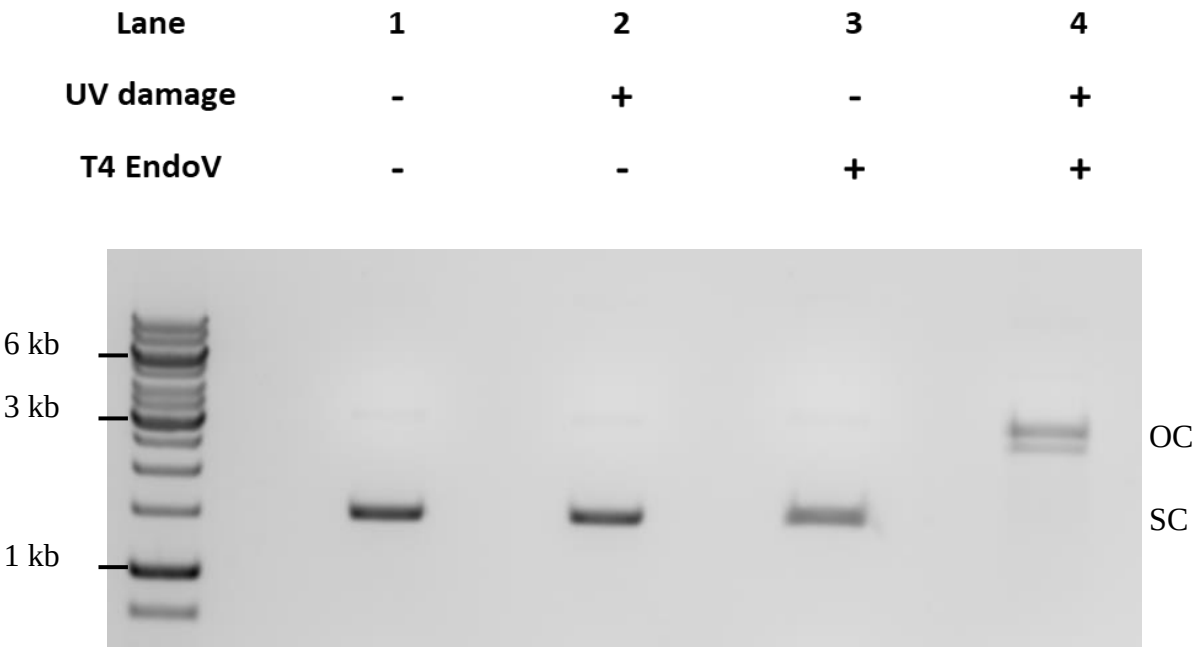


Figure 3.6 T4 EndoV incision assay
To demonstrate damage is present in the ATPase assay, pUC18 plasmid (2686 bp) was irradiated to 300 J/m^2 at 254 nm. To confirm the presence of CPD lesions, UV irradiated (lane 4) or undamaged pUC18 (lane 3) was digested with T4 EndoV and run on a 1% agarose gel in TAE. A high concentration of ethidium bromide (10 µg/ml) was used to induce plasmid supercoiling (SC (305)). T4 EndoV nicks at CPD lesion sites, causing a loss of plasmid supercoiling to open circle (OC (lane 4)). A clear loss of plasmid supercoiling is seen upon UV irradiation of pUC18 (lane 4) but not in the control (lane 3) showing that CPDs are present in the DNA at this level of irradiation.

Lambda DNA is linear so cannot be assayed in the same way. To overcome this, an assay that uses T4EV and a single-stranded nuclease (S1 Nuclease) that cuts at DNA nicks was devised. UV damaged lambda DNA was incubated with either S1 nuclease alone, or T4EV and S1 nuclease and run on an agarose gel (**figure 3.7**). If CPDs are present, T4EV will nick the DNA 5' of the lesion, which S1 Nuclease will then digest to cause a double-strand DNA break. As CPDs are randomly distributed on DNA, the single band on a gel will be lost, creating a smear of products in the lane (**figure 3.7**, lane 7). This assay also shows there are some nicks in native lambda DNA.

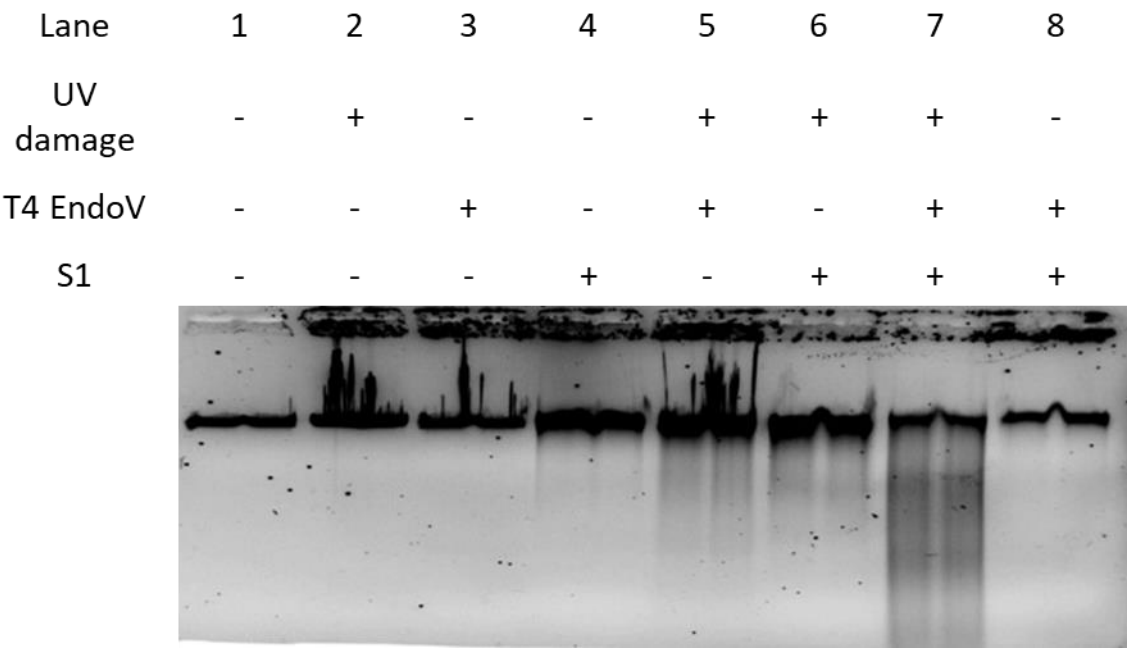


Figure 3.7 T4 EndoV and S1 nuclease assay in lambda DNA. Undigested lambda DNA runs as a high molecular weight product in an agarose gel. UV irradiation generated CPD lesions in the DNA at random locations, these are able to be cleaved by T4 EndoV to cause single-stranded nicks. S1 nuclease is able to cut at a single-stranded nick to generate a double-strand break. When UV irradiated lambda is cut with either T4 EndoV (lane 5) or S1 nuclease (lane 6) alone there is only slight smearing, showing UV irradiation is causing some nicks and breaks in the DNA. There is very little damage in lambda without irradiation (lanes 1, 3, 4, 8), or in the absence of either enzyme (lane 2). When T4 EndoV and S1 nuclease are incubated with UV damaged DNA there is a large increase in DNA smearing, demonstrating presence of CPD lesions in the irradiated DNA, when compared to the absence of UV damage (lane 7 vs lane 8).

These assays confirmed the presence of UV damage in the substrates used for the ATPase assay. We then estimated the number of damage sites per DNA molecule using previously published methods. The number of CPD lesions on a substrate was calculated using a formula of 0.0221 CPDs/kb/Jm^2 at 254 nm (306). The mass of DNA in the ATPase assays was kept constant at 50 ng, meaning both lambda and pUC18 have the same ratio of undamaged and damaged base pairs. F26,50 contains a fluorescein attached to the 26th base, giving one damage site per 50 bp oligonucleotide. When lambda and pUC18 were UV irradiated to 1000 J/m^2 it is predicted lambda DNA (48,502 bp) contains ~1000 CPDs per molecule while pUC18 (2686 bp) contains ~60 CPDs per molecule (306). In the ATPase experiments the concentration of damage in the reaction is approximately equal to the concentration of UvrA, and is unlikely to be saturated by UvrA.

Substrate	Base pairs	UVC dose (J/m^2)	CPDs/molecule	Undamaged base pair concentration (nM)	DNA damage concentration (nM)
pUC18	2686	1000	59	150	17
Lambda (λ DNA)	48502	1000	1072	150	17
F26,50	50	X	1	150	15.4

Table 3.1 Concentrations of DNA damage in the steady state ATPase assay. For each substrate, a fixed mass of DNA was used per reaction, equating to the same number of base pairs. The number of CPDs per molecule was calculated using the method of Douki et al (306). F26,50 contains a single fluorescein modified thymine in the centre of the oligonucleotide and is not subject to the random location of damage induced by UV irradiation.

3.9 Damage discrimination is marked by a change in affinity for ATP

To ensure purified UvrA-mScarlet is able to discriminate damage we measured UvrA's affinity for ATP in the presence of either undamaged or UV damaged pUC18 DNA. Using the NADH-linked assay, ATP was titrated from 50 µM to 2 mM and the ATPase rate measured at each concentration. Plotting these data and fitting them to a Michaelis-Menten relationship gives the maximum velocity of ATP hydrolysis (Vmax (k_{cat})) and K_M, the affinity of UvrA for ATP. **Figure 3.8** shows a clear decrease in UvrA's K_M for ATP from 195 µM in the presence of undamaged DNA to 60 µM with damaged DNA; however, k_{cat} remained unchanged at ~1.4 s^{-1}. The increase in affinity for ATP is approximately three-fold, and suggests that UvrA is able to discriminate DNA damage, and this is linked with a change in affinity for ATP. However, we observe no difference in the steady state ATPase between undamaged and damaged DNA.

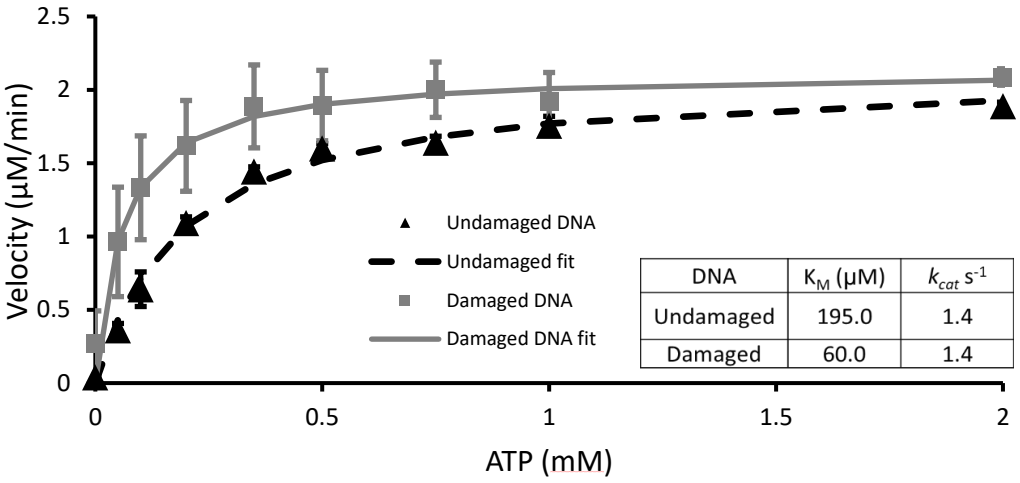
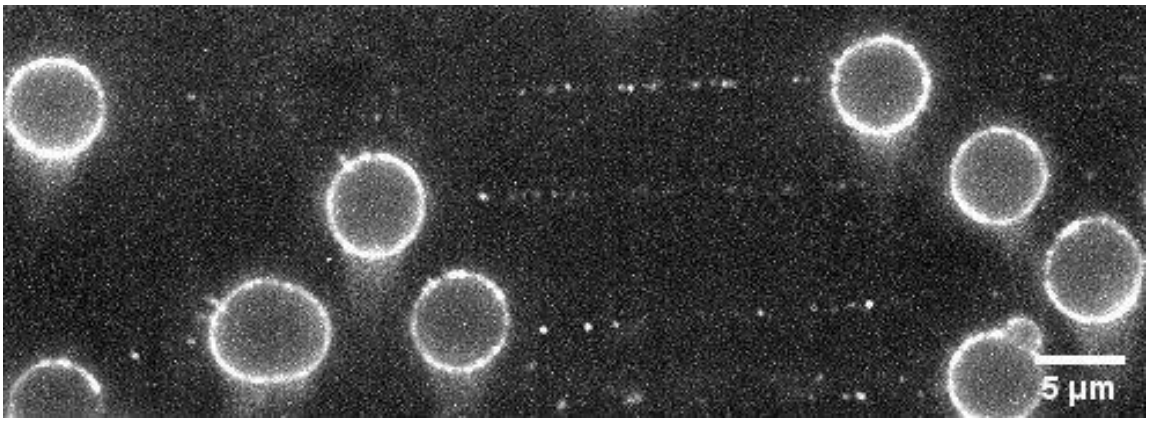


Figure 3.8 Michaelis-Menten plot of UvrA's affinity for ATP. ATP was titrated in the presence of UvrA with either undamaged (black triangles) or UV damaged pUC18 DNA (grey squares). Fitting the data to the Michaelis-Menten relationship (dashed and solid lines) gives a maximum velocity (k_{cat}) and an affinity of UvrA for ATP (K_M) in each condition. The parameters from each fit are shown in the table. In the presence of damage the K_M is decreased 3-fold compared with undamaged DNA, but k_{cat} is unchanged between conditions. Data are averages of 3 repeats plotted with standard deviation as error bars.

3.10 Single molecule imaging shows individual UvrA molecules binding to DNA

As we know UvrA's ATPase is accelerated by the presence of DNA, but not affected by the presence of damage we wanted to directly visualise UvrA's attachment to DNA. Using the *in vitro* DNA tightrope assay we investigated how UvrA's interaction with DNA is affected by the presence of ATP or UV damage. As we know the rate limiting step of ATP hydrolysis occurs on DNA, different nucleotides can substituted in the tightrope assay to attempt to recreate various kinetic states that cannot be done in the NADH-linked biochemical assay. The *in vitro* DNA tightrope assay is described in section 2.10 of this thesis, as well as extensively in Springall et al (283). Briefly, individual lambda DNA molecules are elongated and captured over silica beads immobilised in a microfluidic chamber. The UvrA construct used here contains a C-terminal mScarlet fluorescent protein that can be excited by a 561 nm laser, enabling spatial location of individual UvrA molecules on DNA, shown in **figure 3.9**.

A



B

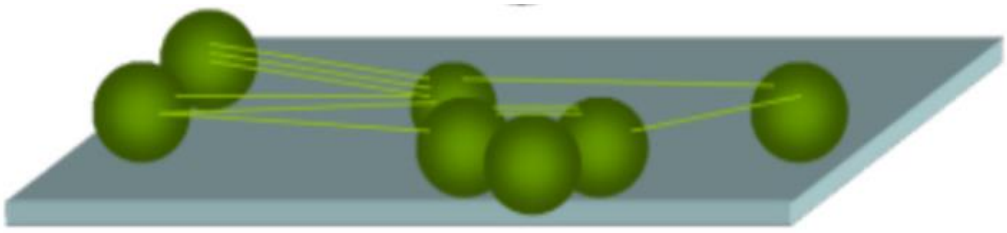


Figure 3.9 UvrA molecules bound to DNA tightropes. A) A Z-stack of frames combined show the position of multiple UvrA binders on DNA tightropes in a single image. B) A schematic of DNA tightropes formed between beads on a coverslip to highlight the likely position of the DNA seen in A).

3.11 UvrA uses a three-dimensional search on DNA

Videos of individual UvrA molecules binding to DNA were collected and transformed into kymographs (See section 2.16). It was apparent from the kymographs that UvrA molecules were binding and releasing from the DNA at random locations. Binding and releasing behaviour, displayed as horizontal streaks on a kymograph (**figure 3.10**) show that UvrA uses a three-dimensional search to find DNA damage. This is consistent with studies using thermophilic UvrA (109) and *in vivo* observations (119). In this mechanism, UvrA binds to DNA, remains bound for a period (attached lifetime) and then detaches from DNA back into solution. This is distinct from a one-dimensional search whereby proteins slide along the DNA to find their target; this would appear as vertical movement of a streak on the position axis over time. *In vivo* it is expected that proteins use a combination of three-dimensional searching and one-dimensional sliding to find their target on DNA (discussed in section 1.8), although we did observe a small number of UvrA molecules displaying vertical movement in kymographs, this occurred at a very low frequency and was not pursued further. This behaviour is consistent with UvrA potentially sliding on DNA *in vivo*, but likely occurs faster, or on a smaller scale than the resolution of these experiments.

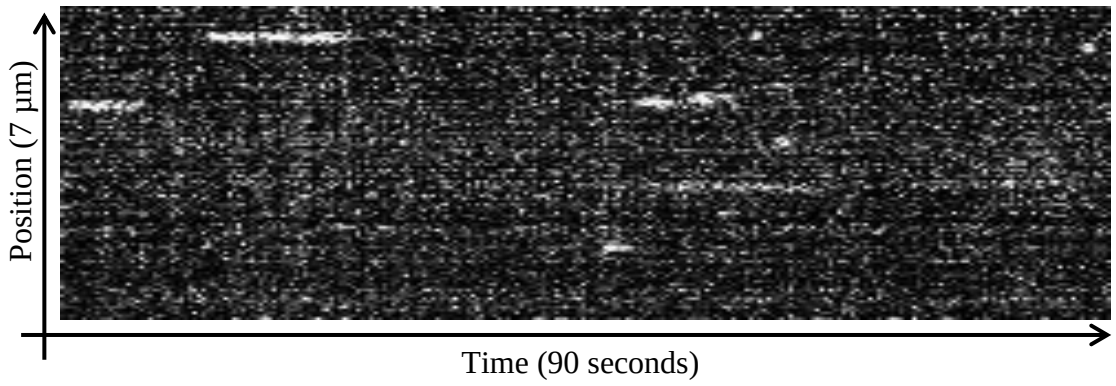


Figure 3.10 Kymograph of UvrA binding and releasing from DNA. A kymograph is a transformation of a video into position (Y) over time (X). Streaks on the Y axis show molecules binding along a position on DNA. Movement in the Y axis indicates diffusion along DNA. Streaks in this kymograph start and end at a single position and occur randomly across the DNA tightrope, showing UvrA uses a random three-dimensional search on DNA where UvrA binds, remains statically bound for a lifetime (length of streak) and then release back into solution.

3.12 UvrA's lifetime on DNA is affected by UV damage

The attached lifetime of individual proteins interacting with DNA can be determined by measuring the streak length and then converted into time by multiplying by the exposure time. Proteins releasing from DNA are independent from each other, i.e. one protein releasing does not affect another protein releasing, and the rate of occurrence is constant. These are criteria for a Poisson probability distribution, commonly used for modelling stochastic processes (109). In order to obtain the rate constant of UvrA's release from DNA, individual protein lifetimes were plotted as cumulative frequency diagrams and fit with exponentials (see section 2.17), appropriate for this stochastic Poisson process. The biochemically measured steady state ATPase had a k_{cat} of 1.07 ATP·UvrA^{-1} s^{-1}. In the single molecule tightrope assay, the detachment rate constant of UvrA on undamaged DNA is 1.24 ± 0.038 s^{-1} (**figure 3.11A, C**). These values are very similar and suggest UvrA's attachment to DNA is governed by ATP. The single molecule data fit better to a double exponential, suggesting two populations of UvrA are attached to the DNA. The necessity of a double exponential was determined using an F-test (F-test = 204, see section 2.17), and supported by the exponential residuals plot (**figure 3.11B**) and the non-linearity of the data on a log(Y) axis (not shown). However, the detachment rate constant *k1* has an amplitude of 94%, representing the vast majority of proteins detaching from undamaged DNA, with *k2* representing just 6%, most likely as a result of non-specific DNA damage occurring on DNA as shown in the gel assays.

An advantage of the DNA tightrope assay is the flexibility to string up different DNA substrates to investigate the effect on protein behaviour. As there is no change in the biochemically measured steady-state ATPase in the presence of UV damage, we recreated this substrate at the single molecule level. To do this, lambda DNA was UV irradiated with 1000 J/m^2 of 254 nm light immediately before creating DNA tightropes. When lifetimes of UvrA molecules in the presence of damaged DNA and ATP were plotted, they also fitted better to a double exponential (F-test = 154 (**figure 3.11B**)). However, the detachment rate constant *k1* showed a striking decrease to 0.411 ± 0.023 s^{-1}, equivalent to a three-fold increase in attached lifetime compared to undamaged DNA (**figure 3.11A, C**). Again, there was a small contribution of *k2* to the fit (4%) (**figure 3.11**). The *k2* rate constant might be expected to be from UvrA interacting with undamaged DNA, however the amplitude is likely too small to represent this population. The observation that we do not observe an

undamaged rate constant contributing as *k2* suggests that UvrA is saturated by the concentration of DNA damage, but it could also be that *k1* is composed of both undamaged and damaged DNA bound lifetimes; this may also explain the poorer fit of UvrA molecules in the presence of UV. However, the fit *k2* rate constant likely represents a small population of proteins that are long lived on DNA (> 10 seconds), and this species was not investigated further. These data show that on undamaged DNA UvrA remains bound for approximately 1 second, in good agreement with the biochemically measured ATPase. However, in the presence of UV damage, UvrA remains bound for around 2 seconds, but the biochemically measured ATPase remains unchanged (**figure 3.11C**).

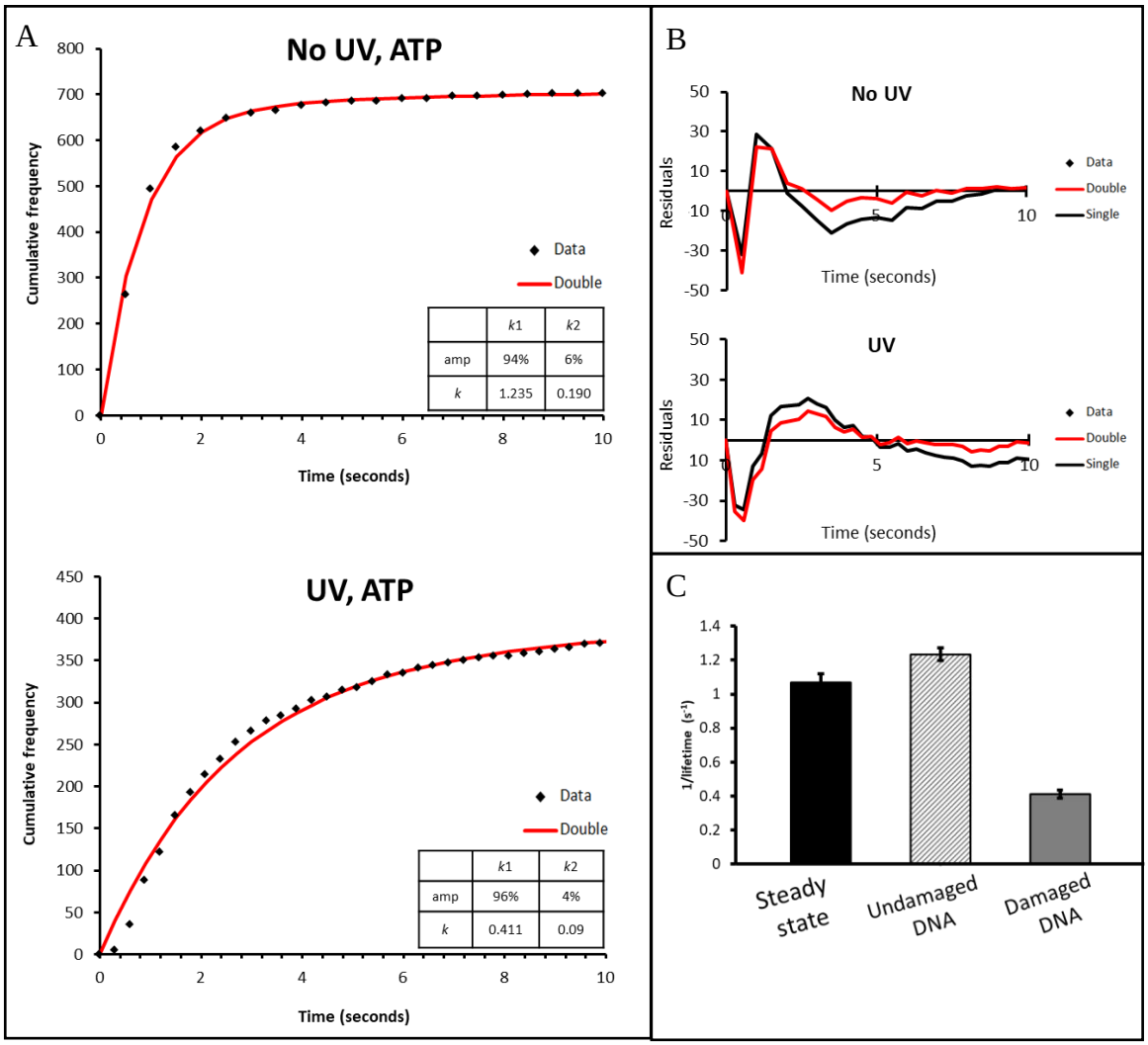


Figure 3.11 UvrA's lifetime on DNA is affected by the presence of UV damage.
A) Attached lifetimes of individual UvrA molecules (black diamonds (see figure 3.14)) plotted as cumulative frequency and fit to a double exponential (red line). The rate constants *k1* and *k2* describe UvrA's average detachment rates. This was repeated in the absence (top) and presence of UV damaged DNA. There is a clear change in the fast phase rate constant *k1*that indicates UvrA remains bound to DNA for longer in the presence of UV damage.

B) A plot of the residuals from a single (black) or double (red) exponential fit to both 'No UV' or 'UV' damage cumulative frequency diagrams. The data plotted here show a characteristic trough and peak, however the double exponential remains closer to 0 in both plots, indicating double exponentials are more appropriate to fit these data.

C) The single molecule detachment rate constant for undamaged DNA (grey stripes) is in excellent agreement with the steady state ATPase (solid black). However in the steady state there is no change in ATPase upon the addition of DNA damage, whereas at the single molecule the detachment rate constant decreases by around half (solid grey). Error bars show the error of the fits determined in Sigma plot.

3.13 ATP hydrolysis is coupled to detachment from DNA

UvrA's increased lifetime on damaged DNA may suggest that another process on DNA is occurring, or a process is being inhibited in the presence of DNA damage. To examine if ATP binding or ATP hydrolysis are required for UvrA's attachment to DNA, a non-hydrolysable ATP analogue ATPγS was used. ATPγS has a vastly slower rate of turnover compared to ATP but has previously been shown to permit formation of UvrA dimers and stimulate binding of UvrA to DNA (90, 101, 102, 106). UvrA's detachment rate constant on undamaged DNA in the presence of ATPγS is reduced by six and a half-fold compared to ATP (1.24 ± 0.038 s^{-1} to 0.155 ± 0.002 s^{-1}($P < 0.0001$)) (**figure 3.12**). This significant increase in lifetime shows ATP hydrolysis, and not just ATP binding is necessary for release from DNA. There is no significant difference to the detachment rate constant on UV irradiated DNA 0.127 ± 0.001 s^{-1}, which suggests that ATP hydrolysis is necessary to locate damage, most likely through repeated rounds of binding and releasing in a three-dimensional search. ATPγS bound UvrA displayed abundant binding to DNA (shown in **table 3.1**) which supports evidence that ATP bound UvrA favours DNA binding. Together these data show that ATP hydrolysis is coupled to UvrA's detachment from DNA. Detachment is drastically slower in the presence of ATPγS, implying hydrolysis and formation of subsequent nucleotide states occurs on the DNA, with one of these states presumably being sensitive to DNA damage and delaying UvrA's release from DNA.

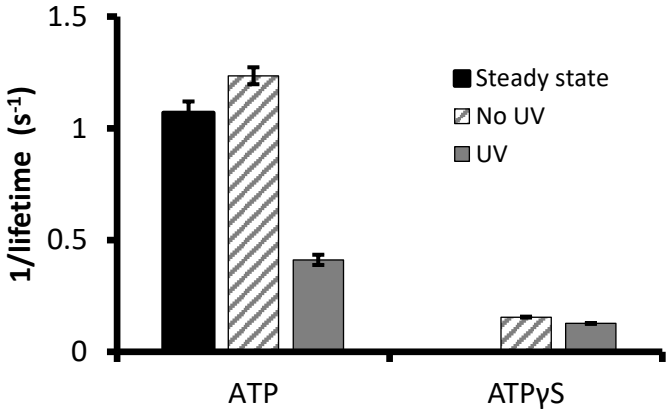


Figure 3.12 ATP hydrolysis is required for UvrA to dissociate from DNA. UvrA molecules in the presence of the non-hydrolysable ATP analogue, ATPγS demonstrate very slow release from DNA. Whereas UvrA molecules in the presence of hydrolysable ATP detach from DNA at a rate of around 1 per second. The presence of ATPγS inhibits this release, suggesting ATP hydrolysis, not just nucleotide binding is necessary for UvrA's release from DNA. The detachment rate constant is also unaffected by the presence of UV damage. Error bars show the error of the fits determined in Sigma plot.

3.14 ADP diminishes UvrA's DNA binding, and increases attached lifetimes

Experiments with ATPγS suggested ATP hydrolysis and subsequent nucleotide states occur on the DNA, and ATP UV experiments suggest a damage sensitive state must exist on DNA. Using DNA tightropes we attempted to recreate the ADP-bound and ADP + P_i-bound states of UvrA and investigate if they were sensitive to DNA damage or displayed a faster detachment rate constant. If a particular state is affected by DNA damage the attachment rate constant should reflect that of UvrA on UV irradiated DNA. Fitting the data from various nucleotide conditions provided dissociation rate constants shown in **figure 3.13**. ADP has previously been shown to facilitate UvrA dimerization, and binding to DNA, possibly in combination with ATP (90, 101, 102, 307). However, in the presence of ADP we observe UvrA molecules displaying vastly reduced DNA binding, with bound molecules remaining attached for a long time. Nevertheless, it was possible to calculate a detachment rate constant for both undamaged and damaged DNA, albeit with a low number of molecules (**figure 3.13** and **figure 3.14**). The ADP-bound state shows no difference between undamaged 0.202 ± 0.006 s^{-1} or UV damaged 0.217 ± 0.016 s^{-1} DNA (P = 0.429), with both of these being approximately five-fold longer than the undamaged ATP condition. The low frequency of protein attachment and long lifetime implies that UvrA is not able to form a stable dimer that can interact with DNA under these conditions, conflicting with previous observations (90, 101, 102, 307). From these data it can also be surmised that ADP-bound UvrA is not a damage sensitive state, but likely one involved in detachment from the DNA.

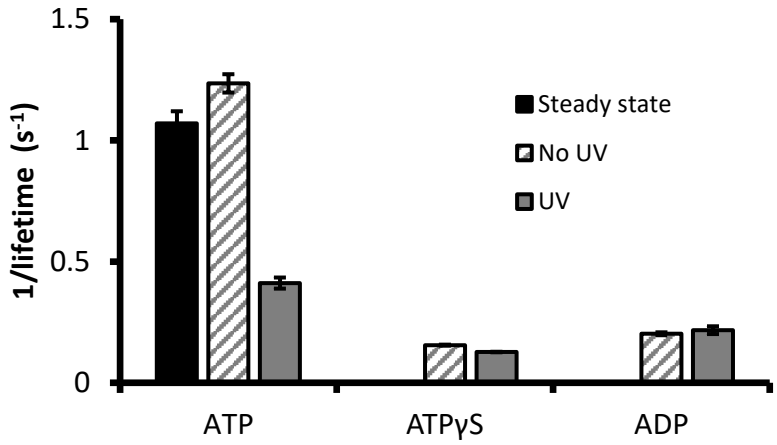
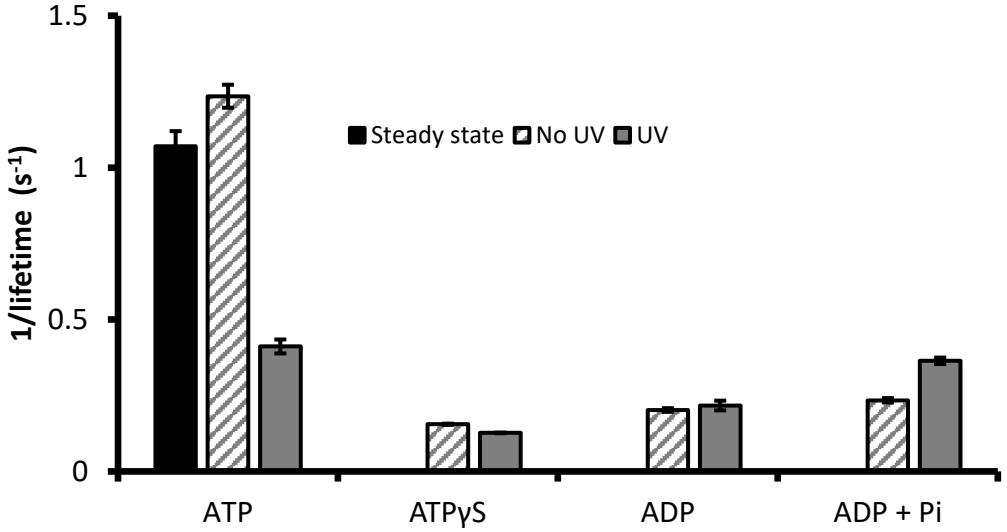


Figure 3.13 UvrA-ADP has low affinity for DNA. UvrA molecules in the presence of ADP show drastically reduced binding to the DNA (table 3.1); however detachment rate constants were determined. UvrA in the presence of ADP remained DNA bound, almost as long as in the presence of ATPγS. The ADP bound state of UvrA was also unaffected by the presence of UV damage. Error bars show the error of the fits determined in Sigma plot. Neither ATPγS or ADP bound UvrA show similar rate constants to the presence of ATP, and likely represent poorly populated states *in vivo*.

3.15 Addition of phosphate to UvrA-ADP recreates the damage sensitive state

If ATP-bound UvrA is the DNA binding state, and ADP-bound is the detachment state it is possible the intermediate ADP + P_i bound UvrA might be sensitive to DNA damage. To recreate this condition an excess of free phosphate was added to the imaging buffer, balancing the change in ionic strength with a reduction in KCl (284). On undamaged DNA, the addition of phosphate to the ADP state did increase UvrA's detachment rate constant slightly $0.202 \pm 0.006\ s^{-1}$ to $0.234 \pm 0.007\ s^{-1}$, but does not resemble the ATP condition ($1.07 \pm 0.019\ s^{-1}$) and is therefore not rate limiting for detachment (**figure 3.14**). Similar to the ADP-bound state, few UvrA molecules were seen interacting with the DNA in this condition, suggesting that phosphate does not restore UvrA's affinity for undamaged DNA (**table 3.2**). On UV damaged DNA the situation was starkly different, the detachment rate constant increased significantly from $0.217 \pm 0.016\ s^{-1}$ to $0.364 \pm 0.011\ s^{-1}$ ($P < 0.0001$), this value is not significantly different from the rate constant for release of UvrA-ATP on damaged DNA ($0.411 \pm 0.023\ s^{-1}$ ($P = 0.501$). This decrease in lifetime suggests that UvrA molecules bound with ADP + P_i represent a damage sensitive state.

A



B

Nucleotide, no damage	Detachment rate constant s^{-1}	Fit error	N molecules
ATP	1.235*	0.038	706
ATPγS	0.155	0.002	166
ADP	0.202	0.006	33
ADP Pi	0.234	0.007	53
Nucleotide, UV damage	**Detachment rate constant s^{-1}**	**Fit error**	**N molecules**
ATP	0.411**	0.023	394
ATPγS	0.127	0.001	152
ADP	0.217	0.016	21
ADP Pi	0.364	0.011	43

Figure 3.14 The addition of phosphate to ADP-UvrA mimics a damage sensitive state. In an attempt to recreate the ADP P_i state of UvrA, an excess of free phosphate was added to the buffer. The addition of phosphate did not restore UvrA's affinity for the DNA, nor did it recreate the ATP-bound lifetime in the absence of any DNA damage. A) When the DNA was UV irradiated, the detachment rate constant increased to a similar value of ATP-UvrA on damage, suggesting ADP P_i is a damage sensitive state. Error bars show the error of the fits determined in Sigma plot. B) A table of all detachment rate constants is shown for various conditions with the error of the fit (determined in Sigma plot) and the N of individual molecules counted from at least three flow cells per condition. *Double exponential, amp1 94%, *k1* 1.235±0.038s^{-1};amp26%, *k2* 0.190 ± 0.044 s^{-1}, **Double exponential, amp1 96%, *k1* 0.411 ± 0.023 s^{-1}; amp2 4%, k2 0.098 ± 0.030s^{-1}. These data are shown graphically in A).

Although a quantitative measure of DNA binding frequency (on rate) is not straightforward, a clear qualitative difference is seen between different nucleotide conditions, shown in (**table 3.2**). Although UvrA-ADP + P_i resembles the lifetime of UvrA-ATP on damaged DNA, the addition of phosphate to UvrA-ADP did not improve the number of binders on UV damaged DNA, and was far lower than UvrA-ATP.

No damage	Qualitative binding frequency
ATP	+++++
ATPγS	+++
ADP	+
ADP Pi	+
UV Damage	**Qualitative binding frequency**
ATP	+++++
ATPγS	+++
ADP	+
ADP Pi	+

Table 3.2 Qualitative binding frequency of UvrA in different nucleotide conditions. As an unbiased, quantifiable approach to determine the attachment rate constant was not possible, the relative binding was judged qualitatively. UvrA in the presence of ATP demonstrated the most DNA decoration, as did ATPγS. Both ADP and ADP Pi showed very poor decoration on DNA, suggesting these states have a low affinity for DNA.

3.16 Photobleaching does not affect measured lifetimes

To ensure photobleaching is not affecting the observed lifetimes of proteins on DNA the longest DNA-bound complex, UvrA-ATPγS was imaged at increasing laser powers. By plotting the lifetimes at different laser powers it is clear that only at laser powers > 10 mW is the attached lifetime affected (**figure 3.15**). This indicates that between 1-5 mW, where videos were collected, photobleaching is not defining the protein's attached lifetime.

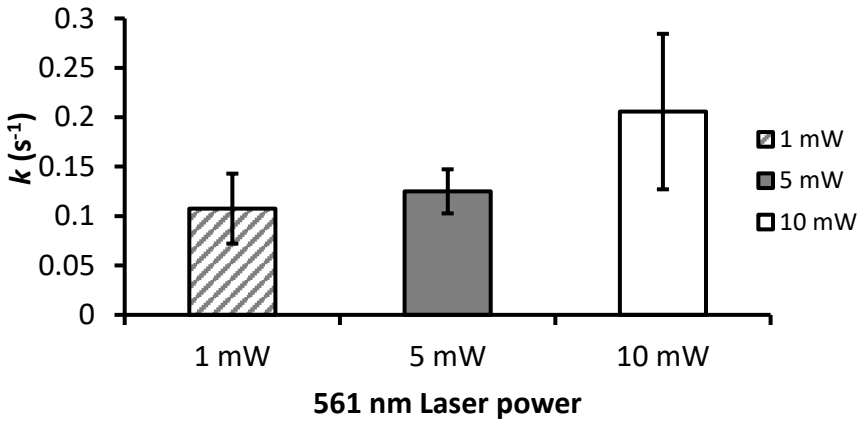


Figure 3.15 Photobleaching does not affect detachment rate constants.
Lifetimes of UvrA in the presence of ATPγS and UV damaged DNA give the slowest detachment rate constant, therefore remain bound to DNA for the longest time. Measuring the lifetime with different laser power shows that only at > 10 mW is the lifetime significantly affected. Lifetimes presented in this thesis were collected using between 1-5 mW of laser power. Error bars show the error of the fits determined in Sigma plot.

3.17 Crystal structures of UvrA suggest a mechanism of negative cooperativity

To try and understand how the different nucleotides affect UvrA's ability to interact with DNA we looked at the crystal structure of *Thermatoga maritima* UvrA in complex with fluorescein modified DNA (281). From the structure of UvrA it is feasible that the position of the bound nucleotide has the capability of directly modulating the position of the zinc fingers (302) that interact with the DNA, and that this would be disrupted by ATPase mutants. When examining the structure, a tunnel between the N- and C-terminal ATPase domains of each UvrA monomer was observed (**figure 3.16**). Using the cavity search feature in PyMol, a clear cavity connecting both ATPase domains in each monomer together was identified (**figure 3.16**). The DNA binding interface of UvrA runs across the surface of the dimer, with the two ATPase sites located at each end, immediately beneath the DNA. The tunnel that connects the two ATPase sites together may be important for the cooperative nature of the ATPase sites, possibly enabling P_i to be passed from one site to the other. This may activate the second site to turn over and lead to the UvrA homodimer falling apart and causing detachment from DNA. The structure also shows how distorted DNA or lesions could easily influence the ATPase sites and the tunnel connecting them, and may be important in understanding how the nucleotide bound to UvrA influences damage recognition.

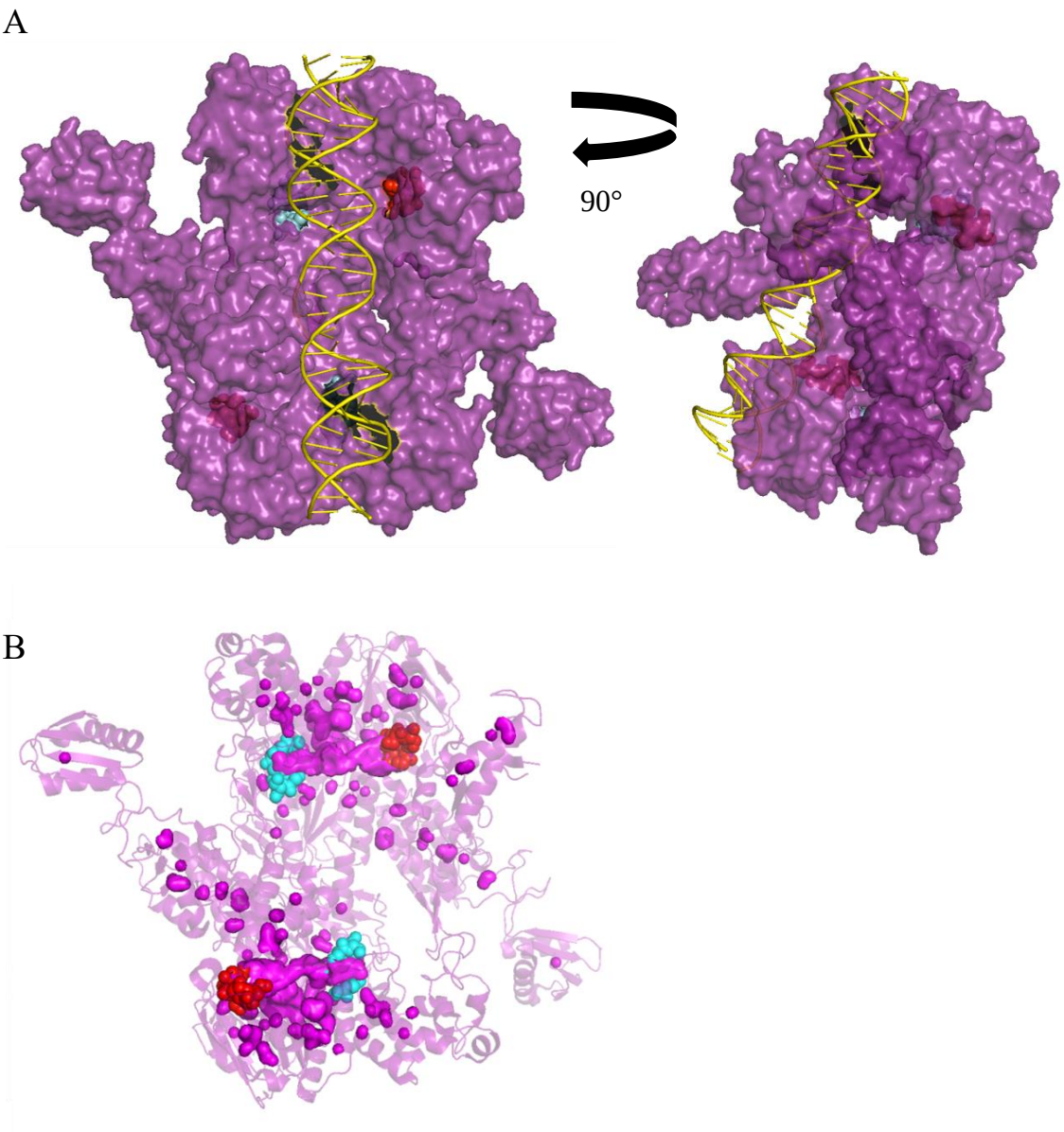


Figure 3.16 Crystal structure of *Thermotoga maritima* UvrA showing the phosphate tunnel.
A) Surface plot of dimeric UvrA (purple) with the N-terminal (cyan) and C-terminal (red) ATPase sites shown behind DNA (yellow). A 90° rotation of the surface plot shows the cavity beneath the DNA.
B) PyMol cavity search of structure reveals a clear tunnel (magenta) between the ATPase sites. UvrA is shown as a transparent cartoon with all pink surface objects showing solvent cavities. Figure created in PyMol using PDB structure (3PIH) (281).

Discussion

The results in this chapter have shown that UvrA has a DNA-stimulated ATPase, but damaged DNA does not further enhance the ATPase activity. We find excellent agreement between bulk biochemical measurements and single molecule kinetics to suggest that UvrA's detachment from DNA is coupled to ATP hydrolysis. However, UvrA's lifetime on damaged DNA is increased, whereas the ATPase rate stays the same between damaged and undamaged DNA, presenting a rate constant paradox. Using a range of different nucleotide conditions to recreate different kinetic intermediates of UvrA, we find that ATP bound UvrA favours DNA attachment, ADP and P_i bound UvrA is likely the damage sensitive state, and ADP bound UvrA is the release state.

3.18 UvrA's ATPase is stimulated by DNA, but is not further affected by the presence of damage

When measuring the steady state ATPase of UvrA, a two-fold increase in ATPase hydrolysis was seen upon the addition of various undamaged DNA substrates. A short oligonucleotide, a plasmid, and lambda DNA all stimulated UvrA's ATPase equally, from 0.71 ATP·UvrA^{-1} s^{-1} to 1.07 ATP·UvrA^{-1} s^{-1} (combined data). Whilst we see an increase in ATPase activity, Wagner et al see a ~50% drop in activity upon introducing undamaged DNA (120), this may be due to the use of different assays or buffers. Our NADH-linked assay has no ADP accumulation, whereas measuring γ^{32}P allows ADP to build up. UvrA has a high affinity for ADP and has been shown to competitively inhibit UvrA's ATPase (96, 97, 101). As described in the introduction, many studies on UvrA's ATPase conflict, but several papers support the DNA stimulation observed here (97, 102, 303).

When DNA substrates were damaged by UV or by incorporation of a fluorescein moiety, there was no additional increase in steady state ATPase 0.71 ATP·UvrA^{-1} s^{-1} to 1.07 ATP·UvrA^{-1} s^{-1} (P = 0.9594 combined data sets), but no additional stimulation compared to undamaged DNA (P = 0.0001). This is surprising considering the literature linking ATP hydrolysis to damage discrimination. However, we confirm that the damage substrates used contain UV damage, and show UvrA-mScarlet has an increased affinity for ATP in the presence of damaged DNA, showing damage discrimination is occurring. Interestingly, the increase in affinity for ATP is approximately the same as the change in lifetime we observe, and is similar to UvrA's affinity for damaged over undamaged DNA (112, 281). In both the ATPase and single molecule experiments the concentration of damage was approximately the concentration of protein, or higher. In the single molecule experiments it is difficult to determine the volume of the system and how much DNA is present, however the large number of damaged bases per molecule, higher affinity for damaged DNA, and low protein concentrations minimise the possibility of seeing a mixed population of undamaged and damaged binding. There was no difference in ATP turnover between fluorescein or UV damage, so if UvrA is capable of discriminating damaged DNA, why does the steady state ATPase not change upon the addition of damage?

3.19 UvrA molecules show a change in attached lifetime on damaged DNA

Directly imaging UvrA molecules binding to DNA enabled the correlation of attached lifetimes to the biochemically measured ATPase rates. In the presence of ATP the biochemical and single molecule experiments are in good agreement with detachment rate constants of 1.07 s^{-1} and 1.23 s^{-1} respectively. However, unlike the biochemical experiments, single molecule experiments show UV damage increases the attached lifetime of UvrA molecules significantly, two to three-fold. These values are approximately the same as values determined *in vivo* (119). This presents a rate constant paradox, as from the ATPγS experiments we know ATP hydrolysis is coupled to detachment from DNA, but we see no change in the steady state ATPase to suggest that additional hydrolysis is taking place to keep UvrA bound to the DNA.

3.20 Reconciling the rate constant paradox with a sequential hydrolysis mechanism

There appears to be tight coupling between ATP turnover and dissociation from DNA, meaning that one ATP is hydrolysed per UvrA monomer, per second on undamaged DNA. However, in the presence of damage the biochemical k_{cat} does not change, but the lifetime on DNA increases approximately three-fold – the same degree as the tightening in affinity for ATP on damaged DNA. Knowing that the ATP turnover and dissociation are linked, this suggests that on damaged DNA two ATPs are hydrolysed per second per dimer, but four ATPs are hydrolysed overall in the interaction with the damaged DNA. To reconcile this paradox, we suggest a model where one ATPase site hydrolyses on undamaged DNA, and that hydrolysis at this site is able to dissociate UvrA from the DNA if no damage is detected. Previous literature using the ATPase mutants would support this site being the C-terminal ATPase site (99, 120). However, if DNA damage is detected, we hypothesise a conformational change takes place to keep UvrA bound to the DNA, triggering hydrolysis at the second ATPase site (N-terminal), but importantly this hydrolysis can only take place once the first site has finished turning over. This assumes that the two sites hydrolyse at approximately the same rate, which is likely as they have been duplicated during evolution and are structurally very similar (295). The overall effect of this is that the steady state ATPase rate remains the same, but the lifetime increase on damaged DNA can be explained by activation of the second site. This sequential action is indicative of negative cooperativity between the sites that is tightly coupled to DNA damage recognition. Similar mechanisms of negative cooperativity have been noted for other ABC ATPase superfamily members, including the bacterial mismatch repair protein MutS (308).

3.21 Identification of the attachment, damage sensitive, and detachment states of UvrA

Although the bulk biochemical ATPase assay gives a sensitive measure of ATP turnover, other nucleotides cannot be substituted in the assay. DNA tightropes allow manipulation of the imaging conditions, including the use of UV damaged DNA and different nucleotides. UvrA-ATPγS demonstrated comparable DNA binding to UvrA-ATP, but was unable to release from the DNA, remaining bound for a long time on both undamaged and damaged DNA. By contrast, UvrA-ADP had virtually no affinity for DNA, but also remained bound for a long time. Both of these states have lifetimes that are too long to be related to the steady state ATPase rate, and neither were affected by DNA damage, and therefore they have no role in UvrA's discrimination of damage. However, the long attachment of UvrA in the presence of ATPγS demonstrates UvrA-ATP is the DNA binding state, but requires hydrolysis to release from DNA (90, 101, 102, 120) whereas ADP-bound UvrA is most likely the detachment state from DNA; marked by the low binding affinity. Several groups have demonstrated a mix of ADP + ATP is optimal for UvrA's activity (90, 106, 303). We show that with ADP alone, UvrA has almost no affinity for DNA, and is not capable of releasing, supporting the suggestion that ATP is required in at least one ATPase site for UvrA to be able to release from DNA; however we cannot say which site this is. To recreate the intermediate kinetic state between ATP and ADP bound, an excess of free P_i was added to ADP bound UvrA. The addition of P_i did not restore UvrA-ADP's low affinity for DNA but there was a change in lifetime. On undamaged DNA, ADP + P_i behaved similarly to ADP alone, (0.202 s^{-1} vs 0.234 s^{-1}). However, on damaged DNA there is a significant decrease in the detachment rate constant from 0.234 s^{-1} to 0.364 s^{-1}, with the latter not being significantly different from the detachment rate constant of UvrA-ATP on damaged DNA (0.411 s^{-1}). These data suggest that ADP + P_i bound UvrA populates a damage sensitive state on DNA, but interestingly it doesn't restore UvrA's affinity for DNA. The addition of P_i attempts to push the pathway backwards, if release of ADP + P_i is rate limiting for detachment from DNA, addition of P_i should accelerate the detachment rate constant. In our assay there was no acceleration of the undamaged DNA rate constant, suggesting the C-terminal ATPase site is unable to be pushed backwards to a P_i and ADP bound state. However, on damaged DNA, the rate is accelerated, suggesting that either one or both ATPase sites are able to reload P_i and accelerate the rate limiting step in the presence of damage.

3.22 The crystal structure of UvrA suggests a mechanism of cooperativity

When examining the crystal structure of UvrA it was apparent that the two ATPase sites are connected by a tunnel (**figure 3.16**). This cavity joining the adjacent ATPase sites of each monomer in the dimer suggests a possible mechanism of how one site might regulate the other. Their close proximity makes it plausible that P_i released from the first site regulates the binding, or hydrolysis of bound nucleotide at the second site. It would be interesting to mutate some of the internal tunnel cavity residues to disrupt the cooperativity between domains, allowing a sequential mechanism to be proved. On undamaged DNA, we propose the C-terminal ATPase site hydrolyses one ATP, per monomer, per second. Evidence suggesting the identity of the first and second sites was revealed using ATPase mutants and *in vivo* experiments (119, 293). Previous observations that nucleotide needs to be bound to the C-terminal ATPase site for damage recognition suggest this is the first ATPase site described here (307). This leaves the N-terminal site as the site activated upon binding damaged DNA. On damaged DNA, the first site hydrolyses ATP, and the second site, after the first site has turned over, hydrolyses another ATP per monomer per second. The overall effect of this is that the steady state ATPase remains constant, but the lifetime increases on DNA. The N-terminal ATPase site has been suggested as necessary for the recruitment of UvrB, the damage verifying helicase to the lesion and loading it onto DNA (99, 119, 293, 302), which supports the role of a damage activated site described here. Recognition of DNA damage by UvrA involves the zinc fingers found in close proximity to the ATPase sites (137, 301, 302). The various zinc finger structures are detailed in the introduction of this chapter, but there is a clear conformational change that either supports a closed conformation (DNA bound) or an open conformation (not bound to DNA) and it is plausible this is linked with ATP hydrolysis at the N-terminal ATPase site. Due to the DNA binding site across the face of the protein sitting above the ATPase sites, it suggests the presence of DNA damage will affect the cooperativity and conformation of the two ATPase sites and this in turn is likely to affect the zinc fingers.

3.23 Conclusions

This chapter has described the investigation of UvrA's ATPase, using both bulk phase biochemical studies and single molecule fluorescence techniques. The role of different nucleotides and DNA substrates on UvrA's activity has been explored. Strong agreement between results from both methods provides evidence that UvrA's ATPase is tightly coupled to attachment to DNA. On undamaged DNA UvrA hydrolyses two ATPs per dimer per second, most likely at the C-terminal ATPase site that lead to dissociation. On damaged DNA, we find the steady state ATPase does not change, however the lifetime on DNA does. This rate constant paradox can be reconciled with the model that the N- and C-terminal ATPase sites fire in sequence, i.e. two ATPs are hydrolysed by the C-terminal site, but the N-terminal ATPase site is activated by damage, causing two ATPs to be hydrolysed at the N-terminal site, but only after the C-terminal site has turned over. This accounts for the observed kinetics and suggests a tight coupling model due to strong negative cooperativity between the UvrA ATPase sites.

3.24 Difficulty proving the sequential model

Throughout this work there have been difficulties; purification, solving the rate constant paradox, and proving the sequential mechanism. The latter of which is still partially unresolved. Previous studies used catalytic lysine mutants of each ATPase site to try and delineate the role of each site (discussed in section 3.3). However, the evidence that mutating one site abolishes all ATPase activity makes accurate interpretation of the various results complex. Therefore using the two ATPase mutants here would not clearly show the model, but likely complicate matters further. The difficulty in proving the sequential hydrolysis hypothesis is that the steady state ATPase does not change, so any measure of this will show no difference between undamaged and damaged DNA. Several different methods were applied to try and prove this hypothesis, but none were able to convincingly provide support; these are briefly explained here.

3.24.1 Malachite green

The malachite green phosphate assay kit (Sigma Aldrich, UK) determines free phosphate concentration in solutions. The association of malachite green dye and free phosphate causes an emission change between 600-660 nm that can be measured. Experiments with UvrA and either undamaged or damaged DNA showed no significant differences, confirming the steady state ATPase result that shows no difference between undamaged and damaged DNA. Experiments carried out with a time course showed that the conditions to stop the assay (high concentration HCl) either caused ATP decay to ADP and P_i, or did not prevent UvrA turning over. Therefore, this experiment does not provide any evidence for sequential hydrolysis and is not discussed further.

3.24.2 Tryptophan fluorescence

E. coli UvrA contains three tryptophan residues per monomer. Tryptophan residues fluoresce with excitation at ~280 nm, and depending on the environment of the residue emit light around 300-350 nm (309). The three residues in UvrA stick out of the structure, almost on top of the DNA. To see if the addition of ATP, or DNA altered the chemical environment of the tryptophan residues, the fluorescence emission wavelength was measured. Initially it seemed there was a good signal from the tryptophans alone, peaking at around 345 nm. On the addition of 1 mM ATP there was a significant drop in fluorescence intensity, which did not change over time. However, the addition of water demonstrated the same drop in signal, likely due to dilution of the tryptophan signal. Interestingly, Kraithong et al used tryptophan fluorescence in their recent study, however this was using thermophilic UvrA, leading us to suspect that structural differences between these proteins gives a tryptophan signal that was not observed here (302).

3.24.3 Stopped flow with fluorescent ATP

Following on from the unsuccessful tryptophan fluorescence experiments, we investigated if Förster resonance energy transfer (FRET) would occur from the tryptophan residues and fluorescently labelled ATP (2'-(or-3')-*O*-(*N*-Methylanthraniloyl) Adenosine 5'-Triphosphate (mant-ATP)) bound to UvrA's ATPase sites. There was a FRET signal that decreased upon the addition of UvrA, and a small decrease in fluorescence decay over time that may suggest a slow conformational change of UvrA-mant-ATP. This was then further investigated using stopped flow pre-steady state kinetics. The association of mant-ATP and UvrA should show any transient decrease over time, and as we know the steady state ATPase is unchanged, it was possible the approach to equilibrium would be different for undamaged and damaged DNA. Binding and hydrolysis of mant-ATP to the ATPase site of UvrA should change the environment of the fluorophore, and provide a fluorescent signal that can be monitored. However, there was only a very small transient change in signal using stopped flow, and as these assays consumed vast volumes of protein they were abandoned. As UvrA's affinity for ATP changes based on the damage state of the DNA this precludes the use of mant-ATP in single molecule studies. Additionally, the nucleotide concentration needs to be far in excess of the K_M, but at this concentration the background fluorescence would obscure any signal from individual molecules binding to UvrA on the DNA.

3.25 Future work on UvrA

To consolidate and support the data presented in this chapter some additional experiments should be carried out. The hypothesis that the phosphate tunnel is responsible for the cooperativity between the ATPase sites in UvrA can be tested using site-specific mutants. Residues that line the cavity can be identified using the crystal structure of UvrA ((PDB: 3PIH) **figure 3.17**) and bulky residues such as tryptophan can be substituted to obstruct the tunnel, affecting the cooperativity without mutating either ATPase site. This uncoupling of the ATPase sites should prevent activation of the second site, even in the presence of damage. At the single molecule level this should give an attached lifetime of ~1 s^{-1} on both undamaged and damaged DNA, this would show that the tunnel is important for regulating the second ATPase site and regulating damage discrimination by UvrA.

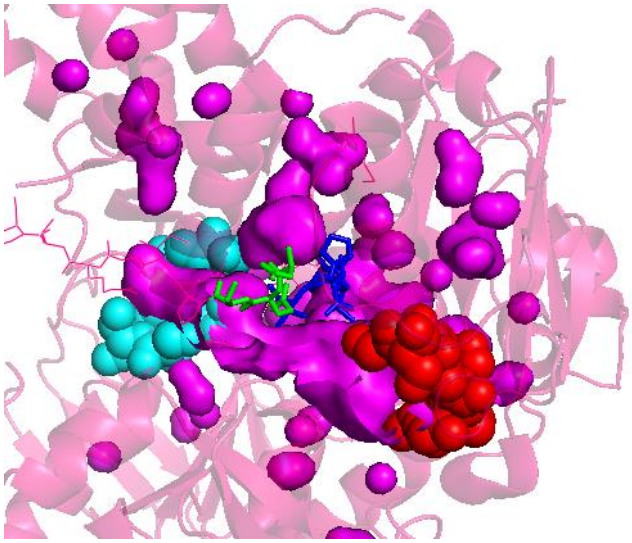


Figure 3.17 Site-specific mutants affecting the tunnel connecting UvrA's ATPase sites. Blue sticks show residues 831-836 (DEPTV) and green sticks show residues 664-666 (QSP) of UvrA sitting just above the cavity (shown in surface as solid magenta) connecting the two ATPase sites. The N-terminal ATPase site is shown in cyan, and the C-terminal ATPase site in red as space filled atoms. These are candidates for mutants that might affect the cavity and disrupt the cooperativity between ATPase sites. Structure created in PyMol using PDB structure 3PIH.

The kinetics of each ATPase site are assumed to be very similar as they likely arose from gene duplication (294, 295). Using the K37A and K646A mutants the kinetics of the other site could theoretically be obtained using the ATPase assay; however, mutating one site seems to abolish all ATPase activity (120, 303). What is really needed are a set of mutants that do not adversely affect the structure or function of UvrA (i.e. UvrA-mutants that are capable of rescuing UV survival in UvrA knockout cells) but affect the ATPase activity of one site. A series of constructs that have mutations in the Walker A motif of both ATPase sites can be screened for *in vivo* complementation and then examined *in vitro* for folding compared to wild-type UvrA using circular dichroism. Two such mutants are used by Stracy et al *in vivo* but these should be characterised for protein folding and biochemical activity to validate them before using them *in vitro* (119). If one ATPase site can be slowed down, this should allow a sequential mechanism to be proved. If the two sites fire at different rates, a mixed population of rate constants will be observed and these can be related to the kinetics shown in this chapter to show both sites have the same rate constant, supporting a sequential model. Taking the results in this chapter further the two-site hypothesis should be expanded to include steps where UvrB is loaded onto the DNA and the kinetics of incision, which is discussed further in section 5.1.

3.25.1 Expanding the two-site hypothesis

The two-site hypothesis should be extended to encompass UvrA's ATPase within the UvrAB complex. It has been shown that UvrAB can detect damage (103, 108), but it is not known how UvrA's ATPase functions when sandwiched by two UvrB molecules. UvrB also turns over ATP when in complex with UvrA, or when bound to damaged DNA (96, 105, 124), preventing the straightforward use of the ATPase assay used in this chapter. Although ATP turnover drastically increases when UvrB is added to UvrA (section 3.6), the contribution of each proteins ATPase cannot be isolated. UvrA can utilise both ATP and GTP, whereas UvrB can only turnover ATP (307). This can then be used to monitor the hydrolysis of the two different proteins turning over nucleotide. This has been performed using thermophilic proteins (137) but should be repeated using the *E. coli* system with a careful analysis of the reaction kinetics. Whether GTP functions in the same way as ATP in the NADH-linked assay is unknown for the *E. coli* proteins but can be tested by replacing ATP with GTP in the presence of UvrA on DNA. The rate constant of UvrA's GTPase compared to the ATPase is also unknown for *E. coli* UvrA and should be determined before examining UvrAB's ATP/GTPase activity. The different turnover of nucleotide in the presence of undamaged

DNA, single-stranded DNA, bubble containing DNA, and UV irradiated DNA should be compared to build on the results in this thesis. These experiments should be supported with a catalytically dead mutant of UvrB's ATPase (K45A (124)), and the beta-hairpin mutant (Y96A (137)). The mutant of UvrB's ATPase should show no ATPase activity, just UvrA's GTPase and if this is stimulated by the presence of UvrB. The beta-hairpin mutant cannot discriminate damage and should cause UvrB's ATPase to continuously turn over without being released from the DNA (137). These biochemical experiments should be combined with single molecule techniques to allow further clarification and characterisation of the kinetics. If UvrB is purified with a fluorescent protein such as mNeonGreen it can be visualised at the same time as UvrA using a multi-channel fluorescence microscope. A complex of UvrA & UvrB can then be imaged on DNA tightropes to get the diffusion characteristics and behaviour, similar to a recent study from our lab using thermophilic proteins (108). The UvrAB complex can then be examined for its response to UV damage; do complexes slide to find the damage sites and then stall? Are DNA damage sites directly bound by UvrAB from solution? The single molecule experiments should also be performed with the beta-hairpin mutant that will prevent UvrB detecting or verifying DNA damage. Finally, UvrA, UvrB, and UvrC should all be tagged with different colour QDOTs via different labelling methodologies (e.g. Biotin-Streptavidin, His-IgG, HA-IgG) that enable all three proteins to be visualised simultaneously binding to DNA tightropes. DNA damage can be introduced at a specific location that can be labelled with a fluorophore such as Cy3 (108). The fluorescently tagged DNA damage might disappear when excised but will probably require the addition of UvrD (labelled or unlabelled). This experiment enables the kinetics of the complete NER excision reaction to be obtained in real-time. Many questions remain about the dynamics of NER and the order of events that can be addressed with this experiment, such as, at which point is UvrA released from the UvrAB:DNA complex? How long does it take from detection to incision? Can we see incision take place without UvrA? If no damage is verified does UvrB detach from the DNA?

Chapter 4

The TFIIH components p44/p62 act as a damage sensor during nucleotide excision repair

Introduction

4.1 TFIIH is the central hub in eukaryotic NER

The major transcription factor TFIIH orchestrates eukaryotic NER. Initial damage detection is carried out by XPC which then recruits the TFIIH complex to the DNA near the lesion (189, 310). The helicases XPD and XPB located within TFIIH then unwind the DNA around the lesion. Critically, XPD must verify the presence of damage before incisions either side of the damage can occur. It is still poorly understood how XPD is capable of discriminating DNA damage, and which other proteins are involved in this step.

4.2 Helicases in TFIIH are regulated by other subunits

XPB and XPD are helicases found at the base of the horseshoe structure in TFIIH (**figure 4.1)**. Both XPD and XPB are part of the SF2 helicase family, but have opposite polarities (190). Many helicases require stimulation by other proteins to conduct their biological function (190). Accordingly, structural and biochemical experiments have shown that both XPB and XPD are regulated by other TFIIH subunits. XPB's ATPase is regulated by p52 and p8 (311-313), and XPD's ATPase is regulated by p44 (156, 207, 233, 240, 314, 315). Several structures of yeast and human TFIIH exist, however flexibility within TFIIH, sample preparation, and imaging artefacts prevent every subunit from being assigned and the interactions being characterised (177, 232, 316, 317). Luo et al performed crosslinking mass spectrometry on human TFIIH to create linkage maps (**figure 4.2)**. Around 70% of the interactions were found in four key domains; 'the anchor' in p62, 'the hub' in p52, 'the lock' in XPB, and 'the latch' in MAT1. These maps show how the subunits within TFIIH interact, and also demonstrate the complex network of protein-protein interactions that hold TFIIH together (177). Several recent cryo-EM structures show clear contacts between the helicases and regulatory subunits in TFIIH (**figure 4.1**) (205, 208, 239). In one cryo-EM structure, every core TFIIH subunit can be unambiguously assigned (239). Extensive contacts between p44, p62, and XPD can clearly be observed (**figure 4.1**) opening up many questions about how these three proteins interact and how this interaction affects DNA repair.

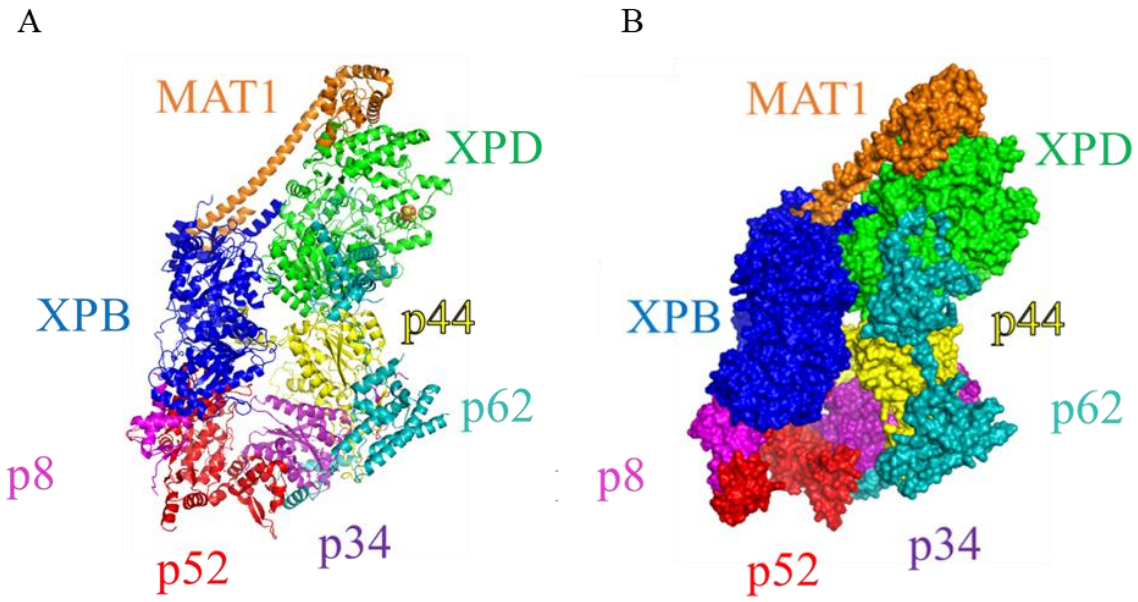


Figure 4.1 Cryo-EM structure of core TFIIH associated with MAT1. PyMol structure 6NMI in A) cartoon and B) surface displays, coloured by protein subunits. The presence of MAT1 disrupts any DNA binding, or helicase activity of XPD. The p62 subunit (teal) can be seen inserted into the XPD DNA binding site. Structure isolated from PDB: 6NMI (239).

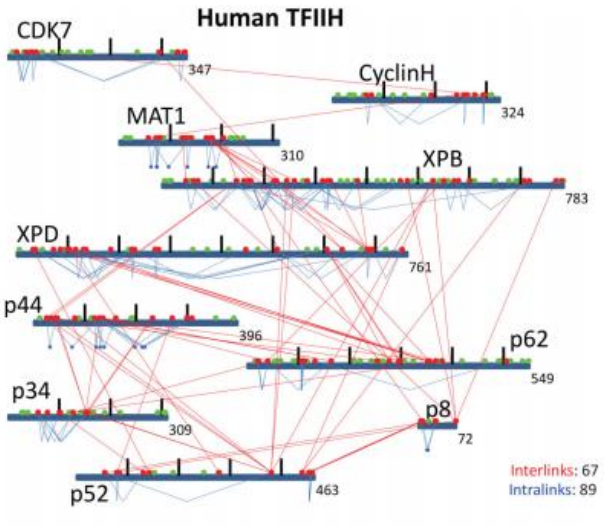


Figure 4.2 Protein-protein contacts within TFIIH. Crosslinking studies on human TFIIH show the complex network of interactions between subunits. Both intralinks (blue) and interlinks (red) can be seen on the crosslink map. Red and green dots show positions of lysine residues. Figure taken from (177).

4.3 Role of p44 and p62 in DNA repair

4.3.1 Structure of p44

Humbert et al first identified and cloned human p44 along with p34 (318). Both proteins were confirmed as subunits of TFIIH and found to possess von Willebrand factor-A like domains (vWA) that have been implicated in forming extensive protein-protein contacts and in dimerization (159, 177, 316, 318-321). In addition to the vWA domains, p44 also contains several zinc finger motifs that form a proposed DNA binding domain (318) that has been shown in several other proteins such as CTCF (322). The different domains of p44 are shown in **figure 4.3** with the cartoon structure of p44 isolated from the Greber 2019 cryo-EM structure. A heavily redundant network of contacts exists between p34 and p44, suggesting this is an important interaction, and may explain the lack of clinical cases showing p44 mutations (321). The p34:p44 interaction was shown to be crucial for TFIIH's architecture, with the vWA domains implicated in binding multiple other TFIIH subunits (177), and being important for TFIIH assembly (177, 319, 321, 323, 324). Mutations in p44 affect the genome stability of yeast, sensitise cells to UV, and cause defects in transcription and DNA repair (325-328). p44 regulates the ATPase activity of the main helicase involved in NER, XPD, and mutations in XPD that disrupt this interaction cause deficiencies in DNA repair and transcription that lead to diseases such as XP (159, 329).

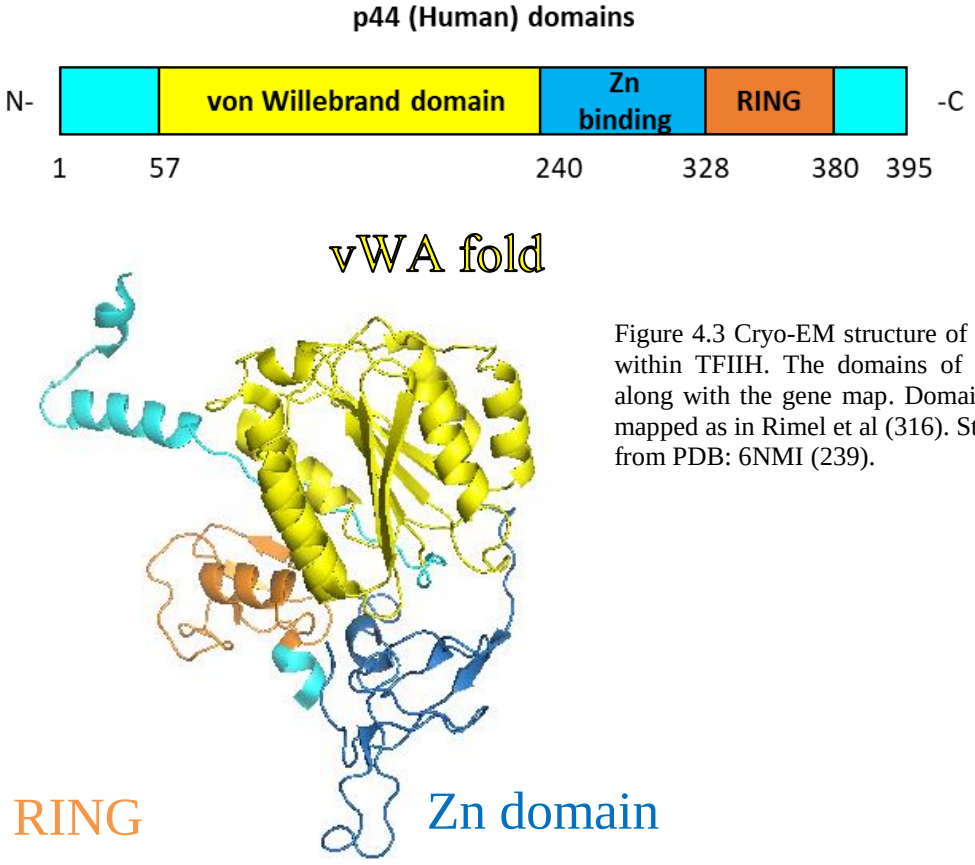


Figure 4.3 Cryo-EM structure of the p44 subunit within TFIIH. The domains of p44 are shown along with the gene map. Domains of p44 were mapped as in Rimel et al (316). Structure isolated from PDB: 6NMI (239).

4.3.2 Structure of p62

Human p62 was found to associate with TFIIH and be involved in transcription (330). p62 is absolutely required for DNA repair, (331) however there appears to be no enzymatic activity associated with p62, and the role in TFIIH is unclear (330, 332, 333). Structurally, p62 contains two BSD domains which have been hypothesised to interact with DNA (316) it also contains a PH-like domain that contains the anchor region which makes important protein-protein contacts within TFIIH (177, 316, 334). The PH domain is not visible in the Greber 2019 structure; however, the other domains are shown in **figure 4.4** as a domain map, and a cartoon structure extracted from the entire human TFIIH structure.

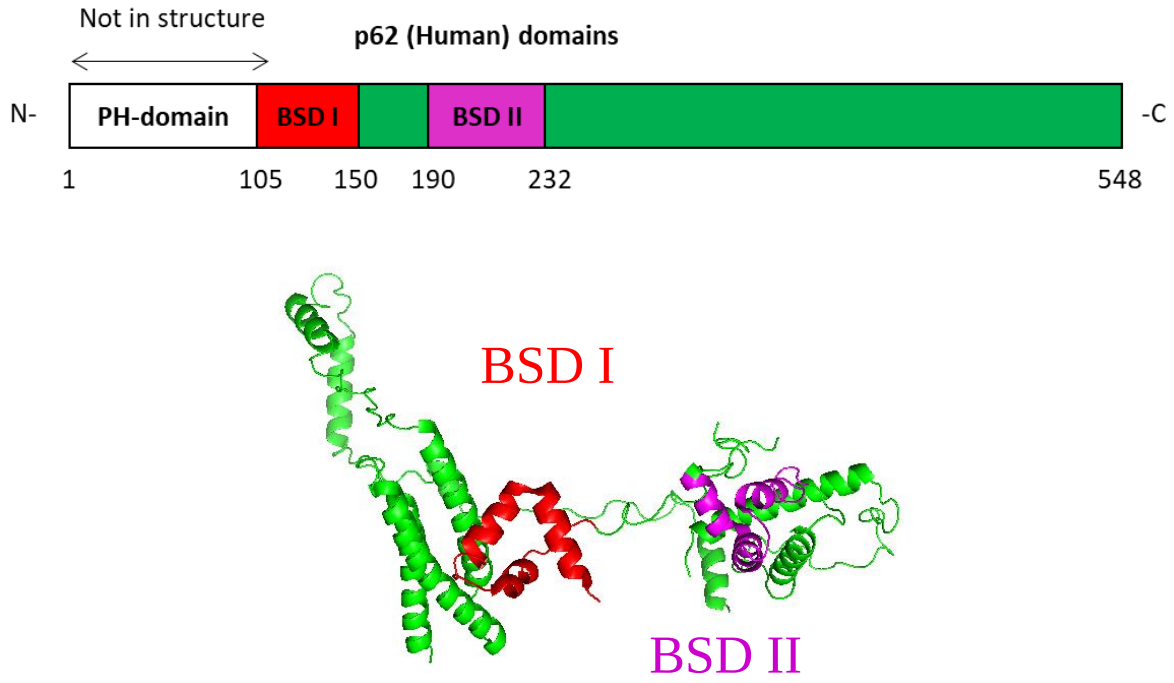


Figure 4.4 Cryo-EM structure of the p62 subunit of core TFIIH. p62 is shown as a cartoon structure. Domains of p62 were identified from the side chains and mapped as in Rimel et al (316). The PH domain is missing in the cryo-EM structure. Structure isolated from PDB: 6NMI (239).

The transcriptional TFIIH cryo-EM structure shows p62 makes extensive contact with the core TFIIH proteins, forming a supportive arch structure (239). However, p62 is not able to be defined in several other TFIIH structures (205, 208, 335), including the most recent cryo-EM structure where TFIIH is bound to DNA, suggesting p62 is flexible when TFIIH is engaged with the DNA, but not in the transcriptional state (208, 239). Biochemically, p62's interactions with TFIIH are complex, with multiple binding partners sharing the same region (177, 332, 336, 337). Studies in yeast have shown XPC binds the PH domain of p62, and XPG shares the same binding site (310, 338). Both of these interactions are crucial for cell survival to UV exposure, and suggest how NER progresses from initial damage recognition to incision. XPC will initially detect the lesion in the DNA, then the XPC:damage complex has to recruit TFIIH, probably through the PH domain of p62 and an interaction with XPB (310, 338). Once TFIIH is loaded on the DNA verification by XPD can take place. If damage is present XPC must be displaced to form the incision complex where XPG then interacts with p62. This displacement may help load XPG in the correct orientation on the damaged strand to enable efficient incision on DNA (177, 314, 331, 339).

4.4 Interactions between p44 & p62 and TFIIH

Several groups have observed interactions between p44 and p62, and this seems to be important for incorporating p62 into TFIIH (156, 185, 323, 324). It has also been shown that p62 contacts the HD2/Fe-S region of XPD and may have a role in regulating XPD's helicase activity. Mutants in XPD where p62 binds affect helicase activity, similar to the interaction between p44 and XPD (177, 207, 239). Furthermore, ~20 residues of p62 are seen to be inserted into the DNA binding cavity of XPD, preventing DNA binding – although this is in a transcriptional state where XPD is not required (**figure 4.5**) (239). Investigating the interactions between p44, p62 and XPD is critical for our knowledge of how XPD discriminates and verifies DNA damage, and understanding the mechanism of how DNA is repaired. While p44 is known to stimulate XPD (159, 207), the role of p62 is almost entirely unknown. Both p44 and p62 have been shown to interact with XPD and possess putative DNA binding domains, but despite this they remain relatively understudied in TFIIH. Using a single-molecule approach we investigated how purified p44 and p62 interact with XPD and contribute to DNA repair.

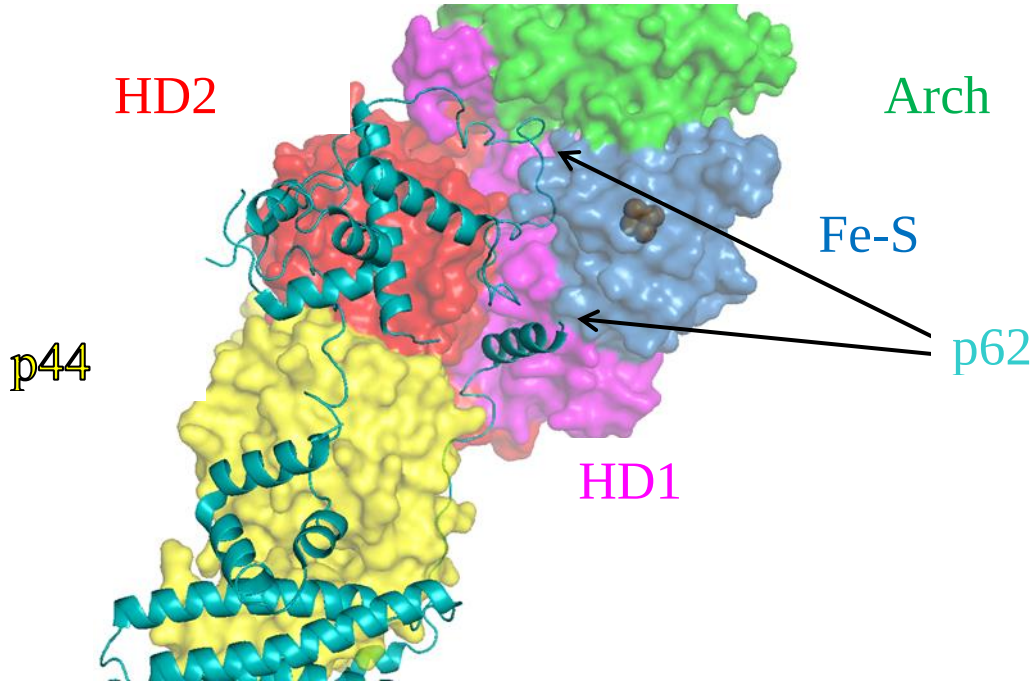


Figure 4.5 Cryo-EM structures of p44, p62 and XPD in TFIIH.
Cryo-EM structure of p44 and p62 contacting XPD. p44 (yellow) and XPD (coloured by domain) are shown as surface plots with p62 shown as a cartoon structure. p62 can be seen making extensive contact with p44, and the helicase domains (HD2 (red) and HD1 (magenta)) of XPD. Furthermore a loop of p62 can be seen inserted into the DNA binding domain of XPD (between the Fe-S (blue) and helicase domains). Structure created in PyMol from PDB: 6NMI, XPD domains were mapped as in Fan et al (239, 248).

Results

4.5 Full length p44 and p62 co-purify as a complex

In collaboration with the Kisker lab, initial studies (performed by Dr Jochen Kuper) discovered that to purify full length active p44, co-purification with p62 is necessary, eluting as a tightly associated complex of p44/p62 (**figure 4.6**). This p44/p62 complex was also capable of forming a ternary complex with XPD (**figure 4.6**). This is the first-time full length p44 has been purified and studied. Previously, p44 has been purified as a truncation consisting of the first 285 residues. This p44 construct contains the vWA domain and the start of the zinc domain, but is missing the C-terminal RING domain, and is referred to as N-p44 hereafter. XPD and p62 are able to be purified independently (207).

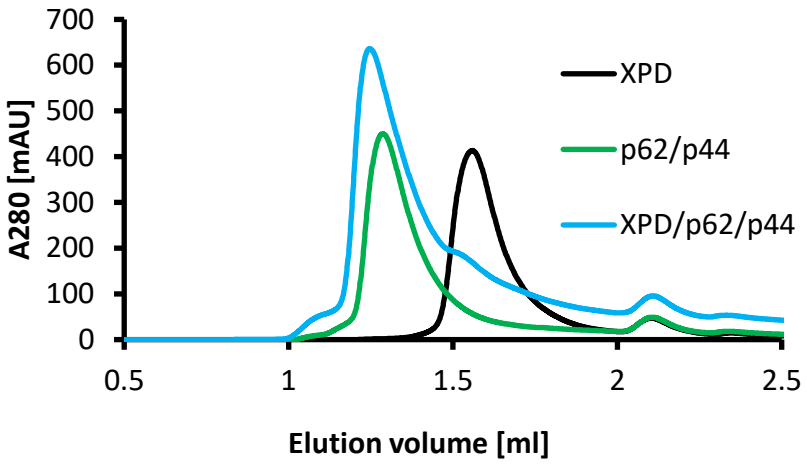


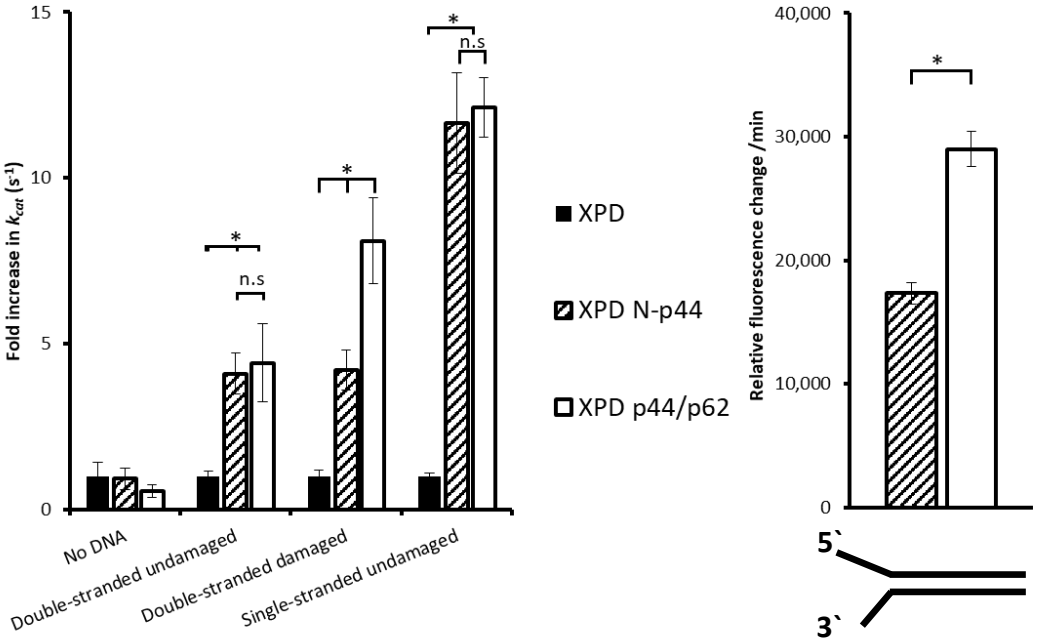
Figure 4.6 p44 and p62 co-purify as a complex that interact with XPD.
Size exclusion chromatography fractions show that p44/p62 (green) co-elute as a complex, while XPD (black) can be eluted alone, or in complex with p44/p62 (blue). Data from Dr Jochen Kuper

4.6 XPD's ATPase is stimulated by p44/p62 on damaged DNA

In the cryo-EM structure of human TFIIH, p62 contacts XPD and p44 (**figure 4.5**) (239). The ATPase of XPD is coupled to the helicase activity, with ATPase activity being stimulated by p44 (159, 207). Mutations in XPD that disrupt interactions with p44 cause disease by disrupting helicase activity – leading to a defect in DNA repair (159, 207). It has been suggested p62 also contacts XPD in the HD domains and the Fe-S cluster that is important for damage discrimination (177, 178). To investigate if p62 was able to stimulate the ATPase of XPD, the turnover of ATP in the presence of different DNA substrates was measured using a NADH-coupled assay. Neither p44 nor p62 have been found to contain any ATPase motifs or ATPase activity *in vitro* (Dr Jochen Kuper, unpublished).

XPD alone had a very slow ATPase, even in the presence of single-stranded DNA (**figure 4.7**). As shown previously (207), N-p44 stimulated XPD's ATPase in the presence of both single-stranded and double-stranded DNA (**figure 4.7**), however p62 alone did not stimulate XPD's ATPase on any substrate. Surprisingly, the p44/p62 complex only accelerated XPD's ATPase on undamaged DNA to the same degree as N-p44, with both N-p44 and p44/p62 showing the greatest enhancement of XPD's ATPase activity in the presence of undamaged single-stranded DNA, consistent with XPD's role as a single-stranded translocase (207). However, when a fluorescein moiety (that NER recognises as damage (200)) was introduced into a double-stranded DNA substrate, p44/p62 accelerated XPD's ATPase two-fold more than N-p44 (**figure 4.7**). The acceleration of XPD's ATPase by N-p44 on damaged DNA is indistinguishable from the stimulation on undamaged DNA, suggesting that additional residues present in p44/p62 are enhancing the stimulation of XPD's ATPase. This is the first evidence that p44/p62 may have a role in modulating XPD's ability to detect or discriminate DNA. These data are supported by helicase data that show p44/p62 enhance XPD's ability to successfully unwind an undamaged fork substrate more so than N-p44 alone. Interestingly, the enhancement of XPD's helicase activity is similar to the fold change between the ATPase activity on damaged DNA, suggesting p44/p62 may enhance XPD's translocation on DNA.

A



B

DNA substrate	XPD k_{cat} values (s^{-1})			
	XPD	XPD p62	XPD N-p44	XPD p44/p62
No DNA	0.046 ± 0.013	-	0.042 ± 0.008	0.025 ± 0.004
Double-stranded undamaged	0.033 ± 0.004	0.023 ± 0.000	0.136 ± 0.013	0.146 ± 0.035
Double-stranded damaged	0.027 ± 0.004	0.024 ± 0.005	0.114 ± 0.006	0.218 ± 0.019
Single-stranded undamaged	0.043 ± 0.003	0.025 ± 0.010	0.504 ± 0.055	0.524 ± 0.016

Figure 4.7 XPD's ATPase is stimulated by p44/p62 and N-p44. A) XPD's ATPase stimulation by p44 (dashed) or p44/p62 (white) are shown as fold changes in k_{cat} from XPD and DNA (black) on different DNA substrates. Experiments containing p62 alone and DNA showed no stimulation of XPD. B) Exact k_{cat} values are shown in the table below with S.E.M as error. C) Helicase data (from Dr. Jochen Kuper) shows p44/p62 accelerate XPD's ability to unwind DNA more so than N-p44 on an undamaged fork substrate (shown below the helicase data and described in section 2.8). Errors are S.E.M from at least 3 repeats. * = $p < 0.007$, n.s = not statistically significant using student's t-test.

4.7 The p44/p62 complex binds and diffuses on DNA

We then investigated how p44/p62 behaved without XPD, and if they are able to interact with DNA independently. To do this we used the single molecule DNA tightrope assay (108, 109, 113). Purified His-tagged p44/p62 was conjugated to a fluorescent QDOT through an Anti-His IgG antibody as described previously and imaged using a fluorescence microscope (109, 287). Strikingly, p44/p62 was observed to bind and diffuse along double-stranded lambda DNA, the first evidence of these factors being able to bind DNA independently of XPD (**figure 4.8**). Approximately 80% of p44/p62 molecules demonstrated diffusion on the DNA (n = 599 total), with several types of behaviour distinguishable when videos were transformed into kymographs. A kymograph is a transformation of the position on the Y axis, through every frame – the X axis (shown in **figure 4.8**). The intensity of molecules bound to DNA in each frame is plotted on the X vs Y coordinates. Streaks moving in the Y dimension indicate the protein-QDOT moving on DNA, linear streaks indicate static DNA binding. Multiple behaviours of the p44/p62 complex were observed. Firstly, the diffusion of this complex was consistent with molecules sliding along the DNA, supported by the inability of molecules to cross one another on the DNA. Secondly, although most complexes diffused, a number were seen to stop, often at the same location on the tightropes. This may indicate a pause at a damage site or a specific sequence. Finally, fluorescence intensity fluctuations of the same molecule over time suggest oligomers might be forming on the DNA.

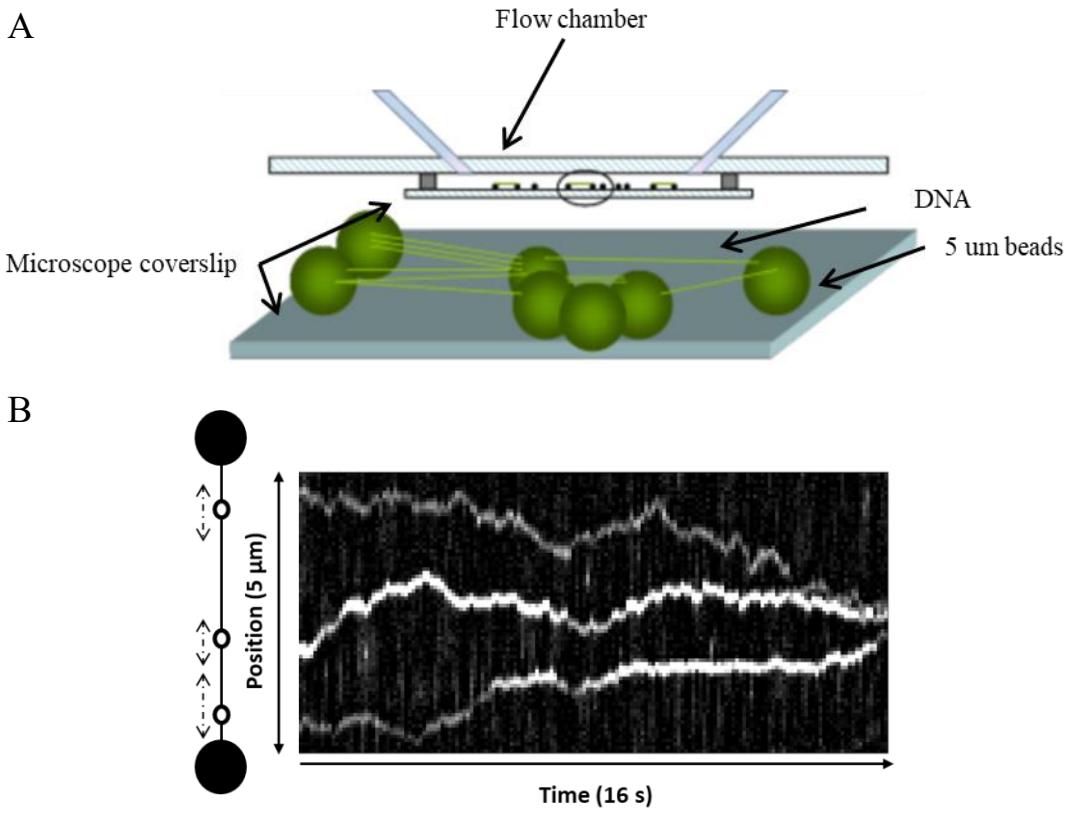


Figure 4.8 Tightrope schematic and kymograph of p44/p62 diffusion
A) A microfluidic flow chamber where 5 µm beads are adhered to the surface, and DNA is flowed over them, forming DNA tightropes (green strands)
B) Kymograph of p44/p62 molecules diffusing on lambda DNA tightropes. Three different binders (also shown on a tightrope for clarity) can be seen diffusing along DNA in a random fashion, but do not cross each other.

4.8 Labelling does not affect XPD's ATPase

To show that p44/p62's activity is not adversely affected by labelling with a QDOT, the ATPase assay was used. The ability of p44/p62 to stimulate XPD's ATPase on single-stranded DNA was tested, as this substrate previously showed the greatest activity (**figure 4.7**). It is not possible to use QDOTs in this assay as they absorb light at 340 nm, preventing the turnover of ATP to be measured. Instead, the primary antibody (penta-His) was used. Firstly, to make sure that XPD was not being affected by labelling, XPD was pre-incubated with penta-His antibody and then unlabelled p44/p62 was added. The stimulation of labelled XPD's ATPase was the same as unlabelled XPD. Then, labelled p44/p62 was added to unlabelled XPD, also showing no difference from the stimulation by unlabelled p44/p62. These results are shown in **figure 4.9** with graphical representations of which protein was labelled. These data show that p44/p62's ability to interact and stimulate XPD's ATPase is not affected by being conjugated to an IgG antibody.

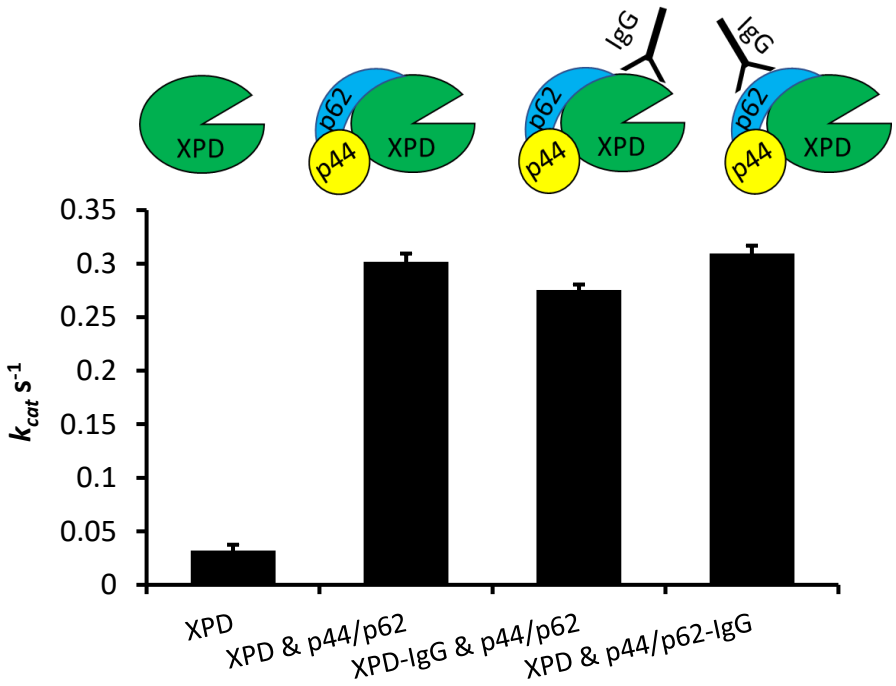


Figure 4.9 IgG labelling does not affect XPD's ATPase. XPD's ATPase is not affected by either XPD being labelled with penta-His antibody, or if the stimulating complex p44/p62 is labelled with penta-His antibody. Mean of three repeats plotted as k_{cat} values, with S.E.M as error bars. Graphic representation shows which protein in the assay is labelled with an antibody.

4.9 p44/p62's interaction with the DNA is salt dependent

The kymographs of p44/p62 molecules interacting with DNA showed that multiple molecules on the same DNA molecule were not observed crossing each other, suggesting they slide along the helical backbone of DNA and cannot pass each other. To investigate this further, we performed experiments changing the salt concentration. As described in the introduction (section 1.8) changing the salt concentration and examining the behaviour of molecules allows sliding to be differentiated from hopping. At higher salt concentrations (100 mM vs 10 mM KCl) fewer molecules of p44/p62 were observed binding to DNA, and of these, a lower percentage diffused (55%, n=58 total). At salt concentrations > 100 mM KCl, almost no molecules bound to DNA and this was not pursued further. The dependence of the p44/p62:DNA interaction on the salt concentration suggests that the electrostatic interaction with the backbone is being affected.

4.10 p44/p62 slide along the DNA helix

Proteins can use multiple modes to scan DNA for their target site, and these can be used in combination (section 1.8). Proteins can either skip along the backbone of the helix (hopping), or it can corkscrew around the backbone following the groove of DNA (sliding). Both of these mechanisms are shown in **figure 4.10**. Hopping generally facilitates a very rapid diffusion constant, as the protein makes transient contacts along the DNA. Sliding however, has the impairment of rotation-coupled drag as the protein follows the turns of the DNA helix. To distinguish between these modes the diffusion constant in different salt concentrations needs to be calculated.

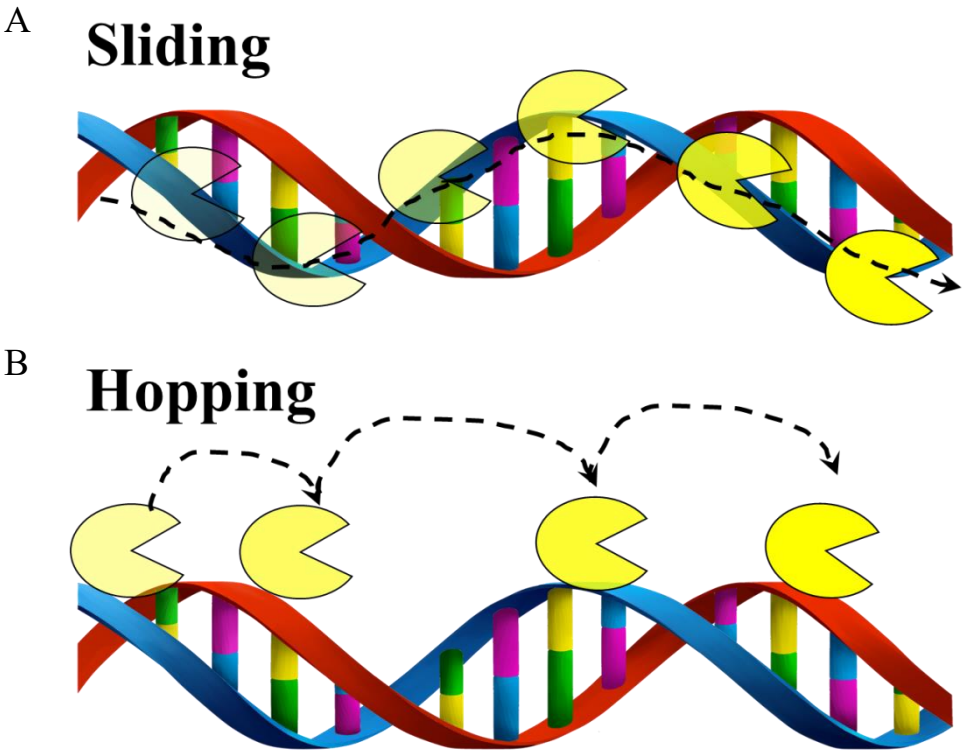


Figure 4.10 Schematic of sliding versus hopping. A) Proteins can slide along the DNA backbone. The yellow major sector represents a protein moving along DNA following the rotation of the helix. The arrow shows the trajectory along the backbone. B) Proteins can hop along the DNA, shown as stepping across the DNA backbone, but not following the rotation of the double helix. The two methods can be separated by their diffusion constants.

To calculate the diffusion constant of p44/p62 molecules in different conditions, each frame of a kymograph is fit to a Gaussian distribution to obtain the position of the molecule with greater resolution above that of the microscope (super-resolution). A sequence of the molecule's positions over time give a trajectory, and these can be used to calculate the mean squared displacement (MSD) – how far a protein has deviated from its starting position over time (described in section 2.18). From fitting the MSD data, a mathematical constant (alpha) which describes diffusion can be obtained; where <1 is constrained diffusion, 1 is unbiased diffusion, and 2 is directed motion. For p44/p62 molecules we find the alpha values ~ 1 (**table 4.1**) which suggests that p44/p62 is performing a random walk on DNA in both salt concentrations. From the MSD we can calculate the diffusion constant, how fast the molecules diffuse on DNA. We obtain a value of 0.067 $\mu m^2/s \pm 0.006$ in 10 mM KCl and 0.042 $\mu m^2/s \pm 0.010$ in 100 mM KCl. A plot of alpha versus diffusion constant is shown in **figure 4.11** with the average values for both axes shown in **table 4.1**. Diffusion constants are distributed over a logarithmic scale and are not significantly different between salt conditions ($P = 0.0457$). These data suggest that p44/p62 molecules slide along the DNA helix, rather than hop along the backbone (260, 263, 264, 266, 267, 269).

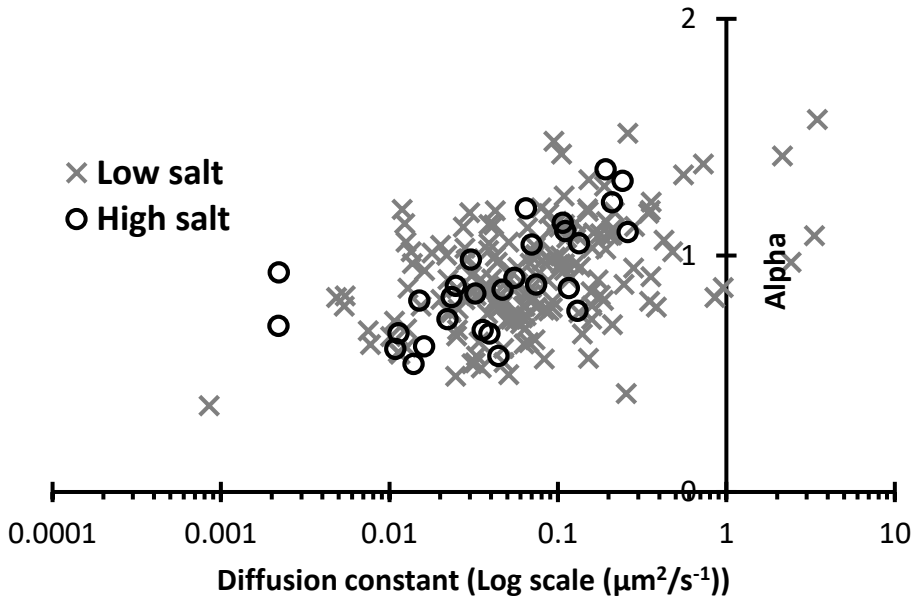


Figure 4.11 Diffusion constant versus alpha plot. Diffusion constants are normally distributed in log space (X axis). Alpha values centre around 0.9, showing p44/p62 perform a random walk in both high salt (circles) and low salt (crosses) buffers. Diffusion constant, alpha values and n's are shown in the table below.

	High salt	Low salt
Alpha	0.890 ± 0.041	0.910 ± 0.016
Diffusion constant	0.042 $\mu m^2/s$ ± 0.010	0.067 $\mu m^2/s$ ± 0.006
n flow cells (molecules)	4 (29)	8 (177)

Table 4.1 Diffusion constant and alpha values for p44/p62 in high and low salt buffers. Values are given ± S.E.M. For the diffusion constant the S.E.M was calculated in log space and propagated to normal space. Student's t-test using flow cell n, shows no significant difference between the diffusion constant in high and low salt (P = 0.0457).

To further confirm that p44/p62 slide along the DNA, the friction of p44/p62 moving along the DNA can be calculated using equation 1.1 where *r* is the radius of the protein in nm (269).

$$\boldsymbol{RC = 1 + \frac{4}{3}(2\pi)^2(\frac{r/1}{3.4})^2} \qquad \textbf{(1.1)}$$

Globular proteins are modelled as spheres to determine the hydrodynamic drag of a complex moving in a solution using Stokes` law. As the dimensions of the p44/p62 complex are not known, they have to be approximated. The molecular mass of a heterodimeric p44/p62 complex is ~106 kDa, similar to a monomeric *E.coli* UvrA molecule (105 kDa) which has a radius of ~ 6 nm; this was used for future calculations (90). In our experiments we label p44/p62 with an anti-His tag antibody, and then an anti-IgG QDOT which increases the radius of the complex to approximately ~13 nm (from manufacturer Thermo scientific). Taking these together and finding the centre of mass of the whole complex (cubing and cube root (109)), the total radius of p44/p62:QDOT complex is 13.3 nm. Using the Stokes-Einstein equation (1.2) the one-dimensional diffusion constant without any rotation can be calculated, this is the value we would expect if p44/p62 hopped along the DNA(269).

$$\boldsymbol{D_{nr} = \frac{k_B T}{6\pi\eta r}} \qquad \textbf{(1.2)}$$

Where k_BT is the Boltzmann constant at room temperature = 4.1 pN nm and η is viscosity (water at 25°C = 0.89), and *r* is radius (340). Using the approximate radius of 13.3 nm in equation 1.2 we would expect a diffusion constant without any rotation of 18.4 $\mu m^2/s$. This is several orders of magnitude faster than the diffusion constants we observe for p44/p62.

If we then use equation 1.1 to calculate the rotation-coupled element (*RC*) and dividing the non-rotational constant (D_{nr}) by this (equation 1.3)

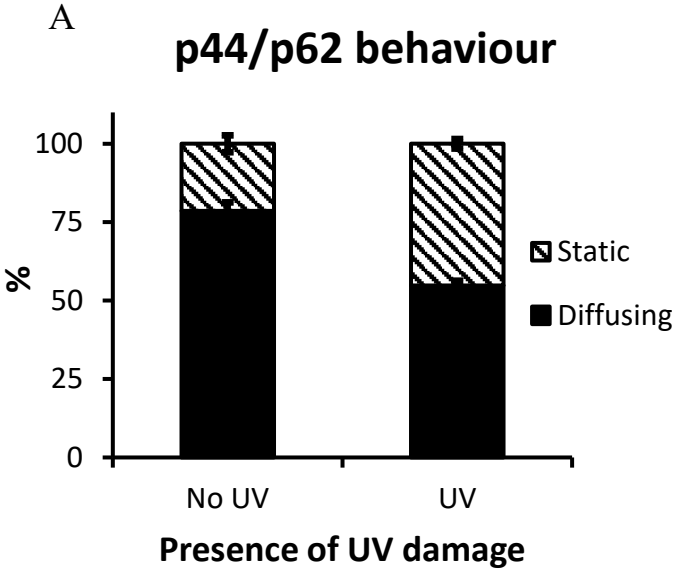
$$\boldsymbol{D_R = {D_{nr}}/{RC}} \qquad \textbf{(1.3)}$$

we obtain 0.023 $\mu m^2/s$, which is much closer to the diffusion constants determined from experimental data for p44/p62 (~0.05 $\mu m^2/s$), and for other proteins reported to slide on DNA (341). Together these data suggest that p44/p62 molecules slide along the DNA helical backbone.

4.11 p44/p62 are capable of detecting DNA damage

As p44/p62 is capable of binding to the DNA and sliding along the backbone, we investigated how the presence of UV damage on the DNA would affect p44/p62's behaviour. Lambda DNA was irradiated with 500 J/m^2 of 254 nm UV light immediately before being used to create DNA tightropes. This method has been shown to randomly induce CPD lesions on the DNA (section 2.11). When imaging p44/p62 molecules in the presence of DNA damage there was a clear shift in the number of molecules that diffused on the DNA. The number of diffusing molecules decreased from ~80% to ~55% in the presence of UV damage (**figure 4.12**). When analysing the kymographs of p44/p62 molecules in the presence of UV damage it was apparent that although some molecules were diffusing, they were restricted in the total excursion distance they could move on the DNA. As described previously, a diffusive exponent, alpha, describes diffusion; where a value of 1 shows unobstructed random walking, $\alpha < 1$ suggests confined diffusion, and $\alpha > 1$ shows directed motion. When p44/p62 molecules alpha values were plotted as a histogram there was a shift to lower alpha values in the presence of UV damage (**figure 4.12**). This shows that when the DNA contains UV damage p44/p62 molecules become restricted in their ability to freely diffuse on DNA, similar to XPC (175, 176). This suggests p44/p62 are capable of recognising damage in double-stranded DNA, and this leads to molecules either stalling or changing from a random walk to confined diffusion around the UV damage.

A



	No UV	UV
Binders	599	217
Flow cells	8	6
% Static	21.37	45.18
% Diffusing	78.63	54.82

B

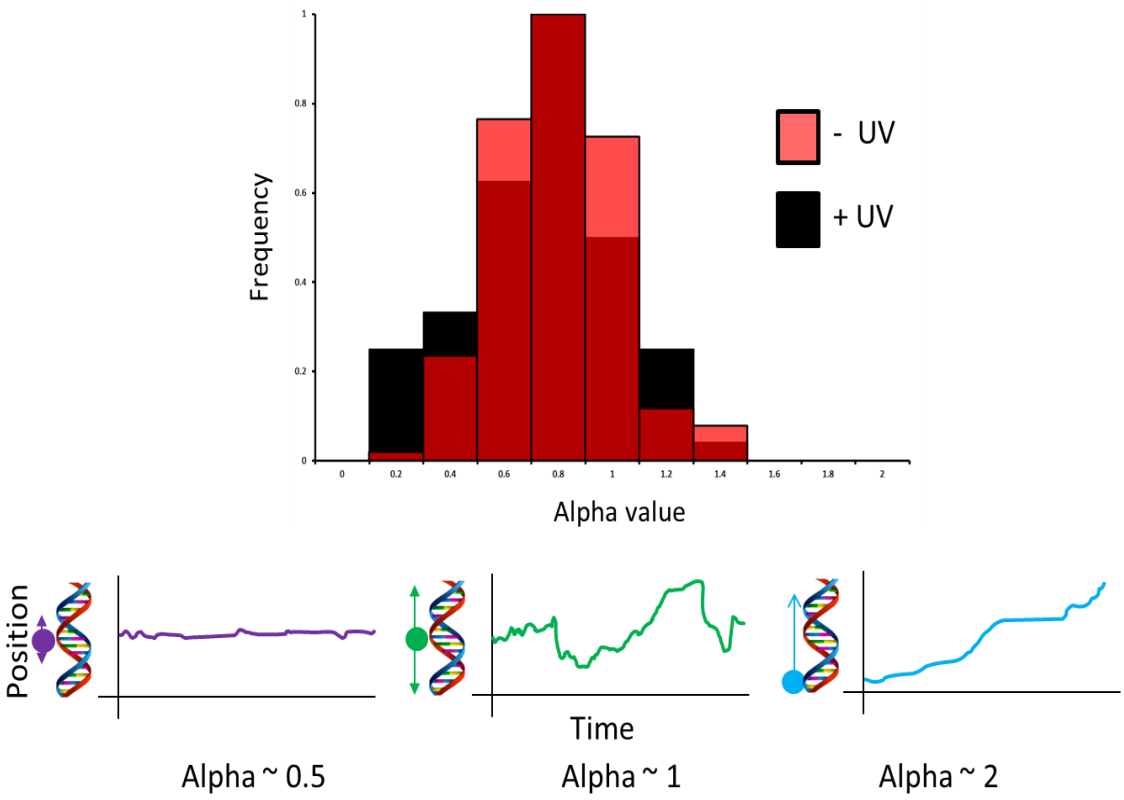


Figure 4.12 The diffusion of p44/p62 is affected by UV damage.
A) The distribution of diffusing and static p44/p62 molecules is affected by the presence of UV damage on the DNA. The values for the figure are shown in the table with average percentages of diffusing and static molecules plotted. Error bars are S.E.M using total flow cell repeats.
B) Histogram of p44/p62 alpha values determined using MSD analysis (described in section 2.18) show a leftward shift when the DNA contains UV damage. Example kymographs below demonstrate protein movement at each alpha value. A total of 150 molecules in the absence of UV, and 72 molecules in the presence of UV damage were used to construct the histogram after being normalised to the total frequency in each condition.

4.12 p44/p62 has a greater affinity for single-stranded DNA

In vivo, TFIIH is bound to a double-stranded: single-stranded DNA junction (ds-ssDNA junction) around a repair bubble (206, 342, 343). To investigate the binding of p44/p62 to this substrate, lambda DNA was ligated with overhanging oligonucleotides to create a region of single-stranded DNA flanked by double-stranded DNA (**figure 4.13**). To localise regions of single-stranded DNA, a ligated oligonucleotide with a biotin moiety covalently attached to the 5' end can be labelled with a streptavidin QDOT. As p44/p62 contains a His-tag this allows us to use two different wavelength QDOTs to examine the interaction with no risk of cross-labelling occurring.

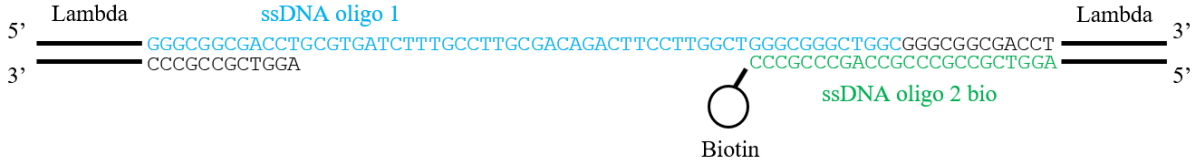
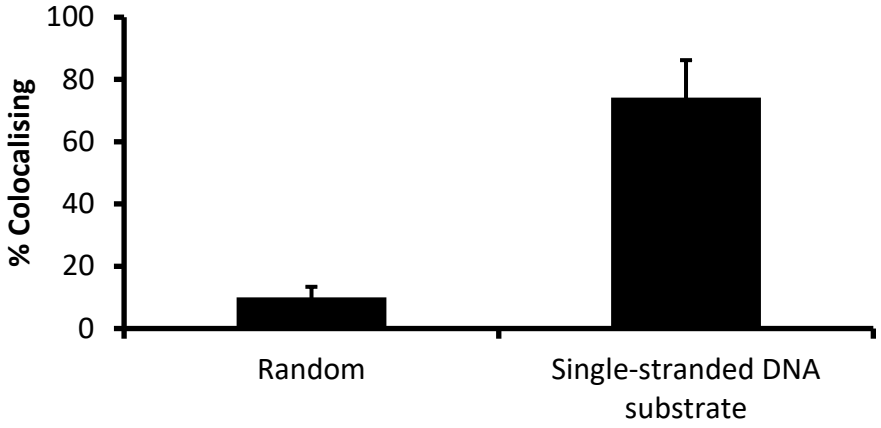


Figure 4.13 The single-stranded lambda substrate. The 12 nucleotide overhang at each end of lambdas cohesive ends are used to ligate a 5' phosphorylated oligonucleotide -ssDNA oligo 1 (blue) and a biotinylated oligonucleotide - ssDNA oligo 2 bio (green) (see section 2.12). The ligated oligonucleotides create a 38 nucleotide single stranded patch. The biotin on the 3' of ssDNA oligo 2 bio allows the site to be labelled with a streptavidin-QDOT and visualised via fluorescence.

In the DNA tightrope assay the probability of a protein colocalising with a random locus on DNA is ~10% (108), anything above this threshold suggests two populations are specifically interacting. When p44/p62 molecules were imaged binding to DNA tightropes that contained a single-stranded region of DNA, the two different coloured QDOTs colocalised 74.2% of the time (**figure 4.14** (n = 31)), significantly higher than the 10% threshold, and demonstrating that p44/p62 has a high affinity for single-stranded DNA.

A



B

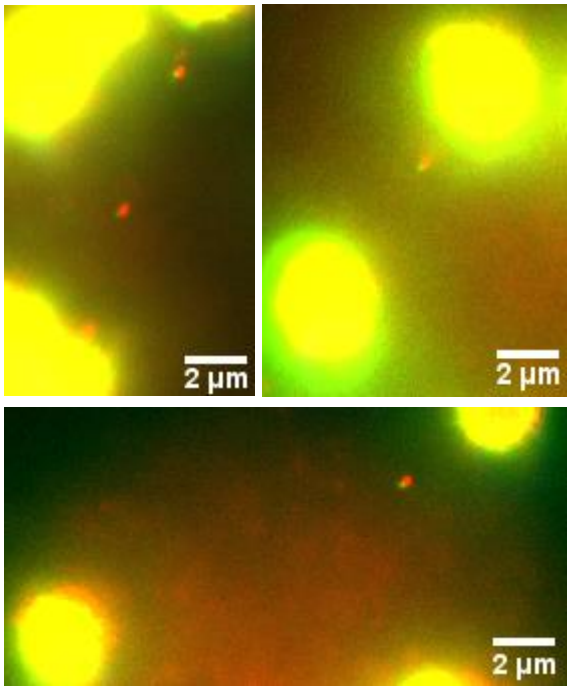


Figure 4.14 p44/p62 preferentially interacts with single stranded DNA.

A) Graph showing the probability of random colocalisation between DNA and protein (108) and p44/p62's colocalisation with single-stranded regions on DNA. Out of 31 single stranded DNA patches observed, 74.2% had a p44/p62 molecule bound. Errors are S.E.M of 3 flow cells.

B) Images of p44/p62 (red) colocalising with the labelled DNA (green). Image overlays are offset to allow clear identification of dual colour complexes on DNA.

Surprisingly, streptavidin QDOTs that are used to identify regions of single-stranded DNA were observed diffusing along the DNA tightrope, even when they did not colocalise with a labelled p44/p62 molecule (there is an excess of unlabelled protein). Movement of a DNA bound oligonucleotide might suggest an unlabelled p44/p62 molecule is able to bind the ligated oligonucleotide while diffusing on lambda DNA. Therefore, to distinguish between an experimental artefact and a discreet activity of p44/p62 this was investigated further.

4.13 Single-stranded DNA tightrope additional experiments

To further investigate the observation that DNA-bound QDOTs were observed moving, a series of experiments were carried out. To ensure diffusion of the DNA-bound oligonucleotide is not light bleeding through from a p44/p62 molecule (besides the fact that not all diffusing molecules colocalised with a labelled p44/p62 molecule) the experiment was repeated with unlabelled p44/p62 protein, however DNA-bound QDOTs were still observed to diffuse. The non-specific binding of QDOTs to lambda DNA was excluded by labelling double-stranded DNA with YOYO-1 (nonspecific DNA dye) that can be excited by a 488 nm laser to show the presence of DNA tightropes. A 561 nm laser was then used to specifically excite QDOTs (655 nm streptavidin with no protein or DNA attached), no QDOTs were observed to bind to lambda DNA, even when the QDOT was conjugated to the single-stranded oligonucleotide. To exclude crossover between QDOT labelling, p44/p62-His was incubated with streptavidin QDOTs and imaged. The streptavidin QDOTs should not bind to the protein and show no binding on DNA. Only one molecule was observed binding to DNA after imaging 55 tightropes, showing that there is no crossing over of labelling. T4 DNA ligase was used to create the single-stranded gap region on DNA, and may be carried over into the flow cell and bind DNA. To eliminate this possibility, T4 DNA ligase was incubated with lambda DNA and strung up in a flow cell, the DNA oligonucleotide-QDOT was introduced, with no binding to the DNA observed. This was repeated with T4 Ligase introduced into the flow cell once the DNA tightropes had already been formed, and then adding the DNA oligonucleotide-QDOT. Finally, the T4 DNA ligase was incubated with the DNA oligonucleotide-QDOT complex and then introduced onto tightropes. In all cases there was no observed binding of the oligo to the DNA, suggesting that T4 DNA ligase does not explain the diffusion observed. This suggested that there is a specific interaction between p44/p62 and the single-stranded-DNA:QDOT. However, the DNA-bound QDOT was observed diffusing even in the absence of p44/p62, although this was not reproducible. At the end of these experiments, there was no plausible explanation for the observation, and this was not pursued further. The colocalisation of p44/p62 and the single-stranded DNA was reproducible therefore the data are included in this thesis.

4.14 p62 alone does not bind DNA, but N-p44 does

Next, we examined the ability of N-p44 and p62 to individually bind DNA in the same experimental conditions used for the p44/p62 complex. Experiments with full length p62 showed no DNA binding in either 10 mM KCl imaging buffer, or 100 mM KCl. Increasing concentration of the protein either at labelling, or at the final concentration (~15 nM) had no effect on DNA binding. N-p44 that was able to stimulate XPD's ATPase did bind DNA. The single molecule experiments involving N-p44 were carried out by Luke Springall. N-p44 molecules diffused on the DNA (56.1% n=110), with a diffusion constant of 0.018 $\mu m^2/s$, similar to that of p44/p62. Together these data show p62 does not bind DNA in the experimental conditions tested here, but the N-terminal truncation of p44 can, and likely contains the DNA binding domain of the p44/p62 complex.

4.15 p44/p62 form oligomeric complexes on DNA

When examining the behaviour of p44/p62 molecules on DNA, we observed that kymographs showed fluorescence intensity fluctuations between molecules bound to the same tightrope (highlighted in **figure 4.15**). QDOTs are known to have varying fluorescence intensities (344) but as p44 and p62 have a well characterised structural role in TFIIH it is possible we are observing multiple molecules interacting on the DNA. If several labelled molecules are in close proximity this places multiple fluorescent QDOTs within a single pixel, giving an integrated brightness that is observed in the kymographs.

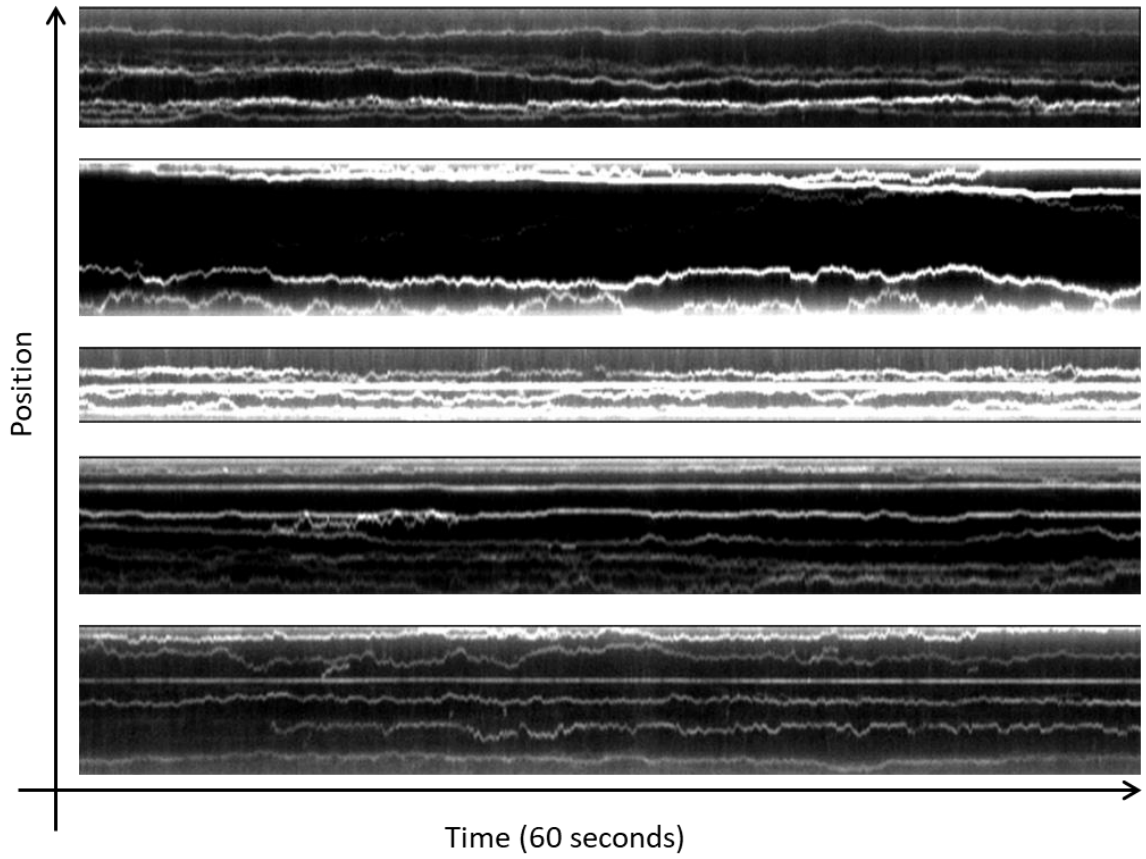


Figure 4.15 Multiple kymographs showing varying intensities of molecules. Five kymographs are aligned and set to the same scale of brightness & contrast using ImageJ. Various types of behaviour can be seen in these kymographs, including the fluctuations in intensity across a kymograph.

To investigate if p44/p62 formed dimers, one population of p44/p62 molecules was labelled with a 565 nm emission QDOT (green), and one population with a 655 nm emission QDOT (red) to enable each population to be spectrally separated. Both protein conjugates were mixed together and introduced into the flow chamber. However, because unlabelled protein represents 2/3 of the sample for each colour, it is difficult to predict the distributions of species we expect to observe if 100% of molecules form dimers. Dark protein can form dimers with other dark protein, or with a green or red protein. We cannot observe these dark dimers (DD) so they are removed from the probability table. However, this over populates the predicted observations of GD or RD complexes because every complex is assumed to be a dimer, whereas in reality there will be some single colour proteins. This under populates the RG population, but the predicted outcomes are show in **table 4.2**.

A

Probability table	**Red**		**Dark**	**Dark**
Green	**R**	**G**	GD	GD
Dark	DR		DD	DD
Dark	DR		DD	DD

B

Complex	**Probability table**	**Dark-Dark removed**	**Observed**
GD	22.25%	40%	41.67%
DR	22.25%	40%	33.89%
RG	11.25%	20%	24.44%
DD	44.25%	N/A	N/A

Table 4.2 Predictions of complexes formed in a dual colour experiment. A) Assuming all p44/p62 molecules form dimers, the distribution of complexes observed can be predicted. As only 1/3 of the protein of each colour is fluorescently labelled, this must be accounted for as dark protein. Because dark-dark complexes cannot be observed they are removed from the probability table in B). This increases the GD, DR, and RG populations (Dark-Dark removed column). The experimentally observed outcomes are shown in the table. n = 180 total binders from 3 flow cells.

In the experiment 24.44% of the total binders colocalise as a dual colour complex on DNA shown in **figure 4.16**, this is highly consistent with p44/p62 forming dimers on the DNA, but we are unable to certainly say this because of the dark protein present. However, supporting the fact we observe interactions between p44/p62 molecules, we observe single colour complexes breaking apart and reforming on the DNA (**figure 4.16**), suggesting the interaction between p44/p62 molecules is dynamic under these conditions.

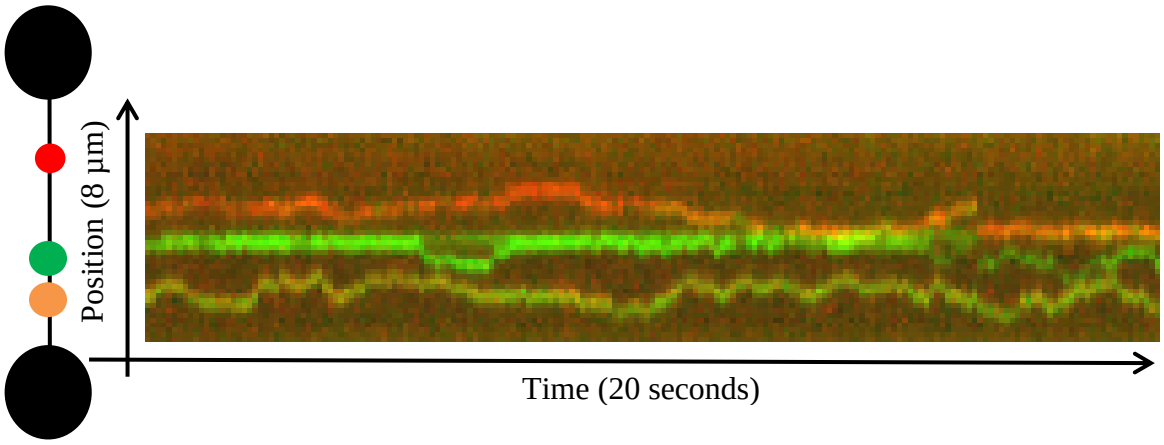


Figure 4.16 Kymograph showing dual colour complexes on DNA.
One population of p44/p62 is labelled with a 565 nm emission QDOT (green), and one population with a 655 nm emission QDOT (red). The kymograph shows a red-green complex diffusing on DNA. A green-green complex can also be observed with molecules breaking apart, and then re-joining. Colocalisation performed using a fiducial marker present in all channels (section 2.16), and a custom written ImageJ script to overlay multiple channels.

To examine if p44/p62 can form higher oligomeric complexes, a third colour complex was added to the experiment. Triple colour complexes were observed diffusing together and dynamically splitting apart and reassembling on DNA (**figure 4.17**). In total, 57 binders were observed, and 16 of these had colocalisation. Here the probability of observing a triple colour population is more complicated and was not predicted. However, the qualitative observation that three different coloured populations overlay and diffuse together demonstrated that p44/p62 molecules are able to form higher order oligomeric complexes on DNA.

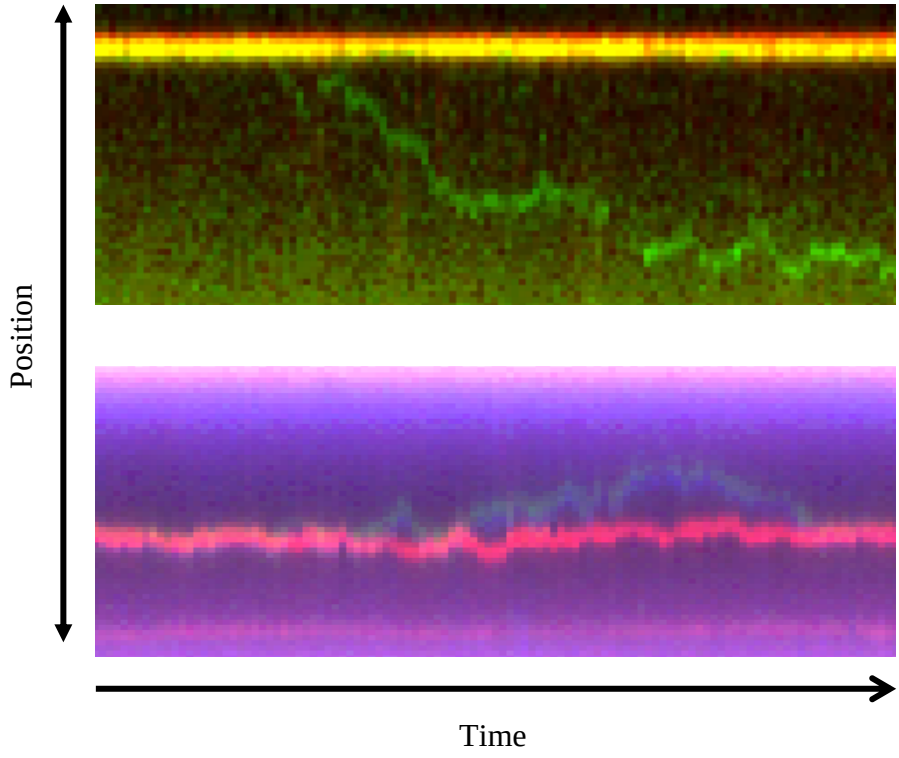


Figure 4.17 Kymographs showing triple colour complexes can form on DNA.
One population of p44/p62 is labelled with a 565 nm emission QDOT (green), one population with a 655 nm emission QDOT (red), and a third population with a 605 nm emission QDOT (blue here). The top kymograph shows a triple colour complex static on DNA, and then a green labelled molecule detaches and starts diffusing on the DNA while a red-green complex remains statically bound. The bottom kymograph shows a triple colour complex of red-green-blue diffusing on DNA together, then the green-blue complex splits apart and diffuses separately from the red labelled molecule, the three then rejoin and diffuse together again. Colocalisation performed using a fiducial marker present in all channels, and a custom written ImageJ script to overlay multiple channels.

4.16 p44/p62 form a complex with XPD on DNA

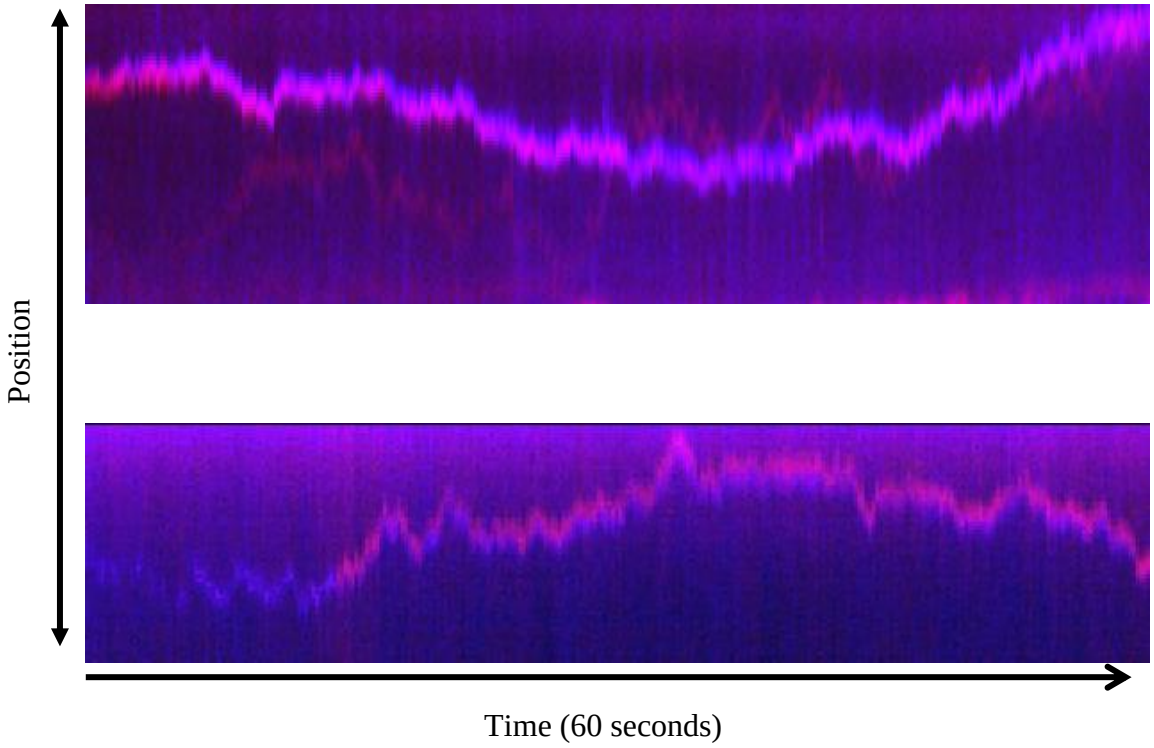
We have observed the stimulation of XPD's ATPase by p44/p62, showing they are able to interact together under these conditions. We wanted to visualise this interaction at the single molecule level and investigate how p44/p62 affect the behaviour of XPD on DNA.

In two independent experiments, XPD was observed colocalising with p44/p62 on DNA (**figure 4.18**). XPD containing the R323/324S mutation (XPD^{MUT} hereafter) was used in this experiment as it demonstrated the most consistent and reproducible results in the DNA tightrope assay. XPD^{MUT} is unable to bind MAT1 from the CAK complex, but still maintains DNA binding. Several colocalising complexes of XPD^{MUT} and p44/p62 were observed; both static and diffusing (**figure 4.18**). Of 11 observed XPD^{MUT} molecules, 6 were in complex with p44/p62, and of these, only 2 were static (**table 4.3**). Further experiments were unable to be performed due to the reproducibility of XPD^{MUT} and wild-type XPD's ability to bind and diffuse on DNA, and these data are not included in this thesis.

	Number of molecules
XPD^{MUT} binders	11
p44/p62 binders	19
Colocalising XPD^{MUT} :p44/p62	6*

Table 4.3 Number of molecules forming a complex between p44/p62 and XPD^{MUT}.
* = 4 of the colocalising molecules were static.

A



B

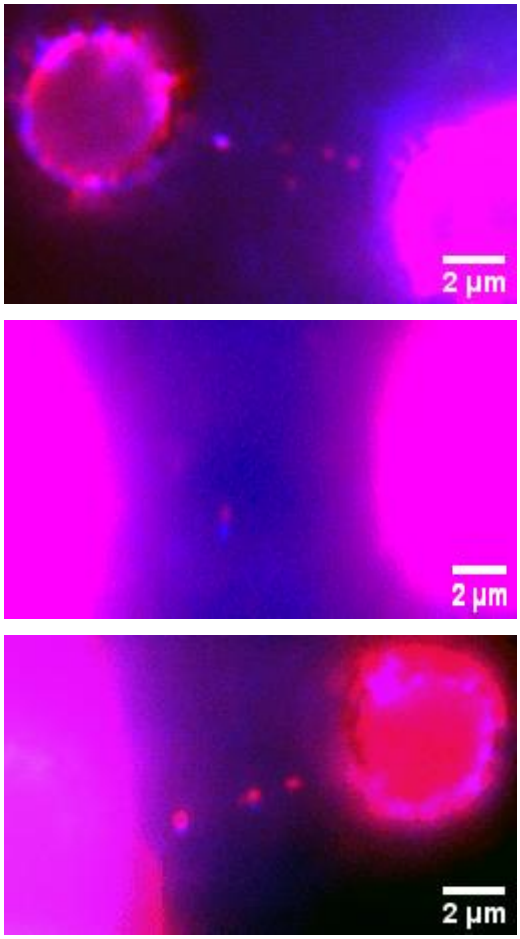


Figure 4.18 Kymographs and images of p44/p62 and XPD^{MUT} colocalising on lambda DNA.

A) Kymographs showing diffusion of XPD^{MUT} (R323/324S) molecules in a complex with p44/p62 molecules. The top kymograph shows a dual colour complex, and a single colour (red) p44/p62 molecule diffusing on DNA. The bottom kymograph shows a XPD^{MUT} (blue) molecule diffusing, and a p44/p62 molecule (red) binding from solution, and the two diffusing together as a heterotrimeric complex.

B) Images showing p44/p62 (red) and XPD^{MUT} molecules colocalising on lambda DNA. Several multicolour complexes can be seen in the images, as well as multiple single colour images. Both XPD^{MUT} and p44/p62 possess his tags, so cross-over between the labels cannot be excluded. All images shown here were formed in the absence of ATP, as this increased the diffusion of XPD^{MUT}.

Discussion

In this chapter a complex of two relatively unstudied subunits of fungal TFIIH has been observed. To our surprise, the p44/p62 complex diffuses on double-stranded DNA and preferentially colocalises to single-stranded DNA. Additionally, when the DNA is damaged by UV irradiation we see fewer p44/p62 molecules diffusing on the DNA. Our data suggest that the p44/p62 complex might act as an anchor to keep TFIIH tethered to the DNA repair bubble and may also be involved in XPD's ability to discriminate between damaged and undamaged DNA. The interaction and physiological relevance of the p44/p62:XPD interaction is currently under further investigation.

4.17 p44/p62 stimulate XPD's ATPase and aids damage verification

The observation that p44/p62 was able to stimulate the ATPase of XPD to the same level as N-p44 on undamaged DNA substrates was unexpected. The N-p44 truncation was previously observed to accelerate XPD's ATPase on some substrates (207). However, the p44/p62 construct used here contains full length, active p44 in complex with p62. This complex contains additional p44 residues not present in N-p44, it was therefore expected these residues may affect the acceleration of XPD's ATPase (177, 207). The physical interaction between p44 and XPD can be seen in **figure 4.19** where the vWA domain of p44 contacts the helicase domains of XPD.

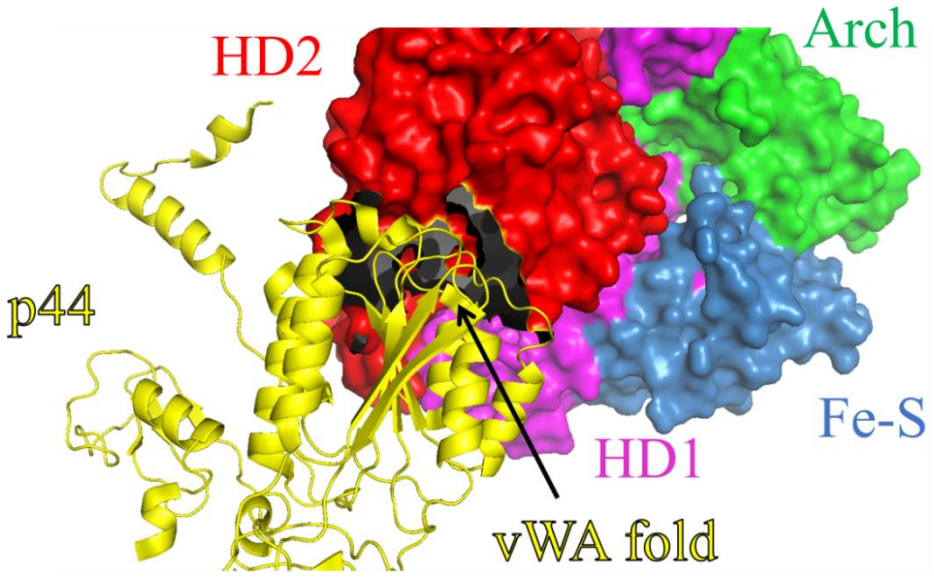


Figure 4.19 Cryo-EM structure of human p44 and XPD in TFIIH. XPD is shown as a surface plot and coloured by domains as seen previously. p44 is shown in cartoon representation, with the vWA fold (highlighted) contacting XPD in the HD2 domain. Structure created in PyMol from PDB: 6NMI (239).

p44/p62 stimulated XPD's ATPase two-fold more than N-p44 alone on damaged, double-stranded DNA. Whereas, N-p44's acceleration of XPD's ATPase is unchanged between damaged and undamaged DNA. This suggests that the addition of p62 to p44 enhances either damage detection or damage verification by XPD. The origin of this enhancement is unclear; it may result from the full length p44, p62, or be a novel activity associated with the p44/p62 complex as an entity – similar to the stimulation of XPB's ATPase by p52 and p8 (184). p62 alone can be purified, but its activity is not well defined, we observe no stimulation of XPD's ATPase, and no DNA binding in our assays. This could be because p62 requires the presence of p44 or other proteins to successfully fold. Purification of active p44 requires co-purification with p62 and this may suggest both of these proteins require assistance to fold correctly. Intrinsically disordered proteins have been shown to form stable 3D structures upon binding to a specific target ligand (345) and this may explain why the co-purified p44/p62 complex can act as an entity to stimulate XPD whereas p62 alone cannot.

4.18 p44/p62 bind to DNA independently of TFIIH

Both p44 and p62 contain putative DNA binding domains (318, 334). Here, the ability of N-p44 to bind DNA in the tightrope assay has been shown (Luke Springall, unpublished observations) while we find p62 does not bind DNA under our experimental conditions. We show that a full length complex of p44/p62 is capable of binding to double-stranded DNA where it slides along the helical backbone. This observation suggests p44/p62 contacts DNA through the p44 subunit; however it is possible that additional DNA binding sites are formed when these proteins are fully folded together as a complex. Furthermore, experiments using single-stranded DNA showed the oligonucleotide-bound QDOT diffusing on the DNA, although the cause of this was not identified, it is possible the diffusion was caused by an unlabelled p44/p62. The observation that these subunits are capable of interacting with DNA may be relevant for the formation of TFIIH. It has previously been suggested that p44 and p34 form the base of TFIIH that recruits the other subunits to form the full complex (177, 321). Our data would suggest that if p44 is important for the formation of TFIIH, then this likely involves p44's partner, p62, which has also been shown to form many protein-protein contacts from the anchor domain (177).

4.19 p44/p62 may recognise disrupted base pairing

In vivo, the roles of p44 and p62 are not clear, evidence shows they are necessary for repair, stimulating XPD's ATPase, and forming protein-protein contacts, but these are not well defined. We observe the p44/p62 heterodimer sliding along DNA alone, and in a complex with XPD. This p44/p62-XPD complex may be capable of scanning the DNA for damage. When DNA tightropes are damaged with UV irradiation we observe a clear change in p44/p62's behaviour. There is a decrease in the number of molecules diffusing and the appearance of a population that undergoes confined diffusion, similar to XPC; XPC has been shown to switch to confined diffusion upon detecting disrupted base pairing, and this is thought to recruit TFIIH to the DNA (175, 176). The interaction of p44/p62 with single-stranded DNA is similar to the observations with UV damaged DNA and could be p44/p62 recognising the single-stranded DNA exposed around the distorted base pairing, again similar to XPC (175, 176). The colocalisation of p44/p62 with single-stranded regions of DNA shows the complex has a high affinity for a substrate similar to the repair bubble created when XPD unwinds DNA around a lesion. These data, supported by the acceleration of XPD's ATPase on a damaged or single-stranded DNA, and the observed interaction with XPD suggest p44/p62 is recognising disrupted base pairing and stalling. This stalling may then recruit XPD to the DNA near the lesion. However, there is currently no evidence that p44/p62 or XPD exist outside of the TFIIH, only the fact that in archaea TFIIH does not exist, but XPD is capable of detecting and verifying damage alone (346).

4.20 p44/p62 binds as a dynamic complex to the DNA

Labelling multiple populations of p44/p62 allowed us to visualise oligomers and the dynamics of this complex on DNA. The ability of p44 and p62 to dimerise is not unexpected, due to the propensity of the vWA and anchor domains to form protein-protein contacts within TFIIH; with the vWA domain having previously been implicated in dimerization (159, 177, 316, 318-321). Here, where the p44/p62 complex is isolated they are left 'free', enabling them to interact with each other whereas in TFIIH they would be bound to their partners. Multiple molecules of p44/p62 were observed coming together and diffusing as a complex, and then breaking apart, diffusing on their own and then re-joining other molecules in a complex (**figure 4.20**). The dynamic behaviour on DNA may be related to recruiting TFIIH factors to the DNA and is similar to the telomeric maintenance proteins TRF1 and TRF2 that search DNA independently, but have a higher affinity for their target when they are in a complex together (347).

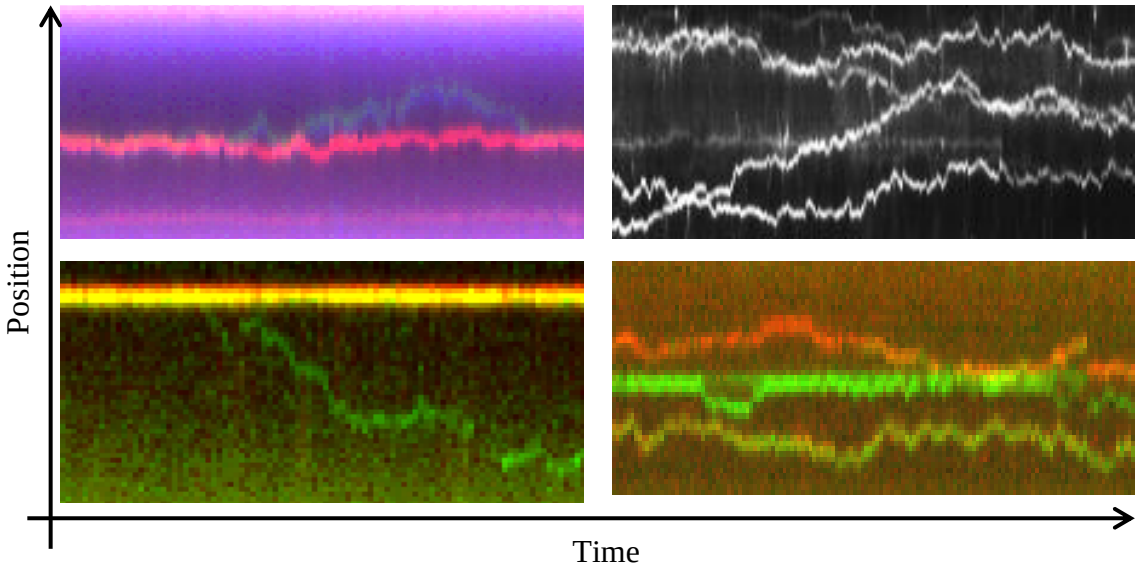


Figure 4.20 dynamic oligomers of p44/p62. Four separate kymographs showing the dynamic behaviour of p44/p62 molecules on DNA. Kymographs have been taken from several figures in this chapter that show the dynamic behaviour of binding and releasing from a complex is clear

The observation that only ~24% of imaged complexes are dual colour can be explained by the excess of unlabelled protein, where the dark protein forms the majority of dimers that we cannot observe in our assay. Dark protein will also form dimers with fluorescent protein, which would then appear as monomeric. There is also the unknown equilibrium transition from monomer to dimer, and higher oligomeric states that contribute to a low overall % of dual colour complexes on DNA. Fortunately, the dynamic behaviour of p44/p62 is an advantage as we can see complexes of 'one colour' splitting apart and joining other molecules on DNA – further confirming these complexes contain multiple p44/p62 molecules.

4.21 p44/p62 slide along the backbone

Videos of p44/p62 diffusing on DNA allowed us to analyse the behaviour of individual protein molecules. By calculating the diffusion constant and alpha values we can describe the interaction of p44/p62 with DNA. The diffusion constant was unchanged between high and low salt conditions, suggesting the protein remains in contact with the DNA and does not bind and release as it would in a hopping-type interaction. As we would expect for an electrostatic interaction with DNA, increasing the salt concentration decreases the number of proteins observed binding to DNA, by disrupting the interaction between the protein and the DNA backbone (263). The alpha value tells us that p44/p62 freely diffuses on DNA in a random walk, independent of the salt concentration (109). Using approximate values for the dimensions of QDOT-labelled p44/p62 and the Schurr model, we can predict the diffusion constant if p44/p62 hops along DNA or if it follows the helical backbone of DNA (269). We find a rotation-coupled mechanism fits our data well, with the hopping diffusion constant several orders of magnitude faster than we observe. If the drag from the large QDOT label is removed from the diffusion constant, the speed of unlabelled p44/p62 can be determined (well described elsewhere (348)). Using the equations in section 4.19 and removing the added radius of a QDOT to model p44/p62 alone, the diffusion constant accounting for rotation-coupled diffusion is approximately 0.26 $\mu m^2/s$. This value is an order of magnitude larger than the values we observe when labelled with a QDOT, but is still far closer to our data than a hopping mechanism. Although the QDOT label does affect the speed the protein diffuses, we have demonstrated it does not affect the stimulation of XPD's ATPase (section 4.8), nor does it affect the ability to bind DNA.

Interestingly, N-p44 was able to bind and diffuse on double-stranded DNA, with a diffusion constant similar to p44/p62. This suggests that part of the zinc binding domain in p44 is capable of binding to DNA, and this is the mechanism for p44/p62 binding to DNA. The missing domains in p44 may alter the interaction with DNA, and we are unable to distinguish this because of the difficulty purifying active, full length p44 without p62. Imaging p62 alone we did not observe any DNA binding under our assay conditions but this may be due to incorrect protein folding as discussed in section 4.17.

4.22 XPD forms a complex with p44 & p62

We attempted to show that XPD and p44/p62 are able to form a complex on DNA tightropes, but this was difficult due to the poor reproducibility of both XPD and XPD^{MUT}. Both XPD^{MUT} and p44/p62 have His-tags and are labelled using anti-His antibody, and IgG QDOTs – therefore we cannot exclude the possibility of crossover between these molecules being observed instead of a real XPD:p44/p62 complex. Additionally, there were great difficulties getting this experiment to work. The behaviour of XPD on DNA tightropes was not reproducible, with batch-to-batch variation of protein and its ability to bind DNA. XPD^{MUT} appeared to be more reliable and was therefore used in this experiment. The mutation in the arch domain of XPD^{MUT} is not thought to affect any interaction with p44, p62, or compromise XPD's ability to bind DNA. Although the statistics suggest a high percent colocalisation of XPD and p44/p62, this is skewed by the fact that few videos were captured on DNA, and these observations do not detract from complexes actually observed – i.e. you cannot observe what you cannot see. The few observations of XPD:p44/p62 may be due to the low protein concentrations used in single molecule experiments. The K_D value for the interaction is currently unknown, but could explain the few complexes observed. In the bulk biochemical assays the concentrations of proteins are much higher (100 nM) and here a complex may be forming more readily, explaining the stimulation of XPD's ATPase. Additionally, interactions between XPD and p44/p62 are seen using higher concentrations of protein in gel based assays and during purification (Dr. Jochen Kuper), which along with the ATPase, helicase, and single molecule data support the existence of a XPD:p44/p62 complex.

4.23 p62's position in DNA-bound TFIIH

In the most recent TFIIH:DNA structure the orientation of the DNA being passed from XPD to XPB exposes a single-stranded region towards the HD2 domain of XPD (**figure 4.21**) (208), this could allow p44/p62 to interact with the translocating strand of single-stranded DNA where they can stall upon detecting damage and trigger XPD to verify the presence of damage.

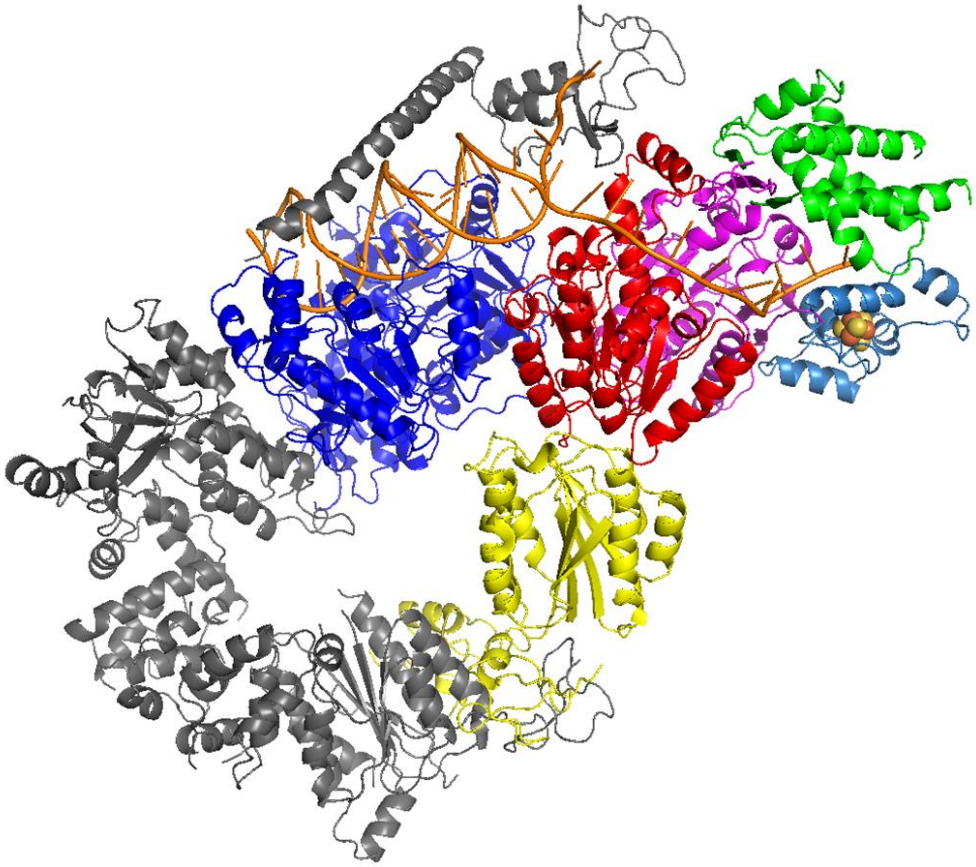


Figure 4.21 Cryo-EM structure showing DNA passed from XPD to XPB. The DNA-bound cryo-EM structure of TFIIH shows the DNA (orange) being passed from XPD's helicase domains (HD1 magenta, HD2 red) to XPB (dark blue) after being unwound by the Fe-S domain (light blue). The arch domain of XPD is shown in green and the iron-sulphur cluster is shown as space fill atoms. The remaining TFIIH subunits apart from p44 (yellow) are shown in grey for clarity. Structure created in PyMol using PDB:6RO4 (208).

In the Greber 2019 structure ~ 20 residues of p62 are occluding the DNA and nucleotide binding residues in XPD's helicase domains (**figure 4.22**); although this is in transcriptional TFIIH where MAT1 is inhibiting XPD's helicase activity via the Arch domain (239). However, this may indicate that p62 has a role in regulating either DNA or nucleotide binding to XPD. Unfortunately, in the TFIIH structure loaded on DNA, the position of p62 is not assigned in the final structure (208). However, in the initial electron density maps there is a density that overhangs the HD2 domain of XPD but is removed in the final structure (**figure 4.22**). It is possible this is part of p62, the position of the density would place p62 in a region where it would be able to interact with the helicase domains of XPD, and the single-stranded DNA. It is therefore conceivable that p62 contacts the DNA and XPD, while p44 makes extensive contacts with the HD2 domain of XPD (**figure 4.22**). In this model, we suggest that p62 would chaperone and support the binding of DNA to XPD, while also contacting the helicase and arch domains, all of which are important for XPD's helicase activity (159, 207). These contacts provide an explanation for p44/p62 interacting with single-stranded DNA, XPD, and stimulating XPD's ATPase. Additionally, the helicase assay data show that on undamaged DNA XPD successfully unwinds more substrate with p44/p62 than with N-p44. This could be due to the addition of p62 keeping XPD engaged with the DNA, increasing the number of successful unwinding events, Further supporting this, p62 is required for DNA repair (331) and the yeast homologue, Tfb1 is necessary for cell viability. Truncations in Tfb1 cause a specific defect in NER and sensitize yeast to UV irradiation (326, 349). Mutations in XPD, p44 or p62 could therefore disrupt the vital damage verification step in NER and cause disease.

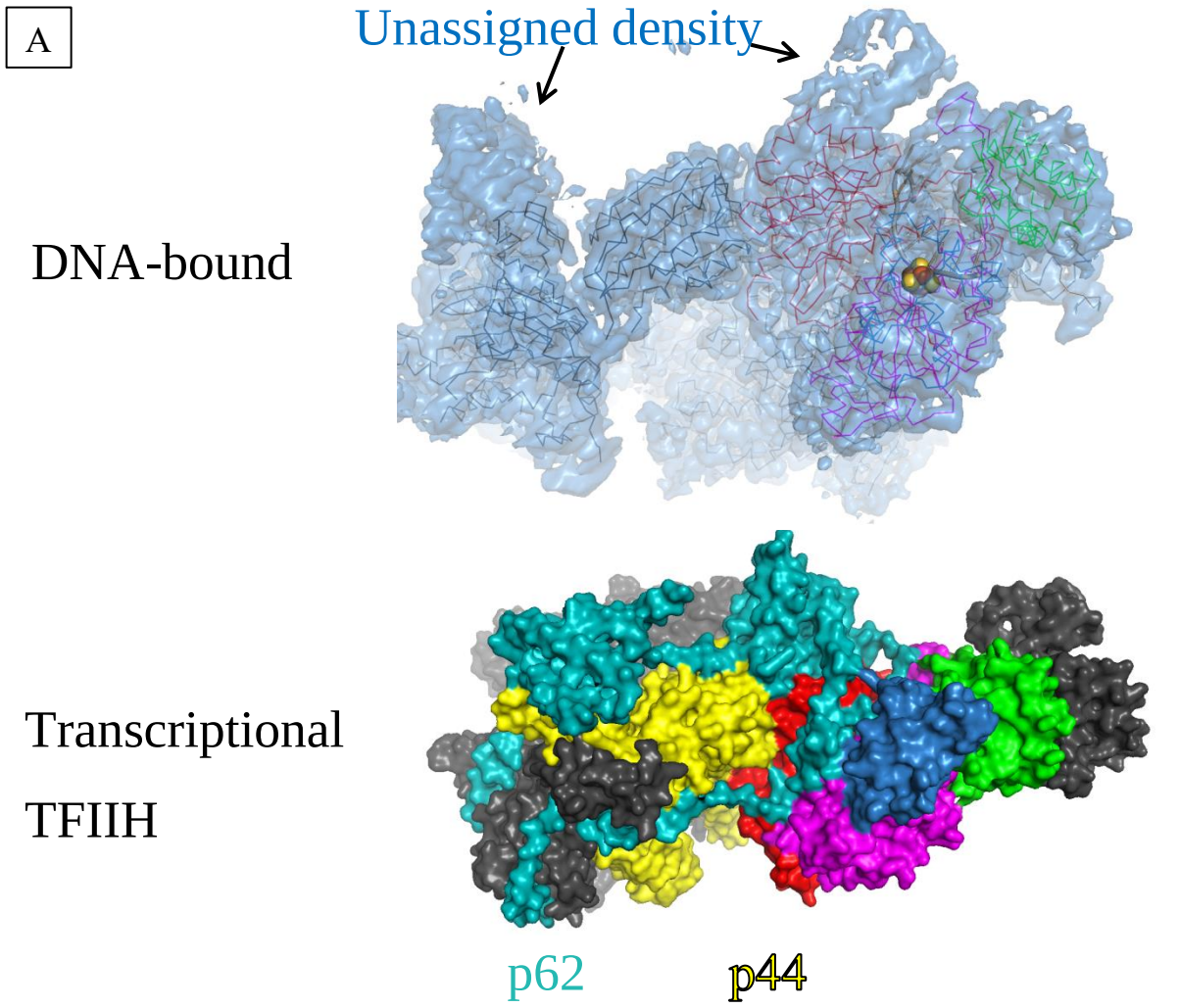

A
Unassigned density
DNA-bound
Transcriptional
TFIIH
p62
p44
B
DNA-bound
Transcriptional TFIIH

Figure 4.22 Comparing the recent cryo-EM structures of TFIIH. A) When TFIIH is in the DNA bound state p62 cannot be assigned (208). However, when examining the unassigned electron density map of the DNA-bound structure (transparent blue shown as surface density) this looks similar to the position of p62 (teal) in transcriptional TFIIH (239). This suggests this electron density represents a part of p62.

B) Closely comparing the two structures, the DNA bound TFIIH shows an unassigned electron density (transparent blue shown as a surface density) overhanging XPD's HD2 domain (red) and interacting with a loop of HD1 (magenta) and close to the Arch domain (green). In this position the electron density is in close proximity to the single-stranded DNA (orange) bound to XPD's helicase domains. In the transcriptional TFIIH structure a loop of p62 (teal) sits in the DNA binding site of XPD where it would block the path of DNA. The HD1 loop is in close proximity to p62 and this seems to agree well between the two structures.

The DNA-bound TFIIH structure was constructed using the electron density map from PDB:6RO4 as a transparent surface plot overlaid with the final structure shown in cartoon format and coloured by XPD domain (208, 248). The transcriptional TFIIH structure was constructed using PDB:6NMI as published in (239); shown as either surface density or as cartoon. XPD is coloured by domain, p44 is shown in yellow, and p62 in teal. Remaining TFIIH subunits of both structures are shown in grey for clarity, or removed from the figures here.

4.24 Summary & hypothesis

In this chapter the contribution of p44 and p62 to DNA repair has been examined. We find p44/p62 form a dynamic complex that is able to interact with DNA and XPD. Using bulk biochemical studies we show p44/p62 stimulates XPD's ATPase more on damaged DNA than undamaged DNA, demonstrating p44/p62 has an active role in DNA repair. Using DNA tightropes we have shown p44/p62 slides along the DNA backbone at a velocity limited by rotation-coupled diffusion, but p44/p62 molecules stall upon interacting with single-stranded DNA or UV damage. These results identify a novel DNA binding domain within the TFIIH complex and support a hypothesis where p44/p62 aid XPD in detecting and verifying DNA damage in a single-stranded substrate as it is unwound and passed to XPB.

4.25 Future work on p44/p62

To further demonstrate the existence of a p44/p62:XPD heterotrimer a different purification tag needs to be used for XPD. Section 4.16 showed colocalisation of XPD^{MUT} and p44/p62, but these proteins both have His-tags meaning that cross-labelling can occur. To improve this experiment, wild-type XPD should be purified with an Avi-Tag that can be labelled with biotin, and then conjugated to a streptavidin QDOT. This will allow proper colocalisation of these proteins on the DNA tightropes. Will p44/p62 be able to load XPD onto the DNA? Will p44/p62 act catalytically in loading XPD on the DNA? Will p44/p62:XPD form static complexes engaged with the DNA but not able to diffuse? Will p44/p62 enhance XPD's ability to recognise damage? These are all exciting questions that have the potential to be answered using differentially labelled XPD and p44/p62.

We observe the p44/p62 complex independently binding to DNA, identifying a novel DNA-binding entity in TFIIH. How p44 or p62 contact the DNA is currently unknown, both have putative DNA binding domains (316, 318), however, N-p44 is observed binding to DNA, suggesting this domain contains the DNA binding site. However, in the Kokic et al cryo-EM structure of TFIIH p44 is some distance from the DNA, and it is not clear how p44 might interact with the DNA from this structure. An electron density discussed in section 4.23 might define p62 in a position where it can interact with the single-stranded DNA bound to XPD, however further work including more structures is needed to confirm this. Interestingly, p62 alone does not bind DNA but this is possibly because p62 needs additional proteins to correctly fold, and this may expose additional DNA binding domains that will fit in with the TFIIH-DNA structure.

The structure of p44/p62 needs to be broken down into fragments to identify which parts could interact with DNA, and which residues contact XPD in the heterotrimeric complex. Using the recent cryo-EM structures and crosslinking data, positively charged residues can be identified and mapped onto the structures to see if they are in close proximity to the DNA. Once specific parts of p44/p62 have been associated with activities such as DNA binding and crucial for stimulating XPD, site-specific mutants should be designed and then reconstituted into TFIIH. This will allow the individual activities observed to be placed in context of TFIIH. This will really expose the direct role of p44/p62 and the consequences for NER and transcription if these subunits are mutated.

The helicase assay works by having Cy3 and a fluorescence quencher (Dabcyl) in close proximity at the 3' end of the translocating strand, as XPD unwinds the DNA, the Cy3 and Dabcyl are separated, and the quenching of Cy3 is relieved, causing an increase in fluorescence output. To expand on the helicase assay several additional experiments should be performed. We observe an increase in XPD's ability to unwind the DNA in the presence of p44/p62, versus N-p44. We suggest this is most likely due to p44/p62 supporting DNA bound to XPD and a subsequent increase in processivity. Unfortunately, when the DNA substrate is damaged there is no unwinding observed, probably because XPD is stalled on the lesion, preventing further translocation and unwinding of the DNA. To investigate XPD's interaction with damage in this assay the Fe-S cluster of XPD can be used instead of the Dabcyl. Previously it has been shown that the Fe-S cluster quenches fluorescence in a distance dependent manner (197, 209). On a DNA substrate where Cy3 is located several bases away from the damage, the position of XPD can be determined by the fluorescence intensity. This assay can be performed in bulk, or at the single molecule level with DNA immobilised on a glass coverslip. This assay will be able to show if XPD is stalled on the lesion, and how N-p44, or p44/p62 affect XPD's interaction with damage.

Furthermore wild-type XPD should be investigated for its ability to recognise a fluorescein lesion placed on DNA tightropes, this can be placed either in double-stranded DNA, a bubble substrate, or potentially in single-stranded DNA. The ability of XPD to recognise damage in these substrates can then be directly observed. This has previously been investigated in bulk assays using archaeal XPD (200), however using fungal XPD and using single molecule techniques the kinetics of XPD's interactions with the DNA can be determined, such as lifetime, relative affinity for damage in different substrates, and how this is affected by additional proteins such as p44/p62. Does XPD alone stall at a lesion only in a bubble? Does the addition of p44/p62 increase XPD's stalling at a lesion in a specific substrate? Do p44/p62 find the lesion first and then recruit XPD from solution, or is XPD able to find the damage without p44/p62?

These experiments will support the results and allow further discussion of the data shown in this chapter in context of TFIIH, the broader picture of the data in this chapter is discussed further in section 5.3.

Chapter 5

Overall discussion

5.1 Combining techniques enhances our understanding of the data

The results in this thesis have examined the mechanism of how DNA is repaired by the NER pathway in both prokaryotes and eukaryotes. Using a powerful combination of biochemical and single molecule methodologies has allowed us to cross-examine the data and use this to suggest models of how these NER proteins detect damage and contribute to DNA repair.

5.2 Discussion of UvrA's ATPase

5.2.1 Clearing up controversy

In chapter 3 the conflicting studies on UvrA's ATPase are discussed (section 3.3), most literature in the field has drawn conclusions based on catalytically dead mutants of each ATPase site; this likely explains conflicts between these studies. For example, the K646A mutant clearly affects UvrA's structure, causing aggregation and inclusion body formation (106, 293) but despite this, it is still used. We avoid the use of both K37A and K646A mutants due to these inconsistencies; instead using full length UvrA that is active in DNA repair. Using single molecule fluorescence imaging we directly visualised individual UvrA molecules binding to DNA, and how this is affected by nucleotide and the presence of UV damage. Combining the single molecule data with bulk phase biochemical experiments enabled us to show ATP turnover is necessary for UvrA releasing from DNA, previously shown using filter-binding assays (90, 101, 102). While it has been known that ATP is crucial for NER to occur (99, 120, 293), the exact role in each step of the pathway is unknown. We used a more direct approach to examine this. We showed that ATP is required for UvrA to find damage through repeated cycles of turnover at the C-terminal ATPase site, and that discriminating damage requires turnover at the N-terminal ATPase site (figure 5.1). We then used nucleotides such as ADP and ADP P_i that cannot be studied *in vivo*. We found that in the presence of ADP or ADP P_i UvrA can still bind DNA, consistent with previous studies (90, 101, 102, 106, 307), however, we observed few bound molecules, and these remained attached for a long time, suggesting UvrA's dimerization is impaired, or that UvrA-ADP has a very low affinity for DNA.

However, despite the use of biochemical and single molecule experiments we were unable to explain the changes in UvrA's ATPase upon the addition of undamaged DNA. Some groups observe an increase (97, 102) and some a decrease (99, 120). We observed a DNA-stimulated ATPase, but no further stimulation by damaged DNA, whereas a recent study showed additional stimulation on damaged DNA (303). It is interesting how this might occur in relation to the lifetimes we observed at the single molecule level. A decrease in ATPase activity on undamaged DNA suggests UvrA would remain bound to the DNA for longer (99, 120), perhaps caused by incidental damage to the DNA or the use of thermophilic protein. Alternatively, this suggests UvrA's hydrolysis of nucleotide is inhibited while bound to the DNA; we observed the opposite, with good agreement between our single molecule and biochemical data. The further acceleration of UvrA's ATPase in the presence of damaged DNA (303) may suggest a second site is activated, but that this hydrolyses ATP at a faster rate than the first site, but again we do not see any suggestion of this here, and seems unlikely considering the similarity of the ATPase sites (294, 295).

5.2.2 Support for the two-site hypothesis

We observed a longer attached lifetime in the presence of UV damaged DNA but the same rate of biochemical ATP turnover. This suggests the second ATPase site is being inhibited by the first site on undamaged DNA and highlights the negative cooperativity shown previously (120, 292, 293, 297). Reconciling our data with the literature suggests the site that turns over ATP on undamaged DNA can be assigned to the C-terminal site and that the damage activated ATPase site is the N-terminal site (**figure 5.1**). Instead of DNA damage directly accelerating UvrA's ATPase there is likely a conformational change that leads to the second ATPase site being activated. This is supported by UvrA's broad damage detection capability, the two-site hypothesis and a recent study that crosslinks UvrA's zinc fingers (302). Furthermore, we directly observe a behavioural change in UvrA that only occurs when the DNA has been UV irradiated; it will be interesting to see if UvrA's ATPase changes when interacting with different types of DNA damage, do more distorting lesions accelerate turnover? Or cause a larger conformational change that accelerates the recruitment of UvrB? If this is the case it may explain some of the discrepancies between studies that use different DNA damage substrates.

Assuming the N-terminal ATPase site is activated in the presence of damage, this site would then be responsible for signalling damage to UvrB (figure 5.1). The recruitment and loading of UvrB is still disputed and should be investigated (discussed in section 3.25). The presence of UvrB on DNA is a critical step in NER and is necessary for the incision of DNA damage from the DNA (129). The strong negative cooperativity between the two ATPase sites suggests that turnover at the second site is tightly controlled, possibly to prevent excessive loading of UvrB onto the DNA. If UvrB is continually loaded onto the DNA in an attempt to verify damage this may trigger a stress response from the cell (96, 105, 124), as well as causing longer lived roadblocks that affect transcription and DNA replication. This negative cooperativity also provides a point of regulation, so that only when base pairing is significantly disrupted does this facilitate ATP hydrolysis and lead to UvrB loading. These data fit well with the two-site hypothesis proposed by Myles and Sancar and strengthens this as the current model (119, 120, 293).

5.2.3 ATP in the UvrAB complex

So far we have studied UvrA alone; however UvrA is capable of forming a complex with UvrB. It is still not entirely clear if the UvrAB complex exists in solution, or forms on the DNA once damage has been detected by UvrA (96, 102-104, 109). Furthermore, it is not known if the UvrAB complex searches for damage *in vivo.* Although these questions have not been resolved, the data in this thesis open up many questions about the UvrAB complex, for example; how does UvrA's C-terminal ATPase contribute to a UvrAB complex that electrostatically slides along the DNA? Within the AB complex, does the second site trigger UvrA's release from the DNA? Future work should try and resolve these questions through careful kinetic analysis, and the generation of new mutants (see section 3.25), these questions will give a much deeper understanding of how UvrA's ATPase functions in damage detection and into verification of damage.

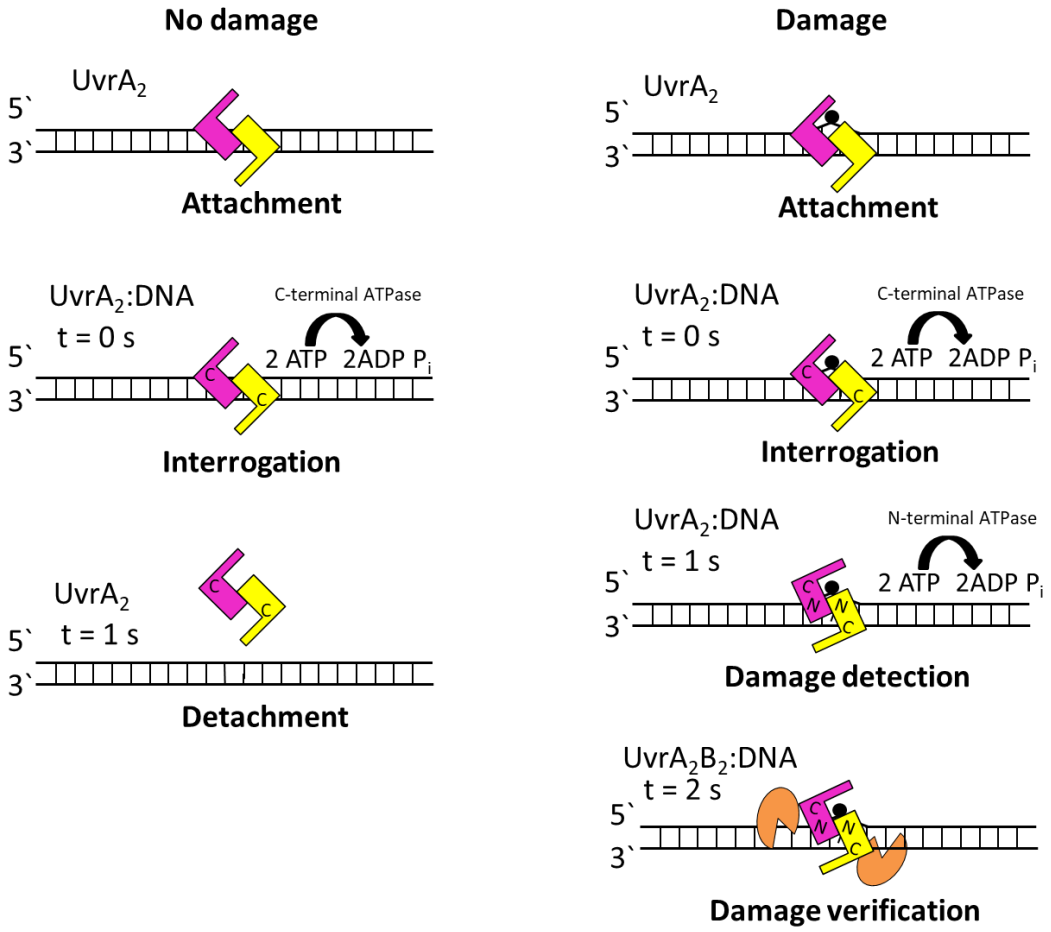


Figure 5.1 Model of UvrA's ATPase and the loading of UvrB onto DNA.
In the absence of damage UvrA turns over ATP at the C-terminal site (in each monomer) and then detaches from DNA in an interaction lasting around 1 second. In the presence of DNA damage UvrA hydroylses ATP at the C-terminal site, but the presence of DNA damage activates the N-terminal ATPase site to turnover (again, in each monomer) and this keeps UvrA attached to the DNA. The second round of hydrolysis occurs after the first site has turned over, and is likely important for recruiting and loading UvrB onto the DNA

5.2.4 UvrA's sequential ATPase

UvrA's ATPase presents a difficult kinetic problem. Because both ATPase sites were copied through evolution (294, 295) they are structurally similar and seem to turn over with a very similar rate constant, supported by the steady state ATPase data. This makes separating the function of each site difficult, as mutating one site seems to abolish all ATPase activity, or severely affect protein function (120, 293, 303). When there was no increase in ATPase activity on damaged DNA it was difficult to rationalise what might be occurring. It suggested that either UvrA cannot discriminate damage, or that DNA damage does not affect the ATPase, which seems unlikely considering we showed a change in affinity for ATP (section 3.9). However, using a combination of bulk and single molecule techniques we showed the lifetime on DNA changes most likely due to activation of the damage sensitive ATPase site, whereas on undamaged DNA the lifetime equates to a single turnover of ATP in the C-terminal ATPase site before detaching from DNA. The results presented in this chapter show an example of a DNA repair protein with a negatively cooperative ATPase. This has also been shown for MutS (350), and therefore might represent a common evolutionary path for regulating the recruitment of further verification factors to the DNA that might be harmful if not tightly controlled.

5.3 Discussion of p44/p62

5.3.1 p44 and p62 have a critical role in TFIIH

Chapter 4 investigates two subunits of the major transcription factor TFIIH and their role in regulating XPD, the key helicase in eukaryotic NER that verifies damage before excision takes place. Mutations that disrupt XPD's activity cause diseases such as XP (157, 158, 329). We are not currently aware of any clinical cases exclusively caused by a mutation in either p44 or p62. This lack of clinical cases is likely caused by a lethal phenotype, although p62 mutants have been shown to cause UV sensitivity in flies (351) and reduce the cellular concentration of TFIIH (352). It is known p44 is critical for XPD's activity, and disruption of the XPD-p44 interaction causes XP (159, 207). Here, we show that p44/p62 provide additional stimulation to XPD and most likely aid with damage verification, this suggests p62 is important and might show a lethal phenotype when mutated, especially as it forms many protein-protein interactions that stabilise TFIIH (177, 352). It would be interesting to take the crosslinking studies from Luo et al and the two high resolution cryo-EM structures (208, 239) to define the interacting residues between p44, p62 and XPD and p44/p62 to see if these are observed to be mutated in clinical cases and what the phenotype is. This may also provide evidence that p44 and p62 are important for transcription, as mutations affecting TFIIH's role in transcription are frequently associated with neurological defects (247).

5.3.2 Do p44/p62 have a role in transcription?

In the transcriptional TFIIH structure a loop of p62 sits in XPD's DNA binding site, presumably preventing XPD interacting with DNA; which is not required for transcriptional activities of TFIIH (239, 253). This suggests p62 is involved in regulating XPD's interaction with DNA, or possibly XPD's ATPase. p62 sits in close proximity to the ATP binding site within the helicase domains and this may be a mechanism of regulating DNA binding to XPD. This interaction should be explored further in the future using purified TFIIH proteins in a transcriptional assay, where the effect of p44 and p62 on XPD can be directly measured. Eventually a high resolution map of TFIIH bound to DNA in both NER and transcriptional roles will presumably answer many questions discussed here and provide a basis to generate mutants that can be reconstituted in TFIIH to examine p62's specific role in repairing DNA and in transcription.

5.3.3 p44/p62 may act as a damage sensor in TFIIH

When DNA used to make tightropes was damaged by exposure to UV irradiation, p44/p62 molecules stalled at specific locations, suggesting they could be recognising damage. This stalling also occurred on single-stranded regions of DNA. These features are remarkably similar to XPC, the main damage recognition protein in eukaryotic GGR (83, 161-163). These data might suggest that p44/p62 is capable of sliding along DNA and stalling at exposed single-stranded DNA or damage. A recent study showed XPC performs confined diffusion around a lesion, and this has been shown to enhance recruitment of TFIIH to the DNA (163, 175, 176, 178). We also observed a population of p44/p62 molecules that performed confined diffusion when the DNA contained UV damage. From the electron density maps it seems p62 contacts XPD's helicase domains (208). Here p62 could chaperone the path of the DNA through to XPB, this could prevent the DNA from being released by XPD, and explain why helicase activity was accelerated in the presence of p44/p62 but not N-p44. As there is no evidence p44/p62 would exist outside of the TFIIH complex, we suggest a model where p44/p62 act as a damage sensor within TFIIH. Here, p44/p62 could lock the DNA into XPD and detect damage in the unwound DNA, the presence of damage would stall p62 and trigger XPD to verify the presence of damage.

5.3.4 p44 and p62 are more than just structural

The results from this chapter show that p44, and more so p62 are not just structural subunits in TFIIH, but they likely have a role in modulating TFIIH's ability to verify DNA damage. These novel interactions within TFIIH may be relevant for our understanding of how DNA damage is detected and verified before incision takes place. However, much more work on the mechanism behind this, and the potential relevance for this interaction in transcription still needs to be explored. In general, the discovery that these proteins are capable of binding DNA and can modulate the activity of XPD is intriguing, having previously been designated as structural subunits. This therefore calls for further investigation of other subunits in TFIIH, such as p52 and p8 that have been shown to regulate XPB in a similar manner (184), but also for other multiprotein complexes where subunits are designated as structural without being properly characterised.

5.4 The conserved mechanisms between NER in both organisms

Single molecule studies use a reductionist approach to look at complex systems. By looking at the individual components, details about the proteins activity and function can be understood, and then built up into the overarching pathway. Both the prokaryotic and eukaryotic NER pathways share a common mechanism, but the eukaryotic pathway involves more proteins. The eukaryotic proteins have more defined roles whereas the prokaryotic proteins generally carry out multiple functions conveyed by their eukaryotic counterparts; these multi-functional proteins allow resources to be conserved in these simpler organisms. For example, UvrA finds damage in the genome and then recruits UvrB to verify the damage. In eukaryotes damage is found by the XPC-Rad23b complex, but requires UV-DDB to locate damage in chromatin (353). These proteins then recruit the multi-subunit TFIIH complex to the DNA. Verification of the damage is similar as UvrB is analogous to XPD, sharing helicase motifs and both being important in controlling incision by endonucleases. However, XPD requires p44, and as we propose here, p62 to efficiently locate or verify DNA damage. Interestingly, archaeal XPD functions without TFIIH, and instead acts independently to verify damage, similar to UvrB (159, 196, 207, 354). The contrasting delegation of roles from prokaryotic to eukaryotic likely evolved through time, necessary for dealing with a broad range of DNA damage, but requiring careful regulation to prevent unwarranted DNA incision that is harmful to cells.

In conclusion, the work in this thesis sets the basis for a deeper investigation into how nucleotide excision repair proteins verify damage and the underlying mechanism of this, and overall how this contributes to cellular survival**.** This is not the end. It is not even the beginning of the end. But it is, perhaps, the end of my PhD.

Appendix I

T4 DNA Ligase's dissociation from DNA

Appendix 1 – T4 DNA Ligase's dissociation from DNA

Unpublished data from our lab (collected by Dr Nicola Don) showed that to prevent T4 DNA Ligase from ligating DNA ends together, the salt concentration of the solution needs to be raised above 800 mM (**figure A.1**). In experiments where T4 Ligase was used to ligate lambda molecules or oligonucleotides together before creating DNA tightropes, the solution was made to above 800 mM salt to dissociate the ligase before creating DNA tightropes.

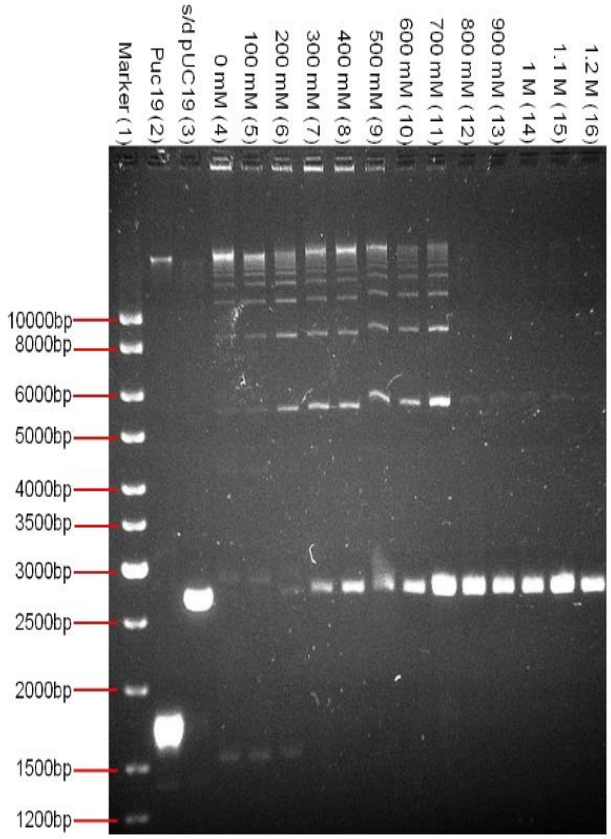


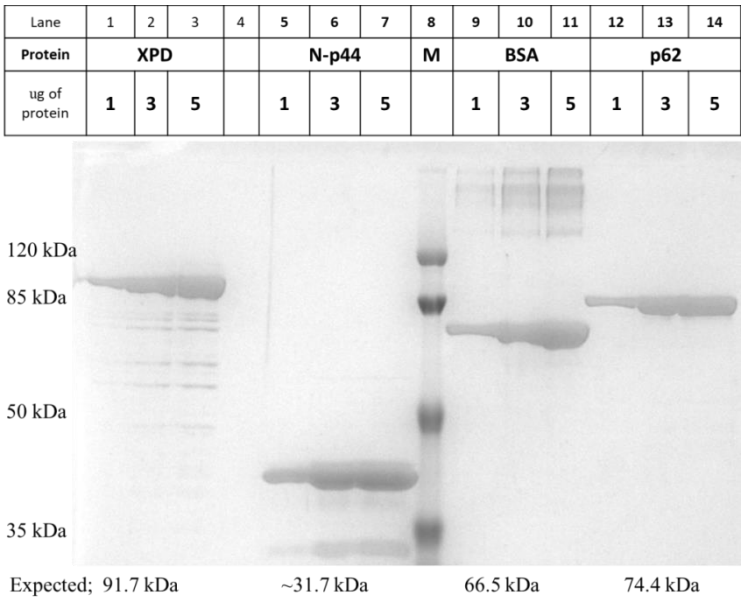
Figure A.1 T4 DNA Ligase dissociation from DNA.
Agarose gel showing the concatemerisation of linearised pUC19 plasmid DNA by T4 DNA Ligase, and the inhibition of this activity at salt concentrations above 800 mM (lane 12). 350 ng of DNA per lane after 30 minutes of ligation in various salt concentrations in the presence of 15% PEG_{6000}. Gel and experiment performed by Dr. Nicola Don.

Appendix II

Purity of TFIIH proteins used in this thesis

Appendix II – Purity of TFIIH proteins used in this thesis

The purity of all TFIIH proteins used in this thesis are shown below in SDS-PAGE gels. BSA is included on each gel as a concentration standard.



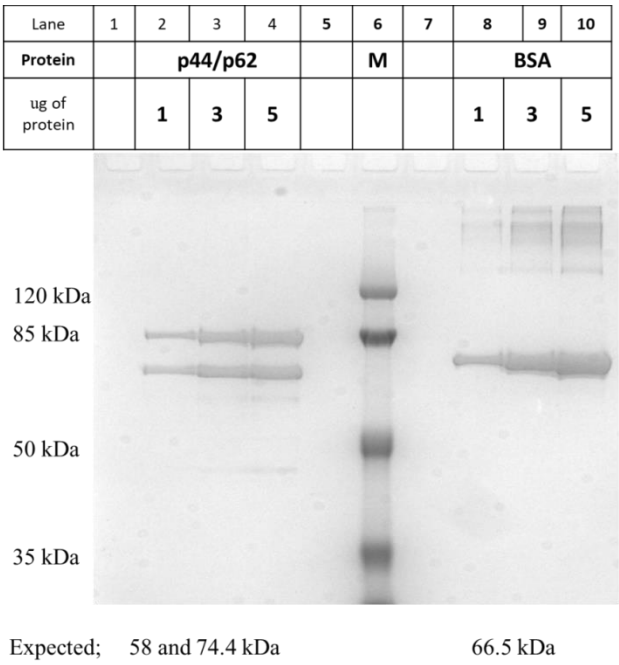


Figure A.2 SDS-PAGE gels of purified TFIIH proteins.
Increasing masses of each TFIIH protein used in this thesis were run on a 10% SDS-PAGE gel and stained with coomassie. The molecular weight standards are indicated along with the expected mass of each product, and a standard of BSA for comparing relative concentrations.

Appendix III

Publication associated with chapter 3

THE FASEB JOURNAL • RESEARCH • www.fasebj.org

Understanding the coupling between DNA damage detection and UvrA's ATPase using bulk and single molecule kinetics

Jamie T. Barnett and Neil M. Kad[1]
School of Biological Sciences, University of Kent, Canterbury, United Kingdom

ABSTRACT: Nucleotide excision repair (NER) protects cells against diverse types of DNA damage, principally UV irradiation. In *Escherichia coli*, damage is recognized by 2 key enzymes: UvrA and UvrB. Despite extensive investigation, the role of UvrA's 2 ATPase domains in NER remains elusive. Combining single-molecule fluorescence microscopy and classic biochemical methods, we have investigated the role of nucleotide binding in UvrA's kinetic cycle. Measurement of UvrA's steady-state ATPase activity shows it is stimulated upon binding DNA (k_{cat} 0.71–1.07/s). Despite UvrA's ability to discriminate damage, we find UV-damaged DNA does not alter the steady-state ATPase. To understand how damage affects UvrA, we studied its binding to DNA under various nucleotide conditions at the single molecule level. We have found that both UV damage and nucleotide cofactors affect the attached lifetime of UvrA. In the presence of ATP and UV damage, the lifetime is significantly greater compared with undamaged DNA. To reconcile these observations, we suggest that UvrA uses negative cooperativity between its ATPase sites that is gated by damage recognition. Only in the presence of damage is the second site activated, most likely in a sequential manner.—Barnett, J. T., Kad, N. M. Understanding the coupling between DNA damage detection and UvrA's ATPase using bulk and single molecule kinetics. FASEB J. 33, 763–769 (2019). www.fasebj.org

KEY WORDS: cooperativity · fluorescence imaging · DNA repair · DNA tightropes

Genomic DNA is constantly damaged by both exogenous and endogenous sources and must be repaired efficiently to maintain genome integrity. Nucleotide excision repair (NER) is an evolutionarily conserved DNA repair mechanism across all kingdoms of life. NER primarily repairs bulky lesions, including UV-induced damage, such as cyclobutane pyrimidine dimers and 6–4 photoproducts. However, NER also acts promiscuously to recognize and repair myriad lesion types (1). The classic description of NER involves the processing of the lesion by multiple enzymes and requires detection, incision, and repair resynthesis (2). Damage verification requires UvrA (a homodimer that is referred to as UvrA here for clarity) to locate damage with UvrB. Upon location of damage, UvrA is ejected from the UvrAB "preincision" complex, leaving UvrB alone on DNA. UvrC is then recruited to perform 2 DNA incisions on either side of the lesion on the same strand. Subsequently, UvrD and DNA polymerase I carry out downstream damaged oligonucleotide removal and repair resynthesis (3, 4). Finally, DNA ligase seals the remaining nick in the DNA backbone.

Successful lesion excision in *Escherichia coli* requires ATP (3, 5–7). Analysis of UvrA's sequence reveals 2 type A Walker motif sites, one at residues 31–45 (N-terminal site) and one at residues 640–654 (C-terminal site) (8). In the 3-dimensional structure of dimeric UvrA, the N-terminal ATPase domain of 1 UvrA is close to the C-terminal ATPase site of the other. Furthermore, all 4 are positioned beneath the DNA binding cleft that runs across the face of the protein (9, 10). The role of UvrA's ATPase activity is uncertain but has been linked to dimer formation and lesion searching (5, 11, 12). Previous studies that mutated catalytic lysine residues in each ATPase site observed a drastic loss of overall ATPase activity (13). Such mutational studies have indicated a role for ATPase in loading UvrB onto DNA (13, 14). Because many of these studies conflict in their outcomes, a clearly defined ATPase mechanism and definition of its role in UvrA's function is needed.

UvrA's ATPase sites are thought to have distinct roles (13, 15, 16); the C-terminal site discriminates DNA damage, and the N-terminal site is implicated in UvrA dimer formation and UvrB binding (17, 18). Using the *in vitro*

ABBREVIATIONS: ATPγS, adenosine 5′-O-(3-thio)triphosphate; NER, nucleotide excision repair; $UvrA_{WT}$, UvrA wild-type

[1] Correspondence: School of Biological Sciences, University of Kent, CT2 7NH Canterbury, United Kingdom. E-mail: n.kad@kent.ac.uk

doi: 10.1096/fj.201800899R
This article includes supplemental data. Please visit *http://www.fasebj.org* to obtain this information.

DNA tightrope assay in conjunction with mScarlet-labeled *E. coli* UvrA (19), we have investigated, at the single-molecule level, how different nucleotide cofactors affect UvrA's interaction with DNA. UvrA was observed to perform a 3-dimensional search on DNA, with a lifetime altered by the nucleotide condition, and the presence of UV damage. The rate-limiting step for ATP turnover occurs on DNA, and the presence of damage alters UvrA's DNA bound lifetime but not its steady-state ATPase rate. This paradox is resolved if UvrA sequentially hydrolyzes ATP in its 2 sites, suggesting that UvrA has a negatively cooperative ATPase that is tightly coupled to damage recognition.

MATERIALS AND METHODS

Unless stated otherwise, all chemicals were purchased from Thermo Fisher Scientific (Waltham, MA, USA), and DNA oligonucleotides were from MilliporeSigma (Burlington, MA, USA). All *in vitro* experiments were carried out at room temperature in ABC buffer [50 mM Tris–HCl (pH 7.5), 50 mM KCl, 10 mM $MgCl_2$, and 10 mM DTT].

Protein expression and purification

The gene for *E. coli* UvrA, obtained from National Bioresource Project (NIG, Kyoto, Japan), was engineered onto a C-terminal mScarlet fluorescent protein separated *via* a flexible linker (20). Further C-terminal to mScarlet, we placed a His tag for purification and induced expression of this pET21a construct overnight at 18°C using isopropyl β-D-1-thiogalactopyranoside. Purification was performed using nickel affinity followed by heparin chromatography (GE Healthcare, Chicago, IL, USA). To avoid protein precipitation, we kept KCl concentrations >200 mM. Once purified, the UvrA concentration was determined using mScarlet's extinction coefficient ($M^{-1}cm^{-1}$) at 569 nm and stored in 50% glycerol 50 mM Tris (pH 7.5), 500 mM KCl, 0.1 mM EDTA, and 10 mM DTT at −20°C (18).

Complementation assays

To investigate whether our UvrA-mScarlet construct can rescue UvrA wild-type ($UvrA_{WT}$) knockout cells, we performed *in vivo* UV complementation assays. $UvrA_{WT}$ knockout (Keio) cells were obtained from the National Bioresource Project (NIG) and transformed with UvrA-mScarlet or empty mScarlet vector. Previously, we have shown $UvrA_{WT}$ knockout cells complemented with ectopically expressed $UvrA_{WT}$ survive UV irradiation at 5 J/m^2 (21). Transformed cells were grown in Luria-Bertani medium with 25 μg/ml chloramphenicol to OD_{600} of 0.5. Five microliters of undiluted and three 10-fold serial dilutions were plated on Luria-Bertani agar and subjected to either no UV or 5 J/m^2 UV (254 nm) irradiation and incubated overnight at 37°C in the dark. $UvrA_{WT}$ and UvrA-mScarlet fully restored UV survival compared with $UvrA_{WT}$ knockout and $UvrA_{WT}$ knockout cells ectopically expressing mScarlet alone.

NADH-linked ATPase assay

ABC buffer supplemented with 0.5 mM phosphoenol pyruvate solution was stored at −20°C; 1 mM DTT was added upon thawing. The phosphoenol pyruvate solution with DTT was blanked at 340 nm in a spectrophotometer, and then 10 μl of pyruvate kinase (600–1000 U/ml) and lactate dehydrogenase (900–1400 U/ml, premixed stock from MilliporeSigma) per 500 μl reaction were added to a cuvette with 210 μM NADH. The change in OD_{340} was fitted linearly to calculate loss of NADH (6220 $M^{-1}cm^{-1}$ at 340 nm), enabling calculation of k_{cat}. Reactions were repeated 3 times, and the error represents the SD.

DNA substrates

F26,50 (5′-GACTACGTACTGTTACGGCTCCATC[FlcdT]CTACCGCAATCAGGCCAGATCTGC-3′) containing a fluorescein adduct opposite a mismatched base was previously demonstrated as a target for NER (22). Double-stranded substrate was produced by mixing equimolar concentrations of the reverse, complementary nondamaged oligonucleotide in TE buffer [10 mM Tris-HCl, 1 mM EDTA (pH 8)] at 95°C and left to cool slowly to room temperature. Cyclobutane pyrimidine dimers pUC18 and λ DNA (New England Biolabs, Ipswich, MA, USA) were irradiated to 1000 J/m^2 with a calibrated 254 nm lamp (ENF-240C/FE; Spectronics, Westbury, NY, USA) immediately before being used in the ATPase assay.

DNA tightropes

DNA tightropes were constructed as previously described (21). Nucleotides [ATP, ADP, adenosine 5′-O-(3-thio)triphosphate (ATPγS)] were used at 1 mM final concentration. For experiments with P_i, 1 mM ADP with 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, and 10 mM DTT was supplemented with 19 mM sodium phosphate (freshly autoclaved to remove pyrophosphate), and ionic strength was balanced by the removal of 50 mM KCl (23). For experiments with damaged λ DNA, the DNA was exposed to 1000 J/m^2 of 254 nm light immediately before creating tightropes.

Single molecule imaging and analysis

Imaging was performed using a custom-built fluorescence microscope capable of oblique-angle fluorescence excitation and multichannel emission (21). mScarlet was excited using a 561 nm diode OBIS LS laser (Coherent, Santa Clara, CA, USA) at 5 mW. To determine if photobleaching affected our observed attached lifetimes, we measured durations of attachment for the longest associating nucleotide condition, UvrA-ATPγS, at different laser powers. The attached lifetime was reduced only above 10 mW, indicating that, at the laser powers used in our experiments, photobleaching is not defining the lifetime. Videos were collected with a 300 ms exposure time for 90 s using 2 × 2 binning and were transformed into kymographs using ImageJ.

Attached lifetimes in each condition were plotted as cumulative frequency histograms to remove any bin size dependence. Cumulative frequencies were fitted to exponentials, and *F* tests were used to determine the requirement for single *vs.* double exponential fits.

RESULTS

DNA accelerates the steady-state ATP hydrolysis rate of UvrA

Although it is known that *E. coli* nucleotide excision repair requires ATP (3, 6, 7), its role in UvrA's function has not been fully determined. Using an NADH-linked assay, the steady-state UvrA-mScarlet (UvrA hereafter) ATP turnover rate was measured with and without DNA. Absorbance at 340 nm was used to linearly quantify ADP release because

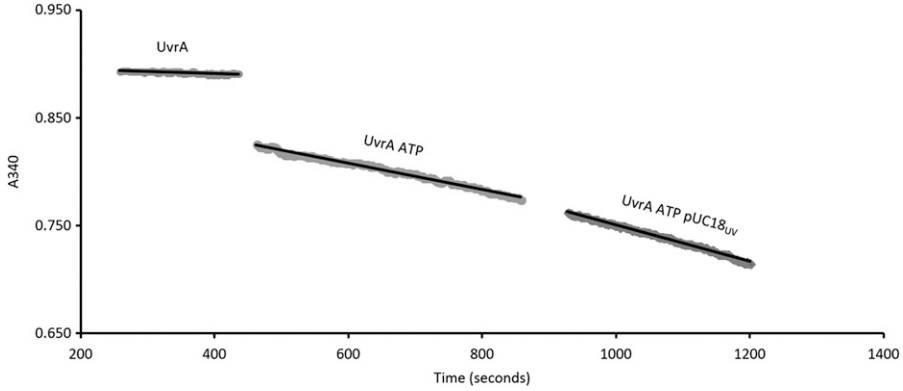


Figure 1. Real-time NADH-coupled ATPase assay. UvrA alone, UvrA with 1 mM ATP added and after the addition of 50 ng 1000 J/m^2 irradiated pUC18. Rates of NAD^+ production include: UvrA, 0.002 μM/s; UvrA + ATP, 0.020 μM/s; and UvrA + ATP + DNA, 0.027 μM/s. Reactions were carried out at room temperature in ABC buffer with 25 nM UvrA-mScarlet; no ADP accumulates in this coupled assay.

1 NADH is oxidized to NAD^+ for each ADP released. Linear fits gave the rates of ATP hydrolysis shown in **Fig. 1**.

This assay was repeated for various damaged and undamaged DNA substrates (**Fig. 2*A***). Undamaged DNA substrates stimulate the ATPase activity of UvrA significantly (k_{cat} UvrA alone, 0.71 ± 0.05 ATP/UvrA/second *vs.* UvrA + undamaged DNA, 1.07 ± 0.08 ATP/UvrA/second; $P < 0.0001$). The ATPase rate was also stimulated by a fluorescein containing oligonucleotide F26,50 (1.06 ± 0.142 ATP/UvrA/second), UV damaged pUC18 (1.13 ± 0.05 ATP/UvrA/second), or λ phage DNA (1.03 ± 0.12 ATP/UvrA/second) compared with ATP alone ($P < 0.0001$). However, damaged DNA does not affect the maximal ATPase activity relative to undamaged equivalents ($P = 0.9594$) (Fig. 2*A*, combined data). To investigate if damage discrimination affects UvrA's ATPase, we titrated ATP (Fig. 2*B*) in the presence of undamaged DNA or UV-irradiated DNA. The K_m for ATP decreased from 195 μM in the presence of undamaged DNA to 60 μM with damaged DNA. The ATP titration independently confirmed that the k_{cat} (per UvrA monomer) was unchanged, remaining at 1.4/s. Because ATP hydrolysis is accelerated by DNA binding, this indicates that the rate-limiting kinetic process occurs while bound to DNA. If the rate-limiting step did not occur on DNA, there would be no change in steady-state ATPase upon addition of DNA.

Determining the kinetic state affected by DNA binding

To identify which nucleotide state is affected by DNA binding, we measured the DNA-bound lifetime of UvrA using single-molecule fluorescence imaging. To do this, we used an *in vitro* DNA tightrope assay, which comprises individual DNA molecules suspended between silica beads in a microfluidic chamber (24, 25). Oblique angle illumination is used to limit the background from molecules in solution, leading to clear images of mScarlet-labeled protein binding to the DNA tightropes. The effects of damage were studied using DNA exposed to UV light (254 nm) immediately before the formation of tightropes. Videos of UvrA molecules binding to DNA were transformed into kymographs (**Fig. 3*A***) for dwell-time analysis and determination of the search mechanism. One-dimensional sliding on DNA will appear as movement on the *y* axis (position) over time, whereas a 3-dimensional search will appear as horizontal streaks (24) (Fig. 3*A*). The length of any continuous interaction corresponds to the dwell time. It was clear from the data that UvrA uses a 3-dimensional search mechanism because no positional movement was seen. The individual attached lifetimes were plotted as cumulative frequency histograms and fit to exponentials, which are appropriate for stochastic interactions (Fig. 3*B*). These fits across various nucleotide and DNA damage conditions provided dissociation rate constants (Fig. 3*C* and **Table 1**).

The bulk-phase steady-state ATPase k_{cat} was determined to be 1.07 ATP/UvrA per second in the presence of damaged or undamaged DNA (Fig. 2). This is in excellent agreement with the single-molecule detachment rate constant of 1.24 ± 0.038/s for UvrA on undamaged DNA with ATP. These data fit better to a double exponential (F test = 204) than a single, suggesting 2 populations of attachments, the vast majority of which (amp1 = 94%) reflect molecules binding to undamaged DNA. The remainder most likely represents molecules that bind to nonspecific damage on λ DNA, as previously shown (24). In the presence of UV-damaged DNA and ATP, UvrA shows a dramatic decrease in the detachment rate constant (0.411 ± 0.023/s; *i.e.*, a 3-fold increase in attached lifetime). The UV damage cumulative frequency histogram also fits better to a double exponential (F test = 154; amp1 96%, $k1$ 0.411 ± 0.023/s; amp2 4%, $k2$ 0.098 ± 0.030/s; Fig. 3*B*), consistent with 2 populations of attached proteins. However, the small contribution (amp2 = 4%) of $k2$ means very few complexes are long lived on DNA. We did not investigate the origin of this species further.

To determine the contribution of the ATP binding to the attached lifetime of UvrA, we used ATPγS, a non-hydrolyzable analog of ATP. ATPγS has been shown to permit formation of UvrA dimers and stimulate binding of UvrA to DNA (5). Figure 3*C* shows that ATPγS slows the release of UvrA from DNA (0.155 ± 0.002/s), equivalent to a 6.5-fold increase in attached lifetime. DNA damage does not affect the dissociation kinetics of UvrA in the presence of ATPγS (0.127 ± 0.001/s), suggesting that ATP-bound UvrA is not involved in damage recognition (5, 12). The slow rate constant for ATP-bound detachment implies that hydrolysis and formation of subsequent nucleotide states occur on DNA; consequently, one of these states will be sensitive to damage.

A possible candidate for controlling the release of UvrA from DNA is the ADP-bound state. ADP facilitates UvrA

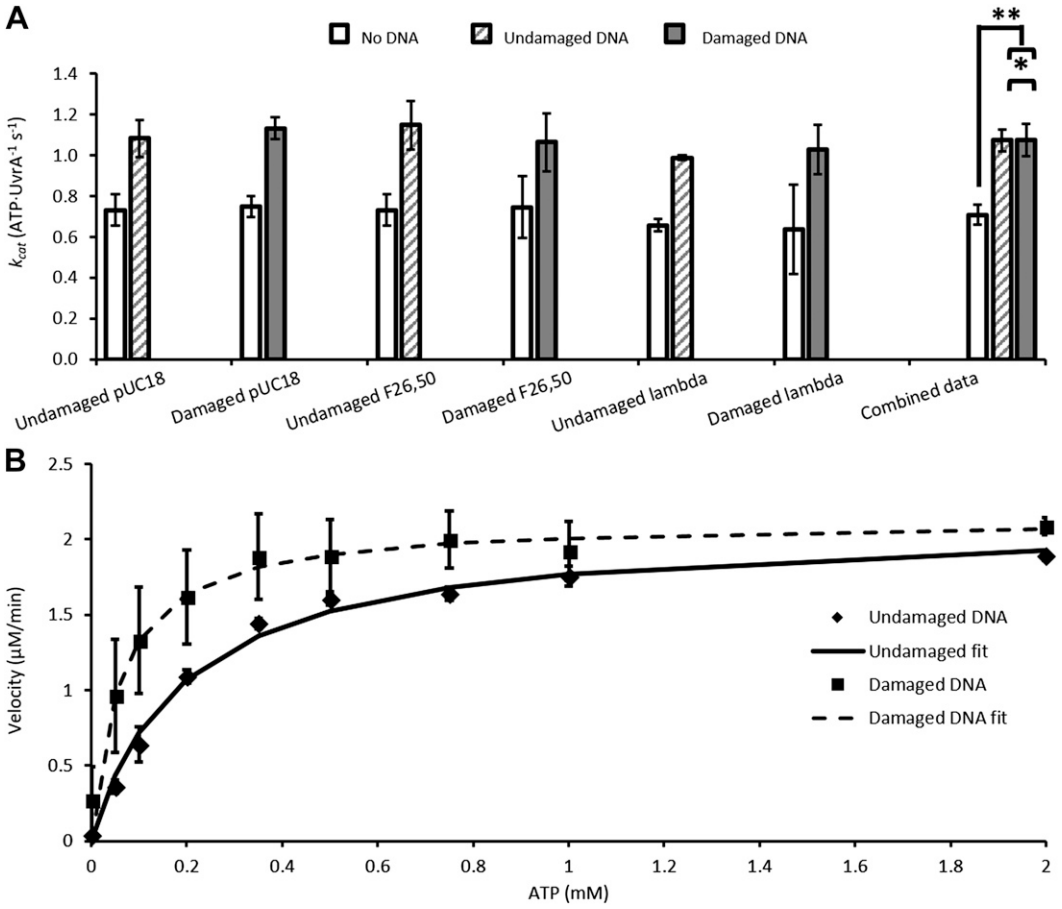


Figure 2. Steady-state ATPase data for UvrA. *A*) Summary of the k_{cat} per UvrA monomer for no DNA (white), undamaged DNA (striped gray), and 1000 J/m^2 UV-damaged DNA (solid gray). All data were collected using the linked assay (Fig. 1). Error bars are SD for 3 repeat determinations. Combined data were obtained from the mean of all non-DNA, undamaged DNA, and damaged DNA ATPase rates. Combined data show that damaged DNA (1.074 ATP/UvrA per second) does not differ from undamaged DNA (1.073 ATP/UvrA per second). *P = 0.9594, **$P < 0.0001$ (Student's t test). *B*) ATP titration of UvrA ATPase with undamaged or UV damaged DNA. In the presence of damage (squares), the K_m is decreased 3-fold compared with undamaged DNA (diamonds). The k_{cat} is unchanged between conditions. Data are averages of 3 repeats, and error is SD.

dimerization and binding to DNA (4, 11, 12). Surprisingly, UvrA in the presence of ADP did not substantially bind to DNA tightropes compared with identical concentrations of UvrA with ATP or ATPγS. Nonetheless, from the attachments observed it was possible to determine the lifetime as 4.9 s (0.20 ± 0.006/s), which was ~5-fold longer than in the presence of ATP. The low level of attachment suggests that few UvrA dimers exist in the presence of ADP (5), conflicting with previous observations (12). With UV-damaged DNA, the attached lifetime remains unchanged at 4.6 s (equivalent to 0.22 ± 0.016/s). It is likely that UvrA does not dimerize with only ADP bound and that the few dimers that form are strongly bound to the DNA. Given the low abundance of this state, it is likely that detachment occurs from the ADP state. Furthermore, the similar rates of release in the presence and absence of DNA damage suggest that UvrA-ADP is not a damage sensing state.

The step between ATP binding and formation of the UvrA ADP-bound state is therefore the most likely to detect damage. To investigate this, we attempted to create the ADP.P_i state of UvrA by adding a large excess of free phosphate to UvrA and ADP. This required careful consideration of ionic strength changes that were balanced by a reduction in the buffer salt (26). On undamaged DNA, the addition of phosphate to the ADP state increases the detachment rate constant slightly from ADP alone (0.202 ± 0.006 to 0.234 ± 0.007/s) but does not resemble the ATP condition (1.07 ± 0.019/s), nor does phosphate restore UvrA's affinity for DNA; the number of binding events was also low, similar to the ADP-bound state. It is clear that P_i addition does not reverse hydrolysis and restore the rate-limiting step for detachment. However, in the presence of UV-damaged DNA, UvrA's detachment rate constant significantly increases (from 0.217 ± 0.016 to 0.364 ± 0.011/s; $P < 0.0001$), resembling the rate constant for release in the presence of UV damage and ATP (0.411 ± 0.023/s; P = 0.501). This change in lifetime suggests that UvrA with P_i and ADP is populating the damage-sensitive state, which is rate limiting for UvrA detachment from DNA.

DISCUSSION

The role of ATP in the mechanism of NER is not clearly understood. Here, we compare the kinetics of ATP turnover by UvrA at the bulk and single-molecule levels. We

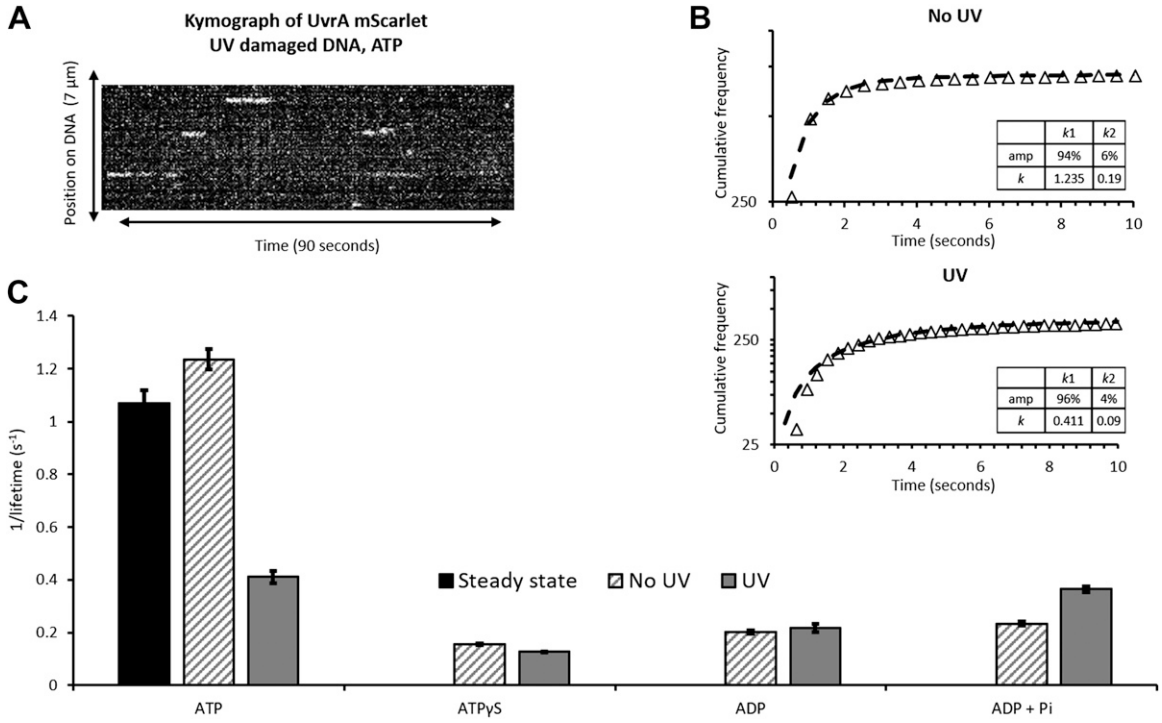


Figure 3. The single-molecule binding kinetics of UvrA. *A*) A kymograph transformation from a video showing UvrA-mScarlet binding and releasing from DNA in a 3-dimensional search. *B*) Lengths of the linear streaks in *A* were compiled into cumulative frequency histograms. The data for UvrA with ATP and DNA, either undamaged (top) or UV irradiated (bottom), are shown as logarithmic cumulative frequency against time; the nonlinearity indicates 2 processes. These were fit to double exponentials with rate constants given in the inset. *C*) A comparison of the rate constants obtained for each nucleotide and DNA condition (Table 1). The steady-state k_{cat} (Fig. 2) is shown as a solid black bar. Error bars indicate the SEM fit error.

have found that the bulk-phase ATPase k_{cat} of UvrA was stimulated by the presence of DNA; however, there was no further change with damaged DNA. This strongly contrasted with the attached lifetimes of UvrA on DNA. The presence of damage increased the attached lifetime 2- to 3-fold in the presence of ATP. After determining the attached lifetimes of UvrA molecules on either damaged or undamaged DNA under various nucleotide conditions, we infer that UvrA most likely functions as a negatively cooperative ATPase, using its 2 ATP sites in sequence.

TABLE 1. *Summary of rate constants for various nucleotide conditions on undamaged or UV-damaged DNA*

Nucleotide	Detachment rate constant per second	Fit error	*N*
No damage			
ATP	1.235[a]	0.038	706
ATPγS	0.155	0.002	166
ADP	0.202	0.006	33
ADP Pi	0.234	0.007	53
UV damage			
ATP	0.411[b]	0.023	394
ATPγS	0.127	0.001	152
ADP	0.217	0.016	21
ADP Pi	0.364	0.011	43

N refers to molecules analyzed from at least 3 flow cells per condition. [a]Double exponential, amp1 94%, *k*1 1.235 ± 0.038/s; amp2 6%, *k*2 0.190 ± 0.044/s. [b]Double exponential, amp1 96%, *k*1 0.411 ± 0.023/s; amp2 4%, *k*2 0.098 ± 0.030/s.

UvrA's second ATPase site is activated by DNA damage

The steady-state ATPase rate for UvrA increased ~50% upon addition of DNA, demonstrating that UvrA is a DNA-stimulated ATPase. However, no further change in the DNA-stimulated ATPase was observed with damaged DNA. Similar changes in bulk-phase ATPase using the linked assay have been reported (17). We used damage substrates at various damage densities and different types of substrate (see Supplemental Information); however, in all conditions the type or density of damage had no effect on the observed results. We also ensured the DNA damage concentration was higher than the UvrA concentration used. Furthermore, linear and plasmid DNA substrates stimulated the ATPase of UvrA equally, indicating that the DNA ends play no part in the observed activation. This contrasts with previous studies, where undamaged DNA was seen to inhibit the ATPase and damage returned the activity to that seen without DNA (16). This difference may be due to ADP's strong affinity for UvrA (14), which may inhibit the ATPase; our use of the linked assay ensures nucleotide is recycled, abrogating any possible issues arising from such product inhibition.

By directly imaging single UvrA molecules binding to DNA, we were able to correlate attached lifetimes with the biochemically measured ATPase rates. With ATPγS, ample binding to DNA was observed; however, with ADP this was virtually abolished, suggesting that UvrA binds DNA

in the ATP-bound state and releases from DNA in the ADP-bound state. In both cases the attached lifetimes are too long to correlate with the steady-state ATPase rate and were not affected by DNA damage; therefore, these states are unlikely to be substantially populated, and their role in the damage-modulated ATPase of UvrA is irrelevant. The most interesting observation derives from a comparison of the bulk k_{cat} of 1/s *vs.* the lifetime in the presence of ATP (~1/s); these values are identical. The k_{cat} was calculated per UvrA monomer; therefore, we infer that on DNA where UvrA is a dimer (4, 12, 24, 27) 2 ATP molecules are hydrolyzed and that attachment to DNA is tightly coupled to ATP turnover. However, in the presence of damage, although k_{cat} is unchanged, the attached lifetime increases substantially. The ~3-fold increased affinity for ATP on damaged DNA corresponds well to the 3-fold increase in lifetime on damaged DNA, demonstrating that, in the presence of damage, there remains tight coupling between ATP turnover and UvrA release from DNA. Therefore, to reconcile the unchanged k_{cat} with the increased lifetime, we suggest that when damage is detected during turnover of the first ATP, rather than detaching from DNA, the second site is activated. This second site has a similar ATP hydrolysis rate, thus keeping the total number of ATPs hydrolyzed per second the same but increasing the attached lifetime. This sequential action is indicative of negative cooperativity between the sites that is tightly coupled to DNA damage recognition. Similar mechanisms of negative cooperativity have been noted for other ABC ATPase superfamily members, including the bacterial mismatch repair protein MutS (28) and its eukaryotic homolog (29, 30).

UvrA's rate-limiting step is affected by the presence of DNA damage

Because the rate-limiting step is not ADP release, we sought to determine whether hydrolysis and subsequent release of P_i controls the turnover rate. To push the pathway backward, P_i was added to ADP and UvrA. If ATP hydrolysis and release of ADP + P_i is the rate-limiting step, the addition of P_i to ADP should restore the lifetime to the ATP condition. With undamaged DNA, the detachment rate constant did not resemble that of UvrA in the presence of ATP (0.23 *vs.* 1.2/s). Therefore, the addition of phosphate does not reverse hydrolysis of the first site. However, with damaged DNA the detachment rate constant was very similar to that of UvrA with ATP and UV damage (0.36 *vs.* 0.41/s). We surmised that the 0.41/s detachment rate constant was derived from sequential ATP hydrolysis at both sites. This suggests that, in the presence of damage, both nucleotide binding sites can reload P_i and that the rate-limiting steps are recapitulated, possibly through reversal of hydrolysis. The identity of the 2 sites was revealed using ATPase mutants (14) and *in vivo* (13). Based on these previous observations that nucleotide needs to be bound to the C-terminal (or distal) ATPase site for damage recognition, we suggest this is the first ATPase site described here. Therefore, the second ATPase site (N-terminal or proximal) is activated after damage recognition at the first site. This confirms the 2-site hypothesis, which suggests the N-terminal site is necessary for recruitment of UvrB (13, 14, 17).

Communication between the ATPase sites is likely facilitated by their proximity. The N-terminal site of 1 ATPase site is adjacent to the C-terminal site. Interestingly, these sites are connected by a tunnel (**Fig. 4**); it is plausible that P_i release occurs through this tunnel and gates the recruitment of nucleotide or hydrolysis at the second site. Future structural investigations of the transition state are imperative to reveal how P_i is released, whether through the tunnel or by monomerization.

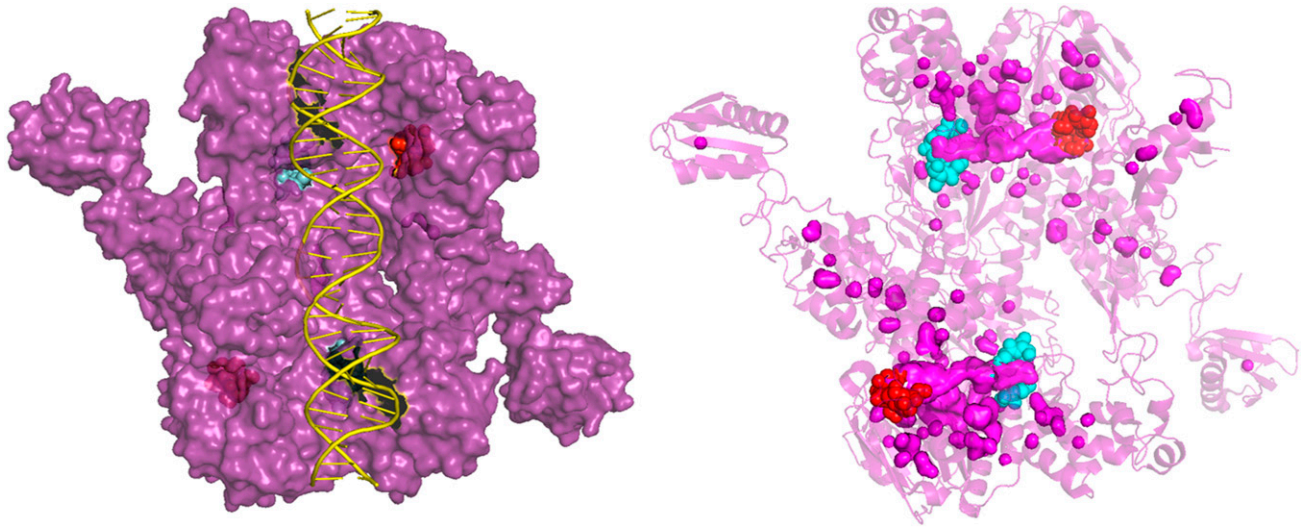


Figure 4. Crystal structure of *Thermotoga maritima* UvrA showing the phosphate tunnel. Left: Surface plot of dimeric UvrA (purple) with the N-terminal (cyan) and C-terminal (red) ATPase sites shown behind DNA (yellow) forming a tunnel. Right: PyMol cavity search of structure revealing a clear tunnel (pink) with the ATPase sites (cyan and red) on either side. UvrA is shown as a transparent cartoon with all pink surface objects showing solvent cavities. Image created in PyMol using PDB structure (3PIH) (9).

CONCLUSIONS

In summary, we have combined data from single-molecule and bulk-phase studies of the UvrA ATPase and DNA interaction. We find excellent correlation between the kinetics using both methods and that UvrA's interaction with DNA is tightly coupled to its ATPase rate-limiting step. Two ATPs are hydrolyzed per dimer per second when bound to undamaged DNA. However, in the presence of damage, although the steady-state ATPase is unaltered, the attached lifetime increases. We suggest this occurs through a sequential mechanism whereby if damage is located, the second ATPase site is activated, resulting in another 2 ATPs per dimer per second being hydrolyzed. This accounts for the observed kinetics and suggests a tight coupling model due to strong negative cooperativity between the UvrA ATPase sites. Therefore, UvrA couples its ATPase activity to its damage-sensing function by activating the second ATPase site. FJ

ACKNOWLEDGMENTS

The authors thank the Kad group for discussion; Bennett Van Houten (University of Pittsburgh, Pittsburgh, PA, USA), Chris Mulligan, Alex Moores, and Michael Geeves (all of the University of Kent) for reading the manuscript; the National Institute of Genetics, Japan for access to their excellent collection of materials; and Alex Moores for constructing the empty mScarlet vector used in this study. This work was supported by Biotechnology and Biological Sciences Research Council Grants BB/P00847X/1, BB/M019144/1, and BB/I003460/1 (to N.M.K.) and BB/M01603X/1 (to J.T.B.). The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

J. T. Barnett collected data; and J. T. Barnett and N. M. Kad designed the experiments, analyzed the data, and wrote the manuscript.

REFERENCES

1. Truglio, J. J., Croteau, D. L., Van Houten, B., and Kisker, C. (2006) Prokaryotic nucleotide excision repair: the UvrABC system. *Chem. Rev.* **106**, 233–252
2. Van Houten, B., Croteau, D. L., DellaVecchia, M. J., Wang, H., and Kisker, C. (2005) 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. *Mutat. Res.* **577**, 92–117
3. Sancar, A., and Rupp, W. D. (1983) A novel repair enzyme: UVRABC excision nuclease of Escherichia coli cuts a DNA strand on both sides of the damaged region. *Cell* **33**, 249–260
4. Orren, D. K., and Sancar, A. (1989) The (A)BC excinuclease of Escherichia coli has only the UvrB and UvrC subunits in the incision complex. *Proc. Natl. Acad. Sci. USA* **86**, 5237–5241
5. Seeberg, E., and Steinum, A. L. (1982) Purification and properties of the uvrA protein from Escherichia coli. *Proc. Natl. Acad. Sci. USA* **79**, 988–992
6. Yeung, A. T., Mattes, W. B., Oh, E. Y., and Grossman, L. (1983) Enzymatic properties of purified Escherichia coli uvrABC proteins. *Proc. Natl. Acad. Sci. USA* **80**, 6157–6161
7. Thiagalingam, S., and Grossman, L. (1991) Both ATPase sites of Escherichia coli UvrA have functional roles in nucleotide excision repair. *J. Biol. Chem.* **266**, 11395–11403
8. Husain, I., Van Houten, B., Thomas, D. C., and Sancar, A. (1986) Sequences of Escherichia coli uvrA gene and protein reveal two potential ATP binding sites. *J. Biol. Chem.* **261**, 4895–4901
9. Jaciuk, M., Nowak, E., Skowronek, K., Tańska, A., and Nowotny, M. (2011) Structure of UvrA nucleotide excision repair protein in complex with modified DNA. *Nat. Struct. Mol. Biol.* **18**, 191–197
10. Pakotiprapha, D., Inuzuka, Y., Bowman, B. R., Moolenaar, G. F., Goosen, N., Jeruzalmi, D., and Verdine, G. L. (2008) Crystal structure of Bacillus stearothermophilus UvrA provides insight into ATP-modulated dimerization, UvrB interaction, and DNA binding. *Mol. Cell* **29**, 122–133
11. Oh, E. Y., Claassen, L., Thiagalingam, S., Mazur, S., and Grossman, L. (1989) ATPase activity of the UvrA and UvrAB protein complexes of the Escherichia coli UvrABC endonuclease. *Nucleic Acids Res.* **17**, 4145–4159
12. Mazur, S. J., and Grossman, L. (1991) Dimerization of Escherichia coli UvrA and its binding to undamaged and ultraviolet light damaged DNA. *Biochemistry* **30**, 4432–4443
13. Stracy, M., Jaciuk, M., Uphoff, S., Kapanidis, A. N., Nowotny, M., Sherratt, D. J., and Zawadzki, P. (2016) Single-molecule imaging of UvrA and UvrB recruitment to DNA lesions in living Escherichia coli. *Nat. Commun.* **7**, 12568
14. Myles, G. M., Hearst, J. E., and Sancar, A. (1991) Site-specific mutagenesis of conserved residues within Walker A and B sequences of Escherichia coli UvrA protein. *Biochemistry* **30**, 3824–3834
15. Myles, G. M., and Sancar, A. (1991) Isolation and characterization of functional domains of UvrA. *Biochemistry* **30**, 3834–3840
16. Wagner, K., Moolenaar, G. F., and Goosen, N. (2010) Role of the two ATPase domains of Escherichia coli UvrA in binding non-bulky DNA lesions and interaction with UvrB. *DNA Repair (Amst.)* **9**, 1176–1186
17. Kraithong, T., Channgam, K., Itsathitphaisarn, O., Tiensuwan, M., Jeruzalmi, D., and Pakotiprapha, D. (2017) Movement of the β-hairpin in the third zinc-binding module of UvrA is required for DNA damage recognition. *DNA Repair (Amst.)* **51**, 60–69
18. Croteau, D. L., DellaVecchia, M. J., Wang, H., Bienstock, R. J., Melton, M. A., and Van Houten, B. (2006) The C-terminal zinc finger of UvrA does not bind DNA directly but regulates damage-specific DNA binding. *J. Biol. Chem.* **281**, 26370–26381
19. Bindels, D. S., Haarbosch, L., van Weeren, L., Postma, M., Wiese, K. E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M. A., and Gadella, T. W., Jr. (2017) mScarlet: a bright monomeric red fluorescent protein for cellular imaging. *Nat. Methods* **14**, 53–56
20. Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. *DNA Res.* **12**, 291–299
21. Springgall, L., Hughes, C. D., Simons, M., Azinas, S., Van Houten, B., and Kad, N. M. (2018) Recruitment of UvrBC complexes to UV-induced damage in the absence of UvrA increases cell survival. *Nucleic Acids Res.* **46**, 1256–1265
22. DellaVecchia, M. J., Croteau, D. L., Skorvaga, M., Dezhurov, S. V., Lavrik, O. I., and Van Houten, B. (2004) Analyzing the handoff of DNA from UvrA to UvrB utilizing DNA-protein photoaffinity labeling. *J. Biol. Chem.* **279**, 45245–45256
23. Bers, D. M., Patton, C. W., and Nuccitelli, R. (2010) A practical guide to the preparation of Ca(2+) buffers. *Methods Cell Biol.* **99**, 1–26
24. 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. *Mol. Cell* **37**, 702–713
25. Springgall, L., Inchingolo, A. V., and Kad, N. M. (2016) DNA-protein interactions studied directly using single molecule fluorescence imaging of quantum dot tagged proteins moving on DNA tightropes. *Methods Mol. Biol.* **1431**, 141–150
26. Kad, N. M., Trybus, K. M., and Warshaw, D. M. (2008) Load and Pi control flux through the branched kinetic cycle of myosin V. *J. Biol. Chem.* **283**, 17477–17484
27. Malta, E., Moolenaar, G. F., and Goosen, N. (2007) Dynamics of the UvrABC nucleotide excision repair proteins analyzed by fluorescence resonance energy transfer. *Biochemistry* **46**, 9080–9088
28. Lamers, M. H., Winterwerp, H. H., and Sixma, T. K. (2003) The alternating ATPase domains of MutS control DNA mismatch repair. *EMBO J.* **22**, 746–756
29. Iaccarino, I., Marra, G., Palombo, F., and Jiricny, J. (1998) hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutSalpha. *EMBO J.* **17**, 2677–2686
30. Studamire, B., Quach, T., and Alani, E. (1998) Saccharomyces cerevisiae Msh2p and Msh6p ATPase activities are both required during mismatch repair. *Mol. Cell. Biol.* **18**, 7590–7601

Received for publication May 8, 2018.
Accepted for publication July 2, 2018.

Appendix IV

Publication associated with chapter 4

The TFIIH components p44/p62 act as a damage sensor during nucleotide excision repair

Barnett JT[1], Kuper J[2], Koelmel W[2], Kisker C[2] and Kad NM[1,*]

[1] School of Biological Sciences, University of Kent, Canterbury, United Kingdom

[2] Rudolf Virchow Center for Experimental Biomedicine, Institute for Structural Biology, University of Würzburg, 97080 Würzburg, Germany.

[*] To whom correspondence may be addressed. Email: n.kad@kent.ac.uk Tel: +44 1227 816151

Keywords; DNA repair, ATPase, fluorescence imaging, single molecule, protein complexes

Classification;
Biological sciences
Biochemistry

Abstract

Nucleotide excision repair (NER) protects the genome following exposure to diverse types of DNA damage, including UV light and chemotherapeutics. Mutations in mammalian NER genes lead to diseases such as xeroderma pigmentosum, trichothiodystrophy, and Cockayne syndrome. In eukaryotes, the major transcription factor TFIIH is the central hub of NER. The core components of TFIIH include the helicases XPB, XPD, and the five core 'structural' subunits. Two of these core-TFIIH proteins, p44 and p62 remain relatively unstudied; although p44 is known to regulate the helicase activity of XPD during NER. p62's role is thought to be structural; however, a recent cryo-EM structure shows p44, p62, and XPD making contacts with each other, implying a more extensive role in DNA repair beyond the structural integrity of TFIIH. Here, we show that p44 stimulates XPD's ATPase, but upon encountering DNA damage further stimulation is only observed when p62 is in the ternary complex; supporting a role for p44/p62 in TFIIH's mechanism of damage detection. More significantly, we show that the p44/p62 complex binds DNA independently of XPD and diffuses along its backbone, indicating a novel DNA-binding entity in TFIIH. This revises our understanding of TFIIH and prompts more extensive investigation of all of the core subunits, for an active role during both DNA repair and transcription.

Significance

DNA repair is crucial for the survival of all organisms, however, given limited cellular resources correct identification of damage is essential. In Nucleotide Excision Repair this is a two-step process; with initial detection by patrolling proteins preceding hand off to the multi-protein complex TFIIH. How this complex verifies damage before initiating repair remains uncertain. Here, using single molecule and bulk-phase techniques we show the two 'structural' subunits p44 & p62 together enhance XPD's ability to detect DNA damage by directly interacting with DNA. These observations redefine how structural components contribute to DNA repair, opening questions about their role in transcription and beyond. Our findings suggest a more nuanced understanding of protein complexes beyond and including DNA repair is warranted.

Introduction

Cell survival relies on accurate replication and transcription of genetic material. When DNA is damaged, rapid repair is essential to ensure genome integrity. Defects in the nucleotide excision repair (NER) machinery lead to diseases such as x*eroderma pigmentosum* (XP), which is phenotypically characterized by UV hyper-sensitivity and increased incidence of skin cancers (1). Mutations found in many components of the NER pathway, including TFIIH can lead to XP. Core TFIIH comprises two helicases, one is enzymatically involved in both transcription and repair (XPB), whereas the catalytic activity of XPD is necessary only for repair (2, 3). In addition, TFIIH possesses five other core subunits: p44, p62, p34, p8, and p52 (4, 5 1987b, 6-14). p44 and p52 stimulate XPD and XPB's helicase activities respectively (9, 15, 16), likely on a structural level. However, as yet no other activity is assigned to any of the core TFIIH subunits. The recent cryo-EM structure of TFIIH (17) reveals numerous interactions between p44, p62, and XPD, possibly suggesting a direct involvement in DNA repair. In this study, we show that p44 stimulates the ATPase activity of XPD, however the co-purified complex of p62 and p44 further enhances XPD's ATPase on damaged DNA. These data suggest that p44/p62 may play an important role in damage discrimination within TFIIH. Moreover, using single molecule fluorescence microscopy we directly confirm the p44/p62 complex binds to and diffuses along double-stranded DNA (dsDNA); identifying it as a novel DNA binding unit in TFIIH. These first glimpses of p44/p62 activity suggest it is directly relevant to the mechanism of damage detection, playing more than just a structural role.

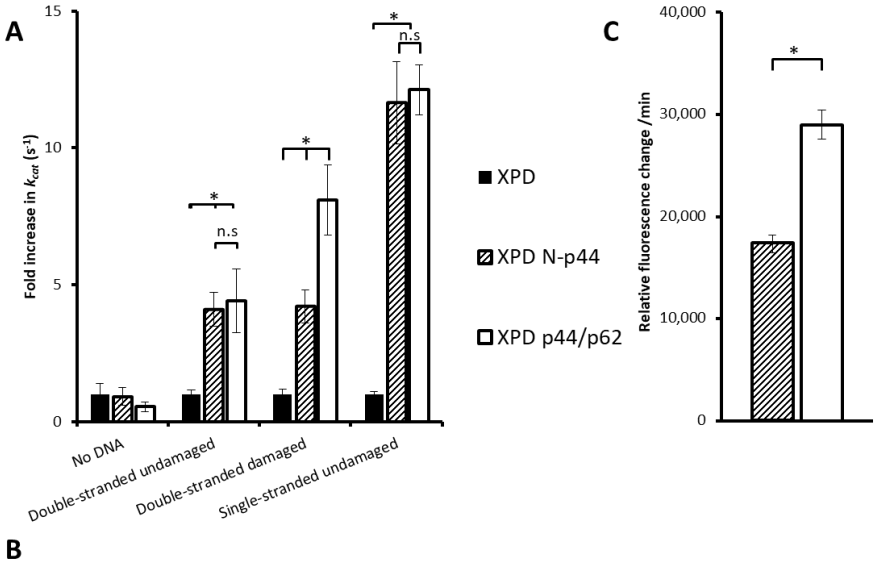
Results & discussion

XPD's ATPase is stimulated by p44/p62

p44 contacts both p62 and XPD in TFIIH (9, 17, 18), and mutations in XPD that disrupt p44 or p62 binding cause defects in NER and result in disease (3, 9, 16, 18). To investigate if p44/p62 was able to stimulate the ATPase of XPD, the turnover of ATP in the presence of different DNA substrates was measured using an NADH-coupled assay.

In the absence of p44 and p62, XPD's ATPase activity is slow even in the presence of single-stranded DNA (0.043 s^{-1}). However, with an N-terminal p44 fragment (residues 1-285 (N-p44)) containing the von Willebrand domain, XPD's ATPase was significantly stimulated in the presence of both double- and single-stranded DNA (3) (~0.03 s^{-1} to 0.136 s^{-1} and 0.504 s^{-1} respectively, $p < 0.05$ (**Figure 1B**)). No further acceleration of the ATPase was observed with full length p44 co-expressed in complex with p62 (p44/p62). However, remarkably, when damage (a fluorescein moiety shown to proxy for damage (19)) was introduced into a dsDNA substrate, p44/p62 accelerated XPD's ATPase two-fold more than on undamaged DNA (**Figure 1A & B**). N-p44 alone could not accelerate XPD's ATPase in the presence of damage, indicating the ternary complex (p44/p62) is responsible for this further enhancement and thus may play an important role in lesion detection. These results may explain why truncations of the yeast p62 homologue (Tfb1) sensitize the organism to UV irradiation (20-22).

To further investigate the role of p44 and p62 in activating XPD we analyzed XPD's helicase activity on an open fork substrate. Again, p44/p62 is seen to play an active role by enhancing the helicase activity compared to N-p44 alone (**Figure 1C**). Although no damage is present in the open fork substrate, p44/p62 significantly enhances XPD's ability to successfully unwind the DNA substrate (two-fold more than XPD with N-p44), despite no change in ATPase activity (**Figure 1A**).



B

DNA substrate	XPD k_{cat} values (s^{-1}) alone, with N-p44, and p44/p62		
	XPD	XPD N-p44	XPD p44/p62
No DNA	0.046 ± 0.013	0.042 ± 0.008	0.025 ± 0.004
Double-stranded undamaged	0.033 ± 0.004	0.136 ± 0.013	0.146 ± 0.035
Double-stranded damaged	0.027 ± 0.004	0.114 ± 0.006	0.218 ± 0.019
Single-stranded undamaged	0.043 ± 0.003	0.504 ± 0.055	0.524 ± 0.016

Figure 1. Steady-state ATPase and helicase activity of XPD in the presence of various DNA substrates and core TFIIH proteins*. A) The activity of XPD's ATPase is stimulated by both N-p44 (dashed) and p44/p62 (white) on various DNA substrates. Values for k_{cat} are given as a fold change from XPD alone (black). Errors are shown as S.E.M from 3 repeats. B) Table showing k_{cat} values ± S.E.M for XPD's ATPase. C) XPD's helicase activity is stimulated by N-p44 (dashed) and p44/p62 (white) on an open fork substrate. XPD alone displays no helicase activity [8]. Errors are shown as S.E.M from 9 repeats. Statistical significance determined using a student's t-test where * = $p < 0.05$, n.s = not statistically significant.*

The p44/p62 complex directly binds DNA

The role of p44/p62 in the recognition of damage presents the intriguing possibility that this complex could interact with DNA independently from XPD. To investigate this, we used a single molecule DNA tightrope assay (23) (**Figure 2**). Conjugation of a fluorescent quantum dot (QDot) to the poly-histidine purification tag on the p44/p62 complex (24) was achieved using an anti-His IgG antibody. Substantial binding of p44/p62 to dsDNA was observed, and of these approximately 80% could diffuse (n =

599 total) providing the first direct evidence that these factors are able to bind DNA independently of XPD.

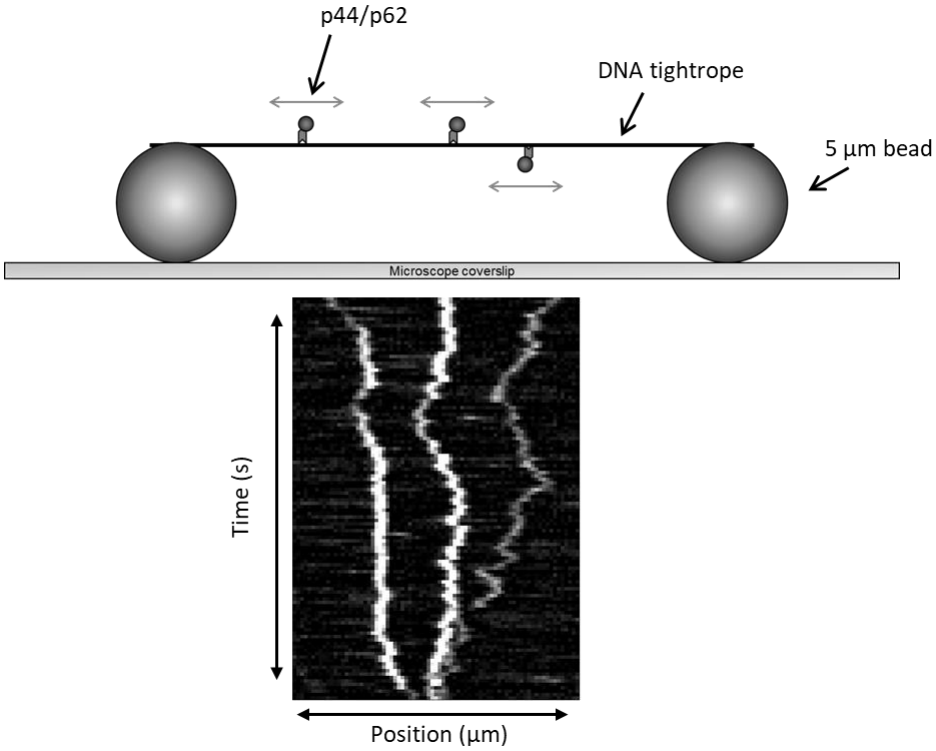


Figure 2. Schematic of a tightrope and kymograph analysis. *DNA tightropes are formed between beads adhered to a coverslip. QDot labelled proteins are then observed binding to the DNA. A video can be transformed into a kymograph by plotting position through time. Diffusing molecules appear as movement in the X axis for a duration of frames (Y axis). The kymograph shown in the lower panel indicates three diffusing p44/p62 molecules.*

The p44/p62 complex displayed multiple types of behavior on DNA. Firstly, we observed complexes randomly diffusing along the DNA, unable to pass one another (**Figure 2**). Secondly while diffusing, pausing was seen, often at the same location on the tightropes. This may indicate a visit to a damage site or a specific sequence. Finally, fluorescence intensity fluctuations of the same molecule over time suggest possible oligomerization. At elevated salt concentrations (100 mM vs 10 mM KCl) fewer molecules bound to DNA, and of these, a lower percentage diffused (55%, n = 58 total). We calculated the diffusion constant using mean-squared displacement analysis (23) and found no significant change ($p > 0.05$) between salt conditions (10 mM KCl 0.067 $\mu m^2/s \pm 0.006$ vs 100 mM KCl 0.042 $\mu m^2/s \pm 0.010$), which suggests

that p44/p62 molecules slide along the DNA helix (25). Based on the estimated size of a p44/p62 complex conjugated to a QDot, the diffusion constant appears limited by rotation-coupled diffusion around the backbone of the DNA helix (26). This is consistent with the inability for complexes to pass one another on the DNA.

In summary, we present the first mechanistic characterisation of the non-helicase TFIIH subunits p44/p62. Complexes formed by these two proteins were observed to bind and slide on dsDNA. Our bulk phase ATPase and helicase data indicate that p44/p62 is involved in damage recognition. One could speculate that p44/p62 actively enhances TFIIH activity towards scanning the opened repair bubble to position TFIIH factors for subsequent excision. Nonetheless, our results clearly show that the p44/p62 complex plays an active and not just a structural role in the TFIIH complex.

Methods

Purification

The genes encoding p44 and p62 were cloned from *C. thermophilum* cDNA. p62 was cloned into the pETM-11 vector (EMBL) without a tag. p44 was cloned into the pBADM-11 vector (EMBL) containing an N-terminal hexa-Histidine tag followed by a TEV cleavage site. p62 and p44 were co-expressed in *E. coli* BL21 CodonPlus (DE3) RIL cells (Agilent) and were co-purified via immobilized metal affinity chromatography (Ni TED, Machery-Nagel), followed by size exclusion chromatography (SEC), and anion exchange chromatography (AEC). SEC was conducted with a HiLoad 16/600 Superdex 200 prep grade column (GE Healthcare) in 20 mM Hepes pH 7.5, 250 mM NaCl, and 1 mM TCEP. AEC was conducted with a MonoQ 5/50 GL column (GE Healthcare). The proteins were eluted via a salt gradient ranging from 50 to 1000 mM NaCl. AEC buffers were composed of 20 mM HEPES pH 7.5, 50/1000 mM NaCl, and 1 mM TCEP. The p62/p44 protein complex was concentrated to approximately 20 mg/ml and flash frozen in liquid nitrogen for storage.

XPD and N-p44 (1-285) from *C. thermophilum* were cloned as described previously (3). XPD was expressed as N-terminally His-tagged proteins in *E. coli* ArcticExpress (DE3)-RIL cells (Agilent). Cells were grown in TB medium at 37°C until they reached an OD_{600} of 0.6-0.8. Expression was started with the addition of 0.05% L-arabinose and performed at 11°C for 20 h. p44 was expressed as N-terminally His-tagged protein in *E. coli* BL21-CodonPlus (DE3)-RIL cells (Stratagene). Cells were grown as described for ctXPD and expression was started by adding 0.1 mM IPTG at 14°C for 18 h. XPD and p44 were purified to homogeneity by metal affinity chromatography (Ni-IDA, Macherey&Nagel) as described previously (3) followed by size exclusion chromatography (SEC) (20 mM HEPES pH 7.5, 200 mM NaCl) and an additional anion exchange chromatography (AEC) step in the case of XPD. AEC was performed using a MonoQ 5/50 GL column (GE Healthcare) with 20 mM HEPES pH 7.5, 50 mM NaCl, and 1 mM TCEP as loading buffer and the same buffer containing 1 M NaCl was used for elution. The final buffer after AEC was 20 mM HEPES pH 7.5, 250 mM NaCl, and 1 mM TCEP. The proteins were concentrated to at least 5 mg/ml based on their calculated extinction coefficient using ProtParam (SwissProt) and then flash frozen for storage at -80°C.

ATPase assay

dsDNA substrates used:

F26,50 contains a fluorescein moiety covalently attached to thymine (*);

5`GACTACGTACTGTTACGGCTCCATCT*CTACCGCAATCAGGCCAGATCTGC 3`

The reverse complementary sequence to F26,50;

5`GCAGATCTGGCCTGATTGCGGTAGCGATGGAGCCGTAACAGTACGTAGTC 3`

F26,50 without the fluorescein moiety;

5`GACTACGTACTGTTACGGCTCCATCTCTACCGCAATCAGGCCAGATCTGC 3`

The NADH-coupled ATPase assay was performed as described previously (27) in plate reader format. Imaging buffer containing the NADH-reaction components was supplemented with 1 mM fresh TCEP, protein (100 nM (equimolar concentrations for XPD N-p44 and XPD p44/p62)), and 50 nM of DNA substrate. The reaction was started with the addition of 1 mM ATP to each well, and the change in OD340 (NADH) was monitored every 8 seconds/well over 30 minutes at room temperature in a Clariostar plate reader. The rates of NADH consumption were used to calculate k_{cat}. Reactions were repeated 3 times, and S.E.M used as errors values.

***In vitro* helicase assay**

Helicase activity was analyzed utilizing a fluorescence-based assay. We used an open fork substrate with a Cy3 label at the 3' end of the translocated strand where unwinding of the DNA substrate reduces quenching of the Cy3 fluorescence.

5`AGCTACCATGCCTGCACGAATTAAGCAATTCGTAATCATGGTCATAGC-Cy3 3`

and a dabcyl modification on the 5' end of the opposite strand

5` Dabcyl-GCTATGACCATGATTACGAATTGCTTGGAATCCTGACGAACTGTAG 3`

Assays were carried out in 20 mM HEPES pH 7.5, 50 mM KCl, 5 mM $MgCl_2$, and 1 mM TCEP. DNA was used at a concentration of 250 nM. Helicase activity was measured with equimolar concentrations of XPD, p44, and/or p62. The mix of reagents, were preincubated at 37°C and the reaction was subsequently started with the addition of 5 mM ATP. Kinetics were recorded with a Flourostar Optima plate reader (BMG labtech). Fluorescence was detected at an excitation wavelength of 550

nm (slit width, 2 nm) and an emission wavelength of 570 nm (slit width, 2 nm). Initial velocities were fitted with the MARS software package (BMG labtech) and represent the averages of at least three different reactions and two independent protein batches.

Single Molecule DNA Tightrope Assay

For a detailed protocol see (24). p44/p62 interactions with DNA were studied in imaging buffer (20 mM Tris pH 8.0, 10 mM KCl (100 mM for high salt), 5 mM $MgCl_2$, 1 mM TCEP). Videos for diffusion analysis were collected between 30 seconds and 5 minutes at 10 frames per second. Video analysis was performed in ImageJ as described previously (23).

Acknowledgements

We would like to thank the members of the Kad group for useful discussions. This work was supported by the Biotechnology and Biological Sciences Research Council BB/P00847X/1, BB/M019144/1, BB/I003460/1 to NMK and BB/M01603X/1 to NMK and JTB and by the German Research Foundation KI-562/7-1 to CK. The authors declare no conflict of interest.

Author contributions

Collected data: JTB, JK, WK. Designed experiments: JTB, JK, CK, NMK. Analysed Data: JTB, JK, NMK. Wrote paper: JTB, JK, CK, NMK.

References

1. Bradford PT, *et al.* (2011) Cancer and neurologic degeneration in xeroderma pigmentosum: long term follow-up characterises the role of DNA repair. *J Med Genet* 48(3):168-176.
2. Coin F, Oksenych V, & Egly JM (2007) Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Mol Cell* 26(2):245-256.
3. Kuper J, *et al.* (2014) In TFIIH, XPD helicase is exclusively devoted to DNA repair. *PLoS biology* 12(9):e1001954.
4. Roy R, *et al.* (1994) The DNA-dependent ATPase activity associated with the class II basic transcription factor BTF2/TFIIH. *J Biol Chem* 269(13):9826-9832.
5. Sung P, Higgins D, Prakash L, & Prakash S (1988) Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and DNA helicase activities but not the ability to bind ATP. *The EMBO Journal* 7(10):3263-3269.
6. Greber BJ, *et al.* (2017) The cryo-electron microscopy structure of human transcription factor IIH. *Nature* 549(7672):414.
7. Sung P, *et al.* (1993) Human xeroderma pigmentosum group D gene encodes a DMA helicase. *Nature* 365:852.
8. Schultz P, *et al.* (2000) Molecular structure of human TFIIH. *Cell* 102(5):599-607.
9. Luo J, *et al.* (2015) Architecture of the Human and Yeast General Transcription and DNA Repair Factor TFIIH. *Mol Cell* 59(5):794-806.
10. Araujo SJ, *et al.* (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 Dev* 14(3):349-359.
11. Svejstrup JQ, *et al.* (1995) Different forms of TFIIH for transcription and DNA repair: Holo-TFIIH and a nucleotide excision repairosome. *Cell* 80(1):21-28.
12. Habraken Y, Sung P, Prakash S, & Prakash L (1996) Transcription factor TFIIH and DNA endonuclease Rad2 constitute yeast nucleotide excision repair factor 3: implications for nucleotide excision repair and Cockayne syndrome. *Proceedings of the National Academy of Sciences* 93(20):10718-10722.
13. Sung P, Prakash L, Matson SW, & Prakash S (1987) RAD3 protein of Saccharomyces cerevisiae is a DNA helicase. *Proceedings of the National Academy of Sciences* 84(24):8951-8955.
14. Sung P, Prakash L, Weber S, & Prakash S (1987) The RAD3 gene of Saccharomyces cerevisiae encodes a DNA-dependent ATPase. *Proceedings of the National Academy of Sciences* 84(17):6045-6049.
15. Jawhari A, *et al.* (2002) p52 Mediates XPB function within the transcription/repair factor TFIIH. *J Biol Chem* 277(35):31761-31767.
16. Coin F, *et al.* (1998) Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nature Genetics* 20(2):184-188.
17. Greber BJ, Toso DB, Fang J, & Nogales E (2019) The complete structure of the human TFIIH core complex. *Elife* 8.

18. Bernardes de Jesus BM, Bjoras M, Coin F, & Egly JM (2008) Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. *Mol Cell Biol* 28(23):7225-7235.
19. Buechner CN*, et al.* (2014) Strand-specific recognition of DNA damages by XPD provides insights into nucleotide excision repair substrate versatility. *J Biol Chem* 289(6):3613-3624.
20. Gervais V*, et al.* (2004) TFIIH contains a PH domain involved in DNA nucleotide excision repair. *Nat Struct Mol Biol* 11(7):616-622.
21. Gileadi O, Feaver WJ, & Kornberg RD (1992) Cloning of a subunit of yeast RNA polymerase II transcription factor b and CTD kinase. *Science* 257(5075):1389.
22. Matsui P, DePaulo J, & Buratowski S (1995) An interaction between the Tfb1 and Ssl1 subunits of yeast TFIIH correlates with DNA repair activity. *Nucleic Acids Research* 23(5):767-772.
23. Kad NM, Wang H, Kennedy GG, Warshaw DM, & Van Houten B (2010) Collaborative dynamic DNA scanning by nucleotide excision repair proteins investigated by single- molecule imaging of quantum-dot-labeled proteins. *Mol Cell* 37(5):702-713.
24. Springall L, Inchingolo AV, & Kad NM (2016) DNA-Protein Interactions Studied Directly Using Single Molecule Fluorescence Imaging of Quantum Dot Tagged Proteins Moving on DNA Tightropes. *Methods Mol Biol* 1431:141-150.
25. von Hippel PH & Berg OG (1989) Facilitated target location in biological systems. *J Biol Chem* 264(2):675-678.
26. Schurr JM (1979) The one-dimensional diffusion coefficient of proteins absorbed on DNA. *Biophysical Chemistry* 9(4):413-414.
27. Barnett JT & Kad NM (2019) Understanding the coupling between DNA damage detection and UvrA's ATPase using bulk and single molecule kinetics. *FASEB journal* 33(1):763-769.

References

1. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. Nature. 2009;461(7267):1071.
2. Lindahl T. Instability and decay of the primary structure of DNA. Nature. 1993;362(6422):709-15.
3. Bustin M, Misteli T. Nongenetic functions of the genome. Science (New York, NY). 2016;352(6286):aad6933.
4. Zhou B-BS, Elledge SJ. The DNA damage response: putting checkpoints in perspective. Nature. 2000;408(6811):433.
5. Abraham RT. Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes & Development. 2001;15(17):2177-96.
6. Brooks PJ. DNA repair in neural cells: basic science and clinical implications. Mutat Res. 2002;509(1-2):93-108.
7. Fishel R, Lescoe MK, Rao M, Copeland NG, Jenkins NA, Garber J, et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell. 1993;75(5):1027-38.
8. Melis JP, van Steeg H, Luijten M. Oxidative DNA damage and nucleotide excision repair. Antioxid Redox Signal. 2013;18(18):2409-19.
9. Hopkins JM, Evans HJ. Cigarette smoke-induced DNA damage and lung cancer risks. Nature. 1980;283(5745):388-90.
10. Atamna H, Cheung I, Ames BN. A method for detecting abasic sites in living cells: age-dependent changes in base excision repair. Proceedings of the National Academy of Sciences. 2000;97(2):686-91.
11. Kingma PS, Osheroff N. Spontaneous DNA damage stimulates topoisomerase II-mediated DNA cleavage. Journal of Biological Chemistry. 1997;272(11):7488-93.
12. Russo MT, De Luca G, Degan P, Bignami M. Different DNA repair strategies to combat the threat from 8-oxoguanine. Mutat Res. 2007;614(1-2):69-76.
13. Hazra TK, Das A, Das S, Choudhury S, Kow YW, Roy R. Oxidative DNA damage repair in mammalian cells: a new perspective. DNA repair. 2007;6(4):470-80.
14. Hegde ML, Hazra TK, Mitra S. Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. Cell Res. 2008;18(1):27-47.

15. van Loon B, Markkanen E, Hübscher U. Oxygen as a friend and enemy: How to combat the mutational potential of 8-oxo-guanine. DNA repair. 2010;9(6):604-16.
16. Modrich P. Mechanisms in eukaryotic mismatch repair. The Journal of biological chemistry. 2006;281(41):30305-9.
17. Iyer RR, Pluciennik A, Burdett V, Modrich PL. DNA mismatch repair: functions and mechanisms. Chem Rev. 2006;106(2):302-23.
18. Strand M, Prolla TA, Liskay RM, Petes TD. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature. 1993;365(6443):274-6.
19. Vidaud M, Fanen P, Martin J, Ghanem N, Nicolas S, Goossens M. Three point mutations in the CFTR gene in French cystic fibrosis patients: identification by denaturing gradient gel electrophoresis. Hum Genet. 1990;85(4):446-9.
20. McKinnon PJ, Caldecott KW. DNA strand break repair and human genetic disease. Annu Rev Genomics Hum Genet. 2007;8:37-55.
21. Wyatt AW, Collins CC. In Brief: Chromothripsis and cancer. J Pathol. 2013;231(1):1-3.
22. Huang H, Zhu L, Reid BR, Drobny GP, Hopkins PB. Solution structure of a cisplatin-induced DNA interstrand cross-link. Science (New York, NY). 1995;270(5243):1842-5.
23. Zhu G, Song L, Lippard SJ. Visualizing inhibition of nucleosome mobility and transcription by cisplatin-DNA interstrand crosslinks in live mammalian cells. Cancer Res. 2013;73(14):4451-60.
24. Hanada K, Budzowska M, Modesti M, Maas A, Wyman C, Essers J, et al. The structure-specific endonuclease Mus81-Eme1 promotes conversion of interstrand DNA crosslinks into double-strands breaks. The EMBO journal. 2006;25(20):4921-32.
25. Gerson SL. MGMT: its role in cancer aetiology and cancer therapeutics. Nat Rev Cancer. 2004;4(4):296-307.
26. Wilson DM, 3rd, Bohr VA. The mechanics of base excision repair, and its relationship to aging and disease. DNA repair. 2007;6(4):544-59.
27. Bessho T, Roy R, Yamamoto K, Kasai H, Nishimura S, Tano K, et al. Repair of 8-Hydroxyguanine in DNA by Mammalian N-Methylpurine-DNA Glycosylase. Proceedings of the National Academy of Sciences of the United States of America. 1993;90(19):8901-4.
28. Almeida KH, Sobol RW. A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification. DNA repair. 2007;6(6):695-711.
29. David SS, O'Shea VL, Kundu S. Base-excision repair of oxidative DNA damage. Nature. 2007;447(7147):941-50.
30. Krokan HE, Bjoras M. Base excision repair. Cold Spring Harb Perspect Biol. 2013;5(4):a012583.
31. Hsieh P, Yamane K. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech Ageing Dev. 2008;129(7-8):391-407.

32. Su SS, Modrich P. Escherichia coli mutS-encoded protein binds to mismatched DNA base pairs. Proceedings of the National Academy of Sciences of the United States of America. 1986;83(14):5057-61.
33. Parker BO, Marinus MG. Repair of DNA heteroduplexes containing small heterologous sequences in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America. 1992;89(5):1730-4.
34. Jiricny J. Mismatch repair: the praying hands of fidelity. Curr Biol. 2000;10(21):R788-90.
35. Bolderson E, Tomimatsu N, Richard DJ, Boucher D, Kumar R, Pandita TK, et al. Phosphorylation of Exo1 modulates homologous recombination repair of DNA double-strand breaks. Nucleic Acids Res. 2010;38(6):1821-31.
36. Russo MT, De Luca G, Casorelli I, Degan P, Molatore S, Barone F, et al. Role of MUTYH and MSH2 in the control of oxidative DNA damage, genetic instability, and tumorigenesis. Cancer Res. 2009;69(10):4372-9.
37. Coverley D, Kenny MK, Lane DP, Wood RD. A role for the human single-stranded DNA binding protein HSSB/RPA in an early stage of nucleotide excision repair. Nucleic Acids Res. 1992;20(15):3873-80.
38. D'Amours D, Desnoyers S, D'Silva I, Poirier GG. Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. Biochem J. 1999;342 (Pt 2):249-68.
39. Mol CD, Hosfield DJ, Tainer JA. Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3' ends justify the means. Mutat Res. 2000;460(3-4):211-29.
40. Karimi-Busheri F, Lee J, Tomkinson AE, Weinfeld M. Repair of DNA strand gaps and nicks containing 3'-phosphate and 5'-hydroxyl termini by purified mammalian enzymes. Nucleic Acids Res. 1998;26(19):4395-400.
41. Anderson DG, Kowalczykowski SC. The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a chi-regulated manner. Cell. 1997;90(1):77-86.
42. Lieber MR, Yu K, Raghavan SC. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. DNA repair. 2006;5(9-10):1234-45.
43. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem. 2010;79:181-211.
44. Khanna KK, Jackson SP. DNA double-strand breaks: signaling, repair and the cancer connection. Nat Genet. 2001;27(3):247-54.
45. Shrivastav M, De Haro LP, Nickoloff JA. Regulation of DNA double-strand break repair pathway choice. Cell Res. 2008;18(1):134-47.

46. Lieber MR. The mechanism of human nonhomologous DNA end joining. The Journal of biological chemistry. 2008;283(1):1-5.
47. Richard DJ, Savage K, Bolderson E, Cubeddu L, So S, Ghita M, et al. hSSB1 rapidly binds at the sites of DNA double-strand breaks and is required for the efficient recruitment of the MRN complex. Nucleic Acids Res. 2011;39(5):1692-702.
48. Lamarche BJ, Orazio NI, Weitzman MD. The MRN complex in double-strand break repair and telomere maintenance. FEBS Lett. 2010;584(17):3682-95.
49. Ashton NW, Bolderson E, Cubeddu L, O'Byrne KJ, Richard DJ. Human single-stranded DNA binding proteins are essential for maintaining genomic stability. BMC Mol Biol. 2013;14:9.
50. Forget AL, Kowalczykowski SC. Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. Nature. 2012;482(7385):423-7.
51. Ragunathan K, Liu C, Ha T. RecA filament sliding on DNA facilitates homology search. eLife. 2012;1:e00067.
52. McHugh PJ, Spanswick VJ, Hartley JA. Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance. The Lancet Oncology. 2001;2(8):483-90.
53. Sarkar S, Davies AA, Ulrich HD, McHugh PJ. DNA interstrand crosslink repair during G1 involves nucleotide excision repair and DNA polymerase zeta. The EMBO journal. 2006;25(6):1285-94.
54. Klug AR, Harbut MB, Lloyd RS, Minko IG. Replication bypass of N2-deoxyguanosine interstrand cross-links by human DNA polymerases eta and iota. Chem Res Toxicol. 2012;25(3):755-62.
55. Niedernhofer LJ, Odijk H, Budzowska M, van Drunen E, Maas A, Theil AF, et al. The structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. Mol Cell Biol. 2004;24(13):5776-87.
56. McHugh PJ, Sones WR, Hartley JA. Repair of intermediate structures produced at DNA interstrand cross-links in Saccharomyces cerevisiae. Mol Cell Biol. 2000;20(10):3425-33.
57. De Silva IU, McHugh PJ, Clingen PH, Hartley JA. Defining the roles of nucleotide excision repair and recombination in the repair of DNA interstrand cross-links in mammalian cells. Mol Cell Biol. 2000;20(21):7980-90.
58. Clauson C, Scharer OD, Niedernhofer L. Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. Cold Spring Harb Perspect Biol. 2013;5(10):a012732.
59. Bakker JL, van Mil SE, Crossan G, Sabbaghian N, De Leeneer K, Poppe B, et al. Analysis of the novel fanconi anemia gene SLX4/FANCP in familial breast cancer cases. Hum Mutat. 2013;34(1):70-3.

60. Petermann E, Helleday T. Pathways of mammalian replication fork restart. Nat Rev Mol Cell Biol. 2010;11(10):683-7.
61. Kisker C, Kuper J, Van Houten B. Prokaryotic nucleotide excision repair. Cold Spring Harb Perspect Biol. 2013;5(3):a012591.
62. Hashimoto S, Anai H, Hanada K. Mechanisms of interstrand DNA crosslink repair and human disorders. Genes Environ. 2016;38:9.
63. Hollaender A. Effect of long ultraviolet and short visible radiation (3500 to 4900A) on escherichia coli. Journal of bacteriology. 1943;46(6):531-41.
64. Setlow RB. The Wavelengths in Sunlight Effective in Producing Skin Cancer: A Theoretical Analysis. Proceedings of the National Academy of Sciences. 1974;71(9):3363-6.
65. Franklin WA, Haseltine WA. Removal of UV light-induced pyrimidine-pyrimidone(6-4) products from Escherichia coli DNA requires the uvrA, uvrB, and urvC gene products. Proceedings of the National Academy of Sciences of the United States of America. 1984;81(12):3821-4.
66. de Gruijl FR, Forbes PD. UV-induced skin cancer in a hairless mouse model. Bioessays. 1995;17(7):651-60.
67. Mitchell PJ, Tjian R. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science (New York, NY). 1989;245(4916):371-8.
68. Chandrasekhar D, Van Houten B. In vivo formation and repair of cyclobutane pyrimidine dimers and 6-4 photoproducts measured at the gene and nucleotide level in Escherichia coli. Mutat Res. 2000;450(1-2):19-40.
69. Yoon JH, Lee CS, O'Connor TR, Yasui A, Pfeifer GP. The DNA damage spectrum produced by simulated sunlight. J Mol Biol. 2000;299(3):681-93.
70. Pfeifer GP, You YH, Besaratinia A. Mutations induced by ultraviolet light. Mutat Res. 2005;571(1-2):19-31.
71. Batista LF, Kaina B, Meneghini R, Menck CF. How DNA lesions are turned into powerful killing structures: insights from UV-induced apoptosis. Mutat Res. 2009;681(2-3):197-208.
72. Muotri AR, Marchetto MC, Suzuki MF, Okazaki K, Lotfi CF, Brumatti G, et al. Low amounts of the DNA repair XPA protein are sufficient to recover UV-resistance. Carcinogenesis. 2002;23(6):1039-46.
73. Yokoyama H, Mizutani R. Structural Biology of DNA (6-4) Photoproducts Formed by Ultraviolet Radiation and Interactions with Their Binding Proteins. International journal of molecular sciences. 2014;15(11):20321-38.
74. Boyce RP, Howard-Flanders P. 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. 1964;51:293-300.

75. Pettijohn D, Hanawalt P. Evidence for repair-replication of ultraviolet damaged DNA in bacteria. Journal of Molecular Biology. 1964;9(2):395-410.
76. Rasmussen RE, Painter RB. EVIDENCE FOR REPAIR OF ULTRA-VIOLET DAMAGED DEOXYRIBONUCLEIC ACID IN CULTURED MAMMALIAN CELLS. Nature. 1964;203:1360-2.
77. Setlow RB, Carrier WL. THE DISAPPEARANCE OF THYMINE DIMERS FROM DNA: AN ERROR-CORRECTING MECHANISM. Proceedings of the National Academy of Sciences of the United States of America. 1964;51:226-31.
78. Sancar A, Rupp WD. A novel repair enzyme: UVRABC excision nuclease of Escherichia coli cuts a DNA strand on both sides of the damaged region. Cell. 1983;33(1):249-60.
79. Van Houten B, Croteau DL, DellaVecchia MJ, Wang H, Kisker C. 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis. 2005;577(1-2 SPEC. ISS.):92-117 (18).
80. Wood RD. DNA damage recognition during nucleotide excision repair in mammalian cells. Biochimie. 1999;81(1-2):39-44.
81. Mitchell DL, Humphrey RM, Adair GM, Thompson LH, Clarkson JM. Repair of (6-4) photoproducts correlates with split-dose recovery in UV-irradiated normal and hypersensitive rodent cells. Mutation Research/DNA Repair Reports. 1988;193(1):53-63.
82. Bohr VA, Smith CA, Okumoto DS, Hanawalt PC. DNA-Repair in an Active Gene - Removal of Pyrimidine Dimers from the Dhfr Gene of Cho Cells Is Much More Efficient Than in the Genome Overall. Cell. 1985;40(2):359-69.
83. Van Hoffen A, Venema J, Meschini R, Van Zeeland A, Mullenders L. Transcription-coupled repair removes both cyclobutane pyrimidine dimers and 6-4 photoproducts with equal efficiency and in a sequential way from transcribed DNA in xeroderma pigmentosum group C fibroblasts. The EMBO journal. 1995;14(2):360-7.
84. Mellon I, Spivak G, Hanawalt PC. Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. Cell. 1987;51(2):241-9.
85. Selby CP, Sancar A. Transcription preferentially inhibits nucleotide excision repair of the template DNA strand in vitro. The Journal of biological chemistry. 1990;265(34):21330-6.
86. Selby CP, Sancar A. Gene- and strand-specific repair in vitro: partial purification of a transcription-repair coupling factor. Proceedings of the National Academy of Sciences of the United States of America. 1991;88(18):8232-6.
87. Selby CP, Sancar A. Transcription-repair coupling and mutation frequency decline. J Bacteriol. 1993;175(23):7509-14.

88. Aamann MD, Muftuoglu M, Bohr VA, Stevnsner T. Multiple interaction partners for Cockayne syndrome proteins: Implications for genome and transcriptome maintenance. Mechanisms of Ageing and Development. 2013;134(5-6):212-24.
89. Zou Y, Van Houten B. Strand opening by the UvrA2B complex allows dynamic recognition of DNA damage. EMBO Journal. 1999;18(17):4889-901.
90. Orren DK, Sancar A. The (A) BC excinuclease of Escherichia coli has only the UvrB and UvrC subunits in the incision complex. Proceedings of the National Academy of Sciences. 1989;86(14):5237-41.
91. Bessho T, Mu D, Sancar A. Initiation of DNA interstrand cross-link repair in humans: the nucleotide excision repair system makes dual incisions 5' to the cross-linked base and removes a 22- to 28-nucleotide-long damage-free strand. Mol Cell Biol. 1997;17(12):6822-30.
92. Caron PR, Grossman L. Incision of damaged versus nondamaged DNA by the Escherichia coli UvrABC proteins. Nucleic Acids Res. 1988;16(16):7855-65.
93. Truglio JJ, Croteau DL, Van Houten B, Kisker C. Prokaryotic nucleotide excision repair: the UvrABC system. Chem Rev. 2006;106(2):233-52.
94. Reardon JT, Sancar A. Molecular anatomy of the human excision nuclease assembled at sites of DNA damage. Mol Cell Biol. 2002;22(16):5938-45.
95. Sancar A, Wharton RP, Seltzer S, Kacinski BM, Clarke ND, Rupp WD. Identification of the uvrA gene product. J Mol Biol. 1981;148(1):45-62.
96. Kacinski BM, Sancar A, Rupp WD. A general approach for purifying proteins encoded by cloned genes without using a functional assay: isolation of the uvrA gene product from radiolabeled maxicells. Nucleic acids research. 1981;9(18):4495-508.
97. Seeberg E, Steinum A-L. Purification and properties of the uvrA protein from Escherichia coli. Proceedings of the National Academy of Sciences. 1982;79(4):988-92.
98. Yeung AT, Mattes WB, Oh EY, Grossman L. Enzymatic properties of purified Escherichia coli uvrABC proteins. Proceedings of the National Academy of Sciences. 1983;80(20):6157-61.
99. Thiagalingam S, Grossman L. Both Atpase Sites of Escherichia-Coli Uvra Have Functional Roles in Nucleotide Excision Repair. Journal of Biological Chemistry. 1991;266(17):11395-403.
100. Truglio JJ, Karakas E, Rhau B, Wang H, DellaVecchia MJ, Van Houten B, et al. Structural basis for DNA recognition and processing by UvrB. Nature structural & molecular biology. 2006;13(4):360-4.
101. Mazur SJ, Grossman L. Dimerization of Escherichia coli UvrA and its binding to undamaged and ultraviolet light damaged DNA. Biochemistry. 1991;30(18):4432-43.

102. Oh EY, Claassen L, Thiagalingam S, Mazur S, Grossman L. ATPase activity of the UvrA and UvrAB protein complexes of the Escherichia coli UvrABC endonuclease. Nucleic acids research. 1989;17(11):4145-59.
103. Van Houten B, Gamper H, Sancar A, Hearst J. DNase I footprint of ABC excinuclease. Journal of Biological Chemistry. 1987;262(27):13180-7.
104. Shi Q, Thresher R, Sancar A, Griffith J. Electron microscopic study of (A)BC excinuclease. DNA is sharply bent in the UvrB-DNA complex. J Mol Biol. 1992;226(2):425-32.
105. Oh EY, Grossman L. Helicase properties of the Escherichia coli UvrAB protein complex. Proceedings of the National Academy of Sciences of the United States of America. 1987;84(11):3638-42.
106. Wagner K, Moolenaar G, van Noort J, Goosen N. Single-molecule analysis reveals two separate DNA-binding domains in the Escherichia coli UvrA dimer. Nucleic acids research. 2009;37(6):1962-72.
107. Dickerson RE. DNA bending: the prevalence of kinkiness and the virtues of normality. Nucleic Acids Res. 1998;26(8):1906-26.
108. Springall L, Hughes CD, Simons M, Azinas S, Van Houten B, Kad NM. Recruitment of UvrBC complexes to UV-induced damage in the absence of UvrA increases cell survival. Nucleic Acids Res. 2018;46(3):1256-65.
109. Kad NM, Wang H, Kennedy GG, Warshaw DM, Van Houten B. Collaborative dynamic DNA scanning by nucleotide excision repair proteins investigated by single- molecule imaging of quantum-dot-labeled proteins. Molecular cell. 2010;37(5):702-13.
110. Verhoeven EEA, Wyman C, Moolenaar GF, Goosen N. The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands. The EMBO journal. 2002;21(15):4196-205.
111. Malta E, Moolenaar GF, Goosen N. Dynamics of the UvrABC nucleotide excision repair proteins analyzed by fluorescence resonance energy transfer. Biochemistry. 2007;46(31):9080-8.
112. Pakotiprapha D, Samuels M, Shen K, Hu JH, Jeruzalmi D. Structure and mechanism of the UvrA–UvrB DNA damage sensor. Nature Structural & Molecular Biology. 2012;19(3):291-8.
113. Hughes CD, Wang H, Ghodke H, Simons M, Towheed A, Peng Y, et al. Real-time single-molecule imaging reveals a direct interaction between UvrC and UvrB on DNA tightropes. Nucleic Acids Res. 2013;41(9):4901-12.
114. Seeberg E. Reconstitution of an Escherichia coli repair endonuclease activity from the separated uvrA+ and uvrB+/uvrC+ gene products. Proceedings of the National Academy of Sciences of the United States of America. 1978;75(6):2569-73.

115. Wirth N, Gross J, Roth HM, Buechner CN, Kisker C, Tessmer I. Conservation and Divergence in Nucleotide Excision Repair Lesion Recognition. The Journal of biological chemistry. 2016;291(36):18932-46.
116. Verhoeven EE, van Kesteren M, Moolenaar GF, Visse R, Goosen N. Catalytic sites for 3' and 5' incision of Escherichia coli nucleotide excision repair are both located in UvrC. The Journal of biological chemistry. 2000;275(7):5120-3.
117. Yoakum GH, Grossman L. Identification of E. coli uvrC protein. Nature. 1981;292(5819):171-3.
118. Moolenaar GF, Uiterkamp RS, Zwijnenburg DA, Goosen N. The C-terminal region of the Escherichia coli UvrC protein, which is homologous to the C-terminal region of the human ERCC1 protein, is involved in DNA binding and 5'-incision. Nucleic Acids Res. 1998;26(2):462-8.
119. Stracy M, Jaciuk M, Uphoff S, Kapanidis AN, Nowotny M, Sherratt DJ, et al. Single-molecule imaging of UvrA and UvrB recruitment to DNA lesions in living Escherichia coli. Nature communications. 2016;7:12568.
120. Wagner K, Moolenaar GF, Goosen N. Role of the two ATPase domains of Escherichia coli UvrA in binding non-bulky DNA lesions and interaction with UvrB. DNA repair. 2010;9(11):1176-86.
121. Schrodinger, LLC. The AxPyMOL Molecular Graphics Plugin for Microsoft PowerPoint, Version 1.8. 2015.
122. Sancar A, Sancar GB. DNA repair enzymes. Annu Rev Biochem. 1988;57:29-67.
123. Orren DK, Selby CP, Hearst JE, Sancar A. Post-incision steps of nucleotide excision repair in Escherichia coli. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I. The Journal of biological chemistry. 1992;267(2):780-8.
124. Seeley TW, Grossman L. Mutations in the Escherichia coli UvrB ATPase motif compromise excision repair capacity. Proceedings of the National Academy of Sciences of the United States of America. 1989;86(17):6577-81.
125. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res. 1989;17(12):4713-30.
126. Theis K, Chen PJ, Skorvaga M, Van Houten B, Kisker C. Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. The EMBO journal. 1999;18(24):6899-907.
127. Skorvaga M, Theis K, Mandavilli BS, Kisker C, Van Houten B. The beta -hairpin motif of UvrB is essential for DNA binding, damage processing, and UvrC-mediated incisions. The Journal of biological chemistry. 2002;277(2):1553-9.
128. Gordienko I, Rupp WD. UvrAB activity at a damaged DNA site: is unpaired DNA present? The EMBO journal. 1997;16(4):880-8.

129. Orren DK, Sancar A. Formation and enzymatic properties of the UvrB.DNA complex. The Journal of biological chemistry. 1990;265(26):15796-803.
130. Theis K, Skorvaga M, Machius M, Nakagawa N, Van Houten B, Kisker C. The nucleotide excision repair protein UvrB, a helicase-like enzyme with a catch. Mutat Res. 2000;460(3-4):277-300.
131. Lin J-J, Phillips AM, Hearst JE, Sancar A. Active site of (A) BC excinuclease. II. Binding, bending, and catalysis mutants of UvrB reveal a direct role in 3'and an indirect role in 5'incision. Journal of Biological Chemistry. 1992;267(25):17693-700.
132. Visse R, King A, Moolenaar GF, Goosen N, Vandeputte P. Protein-DNA Interactions and Alterations in the DNA-Structure Upon Uvrb DNA Preincision Complex-Formation during Nucleotide Excision-Repair in Escherichia-Coli. Biochemistry. 1994;33(33):9881-8.
133. Machius M, Henry L, Palnitkar M, Deisenhofer J. Crystal structure of the DNA nucleotide excision repair enzyme UvrB from Thermus thermophilus. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(21):11717-22.
134. DellaVecchia MJ, Croteau DL, Skorvaga M, Dezhurov SV, Lavrik OI, Van Houten B. Analyzing the handoff of DNA from UvrA to UvrB utilizing DNA-protein photoaffinity labeling. The Journal of biological chemistry. 2004;279(43):45245-56.
135. Moolenaar GF, Höglund L, Goosen N. Clue to damage recognition by UvrB: residues in the β-hairpin structure prevent binding to non-damaged DNA. The EMBO journal. 2001;20(21):6140-9.
136. Zou Y, Luo C, Geacintov NE. Hierarchy of DNA damage recognition in Escherichia coli nucleotide excision repair. Biochemistry. 2001;40(9):2923-31.
137. Truglio JJ, Croteau DL, Skorvaga M, DellaVecchia MJ, Theis K, Mandavilli BS, et al. Interactions between UvrA and UvrB: the role of UvrB's domain 2 in nucleotide excision repair. The EMBO journal. 2004;23(13):2498-509.
138. Moolenaar GF, Herron MF, Monaco V, van der Marel GA, van Boom JH, Visse R, et al. The role of ATP binding and hydrolysis by UvrB during nucleotide excision repair. The Journal of biological chemistry. 2000;275(11):8044-50.
139. Moolenaar GF, Franken KL, Dijkstra DM, Thomas-Oates JE, Visse R, van de Putte P, et al. The C-terminal region of the UvrB protein of Escherichia coli contains an important determinant for UvrC binding to the preincision complex but not the catalytic site for 3'-incision. The Journal of biological chemistry. 1995;270(51):30508-15.
140. Moolenaar GF, Franken KL, van de Putte P, Goosen N. Function of the homologous regions of the Escherichia coli DNA excision repair proteins UvrB and UvrC in stabilization of the UvrBC-DNA complex and in 3'-incision. Mutat Res. 1997;385(3):195-203.

141. Caron PR, Kushner SR, Grossman L. Involvement of helicase II (uvrD gene product) and DNA polymerase I in excision mediated by the uvrABC protein complex. Proceedings of the National Academy of Sciences of the United States of America. 1985;82(15):4925-9.
142. Husain I, Van Houten B, Thomas DC, Abdel-Monem M, Sancar A. Effect of DNA polymerase I and DNA helicase II on the turnover rate of UvrABC excision nuclease. Proceedings of the National Academy of Sciences. 1985;82(20):6774-8.
143. Matson SW. DNA helicases of Escherichia coli. Progress in nucleic acid research and molecular biology. 40: Elsevier; 1991. p. 289-326.
144. Runyon GT, Bear DG, Lohman TM. Escherichia coli helicase II (UvrD) protein initiates DNA unwinding at nicks and blunt ends. Proceedings of the National Academy of Sciences. 1990;87(16):6383-7.
145. Lehman I, Chien JR. Persistence of Deoxyribonucleic Acid Polymerase I and Its 5'→ 3'Exonuclease Activity in PolA Mutants of Escherichia coli K12. Journal of Biological Chemistry. 1973;248(22):7717-23.
146. Peng Y, Wang H, Santana-Santos L, Kisker C, Van Houten B. Nucleotide excision repair from bacteria to humans: Structure–function studies. Chemical Carcinogenesis: Springer; 2011. p. 267-96.
147. de Laat WL, Jaspers NG, Hoeijmakers JH. Molecular mechanism of nucleotide excision repair. Genes Dev. 1999;13(7):768-85.
148. Aboussekhra A, Biggerstaff M, Shivji MKK, Vilpo JA, Moncollin V, Podust VN, et al. Mammalian DNA Nucleotide Excision-Repair Reconstituted with Purified Protein-Components. Cell. 1995;80(6):859-68.
149. Wilcox DR, Prakash L. Incision and postincision steps of pyrimidine dimer removal in excision-defective mutants of Saccharomyces cerevisiae. J Bacteriol. 1981;148(2):618-23.
150. Schaeffer L, Moncollin V, Roy R, Staub A, Mezzina M, Sarasin A, et al. The Ercc2/DNA Repair Protein Is Associated with the Class-Ii Btf2/Tfiih Transcription Factor. Embo Journal. 1994;13(10):2388-92.
151. Reynolds RJ, Friedberg EC. Molecular mechanisms of pyrimidine dimer excision in Saccharomyces cerevisiae: incision of ultraviolet-irradiated deoxyribonucleic acid in vivo. J Bacteriol. 1981;146(2):692-704.
152. Miller RD, Prakash L, Prakash S. Genetic control of excision of Saccharomyces cerevisiae interstrand DNA cross-links induced by psoralen plus near-UV light. Mol Cell Biol. 1982;2(8):939-48.
153. Park E, Guzder SN, Koken MHM, Jaspersdekker I, Weeda G, Hoeijmakers JHJ, et al. Rad25 (Ssl2), the Yeast Homolog of the Human Xeroderma-Pigmentosum Group-B DNA-Repair Gene, Is Essential for Viability. Proceedings of the National Academy of Sciences of the United States of America. 1992;89(23):11416-20.

154. Gerard M, Fischer L, Moncollin V, Chipoulet JM, Chambon P, Egly JM. Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2. The Journal of biological chemistry. 1991;266(31):20940-5.
155. van Vuuren AJ, Vermeulen W, Ma L, Weeda G, Appeldoorn E, Jaspers NG, et al. Correction of xeroderma pigmentosum repair defect by basal transcription factor BTF2 (TFIIH). The EMBO journal. 1994;13(7):1645-53.
156. Bardwell L, Bardwell AJ, Feaver WJ, Svejstrup JQ, Kornberg RD, Friedberg EC. Yeast RAD3 protein binds directly to both SSL2 and SSL1 proteins: implications for the structure and function of transcription/repair factor b. Proceedings of the National Academy of Sciences of the United States of America. 1994;91(9):3926-30.
157. Compe E, Egly JM. TFIIH: when transcription met DNA repair. Nat Rev Mol Cell Biol. 2012;13(6):343-54.
158. Hwang JR, Moncollin V, Vermeulen W, Seroz T, van Vuuren H, Hoeijmakers JH, et al. A 3′→ 5′ XPB helicase defect in repair/transcription factor TFIIH of xeroderma pigmentosum group B affects both DNA repair and transcription. Journal of Biological Chemistry. 1996;271(27):15898-904.
159. Coin F, Marinoni JC, Rodolfo C, Fribourg S, Pedrini AM, Egly JM. Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. Nat Genet. 1998;20(2):184-8.
160. Zurita M, Merino C. The transcriptional complexity of the TFIIH complex. Trends Genet. 2003;19(10):578-84.
161. Sugasawa K, Ng JMY, Masutani C, Iwai S, van der Spek PJ, Eker APM, et al. Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. Molecular cell. 1998;2(2):223-32.
162. Sugasawa K, Masutani C, Uchida A, Maekawa T, van der Spek PJ, Bootsma D, et al. HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair in vitro. Mol Cell Biol. 1996;16(9):4852-61.
163. Volker M, Mone MJ, Karmakar P, van Hoffen A, Schul W, Vermeulen W, et al. Sequential assembly of the nucleotide excision repair factors in vivo. Molecular cell. 2001;8(1):213-24.
164. Mu D, Wakasugi M, Hsu DS, Sancar A. Characterization of reaction intermediates of human excision repair nuclease. Journal of Biological Chemistry. 1997;272(46):28971-9.
165. Fitch ME, Nakajima S, Yasui A, Ford JM. In vivo recruitment of XPC to UV-induced cyclobutane pyrimidine dimers by the DDB2 gene product. The Journal of biological chemistry. 2003;278(47):46906-10.
166. Scrima A, Koníčková R, Czyzewski BK, Kawasaki Y, Jeffrey PD, Groisman R, et al. Structural basis of UV DNA-damage recognition by the DDB1–DDB2 complex. Cell. 2008;135(7):1213-23.

167. Chu G, Chang E. Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. Science (New York, NY). 1988;242(4878):564-7.
168. Yasuda T, Sugasawa K, Shimizu Y, Iwai S, Shiomi T, Hanaoka F. Nucleosomal structure of undamaged DNA regions suppresses the non-specific DNA binding of the XPC complex. DNA repair. 2005;4(3):389-95.
169. Osakabe A, Tachiwana H, Kagawa W, Horikoshi N, Matsumoto S, Hasegawa M, et al. Structural basis of pyrimidine-pyrimidone (6–4) photoproduct recognition by UV-DDB in the nucleosome. Scientific reports. 2015;5:16330.
170. Hara R, Mo J, Sancar A. DNA damage in the nucleosome core is refractory to repair by human excision nuclease. Mol Cell Biol. 2000;20(24):9173-81.
171. Min JH, Pavletich NP. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. Nature. 2007;449(7162):570-5.
172. Mu H, Geacintov NE, Broyde S, Yeo JE, Scharer OD. Molecular basis for damage recognition and verification by XPC-RAD23B and TFIIH in nucleotide excision repair. DNA repair. 2018;71:33-42.
173. Gunz D, Hess MT, Naegeli H. Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism. The Journal of biological chemistry. 1996;271(41):25089-98.
174. Janićijević A, Sugasawa K, Shimizu Y, Hanaoka F, Wijgers N, Djurica M, et al. DNA bending by the human damage recognition complex XPC–HR23B. DNA repair. 2003;2(3):325-36.
175. Kong M, Liu L, Chen X, Driscoll KI, Mao P, Böhm S, et al. Single-molecule imaging reveals that Rad4 employs a dynamic DNA damage recognition process. Molecular cell. 2016;64(2):376-87.
176. Cheon NY, Kim HS, Yeo JE, Scharer OD, Lee JY. Single-molecule visualization reveals the damage search mechanism for the human NER protein XPC-RAD23B. Nucleic Acids Res. 2019;47(16):8337-47.
177. Luo J, Cimermancic P, Viswanath S, Ebmeier CC, Kim B, Dehecq M, et al. Architecture of the Human and Yeast General Transcription and DNA Repair Factor TFIIH. Molecular cell. 2015;59(5):794-806.
178. Bernardes de Jesus BM, Bjoras M, Coin F, Egly JM. Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. Mol Cell Biol. 2008;28(23):7225-35.
179. Riedl T, Hanaoka F, Egly JM. The comings and goings of nucleotide excision repair factors on damaged DNA. The EMBO journal. 2003;22(19):5293-303.

180. Evans E, Moggs JG, Hwang JR, Egly JM, Wood RD. Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. The EMBO journal. 1997;16(21):6559-73.
181. Tapias A, Auriol J, Forget D, Enzlin JH, Scharer OD, Coin F, et al. Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. The Journal of biological chemistry. 2004;279(18):19074-83.
182. Fan L, Arvai AS, Cooper PK, Iwai S, Hanaoka F, Tainer JA. Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair. Molecular cell. 2006;22(1):27-37.
183. Oksenych V, Bernardes de Jesus B, Zhovmer A, Egly JM, Coin F. Molecular insights into the recruitment of TFIIH to sites of DNA damage. The EMBO journal. 2009;28(19):2971-80.
184. Coin F, Oksenych V, Egly JM. Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. Molecular cell. 2007;26(2):245-56.
185. Feaver WJ, Svejstrup JQ, Bardwell L, Bardwell AJ, Buratowski S, Gulyas KD, et al. Dual roles of a multiprotein complex from S. cerevisiae in transcription and DNA repair. Cell. 1993;75(7):1379-87.
186. Schaeffer L, Roy R, Humbert S, Moncollin V, Vermeulen W, Hoeijmakers JH, et al. DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. Science (New York, NY). 1993;260(5104):58-63.
187. Sung P, Bailly V, Weber C, Thompson LH, Prakash L, Prakash S. Human xeroderma pigmentosum group D gene encodes a DNA helicase. Nature. 1993;365(6449):852-5.
188. Richards JD, Cubeddu L, Roberts J, Luau HT, White MF. The archaeal XPB protein is a ssDNA-dependent ATPase with a novel partner. Journal of Molecular Biology. 2008;376(3):634-44.
189. Sugasawa K, Akagi J, Nishi R, Iwai S, Hanaoka F. Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning. Molecular cell. 2009;36(4):642-53.
190. Singleton MR, Dillingham MS, Wigley DB. Structure and mechanism of helicases and nucleic acid translocases. Annu Rev Biochem. 2007;76:23-50.
191. Rudolf J, Makrantoni V, Ingledew WJ, Stark MJ, White MF. The DNA repair helicases XPD and FancJ have essential iron-sulfur domains. Molecular cell. 2006;23(6):801-8.
192. White MF, Dillingham MS. Iron–sulphur clusters in nucleic acid processing enzymes. Current opinion in structural biology. 2012;22(1):94-100.

193. Fuss JO, Tsai C-L, Ishida JP, Tainer JA. Emerging critical roles of Fe–S clusters in DNA replication and repair. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research. 2015;1853(6):1253-71.
194. Porello SL, Cannon MJ, David SS. A substrate recognition role for the [4Fe-4S] 2+ cluster of the DNA repair glycosylase MutY. Biochemistry. 1998;37(18):6465-75.
195. Cunningham RP, Asahara H, Bank JF, Scholes CP, Salerno JC, Surerus K, et al. Endonuclease III is an iron-sulfur protein. Biochemistry. 1989;28(10):4450-5.
196. Kuper J, Kisker C. Damage recognition in nucleotide excision DNA repair. Curr Opin Struct Biol. 2012;22(1):88-93.
197. Pugh RA, Honda M, Leesley H, Thomas A, Lin Y, Nilges MJ, et al. The iron-containing domain is essential in Rad3 helicases for coupling of ATP hydrolysis to DNA translocation and for targeting the helicase to the single-stranded DNA-double-stranded DNA junction. The Journal of biological chemistry. 2008;283(3):1732-43.
198. Pugh RA, Wu CG, Spies M. Regulation of translocation polarity by helicase domain 1 in SF2B helicases. The EMBO journal. 2012;31(2):503-14.
199. Liu H, Rudolf J, Johnson KA, McMahon SA, Oke M, Carter L, et al. Structure of the DNA repair helicase XPD. Cell. 2008;133(5):801-12.
200. Buechner CN, Heil K, Michels G, Carell T, Kisker C, Tessmer I. Strand-specific recognition of DNA damages by XPD provides insights into nucleotide excision repair substrate versatility. The Journal of biological chemistry. 2014;289(6):3613-24.
201. Cheng K, Wigley DB. DNA translocation mechanism of an XPD family helicase. Elife. 2018;7.
202. Mackintosh SG, Raney KD. DNA unwinding and protein displacement by superfamily 1 and superfamily 2 helicases. Nucleic acids research. 2006;34(15):4154-9.
203. Qi Z, Pugh RA, Spies M, Chemla YR. Sequence-dependent base pair stepping dynamics in XPD helicase unwinding. Elife. 2013;2:e00334.
204. Ali JA, Lohman TM. Kinetic measurement of the step size of DNA unwinding by Escherichia coli UvrD helicase. Science (New York, NY). 1997;275(5298):377-80.
205. Greber BJ, Nguyen THD, Fang J, Afonine PV, Adams PD, Nogales E. The cryo-electron microscopy structure of human transcription factor IIH. Nature. 2017;549(7672):414.
206. Constantinescu-Aruxandei D, Petrovic-Stojanovska B, Penedo JC, White MF, Naismith JH. Mechanism of DNA loading by the DNA repair helicase XPD. Nucleic Acids Res. 2016;44(6):2806-15.
207. Kuper J, Braun C, Elias A, Michels G, Sauer F, Schmitt DR, et al. In TFIIH, XPD Helicase Is Exclusively Devoted to DNA Repair. Plos Biology. 2014;12(9):e1001954.
208. Kokic G, Chernev A, Tegunov D, Dienemann C, Urlaub H, Cramer P. Structural basis of TFIIH activation for nucleotide excision repair. Nature communications. 2019;10(1):2885.

209. Honda M, Park J, Pugh RA, Ha T, Spies M. Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase. Molecular cell. 2009;35(5):694-703.
210. Mathieu N, Kaczmarek N, Ruthemann P, Luch A, Naegeli H. DNA quality control by a lesion sensor pocket of the xeroderma pigmentosum group D helicase subunit of TFIIH. Curr Biol. 2013;23(3):204-12.
211. Kuper J, Wolski SC, Michels G, Kisker C. Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. The EMBO journal. 2012;31(2):494-502.
212. Huang J-C, Svoboda DL, Reardon JT, Sancar A. Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5'and the 6th phosphodiester bond 3'to the photodimer. Proceedings of the National Academy of Sciences. 1992;89(8):3664-8.
213. Evans E, Fellows J, Coffer A, Wood RD. Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. The EMBO journal. 1997;16(3):625-38.
214. O'Donovan A, Wood RD. Identical defects in DNA repair in xeroderma pigmentosum group G and rodent ERCC group 5. Nature. 1993;363(6425):185-8.
215. Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, et al. Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. Cell. 1996;86(5):811-22.
216. de Laat W. Mapping of interaction domains between human repair proteins ERCC1 and XPF. Nucleic Acids Research. 1998;26(18):4146-52.
217. Riedl T, Hanaoka F, Egly JM. The comings and goings of nucleotide excision repair factors on damaged DNA. The EMBO journal. 2003;22(19):5293-303.
218. Moggs JG, Yarema KJ, Essigmann JM, Wood RD. Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. The Journal of biological chemistry. 1996;271(12):7177-86.
219. Wakasugi M, Reardon JT, Sancar A. The non-catalytic function of XPG protein during dual incision in human nucleotide excision repair. The Journal of biological chemistry. 1997;272(25):16030-4.
220. Staresincic L, Fagbemi AF, Enzlin JH, Gourdin AM, Wijgers N, Dunand-Sauthier I, et al. Coordination of dual incision and repair synthesis in human nucleotide excision repair. The EMBO journal. 2009;28(8):1111-20.

221. Hu J, Choi JH, Gaddameedhi S, Kemp MG, Reardon JT, Sancar A. Nucleotide excision repair in human cells: fate of the excised oligonucleotide carrying DNA damage in vivo. The Journal of biological chemistry. 2013;288(29):20918-26.
222. Kemp MG, Reardon JT, Lindsey-Boltz LA, Sancar A. Mechanism of release and fate of excised oligonucleotides during nucleotide excision repair. The Journal of biological chemistry. 2012;287(27):22889-99.
223. Shivji MK, Podust VN, Hubscher U, Wood RD. Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. Biochemistry. 1995;34(15):5011-7.
224. Nichols AF, Sancar A. Purification of PCNA as a nucleotide excision repair protein. Nucleic Acids Research. 1992;20(10):2441-6.
225. Hunting DJ, Gowans BJ, Dresler SL. DNA polymerase delta mediates excision repair in growing cells damaged with ultraviolet radiation. Biochemistry and cell biology. 1991;69(4):303-8.
226. Araujo SJ, Tirode F, Coin F, Pospiech H, Syvaoja JE, Stucki M, et al. 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 Dev. 2000;14(3):349-59.
227. Vermeulen W, Van Vuuren A, Chipoulet M, Schaeffer L, Appeldoorn E, Weeda G, et al., editors. Three unusual repair deficiencies associated with transcription factor BTF2 (TFIIH): evidence for the existence of a transcription syndrome. Cold Spring Harbor symposia on quantitative biology; 1994: Cold Spring Harbor Laboratory Press.
228. Dubaele S, Proietti De Santis L, Bienstock RJ, Keriel A, Stefanini M, Van Houten B, et al. Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. Molecular cell. 2003;11(6):1635-46.
229. Sung P, Prakash L, Matson SW, Prakash S. RAD3 protein of Saccharomyces cerevisiae is a DNA helicase. Proceedings of the National Academy of Sciences. 1987;84(24):8951-5.
230. Sung P, Prakash L, Weber S, Prakash S. The RAD3 gene of Saccharomyces cerevisiae encodes a DNA-dependent ATPase. Proceedings of the National Academy of Sciences of the United States of America. 1987;84(17):6045-9.
231. Svejstrup JQ, Wang Z, Feaver WJ, Wu X, Bushnell DA, Donahue TF, et al. Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. Cell. 1995;80(1):21-8.
232. Schultz P, Fribourg S, Poterszman A, Mallouh V, Moras D, Egly JM. Molecular structure of human TFIIH. Cell. 2000;102(5):599-607.
233. Coin F, Oksenych V, Mocquet V, Groh S, Blattner C, Egly JM. Nucleotide excision repair driven by the dissociation of CAK from TFIIH. Molecular cell. 2008;31(1):9-20.

234. Abdulrahman W, Iltis I, Radu L, Braun C, Maglott-Roth A, Giraudon C, et al. 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. 2013;110(8):E633-E42.
235. Sandrock B, Egly JM. A yeast four-hybrid system identifies Cdk-activating kinase as a regulator of the XPD helicase, a subunit of transcription factor IIH. The Journal of biological chemistry. 2001;276(38):35328-33.
236. Feaver WJ, Svejstrup JQ, Henry NL, Kornberg RD. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. Cell. 1994;79(6):1103-9.
237. Lu H, Zawel L, Fisher L, Egly J-M, Reinberg D. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature. 1992;358(6388):641.
238. Serizawa H, Makela TP, Conaway JW, Conaway RC, Weinberg RA, Young RA. Association of Cdk-activating kinase subunits with transcription factor TFIIH. Nature. 1995;374(6519):280-2.
239. Greber BJ, Toso DB, Fang J, Nogales E. The complete structure of the human TFIIH core complex. Elife. 2019;8.
240. Chang WH, Kornberg RD. Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. Cell. 2000;102(5):609-13.
241. Basset-Seguin N, Moles JP, Mils V, Dereure O, Guilhou JJ. TP53 tumor suppressor gene and skin carcinogenesis. J Invest Dermatol. 1994;103(5 Suppl):102S-6S.
242. Levine AJ, Momand J, Finlay CA. The p53 tumour suppressor gene. Nature. 1991;351(6326):453-6.
243. Cleaver JE. Defective repair replication of DNA in xeroderma pigmentosum. Nature. 1968;218(5142):652-6.
244. Kraemer KH, Lee MM, Andrews AD, Lambert WC. The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer. The xeroderma pigmentosum paradigm. Arch Dermatol. 1994;130(8):1018-21.
245. Cleaver JE. Cancer in xeroderma pigmentosum and related disorders of DNA repair. Nat Rev Cancer. 2005;5(7):564-73.
246. Bradford PT, Goldstein AM, Tamura D, Khan SG, Ueda T, Boyle J, et al. Cancer and neurologic degeneration in xeroderma pigmentosum: long term follow-up characterises the role of DNA repair. J Med Genet. 2011;48(3):168-76.
247. de Boer J, Hoeijmakers JH. Nucleotide excision repair and human syndromes. Carcinogenesis. 2000;21(3):453-60.
248. Fan L, Fuss JO, Cheng QJ, Arvai AS, Hammel M, Roberts VA, et al. XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. Cell. 2008;133(5):789-800.

249. Price VH, Odom RB, Ward WH, Jones FT. Trichothiodystrophy: Sulfur-Deficient Brittle Hair as a Marker for a Neuroectodermal Symptom Complex. JAMA Dermatology. 1980;116(12):1375-84.
250. Itin PH, Sarasin A, Pittelkow MR. Trichothiodystrophy: update on the sulfur-deficient brittle hair syndromes. J Am Acad Dermatol. 2001;44(6):891-920; quiz 1-4.
251. Hashimoto S, Egly JM. Trichothiodystrophy view from the molecular basis of DNA repair/transcription factor TFIIH. Human molecular genetics. 2009;18(R2):R224-30.
252. Nardo T, Oneda R, Spivak G, Vaz B, Mortier L, Thomas P, et al. 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. 2009;106(15):6209-14.
253. Winkler GS, Araujo SJ, Fiedler U, Vermeulen W, Coin F, Egly JM, et al. TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair. The Journal of biological chemistry. 2000;275(6):4258-66.
254. Botta E, Nardo T, Lehmann AR, Egly JM, Pedrini AM, Stefanini M. Reduced level of the repair/transcription factor TFIIH in trichothiodystrophy. Human molecular genetics. 2002;11(23):2919-28.
255. Nance MA, Berry SA. Cockayne syndrome: review of 140 cases. Am J Med Genet. 1992;42(1):68-84.
256. Troelstra C, van Gool A, de Wit J, Vermeulen W, Bootsma D, Hoeijmakers JHJ. ERCC6, a member of a subfamily of putative helicases, is involved in Cockayne's syndrome and preferential repair of active genes. Cell. 1992;71(6):939-53.
257. Kristensen U, Epanchintsev A, Rauschendorf MA, Laugel V, Stevnsner T, Bohr VA, et al. Regulatory interplay of Cockayne syndrome B ATPase and stress-response gene ATF3 following genotoxic stress. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(25):E2261-70.
258. Karikkineth AC, Scheibye-Knudsen M, Fivenson E, Croteau DL, Bohr VA. Cockayne syndrome: Clinical features, model systems and pathways. Ageing Res Rev. 2017;33:3-17.
259. Lehmann AR. DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. Biochimie. 2003;85(11):1101-11.
260. Berg OG, Winter RB, Vonhippel PH. Diffusion-Driven Mechanisms of Protein Translocation on Nucleic-Acids .1. Models and Theory. Biochemistry. 1981;20(24):6929-48.
261. Lohman TM. Kinetics of protein-nucleic acid interactions: use of salt effects to probe mechanisms of interaction. CRC Crit Rev Biochem. 1986;19(3):191-245.
262. Kanada R, Terakawa T, Kenzaki H, Takada S. Nucleosome Crowding in Chromatin Slows the Diffusion but Can Promote Target Search of Proteins. Biophys J. 2019;116(12):2285-95.

263. von Hippel PH, Berg OG. Facilitated target location in biological systems. The Journal of biological chemistry. 1989;264(2):675-8.
264. Riggs AD, Suzuki H, Bourgeois S. lac repressor-operator interaction: I. Equilibrium studies. Journal of molecular biology. 1970;48(1):67-83.
265. Adam G, Delbrück M. Reduction of dimensionality in biological diffusion processes. Structural chemistry and molecular biology. 1968;198:198-215.
266. Berg OG, Blomberg C. Association kinetics with coupled diffusional flows. Special application to the lac repressor--operator system. Biophysical chemistry. 1976;4(4):367-81.
267. Richter PH, Eigen M. Diffusion controlled reaction rates in spheroidal geometry. Application to repressor--operator association and membrane bound enzymes. Biophysical chemistry. 1974;2(3):255-63.
268. Bhattacherjee A, Levy Y. Search by proteins for their DNA target site: 1. The effect of DNA conformation on protein sliding. Nucleic Acids Research. 2014;42(20):12404-14.
269. Schurr JM. The one-dimensional diffusion coefficient of proteins absorbed on DNA. Biophysical chemistry. 1979;9(4):413-4.
270. Esadze A, Stivers JT. Facilitated Diffusion Mechanisms in DNA Base Excision Repair and Transcriptional Activation. Chem Rev. 2018;118(23):11298-323.
271. Berg OG, Winter RB, Von Hippel PH. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. Biochemistry. 1981;20(24):6929-48.
272. Fan J, Leroux-Coyau M, Savery NJ, Strick TR. Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. Nature. 2016.
273. Hughes CD, Simons M, Mackenzie CE, Van Houten B, Kad NM. Single molecule techniques in DNA repair: A primer. DNA repair. 2014;20:2-13.
274. Lu HP. Single-Molecule Enzymatic Dynamics. Science (New York, NY). 1998;282(5395):1877-82.
275. Wang MD, Yin H, Landick R, Gelles J, Block SM. Stretching DNA with optical tweezers. Biophys J. 1997;72(3):1335-46.
276. Gorman J, Greene EC. Visualizing one-dimensional diffusion of proteins along DNA. Nature Structural &Amp; Molecular Biology. 2008;15:768.
277. Dillingham MS, Spies M, Kowalczykowski SC. RecBCD enzyme is a bipolar DNA helicase. Nature. 2003;423(6942):893-7.
278. Mashanov GI, Tacon D, Knight AE, Peckham M, Molloy JE. Visualizing single molecules inside living cells using total internal reflection fluorescence microscopy. Methods. 2003;29(2):142-52.
279. Blainey PC, van Oijen AM, Banerjee A, Verdine GL, Xie XS. 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. 2006;103(15):5752-7.

280. Bindels DS, Haarbosch L, van Weeren L, Postma M, Wiese KE, Mastop M, et al. mScarlet: a bright monomeric red fluorescent protein for cellular imaging. Nat Methods. 2017;14(1):53-6.
281. Jaciuk M, Nowak E, Skowronek K, Tanska A, Nowotny M. Structure of UvrA nucleotide excision repair protein in complex with modified DNA. Nat Struct Mol Biol. 2011;18(2):191-7.
282. Green MR. Molecular cloning : a laboratory manual / Michael R. Green, Joseph Sambrook. Sambrook J, Cold Spring Harbor L, editors. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press; 2012.
283. Springall L, Inchingolo AV, Kad NM. DNA-Protein Interactions Studied Directly Using Single Molecule Fluorescence Imaging of Quantum Dot Tagged Proteins Moving on DNA Tightropes. Methods in molecular biology (Clifton, NJ). 2016;1431:141-50.
284. Kad NM, Trybus KM, Warshaw DM. Load and Pi control flux through the branched kinetic cycle of myosin V. The Journal of biological chemistry. 2008;283(25):17477-84.
285. Bers DM, Patton CW, Nuccitelli R. A practical guide to the preparation of Ca2+ buffers. Methods in cell biology. 99: Elsevier; 2010. p. 1-26.
286. Wang H, Tessmer I, Croteau DL, Erie DA, Van Houten B. Functional characterization and atomic force microscopy of a DNA repair protein conjugated to a quantum dot. Nano letters. 2008;8(6):1631-7.
287. Dunn AR, Kad NM, Nelson SR, Warshaw DM, Wallace SS. Single Qdot-labeled glycosylase molecules use a wedge amino acid to probe for lesions while scanning along DNA. Nucleic Acids Res. 2011;39(17):7487-98.
288. Tokunaga M, Imamoto N, Sakata-Sogawa K. Highly inclined thin illumination enables clear single-molecule imaging in cells. Nat Methods. 2008;5(2):159-61.
289. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-5.
290. Saxton MJ. Anomalous subdiffusion in fluorescence photobleaching recovery: a Monte Carlo study. Biophys J. 2001;81(4):2226-40.
291. Thiagalingam S, Grossman L. The multiple roles for ATP in the Escherichia coli UvrABC endonuclease-catalyzed incision reaction. J Biol Chem. 1993;268(24):18382-9.
292. Myles GM, Sancar A. Isolation and characterization of functional domains of UvrA. Biochemistry. 1991;30(16):3834-40.
293. Myles GM, Hearst JE, Sancar A. Site-specific mutagenesis of conserved residues within Walker A and B sequences of Escherichia coli UvrA protein. Biochemistry. 1991;30(16):3824-34.

294. Husain I, Van Houten B, Thomas DC, Sancar A. Sequences of Escherichia coli uvrA gene and protein reveal two potential ATP binding sites. Journal of Biological Chemistry. 1986;261(11):4895-901.
295. Doolittle R, Johnson M, Husain I, Van Houten B, Thomas D, Sancar A. Domainal evolution of a prokaryotic DNA repair protein and its relationship to active-transport proteins. Nature. 1986;323(6087):451.
296. Walker JE, Saraste M, Runswick MJ, Gay NJ. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. The EMBO journal. 1982;1(8):945-51.
297. Pakotiprapha D, Inuzuka Y, Bowman BR, Moolenaar GF, Goosen N, Jeruzalmi D, et al. Crystal Structure of Bacillus stearothermophilus UvrA Provides Insight into ATP-Modulated Dimerization, UvrB Interaction, and DNA Binding. Molecular cell. 2008;29(1):122-33.
298. Navaratnam S, Myles G, Strange R, Sancar A. Evidence from extended X-ray absorption fine structure and site-specific mutagenesis for zinc fingers in UvrA protein of Escherichia coli. Journal of Biological Chemistry. 1989;264(27):16067-71.
299. Visse R, De Ruijter M, Ubbink M, Brandsma J, Van de Putte P. The first zinc-binding domain of UvrA is not essential for UvrABC-mediated DNA excision repair. Mutation Research/DNA Repair. 1993;294(3):263-74.
300. Wang J, Mueller KL, Grossman L. A mutational study of the C-terminal zinc-finger motif of the Escherichia coli UvrA protein. Journal of Biological Chemistry. 1994;269(14):10771-5.
301. Croteau DL, DellaVecchia MJ, Wang H, Bienstock RJ, Melton Ma, Van Houten B. The C-terminal zinc finger of UvrA does not bind DNA directly but regulates damage-specific DNA binding. Journal of Biological Chemistry. 2006;281(36):26370-81.
302. Kraithong T, Channgam K, Itsathitphaisarn O, Tiensuwan M, Jeruzalmi D, Pakotiprapha D. Movement of the beta-hairpin in the third zinc-binding module of UvrA is required for DNA damage recognition. DNA repair. 2017;51:60-9.
303. Case BC, Hartley S, Osuga M, Jeruzalmi D, Hingorani MM. The ATPase mechanism of UvrA2 reveals the distinct roles of proximal and distal ATPase sites in nucleotide excision repair. Nucleic Acids Res. 2019;47(8):4136-52.
304. Clingen PH, Arlett CF, Roza L, Mori T, Nikaido O, Green MH. Induction of cyclobutane pyrimidine dimers, pyrimidine (6-4) pyrimidone photoproducts, and Dewar valence isomers by natural sunlight in normal human mononuclear cells. Cancer research. 1995;55(11):2245-8.

305. Gentry AC, Juul S, Veigaard C, Knudsen BR, Osheroff N. The geometry of DNA supercoils modulates the DNA cleavage activity of human topoisomerase I. Nucleic acids research. 2011;39(3):1014-22.
306. Douki T, Perdiz D, Grof P, Kuluncsics Z, Moustacchi E, Cadet J, et al. Oxidation of guanine in cellular DNA by solar UV radiation: biological role. Photochemistry and photobiology. 1999;70(2):184-90.
307. Thiagalingam S, Grossman L. The multiple roles for ATP in the Escherichia coli UvrABC endonuclease-catalyzed incision reaction. Journal of Biological Chemistry. 1993;268(24):18382-9.
308. Lamers MH, Winterwerp HH, Sixma TK. The alternating ATPase domains of MutS control DNA mismatch repair. EMBO J. 2003;22(3):746-56.
309. Ghisaidoobe AB, Chung SJ. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques. International journal of molecular sciences. 2014;15(12):22518-38.
310. Yokoi M, Masutani C, Maekawa T, Sugasawa K, Ohkuma Y, Hanaoka F. 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. 2000;275(13):9870-5.
311. Ranish JA, Hahn S, Lu Y, Eugene CY, Li X-j, Eng J, et al. Identification of TFB5, a new component of general transcription and DNA repair factor IIH. Nature genetics. 2004;36(7):707.
312. Fregoso M, Laine JP, Aguilar-Fuentes J, Mocquet V, Reynaud E, Coin F, et al. DNA repair and transcriptional deficiencies caused by mutations in the Drosophila p52 subunit of TFIIH generate developmental defects and chromosome fragility. Mol Cell Biol. 2007;27(10):3640-50.
313. Jawhari A, Laine JP, Dubaele S, Lamour V, Poterszman A, Coin F, et al. p52 Mediates XPB function within the transcription/repair factor TFIIH. The Journal of biological chemistry. 2002;277(35):31761-7.
314. Iyer N, Reagan MS, Wu KJ, Canagarajah B, Friedberg EC. Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. Biochemistry. 1996;35(7):2157-67.
315. Busso D, Keriel A, Sandrock B, Poterszman A, Gileadi O, Egly JM. Distinct regions of MAT1 regulate cdk7 kinase and TFIIH transcription activities. The Journal of biological chemistry. 2000;275(30):22815-23.
316. Rimel JK, Taatjes DJ. The essential and multifunctional TFIIH complex. Protein Sci. 2018;27(6):1018-37.

317. Gibbons BJ, Brignole EJ, Azubel M, Murakami K, Voss NR, Bushnell DA, et al. Subunit architecture of general transcription factor TFIIH. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(6):1949-54.
318. Humbert S, Vanvuuren H, Lutz Y, Hoeijmakers JHJ, Egly JM, Moncollin V. P44 and P34 Subunits of the Btf2/Tfiih Transcription Factor Have Homologies with Ssl1, a Yeast Protein Involved in DNA-Repair. Embo Journal. 1994;13(10):2393-8.
319. Warfield L, Luo J, Ranish J, Hahn S. Function of Conserved Topological Regions within the Saccharomyces cerevisiae Basal Transcription Factor TFIIH. Mol Cell Biol. 2016;36(19):2464-75.
320. Schmitt DR, Kuper J, Elias A, Kisker C. The structure of the TFIIH p34 subunit reveals a von Willebrand factor A like fold. PLoS One. 2014;9(7):e102389.
321. Radu L, Schoenwetter E, Braun C, Marcoux J, Koelmel W, Schmitt DR, et al. The intricate network between the p34 and p44 subunits is central to the activity of the transcription/DNA repair factor TFIIH. Nucleic Acids Res. 2017;45(18):10872-83.
322. Renda M, Baglivo I, Burgess-Beusse B, Esposito S, Fattorusso R, Felsenfeld G, et al. Critical dna binding interactions of the insulator protein ctcf a small number of zinc fingers mediate strong binding, and a single finger-dna interaction controls binding at imprinted loci. Journal of Biological Chemistry. 2007;282(46):33336-45.
323. Tremeau-Bravard A, Perez C, Egly JM. A role of the C-terminal part of p44 in the promoter escape activity of transcription factor IIH. The Journal of biological chemistry. 2001;276(29):27693-7.
324. Fribourg S, Kellenberger E, Rogniaux H, Poterszman A, Van Dorsselaer A, Thierry JC, et al. Structural characterization of the cysteine-rich domain of TFIIH p44 subunit. The Journal of biological chemistry. 2000;275(41):31963-71.
325. Maines S, Negritto MC, Wu X, Manthey GM, Bailis AM. Novel mutations in the RAD3 and SSL1 genes perturb genome stability by stimulating recombination between short repeats in Saccharomyces cerevisiae. Genetics. 1998;150(3):963-76.
326. Matsui P, Depaulo J, Buratowski S. An Interaction between the Tfb1 and Ssi1 Subunits of Yeast Tfiih Correlates with DNA-Repair Activity. Nucleic Acids Research. 1995;23(5):767-72.
327. Wang XW, Yeh H, Schaeffer L, Roy R, Moncollin V, Egly JM, et al. p53 modulation of TFIIH-associated nucleotide excision repair activity. Nat Genet. 1995;10(2):188-95.
328. Yoon H, Miller SP, Pabich E, Donahue TF. SSL1, a suppressor of a HIS4 5'-UTR stem-loop mutation, is essential for translation initiation and affects UV resistance in yeast. Genes & development. 1992;6(12b):2463-77.

329. Seroz T, Perez C, Bergmann E, Bradsher J, Egly JM. p44/SSL1, the regulatory subunit of the XPD/RAD3 helicase, plays a crucial role in the transcriptional activity of TFIIH. The Journal of biological chemistry. 2000;275(43):33260-6.
330. Fischer L, Gerard M, Chalut C, Lutz Y, Humbert S, Kanno M, et al. Cloning of the 62-kilodalton component of basic transcription factor BTF2. Science (New York, NY). 1992;257(5075):1392-5.
331. Gervais V, Lamour V, Jawhari A, Frindel F, Wasielewski E, Dubaele S, et al. TFIIH contains a PH domain involved in DNA nucleotide excision repair. Nat Struct Mol Biol. 2004;11(7):616-22.
332. Jawhari A, Boussert S, Lamour V, Atkinson RA, Kieffer B, Poch O, et al. Domain architecture of the p62 subunit from the human transcription/repair factor TFIIH deduced by limited proteolysis and mass spectrometry analysis. Biochemistry. 2004;43(45):14420-30.
333. Marinoni JC, Roy R, Vermeulen W, Miniou P, Lutz Y, Weeda G, et al. Cloning and characterization of p52, the fifth subunit of the core of the transcription/DNA repair factor TFIIH. The EMBO journal. 1997;16(5):1093-102.
334. Di Lello P, Jenkins LMM, Jones TN, Nguyen BD, Hara T, Yamaguchi H, et al. Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. Molecular cell. 2006;22(6):731-40.
335. Schilbach S, Hantsche M, Tegunov D, Dienemann C, Wigge C, Urlaub H, et al. Structures of transcription pre-initiation complex with TFIIH and Mediator. Nature. 2017;551(7679):204-9.
336. Leveillard T, Andera L, Bissonnette N, Schaeffer L, Bracco L, Egly J, et al. Functional interactions between p53 and the TFIIH complex are affected by tumour-associated mutations. The EMBO journal. 1996;15(7):1615-24.
337. Xiao H, Pearson A, Coulombe B, Truant R, Zhang S, Regier JL, et al. Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. Mol Cell Biol. 1994;14(10):7013-24.
338. Okuda M, Kinoshita M, Kakumu E, Sugasawa K, Nishimura Y. Structural Insight into the Mechanism of TFIIH Recognition by the Acidic String of the Nucleotide Excision Repair Factor XPC. Structure. 2015;23(10):1827-37.
339. Lafrance-Vanasse J, Arseneault G, Cappadocia L, Legault P, Omichinski JG. Structural and functional evidence that Rad4 competes with Rad2 for binding to the Tfb1 subunit of TFIIH in NER. Nucleic Acids Research. 2013;41(4):2736-45.
340. Korson L, Drost-Hansen W, Millero FJ. Viscosity of water at various temperatures. The Journal of Physical Chemistry. 1969;73(1):34-9.
341. Kad NM, Van Houten B. Dynamics of lesion processing by bacterial nucleotide excision repair proteins. Prog Mol Biol Transl Sci. 2012;110:1-24.

342. Fuss JO, Tainer JA. XPB and XPD helicases in TFIIH orchestrate DNA duplex opening and damage verification to coordinate repair with transcription and cell cycle via CAK kinase. DNA repair. 2011;10(7):697-713.
343. Rudolf J, Rouillon C, Schwarz-Linek U, White MF. The helicase XPD unwinds bubble structures and is not stalled by DNA lesions removed by the nucleotide excision repair pathway. Nucleic Acids Res. 2010;38(3):931-41.
344. Prost S, Kishen RE, Kluth DC, Bellamy CO. Working with commercially available quantum dots for immunofluorescence on tissue sections. PloS one. 2016;11(9):e0163856.
345. Dyson HJ, Wright PE. Coupling of folding and binding for unstructured proteins. Current opinion in structural biology. 2002;12(1):54-60.
346. Soppa J. Transcription initiation in Archaea: facts, factors and future aspects. Molecular microbiology. 1999;31(5):1295-305.
347. Lin J, Countryman P, Buncher N, Kaur P, E L, Zhang Y, et al. 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. 2014;42(4):2493-504.
348. Bonnet I, Desbiolles P. The diffusion constant of a labeled protein sliding along DNA. Eur Phys J E Soft Matter. 2011;34(3):25.
349. Gileadi O, Feaver W, Kornberg R. Cloning of a subunit of yeast RNA polymerase II transcription factor b and CTD kinase. Science. 1992;257(5075):1389-92.
350. Lamers MH, Winterwerp HH, Sixma TK. The alternating ATPase domains of MutS control DNA mismatch repair. The EMBO journal. 2003;22(3):746-56.
351. Castro J, Merino C, Zurita M. Molecular characterization and developmental expression of the TFIIH factor p62 gene from Drosophila melanogaster: effects on the UV light sensitivity of a p62 mutant fly. DNA repair. 2002;1(5):359-68.
352. Vermeulen W, Bergmann E, Auriol J, Rademakers S, Frit P, Appeldoorn E, et al. Sublimiting concentration of TFIIH transcription/DNA repair factor causes TTD-A trichothiodystrophy disorder. Nat Genet. 2000;26(3):307-13.
353. Yeh JI, Levine AS, Du S, Chinte U, Ghodke H, Wang H, et al. Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. Proceedings of the National Academy of Sciences. 2012;109(41):E2737-E46.
354. Rouillon C, White MF. The evolution and mechanisms of nucleotide excision repair proteins. Res Microbiol. 2011;162(1):19-26.