# The Architecture of Molecular Complexes Involved in UV-damage Recognition

Architectuur van moleculaire complexen betrokken bij de herkenning van UV-schade

**Proefschrift**

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus
Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
donderdag 5 juni 2008 om 11 uur

door

**Ana Janićijević**

geboren te Belgrado, Servie

ERASMUS UNIVERSITEIT ROTTERDAM

**Promotiecommissie**

**Promotor:**    Prof.dr. J.H.J. Hoeijmakers

**Overige leden:**  Dr. W. Vermeulen
            Dr. N. Goosen
            Dr. A.B. Houtsmuller

**Copromotor:**  Dr. C. Wyman

Za moje roditelje, za Mišu...

# Content

# Scope of the thesis

The work presented in this thesis aims to contribute to understanding of the molecular mechanisms that underlie UV-damage recognition. If not recognized and repaired, damage can result in DNA mutation, which might result in uncontrolled cell growth (cancer) or cell death. The biological relevance of UV-damage repair is emphasised by the severe clinical features associated with three rare autosomal-recessive inherited syndromes: xeroderma pigmentosum (XP), Cockayne syndrome (CS), and trichothiodystrophy (TTD). More specifically, it analyzes the architecture of the protein-DNA complexes formed during the repair reaction. To this purpose, individual protein-DNA complexes were visualized by scanning force microscopy (SFM) (Chapter 3 and 4).

Chapter 1 briefly introduces UV-damages and the relevance of UV-damage repair mechanisms. In this chapter we introduce some background important to understand nucleotide excision repair NER, highlighting the roles of every factor involved. In addition, a brief overview of the proteins that are required for damage detection in photoreactivation and NER in humans and *Escherichia coli* is given. In the Chapter 2 the application of SFM to study the conformation of protein complexes and their functional assemblies on DNA in DNA repair processes, imaging strategies and its challenges are presented.

Characterization of the damage recognition factor in NER, XPC-HR32B, binding to DNA with a single lesion is described in Chapter 3. Using trace trajectory method for measuring DNA bend angles it is shown that in the process of binding to DNA XPC-HR23B bends the DNA helix. In Chapter 4 the conformational change in the CPD-containing DNA upon photolyase binding is reported. We employ two methods to analyse DNA bending, the trace trajectory method and end-to-end distance method, and compare their results. Analyses show that photolyase bends DNA when bound at the CPD lesion as well as when bound non-specifically.

For SFM analysis of the repair reaction, we produced DNA substrates containing a single damage at a specific site. Appendix describes two protocols for production of these specific DNA substrates.

*Chapter 1*

Introduction

## 1. *UV-induced DNA damages and syndromes*

Each living orgamism needs to consume energy. On earth, solar radiation is the major driving force of life. However, solar light, particularly the ultraviolet (UV) part of the spectrum can be harmful to biological systems. Important biomolecules, deoxyribose nucleic acid (DNA) and proteins, can absorb energy of UV photons and can be altered (Kochevar 1990; Cadet, Sage et al. 2005; Pfeifer, You et al. 2005). DNA is the carrier of genetic information therefore encoding all essential processes in the cell. A minor change in DNA's chemical structure might alter genetic information and change normal cell processes. The most abundant UV-induced photoproducts are cyclobutane pyrimidine dimers (CPDs) and (6-4) photoproducts (PP). These lesions present a problem for the cell because they can block transcription and replication and cause mistakes if copied into mRNA or DNA.

Every organism has its strategy to deal with the UV radiation climate. Numerous repair systems exist that maintain DNA integrity. They are responsible for repairing DNA photoproducts allowing cells to proceed with their normal activities. If the DNA damages prevail they may block essential cellular processes such as DNA replication and transcription. If this happens, normal proliferating cells typically go into senescence, or die if the damage is present in a gene essential for metabolism. One mechanism to avoid UV damage is to alter the replication machinery so that damage can be bypassed (translesion synthesis, TLS). Most frequently, a specific translesion DNA polymerase incorporates the correct nucleotide, but may also incorporate an incorrect nucleotide opposite the lesion thus introducing a mutation (Lehmann 2005). Mutations in protein coding genes may cause altered cellular methabolism, eventually leading to cell death which may contribute to aging. In addition, when mutations change protein function this may activate oncogenes or inactivate tumor suppressor genes leading to cancerogenesis.

In mammals, nucleotide excision repair (NER) is the major DNA repair system for eliminating UV-induced DNA damage. The biological relevance of NER is emphasized by the severe consequences associated with (rare) human genetic syndromes based on inherited NER deffects: xeroderma pigmentosum (XP), Cockayne syndrome (CS) and trichothiodystrophy (TTD) (Bootsma, Kraemer et al. 2002; Lehmann 2003; Friedberg, Walker et al. 2005). XP patients show an elevated frequency (>1000 fold) of sunlight-induced skin cancer and other cutaneous abnormalities and in some cases display neurological abnormalities (Bootsma, Kraemer et al. 2002). Somatic cell genetics identified seven complementation groups (XPA-G) and the NER-proficient but TLS-deficient variant form, XPV (for review, see (Bootsma, Kraemer et al. 2002)). CS patients have no apparent increase in skin cancer, but display progressive developmental and neurological abnormalities (Nance and Berry 1992; Bootsma and Hoeijmakers 1993; Vermeulen, van Vuuren et al. 1994). TTD patients suffer from photosensitivity, mental retardation and skeletal abnormalities in addition to characteristic sulphur-deficient brittle hair and nails (Itin, Sarasin et al. 2001). The three disorders are associated with different clinical manifestations: several patient groups have clinical features that can largely be understood by defective NER, others suffer from a more complicated clinical profile that might result not

from NER deficiency, but also additional processes reflecting the multiple functions that some proteins implicated in NER have.

**UV-induced DNA damages**

The UV region of the electromagnetic spectrum is subdivided into three parts, each with different biological effects. These three wavelength ranges are UV-A (320-400 nm), UV-B (280 - 320nm) and UV-C (200 - 280nm). 95% of the total solar UV irradiation that reaches the earth surface is long wave UV-A. The major cytotoxic and genotoxic effect of UV-A on various cellular components is indirect. UV-A induces the production of reactive oxygen species (ROS) after absortion by biomolecules. Some of these reactive intermediates can react with other biomolecules, which might result in a change of function (Tyrrell and Keyse 1990; Fu, Jin et al. 2000). While damaged lipids and proteins can be resynthesised, ROS-induced changes of DNA can cause genetic abberations by introducing DNA breaks, abasic sites and DNA-protein cross-links (Sage 1993).

Only a small fraction of sunlight is composed of UV-B, which is responsible for most of the sun's pathological effects. DNA is considered its primary chromophore (Madronich, McKenzie et al. 1998). UV-B and UV-C cause the same types of DNA damages, mostly intrastrand DNA cross-links between neighboring pyrimidine bases (thymine or cytosine): cyclobutane pyrimidine dimers (CPD's) and 6-pyrimidine-4 pirimidone products (6-4)PP (Patrick and Gray 1976; Mitchell and Nairn 1989; Friedberg, Walker et al. 2005). The relative induction of these photoproducts depends on wavelength, DNA sequence, and protein-DNA interactions, but generally CPD's and (6-4)PP are formed in a ratio of about 3:1 (Mitchell and Nairn 1989). Solar UV-C rays are completely absorbed by the atmosphere and have minimal effect on the biosphere. Nevertheless, UV-C is widely used in the laboratory to induce damage and sterilization. UV-C (UVC) is often referred to as UVGI (ultraviolet germicidal irradiation).

CPDs and (6-4)PPs are the most frequently occurring forms of photodamage in cellular DNA. They arise when noncannonic covalent bonds are formed between neighboring pyrimidines on the same DNA strand. The CPD is the major photoproduct induced by UV light (Patrick and Rahn 1976; Cadet and Vigny. 1990). This damage arises when two adjacent pyrimidins become joined in a newly formed four-ring structure resulting from the saturation of their double bonds (C5-C6) (Figure 1). In theory CPDs can be found in 12 isomeric forms but only two (cis-syn and trans–syn) are generated in significant amounts after UV irradiation (Patrick and Gray 1976; Wang 1976). Cis-syn cyclobutane thymine dimers are the most common photoproducts in B form duplex DNA (Smith and Taylor 1993) whereas trans-syn dimers occur less frequently (at a rate of about 2% of the cis-syn isomer) primarily in single-stranded regions of DNA (Patrick and Rahn 1976; Taylor and Brockie 1988).

The other relevant photolesion, the (6-4)PP, is formed by covalent linkage between C6 of the 5'pyrimidine and C4 (carbonyl or imine group) of the 3' pyrimidine (Wang 1976; Rycyna and Alderfer 1985; Mitchell and Nairn 1989). In this case there is only one new covalent bond between adjacent pyrimidines (Figure 1). Subsequently, the (6-4)PP may also undergo a further UVB-dependent conversion to its valence photoisomer, reffered to as Dewar photoproducts
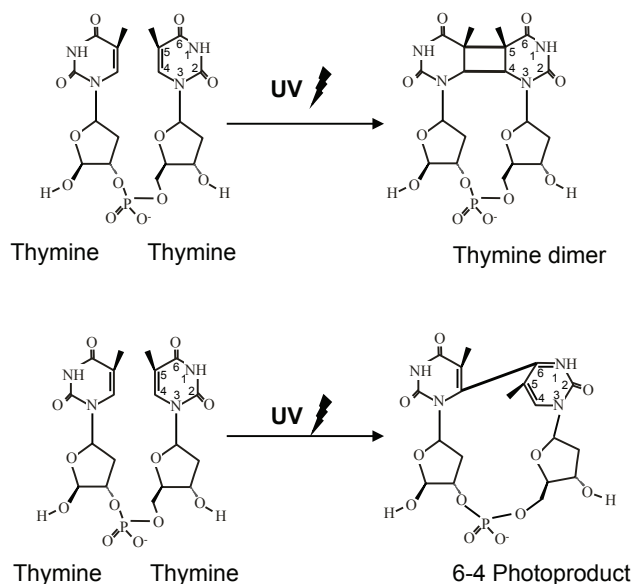
Figure 1. UV-induced DNA lesions A) Structure of cis-syn cyclobutane pyrimidine dimer (CPD) B) Structure of a (6-4) pyrimidone photoproduct ((6-4)PP).

(DewarPP) (Mitchell and Nairn 1989; Taylor, Lu et al. 1990; Matsunaga, Hatakeyama et al. 1993). In experiments with simulated solar light relatively high production of DewarPP was observed, suggesting that DewarPP may be a biologically relevant photolesion (Clingen, Arlett et al. 1995; Perdiz, Grof et al. 2000).

DNA is a dynamic and flexible molecule that changes its conformation by bending, twisting, unwinding and rewinding. These dynamic conformations of DNA are all altered by the presence of UV photoproducts. DNA helix axis bending is recognized by proteins or facilitates their binding. Measuring this feature is one way to detect changes in DNA structure. For instance, (6-4)PPs bend the DNA helix axis by about 44° (Kim and Choi 1995). The (6–4)PP lesion also makes the DNA backbone more flexible and causes slight unwinding of the duplex even at room temperature (Kim and Choi 1995; Wood 1999). The effects of CPDs on DNA conformation described in the literature is not consistent. Different methods and experimental conditions have been used to describe the sutructural influence of CPDs on DNA. Circularisation assay (ligase- mediated circlarisation of DNA), gel electrophoresis and nuclear magnetic resonance (NMR) techniques provide information about the average DNA bend angle in the population of molecules. The CPD-induced DNA bend was reported to be as large as 30° in circularization assay (Husain, Griffith et al. 1988), whilst gel electrophoretic studies show 7-9° (Wang and Taylor 1991; Wang and Taylor 1993; Kim, Patel et al. 1995).However, extensive

NMR studies report a variety of helix bend angles (Kemmink, Boelens et al. 1987; Kemmink, Boelens et al. 1987; Kan, Voituriez et al. 1988; McAteer, Jing et al. 1998; Lee, Park et al. 2004). One NMR solution study also indicated that the DNA helix is more flexible at a CPD site (Lee, Park et al. 2004). A recent crystallographic study of cis-syn CPDs revealed a bend of about 30° toward the minor groove and 9° unwinding twist in the DNA strand containing the lesion (Park, Zhang et al. 2002). Moreover, the hydrogen bonding between CPDs and opposite bases is weakend but still exist, while hydrogen bonding is totally disturbed in (6-4)PP. By definition, crystallography provides detailed three-dimensional structure of the solid state describing only one out of all possible conformations of the molecules and information about the flexibility of the DNA at the CPD site cannot be obtained. Compared to previous methods, single-molecule techniques provide quantitative information about individual molecules which is inaccessible in bulk experiments. Scanning force microscopy (SFM) and electron microscopy (EM) are, on the other hand, surface techniques, which can analyse single molecules and also provide information about the possible variability within a population. An EM study (complemented with two-dimensional electrophoresis analysis) suggested a bend of ~30° induced by CPDs (Husain, Griffith et al. 1988). The advantage of SFM sompared to EM is a simplified sample preparation without the need to stain the DNA. Using this techique we observed DNA bending of ~20° at the CPD site and no change in variability of the conformation compared to B-DNA (Chapter 4).

## 2. Removal of UV-induced damages

In order to prevent the severe biological consequences associated with light-induced damages in DNA, organisms have evolved different DNA repair mechanisms. In principle, three distinct processes deal wih UV-induced DNA lesions: direct chemical reversal, nucleotide excision repair (NER) and `alternative excision repair, which includes specific DNA endonuclease (designated as UVDE). The first two pathways will be described in further detail given the experimental focus of this thesis on these processes.

### 2.1. Direct reversal of base damages

Direct chemical reversal changes the modified bases without the need for new DNA synthesis, which is required for NER (see below). Direct reversal is usually performed by a single lesion-specific protein. The best known example of an enzyme that reverses DNA lesions is photolyase. Photolyase binds to one of the two types of UV-induced pyrimidine dimers (CPDs and (6-4) PPs) and uses the energy of blue or near-UV light to drive the cleavage of covalent bonds that link the two pyrimidine rings in a process known as photoreactivation. This repair mechanism is found in organisms from all three kingdoms of life, but surprisingly not in placental mammals (Yasui 1998). This pathway is not considered the most effective overall DNA repair mechanism, given the need for many specialized enzymes to deal with each of the various DNA lesions. Direct reversal is however fast and in essence, error-free.

Photolyases are monomeric proteins with a molecular weight between 50-60 kDa

(Sancar 1990). Every photolyase is highly specific for one type of lesion hence each photolyases exclusively repair either CPDs or (6-4)PP, but not both (Rupert 1975; Sancar 1990; Kim, Malhotra et al. 1994; Todo, Ryo et al. 1996). These proteins bind at the site of DNA damage, utilize the energy of visible light to convert the dimer back to its original structure and dissociate from the substrate (Weber 2005). Photolyases contain chromophores, which function as light-harvesting 'antennas'. The chromophore can be either methenyltetrahydrofolate (MTHF) or 8 hydroxy-5-deazariboflavin (8 HDF). The energy absorbed travels via an electron-transfer chain to the catalytic center. This catalytic center contains a redox active cofactor FAD (flavin adenine dinucleotide) which interacts directly with the damaged bases (Aubert, Mathis et al. 1999; Aubert, Vos et al. 2000; Byrdin, Eker et al. 2003).

## 2.2. Nucleotide excision repair

In contrast to direct reversal by lesion-specific photolyases, NER is a remarkably versatile repair process able to recognize and repair a variety of structurally and chemically unrelated DNA lesions (Friedberg, Walker et al. 2005). Within NER, DNA lesions are excised as part of an oligonucleotide. Excision of damaged DNA requires the synchronized action of many gene products. Damage is initially recognized by one or more protein factors, followed by opening of the DNA helix around the lesion. After DNA has been locally unwound specific endonucleases cut 3' and 5' of the lesion releasing the oligonucleotide containing the lesion from the helix. Repair is completed by general replication factors that fill in the gap using the intact DNA strand as a template to synthesize the removed oligonucleotide and finally the remaining nick is ligated (Gillet and Scharer 2006).

NER is present in a wide range of organisms and is one of the most highly conserved repair pathways. Enzymes that mediate this process appear in prokaryotes, eukaryotes (considered in this thesis) and archaea (Van Houten, Eisen et al. 2002) (Eisen and Hanawalt 1999).

## 3. *NER in mammals*

NER is an important DNA repair pathway in mammals that eliminates a wide variety of DNA alterations induced by UV light and chemical agents. In placental mammals NER is the only process known to remove UV-lesions. This pathway which involves a complex network of approximately 30 proteins, has been reconstituted in a cell-free system with highly purified or recombinant proteins and damaged DNA as a template (Aboussekhra, Biggerstaff et al. 1995; Guzder, Habraken et al. 1995; Araujo, Tirode et al. 2000). NER studies *in vitro* and in cultured cells suggest a strict order of assembly of NER factors on DNA lesions (Volker, Mone et al. 2001; Riedl, Hanaoka et al. 2003). Sequential complex assembly ensures that previous steps of the reaction were correct and generally serves as a control mechanism in this process. There are two distinct subpathways of NER that differ only in their damage recognition mechanism. Transcription-coupled NER (TC-NER) rapidly eliminates lesions from the coding strand of actively transcribed genes. Global-genome NER (GG-NER) removes UV-damage throughout
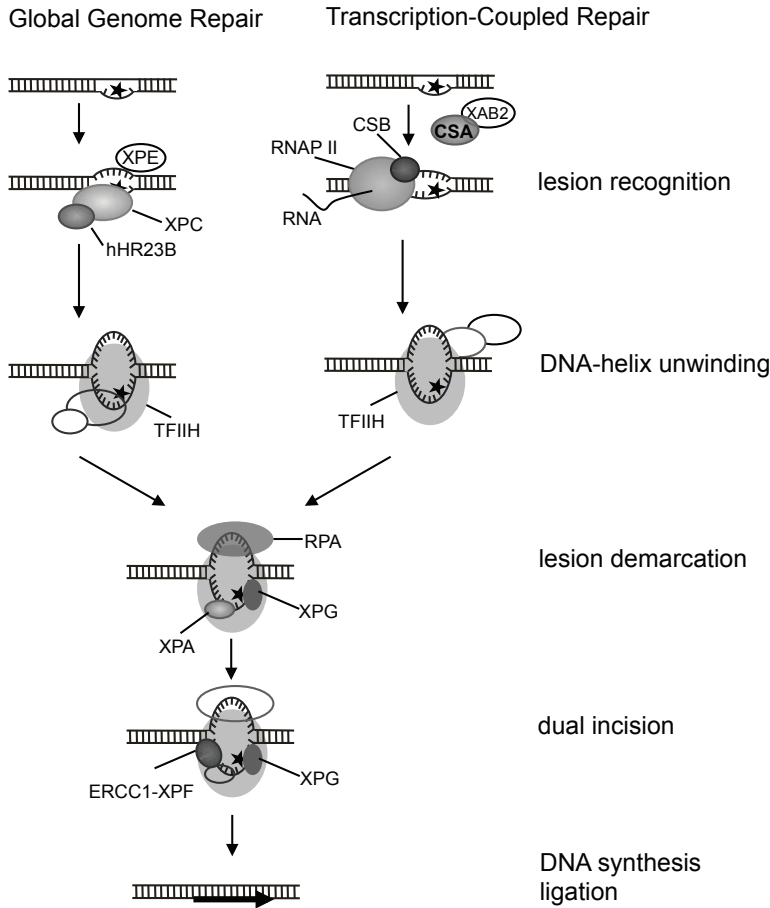
Figure 2. Model for global genome repair (GG-NER) and transcription-coupled repair (TC-NER) in humans. In GG-NER, XPC-hHR23B recognizes the helix distorting lesion (assisted by the XPE complex). Repair in the transcribed strand is initiated by stalling of an RNA Polymerase II elongation complex at a lesion. CSA, CSB and XAB2 are implicated in the initial stages of TC-NER. Upon detection of a lesion, the mechanism of the damage removal is similar in both pathways. TFIIH is required to open up the DNA helix and XPG is sequestered to stabilize this initial complex. Subsequently, XPA and RPA bind to the lesion. The presence of XPA will enable ERCC1-XPF binding to the damaged strand 5` to the lesion and activation of XPG. Two nucleases excise part of the damaged strand. Finally, the replication machinery performs gap-filling DNA synthesis and the nick remaining is sealed.

the entire genome and is therefore responsible for removing of the majority lesions. Removal of some types of damage by GG-NER, such as CPDs, is slower when compared to TC-NER (Mitchell and Nairn 1989; Hanawalt 1995). The efficiency and kinetics of UV-photoproduct repair depends not only on the type of lesion but also on its genomic localization, the cell type

and its developmental state (Sage 1993). After damage recognition, both GG-NER and TC-NER sequentially unwind the DNA helix in the region of the damage, excise the lesion within an approximately 30 nucleotide fragment and *de novo* synthesize DNA using the undamaged strand as a template. A simplified model of steps in NER is shown in Figure 2.

### 3.1. Recognizing lesions in DNA

Damage recognition proteins must be very efficient at distinguishing photolesions from the vast background of undamaged base pairs. In NER, a specific series of protein-DNA complexes need to assemble and change to achieve the specificity, control and accuracy needed during the process. GG-NER and TC-NER employ different sets of enzymes during damage recognition: in GG-NER, the XPC protein complex is the principal damage recognition factor, while stalled RNA polymerase II at a damage site is the recognition signal in TC-NER (Friedberg, Walker et al. 2005).

      XPC is part of a heterotrimeric protein complex that also includes HR23B (or HR23A) and centrin 2 (CEN2). In mammalian cell extracts, the 125 kDa XPC protein is found tightly associated with the 58 kDa HR23B (a mammalian homolog of yeast RAD23) or less frequently with HR23A (Masutani, Sugasawa et al. 1994; Sugasawa, Ng et al. 1997). In the reconstituted *in vitro* nucleotide excision repair reaction on naked DNA, XPC is fully functional, however addition of HR23B increases repair 2-4 fold (Sugasawa, Masutani et al. 1996; Sugasawa, Ng et al. 1997). *In vitro* both HR23B and HR23A can interact with XPC and stimulate its repair activity, suggesting some functional redundancy. In the cell HR23B is more abundant than XPC. The majority of HR23B is in a free form or in complex with other proteins (Sugasawa, Masutani et al. 1996; van der Spek, Eker et al. 1996; Hiyama, Yokoi et al. 1999). *In vivo* HR23B is required for GG-NER and stabilization of XPC and might therefore have a regulatory role in DNA damage repair (Ng, Vermeulen et al. 2003). Recently, another protein, centrin 2 was found to be an additional stabilizing factor for XPC *in vivo* (Araki, Masutani et al. 2001; Nishi, Okuda et al. 2005), but can be ommited from *in vitro* rections. Centrin 2 is one of the three human isoforms of centrins that are also found in centrosomes and has a role in cell cycle control (Salisbury, Suino et al. 2002). XPC is the NER-specific part of the above described complex. It binds DNA by itself and shows higher affinity for damaged DNA (Kusumoto, Masutani et al. 2001; Sugasawa, Okamoto et al. 2001). XPC recognizes a broad spectrum of structurally unrelated DNA lesions. Its function in damage recognition is described in more detail in Chapter 3 and Appendix. Very recently, the structure of the yeast XPC orthologue (Rad4) bound to CPD-containing DNA was determined by x-ray crystallography (Min and Pavletich 2007). The structure showed protein binding to the undamaged strand opposite the lesion, whereas the CPD itself is flipped out from the helix and is replaced by a β-hairpin of the Rad4. The protein does not interact with the damage per se, which might explain the wide substrate specificity of GG-NER.

      For some lesions, like CPDs, XPC has a low binding affinity (Hwang, Ford et al. 1999) even though they are repaired in an XPC-dependent manner. This indicates that recognition of this type of damage requires assistance of additional factors. The UV-DDB (XPE) complex

appears to play this additional role in CPD recognition (Wakasugi, Kawashima et al. 2002). Recently, it was shown that UV-DDB also plays a role in efficient repair of (6-4)PP (Moser, Volker et al. 2005). This complex is composed of two subunits, p125 and p48, also known as DDB1 and DDB2 (XPE). Besides being necessary for recognition of CPDs, UV-DDB may also have a role in detecting damage within chromatin (Datta, Bagchi et al. 2001; Martinez, Palhan et al. 2001; Rapic-Otrin, McLenigan et al. 2002). Recent evidence suggests that DDB1 and DDB2 are part of a larger complex containing ubiquitin ligase activity which includes cullin 4A, Roc1 and the signalosome complex (Groisman, Polanowska et al. 2003).

In TC-NER, damage is recognized when an elongating RNA polymerase II stalls upon encountering a lesion. At this stage two proteins, CSB and CSA, are required to assist in recruiting other NER factors and in displacing the stalled RNA polymerase to allow NER proteins access to the lesion. The exact role of these two proteins and their mechanism of action remain unclear. The amino-acid sequence of CSA includes five WD-40 repeats, which are implicated in protein–protein interaction. So far it is known that CSA is found in close contact with other NER proteins, including TFIIH, CSB (Henning, Li et al. 1995) and XAB2 (Nakatsu, Asahina et al. 2000). CSB belongs to the SWI/SNF family of DNA-dependent ATPases. Other SWI/SNF proteins are involved in chromatin remodeling during transcriptional activation of many genes. CSB was shown to act as a chromatin remodeler in biochemical assays(Citterio, Van Den Boom et al. 2000; Beerens, Hoeijmakers et al. 2005). CSB resides in the RNA pol-containing elongation complex and interacts with XPG (Iyer, Reagan et al. 1996; van den Boom, Citterio et al. 2004). CSA and CSB may also be needed after repair has been completed for TFIIH transition from NER function back to a transcription function with RNA Pol II elongation complex (van Oosterwijk, Filon et al. 1998; de Laat, Jaspers et al. 1999; Balajee and Bohr 2000; Fousteri, Vermeulen et al. 2006).

## 3.2. NER open complex formation

After recognition by either GG-NER or TC-NER machinery, the damaged site is processed in the same manner. An opening of about 30 nucleotides around the leasion is needed prior to dual incision where opening requires recruitment of NER factors; TFIIH, XPA, RPA and XPG.

**TFIIH** is likely to be the next NER factor that arrivs at the damage site after the lesion has been identified by XPC-hHR23B or RNA polymerase II (Yokoi, Masutani et al. 2000; Volker, Mone et al. 2001). TFIIH is a general transcription initiation factor but is also engaged in NER and probably cell cycle regulation and apoptosis (Bhatia, Wang et al. 1996; de Laat, Jaspers et al. 1999). Its function in NER is to unpair bases in the DNA helix around the lesion and to recruit other NER proteins. TFIIH is a large complex consisting of ten subunits (Giglia-Mari, Coin et al. 2004). In humans, TFIIH is organized into two distinct structural and functional units: core-TFIIH is composed of XPB, p62, p52, p44 and p34 proteins; and TFIIH-assotiated trimeric CAK complex is composed of cdk7, cyclin H and MAT1. The XPD subunit can be found associated with either the core-TFIIH or CAK (Drapkin, Le Roy et al. 1996; Reardon, Ge et al. 1996; Rossignol, Kolb-Cheynel et al. 1997) and links these two subcomplexes, thus stabilizing TFIIH (Coin, Bergmann et al. 1999). Electron microscopy analysis showed that core-TFIIH has

a ring-like structure with CAK attached as a knob-like feature (Schultz, Fribourg et al. 2000). Recently a new subunit of TFIIH, p8, has been identified (Giglia-Mari, Coin et al. 2004) that plays an important role in stabilizing the TFIIH complex.

TFIIH possesses several enzymatic activities and also has a general role in organizing the assembly of NER intermediates prior to incision. TFIIH associated enzymatic activities include XPD and XPB, DNA-dependent ATPases which display helicase activity with opposite directions (5'-3' and 3'-5', respectively) (Schaeffer, Roy et al. 1993) (Sung, Bailly et al. 1993; Ma, Siemssen et al. 1994; Schaeffer, Moncollin et al. 1994). DNA binding motifs are found in the p34 and p44 subunits (Humbert, van Vuuren et al. 1994). The N-terminal part of p44 is important for NER since it regulates the helicase activity of XPD (Coin, Marinoni et al. 1998) (Seroz, Perez et al. 2000), while the C-terminal part is important for transcription given its participation in promoter escape (Tremeau-Bravard, Perez et al. 2001). Recently, ubiquitin ligase activity was linked to the p44 subunit of TFIIH (Takagi, Masuda et al. 2005). The p52 and p62 subunits are essential for recruiting NER endonucleases XPF and XPG, respectively to the incision complex (Marinoni, Roy et al. 1997; Jawhari, Laine et al. 2002; Gervais, Lamour et al. 2004). Whilst the CAK complex is dispensable for NER activity *in vitro* (Svejstrup, Wang et al. 1995; Mu, Hsu et al. 1996) it is however, essential in basal transcription because CDK7 is the serine/threonine kinase that phosphorylates RNA polymerase II (Akoulitchev, Makela et al. 1995; Makela, Parvin et al. 1995; Fisher 2005). CAK is also linked to the cell cycle via cyclin-activating kinase complex (Fisher 2005).

One role of TFIIH in NER is to produce single-stranded DNA around the lesion. Two of the TFIIH subunits, XPB and XPD, are helicases vital for NER. They are ATP-dependent helicases with opposite polarity which work synergistically to unwind DNA around the damage and produce substrates for structure-specific nucleases. XPB has 3'-5' helicase activity that is essential for both repair and transcription, whereas the XPD's 5'-3' helicase activity is required only for repair and is not required for *in vitro* basal transcription. XPB was found to interact with one of the endonucleases, XPG (Iyer, Reagan et al. 1996; Dunand-Sauthier, Hohl et al. 2005).

**XPA** is a small, hydrophilic protein of 31 kDa that binds UV-damaged DNA with greater affinity than non-damaged DNA *in vitro* (Asahina, Kuraoka et al. 1994) (Jones and Wood 1993). Two structural features of XPA are important for repair: a zinc finger motif and a glutamic acid cluster (Nagai, Saijo et al. 1995; Morikawa and Shirakawa 2000) (Li, Peterson et al. 1995; Kuraoka, Morita et al. 1996; Buchko, Ni et al. 1998; Ikegami, Kuraoka et al. 1998). XPA deficient cells do not show any NER activity (Koberle, Roginskaya et al. 2006) indicating that it is required for both GG-NER and TC-NER.

XPA interacts with other NER factors, including TFIIH binding (DNA-dependent) (Nocentini, Coin et al. 1997), anchors ERCC1-XPF to DNA damage (Li, Elledge et al. 1994; Nagai, Saijo et al. 1995) and may be necessary for the endonuclease activity of XPG. Purified XPA forms dimers in solution. Together with RPA, the XPA$_2$-RPA complex shows enhanced affinity for damaged DNA (Yang, Liu et al. 2002) however, this XPA-RPA interaction has not been detected in diffusion experiments with labeled proteins *in vivo* (Rademakers, Volker et al. 2003). The precise role of XPA in NER remains unclear. It stabilizes an open DNA complex,

but may also function in double checking for the presence of damage, alone or in complex with RPA. In addition, two novel XPA interacting partners, XPA binding protein 1 (XAB1) and the XPA binding protein 2 (XAB2), have been identified, although their functions in NER have not been established. XAB1 is a cytoplasmic GTP-activating protein (Nitta, Saijo et al. 2000) and seems to be important for XPA translocation to the nucleus. XAB2 interacts with TC-NER-specific proteins CSA, CSB proteins, as well as RNA pol II *in vivo* and may be a component for both repair and transcription (Nakatsu, Asahina et al. 2000).

**RPA** (replication protein A) is a single-strand DNA binding protein with multiple functions in the cell. It was first described as a factor essential for DNA replication, but it is now known to be essential for many processes including NER (Patrick, Oakley et al. 2005; Fanning, Klimovich et al. 2006). RPA is a trimeric protein composed of tightly associated p70, p34 and p14 kDa subunits**.** It has been shown *in vitro* that RPA covers around 30 nucleotides of single-stranded DNA (Kim, Snyder et al. 1992; Kim, Paulus et al. 1994) which is the length of DNA melted in the fully open NER intermediate. This might indicate that only one RPA binds to the DNA strand across from the lesion in the open complex (Bochkarev, Pfuetzner et al. 1997; de Laat, Appeldoorn et al. 1998; Schweizer, Hey et al. 1999). In addition, RPA shows protein-protein interactions with many other NER proteins. RPA was found to interact with XPA via p70 and p34 subunits *in vitro* (He, Henricksen et al. 1995; Li, Lu et al. 1995; Matsuda, Saijo et al. 1995). It is also important for positioning the two endonucleases: the side of the RPA facing a 3'-end of duplex-ssDNA junction binds ERCC1-XPF, while the other side of RPA facing the 5'-end of the open complex stably binds XPG (Matsunaga, Park et al. 1996; de Laat, Appeldoorn et al. 1998).

### 3.3. Damage excision

Excision of the damaged segment of DNA is performed by the concerted action of, XPG and ERCC1–XPF, which are positioned at the borders of the fully open DNA intermediate.

**XPG** is a member of the FEN-1 family of structure-specific nucleases. It cleaves branched DNA at the 3'side of the damage (O'Donovan, Davies et al. 1994; Matsunaga, Park et al. 1996). XPG also has noncatalytic roles in repair. Its presence is needed for open complex formation and for the catalytic activity of other endonucleases (Wakasugi, Reardon et al. 1997; Constantinou, Gunz et al. 1999). XPG has additional roles in RNA pol II transcription and base excision repair (Sarker, Tsutakawa et al. 2005).

The **XPF-ERCC1** protein complex is the other structure-specific endonuclease in NER. XPF-ERCC1 cuts the damaged DNA strand 5' to the lesion (Matsunaga, Mu et al. 1995; Sijbers, de Laat et al. 1996). It does not bind damaged DNA by itself, but is engaged in NER via interaction with XPA and RPA (through ERCC1 and XPF, respectively) (Li, Elledge et al. 1994; Li, Lu et al. 1995; Park, Bessho et al. 1995). Besides acting in NER, XPF-ERCC1 is involved in cross-link repair (Sijbers, de Laat et al. 1996) and homologous recombination (Motycka, Bessho et al. 2004; Niedernhofer, Odijk et al. 2004). Recruitment and function of both endonucleases is dependent on prior assembly of XPA and RPA at the damage site (Li, Peterson et al. 1995; Volker, Mone et al. 2001; Mu, Hsu et al. 1996; Bessho, Sancar et al.

1997; de Laat, Appeldoorn et al. 1998).

Incisions only occur when the enzymes are properly orientated. The endonucleases cut the damaged strand asymmetrically regarding the lesion. XPG cuts at the 3′ side, 0-2 nucleotides away from the ssDNA-dsDNA junction of the open DNA intermediate, which is about 2-8 nucleotides away from the lesion. On the other side, ERCC1–XPF carries out the 5′-incision, approximately 15–24 nucleotides away from the lesion. This variation in incision locations may depend on the type of lesion. The excised piece of DNA dissociates leaving a gap behind. This might also be the moment when most of the NER proteins dissociate from DNA. Little is known about this disassembly step though it is assumed to allow subsequent repair synthesis as well as recycling NER machinery.

### 3.4. DNA repair synthesis

Excision of the oligo with a lesion leaves a gap in the DNA strand. The exposed 3′- hydroxyl(OH)-end is a natural primer for DNA polymerases, and it does not need any extra modification prior to DNA synthesis. Information needed to restore the original DNA sequence is present in the opposite strand, to be used as the template. Either of two DNA polymerases Pol δ or Pol ε can carry out the repair patch synthesis in NER. These polymerases need accessory factors: proliferating cell nuclear antigen (PCNA) and replication factor C (RFC) (Shivji, Kenny et al. 1992; Aboussekhra, Biggerstaff et al. 1995). RFC binds first at the free 3′-end of the DNA and facilitates loading of PCNA. When the complete DNA information is restored, eighter DNA ligase 1 or ligase III/XRCC1 seals the remaining nick at the 5′-end of the newly synthesized sequence (Barnes, Tomkinson et al. 1992; Aboussekhra, Biggerstaff et al. 1995; Moser, Kool et al. 2007).

## 4. *Damage detection*

Recognition of UV-induced DNA lesions represents the first crucial step in repair mechanisms. The repair systems recognize either the chemical structure of DNA photoproduct directly or a change in the conformation of the DNA helix due to the presence of damage. Many studies have investigated DNA-protein interactions and changes in DNA conformation when bound by damage recognition proteins. In this section, the characteristics of repair factors described above will serve as a basis to build up models for damage recognition in photoreactivation, bacterial and human nucleotide excision repair.

### 4.1. Damage detection by photolyases

Photolyases are DNA repair enzymes that rapidly restore UV lesions back to the original undamaged bases in a simple enzymatic reaction using energy from the blue or near-UV part of the spectrum. They are examples of a single protein recognizing and repairing a single

damage type.

Various biological and spectroscopic experiments have been carried out to understand how photolyases recognize damage. The structure of the photolyase-substrate complex is critically important as this can explain the binding and repair mechanism of this enzyme. The three-dimensional structure for three photolyases (*Escherichia coli*, *Anacystis nidulans* and *Thermus thermophilus*) have been determined by X-ray analysis (Tamada, Kitadokoro et al. 1997; Komori, Masui et al. 2001; Park, Zhang et al. 2002). The crystal structure of the protein revealed a domain rich in basic residues on its surface where DNA could dock. In the middle of that positively charged region is a cavity housing FAD cofactor, which is responsible for catalytic redox activity. The catalytic cavity of the photolyase is a hydrophobic environment suitable for binding the DNA bases. The cavity is large enough to accommodate the damaged bases when they are flipped out of the helix. Recently, the crystal structure of a photolyase- CPD-DNA complex was determined, providing detailed information on DNA-protein interactions (Mees, Klar et al. 2004). In this study, binding of the photolyase was accompanied by bending of the DNA helix by ~50° and, as suggested earlier, flipping out the CPD from the helix into the catalytic cavity containing the FAD cofactor (Figure. 3). Using scanning force microscopy (SFM, also known as atomic force microscopy, AFM) we showed that photolyase bends DNA by about 40° when bound to damage but also to un-damaged DNA sites (Chapter 4). From the functional point of view this conformational change might be needed to present the damage to the active
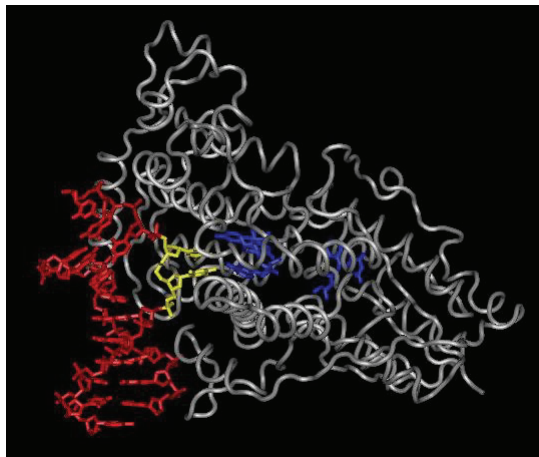


Figure 3. Alpha carbon backbone model of *Anacystis nidulans* photolyase in complex with CPD-DNA as resolved by X-ray crystallography (structure in PubMed: 1TEZ). The protein is shown in light gray and the DNA is presented in red. The CPD (highlighted with yellow) that is flipped out from the helix is inside the protein (in contrast with the NER proteins). Non-covalently bound cofactors (FAD and 8-HDF) are shown in blue.

center of the protein. However, an earlier SFM study (van Noort, Orsini et al. 1999) reported that photolyase does not bend DNA when bound to an undamaged substrate. Comparison of previous studies with ours and possible explanations for the observed discrepancies are described in Chapter 4.

Photolyase can bind to undamaged DNA and apparently slide along the duplex (van Noort, van der Werf et al. 1998). At the damaged site, the protein has several interactions with

the DNA bases flanking the lesion (Mees, Klar et al. 2004). The majority of the interactions are short-range and non-specific (including hydrogen bonding and van-der Waals forces). At present there is no detailed high-resolution structure of an undamaged DNA-photolyase complex, but nevertheless we hypothesize that similar interactions are present when photolyase binds undamaged double-stranded DNA. The presence of a CPD with a local kink in the DNA helix, weakened stacking interactions with adjacent bases, weakened hydrogen bonding and altered geometry (Park, Zhang et al. 2002) would permit formation of additional interactions that significantly stabilize the bent state. Photolyase bends the dimer out of the DNA helix whereby a number of additional interactions that contribute to substrate binding and specificity occur (Figure 3). In the absence of the base dimer, normal hydrogen bonding in the DNA helix strongly opposes moving residues into the catalytic cavity of the protein.

## 4.2. Damage detection in *Escherichia coli* NER

NER in prokaryotes is a multistep and complex process. Compared to photoreactivation, damage recognition and strand incision are mediated by three proteins: UvrA, UvrB and UvrC. In the first step, complex between damaged DNA, UvrA and UvrB is formed. After the damage identification, UvrA dissociates whereas UvrB remains bound to DNA and forms a stable preincision complex. In the last step, UvrC binds to this complex and triggers the incision on both sides of the damaged region.

UvrB is a pivotal protein in *E. coli* NER which is required for damage recognition, strand excision and repair synthesis and interacts with all components of excision repair, i.e. UvrA, UvrC, UvrD (helicase II) and DNA polymerase I (Sancar and Sancar 1988; Orren, Selby et al. 1992). UvrB shares high structural homology with the helicases PcrA, NS3, and Rep (Bird, Subramanya et al. 1998) but has little helicase and translocating activity itself. The most prominent feature of UvrB is a highly conserved and flexible β-hairpin, which is not observed in other helicases. This element is rich in aromatic and hydrophobic residues and it was proposed that this motif is inserted between DNA strands (Theis, Chen et al. 1999). Based on similarity with the crystal structures of a related DNA-helicase co-complex (Machius, Henry et al. 1999; Nakagawa, Sugahara et al. 1999; Theis, Chen et al. 1999), the DNA path along UvrB was modeled close to the beta-hairpin where there is a positively charged patch on the protein. This domain is also conserved between different organisms (Theis, Chen et al. 1999). Damage recognition in the bacterial NER pathway requires both UvrA and UvrB. For a long time it was thought that a trimeric complex composed of two UvrA molecules and a single UvrB molecule formed the recognition unit. Measuring the size of active UvrB bound to damage by SFM showed that actually there are two UvrBs present in the preincision complex (Verhoeven, Wyman et al. 2002).

Solving the crystal structure of UvrB with a molecule of single-stranded DNA was a big step forward in understanding the mechanism of damage recognition (Waters, Eryilmaz et al. 2006). As opposed to photolyase, UvrB does not have an active site 'pocket' to accommodate the lesion but instead a flexible β-hairpin (Moolenaar, Hoglund et al. 2001; Skorvaga, Theis et
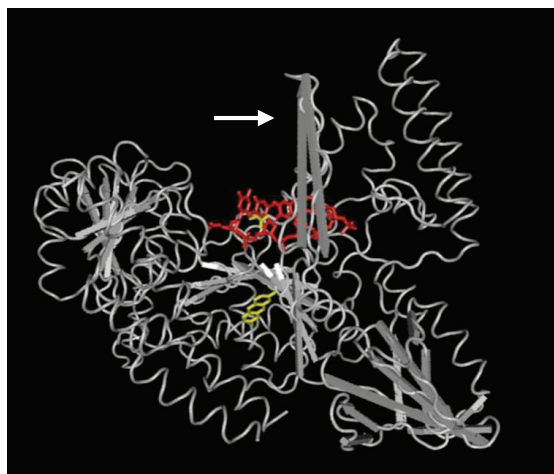
Figure 4. Alpha carbon backbone model of UvrB (*Bacillus subtilis*) in complex with DNA containing a single fluorescein-adducted thymine as resolved by x-ray crystallography (structure in PubMed: 2NMV). The protein is shown in gray and DNA is indicated in red. The extra-helical position of the lesion (highlighted in yellow) indicates that the damage is excluded away from the UvrB molecule into the solvent by the β- hairpin (indicated with arrow). The lesion is T-fluorescein, attached through a linker (which is absent from the density map) to a thymine base.

al. 2002; Malta, Moolenaar et al. 2006; Waters, Eryilmaz et al. 2006). This structural element contains four vital tyrosines, Y92/93 and Y95/96. In the current model of damage recognition the damaged DNA is originally bound to an $UvrA_2B_2$ complex. It is proposed that ATP is required in this step to promote formation of a complex with DNA wrapped around one of the UvrB (Verhoeven, Wyman et al. 2001). The DNA wrap would partially open the helix and allow the β-hairpin of UvrB to probe for damage (Moolenaar, Hoglund et al. 2001). Recent x-ray analysis of a UvrB-DNA complex suggests a model where the protein translocates the DNA strand behind the β-hairpin until it clashes into the damage (Truglio, Karakas et al. 2006; Waters, Eryilmaz et al. 2006) and halts. The arrest causes extrusion of the lesion from the β-hairpin to the area of the protein exposed to solvent (Figure 4). Nucleotides adjacent to the lesion make stabilizing interactions with aromatic amino-acids at the base of the hairpin by hydrophobic stacking (Moolenaar, Hoglund et al. 2001; Waters, Eryilmaz et al. 2006). After damage recognition, UvrA dissociates resulting in a UvrB-DNA complex which is very stable, surviving *in vitro* even in solutions of high ionic strength (Orren and Sancar 1989). The possible role of UvrA in damage recognition is not well understood. It was suggested that UvrA catalyzes the slow process of docking UvrB on DNA (Verhoeven, Wyman et al. 2002; Truglio, Karakas et al. 2006).

Damage recognition in NER in both prokaryotes and eukaryotes differs from photoreactivation in two aspects: lesion recognition involves multiple steps and a broad substrate range. The fact that the different types of damage can be recognized by the same proteins suggests that repair proteins recognize damaged regions based on an altered conformation from the normal DNA helix rather than the altered chemical structure of the DNA photoproduct.

### 4.3. DNA damage recognition in NER in humans

The general process of NER in human cells resembles that in *E. coli*, but there are many enzymatic differences in the details. The most striking difference is the greater number of proteins participating in each step in eukaryotic NER (Eisen and Hanawalt 1999). In human cells there are two distinct subpathways, TC-NER and GG-NER, which show a mechanistic overlap differing only in damage recognition (Friedberg, Walker et al. 2005).

### 4.3.1. Damage recognition in TC-NER

TC-NER has evolved to rapidly repair lesions that obstruct RNA polymerase II elongation from actively transcribed DNA regions. Damage encountered during transcription causes RNA pol II to stall. A stalled RNA pol II signals the TC-NER machinery to initiate repair (Hanawalt 1994) (Tremeau-Bravard, Riedl et al. 2004). The partially unwound region produced by the stalled polymerase is believed to provide access to TFIIH. The fact that the stalled polymerase produces a partially unwound region on its own may be one reason why XPC is not required in human TC-NER (Fousteri, Vermeulen et al. 2006).

Two proteins, CSA and CSB, are specifically required for TC-NER. They work together to couple RNA pol II transcription arrest at the damaged sites to the unwinding event in NER (van den Boom, Jaspers et al. 2002; Svejstrup 2003). CSA can interact with XAB2, CSB and the p44 subunit of TFIIH *in vitro* (Henning, Li et al. 1995; Nakatsu, Asahina et al. 2000) but does not directly bind to RNA pol II (Tantin 1998). *In vivo* however, CSB interacts with RNA pol II (van Gool, Citterio et al. 1997) and was recently shown to inhibit RNA pol II from backing up when it reaches a lesion (Tremeau-Bravard, Riedl et al. 2004). In the course of NER the role of CSA and CSB are to assist in displacing RNA pol II and the nascent RNA strand from the damaged region and subsequently to recruit other NER proteins. Live cell studies show that CSB is more stably bound to RNA pol II after damage recognition. Recent models of the TC-NER pathway (Sarasin and Stary 2007) propose that CSB plays a critical role since it interacts with RNA pol II and appears to be crucial for the further recruitment of the other 'core-NER' factors (all NER factors except XPC-HR23B, centrin 2 and UV-DDB) to the damage site (Fousteri, Vermeulen et al. 2006).

### 4.3.2. Damage recognition in GG-NER

In human global genome NER XPC plays a key role in damage recognition. XPC–HR23B can initiate the GG-NER process *in vitro* (Sugasawa, Ng et al. 1998). In kinetic repair assays the XPC complex is the initial protein needed in the GG-NER reaction. Evidence that XPC is the damage recognition protein *in vivo* came from studies using fluorescently tagged antibodies against various NER proteins to determine which associate with regions of DNA damage in cells lacking other NER proteins (Volker, Mone et al. 2001). XPC-HR23B was found to be the first NER factor arriving at the damaged region. Others have shown that DDB2 is even faster and has a role in CPD and also (6-4)PP recognition (Luijsterburg, Goedhart et al. 2007).

*In vitro*, XPC-HR23B has a strong binding affinity for various structurally unrelated DNA substrates. It can specifically bind small defined DNA lesions including 6-4 PP but also bulky DNA modifications which do not have a common chemical structure (including cisplatin, acetylaminofluorene (AAF) adduct, cholesterol moiety and benzo[a]pyrenyl DNA induced lesions (Sugasawa, Ng et al. 1998; Batty, Rapic'-Otrin et al. 2000; Kusumoto, Masutani et al. 2001; Sugasawa, Okamoto et al. 2001; Janicijevic, Sugasawa et al. 2003; Mocquet, Kropachev et al. 2007)). The affinity of XPC-HR23B for a CPD is not much higher than for undamaged DNA. However, if a mismatch is introduced opposite the damage, the affinity of XPC for CPDs is significantly increased (Batty, Rapic'-Otrin et al. 2000; Sugasawa, Okamoto et al. 2001). XPC-HR23B showed significant affinity for small 'bubbles' in DNA structure and
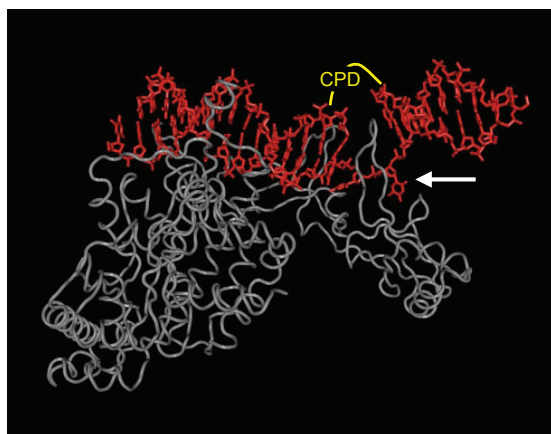


Figure 5. Alpha carbon backbone model of Rad4-Rad23 (*Saccharomyces cerevisiae*) in complex with DNA containing a CPD as resolved by X-ray crystallography (structure in PubMed: 2NMV). The protein is shown in gray, DNA is indicated in red, the position of the CPD is schematically presented in yellow. The structure shows that the position of the CPD is occupied by a β- hairpin (indicated with arrow) of Rad4 which displaces the lesion into the solvent area. The position of the CPD is disordered in the crystal structure and is, most probably, extrahelical.

some specific secondary structures of DNA, particularly single– and double-strand junctions (Sugasawa, Shimizu et al. 2002). Moreover, XPC-HR23B preferentially binds single-stranded DNA over double-stranded DNA (Sugasawa, Shimizu et al. 2002) and can bind to small bubbled structures (3' to 5' nucleotide bubble) with and without damage (Sugasawa, Okamoto et al. 2001). Based on competition experiments, it was assumed that the single-stranded character of the DNA is not the feature that XPC-HR23B recognizes (Masutani, Sugasawa et al. 1994; Shivji, Eker et al. 1994; Reardon, Ge et al. 1996; Batty, Rapic'-Otrin et al. 2000), but recently this has been challenged. There is now a large body of data showing that XPC binds to the undamaged strand opposite the lesion including: biochemical and functional analysis of XPC mutants (Bunick, Miller et al. 2006; Maillard, Solyom et al. 2007), reconstituted NER using synthetic DNA substrates and cell extracts (Buterin, Meyer et al. 2005) and very recently reported crystal structure determination of the *Saccharomyces cerevisiae* XPC ortholog, Rad4, bound to damaged DNA (Min and Pavletich 2007). The crystal structure showed that the protein makes many contacts with DNA in the region close to the CPD, but not directly to the damaged nucleotides. Analogous to bacterial damage recognition by UvrB, Rad4 also has a β-hairpin that is inserted through the DNA helix, whereas the lesion itself is entirely flipped

out from the helix (Figure 5). These findings strongly support the model in which XPC does not recognize the chemical composition of damaged bases, rather a distortion of the Watson-Crick helix induced by lesions and/or unpaired bases (Batty, Rapic'-Otrin et al. 2000; Sugasawa, Shimizu et al. 2002). The ability of XPC to bind such apparently diverse structures may be a great advantage for NER to recognize and remove different lesions. XPC binds asymmetrically (3` side with respect to the distorted site), as demonstrated with DNase I footprint analysis (Sugasawa, Shimizu et al. 2002) and in the crystal structure of the Rad4-CPD-DNA complex (Min and Pavletich 2007). Asymmetry of XPC binding might be projected later in the asymmetric incisions around the lesion. Upon binding, XPC-HR23B bends DNA (Janicijevic, Sugasawa et al. 2003), creating a specific DNA-protein architecture that can be recognized by other NER factors.

XPC can induce conformational changes in the DNA helix. In the process of binding XPC unwinds a small DNA region around the damage (Tapias, Auriol et al. 2004). *In vitro*, when special substrates are used which contain lesions within a small bubble or next to a small bubble of melted DNA, excision can be performed without XPC-HR23B (Mu and Sancar 1997; Sugasawa, Ng et al. 1997; Sugasawa, Shimizu et al. 2002). This suggests that unpaired DNA may represent a naturally occurring intermediate, which is the result of interaction between DNA and the damage recognition proteins.

The accumulated data indicate that XPC bound to the damage is recognized by **TFIIH**. Strong co-operativity of these two proteins was demonstrated in co-immunoprecipitation experiments (Yokoi, Masutani et al. 2000; Araujo, Nigg et al. 2001) and recently with permanganate footprinting analysis (Tapias, Auriol et al. 2004). The association of TFIIH with the XPC-DNA complex is ATP independent (Riedl, Hanaoka et al. 2003; Tapias, Auriol et al. 2004). ATP is required for the helicase function of TFIIH (XPB and XPD) to extend the unwound region to about 20-30 nucleotides, preparing for the arrival of other repair proteins (Roy, Schaeffer et al. 1994; Riedl, Hanaoka et al. 2003). Because TFIIH is required for dual incision even on artificial templates with unpaired DNA around the lesion, TFIIH must play an additional critical role in NER besides unwinding DNA. This role may be in organizing the multiple assembly steps and controlling the activity of the next factor and/or accuracy.

The precise function of **XPA** in damage recognition is still unknown. For a long time it was considered to be the primary damage recognition factor in NER because of its high binding affinity for damaged DNA. XPA was also considered to play a central role in NER, most likely in damage recognition because XPA-deficiency is the most severe NER phenotype. XPA arrives at the lesion after TFIIH as a separate factor (Rademakers, Volker et al. 2003; Riedl, Hanaoka et al. 2003; Tapias, Auriol et al. 2004). It can directly interact with many NER factors like RPA, TFIIH, XPF and ERCC1. In agreement with this, XPA might stabilize intermediates in the open complex by bridging large protein complexes to correctly assemble at the damaged site.  For the damage recognition step, the interaction between XPA and RPA is of interest. *In vitro* these two proteins form a stable complex that can bind rigidly kinked DNA (Stigger, Drissi et al. 1998) (Wang, Mahrenholz et al. 2000). It has been proposed that XPA alone or in complex with RPA might double-check the opened region to confirm the presence of damage (often referred in the literature as "damage verification"). Although it has generally accepted

that XPA plays a crucial role at an early stage of both GG-NER and TC-NER, its precise function remains unclear.

As stated above, CPDs are poorly recognized by XPC-HR23B (Kusumoto, Masutani et al. 2001) but this type of lesion is repaired by GG-NER in an XPC-dependent manner. CPD excision does not occur *in vitro* using protein extracts from cells lacking functional XPC (Kusumoto, Masutani et al. 2001; Sugasawa 2006). Although XPC is required, other factor(s) may be involved in recognition of this lesion. UV-DDB is considered to be the initial recognition factor for this particular type of lesion and assists subsequent binding of XPC-HR23B (Hwang, Ford et al. 1999) (Tang, Hwang et al. 2000) (Kusumoto, Masutani et al. 2001).

Based on data obtained so far a multi-step model of damage recognition in GG-NER has been proposed (Sugasawa, Okamoto et al. 2001). According to the model XPC recognizes physical distortions of the helix (which is, in fact, the only common feature in all lesions) caused by the damage and binds to the opposing strand. This binding induces local unwinding of DNA for several nucleotides. TFIIH recognizes the XPC-damaged unwound-DNA complex. With its helicase activity TFIIH further opens the helix in this region. Subsequently, XPA, XPA-RPA or XPG binds to the damage site. The presence of damage is presumably confirmed in this step. TFIIH, XPA, RPA are candidates that can alone, or in combination with other factors verify the presence of the lesion prior to excision. For some types of the lesions, including CPDs, UV-DDB is the initial damage recognition factor that recruits XPC-HR23B with the same sequence of subsequent steps.

## *References*

Aboussekhra, A., M. Biggerstaff, et al. (1995). "Mammalian DNA nucleotide excision repair reconstituted with purified protein components." Cell **80**(6): 859-668.

Akoulitchev, S., T. P. Makela, et al. (1995). "Requirement for TFIIH kinase activity in transcription by RNA polymerase II." Nature **377**(6549): 557-560.

Araki, M., C. Masutani, et al. (2001). "Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair." J Biol Chem **276**(22): 18665-18672.

Araujo, S. J., E. A. Nigg, et al. (2001). "Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome." Mol Cell Biol **21**(7): 2281-2291.

Araujo, S. J., F. Tirode, 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.

Asahina, H., I. Kuraoka, et al. (1994). "The XPA protein is a zinc metalloprotein with an ability to recognize various kinds of DNA damage." Mutat Res **315**(3): 229-237.

Aubert, C., P. Mathis, et al. (1999). "Intraprotein electron transfer between tyrosine and tryptophan in DNA photolyase from Anacystis nidulans." Proc Natl Acad Sci U S A **96**(10): 5423-5427.

Aubert, C., M. H. Vos, et al. (2000). "Intraprotein radical transfer during photoactivation of DNA photolyase." Nature **405**(6786): 586-590.

Balajee, A. S. and V. A. Bohr (2000). "Genomic heterogeneity of nucleotide excision repair." Gene **250**(1-2): 15-30.

Barnes, D. E., A. E. Tomkinson, et al. (1992). "Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents." Cell **69**(3): 495-503.

Batty, D., V. Rapic'-Otrin, et al. (2000). "Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites." J Mol Biol **300**(2): 275-290.

Beerens, N., J. H. Hoeijmakers, et al. (2005). "The CSB protein actively wraps DNA." J Biol Chem **280**(6): 4722-4729.

Bessho, T., A. Sancar, et al. (1997). "Reconstitution of human excision nuclease with recombinant XPF-ERCC1 complex." J Biol Chem **272**(6): 3833-3837.

Bhatia, P. K., Z. Wang, et al. (1996). "DNA repair and transcription." Curr Opin Genet Dev **6**(2): 146-150.

Bird, L. E., H. S. Subramanya, et al. (1998). "Helicases: a unifying structural theme?" Curr Opin Struct Biol **8**(1): 14-18.

Bochkarev, A., R. A. Pfuetzner, et al. (1997). "Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA." Nature **385**(6612): 176-181.

Bootsma, D. and J. H. Hoeijmakers (1993). "DNA repair. Engagement with transcription." Nature **363**(6425): 114-115.

Bootsma, D., K. H. Kraemer, et al. (2002). The genetic basis of human cancer. Nucleotide excision repair syndromes: xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy. B. Vogelstein and K. W. Kinzler. New York, McGraw-Hill 2002.

Buchko, G. W., S. Ni, et al. (1998). "Structural features of the minimal DNA binding domain (M98-F219) of human nucleotide excision repair protein XPA." Nucleic Acids Res **26**(11): 2779-2788.

Bunick, C. G., M. R. Miller, et al. (2006). "Biochemical and structural domain analysis of xeroderma pigmentosum complementation group C protein." Biochemistry **45**(50): 14965-14979.

Buterin, T., C. Meyer, et al. (2005). "DNA quality control by conformational readout on the undamaged strand of the double helix." Chem Biol **12**(8): 913-922.

Byrdin, M., A. P. Eker, et al. (2003). "Dissection of the triple tryptophan electron transfer chain in Escherichia coli DNA photolyase: Trp382 is the primary donor in photoactivation." Proc Natl Acad Sci U S A **100**(15): 8676-8681.

Cadet, J., E. Sage, et al. (2005). "Ultraviolet radiation-mediated damage to cellular DNA." Mutat Res **571**(1-2): 3-17.

Cadet, J. and P. Vigny. (1990). The photochemistry of Nucleic Acids In Bioorganic Photochemistry. Photochemistry and the Nucleic Acids. New York, John Wiley & Sons, Inc.

Citterio, E., V. Van Den Boom, et al. (2000). "ATP-dependent chromatin remodeling by the Cockayne syndrome B DNA repair-transcription-coupling factor." Mol Cell Biol **20**(20): 7643-7653.

Clingen, P. H., C. F. Arlett, et al. (1995). "Induction of cyclobutane pyrimidine dimers, pyrimidine(6-4) pyrimidone photoproducts, and Dewar valence isomers by natural sunlight in normal human mononuclear cells." Cancer Res **55**(11): 2245-2248.

Coin, F., E. Bergmann, et al. (1999). "Mutations in XPB and XPD helicases found in xeroderma

pigmentosum patients impair the transcription function of TFIIH." <u>Embo J</u> **18**(5): 1357-1366.

Coin, F., J. C. Marinoni, 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." <u>Nat Genet</u> **20**(2): 184-188.

Constantinou, A., D. Gunz, et al. (1999). "Conserved residues of human XPG protein important for nuclease activity and function in nucleotide excision repair." <u>J Biol Chem</u> **274**(9): 5637-5648.

Datta, A., S. Bagchi, et al. (2001). "The p48 subunit of the damaged-DNA binding protein DDB associates with the CBP/p300 family of histone acetyltransferase." <u>Mutat Res</u> **486**(2): 89-97.

de Laat, W. L., E. Appeldoorn, et al. (1998). "DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair." <u>Genes Dev</u> **12**(16): 2598-2609.

de Laat, W. L., N. G. Jaspers, et al. (1999). "Molecular mechanism of nucleotide excision repair." <u>Genes Dev</u> **13**(7): 768-785.

Drapkin, R., G. Le Roy, et al. (1996). "Human cyclin-dependent kinase-activating kinase exists in three distinct complexes." <u>Proc Natl Acad Sci U S A</u> **93**(13): 6488-6493.

Dunand-Sauthier, I., M. Hohl, et al. (2005). "The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity." <u>J Biol Chem</u> **280**(8): 7030-7037.

Eisen, J. A. and P. C. Hanawalt (1999). "A phylogenomic study of DNA repair genes, proteins, and processes." <u>Mutat Res</u> **435**(3): 171-213.

Fanning, E., V. Klimovich, et al. (2006). "A dynamic model for replication protein A (RPA) function in DNA processing pathways." <u>Nucleic Acids Res</u> **34**(15): 4126-4137.

Fisher, R. P. (2005). "Secrets of a double agent: CDK7 in cell-cycle control and transcription." <u>J Cell Sci</u> **118**(Pt 22): 5171-5180.

Fousteri, M., W. Vermeulen, et al. (2006). "Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo." <u>Mol Cell</u> **23**(4): 471-482.

Friedberg, E. C., G. C. Walker, et al. (2005). <u>DNA Repair and Mutagenesis</u>. Washington, DC, ASM Press.

Fu, Y. C., X. P. Jin, et al. (2000). "Ultraviolet radiation and reactive oxygen generation as inducers of keratinocyte apoptosis: protective role of tea polyphenols." <u>J Toxicol Environ Health A</u> **61**(3): 177-188.

Gervais, V., V. Lamour, et al. (2004). "TFIIH contains a PH domain involved in DNA nucleotide excision repair." <u>Nat Struct Mol Biol</u> **11**(7): 616-622.

Giglia-Mari, G., F. Coin, et al. (2004). "A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A." <u>Nat Genet</u> **36**(7): 714-719.

Gillet, L. C. and O. D. Scharer (2006). "Molecular mechanisms of mammalian global genome nucleotide excision repair." <u>Chem Rev</u> **106**(2): 253-276.

Groisman, R., J. Polanowska, et al. (2003). "The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage." <u>Cell</u> **113**(3): 357-367.

Guzder, S. N., Y. Habraken, et al. (1995). "Reconstitution of yeast nucleotide excision repair with purified Rad proteins, replication protein A, and transcription factor TFIIH." <u>J Biol Chem</u> **270**(22): 12973-12976.

Hanawalt, P. C. (1994). "Transcription-coupled repair and human disease." Science **266**(5193): 1957-1958.

Hanawalt, P. C. (1995). "DNA repair comes of age." Mutat Res **336**(2): 101-113.

He, Z., L. A. Henricksen, et al. (1995). "RPA involvement in the damage-recognition and incision steps of nucleotide excision repair." Nature **374**(6522): 566-569.

Henning, K. A., L. Li, et al. (1995). "The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH." Cell **82**(4): 555-564.

Hiyama, H., M. Yokoi, et al. (1999). "Interaction of hHR23 with S5a. The ubiquitin-like domain of hHR23 mediates interaction with S5a subunit of 26 S proteasome." J Biol Chem **274**(39): 28019-28025.

Humbert, S., H. van Vuuren, et al. (1994). "p44 and p34 subunits of the BTF2/TFIIH transcription factor have homologies with SSL1, a yeast protein involved in DNA repair." Embo J **13**(10): 2393-2398.

Husain, I., J. Griffith, et al. (1988). "Thymine dimers bend DNA." Proc Natl Acad Sci U S A **85**(8): 2558-2562.

Hwang, B. J., J. M. Ford, et al. (1999). "Expression of the p48 xeroderma pigmentosum gene is p53-dependent and is involved in global genomic repair." Proc Natl Acad Sci U S A **96**(2): 424-428.

Ikegami, T., I. Kuraoka, et al. (1998). "Solution structure of the DNA- and RPA-binding domain of the human repair factor XPA." Nat Struct Biol **5**(8): 701-706.

Itin, P. H., A. Sarasin, et al. (2001). "Trichothiodystrophy: update on the sulfur-deficient brittle hair syndromes." J Am Acad Dermatol **44**(6): 891-920; quiz 921-4.

Iyer, N., M. S. Reagan, et al. (1996). "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 **35**(7): 2157-2167.

Janicijevic, A., K. Sugasawa, et al. (2003). "DNA bending by the human damage recognition complex XPC-HR23B." DNA Repair (Amst) **2**(3): 325-336.

Jawhari, A., J. P. Laine, et al. (2002). "p52 Mediates XPB function within the transcription/repair factor TFIIH." J Biol Chem **277**(35): 31761-31767.

Jones, C. J. and R. D. Wood (1993). "Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA." Biochemistry **32**(45): 12096-12104.

Kan, L., L. Voituriez, et al. (1988). "Nuclear magnetic resonance studies of cis-syn, trans-syn, and 6-4 photodimers of thymidylyl(3'-5')thymidine monophosphate and cis-syn photodimers of thymidylyl(3'-5')thymidine cyanoethyl phosphotriester." Biochemistry **27**(15): 5796-5803.

Kemmink, J., R. Boelens, et al. (1987). "Two-dimensional 1H NMR study of two cyclobutane type photodimers of thymidylyl-(3'----5')-thymidine." Eur Biophys J **14**(5): 293-299.

Kemmink, J., R. Boelens, et al. (1987). "Conformational changes in the oligonucleotide duplex d(GCGTTGCG) x d(CGCAACGC) induced by formation of a cis-syn thymine dimer. A two-dimensional NMR study." Eur J Biochem **162**(1): 37-43.

Kim, C., B. F. Paulus, et al. (1994). "Interactions of human replication protein A with oligonucleotides." Biochemistry **33**(47): 14197-14206.

Kim, C., R. O. Snyder, et al. (1992). "Binding properties of replication protein A from human and yeast cells." Mol Cell Biol **12**(7): 3050-3059.

Kim, J. K. and B. S. Choi (1995). "The solution structure of DNA duplex-decamer containing the (6-4) photoproduct of thymidylyl(3'-->5')thymidine by NMR and relaxation matrix refinement." Eur J Biochem **228**(3): 849-854.

Kim, J. K., D. Patel, et al. (1995). "Contrasting structural impacts induced by cis-syn cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity." Photochem Photobiol **62**(1): 44-50.

Kim, S. T., K. Malhotra, et al. (1994). "Characterization of (6-4) photoproduct DNA photolyase." J Biol Chem **269**(11): 8535-40.

Koberle, B., V. Roginskaya, et al. (2006). "XPA protein as a limiting factor for nucleotide excision repair and UV sensitivity in human cells." DNA Repair (Amst) **5**(5): 641-8.

Kochevar, I. E. (1990). "UV-induced protein alterations and lipid oxidation in erythrocyte membranes." Photochem Photobiol **52**(4): 795-800.

Komori, H., R. Masui, et al. (2001). "Crystal structure of thermostable DNA photolyase: pyrimidine-dimer recognition mechanism." Proc Natl Acad Sci U S A **98**(24): 13560-5.

Kuraoka, I., E. H. Morita, et al. (1996). "Identification of a damaged-DNA binding domain of the XPA protein." Mutat Res **362**(1): 87-95.

Kusumoto, R., C. Masutani, et al. (2001). "Diversity of the damage recognition step in the global genomic nucleotide excision repair in vitro." Mutat Res **485**(3): 219-27.

Lee, J. H., C. J. Park, et al. (2004). "NMR structure of the DNA decamer duplex containing double T*G mismatches of cis-syn cyclobutane pyrimidine dimer: implications for DNA damage recognition by the XPC-hHR23B complex." Nucleic Acids Res **32**(8): 2474-2481.

Lehmann, A. R. (2003). "DNA repair-deficient diseases, xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy." Biochimie **85**(11): 1101-1111.

Lehmann, A. R. (2005). "Replication of damaged DNA by translesion synthesis in human cells." FEBS Lett **579**(4): 873-876.

Li, L., S. J. Elledge, et al. (1994). "Specific association between the human DNA repair proteins XPA and ERCC1." Proc Natl Acad Sci U S A **91**(11): 5012-5016.

Li, L., X. Lu, et al. (1995). "An interaction between the DNA repair factor XPA and replication protein A appears essential for nucleotide excision repair." Mol Cell Biol **15**(10): 5396-5402.

Li, L., C. A. Peterson, et al. (1995). "Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair." Mol Cell Biol **15**(4): 1993-1998.

Luijsterburg, M. S., J. Goedhart, et al. (2007). "Dynamic in vivo interaction of DDB2 E3 ubiquitin ligase with UV-damaged DNA is independent of damage-recognition protein XPC." J Cell Sci **120**(Pt 15): 2706-2716.

Ma, L., E. D. Siemssen, et al. (1994). "The xeroderma pigmentosum group B protein ERCC3 produced in the baculovirus system exhibits DNA helicase activity." Nucleic Acids Res **22**(20): 4095-4102.

Machius, M., L. Henry, et al. (1999). "Crystal structure of the DNA nucleotide excision repair enzyme UvrB from Thermus thermophilus." Proc Natl Acad Sci U S A **96**(21): 11717-11722.

Madronich, S., R. L. McKenzie, et al. (1998). "Changes in biologically active ultraviolet radiation reaching the Earth's surface." J Photochem Photobiol B **46**(1-3): 5-19.

Maillard, O., S. Solyom, et al. (2007). "An aromatic sensor with aversion to damaged strands confers versatility to DNA repair." PLoS Biol **5**(4): e79.

Makela, T. P., J. D. Parvin, et al. (1995). "A kinase-deficient transcription factor TFIIH is functional in basal and activated transcription." Proc Natl Acad Sci U S A **92**(11): 5174-5178.

Malta, E., G. F. Moolenaar, et al. (2006). "Base flipping in nucleotide excision repair." J Biol Chem **281**(4): 2184-2194.

Marinoni, J. C., R. Roy, et al. (1997). "Cloning and characterization of p52, the fifth subunit of the core of the transcription/DNA repair factor TFIIH." Embo J **16**(5): 1093-1102.

Martinez, E., V. B. Palhan, et al. (2001). "Human STAGA complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors in vivo." Mol Cell Biol **21**(20): 6782-6795.

Masutani, C., K. Sugasawa, et al. (1994). "Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23." Embo J **13**(8): 1831-43.

Matsuda, T., M. Saijo, et al. (1995). "DNA repair protein XPA binds replication protein A (RPA)." J Biol Chem **270**(8): 4152-7.

Matsunaga, T., Y. Hatakeyama, et al. (1993). "Establishment and characterization of a monoclonal antibody recognizing the Dewar isomers of (6-4)photoproducts." Photochem Photobiol **57**(6): 934-40.

Matsunaga, T., D. Mu, et al. (1995). "Human DNA repair excision nuclease. Analysis of the roles of the subunits involved in dual incisions by using anti-XPG and anti-ERCC1 antibodies." J Biol Chem **270**(35): 20862-9.

Matsunaga, T., C. H. Park, et al. (1996). "Replication protein A confers structure-specific endonuclease activities to the XPF-ERCC1 and XPG subunits of human DNA repair excision nuclease." J Biol Chem **271**(19): 11047-50.

McAteer, K., Y. Jing, et al. (1998). "Solution-state structure of a DNA dodecamer duplex containing a Cis-syn thymine cyclobutane dimer, the major UV photoproduct of DNA." J Mol Biol **282**(5): 1013-32.

Mees, A., T. Klar, et al. (2004). "Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair." Science **306**(5702): 1789-93.

Min, J. H. and N. P. Pavletich (2007). "Recognition of DNA damage by the Rad4 nucleotide excision repair protein." Nature **449**(7162): 570-5.

Mitchell, D. L. and R. S. Nairn (1989). "The biology of the (6-4) photoproduct." Photochem Photobiol **49**(6): 805-19.

Mocquet, V., K. Kropachev, et al. (2007). "The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions." Embo J **26**(12): 2923-2932.

Moolenaar, G. F., L. Hoglund, et al. (2001). "Clue to damage recognition by UvrB: residues in the beta-hairpin structure prevent binding to non-damaged DNA." Embo J **20**(21): 6140-6149.

Morikawa, K. and M. Shirakawa (2000). "Three-dimensional structural views of damaged-DNA recognition: T4 endonuclease V, E. coli Vsr protein, and human nucleotide excision repair factor XPA." Mutat Res **460**(3-4): 257-275.

Moser, J., H. Kool, et al. (2007). "Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner." Mol Cell **27**(2): 311-323.

Moser, J., M. Volker, et al. (2005). "The UV-damaged DNA binding protein mediates efficient targeting of the nucleotide excision repair complex to UV-induced photo lesions." DNA Repair (Amst) **4**(5): 571-582.

Motycka, T. A., T. Bessho, et al. (2004). "Physical and functional interaction between the XPF/ERCC1 endonuclease and hRad52." J Biol Chem **279**(14): 13634-13639.

Mu, D., D. S. Hsu, et al. (1996). "Reaction mechanism of human DNA repair excision nuclease." J Biol Chem **271**(14): 8285-8294.

Mu, D. and A. Sancar (1997). "Model for XPC-independent transcription-coupled repair of pyrimidine dimers in humans." J Biol Chem **272**(12): 7570-7573.

Nagai, A., M. Saijo, et al. (1995). "Enhancement of damage-specific DNA binding of XPA by interaction with the ERCC1 DNA repair protein." Biochem Biophys Res Commun **211**(3): 960-966.

Nakagawa, N., M. Sugahara, et al. (1999). "Crystal structure of Thermus thermophilus HB8 UvrB protein, a key enzyme of nucleotide excision repair." J Biochem (Tokyo) **126**(6): 986-990.

Nakatsu, Y., H. Asahina, et al. (2000). "XAB2, a novel tetratricopeptide repeat protein involved in transcription-coupled DNA repair and transcription." J Biol Chem **275**(45): 34931-34937.

Nance, M. A. and S. A. Berry (1992). "Cockayne syndrome: review of 140 cases." Am J Med Genet **42**(1): 68-84.

Ng, J. M., W. Vermeulen, et al. (2003). "A novel regulation mechanism of DNA repair by damage-induced and RAD23-dependent stabilization of xeroderma pigmentosum group C protein." Genes Dev **17**(13): 1630-1645.

Niedernhofer, L. J., H. Odijk, et al. (2004). "The structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks." Mol Cell Biol **24**(13): 5776-5787.

Nishi, R., Y. Okuda, et al. (2005). "Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein." Mol Cell Biol **25**(13): 5664-5674.

Nitta, M., M. Saijo, et al. (2000). "A novel cytoplasmic GTPase XAB1 interacts with DNA repair protein XPA." Nucleic Acids Res **28**(21): 4212-4218.

Nocentini, S., F. Coin, et al. (1997). "DNA damage recognition by XPA protein promotes efficient recruitment of transcription factor II H." J Biol Chem **272**(37): 22991-22994.

O'Donovan, A., A. A. Davies, et al. (1994). "XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair." Nature **371**(6496): 432-435.

Orren, D. K. and A. Sancar (1989). "The (A)BC excinuclease of Escherichia coli has only the UvrB and UvrC subunits in the incision complex." Proc Natl Acad Sci U S A **86**(14): 5237-5241.

Orren, D. K., C. P. Selby, et al. (1992). "Post-incision steps of nucleotide excision repair in Escherichia coli. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I." J Biol Chem **267**(2): 780-788.

Park, C. H., T. Bessho, et al. (1995). "Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease." J Biol Chem **270**(39): 22657-22660.

Park, H., K. Zhang, et al. (2002). "Crystal structure of a DNA decamer containing a cis-syn thymine dimer." Proc Natl Acad Sci U S A **99**(25): 15965-15970.

Patrick, M. H. and D. M. Gray (1976). "Independence of photoproduct formation on DNA conformation." Photochem Photobiol **24**(6): 507-513.

Patrick, M. H. and R. Rahn (1976). Photochemistry and Photobiology of Nucleic Acids. S. Y. Wang. New York, Academic Press. **11:** 35-95.

Patrick, S. M., G. G. Oakley, et al. (2005). "DNA damage induced hyperphosphorylation of replication protein A. 2. Characterization of DNA binding activity, protein interactions, and activity in DNA replication and repair." Biochemistry **44**(23): 8438-8448.

Perdiz, D., P. Grof, et al. (2000). "Distribution and repair of bipyrimidine photoproducts in solar UV-irradiated mammalian cells. Possible role of Dewar photoproducts in solar mutagenesis." J Biol Chem **275**(35): 26732-26742.

Pfeifer, G. P., Y. H. You, et al. (2005). "Mutations induced by ultraviolet light." Mutat Res **571**(1-2): 19-31.

Rademakers, S., M. Volker, et al. (2003). "Xeroderma pigmentosum group A protein loads as a separate factor onto DNA lesions." Mol Cell Biol **23**(16): 5755-5767.

Rapic-Otrin, V., M. P. McLenigan, et al. (2002). "Sequential binding of UV DNA damage binding factor and degradation of the p48 subunit as early events after UV irradiation." Nucleic Acids Res **30**(11): 2588-2598.

Reardon, J. T., H. Ge, et al. (1996). "Isolation and characterization of two human transcription factor IIH (TFIIH)-related complexes: ERCC2/CAK and TFIIH." Proc Natl Acad Sci U S A **93**(13): 6482-6487.

Riedl, T., F. Hanaoka, et al. (2003). "The comings and goings of nucleotide excision repair factors on damaged DNA." Embo J **22**(19): 5293-5303.

Rossignol, M., I. Kolb-Cheynel, et al. (1997). "Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH." Embo J **16**(7): 1628-1637.

Roy, R., L. Schaeffer, 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.

Rupert, C. S. (1975). "Enzymatic photoreactivation: overview." Basic Life Sci **5A**: 73-87.

Rycyna, R. E. and J. L. Alderfer (1985). "UV irradiation of nucleic acids: formation, purification and solution conformational analysis of the '6-4 lesion' of dTpdT." Nucleic Acids Res **13**(16): 5949-5963.

Sage, E. (1993). "Distribution and repair of photolesions in DNA: genetic consequences and the role of sequence context." Photochem Photobiol **57**(1): 163-174.

Salisbury, J. L., K. M. Suino, et al. (2002). "Centrin-2 is required for centriole duplication in mammalian cells." Curr Biol **12**(15): 1287-1292.

Sancar, A. and G. B. Sancar (1988). "DNA repair enzymes." Annu Rev Biochem **57**: 29-67.

Sancar, G. B. (1990). "DNA photolyases: physical properties, action mechanism, and roles in dark repair." Mutat Res **236**(2-3): 147-160.

Sarasin, A. and A. Stary (2007). "New insights for understanding the transcription-coupled repair pathway." DNA Repair (Amst) **6**(2): 265-269.

Sarker, A. H., S. E. Tsutakawa, et al. (2005). "Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: insights for transcription-coupled repair and Cockayne Syndrome." Mol Cell **20**(2): 187-198.

Schaeffer, L., V. Moncollin, et al. (1994). "The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor." Embo J **13**(10): 2388-2392.

Schaeffer, L., R. Roy, et al. (1993). "DNA repair helicase: a component of BTF2 (TFIIH) basic

transcription factor." <u>Science</u> **260**(5104): 58-63.

Schultz, P., S. Fribourg, et al. (2000). "Molecular structure of human TFIIH." <u>Cell</u> **102**(5): 599-607.

Schweizer, U., T. Hey, et al. (1999). "Photocrosslinking locates a binding site for the large subunit of human replication protein A to the damaged strand of cisplatin-modified DNA." <u>Nucleic Acids Res</u> **27**(15): 3183-3189.

Seroz, T., C. Perez, et al. (2000). "p44/SSL1, the regulatory subunit of the XPD/RAD3 helicase, plays a crucial role in the transcriptional activity of TFIIH." <u>J Biol Chem</u> **275**(43): 33260-33266.

Shivji, K. K., M. K. Kenny, et al. (1992). "Proliferating cell nuclear antigen is required for DNA excision repair." <u>Cell</u> **69**(2): 367-374.

Shivji, M. K., A. P. Eker, et al. (1994). "DNA repair defect in xeroderma pigmentosum group C and complementing factor from HeLa cells." <u>J Biol Chem</u> **269**(36): 22749-22757.

Sijbers, A. M., W. L. de Laat, et al. (1996). "Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease." <u>Cell</u> **86**(5): 811-822.

Skorvaga, M., K. Theis, et al. (2002). "The beta -hairpin motif of UvrB is essential for DNA binding, damage processing, and UvrC-mediated incisions." <u>J Biol Chem</u> **277**(2): 1553-1559.

Smith, C. A. and J. S. Taylor (1993). "Preparation and characterization of a set of deoxyoligonucleotide 49-mers containing site-specific cis-syn, trans-syn-I, (6-4), and Dewar photoproducts of thymidylyl(3'-->5')-thymidine." <u>J Biol Chem</u> **268**(15): 11143-11151.

Stigger, E., R. Drissi, et al. (1998). "Functional analysis of human replication protein A in nucleotide excision repair." <u>J Biol Chem</u> **273**(15): 9337-9343.

Sugasawa, K. (2006). "UV-induced ubiquitylation of XPC complex, the UV-DDB-ubiquitin ligase complex, and DNA repair." <u>J Mol Histol</u> **37**(5-7): 189-202.

Sugasawa, K., C. Masutani, et al. (1996). "HHR23B, a human Rad23 homolog, stimulates XPC protein in nucleotide excision repair in vitro." <u>Mol Cell Biol</u> **16**(9): 4852-4861.

Sugasawa, K., J. M. Ng, et al. (1998). "Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair." <u>Mol Cell</u> **2**(2): 223-232.

Sugasawa, K., J. M. Ng, et al. (1997). "Two human homologs of Rad23 are functionally interchangeable in complex formation and stimulation of XPC repair activity." <u>Mol Cell Biol</u> **17**(12): 6924-6931.

Sugasawa, K., T. Okamoto, et al. (2001). "A multistep damage recognition mechanism for global genomic nucleotide excision repair." <u>Genes Dev</u> **15**(5): 507-521.

Sugasawa, K., Y. Shimizu, et al. (2002). "A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex." <u>DNA Repair (Amst)</u> **1**(1): 95-107.

Sung, P., V. Bailly, et al. (1993). "Human xeroderma pigmentosum group D gene encodes a DNA helicase." <u>Nature</u> **365**(6449): 852-855.

Svejstrup, J. Q. (2003). "Rescue of arrested RNA polymerase II complexes." <u>J Cell Sci</u> **116**(Pt 3): 447-451.

Svejstrup, J. Q., Z. Wang, et al. (1995). "Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome." <u>Cell</u> **80**(1): 21-28.

Takagi, Y., C. A. Masuda, et al. (2005). "Ubiquitin ligase activity of TFIIH and the transcriptional response to DNA damage." <u>Mol Cell</u> **18**(2): 237-243.

Tamada, T., K. Kitadokoro, et al. (1997). "Crystal structure of DNA photolyase from Anacystis nidulans." <u>Nat Struct Biol</u> **4**(11): 887-891.

Tang, J. Y., B. J. Hwang, et al. (2000). "Xeroderma pigmentosum p48 gene enhances global genomic repair and suppresses UV-induced mutagenesis." Mol Cell **5**(4): 737-744.

Tantin, D. (1998). "RNA polymerase II elongation complexes containing the Cockayne syndrome group B protein interact with a molecular complex containing the transcription factor IIH components xeroderma pigmentosum B and p62." J Biol Chem **273**(43): 27794-27799.

Tapias, A., J. Auriol, et al. (2004). "Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors." J Biol Chem **279**(18): 19074-19783.

Taylor, J. S. and I. R. Brockie (1988). "Synthesis of a trans-syn thymine dimer building block. Solid phase synthesis of CGTAT[t,s]TATGC." Nucleic Acids Res **16**(11): 5123-5136.

Taylor, J. S., H. F. Lu, et al. (1990). "Quantitative conversion of the (6-4) photoproduct of TpdC to its Dewar valence isomer upon exposure to simulated sunlight." Photochem Photobiol **51**(2): 161-167.

Theis, K., P. J. Chen, et al. (1999). "Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair." Embo J **18**(24): 6899-6907.

Todo, T., H. Ryo, et al. (1996). "Similarity among the Drosophila (6-4)photolyase, a human photolyase homolog, and the DNA photolyase-blue-light photoreceptor family." Science **272**(5258): 109-112.

Tremeau-Bravard, A., C. Perez, et al. (2001). "A role of the C-terminal part of p44 in the promoter escape activity of transcription factor IIH." J Biol Chem **276**(29): 27693-27697.

Tremeau-Bravard, A., T. Riedl, et al. (2004). "Fate of RNA polymerase II stalled at a cisplatin lesion." J Biol Chem **279**(9): 7751-7759.

Truglio, J. J., E. Karakas, et al. (2006). "Structural basis for DNA recognition and processing by UvrB." Nat Struct Mol Biol **13**(4): 360-364.

Tyrrell, R. M. and S. M. Keyse (1990). "New trends in photobiology. The interaction of UVA radiation with cultured cells." J Photochem Photobiol B **4**(4): 349-361.

van den Boom, V., E. Citterio, et al. (2004). "DNA damage stabilizes interaction of CSB with the transcription elongation machinery." J Cell Biol **166**(1): 27-36.

van den Boom, V., N. G. Jaspers, et al. (2002). "When machines get stuck--obstructed RNA polymerase II: displacement, degradation or suicide." Bioessays **24**(9): 780-784.

van der Spek, P. J., A. Eker, et al. (1996). "XPC and human homologs of RAD23: intracellular localization and relationship to other nucleotide excision repair complexes." Nucleic Acids Res **24**(13): 2551-2559.

van Gool, A. J., E. Citterio, et al. (1997). "The Cockayne syndrome B protein, involved in transcription-coupled DNA repair, resides in an RNA polymerase II-containing complex." Embo J **16**(19): 5955-5965.

Van Houten, B., J. A. Eisen, et al. (2002). "A cut above: discovery of an alternative excision repair pathway in bacteria." Proc Natl Acad Sci U S A **99**(5): 2581-2583.

van Noort, J., F. Orsini, et al. (1999). "DNA bending by photolyase in specific and non-specific complexes studied by atomic force microscopy." Nucleic Acids Res **27**(19): 3875-3880.

van Noort, S. J., K. O. van der Werf, et al. (1998). "Direct visualization of dynamic protein-DNA interactions with a dedicated atomic force microscope." Biophys J **74**(6): 2840-2849.

van Oosterwijk, M. F., R. Filon, et al. (1998). "Lack of transcription-coupled repair of acetylaminofluorene

DNA adducts in human fibroblasts contrasts their efficient inhibition of transcription." J Biol Chem **273**(22): 13599-13604.

Verhoeven, E. E., C. Wyman, et al. (2002). "The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands." Embo J **21**(15): 4196-205.

Verhoeven, E. E., C. Wyman, et al. (2001). "Architecture of nucleotide excision repair complexes: DNA is wrapped by UvrB before and after damage recognition." Embo J **20**(3): 601-611.

Vermeulen, W., A. J. van Vuuren, et al. (1994). "Three unusual repair deficiencies associated with transcription factor BTF2(TFIIH): evidence for the existence of a transcription syndrome." Cold Spring Harb Symp Quant Biol **59**: 317-329.

Volker, M., M. J. Mone, et al. (2001). "Sequential assembly of the nucleotide excision repair factors in vivo." Mol Cell **8**(1): 213-224.

Wakasugi, M., A. Kawashima, et al. (2002). "DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair." J Biol Chem **277**(3): 1637-1640.

Wakasugi, M., J. T. Reardon, et al. (1997). "The non-catalytic function of XPG protein during dual incision in human nucleotide excision repair." J Biol Chem **272**(25): 16030-16034.

Wang, C. I. and J. S. Taylor (1991). "Site-specific effect of thymine dimer formation on dAn.dTn tract bending and its biological implications." Proc Natl Acad Sci U S A **88**(20): 9072-9076.

Wang, C. I. and J. S. Taylor (1993). "The trans-syn-I thymine dimer bends DNA by approximately 22 degrees and unwinds DNA by approximately 15 degrees." Chem Res Toxicol **6**(4): 519-523.

Wang, M., A. Mahrenholz, et al. (2000). "RPA stabilizes the XPA-damaged DNA complex through protein-protein interaction." Biochemistry **39**(21): 6433-6439.

Wang, S. Y. (1976). Photochemistry and Photobiology of Nucleic Acids. S. Y. Wang. New York, Academic Press. **1:** 295-350.

Waters, T. R., J. Eryilmaz, et al. (2006). "Damage detection by the UvrABC pathway: crystal structure of UvrB bound to fluorescein-adducted DNA." FEBS Lett **580**(27): 6423-6427.

Weber, S. (2005). "Light-driven enzymatic catalysis of DNA repair: a review of recent biophysical studies on photolyase." Biochim Biophys Acta **1707**(1): 1-23.

Wood, R. D. (1999). "DNA damage recognition during nucleotide excision repair in mammalian cells." Biochimie **81**(1-2): 39-44.

Yang, Z. G., Y. Liu, et al. (2002). "Dimerization of human XPA and formation of XPA2-RPA protein complex." Biochemistry **41**(43): 13012-20.

Yasui, A. (1998). DNA Photolyase. DNA Damage and Repair. J. Hoekstra and N. Totowa, Humana Press Inc. **2:** 9-32.

Yokoi, M., C. Masutani, et al. (2000). "The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA." J Biol Chem **275**(13): 9870-9875.

# Chapter 2

## The molecular machines of DNA repair:
## Scanning force microscopy analysis of their architecture
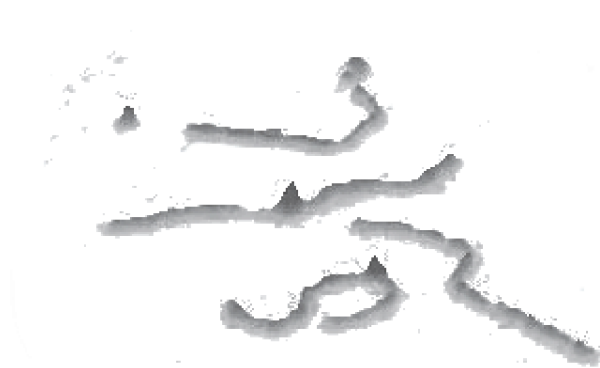
# The molecular machines of DNA repair: scanning force microscopy analysis of their architecture

A. JANIĆIJEVIĆ[1]*, D. RISTIC[1]* & C. WYMAN*†

*Department of Cell Biology and Genetics, Erasmus MC and †Radiation Oncology, Erasmus MC-Daniel, PO Box 1738, 3000 DR Rotterdam, The Netherlands*

## Summary

The application of scanning force microscope (SFM, also called atomic force microscope or AFM) imaging to study the architecture of proteins and their functional assemblies on DNA has provided new and exciting information on the mechanism of vital cellular processes. Rapid progress in molecular biology has resulted in the identification and isolation of proteins and protein complexes that function in specific DNA transactions. These proteins and protein complexes can now be analysed at the single molecule level, whereby the functional assemblies are often described as nanomachines. Understanding how they work requires understanding their structure and functional arrangement in three dimensions. The SFM is uniquely suited to provide three-dimensional structural information on biomolecules at nanometre resolution. In this review we focus on recent applications of SFM to reveal detailed information on the architecture and mechanism of action of protein machinery involved in safeguarding genome stability through DNA repair processes.

## Introduction

The limits of our understanding of biological processes are determined by the technology available to study and define them. Recent advances in microscopy techniques have provided new tools that expand these limits. This review will consider advances in understanding molecular mechanisms of complex genome transactions, specifically repair of DNA damage, that have been made possible by the application of scanning force microscope (SFM also called atomic force

Correspondence to: Claire Wyman. Tel.: +31 10 4088337; fax: +31 10 4089468; e-mail: c.wyman@erasmusmc.nl

[1]These authors contributed equally to this work.

microscope or AFM) imaging to determine molecular architecture. SFM provides three-dimensional (3-D) information on molecular structure at nanometre resolution without the need for external contrast agents. Molecules and complexes are individually analysed providing information on the variety of arrangements possible and their frequency in a mixture. Importantly, this type of single molecule structural analysis allows coherent description of features that would otherwise be lost in the averaging of bulk analysis. In addition, direct observation allows correlation of multiple structural features of individual molecular complexes.

The study of cellular processes has reached the level at which we can identify and isolate the required molecular components for functional analysis. There is intense effort along these lines to understand vital processes such as maintenance of genomic information. One aspect of genome maintenance is the repair of DNA damage that otherwise would result in disruption of cellular function, cell death or mutations. DNA repair pathways are generally classified by the type of damage they correct. The DNA repair pathways that will be considered here are the repair of structural damage, specifically double-strand breaks (DSBs), the removal of chemically modified bases by nucleotide excision repair (NER) and the correction of replication mistakes by mismatch repair (MMR). These various pathways differ widely in required components and mechanism. However, a common salient feature of these pathways is their requirement for the co-ordinated action of several proteins with specific DNA lesions in the genome in order to ensure accurate and efficient repair.

It is at the level of determining the arrangement and functional rearrangement of components that SFM has contributed to our understanding of the mechanisms of several DNA repair pathways. The unique contributions of SFM imaging to understanding mechanisms of DNA transactions can be divided into three categories. First, the size of individual molecules can be accurately determined from SFM images.
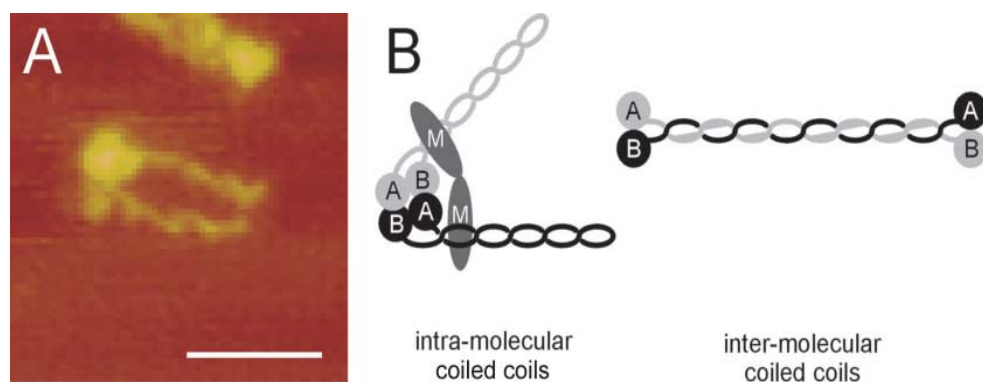
**Fig. 1.** The architecture of the human Rad50/Mre11 complex revealed by SFM imaging. (A) SFM image of a human Rad50/Mre11 complex. Scale bar = 50 nm. Height is represented by dark to light colours, 0–5 nm. (B) Cartoon diagram of the arrangement of Rad50 and Mre11 in the complex. A and B represent halves of the ATPase domain of Rad50 that are located at the N- and C-terminus of the amino acid sequence. M represents Mre11. The human complex shown here is arranged as a heterotetramer of two Rad50s with intramolecular coiled-coils associated near their ATPase domains via an Mre11 dimer. For comparison, the previously accepted model for Rad50 dimerization via intermolecular coiled-coils is shown to the right. (Image courtesy of M. de Jager.)

SFM imaging, like other single molecule techniques, reveals a wealth of information about the variation and distribution of different structures in a mixture. This was especially interesting in the analysis of Rad50/Mre11 structure. The molecules deposited and imaged in air displayed a variety of conformations of the Rad50 coiled-coils (de Jager *et al.*, 2001). These could represent either a mixture of static forms or flexibility of individual molecules. The ability to image molecules in buffer was successfully exploited here to demonstrate that individual Rad50 coiled-coils were indeed flexible. Subsequent high-resolution imaging and quantification of the local flexibility along the Rad50 coiled-coils identified two regions of increased flexibility that correspond to interruptions in the predicted coiled-coil structure (van Noort *et al.*, 2003).

**Protein–DNA complexes**

SFM analysis of protein–DNA complexes involved in a variety of DNA repair pathways has helped to elucidate mechanistic details of these vital cellular reactions. In many cases detailed aspects of molecular structure can be quantitatively described from direct observation by SFM. The examples reviewed include: (1) data on the stoichiometry of DNA-bound proteins, (2) protein-induced changes in DNA structure and (3) complex arrangements of proteins on DNA such as oligomerization and simultaneous interaction of multiple sites on one DNA or multiple DNA molecules.

*Photo-reactivation*

The required first step in any DNA repair pathway is recognition of damage. The ability to image biomolecules in buffer by

SFM has been exploited to study dynamic protein–DNA interactions as a basis for understanding damage recognition. There are two general mechanisms that describe the interaction of proteins with specific DNA sites (Berg *et al.*, 1981). Facilitated 1-D diffusion involves non-specific binding of a protein to DNA followed by translocation of the protein along the DNA strand to find a specific site. The alternative is location of a specific site on DNA by 3-D diffusion of the protein from solution. A simple protein–DNA interaction was studied in SFM experiments in order to optimize imaging parameters and test SFM as a means of distinguishing these two mechanisms of protein location of a specific DNA site. The reaction studied involved bacterial photolyase, a monomeric 55-kDa protein that binds to UV-induced thymidine dimers and uses the energy of visible light to reverse the crosslink chemically (reviewed in Sancar, 1994). By optimizing imaging parameters and modifying the SFM set-up, dynamic interaction of photolyase with partially immobilized DNA could be visualized (van Noort *et al.*, 1998). The DNA in this study was undamaged so all interactions were non-specific. Photolyase association with DNA, disassociation from DNA and sliding over DNA were observed. The later indicates that facilitated 1-D diffusion over DNA to locate damage is at least possible.

Protein–DNA interactions often result in distortion of DNA, which can have important implications for the mechanism of site recognition. SFM imaging has the advantage over most bulk biochemical methods in that protein-induced changes in DNA structure at non-specific sites can be measured and compared to the changes in DNA structure induced at specific binding sites (Erie *et al.*, 1994). For the case of photolyase this has provided some provocative results. It was determined that photolyase bound to damaged DNA induced a bend of 36°

whereas undamaged DNA was not bent by photolyase (van Noort *et al.,* 1999). The surprising finding of this study was that the distribution of bend angles for the specifically bound complexes was larger than that for both protein non-specifically bound and for free DNA. This indicates an increase in flexibility of DNA (Rivetti *et al.,* 1996) with photolyase bound to damaged sites. This puzzling result is, however, consistent with one model for photolyase action, which requires the damaged bases to be flipped out of the helix and could account for increased flexibility of the DNA. However, these imaging experiments were done with a DNA substrate randomly damaged by UV light, and the difference between specifically and non-specifically bound photolyase was based on statistical subtraction of characteristics of known non-specific complexes. Confirmation awaits analysis of photolyase bound to a known specific damaged site.

*Nucleotide excision repair*

The mechanism of DNA damage recognition is particularly intriguing in the NER pathway. NER is responsible for removing a wide variety of chemically distinct lesions from DNA that disrupt transcription and replication and that can lead to cell death or mutagenesis (Wood, 1999; Friedberg, 2003). This is a multistep pathway that is conserved from bacteria to humans. Mechanistically, NER can be divided into a series of steps that involve the assembly and modification of distinct protein complexes on DNA: (1) damage recognition, (2) damage demarcation and opening of the double-stranded DNA around the damage, (3) incision of one DNA strand on both sides of the damage and removal of the damaged oligonucleotide and (4) gap filling DNA synthesis and ligation to restore the correct DNA sequence. In bacteria, damage recognition is accomplished by a complex of the UvrA and UvrB proteins. Subsequently, UvrA is released, UvrC recognizes the UvrB bound to damaged DNA and incision on both sides of the damage occurs. SFM analysis of complexes formed by the Uvr proteins and DNA with a single damaged base at a defined site has provided new insight into the mechanism of bacterial NER. First it was demonstrated that in the damage recognition process DNA is wrapped around UvrB and that this wrapping is dependent on ATP binding by UvrB (Verhoeven *et al.,* 2001). DNA wrapping is obvious in SFM images by a decrease in DNA contour length measured through a bound protein complex. The amount of DNA wrapped was about 70 base pairs and the wrap was asymmetric with respect to the damage. This asymmetry is possibly important for determining the position of the incisions as they occur asymmetrically around the damage. DNA was also wrapped around UvrB when it was bound to undamaged DNA. This suggests that damage recognition depends in part on the fact that damaged DNA can more easily be distorted into a wrap and forms a more stable complex with UvrB (Verhoeven *et al.,* 2001).

It had long been accepted that the bacterial damage recognition complex consisted of two UvrA protomers and one UvrB
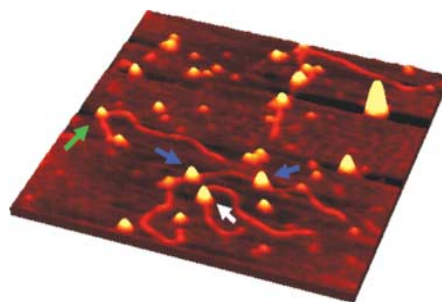


**Fig. 2.** SFM image reveals that UvrB bound to damaged DNA as a dimer. This image shows a combination of DNA protein complexes. The size standard, Ku70/80 (155 kDa), was bound to a 1500-bp DNA fragment. UvrB (76-kDa monomer) was bound to a 1020-bp DNA fragment with a single damaged base at one-third of its length from one end. In this experiment, ATP was washed out after UvrB binding, which releases the DNA wrap and results in two sizes of UvrB complexes bound to the damaged DNA. The larger UvrB complex, white arrow, is about the same size as Ku70/80, blue arrow. The smaller UvrB complex, green arrow, is half the size of Ku70/80. Thus, the larger UvrB complexes can only logically be UvrB dimers. The image is presented at as if it were tilted to emphasize topography. Height is represented as colour from dark to light, 0–3 nm. (Image courtesy of E. Verhoeven and N. Goosen.)

protomer. Volume measurements from SFM images of DNA-bound UvrB complexes clearly indicated the presence of two UvrB protomers (Fig. 2). The damage recognition complex must logically consist of two UvrA protomers and two UvrB protomers, thus correcting a long-standing misconception in the literature (Verhoeven *et al.,* 2002). The size and DNA distortion, in this case a wrap, was simultaneously determined for the individual UvrB–DNA complexes. Using conditions in which both dimers and monomers of UvrB were bound to DNA it was shown that the same amount of DNA was wrapped in both cases. This indicated that DNA was wraped around one UvrB monomer. Furthermore, after addition of UvrC it was shown that one UvrB monomer was released. From these data it was proposed that two UvrB monomers are needed in the pre-incision complex to detect damage in either DNA strand (Verhoeven *et al.,* 2002).

NER in mammalian systems follows the same basic steps as in bacteria, but more than 20 proteins are required (de Laat *et al.,* 1999). The damage recognition step of mammalian NER is also by extension more complex. Damage recognition in mammalian NER is believed to occur either via RNA polymerase stalled by a DNA lesion (de Laat *et al.,* 1999) or by initially binding the XPC-HR23B complex to damaged sites in non-transcribed regions of the genome (Sugasawa *et al.,* 1998; Volker *et al.,* 2001). As a start to understanding the architectural build up of a functional mammalian NER complex, changes in DNA structure induced by XPC-HR23B binding to damaged sites were investigated by SFM. It was shown that

unlike the bacterial UvrB damage recognition protein, XPC-HR23B did not wrap DNA upon binding or damage recognition. Instead, upon binding damaged DNA, XPC-HR23B induced a DNA bend of 39° and a slightly larger bend of 49° when bound to undamaged DNA (Janicijevic *et al.*, 2003). The significance of the difference in these bend angles is not yet certain as it was not statistically tested. It was proposed that the bend angle at the damaged site is an important structural feature required for subsequent build up of an active NER complex.

*Mismatch repair*

Another type of DNA damage is caused by replication errors resulting from the incorporation of mismatched bases as DNA is synthesized. In order to avoid introducing mutations, these mistakes are corrected in the MMR pathway. MMR of course involves recognition of mispaired bases, but correct repair critically depends on distinguishing the newly synthesized DNA strand from the parental DNA strand. After this distinction is made, the newly synthesized DNA including the incorrect base is removed and re-synthesized correctly from the parental template. In bacteria, the proteins required to accomplish MMR have been identified and their mechanistic roles in mismatch recognition, strand discrimination, error removal and re-synthesis have been defined. In eukaryotes, homologues of the bacterial MMR proteins have been identified. It is clear that the overall mechanism has been conserved; however, many more proteins are involved in eukaryotes and mechanistic roles have not yet been defined for all of them (reviewed in Kolodner & Marsischky, 1999; Hsieh, 2001). The availability of purified MMR proteins allows detailed dissection of their mechanistic functions. It has been determined that in eukaryotes the heterodimer complexes Msh2–Msh6 and Msh2–Msh3 are responsible for mismatch recognition (Kolodner & Marsischky, 1999; Hsieh, 2001). The role of additional eukaryotic MMR proteins is still being unravelled and SFM imaging has played an important role here.

After mismatch recognition, the steps of strand discrimination, error removal and re-synthesis have to occur in a co-ordinated fashion. SFM imaging has provided clues to the mechanistic role of the yeast Mlh1–Pms2 heterodimer, which acts downstream of mismatch recognition. Mlh1–Pms2 is known to interact with Msh2-containing complexes bound to mismatched bases, but models of MMR had not included Mlh1–Pms2 interactions with DNA. Biochemical and SFM imaging experiments did reveal high affinity and co-operative binding of yeast Mlh1–Pms2 to double-stranded DNA. Importantly, the SFM images showed long tracts of Mlh1–Pms2 bound to double-stranded DNA (Hall *et al.*, 2001; Drotschmann *et al.*, 2002). This sort of relatively non-specific DNA interaction is difficult to define by bulk biochemical assays but obvious by direct inspection of images. In addition, the SFM images showed that the long tracts of Mlh1–Pms2 often included the association of two separate regions of double-
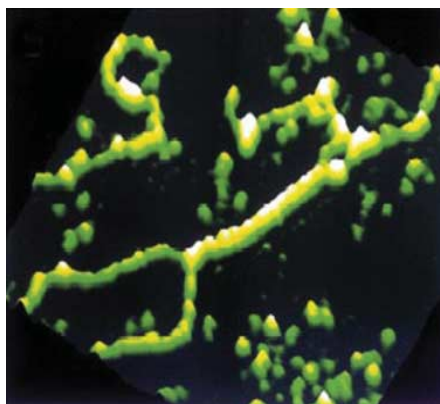


**Fig. 3.** The mismatch repair complex Mlh1–Pms2 binds cooperatively to DNA and can associate two regions of double-stranded DNA. An SFM image of the yeast Mlh1–Pms2 heterodimer bound to linear M13mp2 double-stranded DNA. Two double-stranded DNA molecules can be distinguished in the lower left corner and these are held together in a long tract of bound Mlh1–Pms2. The image is presented as a surface plot to emphasize topography. Height is represented by colour from dark to light, 0–3.5 nm. (Image courtesy of D. Erie and T. Kunkel.)

stranded DNA (Fig. 3). The association of two DNA strands could indicate that Mlh1–Pms2 has two DNA binding sites or that the arrangement in long tracts produces binding sites favouring association of multiple DNA strands. The mechanistic significance of this Mlh1–Pms2 cooperative binding to DNA and association of multiple DNA strands has not yet been defined, but it was suggested that this activity could facilitate strand discrimination by communication with the mismatch and other MMR factors (Hall *et al.*, 2001). These observations will surely spur new directions of investigation into the functional architecture of mammalian MMR complexes.

*DSB repair*

DNA DSBs, arising during genome duplication and from exogenous DNA damaging agents, are among the most toxic DNA lesions. Unrepaired DSBs can be lethal whereas misrepaired DSBs can cause chromosomal rearrangements, genome instability and eventually carcinogenesis. Eukaryotic cells primarily repair DSBs in one of two ways, homologous recombination and non-homologous end joining (reviewed in Kanaar *et al.*, 1998). Homologous recombination repairs DSBs accurately using the information from the undamaged sister chromatid or homologous chromosome. Non-homologous end joining rejoins ends directly in a manner that is more error-prone. Both pathways are complex processes, which involve the co-ordinated action of a large number of proteins. Homologous recombination requires proteins originally identified as the Rad52 epistasis

group of yeast and later found to be conserved in other species: Rad50, Rad51, Rad52, Rad54, Rad55, Rad57, Rad59, Mre11 and Xrs2 (Nbs1 in mammals) (Symington, 2002). The proteins specifically required for non-homologous end joining include Ku70/80, DNA-PK$_{cs}$, DNA ligase IV and XRCC4 (Critchlow & Jackson, 1998). In yeast, the Rad50/Mre11 complex is implicated in non-homologous end joining as well.

Repair of DSBs must begin by recognizing the ends and keeping them close for further processing. Several of the proteins in homologous recombination and non-homologous end joining are known to have important interactions with DNA ends. SFM imaging has helped to characterize and define these interactions as well as provide new ideas about the assembly of complexes that may be needed for DSB repair. One protein complex, the Rad50/Mre11 complex discussed above with respect to the architecture of the proteins, may be involved in both pathways. SFM images of human Rad50/Mre11 bound to DNA revealed structures that could account for a common role in these otherwise mechanistically distinct pathways (de Jager *et al.*, 2001). Rad50/Mre11 was seen to bind to DNA via its globular domain with the long coiled-coils protruding. Large oligomeric assemblies of many Rad50/Mre11 complexes were often observed on linear DNA but never on circular DNA. The large oligomeric DNA-bound Rad50/Mre11 complexes could tether different DNAs apparently via interaction of the protruding coiled-coil domains. It was suggested that this would provide a means of keeping broken DNA ends in close proximity to allow co-ordinated processing and eventual repair via either homologous recombination or non-homologous end joining. SFM images of yeast Rad50/Mre11/Xrs2 complex bound to DNA showed a different picture (Chen *et al.*, 2001). Here protein was observed bound to DNA ends or to internal positions believed to be junctions between short linear fragments. Large oligomers were not observed in this study although the short DNA substrate would have been almost completely obscured if oligomers were formed. In addition, the long coiled-coil structures of Rad50, apparently responsible for tethering DNA by the human protein, were not resolved in this study.

SFM imaging identified large DNA-bound oligomers of Rad50/Mre11 as the functional form in tethering DNA and presumably important for biological activity. The DNA-bound Rad50/Mre11 oligomers are difficult to characterize or quantify by bulk biochemical assays because of their irregular nature and large size. Therefore, SFM imaging was used to test the influence of ATP on formation of DNA-bound Rad50/Mre11 oligomers. Rad50 includes an ATPase activity but the effect of ATP on Rad50 function had not been defined. The formation of large oligomeric complexes was quantified on DNA substrates with different end structures in the presence of ATP or non-hydrolysable ATP analogues. In this way, it was shown that the preference of Rad50/Mre11 to form oligomers on DNA with different end structures was influenced by ATP binding (de Jager *et al.*, 2002). The mechanistic significance of

this ATP-induced change in end binding preference still needs to be defined. One possibility is that it influences the nuclease activity of the Mre11 subunit at different end structures, believed to be important for non-homologous end joining reactions (Hopfner *et al.*, 2001; de Jager *et al.*, 2002).

The other DSB repair proteins that interact with DNA ends, Ku70/80 and DNA-PK$_{cs}$, are involved exclusively in non-homologous end joining and not homologous recombination. Biochemical studies and electron microscopy visualization of DNA bound by Ku70/80 had shown many years ago that this protein needed an end to bind but did not remain at an end (de Vries *et al.*, 1989). This was confirmed by SFM and extended to analyse the interaction of DNA-bound Ku70/80 with DNA-PK$_{cs}$. An SFM study reported frequent end-joining events in the presence of Ku70/80 and DNA-PK$_{cs}$, 16–23% of protein-bound DNA, but the images presented were not very easy to interpret and non-specific aggregations may have occurred (Cary *et al.*, 1997). In another study, it was observed that DNA-PK$_{cs}$ and Ku70/80 could bind separately to DNA, and once bound they could interact to form a complex at a DNA end (Yaneva *et al.*, 1997). This study took advantage of the quantitative 3-D information of SFM images to determine that the size and shape of the protein complex at DNA ends most likely represented DNA-PK$_{cs}$ at the DNA end with an adjacent DNA-internal Ku70/80. In the images presented, DNA ends were joined inter- and intramolecularly in the presence of DNA-PK$_{cs}$ with and without Ku70/80, but this was not quantified. Although not specifically analysed, joining of ends by Ku70/80 alone was not prominent here or in previous electron microscopy studies (de Vries *et al.*, 1989).

In mammals, the Ligase IV-XRCC4 complex is responsible for the final step in non-homologous end joining (Critchlow & Jackson, 1998). The interaction of the Ligase IV-XRCC4 complex with DNA ends with and without Ku70/80 or DNA-PK$_{cs}$ has also been studied by a combination of biochemistry and SFM imaging (Chen *et al.*, 2000). SFM imaging revealed Ligase IV-XRCC4 bound to DNA ends and at the junction of two linear DNA fragments if their ends were complementary, a reassuring position for a ligase. Ligase IV-XRCC4 was also observed together with either Ku70/80 or DNA-PK$_{cs}$ at DNA ends. However, the functional significance of these two interactions cannot be equivalent because biochemical experiments in the same study indicated that Ku70/80 inhibited ligation by Ligase IV-XRCC4 whereas DNA-PK$_{cs}$ shifted the Ligase IV-XRCC4 ligation reaction to favour intermolecular products. The SFM analysis of DNA ends bound by increasingly complex assemblies of end-joining proteins will continue to be important for understanding the complete mechanism of this reaction.

## Proteins working on DNA

Many of the DNA repair pathways include steps that require dynamic changes in DNA structure or large-scale rearrangement and movement of proteins on DNA. It is not always necessary

actually to observe movement of molecules to understand their dynamic interactions. Static SFM images can reveal complicated features of molecules as evidence of dynamic reactions that have occurred. A few examples have already been described. The association of distant DNA sites by both Rad50/Mre11 and Mlh1–Pms2 and association of DNA ends by DNA-PK$_{cs}$ and Ku70/80 were obvious from inspection of SFM images (Yaneva *et al.*, 1997; de Jager *et al.*, 2001; Hall *et al.*, 2001). The dynamic process of DNA wrapping by the bacterial NER damage recognition complex involved ATP binding, and hydrolysis at specific steps was determined from the structures of static complexes formed in defined conditions and with informative mutant components (Verhoeven *et al.*, 2001, 2002).

The homologous recombination reaction that is responsible for accurate repair of DSBs also involves dramatic rearrangement of DNA molecules. The central step in homologous recombination is formation of a joint molecule between a broken DNA processed to a single-stranded end and a homologous sequence in the intact double-stranded template. Homologous recombination can be divided into a series of steps that probably occur in a co-ordinated fashion *in vivo*. In eukaryotic cells the broken DNA end is processed to expose a single-stranded region that is bound by Rad51 in a nucleoprotein filament. The structure of this nucleoprotein filament is much like its bacterial homologue, the RecA filament, for which there are some nice SFM images (Seitz *et al.*, 1998). The Rad51 single-stranded DNA filament then has to invade the double-stranded template and eventually basepair with its complementary strand in the template. This requires melting

of the template double-stand and possible removal of proteins bound to the template that would block the reaction. Rad54 is one of the accessory proteins in Rad51-meadiated joint molecule formation that probably plays a role at these steps. Rad54 can interact with a Rad51 single-stranded DNA filament (Jiang *et al.*, 1996; Clever *et al.*, 1997; Golub *et al.*, 1997; Tan *et al.*, 1999) and has ATPase activity, which is stimulated by double-stranded DNA (Petukhova *et al.*, 1998; Swagemakers *et al.*, 1998). Rad54 can use the energy of ATP hydrolysis to change DNA topology (Tan *et al.*, 1999; Mazin *et al.*, 2000; Van Komen *et al.*, 2000). Introducing superhelical stress into the template double-strand would help in joint molecule formation by either favouring template melting, removal of proteins bound to the template or both (Petukhova *et al.*, 1999; Tan *et al.*, 1999; Mazin *et al.*, 2000; Petukhova *et al.*, 2000; Van Komen *et al.*, 2000; Ristic *et al.*, 2001).

SFM imaging helped to elucidate the mechanism by which Rad54 uses ATP hydrolysis to change DNA topology (Ristic *et al.*, 2001). Images of DNA–Rad54 complexes formed after incubation of human Rad54 (hRad54) with singly nicked circular DNA show small proteins, presumably hRad54 monomers, bound to DNA in the absence of ATP. Addition of ATP to the binding reaction resulted in formation of much larger complexes. However, there was no evidence for protein-constrained supercoils, such as DNA wrapped around protein or DNA stretched in a protein filament. The large hRad54 complexes were observed anchoring the junction between relaxed and plectonemically supercoiled domains of the plasmid (Fig. 4). The occurrence of these structures was demonstrated
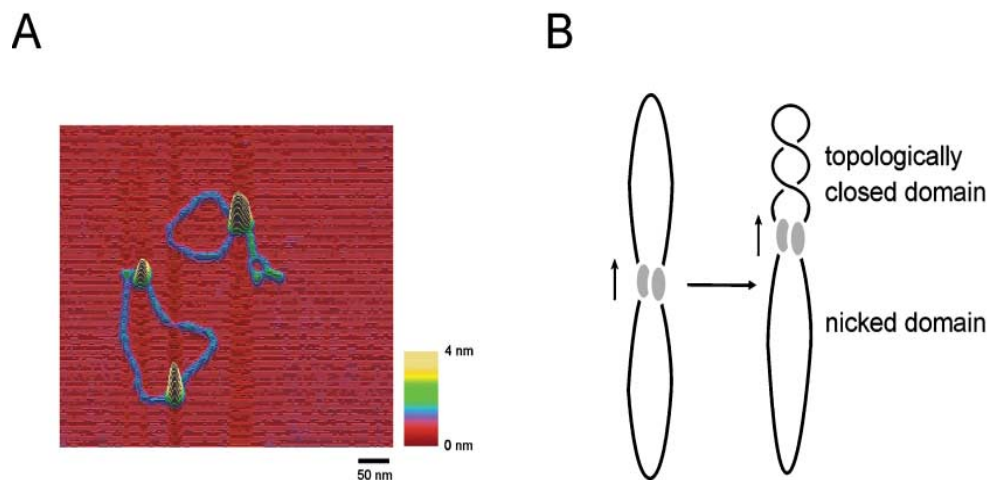


**Fig. 4.** The hRad54 protein anchoring a supercoiled domain in a singly nicked plasmid. (A) Human Rad54 was bound to a 2-kbp singly nicked plasmid in the presence of ATP before deposition for SFM imaging. The relaxed plasmid to the left has two hRad54 complexes bound. The plasmid to the right has a single large hRad54 complex bound at the junction between relaxed and supercoiled domains. (B) A diagram illustrating the association of two DNA-bound proteins (grey ovals). Movement of the associated proteins, indicated by the vertical arrow, along the DNA helix would create supercoils in plasmid domain without a nick.

to require ATP hydrolysis by hRad54. The creation of an hRad54 anchored supercoiled domain in these plasmids was interpreted to be the result of interaction between two DNA-bound hRad54 complexes and movement of one of them along DNA. Volume measurements of the presumably functional large hRad54 complexes bound to DNA in the presence of ATP indicate they are at least trimers and may be as large as hexamers.

## Perspective

Here we have reviewed studies that applied SFM imaging of molecules and complexes in order to better understand the mechanism of DNA repair reactions. Most of these studies used the SFM as a straightforward imaging tool and were able to reveal structural information that would have been difficult to obtain by other methods. Similar information, in some cases, could have been obtained by electron microscopy. However, SFM has the advantage of requiring much simpler sample preparation and being accessible to many more investigators. These advantages and the wealth of interesting protein–DNA structures still to be analysed will ensure that SFM imaging continues to play a productive role in molecular biology.

There are also challenges to overcome in order to learn even more from direct imaging of biomolecules and their functional complexes. As the functional assemblies become more complex and include more components, it will be necessary to develop methods to identify specifically individual proteins within these assemblies. The work reviewed here was all done with purified components combined in defined conditions. There is always the possibility that important components have been omitted. In order to avoid this problem it will also be necessary to develop methods to isolate material from cells or complex mixtures with purity and abundance sufficient for SFM imaging. The combination of SFM imaging with single molecule manipulation techniques, such as magnetic or optical tweezers, will also open new doors for understanding dynamic details of DNA repair and other vital genome transactions.

## Acknowledgements

## References

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

Bustamante, C., Keller, D. & Yang, G. (1993) Scanning force microscopy of nucleic acids and nucleoprotein assemblies. *Curr. Opin. Struct. Biol.* **3**, 363–372.

Cary, R.B., Peterson, S.R., Wang, J., Bear, D.G., Bradbury, E.M. & Chen, D.J. (1997) DNA looping by Ku and the DNA-dependent protein kinase. *Proc. Natl Acad. Sci. USA*, **94**, 4267–4272.

Chen, L., Trujillo, K., Ramos, W., Sung, P. & Tomkinson, A.E. (2001) Promotion of Dnl4-catalyzed DNA end-joining by the Rad50/Mre11/Xrs2 and Hdf1/Hdf2 complexes. *Mol. Cell*, **8**, 1105–1115.

Chen, L., Trujillo, K., Sung, P. & Tomkinson, A.E. (2000) Interactions of the DNA ligase IV-XRCC4 complex with DNA ends and the DNA-dependent protein kinase. *J. Biol. Chem.* **275**, 26196–26205.

Clever, B., Interthal, H., Schmuckli-Maurer, J., King, J., Sigrist, M. & Heyer, W.D. (1997) Recombinational repair in yeast: functional interactions between Rad51 and Rad54 proteins. *EMBO J.* **16**, 2535–2544.

Critchlow, S.E. & Jackson, S.P. (1998) DNA end-joining: from yeast to man. *Trends Biochem. Sci.* **23**, 394–398.

Drotschmann, K., Hall, M.C., Shcherbakova, P.V., Wang, H., Erie, D.A., Brownewell, F.R., Kool, E.T. & Kunkel, T.A. (2002) DNA binding properties of the yeast Msh2-Msh6 and Mlh1-Pms1 heterodimers. *Biol. Chem.* **383**, 969–975.

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

Friedberg, E.C. (2003) DNA damage and repair. *Nature*, **421**, 436–440.

Golub, E.I., Kovalenko, O.V., Gupta, R.C., Ward, D.C. & Radding, C.M. (1997) Interaction of human recombination proteins Rad51 and Rad54. *Nucl Acids Res.* **25**, 4106–4110.

Hall, M.C., Wang, H., Erie, D.A. & Kunkel, T.A. (2001) High affinity cooperative DNA binding by the yeast Mlh1-Pms1 heterodimer. *J. Mol. Biol.* **312**, 637–647.

Henderson, R.M., Schneider, S., Li, Q., Hornby, D., White, S.J. & Oberleithner, H. (1996) Imaging ROMK1 inwardly rectifying ATP-sensitive K+ channel protein using atomic force microscopy. *Proc. Natl Acad. Sci. USA*, **93**, 8756–8760.

Hickson, I.D. (2003) RecQ helicases: caretakers of the genome. *Nat. Rev. Cancer*, **3**, 169–178.

Hopfner, K.P., Karcher, A., Craig, L., Woo, T.T., Carney, J.P. & Tainer, J.A. (2001) Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell*, **105**, 473–485.

Hsieh, P. (2001) Molecular mechanisms of DNA mismatch repair. *Mutat. Res.* **486**, 71–87.

de Jager, M. & Kanaar, R. (2002) Genome instability and Rad50 (S): subtle yet severe. *Genes Dev.* **16**, 2173–2178.

de Jager, M., van Noort, J., van Gent, D.C., Dekker, C., Kanaar, R. & Wyman, C. (2001) Human Rad50/Mre11 is a flexible complex that can tether DNA ends. *Mol. Cell*, **8**, 1129–1135.

de Jager, M., Wyman, C., van Gent, D.C. & Kanaar, R. (2002) DNA end-binding specificity of human Rad50/Mre11 is influenced by ATP. *Nucl Acids Res.* **30**, 4425–4431.

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

Jiang, H., Xie, Y., Houston, P., Stemke-Hale, K., Mortensen, U.H., Rothstein, R. & Kodadek, T. (1996) Direct association between the yeast Rad51 and Rad54 recombination proteins. *J. Biol. Chem.* **271**, 33181–33186.

Kanaar, R., Hoeijmakers, J.H. & van Gent, D.C. (1998) Molecular mechanisms of DNA double strand break repair. *Trends Cell. Biol.* **8**, 483–489.

Kolodner, R.D. & Marsischky, G.T. (1999) Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.* **9**, 89–96.

de Laat, W.L., Jaspers, N.G. & Hoeijmakers, J.H. (1999) Molecular mechanism of nucleotide excision repair. *Genes Dev.* **13**, 768–785.

Mazin, A.V., Bornarth, C.J., Solinger, J.A., Heyer, W.D. & Kowalczykowski, S.C. (2000) Rad54 protein is targeted to pairing loci by the Rad51 nucleoprotein filament. *Mol. Cell*, **6**, 583–592.

Melby, T.E., Ciampaglio, C.N., Briscoe, G. & Erickson, H.P. (1998) The symmetrical structure of structural maintenance of chromosomes (SMC) and MukB proteins: long, antiparallel coiled coils, folded at a flexible hinge. *J. Cell. Biol.* **142**, 1595–1604.

van Noort, J., Orsini, F., Eker, A., Wyman, C., de Grooth, B. & Greve, J. (1999) DNA bending by photolyase in specific and non-specific complexes studied by atomic force microscopy. *Nucl. Acids Res.* **27**, 3875–3880.

van Noort, J., van der Heijden, T., de jager, M., Wyman, C., Kanaar, R. & Dekker, C. (2003) The coiled-coil of the human Rad50 DNA repair protein contains specific segments of increased flexibility. *Proc. Natl Acad. Sci. USA*, **100**, 7581–7586.

van Noort, S.J., van der Werf, K.O., Eker, A.P., Wyman, C., de Grooth, B.G., van Hulst, N.F. & Greve, J. (1998) Direct visualization of dynamic protein–DNA interactions with a dedicated atomic force microscope. *Biophys. J.* **74**, 2840–2849.

Petukhova, G., Stratton, S. & Sung, P. (1998) Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins. *Nature*, **393**, 91–94.

Petukhova, G., Sung, P. & Klein, H. (2000) Promotion of Rad51-dependent ᴅ-loop formation by yeast recombination factor Rdh54/Tid1. *Genes Dev.* **14**, 2206–2215.

Petukhova, G., Van Komen, S., Vergano, S., Klein, H. & Sung, P. (1999) Yeast Rad54 promotes Rad51-dependent homologous DNA pairing via ATP hydrolysis-driven change in DNA double helix conformation. *J. Biol. Chem.* **274**, 29453–29462.

Ratcliff, G.C. & Erie, D.A. (2001) A novel single-molecule study to determine protein–protein association constants. *J. Am. Chem. Soc.* **123**, 5632–5635.

Ristic, D., Wyman, C., Paulusma, C. & Kanaar, R. (2001) The architecture of the human Rad54-DNA complex provides evidence for protein translocation along DNA. *Proc. Natl Acad. Sci. USA*, **98**, 8454–8460.

Rivetti, C., Guthold, M. & Bustamante, C. (1996) Scanning force microscopy of DNA deposited onto mica: equilibrium versus kinetic trapping studied by statistical polymer chain analysis. *J. Mol. Biol.* **264**, 919–932.

Sancar, A. (1994) Structure and function of DNA photolyase. *Biochemistry*, **33**, 2–9.

Seitz, E.M., Brockman, J.P., Sandler, S.J., Clark, A.J. & Kowalczykowski, S.C. (1998) RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev.* **12**, 1248–1253.

Strunnikov, A.V. & Jessberger, R. (1999) Structural maintenance of chromosomes (SMC) protiens. Conserved molecular properties for multiple biological functions. *Eur. J. Biochem.* **263**, 6–13.

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

Swagemakers, S.M., Essers, J., de Wit, J., Hoeijmakers, J.H. & Kanaar, R. (1998) The human RAD54 recombinational DNA repair protein is a double-stranded DNA-dependent ATPase. *J. Biol. Chem.* **273**, 28292–28297.

Symington, L.S. (2002) Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol. Mol. Biol. Rev.* **66**, 630–670.

Tan, T.L., Essers, J., Citterio, E., Swagemakers, S.M., de Wit, J., Benson, F.E., Hoeijmakers, J.H. & Kanaar, R. (1999) Mouse Rad54 affects DNA conformation and DNA-damage-induced Rad51 foci formation. *Curr. Biol.* **9**, 325–328.

Van Komen, S., Petukhova, G., Sigurdsson, S., Stratton, S. & Sung, P. (2000) Superhelicity-driven homologous DNA pairing by yeast recombination factors Rad51 and Rad54. *Mol. Cell*, **6**, 563–572.

Verhoeven, E.E., Wyman, C., Moolenaar, G.F. & Goosen, N. (2002) The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands. *EMBO J.* **21**, 4196–4205.

Verhoeven, E.E., Wyman, C., Moolenaar, G.F., Hoeijmakers, J.H. & Goosen, N. (2001) Architecture of nucleotide excision repair complexes: DNA is wrapped by UvrB before and after damage recognition. *EMBO J.* **20**, 601–611.

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

de Vries, E., van Driel, W., Bergsma, W.G., Arnberg, A.C. & van der Vliet, P.C. (1989) HeLa nuclear protein recognizing DNA termini and translocating on DNA forming a regular DNA-multimeric protein complex. *J. Mol. Biol.* **208**, 65–78.

Wood, R.D. (1999) DNA damage recognition during nucleotide excision repair in mammalian cells. *Biochimie*, **81**, 39–44.

Wyman, C. & Kanaar, R. (2002) Chromosome organization: reaching out to embrace new models. *Curr. Biol.* **12**, R446–R448.

Wyman, C., Rombel, I., North, A.K., Bustamante, C. & Kustu, S. (1997) Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein. *Science*, **275**, 1658–1661.

Xue, Y., Ratcliff, G.C., Wang, H., Davis-Searles, P.R., Gray, M.D., Erie, D.A. & Redinbo, M.R. (2002) A minimal exonuclease domain of WRN forms a hexamer on DNA and possesses both 3′–5′ exonuclease and 5′-protruding strand endonuclease activities. *Biochemistry*, **41**, 2901–2912.

Yaneva, M., Kowalewski, T. & Lieber, M.R. (1997) Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies. *EMBO J.* **16**, 5098–5112.

# Chapter 3

## DNA bending by the human damage recognition complex XPC-HR23B

# DNA bending by the human damage recognition complex XPC–HR23B

Ana Janićijević [a], Kaoru Sugasawa [b,c], Yuichiro Shimizu [b,d], Fumio Hanaoka [b,c,d], Nils Wijgers [a], Miodrag Djurica [e], Jan H.J. Hoeijmakers [a], Claire Wyman [a,f,*]

[a] *Department of Cell Biology and Genetics, Erasmus MC, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands*
[b] *Cellular Physiology Laboratory, RIKEN (The Institute of Physical and Chemical Research),*
*2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[c] *CREST, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[d] *Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamada-oka, Suita, Osaka 565-0871, Japan*
[e] *KPN Research, P.O. Box 421, 2260 AK Leidschendam, The Netherlands*
[f] *Department of Radiation Oncology, Erasmus MC-Daniel, Rotterdam, The Netherlands*

## Abstract

Genome integrity is maintained, despite constant assault on DNA, due to the action of a variety of DNA repair pathways. Nucleotide excision repair (NER) protects the genome from the deleterious effects of UV irradiation as well as other agents that induce chemical changes in DNA bases. The mechanistic steps required for eukaryotic NER involve the concerted action of at least six proteins or protein complexes. The specificity to incise only the DNA strand including the damage at defined positions is determined by the coordinated assembly of active protein complexes onto damaged DNA. In order to understand the molecular mechanism of the NER reactions and the origin of this specificity and control we analyzed the architecture of functional NER complexes at nanometer resolution by scanning force microscopy (SFM). In the initial step of damage recognition by XPC-HR23B we observe a protein induced change in DNA conformation. XPC–HR23B induces a bend in DNA upon binding and this is stabilized at the site of damage. We discuss the importance of the XPC–HR23B-induced distortion as an architectural feature that can be exploited for subsequent assembly of an active NER complex.
© 2002 Elsevier Science B.V. All rights reserved.

*Keywords:* Nucleotide excision repair; Damage recognition; Scanning force microscopy; XPC; DNA bending

## 1. Introduction

Nucleotide excision repair (NER) eliminates of a wide variety of lesions from DNA thereby protecting

---

*Abbreviations:* NER, nucleotide excision repair; GGR, global genome repair; TCR, transcription coupled repair; SFM, scanning force microscopy; CFP, cyan fluorescent protein

* Corresponding author. Tel.: +31-10-4088337;
fax: +31-10-4089468.
*E-mail address:* wyman@ch1.fgg.eur.nl (C. Wyman).

virtually all organisms from the severe consequences of DNA injuries including cell death and mutagenesis. In humans, impaired NER activity is responsible for the sensitivity to sunlight, increased incidence of skin cancer and neurodegeneration experienced by xeroderma pigmentosum (XP) patients. The different genes involved in XP, *XPA* through *XPG*, encode products that are required for proper NER. Mechanistically the complete NER reaction can be separated into a series of distinct steps: (1) damage recognition, (2) damage

demarcation and commitment to incision, (3) dual incisions around the lesion and removal of the damaged oligonucleotide, and finally (4) gap filling DNA synthesis and ligation to completely restore the correct DNA sequence [1]. Specific functions in this pathway have been identified for many of the required proteins. We can now begin to address questions about their molecular mechanisms of action at each step and the necessary coordination.

Recognizing damaged DNA is the required first step in NER and mechanistically very intriguing [2]. It is remarkable that NER is responsible for repair of a wide variety of chemically distinct DNA lesions, including cyclobutane pyrimidine dimers and (6–4) photoproducts caused by exposure to UV irradiation as well as bulky adducts caused by chemical agents. Not only do many chemically distinct DNA lesions have to be recognized in order to initiate repair but inappropriate repair must be avoided at all other sites. There are two subpathways of NER, transcription coupled repair (TCR) and global genome repair (GGR), that employ different methods of damage recognition. A stalled RNA polymerase is the damage sensing signal that initiates repair on the transcribed strand of active genes in TCR. The *XPC* gene product is dispensable for TCR but required for effective GGR which is responsible for repair in the rest of the genome. This suggests that *XPC* plays an important role in GGR damage recognition.

The human *XPC* gene product is a 125 kDa protein found in vivo in complex with HR23B, a 58 kDa protein homologous to *Saccharomyces cerevisiae* Rad23 [3], and centrin 2, an 18 kDa centrosome component [4]. The two component complex including XPC and HR23B is sufficient for complementation of XPC deficiency in a cell free repair reaction [3] and reconstitution of the GGR reaction in vitro [5,6]. XPC accumulates in vivo at sites of DNA damage and is required for the subsequent accumulation of the other NER proteins, indicating an obligatory role early in NER [7]. Order of addition experiments have defined XPC–HR23B as the necessary and sufficient first factor for initiation of NER on damaged substrates in vitro [8]. Purified XPC–HR23B also binds preferentially to a variety of DNA lesions such as UV-induced (6–4) photoproducts, AAF adducts, and artificial cholesterol moieties [8–11]. In addition, XPC–HR23B is the only NER factor to bind

damaged DNA with sufficient specificity and affinity to be detected in DNase footprinting experiments. Interestingly, though XPC–HR23B binding initiates repair, this alone is not sufficient to assure progress to a complete NER reaction [9].

The XPA protein is absolutely required for both TCR and GGR and was previously thought to be a damage sensor. Although XPA binds with a slight preference to damaged DNA in vitro it does not display sufficient discrimination to account for the in vivo accuracy of NER [12–15]. Intriguingly, XPA has a preference for binding to artificially distorted DNA which increases substantially when it is combined with RPA [16]. These artificially distorted substrates, including a kinked cis-Pt adduct as well as three- and four-way junctions, share the architectural feature of presenting two strands emerging from a central bend. Thus it is possible that the XPA–RPA complex recognizes specific architectural features of DNA, perhaps induced by damage recognition proteins, as a necessary step in both GGR and TCR. XPA interacts with many other NER proteins [8–10] and is likely required after initial damage recognition in a step that is common to both GGR and TCR.

A specific architectural arrangement of the required proteins on damaged DNA would account for the precision of the NER reaction. A functional architecture would be achieved through orchestrated changes in both DNA and protein arrangement. The first changes are expected to occur upon damage induction in DNA and its recognition by NER proteins. DNA lesions that are repaired by NER are often described as helix distorting [17], however, the means by which this distortion is exploited by the NER machinery have not been defined. We set out to identify changes in DNA structure that might be stabilized or induced upon damage recognition by XPC–HR23B binding. We used scanning force microscopy (SFM) to directly observe NER proteins bound to a DNA substrate with a single damaged base at a defined position. Protein bound to the damaged site and to non-damaged sites could be identified and both types of complexes were characterized. We observe that XPC–HR23B bound preferentially to the site of damage and that XPC–HR23B induces a bend in DNA both at damaged and non-damages sites. XPA did not exhibit a strong preference for binding to the damaged site but did induce a bend in DNA upon binding. The

importance of this altered DNA architecture for subsequent progression of the NER reaction is discussed.

## 2. Materials and methods

### 2.1. Preparation of DNA containing a defined cholesterol lesion

The damaged DNA substrates were made essentially as described [18] and shown schematically in Fig. 1, with the following modifications. A shorter PCR fragment was designed using biotinylated primer 5′ biotin (TTTCCCGGGGGGGCCCGGGTTCTAT-ACTGTTGACC) to synthesize a bottom-strand of 812 bp. The PCR template was a cloned portion of
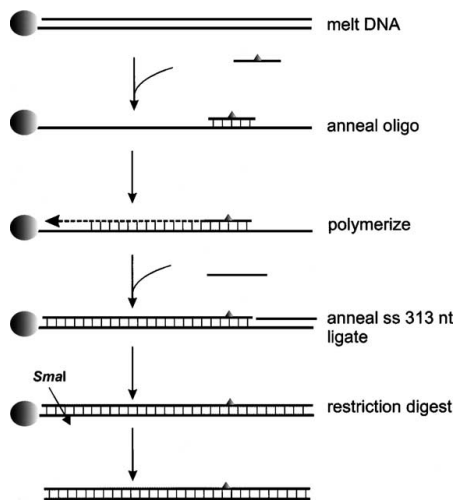


Fig. 1. Outline of the synthesis of a DNA substrate with a single NER lesion at a defined position. A PCR product with a 5′ biotin on one strand is bound to streptavidin coated magnetic beads. The top, non-biotinylated, strand is melted off with NaOH and removed. Then an oligonucleotide including the cholesterol lesion is annealed to the single stranded DNA on the beads, the top strand is synthesized toward the bead using the oligo as a primer. The remaining part of the top strand, obtained by melting and neutralizing a top strand from a shorter biotinylated PCR product, is annealed and ligated to make a complete double stranded DNA. DNA is released from the beads by digestion with a restriction enzyme.

the *S. cerevisiae URA-3* gene. PCR products were purified with GFX$^{TM}$ column (Amersham). The isolated fragment was incubated with paramagnetic streptavidin-coated beads (Dynal M-280) in BW buffer (1 M NaCl, 5 mM Tris–HCl pH 7.5 and 0.5 mM EDTA) for 1 h at room temperature. The top strand was melted by incubating the bead bound DNA in 0.1 M NaOH for 3 min. The cholesterol containing oligo CH22 was purchased from Eurogentec and included a cholesterol-A moiety attached via propanol to the backbone at the position of nucleotide 11. Oligo CH22 was annealed to single stranded DNA (bottom-strand) attached to the beads in Sequenase reaction buffer RB (40 mM Tris–HCl pH 7.5, 20 mM MgCl$_2$ and 50 mM NaCl). After hybridization, the immobilized fragment was incubated with 0.2 U of Sequenase Version 2.0 DNA polymerase (USB) and dNTPs. To complete the double strand a 312-nt fragment was isolated as the melted top-strand from another bead bound biotinylated PRC product as described [18] and annealed to the partially double-stranded DNA on the beads, followed by ligation to seal the nick. The final DNA product of 796 bp with the cholesterol moiety 325 bp from one end was released from the beads by *Sma*I digestion.

To produce a DNA fragment with a centrally located cholesterol moiety we used different PCR primer for synthesis of the bottom-strand, CenCH oligo (5′ biotin—TTTCCCGGGGGGCCCGTCAACAGT-ACCCTTAGTATATTC). Other steps were performed as described above. The final DNA fragment was 649 bp in size with the cholesterol at nucleotide position 325.

### 2.2. Protein preparation

XPC–HR23B was prepared as previously described [9]. His$_9$HA–eCFP tagged XPA was produced and purified essentially as described for eCFP–XPA [19]. Briefly, recombinant protein expression in *Escherichia coli* was induced by IPTG, after which the cells were collected and lysed by sonication. The clarified cell extract was loaded onto a Heparin–Sepharose column equilibrated with buffer A (PBS, 0.01% NP40, 1 mM β-mercaptoethanol, 10% glycerol) and eluted a linear gradient of buffers A to B (buffer A + 1 M NaCl). Peak fractions were pooled and subsequently bound to 1 ml nickel–nitrilotriacedic acid agarose beads

(Ni–NTA, Qiagen) (pre-washed with PBS, 5 mM imidazole), washed with five times column volume of buffer C (25 mM Hepes, 0.01% NP40, 10% glycerol, pefabloc, 1 mM β-mercaptoethanol, 1 M NaCl) containing, respectively, 5, 20 and 60 mM imidazole. The eCFP–XPA was eluted with buffer C containing 200 mM imidazole.

### 2.3. Protein–DNA binding reaction

XPC–HR23B was bound to DNA in reaction mixtures with a final volume of 10 µl, containing 140 mM KCl, 40 mM HEPES, pH 7.5, 10 mM MgCl$_2$, 0.01% NP-40, 1 mM DTT, 10% glycerol, 110 mM XPC–HR23B and 30 mM DNA. Reaction mixtures were incubated at room temperature for 30 min. Samples of the reaction (0.5 µl) were prepared for SFM imaging by diluting into 14 µl deposition buffer (5 mM HEPES, pH 7.5 and 10 mM MgCl$_2$), and spotted onto freshly cleaved mica. After about 1 min the mica was rinsed with distilled water (HPLC grade, Sigma) and dried in the stream of air. CFP–XPA was bound to DNA in the same reaction conditions described for XPC–HR23B and prepared for SFM in the same way.

### 2.4. Scanning force microscopy

The SFM images were obtained with a Nanoscope IIIa (Digital Instruments, Santa Barbara, CA) operating in tapping mode in air with silicon nanotips[TM] (NANOPROBE). Images (512 pixels × 512 pixels) were collected as 2 µm × 2 µm scans. The raw image data was processed only by flattening to remove background slope using NanoScope software.

### 2.5. Image analysis

Only DNA molecules that were completely visible on the image and could be unambiguously traced were taken for analysis. DNA contour lengths were measured using the Alex toolbox [20] in Matlab (The Math Works Inc.). The position of XPC was plotted as the shortest distance between protein and the end of the DNA. DNA bend angles were measured using Nanoscope IIIa software. The angles were defined by manually determining the path of the DNA segments exiting the protein within about 15 nm on each sides of the protein. Angle apex was chosen as the point

were the DNA path trajectories would intersect. By convention DNA bend angle is defined as deviation form straight. DNA molecules with protein bound at their end or within one persistence length from their end were not included in the analysis. The bend angle of DNA alone at the cholesterol site was determined using the DNA fragments with a cholesterol moiety positioned in the middle. A circular mask about the size of XPC–HR23B was placed over the image of DNA at its center and the DNA bend angle at this position was determined as described above.

## 3. Results

For SFM analysis of the architecture of DNA bound by XPC–HR23B we prepared a DNA substrate with a single lesion at the defined position. DNA fragments of 500–1000 bp are convenient for SFM imaging and analysis of DNA–protein complexes. This length of DNA is long enough to deposit efficiently onto mica for imaging and allow accurate length measurements. The method for producing the specifically damaged substrate is outlined in Fig. 1. Essentially it is a solid phase synthesis involving biotinylated PCR products attached to streptavidin magnetic beads. After denaturing the bead-bound DNA the top strand is replaced in two stages, first by annealing an oligo including the damaged site and polymerization from this oligo primer, then the remaining top strand is supplied as a long single-stranded complement which is annealed and ligated. The final product is an 796 DNA fragment with a cholesterol-A moiety incorporated 325 bp from one end. A sample from each batch of DNA product was end-labeled and checked by native and denaturing electrophoresis for correct length and complete ligation. The presence of the damage was confirmed by inhibition of a restriction enzyme that cuts the undamaged sequence. The cholesterol modification used is recognized by XPC–HR23B and excised in vitro in a manner indistinguishable from authentic NER lesions [10,21].

Purified XPC–HR23B was mixed with this DNA in reaction conditions that are compatible with a complete in vitro repair reaction and deposited onto mica for SFM imaging. Fig. 2A shows a field of molecules deposited from such a binding reaction. Several 796-bp DNA fragments can be seen, most
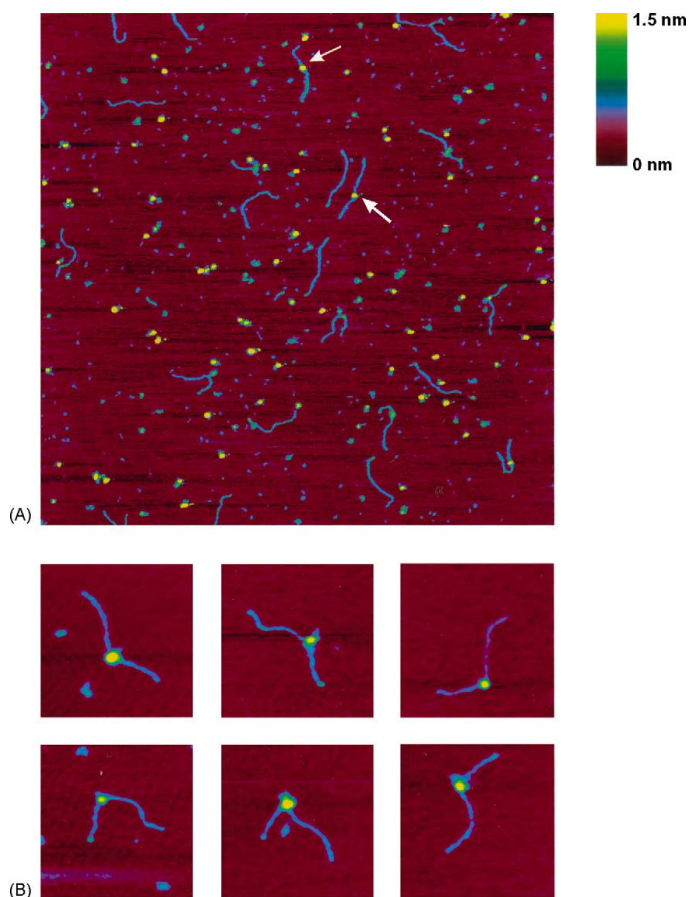
Fig. 2. (A) Scanning force microscope image of a field of molecules deposited from a DNA–XPC–HR23B binding reaction. The arrow indicates protein bound to DNA at the expected site of cholesterol damage. Scan area $2\,\mu m \times 2\,\mu m$. (B) Zoomed images of example of complexes with XPC bound at the expected position of cholesterol. Height is indicated by color as shown by the key to the right (0–1.5 nm, red to yellow).

of which do not have bound protein. The 185-kDa XPC–HR23B complex bound to DNA can easily be identified as a higher and wider feature along the DNA. There are proteins bound approximately at the position of the damaged site (arrow in Fig. 2A and zoomed views in Fig. 2B) and at other non-damaged sites. From one data set, 2012 DNA molecules were observed of which 17.4% had protein bound. The position of the bound protein along the DNA was measured and expressed as a proportion of the length of the fragment. This was determined by measuring the length of the DNA from the protein to the closest end and dividing this by the complete length of the same DNA molecule. The distribution of bound
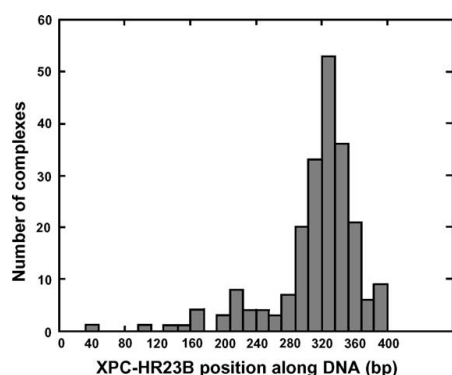
Fig. 3. Histogram of the position of XPC–HR23B binding along DNA. The cholesterol is 325 bp from one end of the DNA. The histogram is truncated at 400 bp because we measure the shortest distance from the protein to a DNA end. Most of the protein binds DNA at or near the damage site. The *X*-axis is length along the DNA in bp. The *Y*-axis is the number of molecules with protein bound in a given length window. Total number of molecules measured was 215.

protein along the DNA is shown in Fig. 3. Data in this histogram represents a conservative subset of the total DNA–protein complexes because it does not include those that could not be unambiguously traced or that had unusually large bound protein. Most of the protein is bound around the position corresponding to the expected site of the cholesterol moiety. To define a margin of error for our measurements we determined a standard deviation of the free DNA contour length (DNA not bound by protein measured from the same images, see below) and expressed this as a fraction of the total length (3%). Protein bound within the range of 315 to 335 bp (325 bp $\pm$ 3%) is assumed to be specifically bound to the cholesterol lesion. This includes 67% of the bound XPC–HR23B and accounts for 11.4% of the total DNA molecules analyzed having protein bound at the damaged site.

SFM images allow simultaneous analysis of multiple features of single protein–DNA complexes. We observe protein bound both to the damage site and non-damaged sites in the same reaction mixture and analyzed distortion of DNA in both cases. The bacterial NER protein UvrB, responsible for damage recognition, wraps DNA around itself both when bound

to damage and to non-damaged sites [18]. Wrapping of DNA is evident in SFM images of protein–DNA complexes as a decrease in the contour length of DNA bound by protein. The length of the DNA molecules with and without bound XPC–HR23B was very similar, having a mean length and standard deviation of $782 \pm 23$ bp for DNA alone and $774 \pm 40$ bp for DNA with bound XPC–HR23B. More than 200 molecules were measured in each class. These data reveal a mechanistic difference in DNA damage recognition between bacterial and eukaryotic NER. While UvrB wraps DNA around its surface, XPC–HR23B does not.

However, XPC–HR23B binding to DNA did appear to alter DNA structure by inducing a bend. To quantify this DNA distortion we measured the bend angle of DNA at the site of XPC–HR23B binding, both for protein bound to the site of the cholesterol modification and to non-damaged DNA. Analysis of bending by protein bound to non-damaged sites was done by selecting complexes with XPC–HR23B bound unambiguously away from the cholesterol site and pooling data collected from several independent binding reactions. The distributions of DNA bend angles with protein bound at the site of cholesterol and bound at non-damaged sites are shown in Fig. 4A and 4B respectively. Because we cannot distinguish the direction of bending from SFM images all angles were assigned positive values. The distribution of bend angles should be Gaussian but is effectively reflected at zero in the histograms presented. The true distribution can best be fitted to a modified Gaussian distribution as described in Schultz et al. [22]. A mean bend angle and its standard deviation were determined from fitting this modified Gaussian to the measured distributions. The standard deviation in these measurements is a representation of the collection of different arrangements of DNA and proteins in the complexes. This is related to the bending rigidity of DNA for measurements done in the absence of protein. For protein bound sites on DNA the standard deviation represents multiple states of the complexes, multiple projections of a bend onto the surface, DNA flexibility in the complex or all of these [23]. With XPC–HR23B bound to the site of cholesterol the DNA was bent by an average of $39 \pm 24°$ (Fig. 4A). With XPC–HR23B bound to non-damaged sites the DNA was bent by an average of $49 \pm 34°$ (Fig. 4B). These data demonstrate
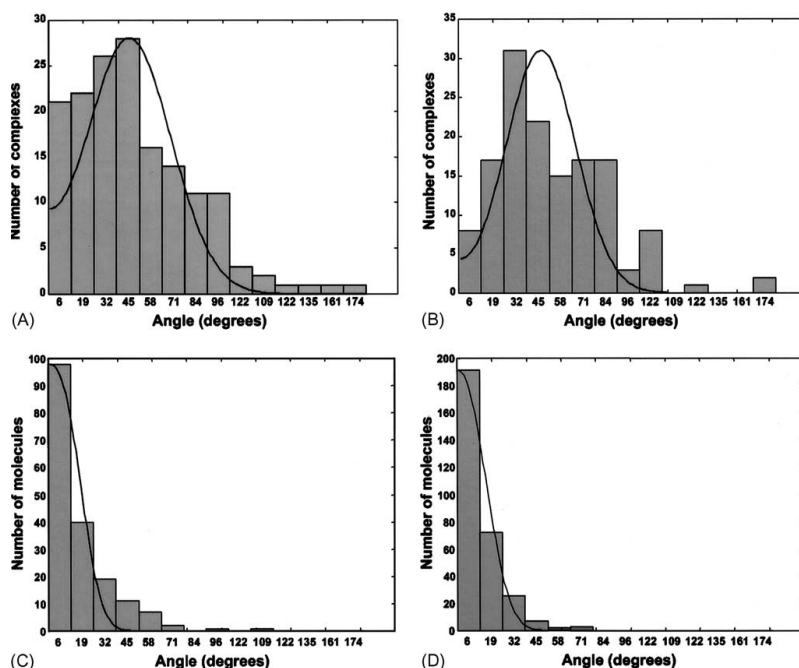
Fig. 4. DNA bend angle distributions. (A) Histogram of bend angle distribution for DNA with XPC–HR23B bound to site of cholesterol. A total of 158 molecules were measured. (B) Histogram of bend angle distribution for DNA with XPC–HR23B bound to non-cholesterol positions. A total of 141 molecules were measured. (C) Histogram of bend angle distribution of DNA alone at the site of cholesterol. A total of 179 molecules were measured. (D) Histogram of bend angle distribution of DNA alone at unmodified positions. A total of 300 molecules were measured. The solid line curves represent a fit of the data to a modified Gaussian distribution as defined in Eq. (2) of Schulz et al. [22].

that DNA bound by XPC–HR23B was significantly bent whether or not it included a lesion.

The cholesterol incorporated into our DNA fragment could itself induce a conformational change in the double helix. To check for this possibility, the bend angle of DNA alone at the position of the cholesterol moiety was measured. For this purpose a DNA fragment with the cholesterol exactly in the middle of the sequence was synthesized essentially as described above (Fig. 1). This fragment differs from the one used for XPC–HR23B binding only on one end where it is slightly shorter. The distribution of bend angles measured at this position (Fig. 4C) had a mean and standard deviation of $0 \pm 15°$. This is similar to the distribution of bend angles measured at non-damaged

positions on DNA (Fig. 4D) which had a mean and standard deviation of $0 \pm 14°$. This indicates that cholesterol alone does not induce any detectable bend in DNA. The standard deviation reflects the flexibility of DNA which was similar for the cholesterol site and other positions along this DNA fragment. This indicates that the cholesterol moiety alone did not change the flexibility of DNA. The change in bend angle observed when XPC–HR23B binds to DNA was induced by the protein at both damaged and non-damaged sites and was not intrinsic to the damaged DNA.

The XPA protein is essential for both GGR and TCR pathways of NER. It has only been shown under some conditions that this protein can bind with a slight preference to damaged DNA [12–15]. XPA is

required for in vitro repair of this cholesterol lesion [21] and therefore we tested its ability to recognize the cholesterol-damaged DNA used here. We used the same conditions as for XPC–HR23B binding, because these are compatible with a complete in vitro repair reaction. Even if XPA does not bind preferentially to this damage we can analyze binding induced distortions in DNA. In order to facilitate detection of XPA in SFM images, we used a CFP–XPA fusion. This protein is almost identical to GFP-tagged XPA which is functional as assayed by its ability to correct the

NER defect of XPA-deficient cells upon microinjection [19]. The fusion protein is larger, about 66 kDa compared to the 31 kDa XPA, and thus easier to detect bound to DNA in SFM images.

An SFM image of molecules deposited from a binding reaction including CFP–XPA and the cholesterol-damaged DNA is shown in Fig. 5A. There were several DNA molecules with no protein bound as well as DNA molecules with protein bound at non-damaged sites and possibly at the damaged site (Fig. 5A arrow). Objects larger than expected for a 66 kDa protein
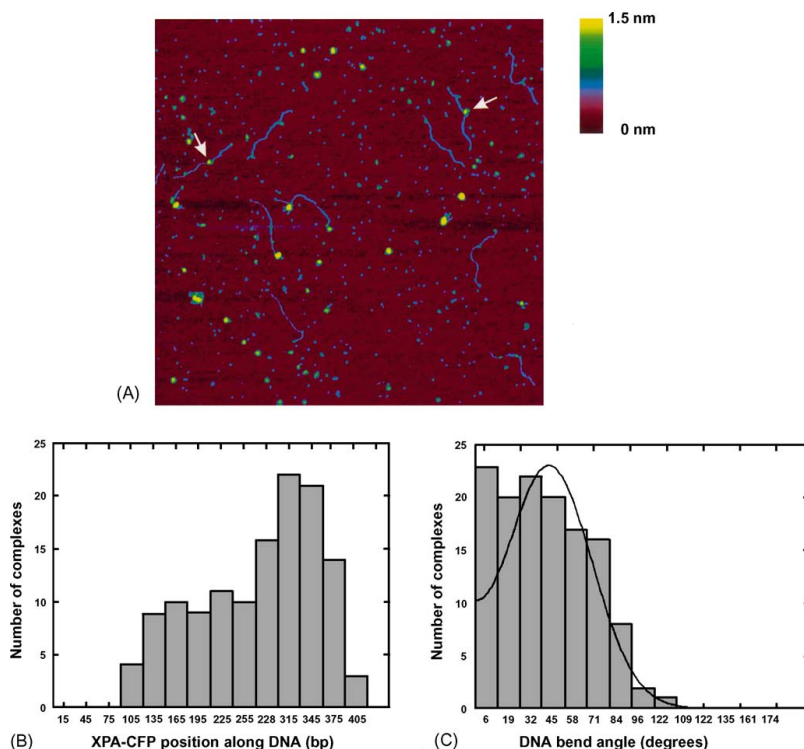


Fig. 5. CFP–XPA binding to cholesterol containing DNA. (A) Scanning force microscope image of a field of molecules deposited from a DNA–CFP–XPA binding reaction. The arrows indicates protein of the expected size bound to DNA. Scan area 1.4 μm × 1.4 μm. Height is indicated by color as shown by the key to the right (0–1.5 nm, red to yellow). (B) Histogram of the position of CFP–XPA binding along DNA. The cholesterol is 325 bp from one end of the DNA. The histogram is truncated at 400 bp because we measure the shortest distance from the protein to a DNA end. (C) Histogram of bend angle distribution for DNA with CFP–XPA bound at any position. A total of 129 molecules were measured. The solid line curve represents a fit of the data to a modified Gaussian distribution as defined in Eq. (2) of Schulz et al. [22].

(Fig. 5A, left center not DNA bound and right center DNA end bound) were probably the result of the known dimerization tendency of CFP. Complexes with these larger proteins were excluded from further measurements. The distribution of CFP–XPA binding position along the DNA is presented in Fig. 5B. The frequency of binding was fairly equal at all positions along the DNA, although there was a slightly increase at the expected site of damage. In addition binding reactions including RPA alone or CFP–XPA and RPA did not result in any preferential protein binding to the site of damage (data not shown). The apparently low frequency of CFP–XPC binding near the end of DNA reflects the exclusion of molecules with protein bound at an end from the measurements as well as the resolution limitations that prevent detecting very short segments of DNA extending beyond a bound protein. We did not observe the same preference for CFP–XPA binding to the damaged site as we did for XPC–HR23B.

The DNA bend angle at the site of CFP–XPA binding was also measured. Because CFP–XPA did not bind preferentially to the cholesterol site the number of complexes that could be classified as having protein bound to the site of damage was too low to give a significant distribution for this subset. Thus, the DNA bend angle distribution included protein bound at any position along the DNA (Fig. 5C). Fitting this data to a modified Gaussian distribution as described above gave an XPA induced DNA bend angle with a mean and standard deviation of $42 \pm 26°$. Thus, XPA induced a bend in DNA upon binding. However, it did not

preferentially bind to damaged DNA. The results of the bend angle measurements are summarized in Table 1.

## 4. Discussion

Damage recognition is the required first step in any DNA repair pathway. For NER the mechanism of damage recognition must exploit features of damaged DNA other than the chemical nature of the damaged bases, since this process removes a variety of lesions that do not share any obvious chemical similarity. The XPC–HR23B complex is responsible for the initial damage recognition step in GGR whereas XPA is required at a step common to both GGR and TCR. Here we have characterized XPC–HR23B/DNA complexes and XPA/DNA complexes by SFM. XPC–HR23B binds preferentially to the site of damage and induces a bend of 39–49° in the DNA whether or not it is bound to a site of damage. XPA on the other hand binds to DNA without any obvious preference for the site of damage in the conditions used here which were chosen to be compatible with a complete in vitro NER reaction. XPA binding induces a bend of 42° in the DNA. The ability to analyze DNA bending by proteins bound to cognate and non-cognate sites is one of the advantages of single molecule analysis such as can be done by SFM. Indeed SFM studies were some of the first to reveal and analyze DNA bending by sequence specific binding proteins at both specific and non-specific sites [23,24]. The preference for XPC–HR23B binding to the damaged site is likely in part due to damaged DNA more easily accommodating the distortion induced by protein binding. The complex formed at the damaged site would therefore be more stable than those formed on undamaged DNA. This is consistent with the observation that the ability to sequester human NER factors correlated with thermodynamic instability of a variety of DNA lesions [17]. XPA also induces a distortion in DNA however this complex is apparently no more stable at the site of damage than at non-damaged sites on DNA.

Commitment to the complete NER reaction requires more than XPC–HR23B binding to DNA as indicated by the observation that XPC–HR23B binds well to DNA with a small region of unpaired bases but this DNA is not processed further in in vitro NER reactions [9]. XPC–HR23B recognizes DNA sites

Table 1
Summary of DNA bend angle determinations

| Protein bound | At position of cholesterol (°) | At other positions (°) |
|---|---|---|
| XPC–HR23B | $39 \pm 24$ | $49 \pm 36$ |
| CFP–XPA[a] | – | $42 \pm 26$ |
| None | $0 \pm 15$ | $0 \pm 14$ |

DNA bend angles are the mean and standard deviation derived from the modified Gaussian fit to the distributions as described in the text and defined in Eq. (2) of Schulz et al. [22]. At least 129 molecules were measured for each distribution.

[a] There were insufficient CFP–XPA proteins bound at the position of cholesterol for a significant distribution of this class. The bend angle distribution was determined for all CFP–XPA molecules of the expected monomer size bound at any position on the DNA.

that are candidates for NER by forming a complex with damaged DNA in which the double helix is distorted, evident as a bend in SFM imaging. This bent DNA/XPC–HR23B complex would then be the substrate for interaction with subsequent NER factors. XPA is required for NER, both GGR and TCR, at a stage described as damage verification, though preferential interaction with damaged DNA has not been demonstrated. XPA and the XPA–RPA complex do preferentially bind to artificial DNAs that share the feature of two DNA strands projecting from a bend [16]. XPA or XPA–RPA might have a preference for binding DNA bent by XPC–HR23B. However, we observed no difference in XPC–HR23B complexes on DNA when the binding reaction included XPC–HR23B and XPA or XPC–HR23B, XPA and RPA. The protein–DNA complexes observed by SFM from all of these binding reactions were similar with respect to the measured frequency of proteins bound to the site of damage, the overall size of the protein complexes bound to DNA and obvious changes in DNA conformation (data not shown). Recent fluorescence anisotropy studies have also failed to show any evidence for ternary interaction between DNA-bound XPC–HR23B and XPA [11]. Thus there is no evidence for direct XPA recognition of the bent DNA complex bound by XPC–HR23B.

The NER factor most likely to interact with DNA bound XPC–HR23B is TFIIH [25,26]. This large complex includes the XPD and XPB helicases that are believed to unwind the DNA around the damaged site. XPA has also been shown to preferentially bind to DNA that can be easily unpaired [27]. Taken together with other data on XPA binding these studies concluded that the interaction of XPA–RPA with DNA requires DNA bending and unwinding. Thus the DNA structure recognized by XPA for further processing in the NER pathway may only be produced after the action TFIIH has resulted in unwinding of DNA around the lesion. Here, it is interesting to note that XPA interacts specifically with several other NER factors, including TFIIH and the late acting structure specific endonucleases XPG and ERCC1–XPF [1]. These interactions are required for assembly of an active repair complex. Specific interactions between XPA and damaged DNA may not be sufficient or even necessary. Because XPA is a small protein with multiple interacting partners it may have an important role in organizing the architecture of the active NER complex. The composition of the nucleoprotein complex(s) including XPA and the arrangement of their components is an important subject for further analysis.

So far we have considered the order of steps in the architectural build up of an active NER complex at the damaged site in DNA. This architectural build up must also account for establishing asymmetry in the complex at some stage. The end result of NER is asymmetric relative to the damaged base in two respects. The damaged base defines the DNA strand on which both incisions are made. In addition incision occurs asymmetrically around the damaged base, 2–8 nt away on the $3'$ side and 15–24 nt on the $5'$ side [1]. Recently it has been shown that XPC–HR23B binds asymmetrically to DNA with respect to an inserted single-stranded loop structure [28]. However, the larger symmetric footprint of XPC–HR23B over a damaged sites suggests that it binds to damage in one of two possible asymmetric orientations, only one of which would be functional for further NER processes. A similar damage recognition scenario has recently been proposed based on SFM studies of the bacterial NER damage recognition complex. The UvrB damage recognition protein wraps DNA and this wrap is asymmetric, including about 70 bp of DNA exclusively from the $3'$ side of the damage [18]. Though DNA is wrapped around a UvrB monomer the protein is loaded onto DNA as a dimer to allow damage recognition in both strands [29]. The asymmetric XPC–HR23B footprint is maximally 15 nt longer on one side [28], a length difference that falls within the error of the SFM measurements of the 796 bp DNA we used here. So we cannot yet say if XPC bind asymmetrically with respect to damage or if there are two population of complexes.

If XPC–HR23B does not form a unique asymmetric complex on damaged DNA, asymmetry must be introduced at a later step. As the protein complex assembled at the damage site becomes larger and more DNA interacts with proteins asymmetry should become detectable by SFM. TFIIH, the NER factor most likely to interact with DNA-bound XPC–HR23B [25,26], is large and includes helicases with the ability to change DNA conformation in an ATP-dependent manner. Interestingly, TFIIH does include components with distinctly asymmetric interactions with damaged DNA. The yeast Rad3 protein, homologue

of the XPD helicase in TFIIH, is a 5′–3′ helicase and is specifically inhibited and sequestered in a stable complex by DNA damage only on the strand to which it binds [30,31]. Important mechanistic insights are likely to be gained by determining the architecture of the complex and possible asymmetry when TFIIH interacts with damage bound XPC–HR23B. We now know that the damaged DNA bound by the human NER factor XPC–HR23B is distortion and expect this distortion to be an important architectural feature required for subsequent assembly of an active NER complex. This XPC–HR23B-induced distortion may also contribute at some stage to the specific asymmetry of the functional complex required for accurate NER.

## Acknowledgements

## References

[1] W.L. de Laat, N.G.J. Jaspers, J.H. Hoeijmakers, Molecular mechanism of nucleotide excision repair, Genes Dev. 13 (1999) 768–785.

[2] R.D. Wood, DNA damage recognition during nucleotide excision repair in mammalian cells, Biochimie 81 (1999) 39–44.

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

[4] M. Araki, C. Masutani, M. Takemura, A. Uchida, K. Sugasawa, J. Kondoh, Y. Ohkuma, F. Hanaoka, Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair, J. Biol. Chem. 276 (2001) 18665–18672.

[5] A. Aboussekhra, M. Biggerstaff, M.K. Shivji, J.A. Vilpo, V. Moncollin, V.N. Podust, M. Protic, U. Hubscher, J.M. Egly, R.D. Wood, Mammalian DNA nucleotide excision repair reconstituted with purified protein components, Cell 80 (1995) 859–868.

[6] D. Mu, C.H. Park, T. Matsunaga, D.S. Hsu, J.T. Reardon, A. Sancar, Reconstitution of human DNA repair excision nuclease in a highly defined system, J. Biol. Chem. 270 (1995) 2415–2418.

[7] M. Volker, M.J. Moné, P. Karmakar, A. van Hoffen, W. Schul, W. Vermeulen, J.H. Hoeijmakers, R. van Driel, A.A. van Zeeland, L.H. Mullenders, Sequential assembly of the nucleotide excision repair factors in vivo, Mol. Cell 8 (2001) 213–224.

[8] K. Sugasawa, J.M. Ng, C. Masutani, S. Iwai, P.J. van der Spek, A.P. Eker, F. Hanaoka, D. Bootsma, J.H. Hoeijmakers, Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair, Mol. Cell 2 (1998) 223–232.

[9] K. Sugasawa, T. Okamoto, Y. Shimizu, C. Masutani, S. Iwai, F. Hanaoka, A multistep damage recognition mechanism for global genomic nucleotide excision repair, Genes Dev. 15 (2001) 507–521.

[10] R. Kusumoto, C. Masutani, K. Sugasawa, S. Iwai, M. Araki, A. Uchida, T. Mizukoshi, F. Hanaoka, Diversity of the damage recognition step in the global genomic nucleotide excision repair in vitro, Mutat. Res. 485 (2001) 219–227.

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

[12] P. Robins, C.J. Jones, M. Biggerstaff, T. Lindahl, R.D. Wood, Complementation of DNA repair in xeroderma pigmentosum group A cell extracts by a protein with affinity for damaged DNA, EMBO J. 10 (1991) 3912–3913.

[13] A.P. Eker, W. Vermeulen, N. Miura, K. Tanaka, N.G.J. Jaspers, Xeroderma pigmentosum group A correcting protein from calf thymus, Mutat. Res. 274 (1992) 211–224.

[14] C.J. Jones, R.D. Wood, Preferential binding of the xeroderma pigmentosum group A complementing protein to damaged DNA, Biochemistry 32 (1993) 12096–12104.

[15] T. Hey, G. Lipps, G. Krauss, Binding of XPA and RPA to damaged DNA investigated by fluorescence anisotropy, Biochemistry 40 (2001) 2901–2910.

[16] M. Missura, T. Buterin, R. Hindges, U. Hübscher, J. Kaspáková, V. Brabec, H. Naegeli, Double-check probing of DNA bending and unwinding by XPA–RPA: an architectural function in DNA repair, EMBO J. 20 (2001) 3554–3564.

[17] D. Gunz, M.T. Hess, H. Naegeli, Recognition of DNA adducts by human nucleotide excision repair. Evidence for a thermodynamic probing mechanism, J. Biol. Chem. 271 (1996) 25089–25098.

[18] E.E. Verhoeven, C. Wyman, G.F. Moolenaar, J.H. Hoeijmakers, N. Goosen, Architecture of nucleotide excision repair complexes: DNA is wrapped by UvrB before and after damage recognition, EMBO J. 20 (2001) 601–611.

[19] G.M.J. Segers-Nolten, C. Wyman, N. Wijgers, W. Vermeulen, A.T.M. Lenferink, J.H.J. Hoeijmakers, J. Greve, C. Otto, Scanning confocal fluorescence microscopy for single

molecule analysis of nucleotide excision repair complexes, Nucleic Acids Res. 30 (2002) 4720–4727.

[20] C. Rivetti, M. Guthold, C. Bustamante, Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis, J. Mol. Biol. 264 (1996) 919–932.

[21] D. Mu, D.S. Hsu, A. Sancar, Reaction mechanism of human DNA repair excision nuclease, J. Biol. Chem. 271 (1996) 8285–8294.

[22] A. Schulz, N. Mucke, J. Langowski, K. Rippe, Scanning force microscopy of *Escherichia coli* RNA polymerase-sigma54 holoenzyme complexes with DNA in buffer and in air, J. Mol. Biol. 283 (1998) 821–836.

[23] D.A. Erie, H.C. Schultz, C. Bustamante, DNA bending by Cro protein in specific and nonspecific complexes: implications for protein site recognition and specificity, Science 266 (1994) 1562–1566.

[24] J. van Noort, F. Orsini, A. Eker, C. Wyman, B. de Grooth, J. Greve, DNA bending by photolyase in specific and non-specific complexes studied by atomic force microscopy, Nucleic Acids Res. 27 (1999) 3875–3880.

[25] M. Yokoi, C. Masutani, T. Maekawa, K. Sugasawa, Y. Ohkuma, F. Hanaoka, The xeroderma pigmentosum group C protein complex XPC–HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA, J. Biol. Chem. 275 (2000) 9870–9875.

[26] S.J. Araujo, E.A. Nigg, R.D. Wood, Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome, Mol. Cell. Biol. 21 (2001) 2281–2291.

[27] N. Buschta-Hedayat, T. Buterin, M.T. Hess, M. Missura, H. Naegeli, Recognition of nonhybridizing base pairs during nucleotide excision repair of DNA, Proc. Natl. Acad. Sci. U.S.A. 96 (1999) 6090–6095.

[28] K. Sugasawa, Y. Shimizu, S. Iwai, F. Hanaoka, A molecular mechanism for DNA damage recognition by the xeroderma pigmentosum group C protein complex, DNA Repair 1 (2002) 95–107.

[29] E.E.A. Verhoeven, C. Wyman, G. F Moolenaar, N. Goosen, The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands, EMBO J. 21 (2002) 4196–4205.

[30] H. Naegeli, L. Bardwell, E.C. Friedberg, The DNA helicase and adenosine triphosphatase activities of yeast Rad3 protein are inhibited by DNA damage. A potential mechanism for damage-specific recognition, J. Biol. Chem. 267 (1992) 392–398.

[31] H. Naegeli, L. Bardwell, E.C. Friedberg, Inhibition of Rad3 DNA helicase activity by DNA adducts and abasic sites: implications for the role of a DNA helicase in damage-specific incision of DNA, Biochemistry 32 (1993) 613–621.

*Chapter 4*

DNA bending by photolyase
in specific and non-specific complexes
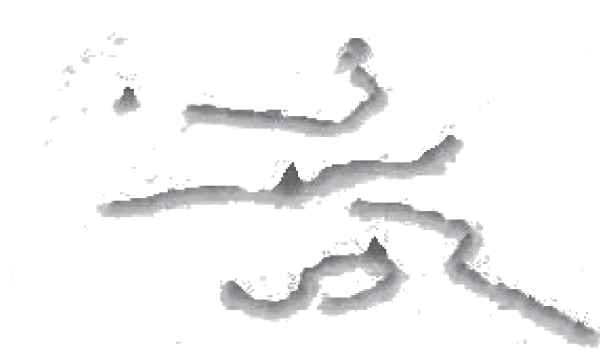
# DNA bending by photolyase in specific and
# non-specific complexes

## Abstract

DNA photolyases are monomeric proteins that use visible light to repair UV-induced DNA damage. A DNA substrate containing a single CPD was constructed and used to bind to photolyase. The architecture of protein-DNA complexes was analyzed by scanning force microscopy (SFM). This analysis shows that photolyase bends DNA by approximately 40 degrees in both specific and non-specific complexes.

## Introduction

Protein binding to DNA is often accompanied by DNA conformational changes. Studies of DNA mechanical properties, especially intrinsic curvature and flexibility, have provided important insights into how proteins function. As an example, in DNA repair in mammals, UV light induces chemical alterations in DNA bases resulting in a local distortion of the DNA helix that is recognized by UV-damage binding proteins. Binding of e.g. XPC protein to lesions induces additional DNA bending and subsequent recruitment of more DNA repair factors (Janicijevic, Sugasawa et al. 2003).

DNA distortions have been studied extensively. In the "classical" Watson-Crick model DNA is regarded as a molecule with a straight helical axis. Many relevant biological processes however, require bending of DNA induced by proteins (e.g. transcription, replication, recombination as well as chromosome condensation (Harrington 1992; van der Vliet and Verrijzer 1993; Ali, Amit et al. 2001; Dame 2005). Some DNA sequences (for example AT-rich sequences) contain intrinsically flexible regions. Evidence for DNA conformational changes comes from a variety of different experimental techniques. Indirect approaches to estimate DNA bending include gel electrophoresis and circularization kinetic experiments. Nuclear Magnetic Resonance (NMR) spectroscopy can provide information about the three-dimensional structure of proteins and nucleic acids at atomic resolution in solids and in solutions with. NMR's inherently low sensitivity requires that the molecule is soluble at sufficiently high concentration and is isotopically labeled. An important limitation of NMR is the size of a macromolecule or complex for full structure determinations where the upper limit is ~ 30 kDa for solution-state NMR. Recent advances in solid-state NMR methodology enabled the analysis of molecules with high molecular weight. X-ray crystallography provides atomic-level DNA structural information, but only one state of DNA present in a given crystal can be determined by this method. Imaging of single molecules by electron microscopy (EM), scanning force microscopy (SFM) or time-resolved fluorescent measurements can yield the mean DNA bend angle and distribution of DNA

bend angles. In addition, SFM and EM allow statistical analysis of a large number of individual molecules and can provide information on the morphological properties and structural changes in different populations of DNAs (Hansma and Hoh 1994; Seong, Kobatake et al. 2002). Unlike EM, which requires a vacuum, and special preparation of biological samples, SFM works in air or in liquid with little or no sample preparation and thus is easier to handle and can be analysed in physiologically relevant conditions when imaging in liquid.

From SFM images, bending of DNA has been determined by several methods. In the trace trajectory method, the bend angle is defined directly as the angle formed between vectors determined by DNA trajectories manually traced from the site of interest. The other two methods are based on a 'wormlike chain' (WLC) model used in polymer statistics to describe the conformation of double stranded DNA molecules in solution (Schellman 1974). In the WLC model, DNA is considered as a semi-flexible rod in which local bends affect the end-to-end distance (EED), the shortest path between two DNA ends (Rivetti, Walker et al. 1998). To determine protein-induced DNA bending Rivetti *et al.* used the mean-squared EED value of a population of molecules. An alternative method, developed by Dame *et al*. (Dame, van Mameren et al. 2005), takes into account the EED distribution instead of the mean value in order to determine DNA bend angles.

Photolyase is an enzyme that uses energy from visible light to repair UV-induced damages. The most abundant UV-damages in DNA are cyclobutane pyrimidine dimers (CPD), which are formed between two adjacent pyrimidine bases (Friedberg, Walker et al. 2005). CPD-photolyase binds to the damaged DNA in a light-independent manner, uses near-UV or visible light to reverse a CPD into the two pyrimidines and dissociates from the repaired DNA (Hearst 1995).

Systematic investigations of photolyase docking to a DNA lesion have been carried out (review by Weber (Weber 2005)). Crystal structures of *E. coli*, *A. nidulans* and *T. thermophilus* photolyase have been obtained (Park, Kim et al. 1995; Tamada, Kitadokoro et al. 1997; Komori, Masui et al. 2001). A recent photolyase-UV-DNA co-crystal structure confirmed a long standing hypothesis that photolyase binds the substrate with concomitant flipping of the dimer out of the DNA helix and into the catalytic pocket of the enzyme (Mees, Klar et al. 2004). In this structure overall the helical axis bends ~ 50º at the site of damage. In 1999, van Noort et al. employed SFM to study DNA bending using UV-irradiated DNA. In this study, no DNA bending was obtained when photolyase bound to non-damaged sites and a bend of ~ 36º upon binding to a dimer site. However, the DNA was exposed to UV to induced damages. This treatment randomly modifies DNA, hence the position of lesions in this substrate is not defined. DNA bending at a dimer site was therefore obtained indirectly by subtracting the distribution of bend angles measured on non-irradiated fragments (containing non-specific bound photolyases) from the distribution obtained with irradiated DNA (containing both specific and non-specific bound photolyases).

Herein we report the preparation of a DNA substrate with a single CPD lesion located at a specific position and the binding properties of *E. coli* and *A. nidulans* photolyase using this substrate. SFM imaging and analysis of individual complexes with this defined substrate allowed direct discrimination between specific and non-specific complexes and thus comparison

of their conformational properties. We used two methods to measure DNA bend angles and compared the results.

## 1.   Materials and Methods

### Proteins and DNA substrates

Protein preparation. *E. coli* (Eker, Yajima et al. 1994) and *A. nidulans* (Eker, Kooiman et al. 1990) DNA photolyases were purified to apparent homogeneity as described before .

DNA substrates with defined damage.

We constructed a 830 bp DNA fragmenth with a single CPD, which is a formacetal deoxyuridine dimer (Appendix, Figure 1A), a structural analog of the natural cys-sin cyclobutane thymidine dimer. This lesion is efficiently recognized and processed by DNA photolyases (Butenandt, Epple et al. 2000). A 22 nt oligonucleotide with the CPD centrally positioned was synthesized by Thomas Carell (Laboratorium für Organische Chemie, ETH, Zurich) (Figure 1A). This oligo was incorporated into a long, double-stranded DNA by successive ligation of three DNA fragments (Figure 1B).

First, a 698 bp DNA fragment was produced by PCR using primers (CCACCCTGGCGCCCAATACGCAAACC) and (GCGTCTGGCCTTCCTGTAGCCAGC) with M13mp18 as a template (Figure 2B, lane 4). After PCR amplification the DNA was concentrated by ethanol precipitation and purified over a GFX™ column (Amersham). The resulting 698 bp DNA was digested with *Tsp*R I to generate a fragment with nine nt overhang. The fragment was separated by electrophoresis on a 1.5 % agarose gel run in TBE buffer (10 mM Tris-Borate, pH 7.5 and 1 mM EDTA, no ethidium bromide (EtBr)). A small part of the gel was cut and stained with EtBr to mark the position of the bands. The stained gel piece was then aligned with the remainder of the gel in order to cut out the unstained fragment. This procedure prevents additional UV-induced DNA damage. The DNA fragment of interest (395 bp was named TF; Figure 2B, lane 6) was recovered from the agarose by gel disruption, phenol extraction of the frozen, crushed gel fragment and precipitation with ethanol. All subsequent DNA purifications were done using this procedure.

Secondly, a 540 bp DNA was amplified by PCR using a plasmid template containing the mRad54 cDNA and primers (GCGTGAGGGCAAGATGAGTGTGTC) and (TAGCAGATCTTCTTTTGACCATCACGCC). The PCR product was concentrated, purified as described above and digested with *Sex*A I to produce two fragments (399 bp and 136 bp) with five nt overhangs. The fragments were separated on an agarose gel and the 399 bp DNA (further in the text referred to as SF) was purified as described above (Figure 2, lane 3).

Hybridization of the CPD-containing oligo and the CPD-complementary oligo (Figure 1A and 2A) was performed in a 10:1 molar ratio, respectively, in T4 ligase buffer L1 (Promega). Hybridization yielded a 22 bp duplex with a nine nt 3' overhang complementary to *Tsp*R I-cut end of the TF fragment at one end and a 5 nt 5' overhang complementary to *Sex*A I-cut end

of the SF fragment (Figure 1A). The CPD is positioned at one (out of three in the final 830 bp DNA) *Bgl* I recognition sites and impairs restriction enzyme activity (Figure 2B, lane 9). We refer to this DNA as short ds DNA.

A
CPD oligo               CCATTCGCCA**UU**CAGGCTGCGC
CPDcomp oligo   CGGTCACGGGGTAAGCGGTAAGTCCGACGCGGGACC

B
540 bp                                              698 bp

*Sex*A I                                            *Tsp*R I
digestion                                          digestion

399 bp                                              395 bp
SF                                                  TF

Hybridisation
and ligation 1                    CPD oligo
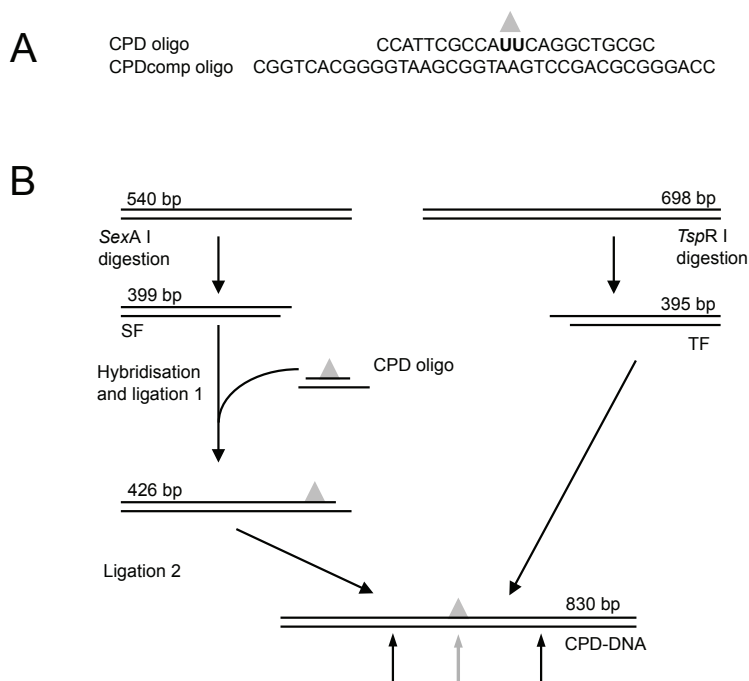
426 bp

Ligation 2

830 bp
CPD-DNA

Figure 1. Construction of the SFM substrates. (A) Sequence of the CPD oligo with the synthetic CPD (printed in bold) hybridized to its complementary oligo. The synthetic CPD analog is introduced at position 11 and 12 (printed in bold and/or denoted with gray triangle). (B) Schematic representation of the construction of the CPD-containing DNA. *Sex*A I and *Tsp*R I enzymes are used to produce two fragments (SF and TF, respectively) with different overhangs. In the first ligation step, the CPD oligo was ligated to the DNA with a *Sex*A I overhang (SF) and the product was purified. This product was ligated to the DNA containing a *Tsp*R I overhang (TF). The final DNA (named CPD-830) was 830 bp with the CPD centrally positioned. *Bgl* I restriction sites are indicated with upward pointing arrows. The grey arrow indicates the *Bgl* I site that can not be cut if the CPD is present.

Ligation of the three DNAs was done stepwise. First, the SF fragment and the short ds DNA (the latter in 5-fold molar excess) were ligated with 1 unit of T4 DNA ligase (Promega) in T4 DNA ligase buffer (Promega) with freshly added ATP at room temperature for 1 h, producing a 426 bp DNA fragment (Figure 1B and Figure 2B, lane 7). The ligation product was separated from unligated oligo on a 2% agarose gel and purified. The second ligation, between the product from the first ligation and the TF fragment, was performed under the same conditions as above with a 3-fold molar excess of the TF fragment. The final product was 830 bp DNA

with the CPD located 315 nt from each end. The final DNA was purified from the agarose gel and stored in TE buffer (1 mM EDTA and 10 mM Tris pH= 7.5). The CPD-DNA preparation was checked for the absence of nicks by $^{32}$P labeling and denaturing gel electrophoresis. No nicked DNA fragments were detected (Figure 2C).

For measuring bend angles of DNA alone, an unmodified version of the CPD oligo was used (containing a TT in place of CPD) to generate undamaged 830 bp DNA. We refer to this DNA as control DNA. As a deposition control, we used a 540 bp DNA generated with PCR and purified.
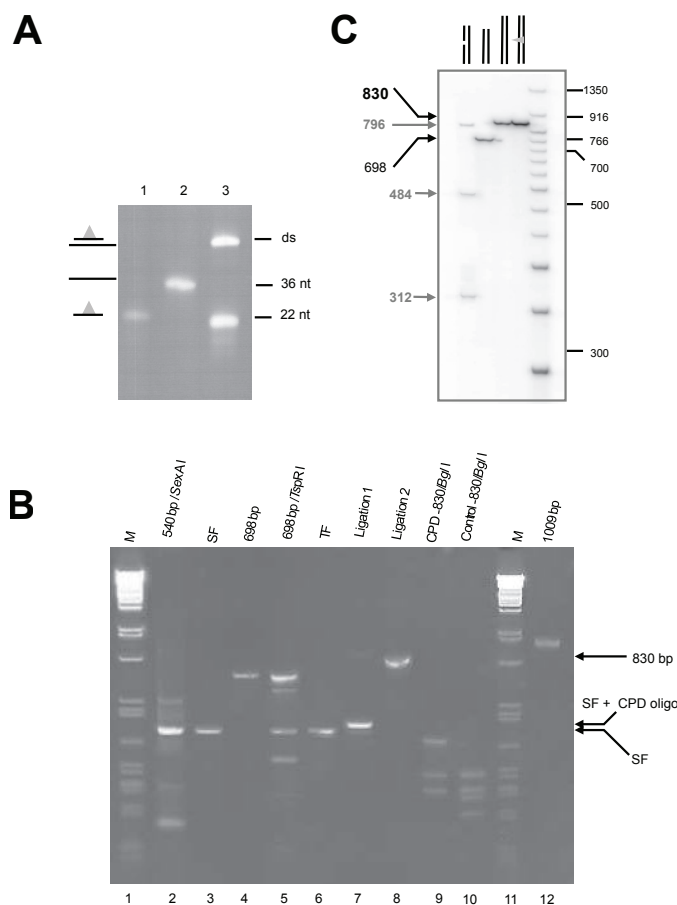


Figure 2. SFM substrates construction. (A) Hybridization of the CPD-containing oligo (lane 1) and the complementary oligo (CPDcomp) (lane 2) yielding a 22 bp duplex (short dsDNA, lane 3). Oligos were resolved by native PAGE and visualized by staining in ethidium bromide (EtBr). (B) DNA fragments used to produce SFM substrates, as indicated (see Figure 1). 540 bp DNA cut with *Sex*A I (lane 2) and purified 399 bp *Sex*A I fragment (SF, lane 3). 698 bp DNA (lane 4) cut with *Tsp*R I (partial digestion) (lane 5). Purified 395 bp *Tsp*R I fragment (TF, lane 6). 22 bp duplex (containing CPD) ligated to the SF (ligation 1, lane 7). The final 830 bp DNA (lane 8) and DNA deposition control (lane 2, 540 bp DNA, top band). CPD-DNA and control-DNA cut with *Bgl* I (lane 9 and 10, respectively). Lane 12 shows the 1009 bp UV-irradiated, and lane 4 shows the 698 bp, non-irradiated, which were also used as substrates for the photolyase binding. *Pst* I-digested lambda DNA size marker (M). (C) Final CPD-DNA and control-DNA were $^{32}$P labeled on their 5'- end and analyzed by denaturing PAGE (lane 4 and 3, respectively). Two DNAs were used: negative control for the presence of nicks: 698 bp DNA (produced by PCR, lane 2) and positive control with a single nick 796 bp DNA (produced as described in Chapter 3, but without the final ligation step) (lane 1, also gray arrows). Lane 5, 50 bp ladder marker (BioLab).

DNA substrates with random damage.

A 1009 bp DNA fragment was produced by PCR using primers (CGCCTGCTGGGGCAAACCAGCG) and (GCAAACAAGAGAATCGATGAACGG) with M13mp18 as a template (Figure 2B, lane 12). The PCR product was purified and irradiated to saturation with a germicidal lamp (S4T4.1, General Electric Company, USA), yielding ~20 CPDs per DNA fragment. This substrate contained UV damage at random locations. The 698 bp PCR product (described before) was used as non-damaged DNA.

**Binding reactions and SFM imaging**.

Photolyase-DNA binding was performed in a 10μl reaction mixture containing 100 mM NaCl, 25 mM HEPES/NaOH pH 7.5, 10 mM MgCl$_2$, 0.01% NP-40, 1 mM DTT, with 0.46 pmols of 830 bp CPD-DNA and 1.37 pmols of photolyase. Reaction mixtures were incubated at room temperature in the dark (to prevent photoreactivation). After 10 min, a sample (1 μl) of the reaction mixture was diluted with 28 μl of deposition buffer DB (5 mM HEPES, pH≈8 and 10 mM MgCl$_2$), containing 0.02 pmols of deposition control DNA, immediately deposited onto a piece of freshly cleaved mica and allowed to bind for 1 minute. Before imaging, the mica surface was rinsed with distilled water (HPLC grade, Sigma) and dried under a stream of filtered air. For the study of UV-DNA-photolyase complexes, 0.23 pmols of UV-irradiated 1009 bp DNA and 0.22 pmols of undamaged 698 bp DNA were mixed in the same binding reaction mixture and diluted 1:15 times in DB before deposition onto mica, as described above. Images were captured with a NanoScope IIIa or NanoScope IV (Digital Instruments, Santa Barbara, CA), operating in tapping mode and using silicon nanotips™ (Nanoprobe). Samples were imaged in air. Fields of 2 x 2 µm were scanned with a resolution of 512 x 512 pixels.

**SFM Image analysis**

Images were processed with NanoScope software by flattening to remove the background slope. Only DNA molecules that were completely visible within a single image and could be traced unambiguously were analyzed. Complexes with photolyase bound to DNA ends were excluded from analysis. Measurements of the DNA contour length, DNA bending angle and end-to end distance (EED) were performed with the program ALEX (Matlab SFM Image Processing Toolbox) (Rivetti, Walker et al. 1998) and volume measurements were done using SXM version 1.69 software, a modified version of NIH Image, provided by Dr Steve Barrett (Surface Science Research Centre, University of Liverpool, Liverpool, UK). DNA bending analysis based on end-to-end distance was done using the simulation-based bending analysis software developed by Dame *et al*. (Dame, van Mameren et al. 2005). Statistical analysis of the experimental data was done using the program StatPlus 2007 Professional 4.1.0. A Wilcoxon rank sum test was used to analyze bend angle distributions (p-values were two tailed), whereas Gaussian-shaped distributions were evaluated by applying a Student's t-test.

Analysis of the photolyase binding sites

The position (p) of photolyase was measured from one end and expressed as the fraction of the total contour length of the DNA (L). The expected location of the CPD damage calculated from the sequence $(p/L)_{expec.}$ is 0.5. In order to discriminate between specific and non-specific binding, this calculated value was accounted for the variation in length measurements, which is reflected through a standard deviation. Measured contour length for 830 bp CPD-DNA was 239 ± 17 nm, i.e. 239 ± 7% nm. Therefore, photolyase-DNA complexes that were within one standard deviation of the calculated value for CPD damage (0.5 – 7%) were designated as specific complexes, i.e. within p/L= 0.5 – 0.465 (Figure 4A and 5) and the complexes with p/L values outside this range were considered to be non-specific.

Measurement of DNA bending angles

To measure the degree of DNA bending we applied the traced trajectory method, as used previously (Garcia, Bustamante et al. 1996; Rivetti, Guthold et al. 1999; Seong, Kobatake et al. 2002; Janicijevic, Sugasawa et al. 2003; Lu, Weers et al. 2003; Rivetti, Codeluppi et al. 2003; Mysiak, Bleijenberg et al. 2004) and compared it with the recently developed method based on measuring end-do-end-distance (Dame, van Mameren et al. 2005). In the trajectory trace method (sometimes referred to as 'tangent method') the measured angle θ is defined as the angle between two vectors intersecting at the position of the protein. The vectors are defined as two trajectories drawn along the center of the protruding DNA arms along 15-20 nm on both sides of the protein. The apex of the angle was defined by the point where the DNA trajectories intersected. The DNA bend angle α is then defined as the deviation from straight (α =180º - θ). For measuring the intrinsic bending of DNA, the center of the molecule was determined as half of its contour length. Then a circular mask with the dimensions of the photolyase was placed at the position of the lesion and the DNA bending angle was measured. To make histograms, the DNA bending angles obtained were grouped into a number of bins that approximated the square root of the number of samples. The angle distributions were fitted using MATLAB, in a self-written software package using the Gaussian equation modified for the effect of measuring always the smallest angles described in (Schulz, Mucke et al. 1998):

$$g(\theta) = A\left[\exp\left(-\frac{\left(\theta - [\theta]^2\right)}{2 \cdot \sigma^2}\right) + \exp\left(-\frac{\left(\theta + [\theta]^2\right)}{2 \cdot \sigma^2}\right)\right] \tag{1}$$

where $\sigma$ is the standard deviation and A is the normalization constant.

The alternative method to obtain DNA bend angles is based on measuring the shortest distance between DNA ends (end-to-end distance, EED). The EED was normalized by DNA contour length for each individual molecule. The resulting distribution was fitted to a curve

based on simulations using least squares minimization according to Dame *et al.* (2005) (Dame, van Mameren et al. 2005).

Complexes formed with UV-irradiated DNA were both specific and non-specific. To obtain the bending of specific complexes, we had to account for the contribution of non-specific complexes on this substrate. For this reason we counted the number of photolyase-DNA complexes and expressed them per kb for both unirradiated and irradiated substrates (Table 3). Obtained values were compared to calculate the contribution of non-specific complexes on irradiated template.

Measurement of protein volume

The volume of the protein is defined by a manually traced area around the protein and its average height, from which the volume of the same size area in the background is subtracted. When measuring the volume of protein bound to DNA, the subtracted background area included a segment of DNA.

## 3. Results

In order to study structural changes in DNA upon photolyase binding, we constructed a linear DNA fragment (830 bp) containing a CPD in the middle (Figure 1 and 2). Binding reaction contained a 3-fold molar excess of photolyase to DNA fragment. This ratio was optimal for image analysis, because at larger excess of photolyase free protein on the mica surface obscures DNA for analysis. Since there is a single damage in the DNA, we could discriminate between photolyase bound at a CPD (specific complex) and photolyase bound at other positions of the DNA template (non-specific complex).

**Deposition of DNA onto mica**

Immobilization of the molecules on a mica surface results in a transition from three-dimensional to a two-dimensional conformation. It is expected that DNA and DNA-protein samples are free to adopt a 2D equilibrium conformation (lowest energy state conformation). The mode of adsorption on a surface can be followed by analyzing the statistical properties of DNA, well defined for undamaged and unbound DNA (Rivetti, Guthold et al. 1996). This DNA is designated here as DNA deposition control because it was added to the reaction mixture prior to deposition as a control for the deposition process. As mentioned, the ability of the DNA molecules to freely equilibrate on the flat surface is often described using polymer chain statistics (Rivetti, Guthold

$$\langle R^2 \rangle = 4PL\left(1 - \frac{2P}{L}\left(1 - e^{-\frac{L}{2P}}\right)\right) \tag{2}$$

et al. 1996). This analysis relates the mean square end-to-end distance $\langle R^2 \rangle$ to the measured contour length (L) of imaged DNA in order to estimate the DNA persistence length (P).

Persistence length is the basic mechanical property of the molecule and can be described as the length over which a DNA molecule does not bend. The persistence length of the completely equilibrated deposited DNA is ~52 nm (Rivetti, Guthold et al. 1996). In our experiments, we measured the persistence length of the 540 bp DNA, which was added to the CPD-DNA-photolyase binding reaction prior to deposition. Using equation 2, the calculated persistence length value of DNA deposition control was 56 nm for 138 DNA molecules in experiments with *A. nidulans* photolyase and 57 nm for 79 DNA molecules in the experiment with *E. coli* photolyase. Under the conditions of the experiment, these DNA persistence length values indicate that DNA and DNA-photolyase complexes could equilibrate on the mica surface (see below). For the purpose of obtaining a larger pool of data, this deposition control was omitted from some depositions. When a mix of non-irradiated and irradiated DNA was used as substrate, we measured the persistence length of the 698 bp DNA molecules. For this fragment we calculated a persistence length of 54 nm (for 119 DNA fragments) in the reaction with *E. coli* photolyase and 57 nm for 60 DNA molecules in the reaction with *A. nidulans* photolyase. The calculated persistence lengths of the equilibrated DNA found in literature measured with similar imaging conditions is between 47 and 58 nm (Rippe, Guthold et al. 1997; Margeat, Le Grimellec et al. 1998; Schulz, Mucke et al. 1998; Rivetti, Guthold et al. 1999; van Noort, Orsini et al. 1999; Podesta, Indrieri et al. 2005). In our experiments the persistence length of the DNA used as a deposition control is in the range of the values reported in literature, indicating that the sample preparation method used allowed adsorbed molecules to equilibrate on the surface.

**Binding of photolyase to CPD-DNA**.

A typical SFM image from a CPD-DNA-photolyase binding reaction is shown in Figure 3. Only molecules that were fully visible and not obstructed by other DNA fragments were used for analysis. By measuring the DNA contour lengths (L) we could easily distinguish the 830 bp CPD-DNA from the 540 bp DNA deposition control. The contour length distribution of free DNA clearly showed two populations of DNA molecules: the CPD-DNA was 241 ± 11 nm and the DNA deposition control 156 ± 9 nm. The average contour length of the CPD-DNA in complexes (239 ± 17 nm) was very similar to the free DNA, which indicates that there is no significant DNA compaction or wrapping around the bound protein (Figure 4A).

Since photolyase is small protein (~55 kDa) in some cases it was not clear by visual inspection if DNA had bound protein. Therefore, we measured the size of free and DNA-bound photolyase which also serves to exclude dimers or multimers. The size of a protein can be approximated by measuring its volume, as described before (Wyman, Rombel et al. 1997) (see also Materials and Methods). Because we intended to compare only the relative sizes of photolyases, volume measurements are presented in arbitrary units (Figure 4B). The volume measurements yielded a Gaussian shaped distribution with means and standard deviations of 9.8 ± 2.4 for free protein, 9.6 ± 2.6 for DNA bound *E. coli* photolyase and 8.9 ± 1.3 and 8.9 ± 2.3 for free and bound *A. nidulans* photolyase, respectively. The t-test showed that volume distributions of free and bound protein had very similar mean values with a 99 % confidence
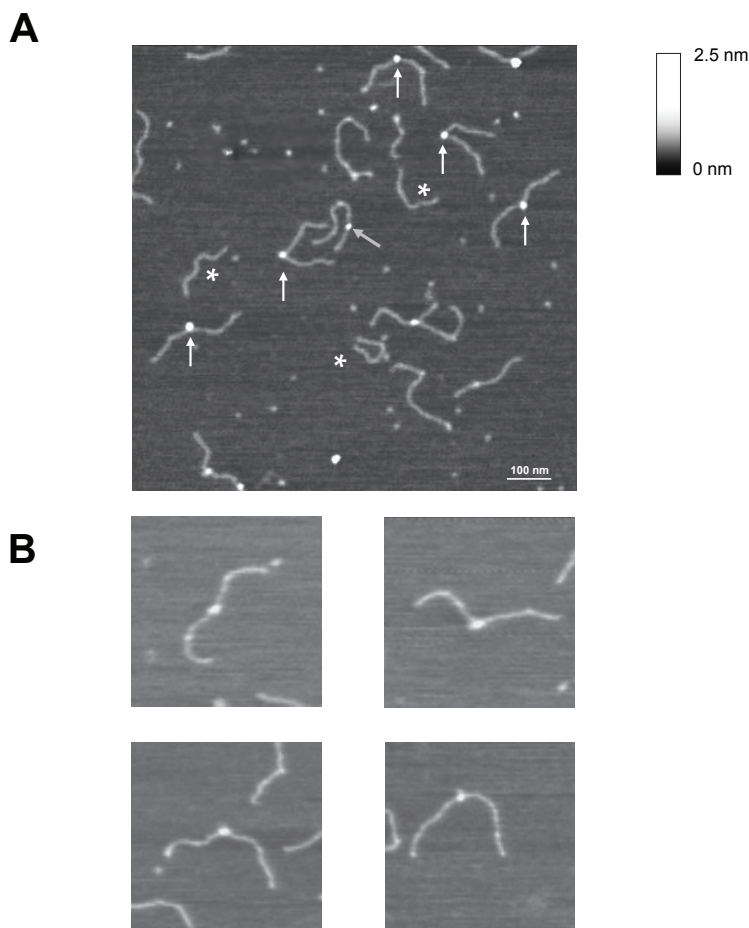
**A**



**B**



Figure 3. SFM images of DNA-photolyase complexes immobilized on mica. (A) Characteristic scan field of *E. coli* photolyase-DNA complexes. White arrows indicate photolyase binding to the position of the CPD and the gray arrow indicates a non-specific complex. Deposition control DNA (540 bp) is indicated with asterixes. Height is represented by gray scale as indicated by the bar at the right. (B) Zoomed images of specific complexes. Image size 300 x 300 nm.

level for both *A. nidulans* and *E. coli* photolyases. For bend angle analysis we included molecules that were within one standard deviation from the average volume measured for free proteins. Proteins bound to the DNA ends and ≤ 20 nm from the DNA end were not taken for analysis. The fraction of this end-bound protein was significant (about 20 %).

The position (p) of the photolyase was determined by measuring the DNA contour length from the center of the protein to the nearest DNA end. The binding position was then normalized with respect to DNA contour length (L) for each individual molecule. The distribution of photolyase binding positions is shown in Figure 5. The calculated p/L value for CPD lesion from the sequence is 0.5. Complexes that were within one standard deviation, of the calculated
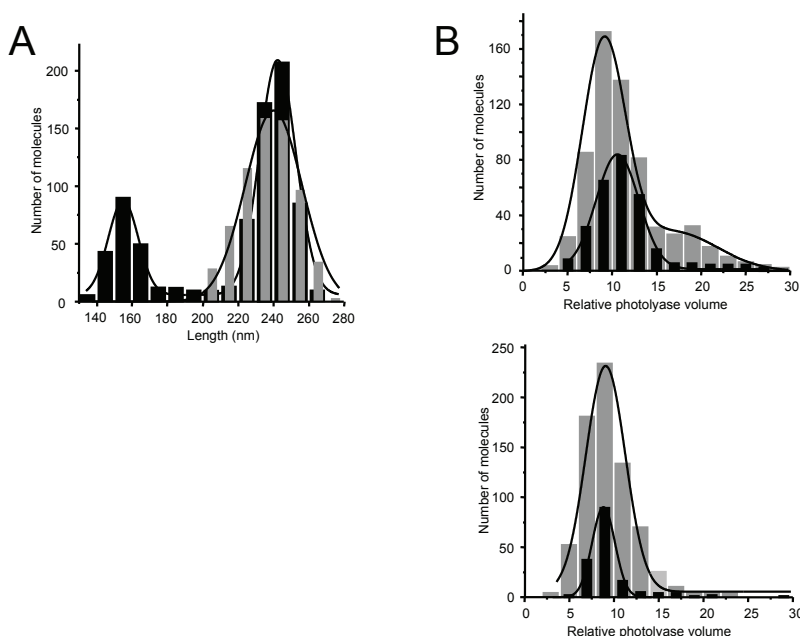
Figure 4. Size of the molecules measured from SFM images. The lines represent the Gaussian fitting of the distributions and values obtained  (average and standard deviation) are presented in the text. (A) Contour length distribution of DNA molecules without bound protein (mixture of the 540 bp DNA deposition control and 830 bp CPD-DNA) (black bars, 788 data) and the CPD-DNA (830 bp) in complexes with *E. coli* photolyase (gray bars, 688 data). The 540 bp molecules were used in the measurement of persistence length. (B) Volume distribution of photolyase molecules. Upper panel shows the histogram of volume distribution of the free *E. coli* photolyase (black bars, 285 data), compared to volume distribution of DNA bound *E. coli* photolyase (gray bars, 633 data). Lower panel shows histogram of the free *A. nidulans* photolyase (black bars, 162 data) and protein bound to DNA  (gray bars, 718 data). Volume is given in arbitrary units.

value for variation in DNA length, of the position of the lesion measured for free DNA molecules (p/L= 0.47-0.5) were classified as specific complexes and those that were outside this range were designated as non-specific complexes (see Materials and Methods). According to this criterion, we observed equal amounts (50%) of specific and non-specific complexes for the *E. coli* photolyase (Figure 5A) and 23% and 77% specific and non-specific complexes respectively for *A. nidulans* photolyase (Figure 5B). The largely non-specific binding of photolyase was unexpected given the high specificity for thymine dimers in DNA (dissociation constants are $1\times10^{-8}$ – $1\times10^{-9}$ M) (Sancar 1994; Yasui, Eker et al. 1994), but this observation was made previously in similar binding experiments imaged in air with SFM (van Noort, Orsini et al. 1999). An underestimation of specific complexes might be the result of dissociation of the protein due to the protocol used to immobilize photolyase-DNA samples on mica (dilution of the reaction mixture in a low salt buffer and rinsing the sample with water; this suggestion was based on the

observation from a study of *E. coli* RNA polymerase, where the number of specific complexes was lower when protein-DNA samples were thoroughly rinsed before drying (Schulz, Mucke et al. 1998)). A substantially larger proportion of non-specific complexes with *A. nidulans* compared to *E. coli* photolyase may indicate protein binding which is more sensitive to the buffer conditions (specifically to low salt concentration and pH).

### *Photolyase-induced DNA bending*

To estimate the degree of DNA bending we employed two methods: the trace trajectory method (for specific and non-specific binding) and the end-to-end distance method (possible
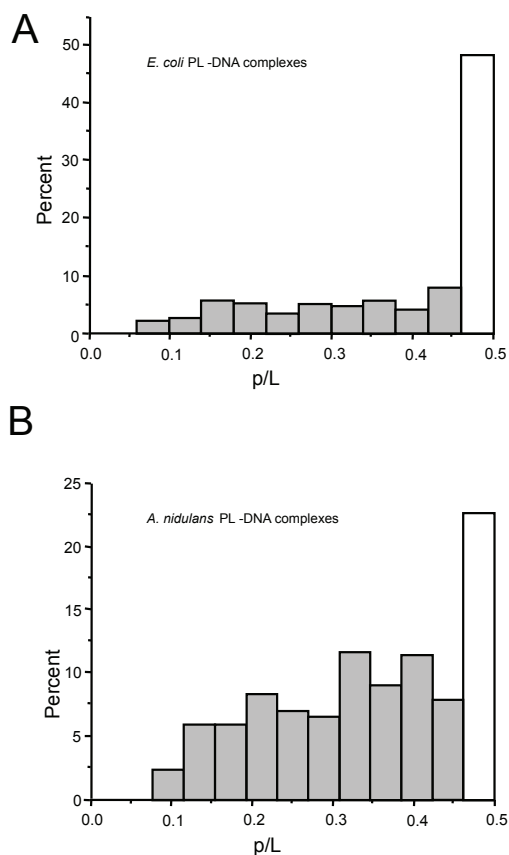


Figure 5 . Distribution of photolyase binding positions (p) along DNA for: (A) *E. coli* (527 molecules) and (B) *A. nidulans* (450 molecules). Position of photolyase was normalized with respect to the DNA contour length (L). The histogram is truncated at 0.5 because the shortest distance from the protein to a DNA end was measured. The CPD is at position 0.5. Positions within 0.47-0.5 were considered as specific (indicated by the white bar).

only for specific binding). For bend angle analysis, we pooled data from several independent experiments. The distribution of bend angles obtained by tracing DNA trajectories was fitted to a modified Gaussian distribution since angle distributions were truncated at 0° (Schulz, Mucke et al. 1998). From this Gaussian fit, the mean bend angle and standard deviation were calculated as summarized in Table 1. The standard deviation does not represent solely the error in the measurements, but also reflects fluctuations in DNA conformation, or flexibility.

Table 1: Comparison of DNA bend angle measurements using trajectory and end-to-end methods for 830 bp DNA.

| photolyase | CPD | trajectory method | | | end-to-end method | |
|---|---|---|---|---|---|---|
| | | Av. (°) | SD(°) | n | Av. (°) | n |
| - | + | 24 | 16 | 390 | 10±10 | 390 |
| | - | 0 | 17 | 154 | - | 154 |
| *E. coli* | + | 48 | 38 | 286 | 30±20 | 169 |
| | - | 40 | 30 | 221 | - | - |
| *A. nidulans* | + | 36 | 32 | 162 | 40±10 | 124 |
| | - | 38 | 28 | 330 | - | - |

Av., average bend angle; SD, standard deviation; n, number of complexes or DNA molecules

To determine whether a CPD alone induces bending in DNA we measured the DNA bend angles at the relative position 0.5 on CPD-DNA and undamaged control 830 bp DNA in the absence of photolyase. For undamaged DNA the distribution of bending angles was centered around 0, as expected for DNA without intrinsic curvature (Figure 6B). However, angles measured at the CPD reveal a bend of 20 degrees (Figure 6A). Based on a Wilcoxon rank sum test, there was a significant difference between bending at the CPD and undamaged DNA (p <0.00001). Upon photolyase binding, the template appears to be systematically bent. From the fit to a modified Gaussian curve DNA bending and standard deviation for the specific complexes of *E. coli* photolyase was 48 ± 38° and for non-specific complexes 40 ± 30° (Figure 6C and D). The distribution of bend angles for specifically bound *A. nidulans* photolyase had an average angle of 36 ± 32° and for non-specifically bound 38 ± 36° (Figure 6E and F). According to a Wilcoxon rank sum test, there was no significant difference (p >0.02) between the bend angle distributions of specific and non-specific complexes for both photolyases (Figure 6C and D and Figure 6E and F), whereas comparing distributions of the DNA-protein complexes to distribution of unbound DNA (Figure 6B) the shift from 20 to ~40° was significant with (p <0.0001).

The standard deviation of the distribution can be related to the variability of the protein-DNA complexes. This is due to the more flexible DNA in complexes and/or different DNA-protein contacts (Erie, Yang et al. 1994; van Noort, Orsini et al. 1999). Both specific and non-specific binding showed much broader distribution than that for the unbound DNA (SD are ~34 vs. ~17 degrees), whereas they are very similar for the unbound DNA with and without CPD (SD is ~17 degrees) (Figure 6 and Table 1). These results may indicate that the bent state of the photolyase-DNA complex is more flexibleor variable than free DNA.
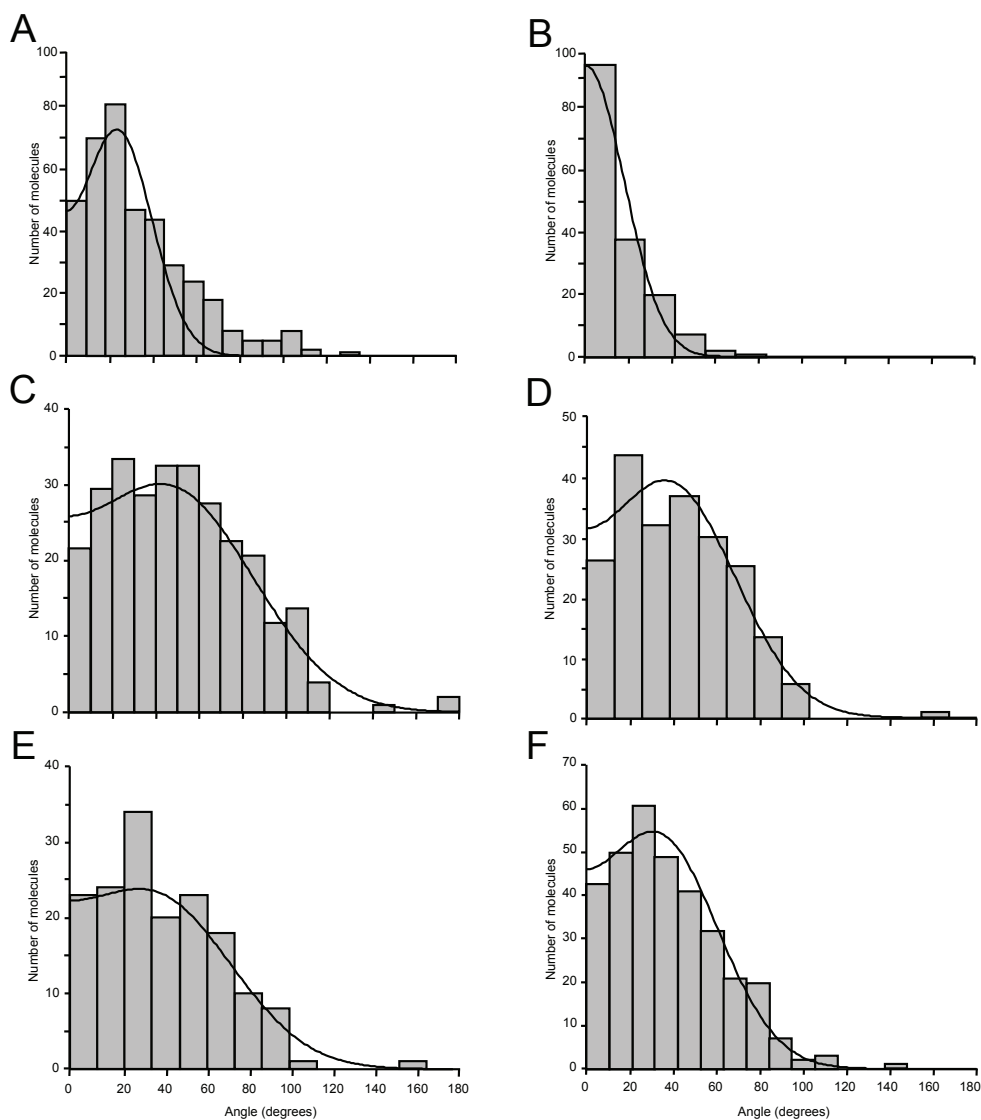
Figure 6. Distribution of the DNA bend angles with CPD-DNA as substrate. (A) Histogram of the bend angle distribution of naked DNA measured at the dimer site and (B) the center of non-damaged DNA. Histogram of the bend angles with protein bound at the position of the CPD for (C) *E. coli* and (E) *A. nidulans* photolyase. Histogram of the bend angles with protein bound at non-dimer sites on DNA for (D) *E.coli* and (F) and *A. nidulans* photolyase. Curved lines represent fit of the data to a modified Gaussian distribution. Data obtained are represented in Table1.
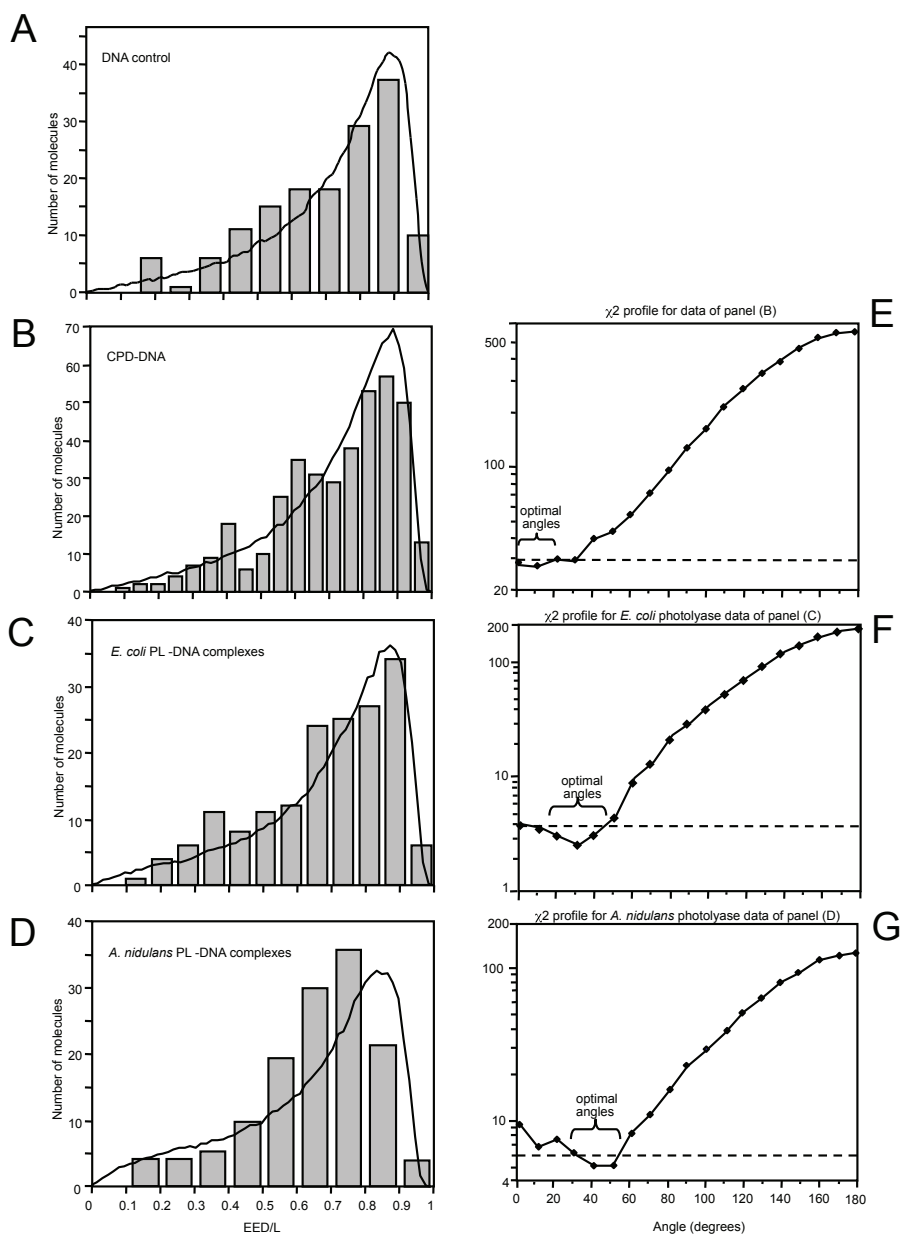
Figure 7. Bend angle determination by EED distribution. (A-D) Histograms of the EED/ L values of the bare DNA molecules and DNA-photolyase complexes and the corresponding fitting to simulated distributions for a given bend angle, obtained with the bending analysis program based on the EED/L distribution [10]. (E-G) χ2 (mean-squared error) profiles (solid line with squares) comparing experimental distribution for the data panels (B-D) to theoretical distributions at different angles. The values below the dashed line indicate the range of angle distributions that most closely fit the data.

We next estimated the DNA bend angle from the same data set using an alternative method, the end-to-end distance (EED) method (Dame, van Mameren et al. 2005). This method is applicable only to a defined site on DNA. Hence, non-specific complexes cannot be analyzed because the distribution depends on the relative location of the binding site along the contour length, which is not consistent for random binding positions. The results obtained by using this method are summarized in Figure 7 and Table 1. In contrast to the trajectory trace method, the standard deviation in these measurements represents the uncertainty in the angle determination. The DNA bend angles, obtained using this method are similar to the values obtained with the trajectory trace method.

The experimental results above differ from the results of Noort *et al.* (1999) for *A. nidulans* photolyase, especially for bending in non-specific complexes where no DNA bending was found (van Noort, Orsini et al. 1999). Therefore, we performed comparable binding experiments. As in the previous experiment, both *E. coli* and *A. nidulans* photolyase were used to bind to DNA. A mix of UV-irradiated (1009 bp) and non-irradiated DNA (698 bp) was used as substrate. We could clearly distinguish these two DNA fragments by their size. The distribution of non-specific complexes was measured on the 698 bp DNA fragment. The distribution obtained from the UV-irradiated 1009 bp fragment included both specific and non-specific complexes because this fragment had randomly distributed damages. In order to obtain the angle distribution for specific complexes we had to correct the angle distribution on the 1009 bp fragment for the contribution of non-specific binding on this fragment. The number of complexes formed on 698 bp and 1009 bp DNA was expressed as the number of complexes per 1 kilo base pairs (kb) for comparison. From our experiments, we estimate that the most significant portion of the complexes on 1009 bp DNA fragments were non-specific (82 % and 85% for *E. coli* and *A. nidulans* photolyase, respectively) (see Table 2). The specific bend angle distribution was calculated by subtracting the contribution of non-specific complexes from the angle distribution on 1009 bp DNA. The distributions of bend angles are shown in Fig. 8 and the results are summarized in Table 3. These photolyase induced bends of ~40° are very similar to those obtained using CPD-DNA as substrate.

Table 2: Summary of number of complexes formed on non-damaged and UV-irradiated DNA substrates

| photolyase | DNA size | UV | No. of DNA molecules | No. of photolyase-DNA complexes | No. of photolyase-DNA complexes/DNA molecule | No. of photolyase-DNA complexes/kb |
|---|---|---|---|---|---|---|
| *E. coli* | 698 bp | - | 186 | 137 | 1.32 | 1.91 |
| | 1009 bp | + | 387 | 176 | 2.19 | 2.18 |
| *A. nidulans* | 698 bp | - | 309 | 209 | 1.48 | 2.11 |
| | 1009 bp | + | 557 | 220 | 2.53 | 2.51 |
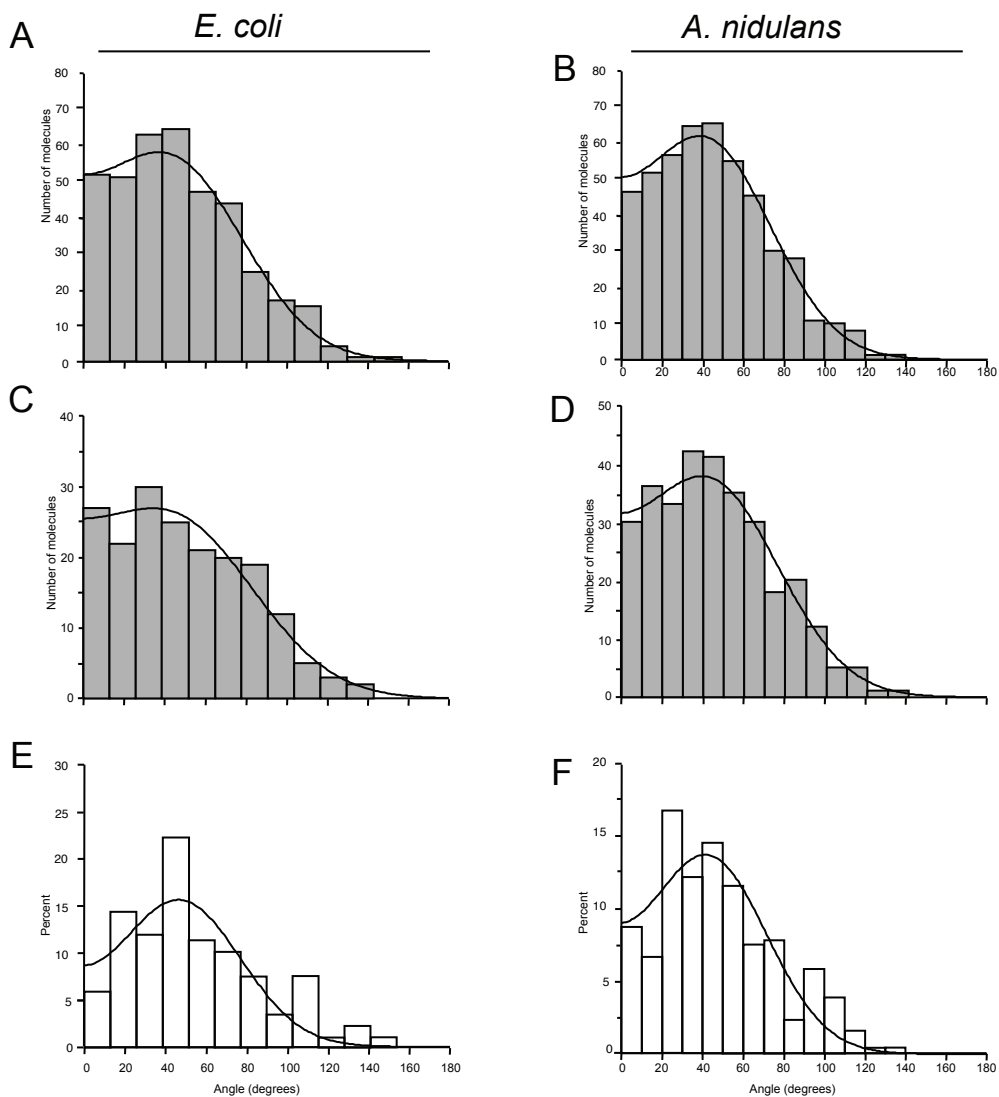
Figure 8. Distribution of the DNA bend angles with UV-irradiated DNA used as a substrate. (A and B) Mixture of specific and non-specific complexes on 1009 bp UV-DNA for *E. coli* and *A. nidulans* photolyases, respectively. (C and D) Non-specific complexes on 698 bp DNA for *E. coli* and *A. nidulans* photolyases, respectively. (E and F) Specific complexes on 1009 bp UV-DNA for *E. coli* and *A. nidulans* photolyases, respectively, obtained by subtracting distribution (C) from (A) (for *E. coli*) and (D) from (B) (for *A. nidulans*). Curved lines represent the fit of the data to a modified Gaussian distribution. Mean and standard deviation obtained from the Gaussian fit and number of molecules analyzed are presented in Table 3.

Table 3: Summary of bending angle distribution, using UV-irradiated DNA

| photolyase | DNA size | type of complexes | trajectory method | | |
|---|---|---|---|---|---|
| | | | Av. (°) | SD (°) | n |
| *E. coli* | 698 bp | nsp | 40 | 27 | 186 |
| | 1009 bp | sp + nsp | 46 | 36 | 387 |
| | (calculated) | sp | 45 | 43 | 58 |
| *A. nidulans* | 698 bp | nsp | 44 | 34 | 309 |
| | 1009 bp | sp + nsp | 36 | 35 | 557 |
| | (calculated) | sp | 45 | 33 | 84 |

Av., average bend angle; SD, standard deviation; n, number of complexes or DNA molecules; nsp, non-specific binding; sp, specific binding

## 4. Discussion

In this study we used SFM to analyze the conformation of DNA molecules in complexes with photolyase. By using a defined substrate with a CPD at a specific site, we could directly discriminate between specific and non-specific binding.

Evidence that a CPD causes alterations in DNA structure was obtained recently from a crystal structure, reporting a bend of ~30 degrees (Park, Zhang et al. 2002). A similar effect of CPD formation on duplex DNA has been observed in several other studies (circularization assay (Husain, Griffith et al. 1988), a recent NMR study (McAteer, Jing et al. 1998) and a theoretical prediction (Raghunathan, Kieber-Emmons et al. 1990)). However, very little or no bending was detected by a phased multimer gel electrophoretic study (Wang and Taylor 1991), some NMR (Taylor, Garrett et al. 1990; Kim, Patel et al. 1995) and theoretical studies (Raghunathan, Kieber-Emmons et al. 1990; Miaskiewicz, Miller et al. 1995; Cooney and Miller 1997). Our results indicate apparent DNA bending induced by CPD of ~20 degrees.

Analysis of photolyase-DNA complexes revealed that this enzyme has a strong impact on DNA structure upon specific and non-specific binding: it bent DNA by about 40°. The average bend angle determined from the distributions at specific and non-specific sites vary slightly but this difference is not statistically relevant (confirmed with Wilcoxon rank sum test). The width of the bend angle distributions (standard deviations) for the bent states were greater than for the unbound DNA (SD are ~35° and 17°, respectively) likely reflecting an ensemble of states in dynamic equilibrium for the protein-bound complexes. The fact that in the product of the crystal structure of CPD-DNA-photolyase one end of the DNA helix was not resolved (Mees, Klar et al. 2004) may indicate increased DNA dynamics in this region.

These results are in contrast to the previous work of Noort *et al.* (1999) (van Noort, Orsini et al. 1999) which found no bending when photolyase was bound to undamaged DNA, and about a 36 degrees bend for *A. nidulans* photolyase bound to damaged DNA. Therefore, we performed experiments under similar conditions using a mix of UV-irradiated and non-

damaged DNA as binding substrates, and analyzed distributions as performed previously. For the purpose of comparison we used photolyases from two different organisms (*E. coli* and *A. nidulans*), which share a high degree of structural homology (Park, Kim et al. 1995; Tamada, Kitadokoro et al. 1997). Examination of (UV) DNA-photolyase complexes showed very similar results compared to those obtained using a defined CPD-DNA (this study), in contrast with the earlier study (van Noort, Orsini et al. 1999). One possible explanation for these discrepancies might be the different methods of data analysis. Previously, a DNA substrate for photolyase was irradiated, which induced different types of UV-lesions at random positions. This required an indirect, subtraction-based method to obtain the distribution of DNA bend angles for specific photolyase-dimer complexes. But, this still does not explain why we do not have the same distribution of bend angles for non-specific binding. This difference may have arisen from the way the bend angle is defined. In the previous work, the center of the protein was chosen for the apex of the angle (in the Figure 9 designated as α), while in the present study the apex was a point where the extrapolated traces of the DNA intersected within the photolyase (in the Figure 9 designated as β) (also see Materials and Methods). On the same sample of 60 randomly chosen non-specific complexes (containing *A. nidulans* photolyase) bend angles were measured in both ways. Figure 9B shows that the obtained values for α are on average smaller than for β (the mean of measured α angles is 35.9 degrees and for β is 43.4 degrees). This implies that the bend angle determination is sensitive to the way the angle is defined. Both previously described measuring methods are used in the literature.
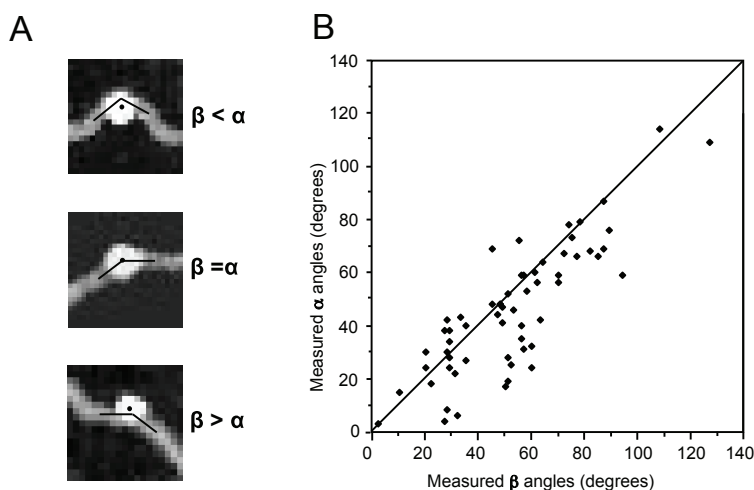


Figure 9. DNA bend angle determination. (A) Zoomed SFM images showing *A. nidulans* photolyase bound to DNA. Straight lines are the vectors used to determine the bend angle. a is the angle with the apex set at the center of the protein (marked with black dot) . β is the angle where the apex is determined at the intersection of the two vectors. Three possibilities are shown: β is smaller than α, α = β and b is bigger than a. (B) Measurement of α and β a set of 60 DNA-photolyase complexes. Each dot represents an individual protein-DNA complex. The straight line represent cases where β = α.

To circumvent discrepancies that can arise from manually defining angles we used a method proposed by Dame *et al.* (Dame, van Mameren et al. 2005). Here only EED and L are measured by defining end points and EED/L distribution is used to estimate the bend angle based on fitting to simulated distribution. This method indicates a ~40° DNA bend for specific binding, very similar to what we obtained using the trajectory method.

Another difference from the previous study of photolyase-DNA complexes by SFM is in the experimental conditions. The binding reaction in this study was performed in the presence of 0.01% NP-40, which was omitted in the previous one. Low concentrations of this non-ionic detergent were used to reduce potential protein-protein interactions. Since the influence of NP-40 on the protein-DNA interaction and binding of complexes to the mica surface has not been systematically investigated, we cannot excluded that some of the differences from the previous study reflect the influence of this detergent.

According to the current model, photolyase flips the dimer out of the DNA helix into a cavity of the protein near the catalytic cofactor FAD. This cavity resides within a stretch of positively charged amino acids on the surface of photolyase. This model is supported by site-directed mutagenesis (van de Berg and Sancar 1998; Chen, Haushalter et al. 2002), crystal structure  (Mees, Klar et al. 2004) and NMR analysis (Torizawa, Ueda et al. 2004). Base flipping is a feature observed in other DNA repair enzymes, such as DNA glycosylases, T4 endonuclease V, AP endonucleases (Yang 2006), and most recently in Rad4-Rad23 (Min and Pavletich 2007).

Photolyase has a very high specificity for pyrimidine dimers in DNA (association constant for UV-irradiated DNA is $K_a$ ~$10^8\,M^{-1}$, and for the non-damaged DNA is Ka ~$10^3$-$10^4\,M^{-1}$). Bending of the DNA in specific and non-specific complexes may play a role in the mechanism of damage recognition. The CPD-DNA-photolyase crystal structure revealed that the majority of the interactions are short-range and non-specific (including hydrogen bonding and van-der Waals forces) over a region of several nucleotides flanking the damage. Similar interactions can be assumed when photolyase binds to undamaged double-stranded DNA and induce DNA bending. Bound protein can now diffuse along the DNA checking for specific contacts. The energy required to propagate a bend of a 40° is 0.16 kcal/mol·bp, which is less then the value of $k$T (0.59 kcal/mol) (Erie, Yang et al. 1994). At the specific site, photolyase can flip the dimer out of the DNA helix. Nearby residues on the concave protein-DNA binding surface (especially tryptophans and methionine within the catalytic center) can form additional (specific) interactions with the dimer and DNA surrounding the damage, thus significantly stabilizing this protein-DNA complex. In the absence of the dimer, the normal hydrogen bonding within the helix reduces the chance for the residues in the catalytic cavity to interact with correctly paired bases. In addition to differential energy costs between binding damaged and non-damaged DNA, the energy cost of bending DNA at the dimer site from ~20° to ~40° is  lower ($\Delta G$ = 0.48 kcal/mol) (Erie, Yang et al. 1994) than bending the B-form DNA from 0° to ~40°($\Delta G$ = 1.94 kcal/mol). Consequently, at the CPD site photolyase needs less energy to distort DNA, which increases binding specificity. Energy needed to bend DNA was estimated using the equation $\Delta G = Pk T\alpha^2/2l$ where k is the Boltzmann constant, T is the absolute temperature, I is the DNA length covered by the protein (assumed to be 12 nm based on the crystal structure (Mees, Klar

et al. 2004) and footprinting analysis (Husain, Sancar et al. 1987; Baer and Sancar 1993)) and P is the persistence length of DNA (estimated here to be 55 nm). Similar binding properties have been reported for proteins like Cro (Erie, Yang et al. 1994), other sequence-specific DNA binding proteins (Bustamante and Rivetti 1996), DNA glycosylase (Chen, Haushalter et al. 2002), Mut S (Wang, Yang et al. 2003) and UvrB, which even wraps DNA at an undamaged site (Verhoeven, Wyman et al. 2002). This tendency to non-specifically bend DNA has been proposed to be a general feature of many DNA binding proteins (Erie, Yang et al. 1994).

As discussed previously, a substantial amount of information exists on the photolyase damage recognition process. Continuing investigations, especially resolving 3D structure of photolyase binding to undamaged DNA and visualizing dynamic properties of damaged-DNA-photolyase using SFM in solution would provide additional information which would address the photolyase binding and damage specificity.

## References

Ali, B. M., R. Amit, et al. (2001). "Compaction of single DNA molecules induced by binding of integration host factor (IHF)." Proc Natl Acad Sci U S A **98**(19): 10658-63.

Baer, M. E. and G. B. Sancar (1993). "The role of conserved amino acids in substrate binding and discrimination by photolyase." J Biol Chem **268**(22): 16717-24.

Bustamante, C. and C. Rivetti (1996). "Visualizing protein-nucleic acid interactions on a large scale with the scanning force microscope." Annu Rev Biophys Biomol Struct **25**: 395-429.

Butenandt, J., R. Epple, et al. (2000). "A comparative repair study of thymine- and uracil-photodimers with model compounds and a photolyase repair enzyme." Chemistry **6**(1): 62-72.

Chen, L., K. A. Haushalter, et al. (2002). "Direct visualization of a DNA glycosylase searching for damage." Chem Biol **9**(3): 345-50.

Cooney, M. G. and J. H. Miller (1997). "Calculated distortions of duplex DNA by a cis, syn cyclobutane thymine dimer are unaffected by a 3' TpA step." Nucleic Acids Res **25**(7): 1432-6.

Dame, R. T. (2005). "The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin." Mol Microbiol **56**(4): 858-70.

Dame, R. T., J. van Mameren, et al. (2005). "Analysis of scanning force microscopy images of protein-induced DNA bending using simulations." Nucleic Acids Res **33**(7): e68.

Eker, A. P., P. Kooiman, et al. (1990). "DNA photoreactivating enzyme from the cyanobacterium Anacystis nidulans." J Biol Chem **265**(14): 8009-15.

Eker, A. P., H. Yajima, et al. (1994). "DNA photolyase from the fungus Neurospora crassa. Purification, characterization and comparison with other photolyases." Photochem Photobiol **60**(2): 125-33.

Erie, D. A., G. Yang, et al. (1994). "DNA bending by Cro protein in specific and nonspecific complexes: implications for protein site recognition and specificity." Science **266**(5190): 1562-6.

Friedberg, E. C., G. C. Walker, et al. (2005). DNA Repair and Mutagenesis. Washington, DC, ASM Press.

Garcia, R. A., C. J. Bustamante, et al. (1996). "Sequence-specific recognition of cytosine C5 and adenine N6 DNA methyltransferases requires different deformations of DNA." Proc Natl Acad Sci U S A **93**(15): 7618-22.

Hansma, H. G. and J. H. Hoh (1994). "Biomolecular imaging with the atomic force microscope." <u>Annu Rev Biophys Biomol Struct</u> **23**: 115-39.

Harrington, R. E. (1992). "DNA curving and bending in protein-DNA recognition." <u>Mol Microbiol</u> **6**(18): 2549-55.

Hearst, J. E. (1995). "The structure of photolyase: using photon energy for DNA repair." <u>Science</u> **268**(5219): 1858-9.

Husain, I., J. Griffith, et al. (1988). "Thymine dimers bend DNA." <u>Proc Natl Acad Sci U S A</u> **85**(8): 2558-62.

Husain, I., G. B. Sancar, et al. (1987). "Mechanism of damage recognition by Escherichia coli DNA photolyase." <u>J Biol Chem</u> **262**(27): 13188-97.

Janicijevic, A., K. Sugasawa, et al. (2003). "DNA bending by the human damage recognition complex XPC-HR23B." <u>DNA Repair (Amst)</u> **2**(3): 325-36.

Kim, J. K., D. Patel, et al. (1995). "Contrasting structural impacts induced by cis-syn cyclobutane dimer and (6-4) adduct in DNA duplex decamers: implication in mutagenesis and repair activity." <u>Photochem Photobiol</u> **62**(1): 44-50.

Komori, H., R. Masui, et al. (2001). "Crystal structure of thermostable DNA photolyase: pyrimidine-dimer recognition mechanism." <u>Proc Natl Acad Sci U S A</u> **98**(24): 13560-5.

Lu, Y., B. D. Weers, et al. (2003). "Analysis of the intrinsic bend in the M13 origin of replication by atomic force microscopy." <u>Biophys J</u> **85**(1): 409-15.

Margeat, E., C. Le Grimellec, et al. (1998). "Visualization of trp repressor and its complexes with DNA by atomic force microscopy." <u>Biophys J</u> **75**(6): 2712-20.

McAteer, K., Y. Jing, et al. (1998). "Solution-state structure of a DNA dodecamer duplex containing a Cis-syn thymine cyclobutane dimer, the major UV photoproduct of DNA." <u>J Mol Biol</u> **282**(5): 1013-32.

Mees, A., T. Klar, et al. (2004). "Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair." <u>Science</u> **306**(5702): 1789-93.

Miaskiewicz, K., J. Miller, et al. (1995). "Molecular dynamics simulations of the effects of ring-saturated thymine lesions on DNA structure." <u>Biopolymers</u> **35**(1): 113-24.

Min, J. H. and N. P. Pavletich (2007). "Recognition of DNA damage by the Rad4 nucleotide excision repair protein." <u>Nature</u> **449**(7162): 570-5.

Mysiak, M. E., M. H. Bleijenberg, et al. (2004). "Bending of adenovirus origin DNA by nuclear factor I as shown by scanning force microscopy is required for optimal DNA replication." <u>J Virol</u> **78**(4): 1928-35.

Park, H., K. Zhang, et al. (2002). "Crystal structure of a DNA decamer containing a cis-syn thymine dimer." <u>Proc Natl Acad Sci U S A</u> **99**(25): 15965-70.

Park, H. W., S. T. Kim, et al. (1995). "Crystal structure of DNA photolyase from Escherichia coli." <u>Science</u> **268**(5219): 1866-72.

Podesta, A., M. Indrieri, et al. (2005). "Positively charged surfaces increase the flexibility of DNA." <u>Biophys J</u> **89**(4): 2558-63.

Raghunathan, G., T. Kieber-Emmons, et al. (1990). "Conformational features of DNA containing a cis-syn photodimer." <u>J Biomol Struct Dyn</u> **7**(4): 899-913.

Rippe, K., M. Guthold, et al. (1997). "Transcriptional activation via DNA-looping: visualization of

intermediates in the activation pathway of E. coli RNA polymerase x sigma 54 holoenzyme by scanning force microscopy." J Mol Biol **270**(2): 125-38.

Rivetti, C., S. Codeluppi, et al. (2003). "Visualizing RNA extrusion and DNA wrapping in transcription elongation complexes of bacterial and eukaryotic RNA polymerases." J Mol Biol **326**(5): 1413-26.

Rivetti, C., M. Guthold, et al. (1996). "Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis." J Mol Biol **264**(5): 919-32.

Rivetti, C., M. Guthold, et al. (1999). "Wrapping of DNA around the E.coli RNA polymerase open promoter complex." Embo J **18**(16): 4464-75.

Rivetti, C., C. Walker, et al. (1998). "Polymer chain statistics and conformational analysis of DNA molecules with bends or sections of different flexibility." J Mol Biol **280**(1): 41-59.

Sancar, A. (1994). "Structure and function of DNA photolyase." Biochemistry **33**(1): 2-9.

Schellman, J. A. (1974). "Flexibility of DNA." Biopolymers **13**(1): 217-26.

Schulz, A., N. Mucke, et al. (1998). "Scanning force microscopy of Escherichia coli RNA polymerase. sigma54 holoenzyme complexes with DNA in buffer and in air." J Mol Biol **283**(4): 821-36.

Seong, G. H., E. Kobatake, et al. (2002). "Direct atomic force microscopy visualization of integration host factor-induced DNA bending structure of the promoter regulatory region on the Pseudomonas TOL plasmid." Biochem Biophys Res Commun **291**(2): 361-6.

Tamada, T., K. Kitadokoro, et al. (1997). "Crystal structure of DNA photolyase from Anacystis nidulans." Nat Struct Biol **4**(11): 887-91.

Taylor, J. S., D. S. Garrett, et al. (1990). "1H NMR assignment and melting temperature study of cis-syn and trans-syn thymine dimer containing duplexes of d(CGTATTATGC).d(GCATAATACG)." Biochemistry **29**(37): 8858-66.

Torizawa, T., T. Ueda, et al. (2004). "Investigation of the cyclobutane pyrimidine dimer (CPD) photolyase DNA recognition mechanism by NMR analyses." J Biol Chem **279**(31): 32950-6.

van der Berg, B.J. and G.B. Sancar (1998). "Evidence for dinucleotide flipping by DNA photolyase." J Biol Chem **273**: 20276-20284.

van der Vliet, P. C. and C. P. Verrijzer (1993). "Bending of DNA by transcription factors." Bioessays **15**(1): 25-32.

van Noort, J., F. Orsini, et al. (1999). "DNA bending by photolyase in specific and non-specific complexes studied by atomic force microscopy." Nucleic Acids Res **27**(19): 3875-80.

Verhoeven, E. E., C. Wyman, et al. (2002). "The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands." Embo J **21**(15): 4196-205.

Wang, C. I. and J. S. Taylor (1991). "Site-specific effect of thymine dimer formation on dAn.dTn tract bending and its biological implications." Proc Natl Acad Sci U S A **88**(20): 9072-6.

Wang, H., Y. Yang, et al. (2003). "DNA bending and unbending by MutS govern mismatch recognition and specificity." Proc Natl Acad Sci U S A **100**(25): 14822-7.

Weber, S. (2005). "Light-driven enzymatic catalysis of DNA repair: a review of recent biophysical studies on photolyase." Biochim Biophys Acta **1707**(1): 1-23.

Wyman, C., I. Rombel, et al. (1997). "Unusual oligomerization required for activity of NtrC, a bacterial enhancer-binding protein." Science **275**(5306): 1658-61.

Yang, W. (2006). "Poor base stacking at DNA lesions may initiate recognition by many repair proteins." DNA Repair (Amst) **5**(6): 654-66.

Yasui, A., A. P. Eker, et al. (1994). "A new class of DNA photolyases present in various organisms including aplacental mammals." Embo J **13**(24): 6143-51.

*Appendix*

Construction of damaged DNA substrates for SFM

# Construction of Damaged DNA Substrates for SFM

We used scanning force microscopy (SFM) to investigate how proteins recognize UV-induced DNA lesions. This technique can provide qualitative and quantitative information of DNA-protein complexes. With SFM, we can visualize a complex consisting of single protein and DNA molecules. This is in contrast to other methods, i.e. electrophoretic mobility shift assays, footprinting and NMR, which measure an averaged feature of the predominant DNA-protein interaction. SFM provides information about geometry and spatial relationship of DNA-protein complexes. For instance, by measuring the position of protein binding on the DNA we can infer the interaction as 'specific' or 'non-specific'. In addition, we can measure protein-induced changes in DNA conformation such as bending and wrapping. Another unique advantage of SFM is the ability to determine  protein-induced changes in DNA even at the so-called 'non-specific' sites.

For SFM analysis, DNA substrates were designed to have lesions incorporated at defined positions. The rest of the DNA effectively provides a 'non-specific' control in the same molecule. Substrates used to study repair reactions can have damage incorporated asymmetrically along a DNA fragment (DNA arms have different lengths). This enables detection of DNA wrap and position of a protein with regard to damage. On the other hand, for measuring bend angles induced by the DNA lesions alone, we designed templates with lesions incorporated in the center of the sequence.

We used two DNA lesions efficiently recognized by specific DNA repair proteins. These lesions are chemically modified nucleotides that mimic specific endogenous DNA damages. The cholesterol-A moiety (Figure 1A) is a structure not found in living cells. However, it is used as a model lesion because it is efficiently removed in NER *in vitro* reactions in an XPC-hHR23B-dependent manner (Matsunaga, Mu et al. 1995; Mu, Hsu et al. 1996; Gomez-Pinto, Cubero et al. 2004). The second damage is a formacetal deoxyuridine dimer (Figure 1B) which is a structural analog of the natural cys-sin cyclobutane thymidine dimer with a formacetal group in place of the central intradimer phosphate group and hydrogen instead of a methyl group on the C5 of uridine. This lesion is efficiently recognized by photolyase and therefore is used for probing the mechanism of this enzyme.

For SFM analysis, two practical requirements must be met. The first requirement is the ability to produce a sufficient amount of DNA for a set of experiments. We wanted to obtain at least 2-5 µg of DNA in order to be able to perform multiple experiments with the same batch of damaged DNA. The second requirement is related to the DNA length. For SFM image analysis, we need to have DNA of sufficient length to measure possible bending, wrapping and other distortions. For instance, the analysis of protein-induced bends requires that DNA segments on each side of the bound protein should be longer than ~330 base pairs (~110 nm), or two times the persistence length. The persistence length of double stranded DNA

(P), describes mechanical properties of DNA and it characterizes the length within which the DNA axis remains stiff and for the DNA molecules in solution is approximately 50 nm (Rivetti, Guthold et al. 1996). This means that DNA has to be at least 660 bp (~220 nm) in length or more, preferably. We have the best experience with DNA size between 600 bp and 2 Kb. In addition, longer DNA (> 600 bp) allows evaluation of deposition by visual inspection and by measuring its end-to-end (EED) distribution to show that deposition has not distorted DNA strands.

**A**



Cholesterol-A

**B**
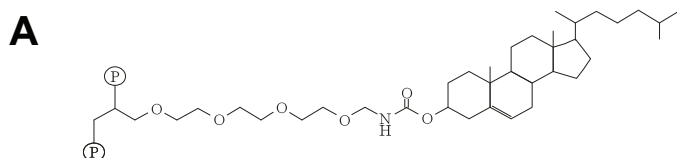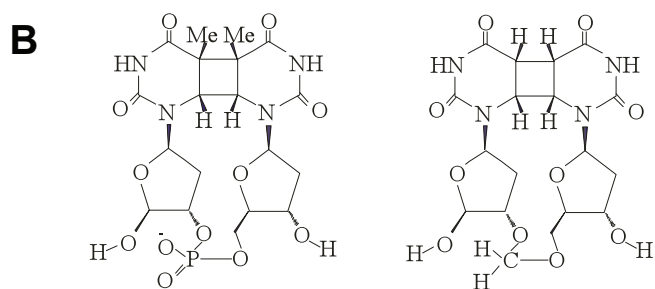


Natural cyclobutane thymine dimer          CPD lesion analogue

**C**

**CH22**- oligo: 5'- ACA CAT GTG G**X**T ATC TTG ATC G -3'.

**X** indicates the cholesterol moiety

**CPD-oligo**:  5'- CCA TTC GCC A**UU** CAG GCT GCG C -3'.

**UU** indicates the CPD

Figure 1. DNA oligonucleotide sequences and modifications. (A) Structure of choles-terol–A moiety. (B) Structure of the cyclobutane uridin dimer and its synthetic lesion analogue. (C) The nucleotide sequences of the oligos containing modified nucleotides.

Synthesis of a long oligonucleotide up to 200 nucleotides using automated synthesizers usually results in a very low yield. In addition, the synthesis of a long oligo may introduce damage to the bases and this increases markedly with oligo length. The most practical way to incorporate specific lesions into DNA is to include them in short oligos during chemical synthesis and then enzymatically incorporate the oligos into longer DNA.

To construct SFM substrates we used short DNA (oligonucleotides) with modified nucleotides chemically incorporated. These were either commercially available or synthesized by collaborators. The cholesterol-A moiety (Figure 1A) was attached via propanol to the backbone of the oligo (purchased from Eurogentec). This oligo was named CH22 and its sequence is part of the URA-3 gene of Saccharomyces cerevisiae. This oligo is 22 bp with cholesterol attached at the position of nucleotide 11 (Figure 1C). The other modification, the formacetal deoxyuridine dimer was incorporated into an oligo of 22 nucleotides at position 11 and 12, as described (Butenandt, Epple et al. 2000). The sequence of this oligo is a part of the M13 genome and is named CPD oligo (Figure 1C). The CPD oligo was a kind gift from Thomas Carell (Laboratorium fur Organische Chemie, Zurich).

First, we tried to incorporate the CPD oligo into double-stranded M13 by using it to prime DNA synthesis from the ssM13 genome. This method is widely used for site directed mutagenesis (developed in 1978 by Clyde Hutchison and Michael Smith). This converts the M13 genome into a nicked circular double-stranded DNA. After ligation, it's possible to cut out a fragment containing the lesion. In practice, DNA ends turned out un-ligatable, presumably due to incomplete or extra nucleotide synthesis. We could not recover sufficient correctly ligated DNA using this method for the purpose of our study.

Below, the two methods developed to make damaged substrates for SFM will be described in detail, including their practical advantages and disadvantages.

**Preparation of damaged DNA on a single-stranded DNA scaffold (Method I)**

An overview cartoon of this method to produce DNA with an internal cholesterol is shown in Chapter 3. This is a modified version of the method for construction of DNA substrate developed by Verhoeven et al. (Verhoeven, Wyman et al. 2001) for SFM. The outline of the synthesis is as follows: (1) An 812 bp double-stranded DNA with one 5' biotin was produced by PCR using one biotinylated primer. (2) This dsDNA was bound to streptavidin coated magnetic beads. (3) Alkaline treatment denatured the DNA, leaving the biotinylated strand (named bottom-strand) attached to streptavidin magnetic beads, thus providing a scaffold on which the strand with the damage will be constructed. (4) First, the oligo with the damage was hybridized to the bottom-strand DNA, (5) which was converted to partially double-stranded DNA by DNA polymerization using the damaged oligo as a primer. (6) A second PCR reaction produced biotinylated DNA, which was equivalent to the first 313 bp nucleotides of the first PCR product. (7) This second PCR was also bound to streptavidin beads. (8) After denaturing with alkali, the strand complementary to the single-stranded region of the partial double-stranded DNA was obtained in the supernatant. (9) This ssDNA was hybridized to the remaining single-stranded region of the scaffold DNA on the beads and (10) ligated to produce a continuous top strand.
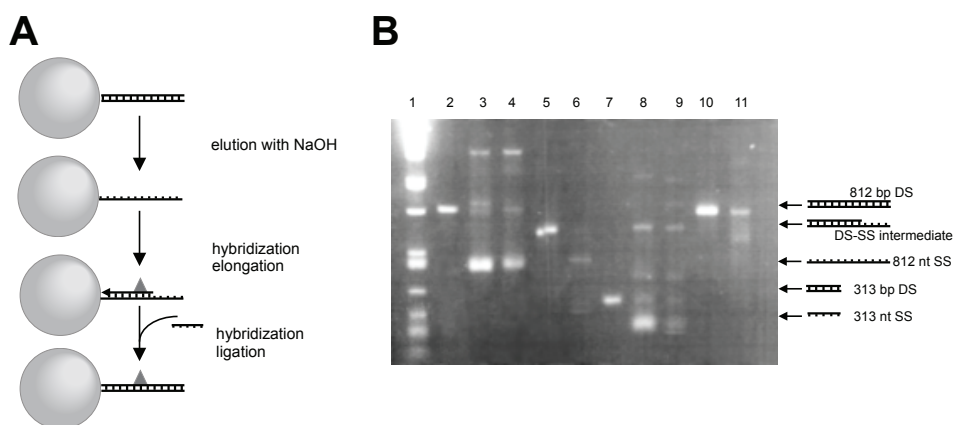
Figure 2. Preparation of damaged DNA on beads. A) Schematic representation of constructing ds DNA with a single lesion B) A portion of each reaction was electrophoresed in an agarose gel parallel with *Pst* I- digested lambda DNA size marker (lane 1). 812 bp PCR-DNA released from beads by *Sma* I digestion (lane 2). Material from to beads eluted with 0.1 M NaOH and neutralized (lane 3) containing top strand of the 812 bp DNA. Bottom strand of the 812 bp DNA (lane 4) obtained by boiling the beads to disrupt streptavidin. Partially single-stranded intermediate (produced by anniling and elongation of an oligo with a damage) and realized by *Sma* I digestion (lane 5). Non-digested intermediate DNA released by boiling (lane 6). 313 bp PCR-DNA (lane 7). Top strand of 313 bp DNA used for ligation (lane 8). Supernatant from reaction solution after anniling and ligation of the 313 bp DNA (lane 9). Final, damage-containing DNA digested by *Sma* I (lane 10). DNA released by boiling after *Sma* I digestion (lane 11).

(11) The final DNA product was released from the beads by digestion with a restriction enzyme that cuts DNA near the end with the biotin-streptavidin link.

To estimate yield, the amount of DNA for each step was compared to the amount of DNA initially bound to 100 µl of beads. Figure 2 gives an example of such an analysis. PCR amplification produced an 812 bp DNA fragment. One of the primers used in the PCR was 5'-labelled with biotin to allow binding to streptavidin-coupled beads. In order to determine the amount of immobilized DNA the streptavidin-biotin interaction was disrupted by boiling for 5 minutes in the presence of 0.01% SDS.

DNA from a 100 µl PCR reaction (~2.3 µg) was purified and coupled to 100 µl of beads, as described in Chapter 3. 100 µl of beads (Dynabeads®, Invitrogen) can bind 1.2 µg of 812 bp DNA. The amount of bound DNA was determined by comparing the amount of DNA added to the beads and the amount of DNA recovered in the supernatant after binding. Bead-bound DNA was denatured in 0.1 M NaOH for 3 minutes. The DNA strand with the biotin remained attached to the beads (bottom strand) while the other (top strand) was removed in the supernatant. The efficiency of denaturing was determined by checking the amount of DNA bound to both beads and the amount of DNA in the alkaline solution supernatant. For this, the beads were washed tree times with TE pH 7, boiled in 0.01 % SDS and allowed to

cool slowly to room temperature. This treatment detached DNA from the beads and allowed hybridizing of any complementary DNA that might remain (Figure 2, lane 4). We also checked the 0.1 M NaOH solution for the presence of single- and double-stranded DNA. A portion of the base solution was neutralized with 1 M HCl, buffered with Tris 7.5 and boiled as mentioned above (Figure 2, lane 4). Hybridization of the oligo (with the damage) to the bottom strand was performed on the beads in Sequenase 2.0 reaction buffer (40 mM Tris-HCl pH 7.7, 50 mM KCl or 50 mM NaCl). Subsequently, Sequenase 2.0 (USB) and dNTPs were added to extend the oligo. This generated partially double-stranded DNA. A portion of this intermediate DNA was detached from the beads by *Sma* I digestion to determine the efficiency of this step (Figure 2, lane 5). In the example shown, 300 ng of this intermediate DNA was bound to 100 µl of beads. After digestion, beads were also checked for the residual DNA by boiling in 0.01% SDS buffer (which disrupts streptavidin and releases the protein-DNA complex from the beads) and then electrophoresed. We detected only a minor portion of partially double-stranded DNA remaining on the beads after *Sma* I digestion. In addition, boiling released in the solution little of the bottom strand (to which the oligo was not hybridized and/or extended) and very little of DNA, which appears on the agarose gel as a faint, undefined band at approximately 260 bp (Figure 2, lane 6). Finally, to make the DNA substrate completely double-stranded a 313 nt single-stranded DNA was hybridized and ligated together with T4 ligase. Both steps were performed in T4 ligase reaction buffer. The final DNA product was released from the beads by digestion with *Sma* I (Figure 2, lane 7). Only a small portion of the DNA remained bound after *Sma* I digestion (~15 ng), as seen after boiling the beads (Figure 2, lane 8). The amount of newly synthesized damaged DNA was 400 ng. This corresponds to the amount of the partially single-stranded intermediate, which has been converted to an 812 bp fully double-stranded product. In this procedure the starting amount of DNA was 1.2 µg (DNA bound to the beads), therefore the overall yield of this protocol is about 25 %.

The most significant DNA loss occurred between DNA binding to beads and hybridization of the 313 nt DNA fragment. One possibility for the loss might be the alkali treatment (in the DNA denaturing step). After this step, both beads and supernatant were checked for DNA as described above. As shown in Figure 2 (lane 3 and 4), most of the DNA is single-stranded, with only a minor fraction of double-stranded DNA removed from beads. Another possibility is that the beads are not inert and that they bind DNA non-specifically. By adding BSA to the oligo-hybridization and elongation step, we saw little improvement when the CPD oligo was used, but higher overall yield when oligo with the CH22 was used (up to 650 ng of the final DNA, which is 50 % of the PCR DNA bound to the beads). In order to produce 2-3 µg of DNA, a large amount of beads is required because of the bead's binding capacity (DYNAL, 1998).

Since the beads might have contributed to the DNA loss, we tested the yield in the reactions without beads. Single-stranded DNA was produced by boiling and separated by denaturing (urea) PAGE. In the first instance, we had to identify 'bottom' and 'top strands' on the polyacrylamide gel. For this purpose, two labels were separately incorporated into the PCR product. First, a PCR reaction was performed with α-P$^{32}$ dATP, thus labeling both strands. In another PCR reaction, one primer was 5' labeled with γ-P$^{32}$ hence labeling only one strand.

Both PCR products were denatured, loaded next to each other on a urea polyacrylamid gel, separated by electrophoresis and detected by a phosphoimager. By comparing the bands on the gel, we could identify the position of the two different strands. Once identified, the strand complementary to the oligo (named 'bottom' strand in the previous method) was then isolated by diffusion of the DNA out of the crushed gel into water (over night at 37 °C) and concentrated with ethanol precipitation. The DNA strand obtained was used as a scaffold to construct the complementary strand with the damage, in principle the same as described in the method with DNA attached to beads. The major drawback of this method is inefficient recovery of DNA from acrylamide gel. Bands of the single-stranded DNA were usually not sharp and elution of DNA from the gel had a very low yield (data not shown).

**Preparation of a damaged-DNA substrate by serial ligation (Method II)**

An alternative method to incorporate lesion-containing oligos into longer DNA is schematically shown in the Figure 1B, Chapter 4. In this method, damaged DNA was produced by successive ligations of three DNA fragments without immobilization on beads. Directed ligations required DNA fragments with complementary, but nonpalindromic overhangs. We used the restriction enzymes *Sex*A I and *Tsp*R I because these create nonpalindromic overhangs of five and nine nucleotides, respectively.

The first fragment was a small, double-stranded DNA containing the damage. This DNA was obtained by annealing the damage containing oligo to its complement, which is longer such that the product has single-stranded extensions complementary to *Sex*A I and *Trp*R I cut fragments. Annealing of the oligos was performed with a 10-fold mol excess of the damaged oligo. Figure 2A in Chapter 4, shows that the whole pool of the complementary oligo was converted into double-stranded DNA. This is important for the subsequent ligation in order to avoid unproductive pairing of only complementary oligo with the 399 bp DNA (SF). In the next step, this short ds DNA was ligated to the longer fragment with a *Sex*A I cut end (399 bp) (Figure 2B, lane 7, Chapter 4). The ligated DNA was separated from unligated DNA by agarose gel electrophoresis and purified. Subsequently, the product of this first ligation step was ligated to a 395 bp DNA fragment with one *Tsp*R I cut end to make the final 830 bp DNA (Figure 4C) (full details of the reaction conditions are included in Material and Methods, Chapter 4).

A similar protocol was also used to produce a 932 bp DNA with a single cholesterol asymmetrically positioned in the molecule. For this purpose, different oligos (including PCR primers) were designed. In the first step, 1009 bp DNA was produced by PCR using primers (CGCCTGCTGGGGCAAACCAGCG) and (GCAAACAAGAGAATCGATGAACGG) with M13mp18 as a template. After PCR amplification and purification, the DNA was digested with *Tsp*R I to generate fragments of 612 bp (designated as TF). In the second step 433 bp DNA was produced by PCR using plasmid including cDNA of the mRad54 as a template and primers (GCGTGAGGGCAAGATGAGTGTGTC) and (GGTTAAGACCACAGCCCCCGGC). This 433 bp DNA was digested with *Sex*A I to produce 298 bp DNA (SF). Hybridization of the cholesterol-containing oligo CH22 (ACACATGTGG<u>X</u>TATCTTGATC, where X represents cholesterol in a place

of the nucleotide), and CH22comp (CCAGGCAGTCAAGATATCCACATGTGTGGCACTGGC) yielded a 22 bp duplex with a nine nt 3' complementary to the *Tsp*R I cut end of the 612 bp fragment (TF) at one end and a 5' overhang complementary to the *Sex*A I cut end of the 298 bp fragment (SF) (data not shown).

During DNA preparation, it was necessary to avoid exposing the DNA to ethidium bromide (EtBr) and UV to prevent random damage. Therefore, the long DNAs (both the restriction fragments used in ligations and the final product) were separated on 1.2 - 2% agarose gels (ultra PURE, Invirogen™) in TBE buffer (10 mM Tris-Borate, pH 7.5 and 1 mM EDTA, no EtBr). A small part of the gel was stained with EtBr to mark the position of the bands. The stained gel piece was then aligned with the rest of the gel in order to cut out the DNA fragment of interest. DNA was purified from the agarose by gel disruption, extraction (by freezing the crushed gel in the phenol) and ethanol precipitation.

A typical yield from this method for the production of cholesterol-DNA: starting with 10 x 100 µl of PCR reactions to amplify the 433 bp and 1009 bp DNA; after digestion with restriction enzymes and gel purification 17 µg of the 298 bp fragment and 15 µg of the 612 bp fragment were obtained; after ligations and gel isolations we obtained 2.4 µg of 932 bp DNA with a damage site incorporated. This is a ~ 10% yield with respect to the amount of the 298 bp PCR product used.

We aimed to develop a method for DNA construction, which would be efficient, reproducible and cost efficient. We described two methods based on using oligos with chemically incorporated damages. In the method where DNA is immobilized on a solid surface (magnetic beads), it is relatively easy to change reaction conditions without losing material. The yield of this method can be up to 50%, regarding the amount of PCR DNA bound to the beads. However, due to the initial low binding of DNA to the beads, a large amounts of beads was required, which apropos are relatively expensive. The yield of the second method, where DNA is prepared by serial ligations is comparatively lower (10%), but it does not require any special reagents and therefore was considered more cost effective than the previous method.

Using both methods to construct modified DNA, virtually any chain length can be obtained. The position of the damage within the sequence can also be varied. In general, any available DNA modification can be in incorporated into DNA using these protocols.

## *References*

Butenandt, J., R. Epple, et al. (2000). "A comparative repair study of thymine- and uracil-photodimers with model compounds and a photolyase repair enzyme." Chemistry **6**(1): 62-72.

Gomez-Pinto, I., E. Cubero, et al. (2004). "Effect of bulky lesions on DNA: solution structure of a DNA duplex containing a cholesterol adduct." J Biol Chem **279**(23): 24552-60.

Matsunaga, T., D. Mu, et al. (1995). "Human DNA repair excision nuclease. Analysis of the roles of the

subunits involved in dual incisions by using anti-XPG and anti-ERCC1 antibodies." <u>J Biol Chem</u> **270**(35): 20862-9.

Mu, D., D. S. Hsu, et al. (1996). "Reaction mechanism of human DNA repair excision nuclease." <u>J Biol Chem</u> **271**(14): 8285-94.

Rivetti, C., M. Guthold, et al. (1996). "Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis." <u>J Mol Biol</u> **264**(5): 919-32.

Verhoeven, E. E., C. Wyman, et al. (2001). "Architecture of nucleotide excision repair complexes: DNA is wrapped by UvrB before and after damage recognition." <u>Embo J</u> **20**(3): 601-11.

Summary

# Summary

The integrity of DNA, the carrier of genetic information, is constantly challenged by numerous endogenous and exogenous agents. DNA damage can arise within a cell as a result of endogenous metabolic processes or by exposure to exogenous agents such as UV and ionizing radiation from the sun and many chemical agents in our surroundings. Persistent DNA damage can become a serious threat for the cell or organism. A number of specialized repair mechanisms have evolved to counteract the deleterious effect of DNA damage, emphasizing the importance of genome maintenance.

Two important repair pathways for removing UV-induced damages from DNA are direct reversal and nucleotide excision repair (NER). Chapter 1 provides an overview of the basic concepts of these two repair pathways and the main factors involved, with emphasis on the NER pathway. Furthermore, proteins involved in UV-damage recognition in *E. coli* and eukaryotes and their mechanism of actions, are discussed.

Scanning force microscopy (SFM) has been used for high-resolution imaging DNA-protein complexes in this thesis. SFM is a relatively new technique yet it has become an essential tool in the study DNA repair mechanisms. In Chapter 2 the relevance of this single molecule technique in studying the structural basis of protein-DNA interactions and the relevance of these interactions in the processes of DNA repair, are presented.

XPC-HR23B is the factor that is essential for UV-damage recognition in a nucleotide excision repair subpathway that operates throughout the genome. To understand how XPC-hHR23B function, interactions of the human XPC-HR23B with DNA containing single damage are reported in Chapter 3. Scanning force microscopy analysis reveals that XPC–HR23B induces a bend on cholesterol-DNA substrates. This architectural feature might be important for the repair reaction.

The conformation of the DNA in specific and non-specific complexes with photolyase, a small DNA repair enzyme, and a DNA fragment containing synthetic CPD is investigated in Chapter 4. From images obtained using scanning force microscopy bend angles were estimated using the trace trajectory method and the end-to-end method. They revealed that photolyase bent specific and non-specific DNA by roughly equivalent amounts, ~40 degrees. The role of DNA bending in the binding specificity and mechanism of specific site recognition are discussed.

To study molecular mechanism of DNA repair on a single molecule level, it is of importance to have specific defined DNA substrates. In the Appendix, two methods for DNA preparation used in this thesis are described in details. They both start with short DNA (oligo-nucleotides) with modified nucleotides chemically incorporated. In first method, damaged DNA is prepared on a single-stranded DNA scaffold. Second method is based on serial ligation of DNA fragments. Efficiency and potential applications of both methods are discussed.

# Samenvatting

De integriteit van het DNA, de drager van genetische informatie wordt voortdurend bedreigd door talloze endogene en exogene stoffen en agentia. DNA beschadigingen kunnen ontstaan binnen een cel als gevolg van endogene metabole processen of na blootstelling aan exogene agentia zoals UV licht van de zon en röntgenstraling en vele chemische verbindingen uit de omgeving. Blijvende DNA beschadigingen kunnen een ernstige bedreiging zijn voor de cel en uiteindelijk het hele organisme. Een aantal gespecialiseerde reparatiemechanismen zijn in de evolutie ontstaan om het nadelige effect van DNA schade te verminderen. Dit laat het belang van genetische stabiliteit zien.

Twee belangrijke reparatieprocessen voor het verwijderen van UV geïnduceerde DNA schade zijn: directe terugvorming van de onbeschadigde DNA basen door het reparatie enzym fotolyase en nucleotide excisie reparatie (NER). Hoofdstuk èèn geeft een overzicht van de basale concepten van deze twee reparatie mechanismen en de belangrijkste factoren die daarbij betrokken zijn met speciale aandacht voor het NER systeem. Daarnaast worden eiwitten besproken die betrokken zijn in de herkenning en het reactiemechanisme van UV DNA schade in E.coli bacteriën en in eukaryoten.

Scanning forse microscopy (SFM) is in dit proefschrift aangewend voor het bestuderen van DNA eiwitcomplexen met een hoge mate van resolutie. SFM is een relatief nieuwe techniek die een belangrijk hulpmiddel is bij het bestuderen van DNA reparatiemechanismen. In Hoofdstuk twee wordt de relevantie besproken van deze techniek die werkt op het niveau van een enkel molecuul.

XPC-HR23B is het eiwit dat van essentieel belang is voor het herkennen van UV beschadigingen in DNA in een NER subproces dat zich richt op het hele genoom. Om inzicht te krijgen hoe het XPC-HR23B zijn functie uitoefent wordt in hoofdstuk drie beschreven hoe het humane XPC-HR23B eiwit een interactie aangaat met een DNA molecuul dat een enkele beschadiging bevat. Met behulp van SFM hebben we gevonden dat het XPC-HR23B een bocht in het DNA induceert op de plek van een cholesterol-DNA substraat. Deze conformatieverandering kan belangrijk zijn voor het vervolg van de reparatiereactie.

De conformatie van het DNA bij specifieke en niet-specifieke complexen van fotolyase en een DNA fragment met een synthetische UV beschadiging is onderzocht in Hoofdstuk vier. Met behulp van trajectmetingen en eind-eindmethoden zijn buigingshoeken bepaald door middel van SFM. Deze studies lieten zien dat fotolyase DNA specifiek en niet-specifiek buigt met om en nabij dezelfde buigingshoeken van ongeveer 40 graden. De rol van DNA buiging in de bindingsspecificiteit en het herkenningsmechanisme van specifieke beschadigingen worden besproken.

Om het moleculaire mechanisme van DNA reparatie onder meer op het niveau van enkele moleculen te bestuderen is het van belang om te beschikken over specifieke DNA fragmenten. In de appendix worden twee methoden in detail beschreven voor het bereiden van DNA moleculen die in deze thesis worden gebruikt. Beide methoden beginnen

met korte stukjes DNA (oligonucleotiden) waarin chemisch gemodificeerde nucleotiden zijn geïncorporeerd. In de eerste methode wordt het beschadigde DNA vervaardigd op een enkelstrengs DNA raamwerk. De tweede methode is gebaseerd op achtereenvolgens achter elkaar plakken van DNA fragmenten. De efficiëntie en de mogelijke toepassingen van beide methoden worden besproken.

*Vertaling: J.H.J. Hoeijmakers*

# Abbreviations

| | |
|---|---|
| AFM | atomic force microscopy |
| ATP | adenosine triphosphate |
| bp | base pair |
| CFP | cyan fluorescent protein |
| CPD | cyclobutane pyrimidine dimer |
| CS | Cockayne sndrome |
| CSA/B | Cockayne sndrome A/B protein |
| DBS | double-strand break |
| DDB1/2 | damaged-DNA binding protein 1/2 |
| DNA | deoxyribonucleic acid |
| dsDNA | double-stranded DNA |
| EED | end-to-end distance |
| EM | electron microscopy |
| FAD | flavin adenine dinucleotide |
| GFP | green fluorescent protein |
| GG-NER | global genome nucleotide excision repair |
| (h)H23B | (human) homolog of *S. cerevisiae* Rad23 |
| Ka | association constant |
| kDa | kiloDalton |
| MMR | mismatch repair |
| NER | nucleotide excision repair |
| NMR | nuclear magntric resonance |
| nsp | non-specific |
| nt | nucleotide(s) |
| PAGE | polyacrylamide gel electrophoresis |
| PCR | polymerase chain reaction |
| (6-4)PP | (6-4) pyrimidine-pyrimidone photoproduct |
| RPA | replication factor A |
| SFM | scanning force microscopy |
| sp | specific |
| ssDNA | single-stranded DNA |
| TC-NER | transcription-coupled nucleotide excision repair |
| TFIIH | transcription (initiation) factor IIH |
| TTD | trichothiodystrophy |
| UV | ultraviolet |
| WLC | wormlike chain (model) |
| XP | xeroderma pigmentosum |
| XPA to -G | xeroderma pigmentosum group A to -G protein |

# *Curriculum Vitae*

Name:          Ana Janićijević

Born:          16 September 1974, Belgrade, Serbia

1993 – 1998:   10th Gymnasium, Belgrade, Serbia.

1993 – 1998:   University of Belgrade, Faculty of Biology, Serbia. Obtained title: M.Sc. Molecular biologist and physiologist.

1996:          Undergraduate student at the Institute for Microbiology, Biochemistry and Genetics, Munich, Germany (supervision Prof. Dr. Christian Koch).

1996 – 1997:   Undergraduate student at the department for quantitative genetics, Belgrade, Serbia (supervision Prof. Dr. Marina Stamenković -Radak).

1998:          University diploma work at the Institute of Molecular Genetics and Genetic Engineering, Belgrade, Serbia (supervision: Dr. Branka Vasiljević).

1998 – 1999:   Practical training period at the department for microbiology, Diversey Lever, Maarsenbroek, The Netherlands.

1999 – 2004:   Ph.D. student at Department of Genetics, Erasmus MC, The Netherlands (supervision: Prof.Dr. J.H.J. Hoeijmakers and Dr. Claire Wyman).

Publications

- Dame RT, van Mameren J, Luijsterburg MS, Mysiak ME, Janicijevic A, Pazdzior G, van der Vliet PC, Wyman C, Wuite GJ. Analysis of scanning force microscopy images of protein-induced DNA bending using simulations. Nucleic Acids Res. 2005 Apr 20; 33(7): e68.

- Janicijevic A, Ristic D, Wyman C. The molecular machines of DNA repair: scanning force microscopy analysis of their architecture. J Microsc. 2003 Dec; 212(Pt 3): 264-72. Review.

- Janicijevic A, Sugasawa K, Shimizu Y, Hanaoka F, Wijgers N, Djurica M, Hoeijmakers JH, Wyman C. DNA bending by the human damage recognition complex XPC-HR23B. DNA Repair (Amst). 2003 Mar 1; 2(3): 325-36.

## Acknowledgments

It was a pleasure for me to work with all the wonderful people at Erasmus...

First of all, I would like to express my gratitude to my supervisor, Jan Hoeijmakers, whose expertise and understanding, added considerably to my graduate experience. I would also like to thank to Claire Wyman for being a great advisor. Claire, your ideas and enthusiasm had a major influence on this thesis. I appreciate your vast knowledge and skill in many areas, and to be honest - your healthy criticism.

I wish to express my warmest thanks to Wim Vermeulen for encouragement and support even for work where you were not directly involved or interested. Another scientist from whom I learnt a good deal is Andre Eker. Besides helping me practically, thank you for reading my thesis and providing constructive comments/criticisms. Many, many thanks go to Jasperina, Rita and Marike, great and efficient secretaries and Sjozef, Ton and Pim for their help with computer problems. I will like to thanks to my colleagues and former colleagues who have provided a range of advice and support during various stages of this thesis: Deborah, Betty, Suzanne, Vincent, Steven, Nils, Angelica, Astrid, Arjan, Chris, Sergey, Lise, P.O., Jay, Dejan, Dorota, Kris, Nancy, Audrey...I also spent sometime working with my friends from Leiden, Ester and Remus - and it was such a pleasure! Thank you both for tolerating my constant presence in that small SFM room and helping me in my early experiments. Ambra, thanks for being not only excellent colleague but also very good friend, for your understanding and advices during all these years. Martijn, you have been always there for me - professionally and personally. I am very thankful for this, and I hope to continue this relationship in a future.  Karl, I knew I could count on you; thanks for taking some time out to read this thesis and to offer many useful comments and suggestions for improvement. I would also like to thank Farzin for unselfish assistance in the C-lab. On a more personal note, I would like to thank Filippo for his friendship and optimism over the years. Sanja, puno hvala za društvo i vesele razgovore. Milena, ili 'naša Malena', kako te interno zovemo, hvala na svojim pozitivnim mislima, na podršci i pomoći svih ovih godina.

And most importantly -  thanks to my family.

Mojim najmilijima: Zaista me je oduvek  oduševljavalo strpljenje, podrška i spremnost da mi budete tu kad god mi zatreba. Poslednjih nekoliko godina su mi bile jedne od težih, ali nikad nisam sumnjala u vaše strpljenje i veru u mene. Hvala (super)mama, tata, hvala Aco i mnogo vas volim. Ivana i Strahinja,  kad budete malo veći i budete ovo čitali nadam se da ćete biti ponosni na vašu mamu kao što je ona na vas dvoje. I naravno, Mišo, tebi hvala za za podršku, inspiraciju i sve komentare (izgovorene i neizgovorene). Ovu tezu posvećujem svima vama uz zahvalnost, koju se ne mogu izkazati ni rečima ni delima već beskrajnom radošću zajedničkog življenja.