



Single molecule studies of DNA lesion search and recognition strategies

Einzel-Molekül-Studien von Strategien
zur DNS-Schadenssuche und –erkennung

Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Biomedicine

submitted by

Claudia Nadine Büchner

from:

Bad Kissingen

Würzburg 2014

Submitted on: 23.12.2014

Members of the Promotionskomitee:

Chairperson: Prof. Dr. Thomas Dandekar

Primary Supervisor: Dr. Ingrid Tessmer

Supervisor (Second): Prof. Dr. Caroline Kisker

Supervisor (Third): Prof. Dr. Markus Sauer

Date of Public Defence:

Date of Receipt of Certificates:

Table of content

1. Summary.....	4
1.1. Summary (English).....	4
1.2. Zusammenfassung (Deutsch)	6
2. Introduction.....	8
2.1 Atomic Force Microscopy (AFM) imaging of biomolecules.....	8
2.1.1 Principles of AFM.....	8
2.1.2 Assets and drawbacks of AFM.....	11
2.1.3 Biological applications of AFM	13
2.2 DNA repair as a fundamental concept for genome maintenance	15
2.2.1 Sources and consequences of DNA damage	15
2.2.2 DNA damage excision repair pathways.....	17
2.3 Nucleotide excision repair (NER).....	18
2.3.1 Mechanism of eukaryotic NER	19
2.3.2 DNA damage verification in eukaryotic NER	21
2.3.3 Archaeal XP homologs as a model for eukaryotic NER	22
2.3.4 Objectives in lesion recognition and verification by XPD.....	24
2.4 Base excision repair (BER)	28
2.4.1 Principles and mechanism of BER	28
2.4.2 DNA lesion search and recognition by DNA glycosylases.....	31
2.4.3 Thymine DNA glycosylase (hTDG)	32
2.4.4 8-oxoguanine DNA glycosylase (hOgg1).....	34
2.4.5 Objectives in BER lesion search and recognition.....	36
3. Materials.....	37
3.1 Equipment	37
3.2 Chemicals.....	39

3.3	Biochemicals.....	41
4.	Methods	44
4.1	Atomic Force Microscopy (AFM) imaging	44
4.1.1	Sample deposition.....	44
4.1.2	AFM image collection and processing.....	46
4.1.3	AFM data analysis.....	46
4.2	Molecular biology.....	49
4.2.1	Transformation and plasmid isolation.....	49
4.2.2	DNA substrate preparation for AFM	49
4.3	Expression and purification of taXPD	54
4.3.1	Protein expression and cell lysis.....	54
4.3.2	Metal affinity chromatography	54
4.3.3	Size exclusion chromatography.....	55
4.4	Protein characterization.....	55
4.4.1	SDS-PAGE analysis	55
4.4.2	Determination of protein concentration by UV-spectrophotometry	56
4.4.3	Circular dichroism (CD).....	56
4.5	Analysis of protein-DNA interactions	57
4.5.1	Gel-based helicase assay	57
4.5.2	Bio-Layer Interferometry (BLI) DNA binding assay.....	57
4.5.3	2-Aminopurine fluorescence based flipping assay.....	59
5.	Results and discussion.....	61
5.1	Substrate versatility of the NER helicase XPD in damage verification DNA.....	61
5.1.1	Purification of taXPD proteins	62
5.1.2	taXPD preferentially loads on unpaired regions within dsDNA	65
5.1.3	Lesion recognition and stalling at lesion sites requires ATP hydrolysis	68

5.1.4	taXPD utilizes different mechanisms for verification of CPD and fluorescein lesions	73
5.1.5	Stable complex formation at the lesion involves ATP binding	77
5.1.6	Model of damage verification by taXPD	83
5.2	Lesion recognition strategies of DNA glycosylases in base excision repair	84
5.2.1	hTDG recognizes G:U and G:T mismatch sites with moderate specificity	84
5.2.2	DNA bending and base flipping by hTDG upon lesion encounter	86
5.2.3	Nonspecific complexes of hTDG on homoduplex DNA	90
5.2.4	AFM volume analysis reveals the stoichiometry of free and DNA-bound hTDG	94
5.2.5	hTDG complex size correlates with DNA bending	96
5.2.6	hTDG damage search and recognition model	98
5.2.7	Generalizability of hTDG lesion search and detection in comparison with hOgg1	101
6.	Final conclusions	104
7.	References	105
8.	Supplement	121
8.1	Abbreviations	121
8.2.	Figure index	124
8.3.	Table index	126
8.4.	Affidavit	127
8.5.	Publications	128
8.6.	Oral presentations and posters	129
8.7.	Curriculum Vitae	130
8.8.	Acknowledgments	131

1. Summary

1.1. Summary (English)

The integrity of our genome is continuously endangered by DNA damaging factors. Several cellular mechanisms have evolved to recognize and remove different types of DNA lesions. Despite the wealth of information on the three-dimensional structure and the catalytic mechanism of DNA repair enzymes, the essential process of target site search and identification remains more elusive. How can a small number of repair proteins find and detect the rare sites of damage rapidly and efficiently over an excess of millions of undamaged bases?

To address this pivotal question in DNA repair, I focused on the central players from the two DNA damage excision repair pathways in my studies: nucleotide excision repair (NER) and base excision repair (BER). As examples for completely different approaches of damage search, recognition and verification, I compared the NER protein Xeroderma pigmentosum group D (XPD) with the BER proteins human thymine DNA glycosylase (hTDG) and human 8-oxoguanine glycosylase (hOgg1).

In particular, the single molecule approach of atomic force microscopy (AFM) imaging and complementary biochemical and biophysical techniques were applied. I established a simple, optimized preparation approach, which yields homogeneous and pure samples of long (several hundreds to thousands of base pairs) DNA substrates suitable for the AFM studies with DNA repair proteins. Via this sample preparation, a single target site of interest can be introduced into DNA at a known position, which allows separate analysis of specific protein-DNA complexes bound to the lesion site and nonspecific complexes bound to non-damaged DNA.

The first part of the thesis investigates the XPD protein involved in eukaryotic NER. In general, the NER mechanism removes helix-distorting lesions – carcinogenic UV light induced photoproducts, such as cyclobutane pyrimidine dimers (CPDs) as well as bulky DNA adducts. The 5'-3' helicase XPD has been proposed to be one of the key players in DNA damage verification in eukaryotic NER, which is still a matter of hot debate. In the studies, I focused on XPD from the archaeal species *Thermoplasma acidophilum* (taXPD), which shares a relatively high sequence homology with the sequence of the human protein and may serve as a good model for its eukaryotic counterpart. Based on AFM experiments and accompanying DNA binding affinity measurements with the biosensor technology Biolayer Interferometry (BLI), a clear role of XPD in damage verification was deciphered. Specifically, the data suggested that the ATP-dependent 5'-3' helicase activity of XPD was blocked by the presence of damage leading to stalled XPD-DNA damage verification complexes at the lesion sites.

Successful damage verification led to ATP-dependent conformational changes visible by a significant transition in DNA bend angles from $\sim 50^\circ$ to $\sim 65^\circ$ at the site of the bound protein. Remarkably, this DNA bend angle shift was observed both in the presence of ATP and ATP γ S (non-hydrolyzable ATP analog) indicating that ATP-binding instead of ATP hydrolysis was sufficient to induce repair competent conformational changes of XPD. Most importantly, detailed protein binding position and DNA bend angle analyses revealed for the first time that XPD preferably recognizes a bulky fluorescein lesion on the translocated strand, whereas a CPD lesion is preferentially detected on the opposite, non-translocated strand. Despite the different recognition strategies for both types of damages, they share a common verification complex conformation, which may serve as a signal for the recruitment of further NER factors.

In the second part of the thesis, AFM imaging and a 2-Aminopurine fluorescence-based base-flipping assay were combined to investigate damage search and recognition by DNA glycosylases in BER. Exemplarily, I chose to study hTDG as a representative of the vast glycosylase family. hTDG excises thymine and uracil from mutagenic G:T and G:U mismatches contributing to cancer and genetic disease. The AFM data suggested that hTDG uses the intrinsic flexibility of G:T and G:U wobble pairs for initial damage sensing, while scanning DNA as a search complex (SC, slightly bent DNA). Remarkably, hTDG has been indicated to continuously switch between the search and interrogation conformation (IC, stronger bent DNA) during damage search. In the IC, target bases are interrogated by extrahelical base flipping, which is facilitated by protein-induced DNA bending and enhanced DNA flexibility at mismatches. AFM and fluorescence analyses revealed that the flipped base is stabilized via hTDG's arginine finger. Correct target bases are perfectly stabilized within the enzyme's catalytic pocket resulting in prolonged residence time and enhanced excision probability. To test for the generalizability of the proposed hTDG damage search model to BER glycosylases, identical studies were performed with a second glycosylase, hOgg1. The data on hOgg1, which removes structurally more stable 8-oxoguanine lesions, supported the hypothesis developed for lesion recognition by hTDG as a common strategy employed by BER glycosylases.

1.2. Zusammenfassung (Deutsch)

Die Stabilität des menschlichen Genoms wird durch DNA-schädigende Faktoren ständig bedroht. Mehrere zelluläre Mechanismen haben sich entwickelt, um verschiedene Typen von DNS-Schädigungen zu erkennen und zu entfernen. Obwohl zahlreiche und vielfältige Informationen über die dreidimensionalen Strukturen und katalytischen Mechanismen von DNS-Reparaturenzymen vorhanden sind, ist der essentielle Prozess der Suche und Identifikation von Läsionen kaum verstanden. Wie ist es möglich, dass eine kleine Anzahl an Reparaturenzymen die seltenen DNS-Schadensstellen unter Millionen von unbeschädigten Basen schnell und effizient finden kann?

Diese zentrale Frage der DNS-Reparatur habe ich mit Hilfe von Schlüssel-Proteinen aus zwei verschiedenen DNS-Reparaturmechanismen untersucht, zum einen aus der Nukleotid- (NER) und zum anderen aus der Basenexzisionsreparatur (BER). Als Beispiel für zwei völlig unterschiedliche Ansätze zur Schadenssuche, -erkennung und -verifizierung, habe ich das NER Protein Xeroderma pigmentosum group D (XPD) mit den BER Proteinen humane Thymin DNA Glykosylase (hTDG) und der humanen 8-Oxoguanin Glykosylase (hOgg1) verglichen.

Im Detail habe ich Einzelmoleküluntersuchungen mittels Rasterkraftmikroskopie (engl. ‚*atomic force microscopy*‘, AFM) und unterstützenden biochemischen und biophysikalischen Techniken angewandt. Ich habe eine einfache und optimierte Probenaufbereitungsmethode etabliert, welche es ermöglicht homogene, hochreine und lange (mehrere 100 Basenpaare) DNS-Substrate herzustellen, die für AFM Studien mit DNS-Reparaturenzymen geeignet sind. Mit Hilfe dieser Probenherstellungs-Technik kann eine einzelne, gewünschte Zielstelle an einer bestimmten Position in diese DNS-Substrate eingefügt werden. Die Verwendung dieser speziellen DNS-Substrate erlaubt eine separate Analyse von spezifischen Protein-DNS-Komplexen, die an bestimmte Läsionen gebunden sind, und unspezifischen Komplexen mit unbeschädigter DNS.

Der erste Teil dieser Arbeit behandelt die Rolle von XPD in der eukaryotischen NER. Der NER Mechanismus entfernt DNS-Schäden, welche die Helix-Struktur der DNS verzerren. Das sind zum einen krebserregende UV-Schäden, wie zum Beispiel Cyclobutan-Pyrimidindimere (CPDs) sowie sperrige DNS-Addukte. Die 5'-3' Helikase XPD wird als eines der Schlüsselenzyme bei der Schadens-Verifizierung gehandelt, was jedoch derzeit noch umstritten ist. In meinen Studien habe ich mich mit dem XPD-Protein aus dem archaischen Organismus *Thermoplasma acidophilum* (taXPD) beschäftigt. Dieses weist eine relativ große Sequenzhomologie mit dem humanen Protein auf und stellt daher ein gutes Modell für das eukaryotische XPD dar. Auf Grund von AFM Experimenten und DNS-Bindungsaffinitätsmessungen mittels Biolayer-Interferometrie (BLI), konnte XPD eine eindeutige Rolle in der DNS-Schadensverifizierung

zugesprochen werden. Meine Ergebnisse zeigten, dass die ATP-abhängige Helikase-Aktivität von XPD durch die Anwesenheit eines DNA-Schadens gehemmt wird, was zur Schadensverifizierung und Bildung von XPD-DNA-Komplexen führt, die von der Schadensstelle ‚aufgehalten‘ wurden.

Eine erfolgreiche Schadensverifizierung führt daraufhin zu ATP-abhängigen Konformationsänderungen, die sich in einer Änderung des DNA-Biegewinkels von $\sim 50^\circ$ zu $\sim 65^\circ$ an der Stelle des gebunden Proteins äußern. Es ist bemerkenswert, dass diese Änderung des DNS-Biegewinkels sowohl mit ATP und als auch mit ATPys (nicht-hydrolysierbares ATP Analog) beobachtet wurde. Dies zeigt, dass bereits die Bindung und anstatt der Hydrolyse von ATP ausreicht, um reparaturkompetente Konformationsänderungen durch XPD zu veranlassen. Darüber hinaus haben meine detaillierten Proteinbindeposition- und DNS-Biegewinkelanalysen zum ersten Mal gezeigt, dass XPD sperrige Fluoreszein-Schäden vor allem auf dem translozierten DNS-Strang erkennt, während CPD-Schäden vor allem auf dem gegenüberliegenden nicht-translozierten Strang erkannt werden. Trotz dieser unterschiedlichen Erkennungsstrategien für die zwei Schadenstypen, nehmen die beiden Schäden die gleiche Konformation im Schadensverifizierungskomplex an, was als Signal für die Rekrutierung weiterer NER Faktoren dienen könnte.

Im zweiten Teil meiner Arbeit kombinierte ich AFM-Experimente mit einem sogenannten ‚base flipping‘ Test, der auf der Fluoreszenz von 2-Aminopurine basiert, um die Schadenssuche und -erkennung durch DNS-Glykosylasen im BER-Mechanismus zu untersuchen. Als Beispiel für die weitläufige Familie der Glykosylasen wählte ich hTDG aus. hTDG schneidet Thymin und Uracil aus mutagenen G:T und G:U Fehlpaarungen heraus. Die AFM Daten zeigten, dass sich hTDG für die initiale Schadenserkennung die intrinsische Flexibilität in G:T und G:U Paaren zu Nutze macht, während die DNA als Suchkomplex geprüft wird (engl. ‚*search complex*‘, SC, leicht gebogene DNS-Struktur). Erstaunlicherweise scheint hTDG dabei kontinuierlich zwischen einem Such- und Prüfkomples umzuschalten (engl. ‚*interrogation complex*‘, IC, stärkere Biegung der DNS). Im IC werden Basen durch „flippen“ außerhalb der DNS-Helix geprüft, was durch die Protein induzierte DNS-Biegung und die erhöhte Flexibilität von Fehlpaarungen in der DNS ermöglicht wird. Die Analyse von AFM- und Fluoreszenzexperimenten brachte zum Vorschein, dass die ‚geflippte‘ Base durch den Arginin-Finger von hTDG stabilisiert wird. Die korrekten Zielbasen passen exakt in die katalytische Tasche des Enzyms und werden dort perfekt stabilisiert, was zu einer längeren Aufenthaltsdauer führt, die wiederum die Wahrscheinlichkeit erhöht, dass die Base herausgeschnitten wird. Um zu testen, ob das für hTDG vorgeschlagene Schadenssuchmodell auch allgemein für andere BER Glykosylasen gilt, habe ich die gleichen Experimente mit einer weiteren Glykosylase (hOgg1) durchgeführt, ein Protein das strukturell stabilere 8-Oxoguanin-Schäden entfernt. Die Daten für hOgg1 untermauern die Hypothese, die für die Schadenssuche von hTDG erarbeitet wurde, als eine gemeinsame Strategie von BER Glykosylasen.

2. Introduction

DNA repair is of vital relevance to maintain the stability of the human genome. Several cellular mechanisms have evolved to detect and remove specific DNA damages. In this thesis, atomic force microscopy (AFM) imaging as well as biochemical and biophysical techniques have been used to investigate two DNA repair mechanisms, nucleotide excision repair (NER) and base excision repair (BER). The NER helicase Xeroderma pigmentosum group D (XPD) and the BER proteins thymine DNA glycosylase (TDG) and human 8-oxoguanine glycosylase (Ogg1) were chosen as examples to compare different approaches for damage search, recognition and verification.

2.1 Atomic Force Microscopy (AFM) imaging of biomolecules

In the major part of my studies, the technique AFM was used to investigate protein-DNA-interactions in DNA repair. AFM can be used for direct visualization and studies of proteins, DNA and their complexes at the single molecule level (Hansma and Hoh, 1994) due to its high lateral resolution of a few nanometers (Seo and Lee, 2008). In this technique, a sharp probe fixed at the end of a flexible cantilever is scanned over a sample surface. Owing to attractive and repulsive forces arising between atoms of the probe and surface, a topographical image can be formed without the need for staining the samples (Bustamante et al., 1992, Bustamante et al., 2000, Scheuring et al., 2001).

2.1.1 Principles of AFM

During the last two decades, biological research has remarkably advanced towards investigation of processes at the level of single molecules. One of the high-resolution imaging methods is AFM, which belongs to the group of scanning probe microscopies (SPM). The basic element in these techniques is a mechanical probe, which scans the topography of a sample deposited onto a smooth surface. The invention of the scanning tunneling microscopy (STM) in 1982 (Binnig et al., 1982) gave rise to the development of a couple of SPM techniques including AFM, which was established four years later by Binnig, Quate and Gerber (Binnig et al., 1986, Santos and Castanho, 2004). As the lateral resolution in AFM is predominantly determined by the dimensions of the probe, AFM tips are very small and sharp, with a diameter of usually lower than 10 nm (Albrecht et al., 1989). The tip is fixed at the bottom end of a flexible cantilever, which is commonly made of silicon nitride (Si_3N_4) and has force constants ranging from 0.01 to 100 Nm^{-1} (Colton et al., 1997, Walters et al., 1997).

In AFM imaging, the sample surface is screened row by row in xy plane either by moving the sample stage via a piezo-scanner system (piezo driver and converters implemented in the AFM controller electronics) or by moving the AFM tip itself. Any molecular interactions arising between atoms of the tip

and the sample surface are detected from the bending of the cantilever towards (for attractive forces) or away from the surface (for repulsive forces). The most relevant interaction forces for AFM are various non-covalent forces, long range electrostatic interactions, short range attractive van der Waals forces, and very short range Pauli repulsion (Allen et al., 1997).

The deflection of the cantilever is commonly detected by a laser beam focused on the back of the cantilever and reflected onto a position sensitive photodiode. Any bending of the cantilever leads to displacement of the laser position on the split photodetector, which is translated into a differential electronic signal. This electric potential is further converted into height information for each pixel position. Usually, a feedback mechanism is applied for keeping a roughly constant distance between AFM tip and the sample stage. The sample stage is connected to a piezo scanner, which enables movement of the sample in direction of the z-axis. A rise or depression of the sample surface induces a change in the interaction force and is compensated by either approaching or removing the sample stage from the tip. This displacement of the sample stage (z value) is measured for each xy position of the sample by the feedback system and a pseudo-three-dimensional topography picture of the sample is produced by the computer software (Ando et al., 2013, Ando, 2013). In contrast to AFM imaging, which provides height information, the deflection of the cantilever is translated into information on interaction forces in AFM force spectroscopy.

Three different modes are commonly used for AFM imaging: contact, intermittent contact and non-contact mode (Figure 1).

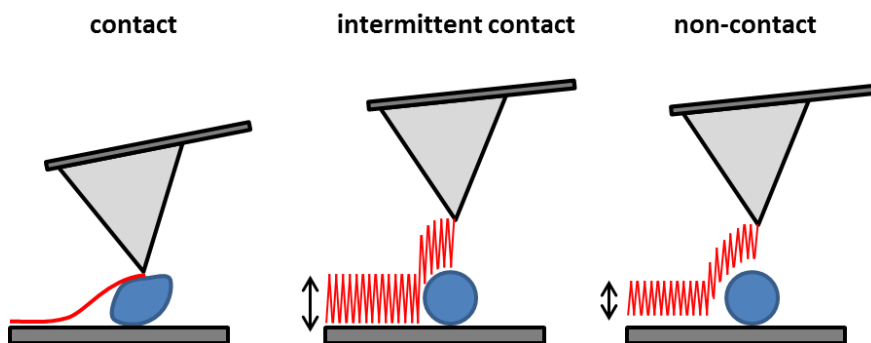


Figure 1: Imaging modes in AFM.

For details please refer to the main text (adapted from Benzanilla et al. (1994)).

In contact mode, the AFM tip is in continuous contact with the molecules in the sample and directly follows the sample topography while scanning the surface. When the AFM tip encounters either a rise or a depression in the structure, the cantilever is bent. The degree of bending is directly translated into sample height by the AFM electronics. The high lateral forces exerted on the sample during the scan are

a major drawback of contact mode imaging. These forces may cause severe damage to soft biological materials such as proteins or lead to their detachment from the surface (Binnig et al., 1986). Thus, intermittent contact mode (alternatively named tapping mode) has been developed for imaging of biological structures. Figure 2 shows the typical instrumental setup of tapping mode AFM.

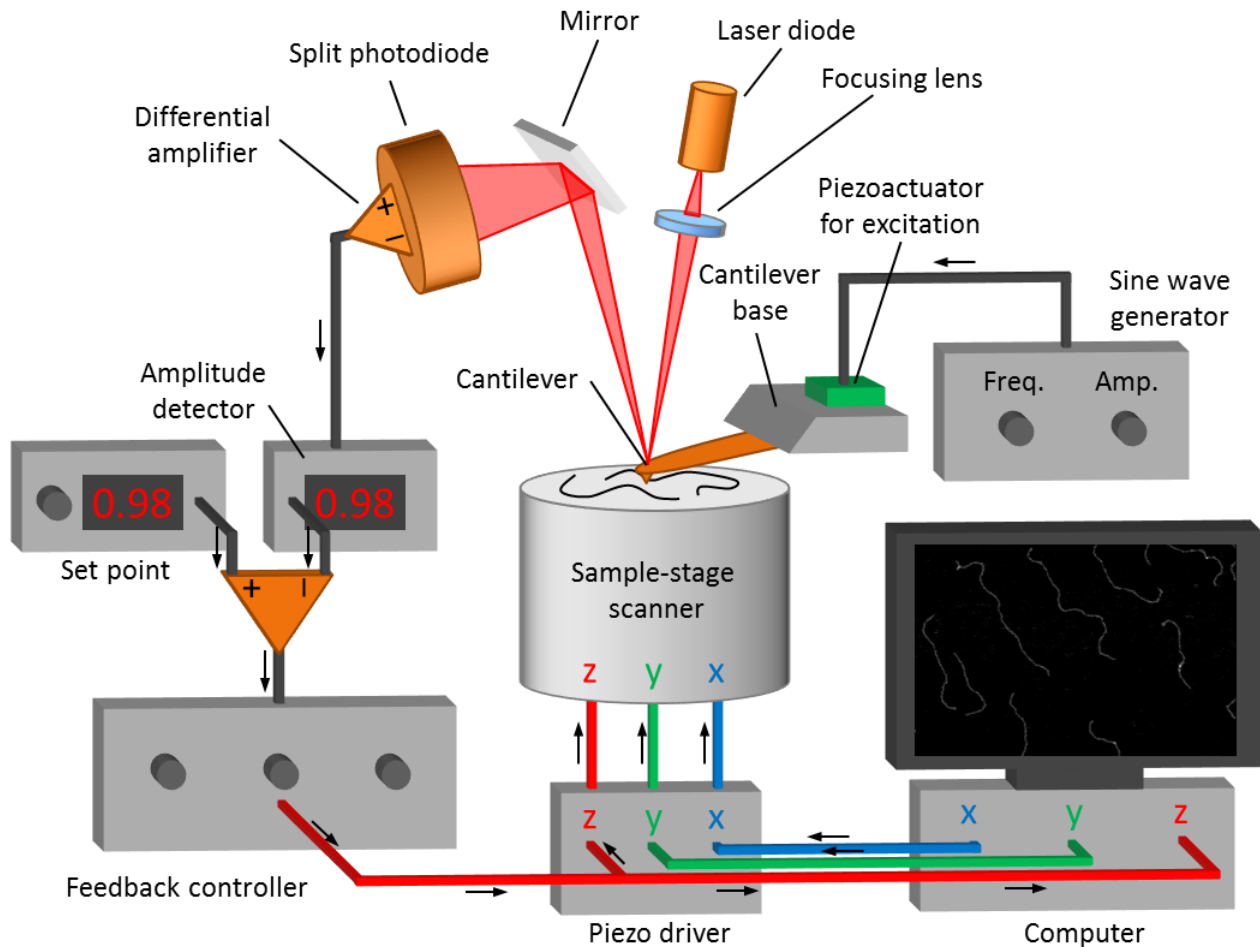


Figure 2: Typical setup of a tapping mode atomic force microscope.

A biological sample (e.g. DNA molecules) deposited onto a flat surface is scanned by a sharp, oscillating AFM tip and a three-dimensional topographical image of the sample is produced on the computer screen (for details see main text) (adapted from Ando et al. (2013)).

In general, tapping mode imaging combines high lateral resolution with minor invasiveness. In this imaging mode, the AFM tip is intermittently brought into contact with the sample surface by oscillating the flexible cantilever in z-direction at its first resonance frequency (Figure 2). To minimize lateral forces, the probe touches the sample surface only in vertical direction and at the minimum of its oscillation amplitude. The oscillating movement of the AFM stylus in a pre-defined frequency, amplitude and height baseline is driven by a piezoactuator, which excites the cantilever and is connected to a sine wave generator. Upon tapping the surface, the deflection of the cantilever is detected as a decrease in the

cantilever's oscillation amplitude (amplitude cutting) leading to a phase shift relative to the excitation signal. Using a feedback control system, the original height of the amplitude (set point value) is restored by moving the sample stage in z-direction. For each x-y-position on the scanned sample surface a corresponding signal of the feedback controller (amplitude) is recorded, which is proportional to the sample height at this specific position. Similar as described above for contact mode imaging, the feedback controller output is used for reconstruction of a pseudo-three-dimensional image of the sample topography (Ando et al., 2013, Ando, 2013). Furthermore, by detection of phase changes the tapping mode provides an additional parameter for characterization of material features (adhesion, friction) of the sample (Stark et al., 2001).

The most gentle imaging mode is non-contact imaging, which uses very small oscillation amplitudes of the AFM tip so that it does not even touch the surface. The detection is based on attractive tip-sample interactions and solely relies on measuring the amplitude increase and the change in oscillation phase. This occurs at the expense of a poor lateral resolution because the thus detected forces are very weak (Allen et al., 1997).

2.1.2 Assets and drawbacks of AFM

Biochemical and biophysical ensemble methods such as electrophoretic mobility shift assays (EMSA), surface plasmon resonance (SPR) and calorimetric assays are powerful techniques for analysis of thermodynamic equilibrium constants of protein–DNA interactions (Carey, 1991, Lohman and Bujalowski, 1991, Myszka, 2000, Oda et al., 1998, Winzor and Wills, 1995). However, they share several major limitations. To begin with, these methods are bulk measurements and their results represent an averaged sum of all molecules (Record et al., 1991), neglecting different states of individual molecules. Second, analysis of protein-DNA interaction often relies on indirect detection of a signal, which is assumed to be linearly proportional to the actual parameter of interest e.g. binding affinities measured as heat in isothermal titration calorimetry (ITC). Although this assumption is often true, in many cases a linear proportionality is not given (Lohman and Bujalowski, 1991). Third, the size of the DNA oligomers typically used in biochemical assays is often limited to a few dozen of base pairs (bp).

The single molecule technique AFM allows direct visualization of individual molecules and is able to resolve different populations of protein-DNA complexes. In contrast to conventional methods including X-ray crystallography, AFM imaging allows application of long DNA substrates of hundreds to thousands of bp length. These DNA substrates more closely resemble naturally occurring DNA strands and are long enough to stabilize tertiary structures e.g. DNA loops, triple helices or G-quadruplexes. Finally, we can

considerably reduce the interference of DNA ends on protein binding by using long DNA fragments (Buechner and Tessmer, 2013).

Besides its great advantages, AFM also holds some limitations, which are evident in comparison with the two major microscopic methods, electron microscopy and fluorescence microscopy (Table 1).

Table 1: Feasibility comparison of AFM with other types of microscopes.

Microscopy	Conditions					Direct imaging
	In-liquid specimen	High spatial resolution	High temporal resolution	Low invasiveness		
Fluorescence microscopy	●	●*	●	●		X
Electron microscopy	⊖	●	⊖	X		●
Atomic force microscopy	●	●	X	⊖		●

● Fully met; ⊖ partially met or conditional; X impossible or extensive efforts required to meet. * requires application of computer assisted superresolution fluorescence microscopy. ‘Direct imaging’ means that the morphology of an object itself is resolved (adapted from Ando (2013)).

Most conveniently, AFM does not require staining of the samples for visualization, which allows direct imaging without background or artifacts introduced by labelling. Due to its high sensitivity with spatial resolution in the range of a few nanometers, AFM offers remarkable insights into structural and functional parameters of protein–DNA complexes. EM imaging affords slightly higher spatial resolution than AFM because it is not limited by convolution of the sample with the AFM tip. Regarding the native-ness of the imaged particles, however, AFM is superior over EM. By means of rapid and easy sample preparation (see below for description), AFM imaging in liquid even enables the observation of fully hydrated molecules under near physiological conditions in buffer solution (Drake et al., 1989).

Major criticism on AFM imaging often concerns the invasiveness of the AFM probe on the scanned sample molecules. By using tapping mode, lateral forces exerted on the samples can be largely avoided. The remaining effect of the tip on the sample is given as the impulsive force, which is defined as product of the tapping force (vertical force induced by AFM tip tapping) and the duration of the acting force. The maximum tapping force exerted on a typical protein (5 nm height) in a conventional AFM setup is estimated to be ~100 pN (Ando, 2013). However, the actual effect of the AFM probe on the sample is minimal because the force acting time is very short (< 100 ns).

Although AFM imaging can provide a wealth of information about the static structure of biomolecules, the temporal resolution of a typical AFM setup is limited (~30 s at its highest level) (Ando, 2013). To resolve the dynamics of fast biological processes, high-speed AFM (HS-AFM) has been developed by Ando and colleagues offering subsecond to sub-100-ms temporal resolution (Katan and Dekker, 2011,

Ando et al., 2013, Ando, 2014, Eghiaian et al., 2014, Lyubchenko, 2014). Combining AFM imaging in liquid with the recently developed technology of HS-AFM may serve to resolve dynamic actions of protein and DNA (either the interaction of both or the two components on their own). In this newly emerged field, some groups have observed e.g. the formation of DNA structures (e.g. supercoiling, condensation), DNA folding into intricate shapes (DNA origami) and the assembly and disassembly of proteins bound and unbound to DNA (Yokokawa et al., 2006, Crampton et al., 2007, Endo et al., 2010, Suzuki et al., 2010, Shlyakhtenko et al., 2012a, Shlyakhtenko et al., 2012b).

2.1.3 Biological applications of AFM

AFM is a widely used approach for studying structure and function of biomolecules. For example, AFM imaging can be used to investigate the structure of supported membranes (for reviews see Dufrene and Lee (2000) and Balashev et al. (2001)) or even the surface of a whole cell or a virus (e.g. Ohnesorge et al. (1997) and Grandbois et al. (2000)). Moreover, AFM is an attractive approach to analyze structure and function of soluble proteins (e.g. muscular fibers of actin (Weisenhorn et al., 1990) myosin (Hallett et al., 1995), β -amyloid fibrilization, (Blackley et al., 2000, Xu et al., 2001) and membrane proteins (e.g. F_1F_0 complex of ATP synthase, (Singh et al., 1996), Bacteriorhodopsin (Moller et al., 2000)). In general, AFM studies can reveal oligomerization states, characteristic substructures of individual proteins, conformational changes, protein-protein interactions (including association constants) and the assembly of membrane proteins (e.g. (Seelert et al., 2000, Scheuring et al., 1999, Fotiadis et al., 2000, Mou et al., 1996)).

To demonstrate the strength of AFM to investigate conformational changes of proteins, the AFM experiments with the protein collybistin performed in a recent collaboration with the group of H. Schindelin will be described here exemplarily (see Figure 3). The brain-specific GDP/GTP exchange factor collybistin acts at inhibitory postsynapses as adaptor protein by linking the scaffold protein gephyrin to plasma membrane phosphoinositides (Papadopoulos and Soykan, 2011, Soykan et al., 2014). Collybistin has been proposed to switch between an open/active and closed/inactive conformation, which regulates the gephyrin content of inhibitory postsynapses. In this protein interaction network, conformational states of collybistin are regulated by the adhesion protein neuroligin-2, which stabilizes the open conformation by competing with an intramolecular interaction in collybistin between the SH3 and DH/PH domains favoring the closed conformation. AFM experiments were performed with collybistin wildtype (wt) and its variant E262A (AFM images in Figure 3), which is indicated to disturb the intramolecular interactions in collybistin and to modulate the equilibrium between the two conformational states towards the open/active conformation. The AFM images displayed mixtures of circular and elongated

shapes for both proteins. To quantify the observed elongation in the variant, the molecules were described by ellipses and the ratio of the minor/major axes of the protein was determined as an indicator for conformational changes. Statistical analyses of the individual particle shapes revealed an equilibrium between three major conformational states of the wt protein and the variant, which were classified by their axial ratios as compact (~ 1.22), intermediate (~ 1.45), and elongated states (~ 1.75). The intermediate (axis ratio 1.45) and the most extended form (axis ratio 1.75) were significantly preferred by the variant suggesting stabilization of the open/active (elongated) conformation in the mutant E262A.

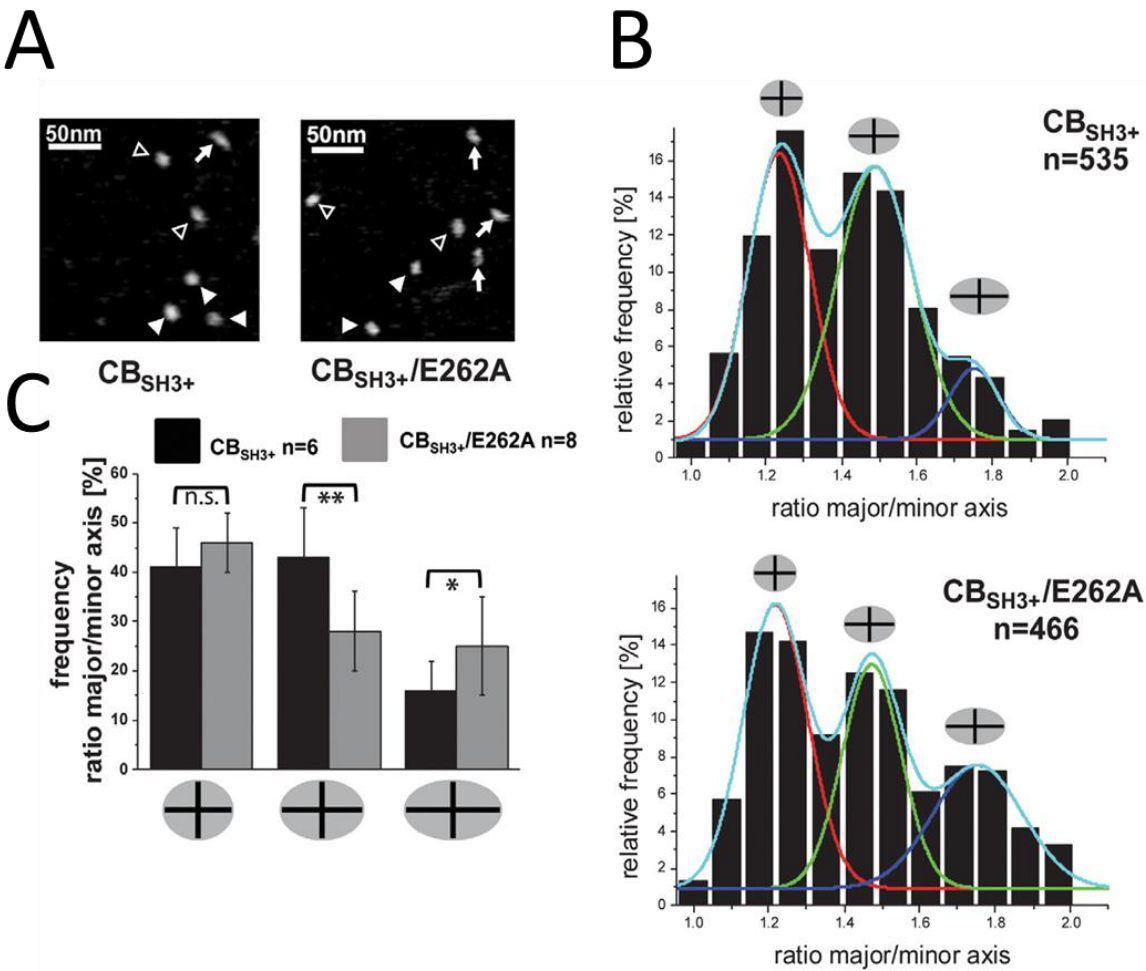


Figure 3: Conformational states of a neuro-protein visualized in AFM experiments.

A) AFM images of CB_{SH3+} wt (left) and the E262A variant (right). Filled arrow heads (compact form), open arrow heads (intermediate form), and arrows (elongated form) point to different particle shapes. B) Relative frequencies of occurrence of CB particle shapes as classified by the different major/minor axis ratios featuring compact (red), intermediate (green), and elongated (blue) states for the WT and E262A samples. One representative experiment is shown for each protein with n representing the number of particles per experiment. C) Quantification of frequencies of compact, intermediate, and elongated particles confirmed that the intermediate form (axis ratio 1.45) and the most extended form (axis ratio 1.75) are significantly preferred by the variant as calculated by pairwise Student's *t*-test. This research was originally published in Soykan et al. (2014) © EMBO, John Wiley & Sons.

Most interesting for the present work are of course AFM imaging experiments on protein-DNA interactions, which can reveal details of the molecular recognition of a specific target site in DNA by a protein. In general, AFM experiments in air with dried samples and AFM imaging in liquid in physiological buffer solutions can be distinguished. In this thesis, the well-established technique of AFM imaging in air was applied to characterize protein-DNA interactions in DNA repair. For more than one decade, many groups have performed extensive AFM studies in air with a diversity of protein-DNA systems involved in essential cellular processes such as transcription, replication and DNA repair (Verhoeven et al., 2001, Chen et al., 2002, Seong et al., 2002, Verhoeven et al., 2002, Janicijevic et al., 2003, Wang et al., 2003, Crampton et al., 2006, Wu et al., 2010, Jiang and Marszalek, 2011, Shlyakhtenko et al., 2011, Shlyakhtenko et al., 2012b, Billingsley et al., 2012).

2.2 DNA repair as a fundamental concept for genome maintenance

2.2.1 Sources and consequences of DNA damage

The integrity of our genome is continuously endangered by DNA damaging factors. It has been estimated that a total of 10^5 DNA lesions are induced in one mammalian cell during one day (Svenberg et al., 2011). In general, the presence of a DNA lesion may affect the stability of our genome twofold (Figure 4).

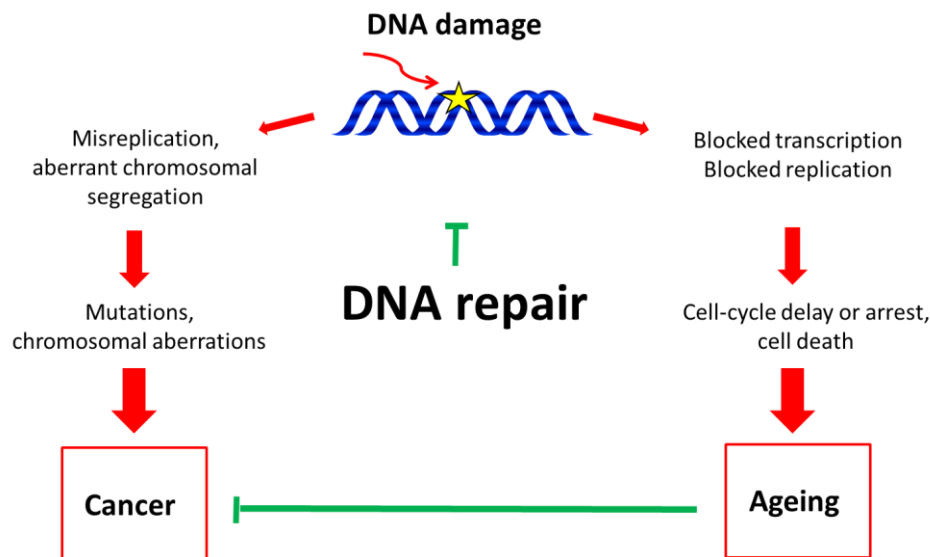


Figure 4: The twofold danger of DNA damages.

For details please refer to the main text (adapted from Hoeijmakers (2009)).

If DNA repair is not working correctly or a damaged template has been replicated, persistent changes in our genetic code can occur in the form of mutations or chromosomal aberrations. These may give rise to activation of proto-oncogenes and inactivation of tumor-suppressor genes increasing the risk of cancer

(Bartek et al., 2007). Alternatively, DNA lesions may block the essential cellular processes transcription and replication leading to cellular senescence and cell death. Such lesion-induced apoptosis is a double-edged sword: on the one hand it is an important mechanism to fight cancer, but on the other hand it may contribute to accelerated aging (Schumacher et al., 2008) (as reviewed by Hoeijmakers (2009)).

DNA damaging factors originate from two main sources. First, exogenous agents from our environment such as UV-irradiation, ionizing radiation and genotoxic substances contained in cigarette smoke and chemotherapeutics result in structural alterations of DNA. Second, the stability of DNA is threatened by endogenous substances from our metabolism including nitrogen species and reactive oxygen species (ROS) produced during respiration and lipid peroxidation. Due to these DNA damaging factors and due to the chemical nature of DNA in an aqueous solution, nucleic acids are highly vulnerable under physiological conditions. Spontaneous reactions (mostly hydrolysis and nucleophilic attack) can lead to formation of abasic sites and deamination of cytosine, adenine or guanine (Hoeijmakers, 2001, Hoeijmakers, 2009). Overall these factors may induce a plethora of DNA damages. The most common types are summarized in Figure 5.

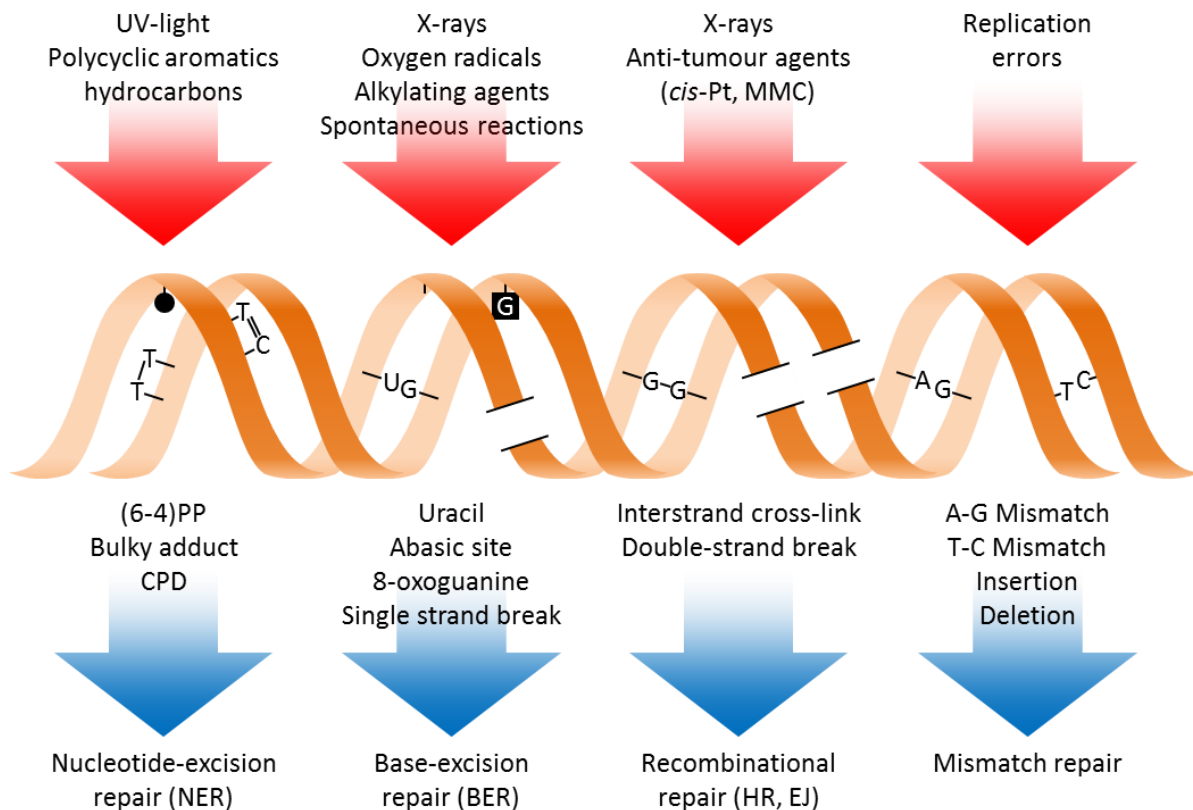


Figure 5: Sources and types of DNA damages.

DNA damaging agents (top) can cause a variety of DNA lesions as displayed by some typical examples (middle). Different DNA repair mechanisms have evolved for removal of a specific subset of lesions (bottom) (adapted from Hoeijmakers (2001)).

Considering the battery of DNA lesion types, a single DNA repair mechanism would be completely overburdened. Thus, several different DNA repair mechanisms have evolved to maintain the stability of a cell's genetic information. Although each of the pathways has specialized for removal of a specific subset of lesions, DNA repair systems are a highly interwoven network with partly overlapping specificities.

2.2.2 DNA damage excision repair pathways

As DNA is the genetic material in all living cells, the problem of DNA injury has existed *ab initio*. In contrast to other biomolecules, damaged DNA cannot be completely replaced by a new molecule, but solely relies on repair of single nucleotides to remain intact. Thus, DNA repair systems must have developed very early in evolution and are conserved in all three kingdoms of life, which are comprised of eukaryotes, prokaryotes and archaea (Woese and Fox, 1977). Remarkably, the main principles of DNA repair are usually similar across the prokaryotic and eukaryotic border. Among eukaryotes, the structure and function of most DNA repair proteins is largely conserved and in some cases even throughout bacteria. This high conservation is a great benefit in studying and understanding the mechanisms of DNA repair (as reviewed by Taylor and Lehmann (1998)).

In eukaryotes, the most relevant DNA repair pathways include: nucleotide excision repair (NER), base excision repair (BER), double-strand break repair, and mismatch repair (see Figure 5). Here, I will focus on the two DNA lesion excision mechanisms of NER and BER. As a common feature, DNA injury in these two processes affects only one of the DNA strands. In general, DNA excision repair pathways operate through 'cut-and-patch' type reactions, in which the DNA damage (with or without some flanking DNA sequences) is excised and removed. The complementary non-damaged DNA strand is used as a template for the repair synthesis to restore the original DNA sequence. The most fundamental steps in the two pathways are DNA lesion search and recognition over the excess of nonspecific bases. These processes will be discussed exemplarily for eukaryotic NER and BER in this thesis.

In NER, a short fragment of single-stranded DNA (24-32 nucleotides) surrounding the DNA lesion is excised and removed. Most DNA substrates for NER originate from exogenous sources (except for some oxidative lesions) and embrace the class of helix-distorting lesions, which may disrupt normal base-pairing and generally obstruct normal transcription and replication. For example, bulky carcinogenic photoproducts induced by UV irradiation, such as cyclobutane pyrimidine dimers (CPD) as well as cis-platin-adducts and benzopyren-adducts caused by chemotherapy and smoking can be removed by the NER pathway (Lindahl, 1993, Gillet and Scharer, 2006). Its extraordinarily wide range of substrate specificities renders NER an excellent model mechanism for the study of recognition strategies for different target sites in DNA repair. The high versatility of NER is based on its two-step damage

recognition strategy (also called 'bipartite substrate discrimination'). The initial damage sensing step circumvents direct recognition of the lesion itself, but instead uses a set of common features shared by many different DNA damages such as helix distortion for lesion discrimination. In a second and ATP-dependent step, the presence of a chemical modification is verified by a multi-protein damage recognition complex, a process also referred to as 'kinetic proofreading'.

The BER mechanism recognizes and removes small, non-helix-distorting chemical base lesions, which predominantly arise from endogenous sources. Alterations in bases can be due to oxidation, alkylation or deamination (Lindahl, 1993), which provide a high miscoding potential, but may not interfere with transcription and replication as rigidly as the NER mechanism. Thus, BER is particularly essential to prevent mutagenesis caused by mispairing or breaks in DNA during replication. BER involves 4-5 steps, with the initial step performed by DNA glycosylases. DNA glycosylases recognize a damaged or mismatched base within the huge excess of normal bases with high specificity by employing a base-flipping mechanism to identify their target site in their active site pocket (Friedman and Stivers, 2010). In addition, these enzymes catalyze base excision by cleaving the glycosidic bond between the base and the DNA sugar-phosphate backbone.

It is remarkable that DNA glycosylases do not require any metal cofactors or exogenous energy source for fulfilling these tasks and solely rely on DNA base pairing energies. Lesion recognition and excision in BER is then followed by a common multi-protein pathway usually involving an AP-endonuclease for processing of the abasic site, a DNA polymerase for repair synthesis and a DNA ligase for nick sealing. In striking contrast to NER, the broad substrate specificity in BER is based on a variety of different DNA glycosylases, which have specialized for a narrow subset of base alterations, instead of a multi-subunit damage recognition complex. Hence, glycosylases are interesting target proteins to compare the two different approaches for damage search and recognition strategies in the two DNA excision repair pathways, BER and NER.

2.3 Nucleotide excision repair (NER)

The general mechanism of NER is conserved in prokaryotes and eukaryotes and can be grouped into four common steps: a) initial damage detection b) damage verification c) damage removal and d) DNA re-synthesis. The prokaryotic NER system mainly involves the UvrABC machinery. The DNA is initially probed for the presence of lesions by the sliding complex of UvrAB. Upon lesion encounter, UvrA loads UvrB onto DNA and then dissociates from the UvrB-DNA complex. After successful damage verification by UvrB, a stable pre-incision complex is formed by recruitment of UvrC, which cuts DNA 3' and 5' to the

lesion. Subsequently, the incised oligonucleotide as well as UvrB and UvrC are released by the helicase UvrD to allow DNA re-synthesis and ligation by DNA polymerase I and DNA ligase, respectively (Van Houten et al., 2005, Truglio et al., 2006).

Accounting for the higher organization in eukaryotes, up to 30 different proteins cooperate in a sophisticated cascade in eukaryotic NER. Due to the complexity of NER itself and its crosstalk with other cellular processes such as transcription and the cell-cycle, defects in NER proteins can lead to inherited, severe diseases: Xeroderma Pigmentosum (XP), Cockayne Syndrome (CS) and Trichothiodystrophy (TTD). Patients suffering from XP show high photosensitivity and high proneness to skin cancer. XP is often coupled to CS, a second syndrome with high UV-light sensitivity that is accompanied by additional symptoms such as neurological abnormalities and short stature. TTD phenotypes are also sensitive to UV-irradiation and associated with neurological disorders and premature ageing (as reviewed by Lehmann (2001), Gillet and Scharer (2006) and Egly and Coin (2011)). The following chapter describes the mechanism of eukaryotic NER in detail with special focus on the process of DNA damage verification.

2.3.1 Mechanism of eukaryotic NER

Eukaryotic NER can be divided in two subpathways, which are schematically depicted in Figure 6. In global genomic repair (GG-NER) the whole genome is scanned for the presence of helix distortions, whereas transcription-coupled repair (TC-NER) removes lesions on actively transcribed genes induced by a blockage of RNA polymerase II (RNA Pol II) (as reviewed by Hoeijmakers (2001), Hoeijmakers (2009) and Marteijn et al. (2014)).

In GG-NER, the heterotrimeric complex of XPC, the RAD23 homologue B (RAD23B) and centrin 2 (CETN2) continuously probes DNA for the presence of helix-distortions (step 1, upper pathway) (Masutani et al., 1994, Sugasawa et al., 1998, Sugasawa et al., 2001, Nishi et al., 2005). Most NER lesions strongly destabilize the DNA duplex causing small unpaired DNA regions (bubbles). Due to its high affinity for junctions between single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) (Min and Pavletich, 2007), XPC binds to the complementary strand opposite the DNA damage, recognizing lesions in an indirect manner. For detection of CPD lesions, which induce only minor helix destabilization (Sugasawa et al., 2001, Reardon and Sancar, 2003) and could not be detected by XPC on its own, XPC is assisted by the ultraviolet radiation DNA damage binding protein complex (UV-DDB, step 2) (Chu and Chang, 1988, Wakasugi et al., 2002, Scrima et al., 2008). By intercalation between the DNA strands with its β -hairpin domain, XPC stabilizes the ssDNA region and was also proposed to enlarge the bubble size to allow binding of the large ring-like transcription factor II H (TFIIH) (Sugasawa et al., 2001, Maillard et al., 2007,

Min and Pavletich, 2007). Prior to recruitment of TFIIH, RAD23B dissociates from the heterotrimeric XPC complex (step 3).

RNA Pol II transiently interacts with UV-stimulated scaffold protein A (UVSSA), ubiquitin-specific-processing protease 7 (USP7) and Cockayne syndrome protein (CSB) during transcription elongation (step 1, lower pathway). Upon damage encounter, the TC-NER pathway is activated (Hanawalt and Spivak, 2008, Vermeulen and Foustieri, 2013) by RNA Pol II stalling at the lesion site thereby enhancing the affinity of CSB (step 2). To allow unhindered DNA repair, CSA is recruited to form a complex with CSB and RNA Pol II (Foustieri et al., 2006, Schwertman et al., 2012), which induces RNA Pol II to move backwards, down-stream from the damage (backtracking) (step 3).

In the next step, the formation of the pre-incision complex, GG-NER and TC-NER merge into a common pathway (step 4). A central player is the ten-subunit TFIIH-complex with the Cdk-activating kinase (CAK) subcomplex (three subunits) (Yokoi et al., 2000, Volker et al., 2001, Riedl et al., 2003). The TFIIH complex fulfils a dual role as it functions in NER as well as in transcription initiation (as reviewed by Compe and Egly (2012)). Upon binding of the TFIIH complex to DNA-bound XPC, the CAK sub-complex readily dissociates (Coin et al., 2008) from the TFIIH core complex (seven subunits). The 3' specific endonuclease XPG is recruited simultaneously either in complex with TFIIH or separately (Dunand-Sauthier et al., 2005, Zotter et al., 2006, Ito et al., 2007). The TFIIH-complex includes the 5'-3'-helicase XPD and the ATPase XPB (Riedl et al., 2003). XPB has been suggested to anchor TFIIH at the site of the damage by its ATPase activity (Oksenyshyn et al., 2009), which has been shown to be stimulated by XPC (Bernardes de Jesus et al., 2008) in the GG-NER pathway. In the pre-incision step, XPB is indicated to undergo ATP-dependent conformational changes to enlarge the DNA bubble size by its ATPase activity (Coin et al., 2007). XPD is proposed to further unwind the DNA around the damage by its helicase activity in 5'-3' direction (Tapias et al., 2004, Compe and Egly, 2012). Concomitantly, the evolving undamaged ssDNA strand is stabilized by the single-stranded DNA binding protein replication protein A (RPA) (de Laat et al., 1998). Apart from this proposed role in NER, XPD has been indicated to fulfil an essential function in damage verification likely supported by the ATPase activity of XPB and the factor XPA, which has a proposed high affinity for single-stranded, chemically modified DNA (Camenisch et al., 2006, Sugawara et al., 2009).

After the presence of a chemical modification is verified successfully, the 'point of no return' is reached. XPA and RPA direct the 5' specific endonuclease XPF-ERCC1 complex to the 5' site of the lesion (step 5). XPF-ERCC1 cleaves at the 5' of the damage (Fagbemi et al., 2011) and the trimeric proliferating cell nuclear antigen (PCNA) ring loads onto DNA. XPG cuts at the 3' side of the lesion and an oligomer with a length of 24-32 nucleotides containing the lesion is excised (step 6) (Huang et al., 1992, Moggs et al., 1996). The missing ssDNA gap is filled by the DNA re-synthesis machinery including DNA Pol δ , DNA Pol κ

or DNA Pol ϵ recruited by PCNA (step 7). Finally, DNA-Ligase I or DNA-Ligase III seals the remaining nicks between the newly synthesized DNA-strand and the original strand (step 8) (Moser et al., 2007, Ogi et al., 2010).

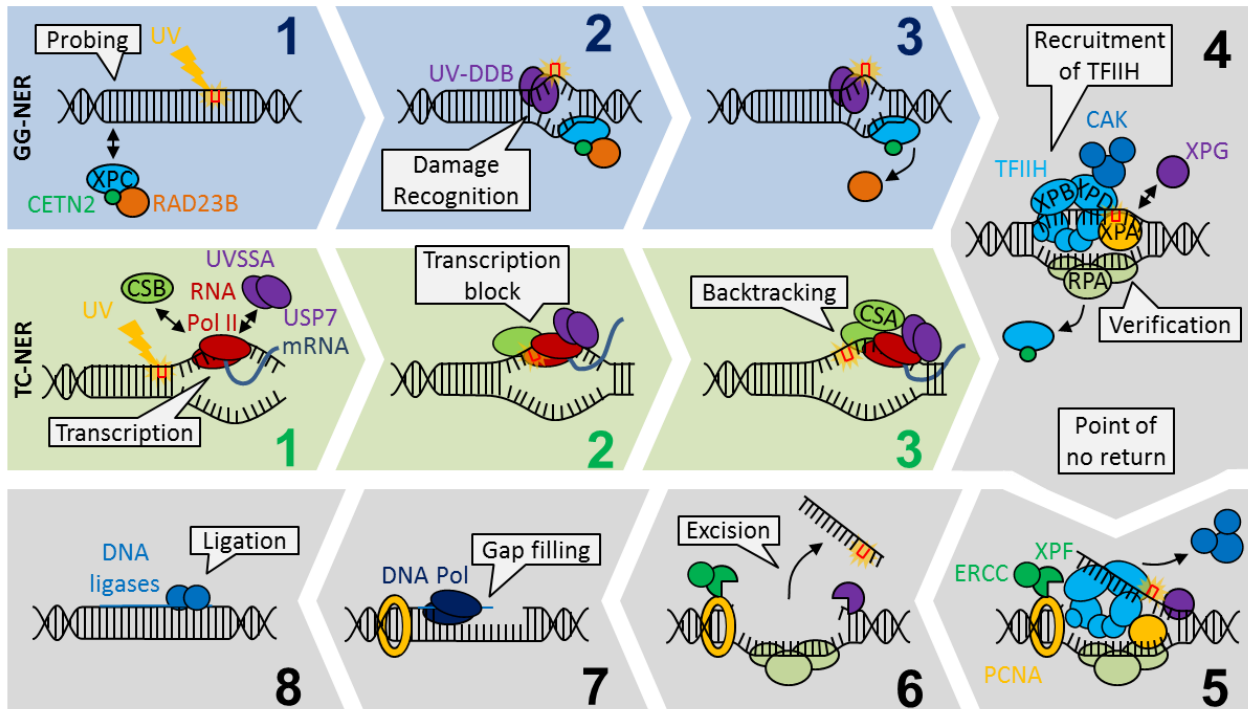


Figure 6: Schematic of the eukaryotic NER mechanism.

DNA lesion search and initial damage sensing (steps 1-3) can either occur through the GG-NER or the TC-NER pathway. After initial damage detection, GG-NER and TC-NER merge into a common pathway. Subsequent to damage verification (step 4), a short stretch of DNA containing the lesion is excised and removed (steps 5-6). Then, the original DNA sequence is restored by the DNA re-synthesis machinery using the non-damaged strand of the DNA duplex as template for repair synthesis (steps 7-8). For details see main text (adapted from Marteijn et al. (2014)).

2.3.2 DNA damage verification in eukaryotic NER

DNA damage verification is the most fundamental step in NER. On the one hand, cytotoxic or mutagenic DNA lesions have to be verified and removed as efficiently as possible and on the other hand futile repair of non-damaged bases has to be avoided (Dip et al., 2004). Currently, it is highly debated whether XPD is involved in damage verification and if this task is fulfilled solely by TFIH or by TFIH in cooperation with further NER factors such as XPA (see above, step 4 in NER pathway) (Camenisch et al., 2006, Sugawara et al., 2009).

In eukaryotes, XPD serves as structural component within the TFIH complex by bridging the TFIH core complex with the CAK complex via its interaction with p44 (structural protein in the core complex) and MAT1 (CAK) (Reardon et al., 1996). More importantly, XPD plays a functional role in NER, but is dispensable for transcription initiation (Tirode et al., 1999). In the pre-incision step, XPD is essential to

enlarge the repair bubble around the lesion by its helicase activity (see above, NER mechanism step 4), which is a prerequisite for dual incision by XPG and XPF.

In general, DNA helicases such as XPD are motor proteins, which translocate on nucleic acids in a predetermined direction utilizing the free energy provided by binding and hydrolysis of nucleoside triphosphates to unwind double-stranded nucleic acids (Singleton et al., 2007). As a common feature, helicases share seven helicase motifs (walker motif I, Ia, II, III, IV, V and VI) constituted of conserved amino-acid sequences. XPD is an ATP-dependent 5'-3' helicase and belongs to the superfamily II (SF2) helicases.

In an early study using a yeast homolog of XPD (Rad3), it has been shown that DNA translocation of XPD was stalled at lesion sites *in vivo* (Naegeli et al., 1992) implicating XPD in damage verification. Recent biochemical studies with an archaeal XPD homolog from *Ferroplasma acidarmanus* (facXPD) confirmed that XPD's helicase activity was blocked by the presence of a lesion *in vitro* and discovered that XPD formed a stable complex with DNA at the site of the damage (Mathieu et al., 2010, Mathieu et al., 2013). These observations argued for a role of XPD in damage recognition as well as verification. In contrast, a different group showed that XPD from the archaeal organism *Sulfolobus acidocaldarius* (saXPD) was not stalled by DNA lesions addressed by the NER pathway (Rudolf et al., 2010). Overall, these studies demonstrate that DNA damage verification by XPD is still a matter of considerable debate and will require further investigation.

2.3.3 Archaeal XP homologs as a model for eukaryotic NER

DNA repair proteins from archaeal homologs have been widely used as model systems for eukaryotic NER in biochemical and structural approaches (see above in 2.3.2 and Nishino et al., 2003, Roberts et al., 2003, Fan et al., 2006, Fan et al., 2008, Liu et al., 2008, Richards et al., 2008, Wolski et al., 2008, Mathieu et al., 2010, Ma et al., 2011), because archaeal proteins are often thermophilic (preferred temperature > 45°C) and thus are more amenable to biochemical investigation than their human counterparts. As DNA repair mechanisms are highly conserved between the three domains of life (see 2.2.2), archaeal and eukaryotic NER proteins often share a high sequence homology.

Archaea were identified as the third kingdom of life (Woese and Fox, 1977) and can be grouped into euryarchaea and crenarchaea based on phylogenetic analysis of their ribosomes (Cox et al., 2008). Considering the natural habitats of archaea in particular those of hyperthermophilic species (preferred temperature >80°C) e.g. living in volcanos and hot springs, archaea have to face tremendous challenges in preserving their genomic stability. Obviously, the question is raised, which genome maintenance mechanisms are employed in archaea (Grogan, 2000, Grogan, 2004).

Bioinformatic homology searches for eukaryotic and prokaryotic DNA repair proteins among the wealth of genomic sequence information, which is available for archaea, revealed that archaeal repair proteins more closely resemble their eukaryotic than their prokaryotic counterparts (Grogan, 2004). Table 2 provides an overview of selected DNA repair proteins among different sequenced archaeal organisms:

Table 2: Distribution of selected DNA repair proteins in sequenced archaeal genomes.

Species	XPF	XPB	XPD	FEN-1	Dpo4	Photo -lyase	UvrAB C	MutS/ MutL	Temp (°C)
<i>Aeropyrum pernix</i>	1	1	1	1					95
<i>Pyrobaculum aerophilum</i>	1	2	1	1					95
<i>Sulfolobus solfataricus</i>	1	2	1	1	1	1			80
<i>Sulfolobus tokodaii</i>	1	2	2	1	1	1			80
<i>Sulfolobus acidocaldarius</i>	1	2	1	1	1	1			75
<i>Methanopyrus kandleri</i>	1			1					100
<i>Pyrococcus furiosus</i>	1	2	1	1					100
<i>Pyrococcus horikoshii</i>	1	2	1	1					100
<i>Pyrococcus abyssi</i>	1	2	1	1					100
<i>Thermococcus kodakarensis</i>	1	1	1	1					85
<i>Nanoarchaeum equitans</i>	1	1	1	1					85
<i>Methanococcus jannaschii</i>	1		1	1					85
<i>Archaeoglobus fulgidus</i>	1	1		1					85
<i>Methanothermobacter thermautotrophicus</i>	1			1		1	•		70
<i>Picrophilus torridus</i>		1	1	1	1	1			60
<i>Thermoplasma acidophilum</i>		1	1	1					60
<i>Thermoplasma volcanium</i>		1	1	1					60
<i>Methanosarcina acetivorans</i>	1	1	1	1	1	1	•	•	<40
<i>Methanosarcina mazei</i>	1	1	1	1	1	1	•	•	<40
<i>Haloarcula marismortui</i>	1	2		1	1	3	•	•	<40
<i>Halobacterium NRC-1</i>	1	2	1	1	1	2	•	•	<40

Archaeal species are grouped by growth temperature. Hyperthermophilic organisms (>80°C) are marked in red, thermophilic in yellow (45-80°C) and mesophilic in blue (<45°C) • indicates the presence of bacterial DNA repair proteins in archaeal organisms (adapted from Kelman and White (2005)).

Only a few archaeal organisms particularly the mesophilic species (temperatures < 45°C preferred, compare Table 2) possess the bacterial UvrABC repair system (Kelman and White, 2005) likely by lateral gene transfer from a bacterial donor (Grogan, 2000). In the majority of archaea, eukaryotic NER proteins were found such as the helicase XPD and the ATPase XPB and the endonucleases XPF and the Flap structure-specific endonuclease 1 (FEN-1) (Table 2), a homolog of the eukaryotic endonuclease XPG,

which also functions in other DNA repair mechanisms (e.g. in the BER mechanism, see 2.4.1) and replication. However, not all archaeal species embrace the same set of XP-homologs and archaeal genomes generally lack XPC and XPA, two essential factors required for eukaryotic NER. It is evident from the overview provided by Table 2 that the distribution of DNA repair proteins is highly diverse among archaea.

The existence of a NER-type patch repair pathway in archaea has been only shown for the archaeon *Methanothermobacterthermautotrophicus*, which contains the UvrABC machinery and exhibited a prokaryotic-type excision pattern of NER (Ogrunc et al., 1998), but not for eukaryotic-type NER so far. Hence, the pure presence of homologous XP-proteins in archaeal organisms does not necessarily argue for their role in NER and it is likely that XP-proteins may function in other cellular mechanisms such as DNA replication in archaea (Kelman and White, 2005, Rouillon and White, 2011). Despite the poor understanding of NER in archaea, the archaeal XP homologs, however, represent attractive NER models due to their significant sequence similarity to the eukaryotic NER proteins (Grogan, 2000, Grogan, 2004).

2.3.4 Objectives in lesion recognition and verification by XPD

As described in the previous section, archaeal XP homologs can be employed as an adequate model for eukaryotic NER. In this thesis, the XPD homolog from the euryarchaeon *Thermoplasma acidophilum* (taXPD), whose complete genome sequence has been published (Ruepp et al., 2000), served as model to investigate damage verification in NER because the protein shares a high sequence identity of 39 % with its human counterpart (as defined by BLAST, NCBI). Compared to other archaeal species (Table 2), *T. acidophilum* contains only a few XP proteins including XPD, XPB and the XPG homolog FEN-1 and completely lacks prokaryotic-type UvrABC proteins or further homologous proteins from other repair systems. In contrast to eukaryotic XPD, which functions as part of the TFIIH complex, taXPD is a monomeric protein in solution and has not been shown to form stable interactions with other protein partners so far (Rouillon and White, 2011). Hence, XPD from *T. acidophilum* may provide a simplistic model for damage verification processes. Due to its thermophilia (growth temperature of ~ 60 °C), the proteins from this archaeal organism are more applicable to biochemical and structural approaches than their human equivalents. Importantly, detailed knowledge about the enzyme's three-dimensional structure can be gained from the crystal structure of taXPD (Wolski et al., 2008). In the crystal structure of taXPD four domains can be distinguished (Figure 7): The two RecA-like domains, (helicase domain 1 and 2, respectively) responsible for helicase activity, a highly conserved arch domain and an iron sulfur cluster domain (Wolski et al., 2008).

In addition, the X-ray structure of taXPD supports the molecular understanding of the point mutations leading to three inherited, severe diseases XP, CS combined with XP and TTD (compare 2.3) in the human homolog. Using mutational and biochemical studies, the implications of disease-related taXPD variants (corresponding to mutations in the human homologue) on the protein's functions such as DNA binding, helicase and ATPase activity were investigated (Wolski et al., 2008, Kuper et al., 2012).

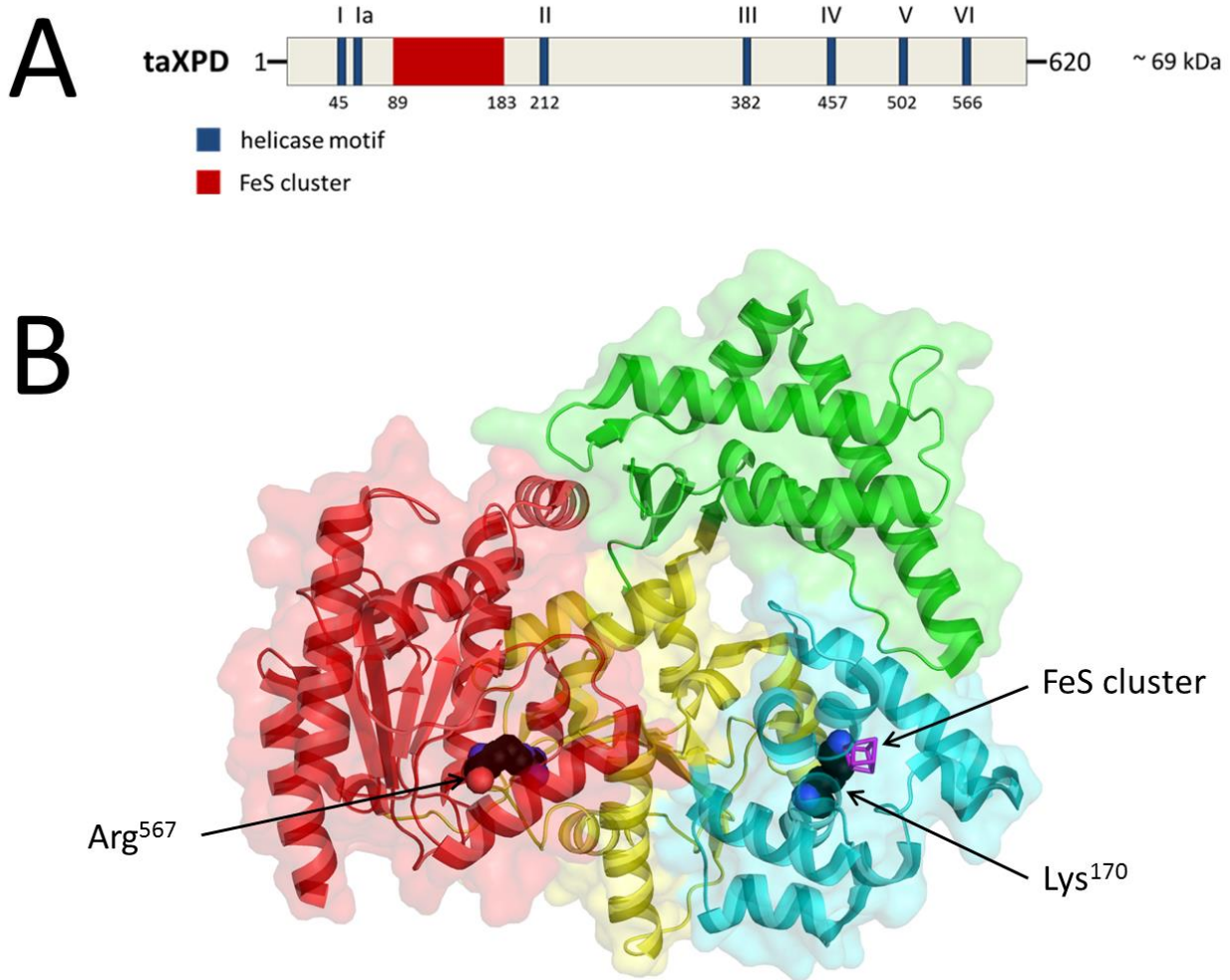


Figure 7: Primary and crystal structure of taXPD showing mutated residues as used in this work.

A) Schematic representing the primary structure of taXPD. The helicase motifs (I-VI), the FeS cluster and the molecular weight of the protein are indicated (adapted from Wolski et al. (2008)). B) The crystal structure of taXPD (PDB-ID: 2VSF) (adapted from Wolski et al. (2008)) is shown in ribbon presentation with the RecA-like helicase domains 1 and 2 in yellow and red, respectively. The FeS cluster domain is depicted in cyan and the arch domain in green. The FeS cluster is shown in purple. The positions of the residues Lys170 and Arg567 (depicted in space-filling representation) are indicated in the structure as the XPD variants K170A and R567W were used as controls in the AFM studies.

Based on a recent co-crystal-structure of taXPD bound to a short fragment of ssDNA, a model of the possible path of DNA across the enzyme could be derived (Kuper et al., 2012). The first three domains (helicase domain 1, the FeS cluster domain and the arch domain) form a donut-shaped structure with a

pore (10 Å diameter). In the XPD-DNA model, the ssDNA is threaded in the pore. The route of the modeled DNA passes by a narrow pocket, which is located in the wall of the central pore directly next to the FeS cluster (Wolski et al., 2008). The pocket is shaped by the strictly conserved residues Arg88 and Tyr166 on the one side and by Tyr185 on the other side. An undamaged base, either a purine or a pyrimidine, would perfectly fit in this narrow hole and could be stabilized via Van-der-Waals interactions with the arginine and the two tyrosine residues. In contrast, a bulky DNA damage such as cis-platin- or benzopyren-adducts would be excluded from this pore. Due to these findings and the close proximity to the FeS cluster, this narrow pocket has been proposed to be involved in damage verification.

Interestingly, a recent study by Barton and co-workers showed that [4Fe-4S] cluster containing proteins (e.g. XPD) may employ such clusters as a damage sensor (Sontz et al., 2012). Within perfectly matched unmodified DNA (Figure 8, left), DNA-mediated charge transfer (CT) (green arrow) can occur between two FeS cluster containing proteins of similar potential (Fe^{3+} , dark red and violet oval), which are capable of performing CT.

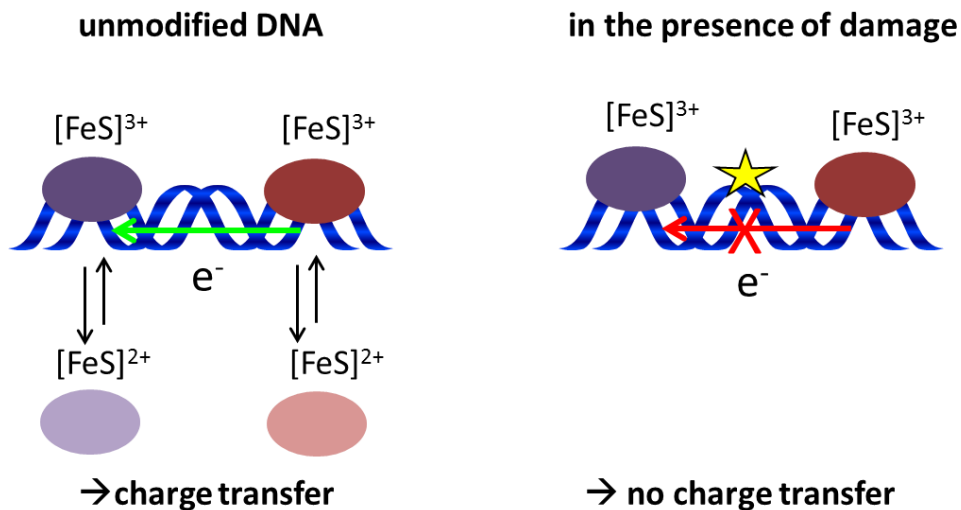


Figure 8: Model for DNA mediated charge transfer.
For details see main text (adapted from Sontz et al. (2012)).

In principle, stacking interactions between base pairs in the DNA double helix (π - π interactions) can serve as a 'molecular wire' for rapid, one-dimensional charge separation (or DNA-mediated CT) (Eley and Spivey, 1962). DNA CT between two DNA-bound proteins is driven by free guanine radicals formed under oxidative stress, which stimulates reduction of the FeS clusters (Fe^{2+} , light red and light violet ovals). Reduced proteins possess lower affinity to DNA and dissociate from DNA. In the presence of an intervening damage or mismatch (Figure 8, right, yellow star), DNA CT is not able to proceed (red arrow) and DNA-bound repair proteins remain in the oxidized state (Fe^{3+} , dark red and violet oval). Instead of

dissociation from DNA, the repair proteins remain bound close to the DNA lesion and can translocate towards the damage for repair on a lower time scale (Romano et al., 2011). Thus, inhibition of DNA-mediated CT by the presence of a lesion can lead to redistribution of repair proteins in the vicinity of the lesion, which has been shown for XPD (from the archaeon *Sulfolobus acidocaldarius*) and Endo III (BER glycosylase from *E. coli*) in AFM imaging experiments on long DNA fragments (>1 kb) (Romano et al., 2011, Sontz et al., 2012). Thus, it can be envisioned that the mechanical readout of DNA bases is coupled to redox sensing by the iron-sulfur cluster.

Although the crystal structures of taXPD can reveal the enzyme's general architecture and residues involved in protein-DNA interactions, the understanding if and how XPD may achieve lesion recognition and verification still remains elusive. To gain more detailed insights into these processes from a mechanistic perspective, XPD-DNA interactions were investigated in single molecule AFM imaging experiments. In contrast to X-ray crystallography, AFM imaging allows application of long DNA substrates of hundreds to thousands of bp length, which more closely resemble the *in vivo* occurring DNA strands. Thus, AFM can resolve different populations of XPD at the molecular level in a single experiment. Furthermore, different experimental conditions e.g. the presence or absence of ATP and various DNA substrates can be more readily tested.

In addition, in my studies I compared DNA interactions of two taXPD variants to the wt protein by AFM. Firstly, the residue Arg⁵⁶⁷, which is located near the ATP-binding site in the crystal structure (Figure 7) was substituted by tryptophane. The resulting taXPD variant R567W has been indicated to be deficient in ATP binding and hydrolysis and corresponds to R666W in the human homolog, which is related to the disease XP (Wolski et al., 2008). Secondly, the taXPD variant K170A was applied, which has recently been suggested to possess helicase hyperactivity and elevated ATPase activity compared to the wt protein (Kuper et al., 2012). As the mutated Lys¹⁷⁰ residue is located in close vicinity to the FeS cluster (Figure 7), which is proposed to be involved in damage verification as described above, the K170A variant can be expected to possess a damage verification phenotype in the AFM studies.

To account for the high substrate versatility of NER, I compared damage verification by XPD for two structurally different NER lesions (Figure 9). First, a CPD damage was used as representative for UV-photoproducts, induced by UV-light-driven dimerization of adjacent thymines. In contrast to other photolesions such as (6-4) pyrimidine-pyrimidone photoproduct (6-4 PP), CPD causes only minor helix distortion (Kim and Sancar, 1995, McAteer et al., 1998, Lee et al., 2004). Second, a fluorescein damage, an artificial substrate often used for damage recognition studies in NER (Rudolf et al., 2010), served as representative of bulky DNA-adducts.



Figure 9: DNA substrates for lesion recognition studies with taXPD.

The NER lesion CPD (left) is caused by UV-light-driven dimerization of adjacent thymines and is related to minor helix distortion. The artificial DNA substrate fluorescein (right) functions as a representative for bulky DNA-adducts. In the DNA substrates used for AFM, the fluorescein damage is attached to a thymine base by a short linker.

2.4 Base excision repair (BER)

2.4.1 Principles and mechanism of BER

A variety of cytotoxic and mutagenic base lesions are initially detected and removed by DNA glycosylases, which employ a base-flipping mechanism to identify a target site within the huge excess of normal bases (Friedman and Stivers, 2010). Based on their catalytic mechanism of base excision, DNA glycosylases can be divided in two classes (Hegde et al., 2008, Schermerhorn and Delaney, 2014).

Monofunctional glycosylases (e.g. TDG) use an activated water molecule for nucleophilic attack of the base-sugar (N-glycosidic) bond of the everted base creating an intact abasic site in the DNA (Figure 10, step 1A). Abasic sites entail a great risk for the cells: they are susceptible to ssDNA breaks and have a tendency to induce intrastrand crosslinks with DNA or crosslinks of DNA with proteins (Sczepanski et al., 2010). Therefore, these cytotoxic DNA sites are protected by monofunctional glycosylases in humans, until a downstream AP endonuclease (APE1) arrives (step 2A) (Vidal et al., 2001). APE1 cleaves the phosphoribose backbone at abasic sites, inducing a nick with 3'-OH and 5'-deoxyribose phosphate (dRP) termini (Barzilay et al., 1995, Erzberger and Wilson, 1999). Following cleavage by APE1, DNA pol β cleans the remaining DNA ends by converting the 5'-dRP to a 5'-phosphate using its dRP lyase activity (step 3A) (Beard and Wilson, 2006, Prasad et al., 2010).

In contrast, bifunctional glycosylases (e.g. NTH1 and OGG1) utilize the ϵ -NH₂ of a lysine or the amino group of an N-terminal proline for nucleophilic attack (step 1B) and harbor an additional intrinsic AP lyase (DNA strand cleavage) activity for processing of the abasic site. In a concerted step, a transient Schiff base between the amino group of the protein and the C1' of deoxyribose is formed both to excise the base lesion and to cleave the DNA backbone (McCullough et al., 1999, Dalhus et al., 2011). Via β -elimination (step 2B), a 3'-phosphor- α,β -unsaturated aldehyde (3'PUA) and a 5'-phosphate terminus are generated. In the next step, 3'-PUA is converted to a 3'-OH group by APE1 (step 3B).

Alternatively, some bifunctional DNA glycosylases (e.g. NEIL1 and NEIL2) also utilize an N-terminal proline as nucleophile for removal of the deoxyribose residue yielding a 3'-phosphate terminus at the nicking site ($\beta\delta$ -elimination, step 2C) (Zharkov et al., 2003). As the product of the $\beta\delta$ -elimination is a poor substrate for APE1, the polynucleotide kinase (PNK), which is highly expressed in mammals, dephosphorylates the 3' phosphate by its 3'-phosphatase activity (step 3C) (Meijer et al., 2002).

Downstream from these steps, BER can continue in two different modes, either as short-patch BER (SP-BER, also named single nucleotide BER (SN-BER)) (Kubota et al., 1996) or as long patch BER (LP-BER) depending on the status of the 5' terminal group and on other factors. For example, if the 5'-dRP group is modified to 5' phosphate, the dRP lyase activity of pol β can be inhibited and the LP-BER pathway is preferred (Horton et al., 2000, Sung et al., 2005).

In LP-BER, polymerase β , δ , or ϵ incorporates 2-8 deoxynucleotides at the gap site (step 4A) (Singhal et al., 1995) introducing a displaced single-stranded flap of DNA. The flap is removed by the 5'-flap endonuclease 1 (FEN-1) (step 5A) (Prasad et al., 2000) leaving a nick, which is sealed by DNA Ligase I (step 6) (Taylor et al., 2011). Pol β , δ , or ϵ , FEN-1 and DNA Ligase I are recruited by the sliding clamp PCNA (see NER mechanism in 2.2.2) (Moldovan et al., 2007), which is loaded onto duplex DNA by the replication factor-C (RFC) (Levin et al., 2000).

In SP-BER, a single nucleotide is incorporated to the 3'-OH of the nick by pol β , which interacts with the X-ray repair cross-complementing protein 1 (XRCC1) (step 4B) (Caldecott, 2003). Finally, the nick is sealed by DNA Ligase III α (Gao et al., 2011, Simsek et al., 2011), which also requires interaction with XRCC1 (step 5B).

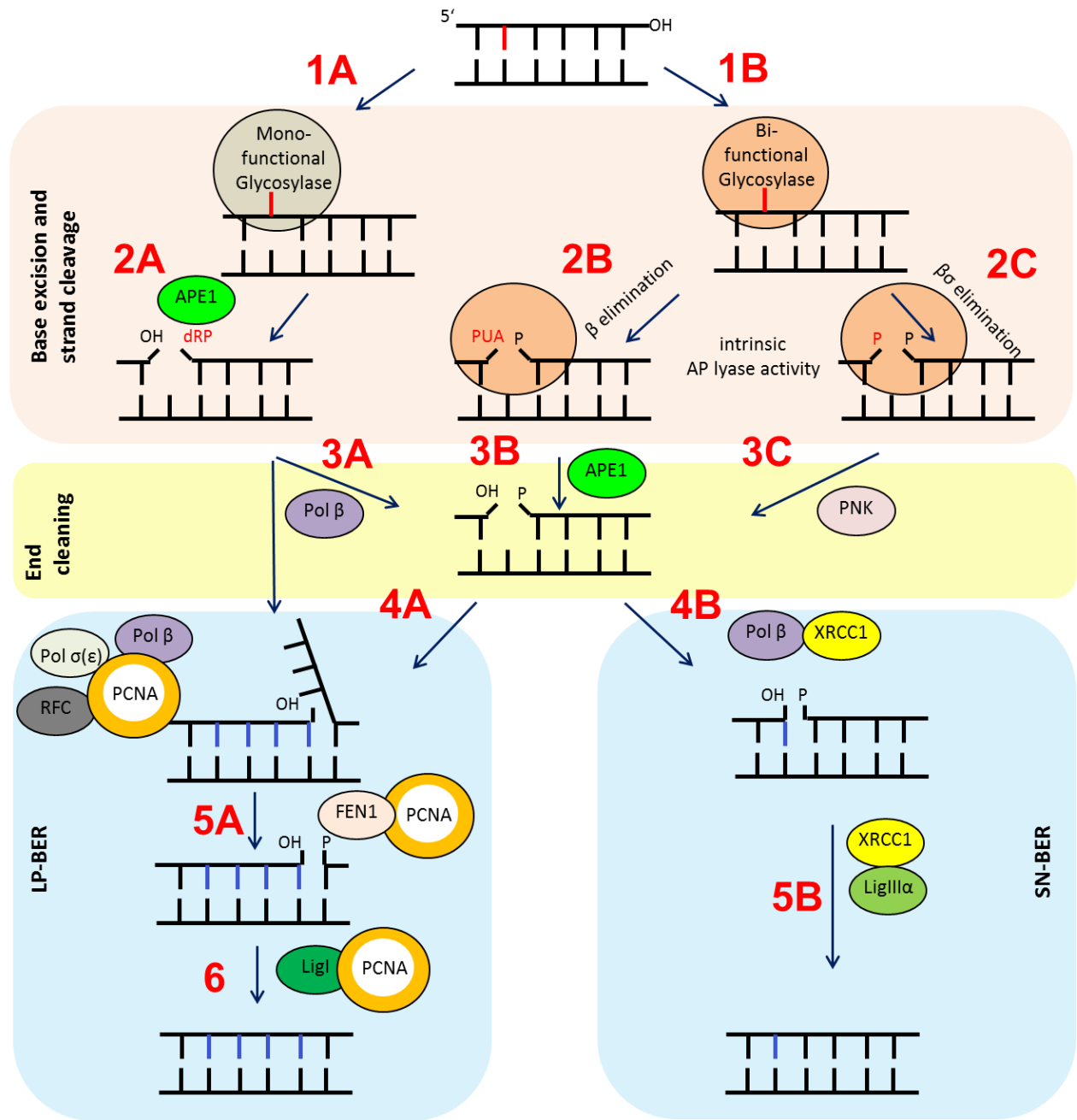


Figure 10: Schematics of BER subpathways.

The base lesion or mismatch is represented by a red line. Divergent pathways for base excision and strand cleavage (red) merge into common steps for end processing of terminal groups ('end cleaning', yellow). DNA repair synthesis (cyan) can either occur via LP-BER or SN-BER. Newly inserted 'correct' nucleotides are depicted as blue lines. For other details please refer to the main text (adapted from Hegde et al. (2008)).

2.4.2 DNA lesion search and recognition by DNA glycosylases

So far, it remained largely controversial how BER enzymes detect their target sites within the huge excess of undamaged bases. In general, DNA glycosylases have to compromise between the needs to minimize the time for lesion search and to maximize the selectivity for removal of their target base (Zharkov et al., 2010). Under the *in vivo* conditions of high DNA concentration, these DNA repair enzymes have developed the ability to bind nonspecific DNA with moderate affinity (Berg et al., 1981, Bruinsma, 2002). One-dimensional diffusion in contact with undamaged DNA (sliding) and three-dimensional transfer among DNA segments (hopping) may then facilitate rapid lesion detection (Halford and Marko, 2004, Blainey et al., 2006, Hu et al., 2006). Final target identification requires contacts with residues in the catalytic center of the enzyme, which are achieved by flipping of the damaged base into an extrahelical state. Many DNA glycosylases (e.g. UDG, hOgg1, or AlkA) (Stivers and Jiang, 2003) display a high density of polar phosphodiester interactions with the damaged strand and only a few contacts with the undamaged strand leading to DNA backbone chain distortion (bending). Such protein-induced DNA bending may function to introduce torsional stress on the damaged base to lower the energetic barrier for base eversion (Ramstein and Lavery, 1988).

Furthermore, in computational studies damaged DNA was found to be intrinsically more flexible, facilitating bending and flipping at damaged sites (Fuxreiter et al., 2002, Seibert et al., 2002, Imhof and Zahran, 2013, Yin et al., 2014). Currently, these processes are less well understood for nonspecific DNA and the question whether DNA glycosylases bend and flip undamaged DNA during lesion search is still a matter of considerable debate (McCullough et al., 1999, Friedman and Stivers, 2010). The mechanisms underlying enzymatic base flipping have been studied for some glycosylases (e.g. UDG, T4-Pdg, hOgg1, AlkA) by different techniques such as equilibrium fluorescence, stopped-flow and quenching experiments and diffusion-decelerated fluorescence correlation spectroscopy combined with molecular dynamics simulations (Chen et al., 2002, Walker et al., 2006, Bellamy et al., 2007, Zharkov et al., 2010, Yin et al., 2014). A major question is whether the proteins fulfill an active (inducing base flipping) or passive (stabilizing the flipped state) role in nucleotide flipping. For instance, the UNG glycosylase, has been reported to capture transiently emerging bases (passive flipping) at non-target sites (Cao et al., 2004, Cao et al., 2006, Parker et al., 2007, Friedman et al., 2009). The glycosylase hOgg1 and its bacterial homolog MutM, however, have been indicated to employ an active extra-helical base interrogation and flipping mechanism (Nelson et al., 2014, Qi et al., 2009). The type of exploited base eversion mechanism (active or passive) may hence vary for different glycosylases. Moreover, not only base extrusion but also the entire process of damage search in BER appears to be optimized for the individual energetic requirements of the particular target sites of a specific glycosylase.

2.4.3 Thymine DNA glycosylase (hTDG)

In my studies, the question of DNA lesion search strategies of DNA glycosylases was addressed by atomic force microscopy (AFM) imaging focusing on the specific example of the human thymine DNA glycosylase (hTDG). The monofunctional DNA glycosylase hTDG belongs to the uracil DNA glycosylase (UDG) protein superfamily (Pearl, 2000) and is involved in BER as well as epigenetic gene regulation being responsible for active DNA demethylation (Wu et al., 2010, Nabel et al., 2012). It shows a strong repair activity for uracil from G:U mismatches as well as for oxidized forms of 5-methylcytosine (mC) created by Tet enzymes (5-formylcytosine, 5fC and 5-carboxylcytosine, 5caC) (He et al., 2011, Maiti and Drohat, 2011). In addition, hTDG restores CpG sites by selective removal of thymine from G:T mismatches caused by deamination of mC. Notably, hTDG avoids futile repair of undamaged DNA, by excising thymine, a *normal* base, with 18000 fold greater activity from G:T mismatches than from A:T pairs (Morgan et al., 2007).

So far, several co-crystal structures of the hTDG catalytic core domain (hTDG_{cat}) in complex with DNA have been solved, which provided valuable insights into the active site of the enzyme. Due to its largely unstructured N-terminal domain, there are no crystal structures of hTDG full length available so far. An early X-ray structure shows hTDG bound to the abasic site analog tetrahydrofuran (THF) simulating the protein in the product state (Maiti et al., 2008). Recent structures of hTDG_{cat} captured the enzyme in the lesion recognition state bound to the substrate analog 2'-deoxy-2'-fluoroarabinouridine (U^F) (Maiti et al., 2012) (Figure 11) or 5cac (Hashimoto et al., 2012), which was flipped into the active site but not cleaved. Interestingly, in two of these structures (Maiti et al., 2008, Maiti et al., 2012) the enzyme crystallized in a 2:1 complex with DNA, with one subunit bound to the target site (to THF or G:U^F, respectively) (specific complex) and one subunit bound to undamaged DNA (nonspecific complex) directly next to the specific complex. Biochemical assays and kinetic studies confirmed dimerization (of the catalytic domain as well as the full length protein) under conditions of high and saturating [hTDG], but showed that a monomer of hTDG was fully capable of DNA lesion recognition and base excision (Maiti et al., 2008, Morgan et al., 2011).

Both hTDG_{cat} structures (bound to an abasic site analog and to G:U^F (Figure 11) revealed several residues essential for nucleotide flipping and catalysis (Maiti et al., 2008, Maiti et al., 2012). The strictly conserved residue Arg²⁷⁵ forms the tip of an arginine finger, which has been implicated in promotion and/or stabilization of nucleotide flipping in TDG. Based on the co-crystallographic studies and further pre-steady-state kinetics experiments, the arginine finger has been proposed to penetrate the DNA minor groove and to fill the void created by base extrusion (Maiti et al., 2008, Maiti et al., 2009, Maiti et al., 2012). Although the chemistry of base excision catalysis has not been studied thoroughly so far, structural and biochemical studies suggested that the strictly conserved residue Asn¹⁴⁰ (in TDG and MUG

enzymes) is essential for the chemical step of this enzymatic reaction. Furthermore, it has been suggested that Asn¹⁴⁰ is not involved in the steps prior to base excision and is dispensable for DNA substrate binding (Maiti et al., 2009).

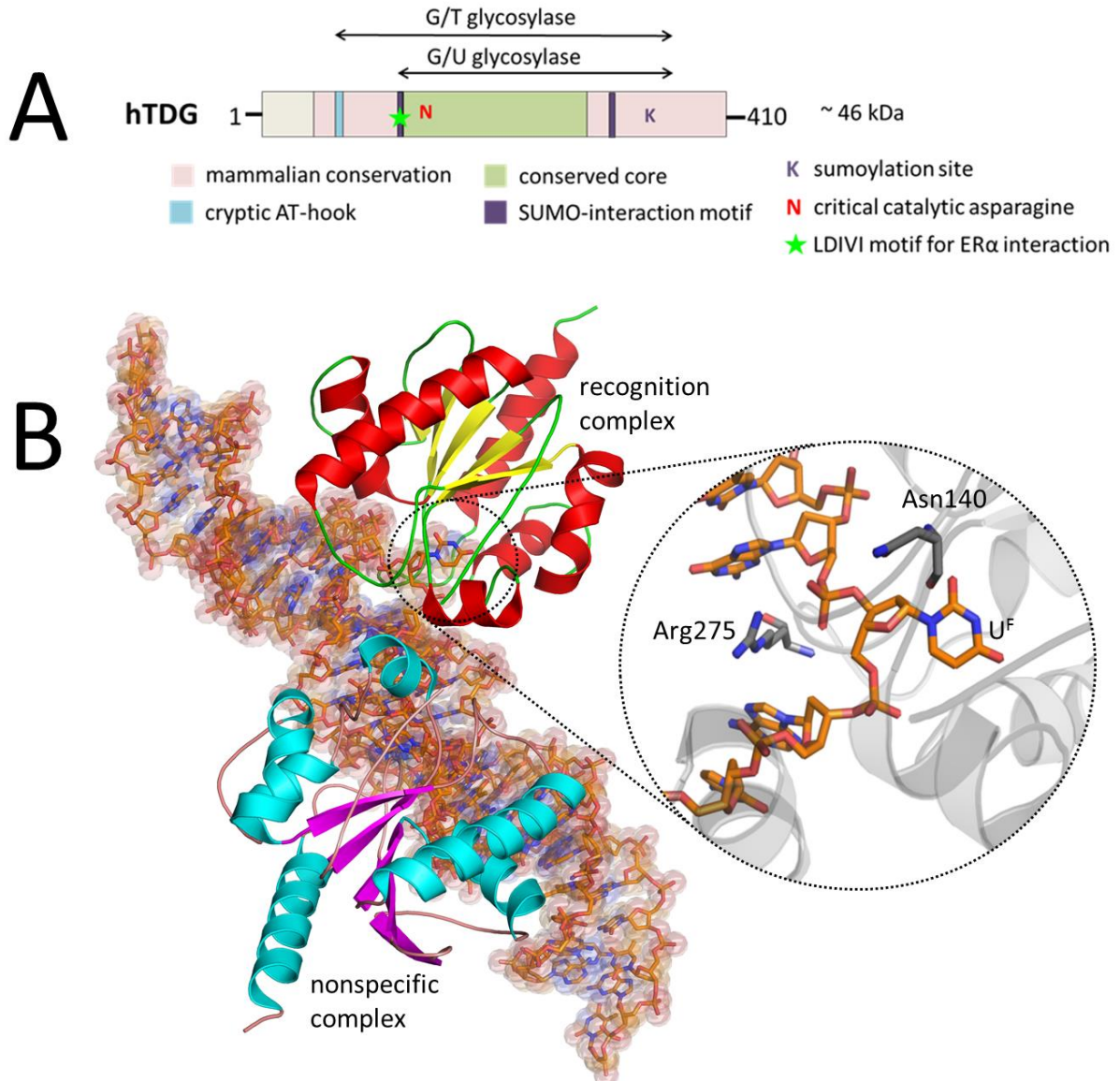


Figure 11: Primary structure and co-crystal structure of hTDG bound to a uracil analog.

A) Schematic representing the primary structure of hTDG. Important structural motifs and residues are indicated by different colors as described in the figure legend (adapted from Cortazar et al. (2007)). B) The co-crystal structure of the catalytic core domain of hTDG bound to DNA (in stick representation) reveals a hTDG homodimer with one subunit bound at the specific site (non-cleavable uracil analog, 2'-deoxy-2'-flouroarabinouridine, U^F) and the second subunit bound in close proximity at nonspecific DNA (PDB-ID: 3UFJ). The close-up view of the catalytic center shows U^F flipped into the active site. The side chain Asn¹⁴⁰ is involved in the catalytic step of base excision. Arg²⁷⁵ penetrates the minor groove and fills the void from the orphaned base. (adapted from Maiti et al. (2012)).

2.4.4 8-oxoguanine DNA glycosylase (hOgg1)

To investigate the generality of this lesion search mechanism among DNA glycosylases, lesion search strategies for hTDG, I further compared them with those of human 8-oxoguanine DNA glycosylase (hOgg1), which addresses target sites with completely different structural features. hTDG and hOGG1 are two representatives of the vast glycosylase family, which address target sites with significantly different structural features. While hTDG predominantly repairs an unstable G:T and G:U wobble pair, hOGG1 mainly removes 7,8-dihydro-8-oxoguanine (8-oxoguanine, oxoG), which can form Watson-Crick like base pairing. OxoG is caused by ROS via spontaneous oxidation of guanine (see 2.2.1) and differs at only two atoms in the chemical composition from the normal base, guanine (Wallace, 2002, Evans et al., 2004). Due to guanine's low redox potential, oxoG is one of the most common DNA oxidation products. In addition, 8oxoG is also one of the most deleterious base lesions (Steenken and Jovanovic, 1997, Burrows and Muller, 1998, Dedon, 2011). Besides its ability to form a Watson-Crick pair with cytosine, oxoG can also pair with adenine as a Hoogsteen pair during replication. A Hoogsteen base pair is an alternative base pairing within DNA duplexes, which can be formed by hydrogen bonds in the major groove and differs from the usual geometry found in Watson-Crick pairs. In this alternative DNA conformation, the distance between the C1' atoms of the two nucleobases is smaller and the angle between the two glycosidic bonds is larger than in regular base pairs (Hoogsteen, 1963, Aishima et al., 2002, Nikolova et al., 2011). Misread as thymine by the replicative DNA polymerization machinery, the resulting oxoG:A may lead to a G/C → T/A transversion mutation after one further round of replication (Shibutani et al., 1991, Michaels and Miller, 1992, Grollman and Moriya, 1993).

To elucidate the impressive task of recognizing and removing oxoG, hOgg1's lesion specificity and catalytic mechanism have been investigated thoroughly in structural and biochemical studies. Thus, this well-characterized glycosylase represents an attractive model for studying lesion detection by DNA glycosylases in AFM. A total of 27 different crystal structures of hOgg1 and its complexes with various DNA substrates have been solved so far revealing essential residues for lesion recognition and catalysis (e.g. Bruner et al., 2000, BJORAS et al., 2002, Norman et al., 2003, Banerjee et al., 2005, Radom et al., 2007, Dalhus et al., 2011, Crenshaw et al., 2012).

The bifunctional glycosylase hOgg1 belongs to the HhH-GPD superfamily, with a helix-hairpin-helix (HhH) motif followed by a Gly/Pro-rich loop terminating in an Asp residue as central structural feature (Thayer et al., 1995, Nash et al., 1996, Bruner et al., 2000, BJORAS et al., 2002). For base excision, the ϵ -NH₂ group of Lys²⁴⁹ attacks the C1' atom of the oxoG nucleotide. Via Schiff base formation and β -elimination, the DNA strand is cleaved at the 3'-side of the abasic site creating 3'-PUA and a 5'-phosphate terminus (see 3.2.1) (Girard et al., 1997, Nash et al., 1997). The co-crystal structure of hOgg1 K249Q, a catalytically

inactive variant with intact substrate recognition activity, bound to oxoG:C, reveals that the DNA is severely kinked by 70° at the target site (Figure 12) (Bruner et al., 2000). The base lesion is extra-helically flipped from the DNA duplex into a dedicated lesion recognition pocket of hOgg1 enabling access to active site residues (Phe³¹⁹, Cys²⁵³, Gly⁴², Gln⁴³, Phe⁴⁵ and Gln³¹⁵) essential for DNA substrate recognition and catalysis. The X-ray structure of a covalently trapped complex of hOgg1 with an undamaged guanine (G-ligand) revealed that the nonspecific base (G) was completely flipped out of the DNA helix, but into an exo-site, which is located in close proximity (0.5 nm distance) from the active site pocket. This finding suggests that hOgg1 discriminates G from oxoG by extrahelical interrogation, while rejecting the undamaged base from the active site. From a biological perspective, this strategy is effective to prevent futile incision of normal bases (Banerjee et al., 2005).

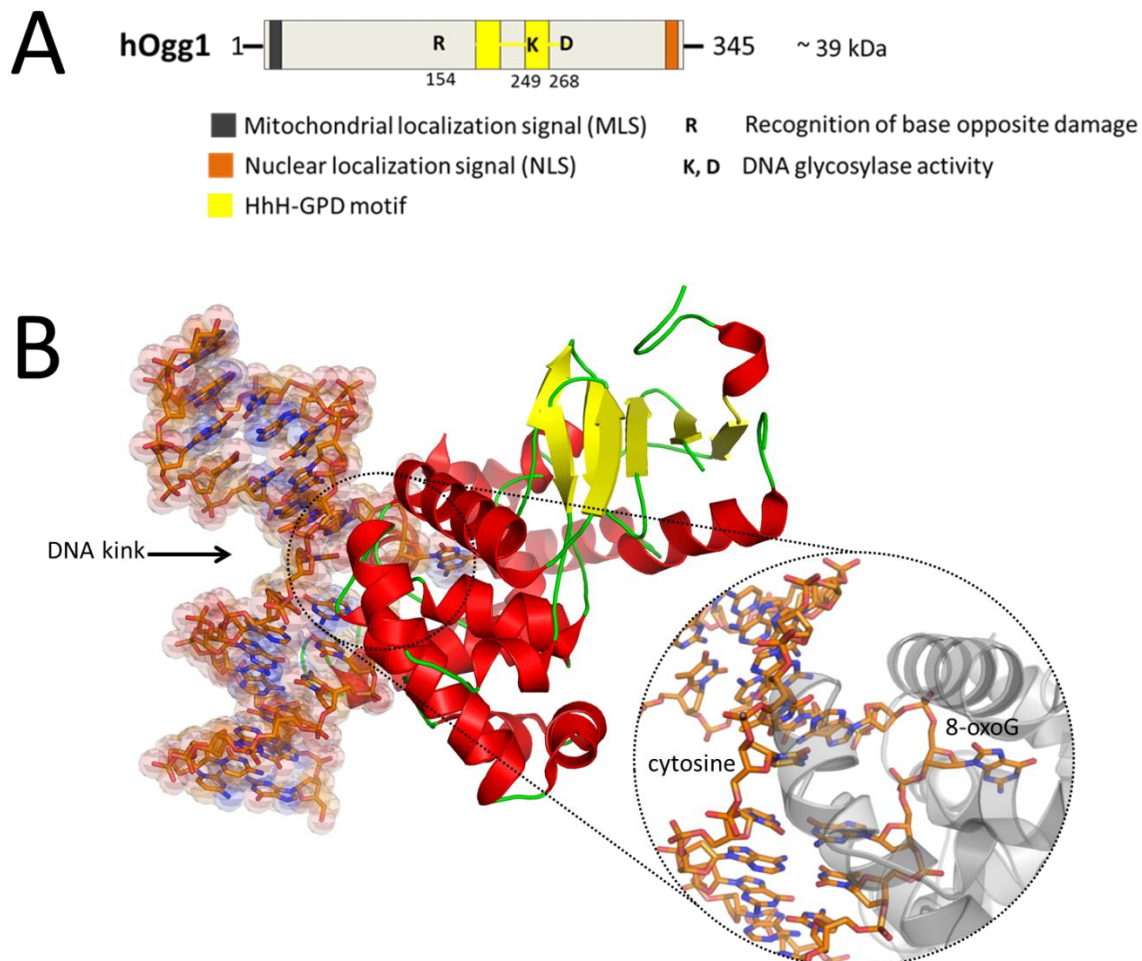


Figure 12: Primary structure and co-crystal structure of hOgg1 with 8oxoG flipped in the active site.

A) Schematic representing the primary structure of hOgg1. Important structural motifs and residues are indicated by different colors as described in the figure legend (adapted from Hashiguchi et al. (2004)). B) hOgg1 in complex with a 15 bp DNA substrate containing a 8oxoG:C lesion (PDB-ID: 1EBM). The target base is completely flipped out of the DNA duplex into the active site of the enzyme, whereas the complementary cytosine remains in the double helix. hOgg1 strongly bends DNA by 70° at the plane of the oxoG:C base pair (adapted from Bruner et al. (2000)).

2.4.5 Objectives in BER lesion search and recognition

Although the crystal structures of hTDG and hOgg1 provide a wealth of information about the residues involved in base flipping, lesion recognition and base excision, they can only present a static snapshot of the enzyme-DNA complex. In contrast, the dynamics of base flipping in particular for undamaged nonspecific bases are less well understood and it remains elusive, for instance, whether hTDG flips DNA bases in an active or passive manner.

In the AFM studies, long DNA substrates (500 – 2000 bp) were applied, either non-damaged homoduplex DNA or DNA fragments that contain a specific lesion at a well-defined position. For hTDG, a G:U and a G:T mismatch in a CpG context were used as target sites for hTDG in AFM (Figure 13).

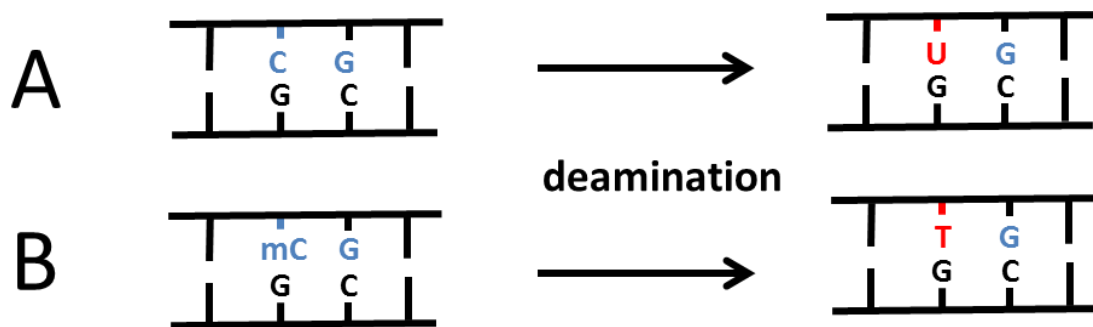


Figure 13: DNA substrates used for AFM studies with hTDG.

A) Deamination of cytosine at CpG sites causes a U:G mismatch. B) Deamination of methylated cytosine (mC) produces a T:G mismatch in a CpG context.

To capture the enzyme in its lesion recognition state, the catalytically inactive hTDG variant N140A was used, which exhibits similar DNA binding affinity as the wt enzyme but only very minor, residual base excision activity for G:U and no detectable activity for G:T mismatches (Maiti et al., *JBC*, 2009; Hardeland et al *JBC* 2000). Nucleotide flipping by TDG was investigated using mutational studies with the hTDG variant R275A and a fluorescence-based nucleotide flipping assay. Identical studies were performed with a second glycosylase, hOgg1, which recognizes structurally different 8oxoG lesions, to generalize the findings for hTDG in the context of DNA lesion search by DNA glycosylases.

3. Materials

3.1 Equipment

Table 3: Instruments.

Equipment	Supplier
Autoclave Systec V-150	Systec
Bacteria shaker ISF-1-W	Kühner
Balances XS 105 Dual Range (0-1 g) XS 6002S Dual Range (0.5-500 g)	Mettler Toledo
Bio-Photometer	Eppendorf
CD Spectropolarimeter J-810	Jasco
Cell disruptor	Constant Systems
Centrifuges Centrifuge 5415 D (0-2 mL) Centrifuge 5810 R (0-50mL) Avanti J-26 XP (0-1000 mL)	Eppendorf Eppendorf Beckmann Coulter
Chromatography system Äkta express	GE Healthcare
Chromatography system Äkta purifier	GE Healthcare
Columns for affinity chromatography Econo-Column 1,5 x 15 cm Econo-Column 2,5 x 20 cm	Biorad
Fluoromax4 fluorescence spectrometer	Horiba Jobin Yvon
Fluorescence ultra-micro cuvette (30 µL)	Hellma Analytics
Fraction collector Frac-950	GE Healthcare
Gelelectrophoresis Mini-Protean 3-cell	Biorad
Gelfiltration HiLoad 26/60 Superdex 200 prep grade	GE Healthcare
Ice machine 94774	Ziegler Eismaschinen
Magnetic stirrer MR 3002	Heidolph
Molecular force probe (MFP) 3D-Bio AFM	Asylum Research
Monitor pH/C-900 (pH, conductivity temperature) Monitor UV900 (absorption)	GE Healthcare
Nanodrop ND 1000 Spectrophotometer	Peqlab
Octet RED System	Forté Bio

PharosFX imager system	Biorad
pH electrode and -meter	Schott
Pipettes (10, 20, 200, 1000 µL)	Rainin
Pipette (2 µL)	Eppendorf
Pipetboy easypet	Eppendorf
Power supply	Biorad
Rotors JLA 25.50 JA 8.1000	Beckmann Coulter
sample loading pump P960	GE Healthcare
Thermomixer	Eppendorf

Table 4: Disposable Materials.

Disposables	Supplier
Blotting paper	Roth
Column material for affinity chromatography Ni-NTA Agarose	Invitrogen
Concentrator MWCO 30,000 Vivaspin 20	Sartorius Stedim Biotech
Filter (0.22 µm pore size)	Roth
Filter Anotop (0.02 µm pore size)	Whatmen
Greinertube 15 mL, 50 mL	Greiner
Mica Grade V	SPI Supplies
Microcon centrifugal filters for DNA concentration 10 kD MWCO 50 kD MWCO 100 kD MWCO	Millipore
Microscope slides (76 x 26 x 1 mm)	Hartenstein
OMCL-AC240TS noncontact/ tapping mode silicon probes	Olympus
Pipette tips 20 µL, 200 µL, 1000 µL	Rainin
PageRuler prestained protein ladder	Fermentas
Pipette tips 2 µL	Eppendorf
Reactiontube 0.5 mL, 1.5 mL	Eppendorf
Streptavidin tips for biolayer interferometry	Forté Bio

Table 5: Kits.

Kit	Supplier
NucleoSpin Extract II	Macherey-Nagel, Dueren, Germany
NucleoBond Plasmid Kit	Macherey-Nagel, Dueren, Germany

Table 6: Computer software.

Software	Author/ Reference
Aekta Software Unicorn 5	GE Healthcare
Adobe Photoshop Elements 8.0	Adobe Systems Incorporated
CD-Spectroscopy, Spectra Manager	Jasco
Corel Paintshop Pro X4	Corel Corporation
Fluoromax Software	Horiba Jobin Yvon
GIMP	Spencer Kimball, Peter Mattis
Igor Pro for MFP3D	Asylum Research
Image J	Wayne Rasband, NIH, USA
Image SXM	Steve Barrett
Nanodrop V3 Software	Peqlab
Octet RED Software Analysis 6.3	Forté Bio
OriginPro 8.5	Origin Lab Corporation
OriginPro 9.0	Origin Lab Corporation
Pubmed	NCBI
Pymol	DeLano 2002
Quantity One (Molecular Imager)	Biorad

3.2 Chemicals

All chemicals and solutions were obtained from Carl Roth (Karlsruhe), Sigma Aldrich (Seelze), Fluka (Neu-Ulm), Hampton Research (Laguna Hills, USA), or Applichem (Darmstadt) if not noted differently. All chemicals were of analytical grade or better.

Table 7: Media for bacterial cell culture.

Name	Composition
LB-Media	10 g/L Tryptone; 5 g/L yeast extract; 10 g/L NaCl
LB-Agar	10 g/L Tryptone; 5 g/L yeast extract; 10 g/L NaCl; 15 g/L Agar

Table 8: Media additives.

Substance	Stock solution	Final concentration
Ampicillin	100 mg/mL in ddH ₂ O	100 µg/mL
Chloramphenicol	34 mg/mL in 70 % Ethanol	34 µg/mL
Isopropyl—thiogalactoside (IPTG)	1 M	0.1 mM

Table 9: Buffers and solutions.**Cell lysis and protein purification of taXPD and variants**

Lysis buffer	Elution buffer	SEC and storage buffer
50 mM Tris pH 8	50 mM Tris pH 8	50 mM Tris pH 8
500 mM NaCl	500 mM NaCl	300 mM NaCl
5 mM Imidazole	500 mM Imidazole	1 mM DTT

Reaction buffers for taXPD

Octet reaction buffer	AFM deposition buffer	AFM incubation buffer
20 mM Tris pH 8	25 mM HEPES pH 7.5	20 mM Tris pH 8
5 mM MgCl ₂	25 mM NaAc	5 mM MgCl ₂
1 mM EDTA	10 mM MgAc	1 mM EDTA
10 mM KCl		10 mM KCl
5 mM DTT		5 mM DTT
0.1 mg/mL BSA		+/- 2 mM ATP
Helicase Assay buffer	Helicase assay stop buffer	Helicase assay running buffer
20 mM Tris pH 8	10 mM Tris pH 8	14.4 g/L Glycine
5 mM MgCl ₂	10 mM KCl	3 g/L Tris
1 mM EDTA	10 mM EDTA	1 mM MgCl ₂
10 mM KCl	0.5 % SDS	
	50 % Glycerol	

Reaction buffers for hTDG

AFM deposition and incubation buffer	2-AP base flipping assay buffer
25 mM HEPES pH 7.5	20 mM HEPES pH 7
25 mM NaAc	0.2 mM EDTA
10 mM MgAc	2.5 mM MgCl ₂
	100 mM NaCl

Additional buffers

2-AP base flipping assay buffer for hOgg1	hOgg1 storage buffer	hTDG storage buffer
50 mM Tris pH 7.5	20 mM Tris pH 8.0	20 mM HEPES pH 7
1 mM EDTA	100 mM NaCl	0.5 mM EDTA
50 mM KCl	40 % Glycerol	100 mM NaCl
1 mM DTT		1 mM DTT
		1% Glycerol

1x TAE buffer for agarose gels	Buffer for DNA ligase	SDS staining solution
40 mM Tris pH 7.6	50 mM Tris pH 7.5	50 % Methanol
20 mM acetic acid	50 mM KCl	10 % Acetic acid
1 mM EDTA	10 mM MgCl ₂	0.1 % Coomassie brilliant blue
	5 mM DTT	
	+ 1 mM ATP	
SDS destaining solution	SDS running buffer (1x)	SDS loading buffer
10 % Methanol	192 mM Glycine	50 mM Tris pH 6.8
5 % Acetic acid	0.1 % SDS	100 mM DTT
	25 mM Tris	2 % SDS
		0.1 % Bromphenol blue
15 % Resolving Gel	5 % Stacking Gel	8% native Gel
15% AA/ BAA	15% AA/ BAA	8% AA/ BAA
375 mM Tris pH 8.8	125 mM Tris pH 6.8	0.5x Trisglycine pH 8
0.1 % SDS	0.1 % SDS	0.2 % APS
0.25 % APS	0.25 % APS	0.1 % TEMED
0.05 % TEMED	0.15 % TEMED	

All buffers were adjusted to the respective pH value by adding either 32 % (w/v) NaOH or 37 % (v/v) HCl.

3.3 Biochemicals

Table 10: Enzymes.

Enzyme	Source
Accl	New England Biolabs
BglII	New England Biolabs
BspQI	New England Biolabs
Desoxyribonuclease I (DNase I)	Invitrogen
hOgg1	MyBioSource
hTDG (wt/ variants N140A and R275A)	A. Maiti (group A. Drohat, University of Maryland)
NdeI	New England Biolabs
Nt. BstNBI	New England Biolabs
SspI	New England Biolabs
T4 DNA Ligase	New England Biolabs
T4 polynucleotide kinase	New England Biolabs
XhoI	New England Biolabs
XPD from <i>T. acidophilum</i> (2 nd batch wt/ variants K170A)	G. Michels/ S. Wolski (group C. Kisker, University of Wuerzburg)

Table 11: DNA substrates for AFM and BLI experiments with taXPD.

	DNA substrate	DNA sequence
A	Bottom strand	GGT CGA CTC TAG AGG ATC AGA TCT GGT ACC TCT AGA CTC GAG GCA TGC
B	Top - / -	GCA TGC CTC GAG TCT AGA GGT ACC AGA TCT GAT CCT CTA GAG TCG ACC
C	Top - / bubble (8nt)	GCA TGC CTC GAG TCT AGA <u>CTC TTT CCA</u> TCT GAT CCT CTA GAG TCG ACC
D	Top F / bubble (8nt)	GCA TGC CTC GAG TCT AGA <u>CTC FTT CCA</u> TCT GAT CCT CTA GAG TCG ACC
E	Top F / bubble (14nt)	GCA TGC CTC GAG TCT AGT <u>CTC GTT CCF ATA TAT</u> CCT CTA GAG TCG ACC
F	Top F / 5' bubble (8nt)	GCA TGC CTC <u>GTC AAA TCT</u> GGT ACC AGA TCT GAT CCT CTA GAG TCG FCC
G	Top F / 3' bubble (8nt)	GCA TGC CFC GAG TCT AGA GGT ACC AGA TCT <u>CTA AAG TAA</u> GAG TCG ACC
H	Top F / -	GCA TGC CTC GAG TCT AGA GGT FCC AGA TCT GAT CCT CTA GAG TCG ACC
I	Top CPD / bubble (8nt)	GCA TGC CTC GAG TCT AGA <u>CTC [CPD] TT CCA</u> TCT GAT CCT CTA GAG TCG ACC
J	Top CPD / 5' bubble (8nt)	GCA TGC CTC <u>GTC AAA TCT</u> GGT ACC AGA TCT GAT CCT CTA GA [CPD] TCG ACC
K	Top CPD / 3' bubble (8nt)	GCA TGC [CPD] TC GAG TCT AGA GGT ACC AGA TCT <u>CTA AAG TAA</u> GAG TCG ACC
L	Top CPD / -	GCA TGC CTC GAG TCT AGA GGT [CPD] CC AGA TCT GAT CCT CTA GAG TCG ACC
M	Top fork (helicase assay)	<u>AGC TAC CAT GCC TGC ACG AAT TAA</u> GCA ATT CGT AAT CAT GGT CAT AGC T- Cy3
N	Bottom fork (helicase assay)	AGC TAT GAC CAT GAT TAC GAA TTG CTT <u>GGA ATC CTG ACG AAC TGT AG</u>
O	Competitor oligo (helicase assay)	GAC CAT GAT TAC GAA TTG

The DNA sequences forming unpaired DNA regions (bubble of fork) are underlined. F = fluorescein adducted thymine; [CPD] = cyclobutane pyrimidine (thymine) dimerized with 3'-thymine. Cy3 = Cyanine dye for gel-based helicase assays. For BLI measurements, substrate A was obtained with a biotin group coupled to the 3'-end and annealed with substrates B, C, D, H, I and L at an equimolar ratio. For DNA substrates used in the AFM studies only the 48 nt insert is shown (see Methods 4.1.2) and the target site is located from 28-34% of DNA fragment length (916 bp). DNA substrates M and N for helicase assays (from Biomers) were annealed at equimolar ratio yielding forked substrates for XPD. Competitor oligonucleotide O for bottom fork N prevented re-annealing of the separated DNA duplex. DNA oligonucleotides containing a fluorescein (substrates D-H) were obtained from Integrated DNA Technologies (IDT) and CPD-containing oligomers (I-L) were provided by Thomas Carell's group, LMU Munich. Undamaged oligonucleotides were purchased from Biomers.

Table 12: DNA substrates for hOgg1 and hTDG.

	DNA substrate	DNA sequence
A	G:U mismatch	GCA TGC CT(dU) GAG TCT AGA GGT ACC AGA TCT GAT CCT CTA GAG TCG ACC
B	G:U ^F mismatch	GCA TGC CT(dU ^F) GAG TCT AGA GGT ACC AGA TCT GAT CCT CTA GAG TCG ACC
C	G:T mismatch	GCA TGC CT(T) GAG TCT AGA GGT ACC AGA TCT GAT CCT CTA GAG TCG ACC
D	8oxoG:C	GCA TGC CTC (8-oxo-G)AG TCT AGA GGT ACC AGA TCT GAT CCT CTA GAG TCG ACC
E	2-AP G:U ^F	ATG CAT GCC (2-AP)(dU ^F)G AGT CTA GAG G
F	2-AP nonspecific	AT GCA TGC C (2-AP)CG AGT CTA GAG G
G	2-AP 8oxoG:C	AT G CA T GC C (2-AP)(8oxoG)G AGT CTA GAG G
H	2-AP Counter 1	CCT CTA GAC TCG TGG CAT GCA T
I	2-AP Counter 2	CCT CTA GAC TCC TGG CAT GCA T

A-D: DNA substrates for AFM experiments. E-I: DNA substrates for fluorescence assays containing 2-Aminopurine (2-AP). For DNA substrates for AFM only the 48 nt insert is shown (see Methods 4.1.2) and the target site is located at 46% of DNA fragment length (549 bp). The bases that form a specific DNA site are in brackets. Target sites for hTDG are G:U, G:U^F (2'-deoxy-2'-fluoroarabouridine, U^F), and G:T mismatches in a CpG context, the target site for hOgg1 is 8oxoG:C. Oligonucleotides E and F were annealed with oligonucleotide H at an equimolar ratio, respectively. Likewise, oligonucleotide G was annealed with Oligonucleotide I. Oligonucleotides B, D, E, G were purchased from Midland Certified Reagents and oligonucleotides H and I from Sigma Aldrich. Oligonucleotides A, F were obtained from Integrated DNA Technologies and oligonucleotides C from Biomers.

Table 13: Bacterial strains.

Name	Genotype	Source
BL21-CodonPlus(DE3)-RIL	<i>E. coli</i> B F ⁻ <i>ompT hsdS</i> (r _B ⁻ m _B ⁻) <i>dcm</i> ⁺ Tet ^r <i>gal</i> λ(DE3) <i>endA</i> Hte [<i>argU ileY leuW Cam</i> ^r]	Stratagene

Table 14: Plasmids.

Name	Vector	Resistance	Tag	Source
taXPD (T3)	pET16b	Ampicillin	N-terminal His ₁₀	J. Truglio (group C. Kisker)
taXPD mutants	pET16b	Ampicillin	N-terminal His ₁₀	S. Wolski (group C. Kisker)
pUC19N	pUC19-based	Ampicillin	/	S. Wilson, NIEHS, USA

4. Methods

4.1 Atomic Force Microscopy (AFM) imaging

Originally, AFM has been developed for visualization of atoms on solid surfaces (Binnig et al., 1986) and has been widely used and accepted in material sciences. However, high-resolution AFM imaging is now valued as highly attractive tool in life sciences as it meets several requirements necessary for studying biomolecules. In the present work, I use AFM as an especially appealing approach to analyze protein-DNA interactions in DNA repair processes.

4.1.1 Sample deposition

For studying biological samples by AFM, molecules have to be deposited onto a substrate surface. Above all, AFM substrates have to be extremely smooth at the atomic level. In addition, the substrate has to bind the particles strong enough to prevent dragging by the tip while preserving the native structure of the sample as far as possible. Adsorption is mainly mediated by physical interactions such as electrostatic, polar or Van der Waals forces, which can be adjusted e.g. by pH, ionic strength and the solution composition (Keller and Bustamante, 1993, Muller et al., 1997).

The AFM substrate most commonly used for biological samples is Mucsovite (mica). Mica is constituted by layers of $(\text{Si,Al})_2\text{O}_5$ ionically bound to a central layer of $\text{Al}_2(\text{OH})_2$ and has a typical surface roughness of 0.05 nm root mean square (RMS) (Fronczek et al., 2011). Conveniently, mica can be easily prepared (and reused) for sample deposition by cleaving a layer with simple scotch tape. However, other materials used as AFM substrates e.g. glass or gold have to be carefully cleaned before sample deposition (Hegner et al., 1993, Keller and Bustamante, 1993, Muller et al., 1997).

DNA was one of the first biological subjects used for AFM imaging. Initial difficulties in deposition and imaging of DNA molecules were overcome by the work of Bustamante and colleagues in 1992 (Bustamante et al., 1992). The deposition of DNA was optimized based on their studies.

For AFM imaging in air, DNA can be deposited onto untreated mica in a suitable buffer solution containing divalent cations (e.g. Mg^{2+} , Ni^{2+}). In air, the negative net charge of mica is compensated by potassium ions, while in an aqueous solution, K^+ ions will dissociate and the negative net charge of mica is neutralized by protons (Herder et al., 1987, Claesson et al., 1986). Thus, for deposition of DNA, mica is rinsed with ultrapure water, which depletes the solution surrounding DNA of free cations and allows stable chelation of the negatively charged DNA molecules to the mica surface by Mg^{2+} in the buffer solution (Lyubchenko and Shlyakhtenko, 2009).

For the AFM imaging experiments described here, all DNA substrates used for AFM were heated at 65° C for 10 min prior to the experiments and slowly cooled down to room temperature to remove possible salt microcrystals formed upon storage on DNA. All DNA and protein samples for AFM imaging were deposited in AFM deposition buffer (see Table 9).

In the following section, incubation conditions for taXPD will be discussed. In incubations with specific 916 bp DNA substrates, 350 nM taXPD (wt, K170A and R567W) were mixed with 15 nM DNA substrates (Table 11, substrates C-L) and incubated for 30 min at 37°C in XPD incubation buffer (see Table 9) \pm 2 mM ATP or ATP γ s. For sample deposition, protein-DNA samples were diluted 8-fold in AFM deposition buffer to a final volume of 20 μ L and immediately pipetted onto freshly cleaved mica (Grade V; SPI Supplies). Depositions were then immediately rinsed with ultra-pure deionized water to remove all buffer components (Lyubchenko and Shlyakhtenko, 2009) and carefully dried in a gentle stream of nitrogen. Likewise, taXPD wt was incubated with 9 nM 1813 bp nonspecific linear DNA fragments or 5 nM plasmid DNA pUC19N (2729 bp) for determination of DNA binding densities in the presence and in the absence of 2 mM ATP.

This section describes experimental conditions for AFM experiments with DNA glycosylases. For analysis of hTDG interactions with lesion specific 549 bp DNA substrates for AFM (Table 12, substrates A-C), 50 nM hTDG (variants N140A and R275A) were mixed with 2.5 nM DNA substrates (G:U, G:U^F, G:T mismatch in a CpG context). The samples were incubated for 20 min at room temperature and a final volume of 20 μ L was deposited onto freshly cleaved mica as described above for taXPD. 50 nM hTDG (wt and variant R275A) were mixed with 1.5 nM undamaged DNA fragments (1813 bp) and 50 nM hTDG N140A were mixed with 2.5 nM undamaged DNA fragments (549 bp) in AFM deposition buffer to analyze nonspecific hTDG-DNA complexes. The samples were incubated and deposited as in incubations with specific DNA substrates. hOgg1 wt (150 nM) was incubated with undamaged DNA fragments (2 nM for 1813 bp DNA and 5 nM for 549 bp DNA) in AFM deposition buffer for 10 min at room temperature. The solution was deposited for 1 min onto freshly cleaved mica and subsequently rinsed and dried as described above for the other protein systems.

To determine intrinsic DNA bend angles from reference images, DNA substrates were imaged in the absence of protein at a deposition concentration of 2.5 nM for 549 bp DNA fragments, 2.0 nM for 916 bp DNA fragments and 1.5 nM for 1813 bp DNA fragments.

Control images for protein in the absence of DNA were collected at a deposition concentration of 25 nM for taXPD, hTDG and hOgg1.

4.1.2 AFM image collection and processing

All AFM images were collected in air with a molecular force probe (MFP) 3D-Bio AFM (Asylum Research, Santa Barbara, CA) in tapping mode using OMCL-AC240TS (Olympus) noncontact/ tapping mode silicon probes with spring constants of ~2 N/m and resonance frequencies of ~75 kHz. AFM micrographs were captured at a scan speed of 2.5 $\mu\text{m/s}$ with image sizes of 1 $\mu\text{m} \times 1 \mu\text{m}$, 2 $\mu\text{m} \times 2 \mu\text{m}$, 4 $\mu\text{m} \times 4 \mu\text{m}$ or 8 $\mu\text{m} \times 8 \mu\text{m}$ and pixel resolution of ~2 nm. Utilizing Asylum Research software on Igor Pro, AFM images were processed by 3rd order plane-fitting and flattening.

4.1.3 AFM data analysis

This section gives an overview of different analytical approaches to extract quantitative information from AFM images of dried protein-DNA samples. Statistical analysis of protein-DNA complexes can offer valuable clues to protein binding affinities on DNA from the number of bound protein molecules. Moreover, determination of protein binding positions on DNA substrates containing a single DNA lesion at a defined position enables the investigation of molecular DNA damage recognition and can reveal information about binding specificities of a protein for its target site in DNA. Finally, conformational properties of protein-DNA complexes and information on oligomeric states of proteins bound and unbound to DNA can be gained from analysis of the bend angles induced into DNA at the site of the bound protein and AFM volumes.

Initially, protein-DNA complexes were manually identified from the AFM images neglecting protein-complexes on DNA aggregates and on DNA fragments that were cut off by the image margins. To quantify DNA coverage with protein, all DNA fragments (bound with protein and unbound) and all protein-DNA complexes were counted. DNA binding densities per base pair (bp) were calculated as:

$$\text{binding density} = \frac{\# \text{ protein peaks}}{(\# \text{ DNA fragments} * \# \text{ basepairs})} \text{ (Equation 1)}$$

AFM experiments for determination of binding densities (Figure 23B) were performed at least in triplicates for each substrate and condition (\pm ATP or ATP γ s), with n = 340 – 876 DNA-bound protein peaks for each substrate type.

Substrate binding specificities S of a protein for its target site were quantified from Gaussian fits to position distributions of protein complexes on DNA as recently described (Tessmer et al., 2008, Buechner et al., 2014). In short, complex positions were determined as the distance of protein peaks to the closer DNA fragment end normed to the full length of the DNA substrate by placing a contour along the DNA backbone using the NIH software Image J. Only DNA fragments with lengths within two standard

deviations from the center of a Gaussian fit to the DNA length distributions were included in the analyses. Positions were plotted as histograms using the software Origin Pro 8.5.

For quantitative analysis of binding specificities, DNA end-bound protein complexes in the histograms were excluded, starting at 5 % of DNA length. A fixed bin size of 4 % was applied for all individual experiments. Lesion specificity for a target site results in a Gaussian shaped peak at the position of the specific site, which can be fitted with a single Gaussian curve (Figure 14).

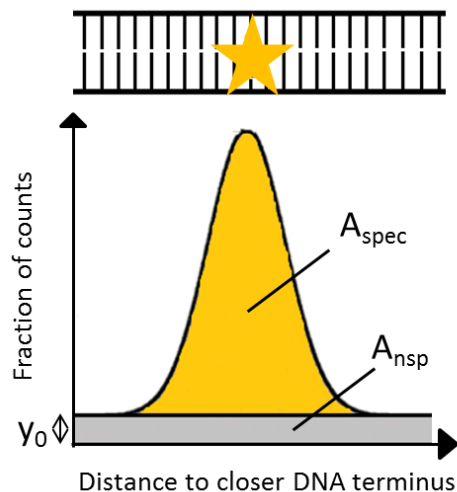


Figure 14: Illustration of the position distribution of protein binding to a single target site in DNA.

The DNA fragment contains a single specific target site (orange star) in the middle. The areas under the position distribution of protein-DNA complexes, A_{nsp} (grey) and A_{spec} (orange), and the height of the background in the Gaussian fit (y_0) are indicated. The parameters derived from the Gaussian distribution can be used for calculation of the lesion specificity of a protein for a specific target site in DNA (Equation 6) (adapted from Yang et al. (2005)).

Integration of the Gaussian peak area (A_{spec}) provides the fraction of specific complexes over the fraction of nonspecific complexes, which is represented by the area of the nonspecific background (A_{nsp}) and is calculated as the product of considered DNA length (5 - 50 % of DNA fragment length, excluding DNA ends) and the height of the background in the Gaussian fit (y_0) (Yang et al., 2005). Lesion specificity S is defined as:

$$S = N * \frac{A_{spec}}{A_{nsp}} + 1 \text{ (Equation 2)}$$

With N being the number of available binding sites on DNA ($N = 824$ for XPD and $N = 494$ for hTDG excluding $2 \times 5\%$ of DNA fragment length). Gaussian widths were all fixed to the width of the Gaussian fit to the DNA substrate with highest specificity (2.6 % for taXPD and fluorescein/bubble with ATP and 3.2 % for hTDG R275A and G:U^F) to facilitate direct comparison of binding specificities for different DNA substrates and conditions. Specificities were calculated as average values from at least three individual experiments with standard deviations (SD) derived from the variations between single experiments. All

position distributions shown in this work represent pooled data from at least three individual depositions.

Based on the protein position distributions, complexes bound at positions within two standard deviations (based on Gaussian widths) from the center of the fit were classified into specific complexes and nonspecific complexes bound elsewhere on DNA. For further analysis (DNA bend angle and AFM volume analysis) specific and nonspecific protein-DNA complexes were treated separately.

DNA bend angles β were defined as the deviation from a straight line through the DNA backbone:

$$\beta = 180^\circ - \alpha \text{ (Equation 3)}$$

α was determined by placing two lines along the DNA backbone on each side of a protein complex by manual tangent overlay using Image J. All DNA bend angle distributions in the presence of protein shown in this work represent pooled data from at least three individual AFM experiments. Bend angle values for protein-DNA complexes were derived as the maxima of single, double or triple Gaussian fits to pooled bend angle distributions from at least three individual experiments with errors estimated from the half-width of the Gaussian peak (σ). To determine percentage areas of different bend angle populations, we fixed widths of the Gaussian fits in such a way yielding highest possible R^2 values. The resulting peak areas were integrated using Origin Pro.

To determine intrinsic DNA bending of specific DNA substrates at the site of the damage, a mask comparable to the size of the protein was moved to the position of the target site (to 30% of DNA fragment length for taXPD and to 46 % for hTDG) DNA bending was measured at this position as described above for protein induced DNA bend angles. To exclude any influence of intrinsic DNA bending behavior due to sequence context at the position itself, intrinsic bending was determined for 549 bp undamaged DNA substrates at this position (46 % of fragment length).

Furthermore, intrinsic DNA bending of nonspecific (undamaged) DNA fragments was measured by moving the mask at regular intervals of 50 nm along DNA. All individual intrinsic DNA bend angle distributions for the different DNA substrates in the absence of protein represent pooled data from two individual AFM experiments.

To gain further information on conformational and oligomeric states of protein molecules, we determined AFM volumes of free (non-DNA-bound) proteins and protein-DNA complexes from the AFM images using the software ImageSXM. The AFM volumes V of proteins can be translated into molecular

weight (MW) (Ratcliff and Erie, 2001) using the empirically derived calibration for the employed AFM set-up (Roth et al., 2012):

$$MW = (V + 5.9)/1.2 \text{ (Equation 4)}$$

These different approaches described above, point out a number of merits in AFM imaging experiments on analysis of protein-DNA interactions.

4.2 Molecular biology

4.2.1 Transformation and plasmid isolation

For transformation, 100 μ L of chemocompetent *Escherichia coli* (*E.coli*) BL21-CodonPlus (DE3)-RIL cells (carrying the resistance for chloramphenicol) (Stratagene) were incubated with 1 μ L of plasmid DNA (~100 ng/ μ L) (pET16b or pUC19N, both carrying the resistance for ampicillin) for 30 min on ice. The cells were subsequently heat shocked at 42°C for 1.5 min and cooled down on ice for 2 min. Prior to incubation, 1 mL sterile Luria Broth (LB) medium were added to the cell suspension. Afterwards, the cells were shaken and incubated at 37°C for 1h (600 rpm; Eppendorf Thermomixer), spread on a LB-agar plate with the appropriate antibiotics (ampicillin (Amp) and chloramphenicol (Cam) at concentrations of 100 μ g/mL and 34 μ g/mL, respectively) and grown overnight at 37°C. For protein expression, a single colony was picked from the agar plate to inoculate a pre-culture.

To amplify plasmid DNA (pUC19N) for DNA substrate preparation for AFM (see 4.1.2), 200 mL LB^{Amp Cam} medium were inoculated with a single colony from the LB-agar plate of a fresh transformation. After shaking and incubation overnight at 37°C and 200 rpm, the bacterial cells were pelleted by centrifugation at 4000 rpm for 10 min at 4°C (centrifuge 5810R, Eppendorf). Cell lysis and DNA isolation was performed using the NucleoBond Plasmid Kit (Macherey Nagel) according to the manufacturer's instructions.

4.2.2 DNA substrate preparation for AFM

DNA repair proteins possess a high binding preference for specific types of DNA lesions. To study molecular recognition of their target sites by AFM imaging experiments, a single DNA damage can be introduced into long DNA substrates at a known position. The next section describes the preparation procedure for production of such DNA substrates for AFM, which has been established and optimized in our laboratory (Buechner and Tessmer, 2013).

The DNA substrate preparation is based on the modified pUC19 based DNA plasmid pUC19N (1, Figure 15) (Hou et al., 2007). pUC19N (2729 base pairs, a kind gift from Samuel Wilson's laboratory NIEHS, USA) contains 6 restriction sites for the nicking endonuclease *Nt.BstNBI* (GAGTC). Three of these restriction

sites are in close vicinity (positions 422, 440 and 470) and serve to produce a ssDNA gap of 48 nt, which can then be replaced by an specific oligonucleotide of choice (Buechner and Tessmer, 2013).

After incubation with Nt.BstNBI for 2h at 55°C (2, nicking, Figure 15), a short ssDNA fragment (48 nt) between the two incision sites was removed by repeated heating at 68°C in the presence of an at least 10-fold excess of complementary oligonucleotide (substrate A, Table 11) (3, gapping, Figure 15). The annealed 48 bp oligomers and the excess of 48 nt ssDNA counter oligonucleotides were filtered out by centrifugation through a molecular weight cut-off filter (125 nt cut-off; Amicon Ultra, Millipore). The resulting ssDNA gap was replaced by incubation with a 20-fold excess of a specific 5'-phosphorylated target site containing DNA insert (DNA sequences listed in Table 11 and Table 12) for 4h at 45 °C (4, annealing, Figure 15). The remaining nicks between the insert DNA and the original strand were ligated by overnight incubation with T4 DNA Ligase I at room temperature (5, nick sealing, Figure 15).

Damage containing plasmids with target sites for taXPD were digested with the restriction enzymes SspI and BspQI yielding linear specific 916 bp DNA fragments and linear nonspecific 1813 bp DNA fragments, which were used as undamaged control substrates for damage search studies (6, linearization, Figure 15). Depending on the position of the bubble (5' or 3' to the damage), target sites for taXPD (fluorescein/CPD) were located at 28-34 % of DNA fragment length. Incubation conditions for the digestion reactions were based on the enclosed product information provided with the restriction enzymes. To ensure complete digestion of DNA, the reactions were performed using twice as high incubation concentrations of the enzymes and reaction times as recommended in the manufacturer's instructions.

To produce target site containing DNA substrates for glycosylases, lesion specific pUC19N was cut by NdeI and BspQI yielding 549 bp specific DNA fragments (containing target sites for DNA glycosylases) and 2180 bp nonspecific fragments similar as described in the previous section. The target sites for hTDG (G:T, G:U or G:U^F mismatch) or hOgg1 (8oxoG:C) were located at 46 % of DNA fragment length. Nonspecific DNA fragments of 549 bp length were generated by digestion of undamaged pUC19N with NdeI and BspQI.

After restriction enzyme digestion, all DNA substrates (specific and nonspecific) used for AFM imaging were purified by extraction of the corresponding band from agarose gel electrophoresis following the instructions of the NucleoSpin Extract II Kits (Macherey-Nagel, Dueren, Germany) (6, gel purification, Figure 15). The UV-irradiation used for visualization of the stained DNA bands in the agarose gels can induce additional UV lesions in DNA, which can interfere with the correct interpretation of the AFM experiments on protein-DNA interactions (Buechner and Tessmer, 2013). To avoid UV light exposure of

the DNA substrates, agarose gels were cut between lane 1 and the remaining lanes and only lane 1 was exposed to UV-irradiation. Upon re-combining the two gel parts, the correct excised DNA band then functions as a marker for the position of the specific DNA substrate in the rest of the gel.

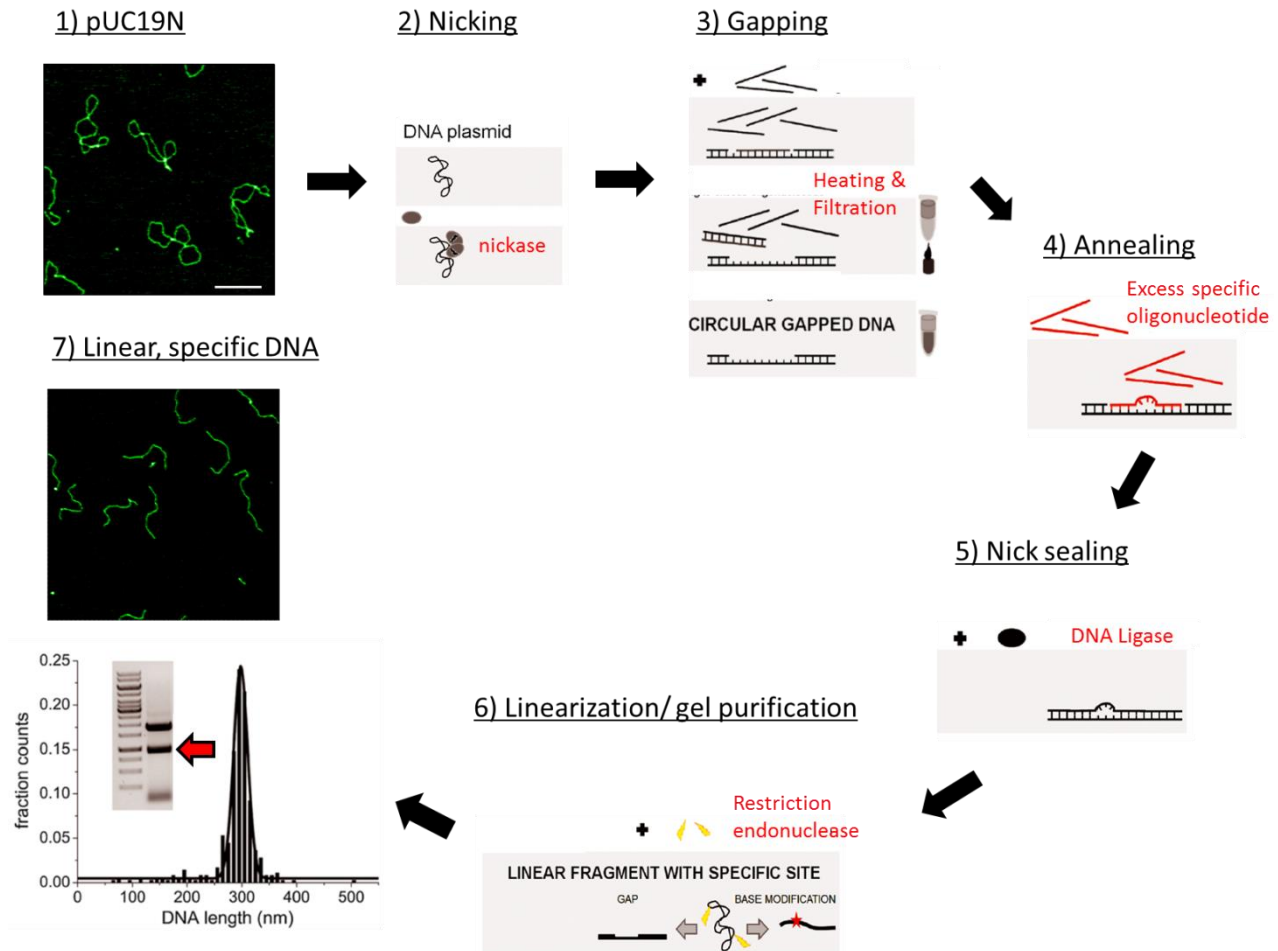


Figure 15: Schematics of DNA substrate preparation for AFM imaging experiments.

1) AFM image of the nonspecific undamaged plasmid pUC19N (2729 bp). The scalebar corresponds to 200 nm. 2) pUC19N, which contains two nicking sites in close vicinity, was incubated with a nickase to generate ssDNA cuts. 3) Gapped DNA was produced by repeated heating and filtration steps in the presence of an excess of a complementary oligonucleotide. 4) An excess of a target site containing oligonucleotide (red) was inserted into the ssDNA gap of the plasmid. 5) The remaining nicks between the insert and the original strand of the vector were sealed by a DNA ligase. 6) Plasmids were cut into linear fragments containing the specific target site at a defined distance from the closer DNA fragment end by digestion with appropriate restriction enzymes and purified by agarose gel extraction. 7) AFM image of linear DNA fragments (916 bp) after gel purification containing a single fluorescein modification within the context of a DNA bubble at 30% of DNA fragment length. DNA length measurements of the purified DNA fragments (n=358) resulted in a distribution with a distinct Gaussian peak with maximum at 298 nm (corresponding to 876 bp DNA). The small deviation from the theoretical DNA length (311 nm with 0.34 nm/bp) of 4% is typical for DNA length measurements from AFM images (Campbell et al., 2008). The insert shows the digestion products of pUC19N upon incubation with the restriction enzymes SspI and BspQI. The red arrow points to the specific target site containing DNA fragment (916 bp). Both AFM images in 1) and 7) are 1 μm x 1 μm. For details see main text (adapted from Buechner and Tessmer (2013)).

Linear, lesion-specific DNA fragments suitable for AFM studies (7, Figure 15) are the end product of the DNA substrate preparation. A successful DNA substrate preparation typically yielded homogenous and pure linear DNA fragments containing a target at a defined position. In the specific example in Figure 15, the DNA fragments were 916 bp long and contained a fluorescein adduct in the context of an 8 nt bubble at 30 % of the DNA fragment length.

The concentration (c_{DNA}) of the purified DNA substrates could be determined from the maximum absorption at 260 nm (A_{260}) using a nanodrop UV/Vis spectrophotometer (PqLab). The DNA concentration is given by Lambert-Beer's law (Pfeiffer and Liebhafsky, 1951):

$$c_{DNA} = \frac{A_{260}}{(\epsilon_{DNA}) * d} \text{ (Equation 5)}$$

with d being the thickness of the cuvette ($d = 1 \text{ cm}$) and ϵ_{DNA} being the extinction coefficient for DNA determined by the nearest neighbor method (Cantor et al., 1970, Fasman, 1975). In the nearest neighbor model, interactions between adjacent bases are considered and measured empirically. Using these nearest neighbor parameters, accurate values for ϵ_{DNA} can be calculated (Table 15) (Warshaw and Tinoco, 1966):

Table 15: Nearest neighbor values for ϵ_{260} of dNTPs.

5' → 3'	dA	dC	dG	dT
dA	27,400	21,200	25,000	22,800
dC	21,200	14,600	18,000	15,200
dG	25,200	17,600	21,600	20,000
dT	23,400	16,200	19,000	16,800

Values are provided in (L/mol)*cm (Warshaw and Tinoco, 1966).

Then, the following equation is applied to calculate the molar extinction coefficient for DNA (Cantor et al., 1970, Cavaluzzi and Borer, 2004):

$$\epsilon_{260} = \sum_1^{N-1} \epsilon_{nearest\ neighbor} - \sum_2^{N-1} \epsilon_{individual\ bases} + \sum_1^N \epsilon_{modifications} \text{ (Equation 6)}$$

Each step of the sample preparation (nicking, gapping and annealing) was controlled by electrophoretic mobility shift assays using agarose gel electrophoresis as previously reported (Buechner and Tessmer, 2013, Buechner et al., 2014). The introduction of nicking sites into pUC19N during the first preparation step is tested by a nicking assay (Figure 16A). A successful nicking procedure leads to relaxation of superstructures in the plasmid DNA. The nicked plasmid possesses a different electrophoretic mobility in agarose gels (Figure 16A, lane 1) compared to untreated, supercoiled pUC19N (Figure 16A, lane 2).

Complete removal of the ssDNA fragment (48 nt) between the nicks in close vicinity (gapping assay, e.g. Figure 16B) and correct insertion of the target site specific oligomer into the ssDNA gap can be tested by incubation with the restriction enzymes XhoI (for XPD substrates A plus C, D, E, H, I, and L, Table 11), AccI (for XPD substrate G, Table 11) and BglII (for XPD substrates F, J, and K, Table 11 and for hTDG and hOgg1 substrates A-D, Table 12). Gapped DNA lacks intact restriction sites and could not be digested by the restriction enzymes (in Figure 16B), whereas successfully replaced DNA substrates could be incised resulting in higher electrophoretic mobility in agarose gels.

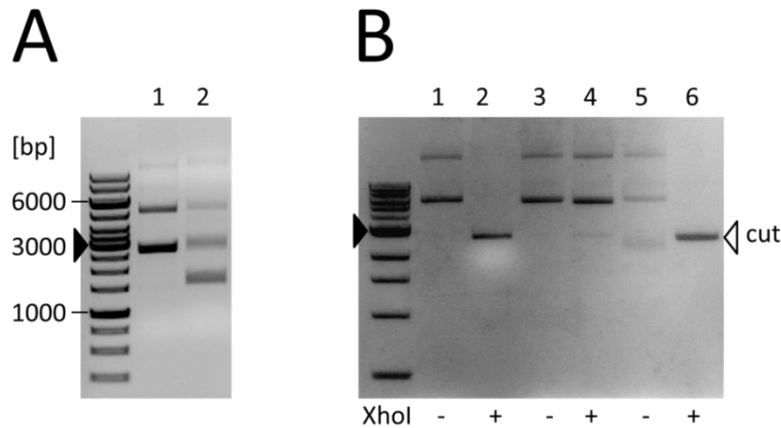


Figure 16: Assaying DNA substrate preparation.

Agarose gel-based electrophoretic mobility shift assays to control successful A) nicking and B) replacement of the ssDNA gap with a target site specific oligonucleotide insert. DNA bands were stained with Midori Green. A) In the nickase assay, a shift in the electrophoretic mobility is observed between the untreated, supercoiled plasmid (lane 1) and the incised (nicked), relaxed plasmid (lane 2). B) Correct removal of the ssDNA fragment between the nicks and replacement with a specific insert can be tested via restriction enzyme digestion e.g. by XhoI. Nicked DNA (lanes 1 and 2) and plasmids replaced with specific oligonucleotides (lanes 5 and 6) still contain the restriction site in the gapped DNA area and were incised upon addition of the enzyme (lanes 2 and 6). The white arrowhead points to the running position of the linearized plasmid, which indicates successful digestion by XhoI. Gapped DNA (lanes 3 and 4) could not be digested by the enzyme due to missing restriction sites. The bold arrowheads in A) and B) indicates the same DNA marker position (3000 bp) (adapted from Buechner and Tessmer (2013)).

Using this simple DNA substrate preparation described above, a single specific target site can be introduced into long DNA substrates (500-3000 bp) at a known position. The homogeneous and pure DNA substrates are suitable for AFM as well as other single molecule techniques such as optical or magnetic tweezers and may represent better mimics of the naturally occurring DNA substrates than the short oligonucleotides typically used in biochemical ensemble methods (Buechner and Tessmer, 2013, Buechner et al., 2014).

4.3 Expression and purification of taXPD

4.3.1 Protein expression and cell lysis

Expression and purification of taXPD proteins is based on an established protocol (Wolski et al., 2008, Kuper et al., 2012). Briefly, bacterial cells from *Escherichia coli* (E.coli) BL21-CodonPlus (DE3)-RIL cells were transformed with pET16b vectors carrying the genes encoding XPD wt or the variant R567W from *T. acidophilum* (the constructs were kindly provided by Caroline Kisker's laboratory) as described above (see 4.1.1). A single colony was picked from the LB-Agar plate to inoculate 200 ml LB^{Amp/Cam} medium. The pre-culture was incubated overnight at 37 °C and 200 rpm and 20 mL cells were transferred into 2 L LB^{Amp/Cam}. For expression, 2 L bacterial cell cultures were grown at 37 °C and 200 rpm to an optical density (OD) of 0.5 to 0.6. To obtain sufficient protein amount, a total volume of 8 – 16 L cell culture was applied for large scale expressions of taXPD and variants. Prior to induction (OD ~ 0.3 to 0.4), the temperature was reduced to 14°C to prevent formation of inclusion bodies during expression. 0.1 mM isopropyl--thiogalactoside (IPTG) was added to induce protein expression. After incubation for 18 h at 14 °C and 200 rpm, bacterial cells were harvested by centrifugation (Rotor JLA 8.1000, 4 °C, 5,000 xg, 20 min; Beckmann Coulter Avanti Centrifuge J-26 XP).

The cell pellet was re-suspended in lysis buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 5 mM imidazole). After addition of DNase I in a final concentration of 2 units per 50 mL suspension, the cell suspension was lysed using a cell disrupter system at a pressure setting of 1.8 kbar. For separation of soluble proteins from insoluble proteins and cell debris, the cell lysate was centrifuged for 1 h at 4 °C and 30,000 xg (Rotor JA 25.50, Beckmann Coulter Avanti Centrifuge J-26 XP). The supernatant contained soluble overexpressed taXPD, which was purified by two-step column chromatography: first by metal affinity chromatography followed by size exclusion chromatography.

4.3.2 Metal affinity chromatography

In the first purification step, the cell lysate was added to the matrix of a Ni²⁺-column consisting of Ni-NTA agarose after equilibration of the column with 25 mL Lysis buffer (5 column volumes). The overexpressed protein in the supernatant contains an N-terminal polyhistidine-tag (His 10-tag), which binds with high affinity to Ni²⁺-ions of the column material. Unbound protein was removed from the column in a washing step with 50 mL lysis buffer. taXPD was then eluted with 500 mM imidazole, which is chemically and structurally similar to histidine residues and strongly competes with protein molecules bound to the column. After affinity chromatography, representative fractions (lysate, wash and elution fractions) were analyzed on a 15 % SDS gel (Laemmli, 1970). Subsequently, all elution fractions containing taXPD were

pooled and concentrated to a final volume of 5 mL with Vivaspin 20 concentrators (MWCO 30,000 g/mol).

4.3.3 Size exclusion chromatography

The concentrated protein solution was further purified by size exclusion chromatography (SEC), which separates proteins according to their molecular weight and shape. In SEC, differently formed proteins show different retention times on the column material. Small, roundish particles are able to enter the pores of the column resulting in longer diffusion ways and longer elution times, whereas larger particles being excluded from the pores elute at earlier time points.

Specifically, I used a Superdex 200XK 26/60 column on an ÄKTA-Purifier-Chromatography-System for purification of taXPD proteins. The Superdex column (~320 mL column volume) is able to separate proteins between 10 kDa and 600 kDa and its column material consists of dextran covalently attached to agarose. After equilibration of the Superdex with SEC buffer (Table 9), the protein solution was injected via a sample loop or the sample loading pump P960. During the flow (1 mL/min for 1.2 CV), the pH-value, the conductivity, and the temperature were detected by a monitor pH/C-900. Via a UV900 monitor, three wavelengths could be detected simultaneously in the size exclusion chromatogram: 260 nm (absorption of nucleic acids), 280 nm (protein absorption) and 410 nm (absorption of the iron sulfur cluster of XPD). The elution fractions were automatically collected by the fraction collector Frac-950 and were checked on a 15 % SDS-Gel as described further down.

4.4 Protein characterization

4.4.1 SDS-PAGE analysis

To identify the elution fractions containing taXPD and to test for the purity of the protein, I performed sodium dodecyl sulfate (SDS) polyacrylamid gel electrophoresis (SDS-PAGE). In SDS-PAGE, proteins can be separated according to their electrophoretic mobility in an electric field under denaturing conditions. For gel analysis, 10 µL of the protein fraction were mixed with 3 µL 5x protein sample buffer containing SDS. The anionic detergent SDS is applied to normalize the intrinsic charge of proteins as it binds with high affinity to proteins. Thus, the charge of SDS-protein complexes is proportional to the molecular weight of the protein. Furthermore, the protein sample buffer contained a reducing agent (dithiothreitol, DTT) to destroy intra- and intermolecular disulfide bonds. Before loading onto the gel, samples were heated for 5 min at 95 °C to linearize the proteins by disruption of secondary and tertiary structures. As a ruler for molecular weight, the PageRuler prestained protein ladder (bands between 10 and 170 kDa) was applied in the first lane of the gel. The SDS gel was run for 45 min at 200 V and afterwards stained

with Coomassie staining solution for 20 min at room temperature. Finally, the unbound dye could be washed out with destaining solution (Laemmli, 1970).

4.4.2 Determination of protein concentration by UV-spectrophotometry

Protein absorption was measured with a nanodrop UV/Vis spectrophotometer using the spectrum mode from 175 nm to 875 nm. The protein concentration (c_{protein}) could be calculated from the maximum absorption ($A_{280 \text{ nm}}$) by Lambert-Beer's law (Pfeiffer and Liebhafsky, 1951) as described above for DNA absorption (see 4.1.2):

$$c_{\text{protein}} = \frac{A_{280 \text{ nm}}}{\epsilon_{\text{protein}} * d} \text{ (Equation 7)}$$

with ϵ being the extinction coefficient of the protein at 280 nm and $d = 1 \text{ cm}$ the thickness of the cuvette. The molar extinction coefficient for taXPD (molecular weight of 72 kDa including his-Tag) at 280 nm considers the number of fluorescent residues n in the monomeric protein (2 tryptophanes, 36 tyrosines and 8 cysteines,) and is expressed by the equation:

$$\epsilon_{280} = (n * \text{Trp} * 5690 + n * \text{Tyr} * 1280 + n * \text{Cys} * 120) \text{ (Equation 8)}$$

For XPD from *T. acidophilum* the extinction coefficient is $65,140 \text{ M}^{-1}\text{cm}^{-1}$.

4.4.3 Circular dichroism (CD)

Chiral molecules in solution such as protein and DNA are able to absorb circular polarized light. The difference in absorbance of right- and left-circularly polarized light is measured by circular dichroism (CD) spectroscopy and can be quantified as ellipticity (in degrees). Protein and nucleic acids contain secondary structure elements that lead to characteristic CD-spectra. Thus, the technique enables us to gain structural information e.g. about the content of α -helices, β -strands and unfolded regions in a protein (Kelly et al., 2005) as well as conformational properties in DNA including B-form, A-form, Z-form, guanine quadruplexes, triplexes etc. (Kypr et al., 2009).

Here, CD spectroscopy was applied to analyze and compare folding of taXPD wt and the variant R567W using a J-810 CD spectropolarimeter. CD spectra were recorded from 260 nm to 190 nm at 20°C in 0.1 cm quartz cuvettes. The proteins were diluted in 200 μL in 50 mM potassium-phosphate buffer (Sørensen buffer) pH 8.0 to a final concentration of 1.25 μM . For optimization of the signal to noise ratio, spectra were accumulated 10 times at a scan speed of 50 nm/min and at a response time of 1 s.

4.5 Analysis of protein-DNA interactions

4.5.1 Gel-based helicase assay

DNA unwinding activity of taXPD wt was investigated by a gel shift assay using forked Cy3 labeled dsDNA substrates (for DNA sequence see Table 11, substrates M and N) (Figure 22). taXPD (0.1, 0.25, 0.5, 0.75, 1 μ M) was mixed with 50 nM DNA substrate and 1 μ M competitor oligonucleotides (Table 11, substrate O) in helicase reaction buffer (see Table 9) in a reaction volume of 50 μ L. Upon addition of 2 mM ATP, the helicase reaction was started and samples were incubated for 20 min at 37 °C. After incubation, 50 μ L stop buffer containing EDTA and SDS (see Table 9) were added to stop the reaction by denaturation of the enzyme, and samples were put on ice. As controls, the DNA substrate without protein (double stranded control), a reaction without ATP (negative control) and boiled DNA substrate (10 min at 95 °C, single stranded control) were pipetted and incubated as described above. Finally, 15 μ L of each sample were loaded on a 8 % native polyacrylamide gel and then the gel was run at 100 V for at least 1 h. Fluorescently labeled DNA substrates were directly visualized by the PharosFX™ imager system.

4.5.2 Bio-Layer Interferometry (BLI) DNA binding assay

DNA binding affinities of taXPD and hTDG were determined by the label-free biosensor technology Bio-Layer Interferometry (BLI) using an Octet RED system (Fortebio, Menlo Park, CA). The Octet instrument emits white light down the sensor and detects the light reflected back from the sensor surface, which is called interference of light (Figure 17). The interference pattern is recorded by the Octet system as a unique spectral signature and is indicated as a relative intensity unit in nm. As the interference depends on the thickness of the biological layer coated on the sensor surface, it can serve as a direct measure for any change in the thickness of the molecular layer induced by binding.

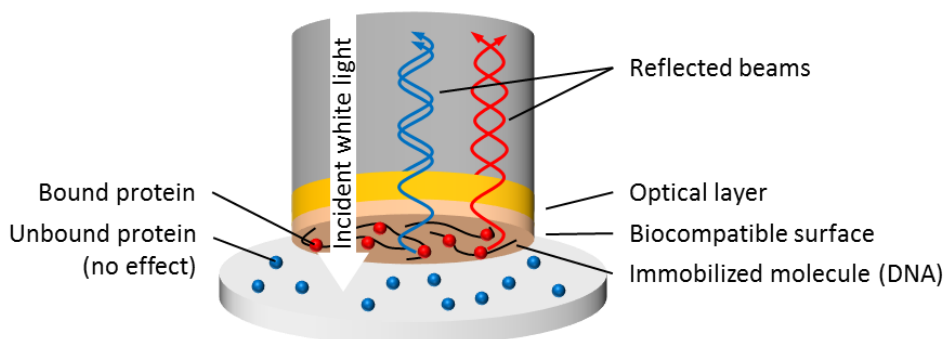


Figure 17: Principle of Biolayer Interferometry.

BLI measures the interference of incident white light emitted from the instrument and reflected from the sensor surface (reflected beams, in blue and red). The optical layer consists of molecules (e.g. biotinylated DNA, black lines) immobilized to the streptavidin coated sensor surface and proteins bound to DNA (red spheres). Unbound proteins (blue spheres) have no effect on the detected shift in the interference pattern induced by protein-DNA interactions (adapted from 'Octet RED' brochure, ForteBio, Menlo Park, CA).

For BLI experiments, biotinylated DNA substrates (ligand) were immobilized on streptavidin coated fiber optic sensors (capture molecules) using a ‘Dip and Read™’ workflow (Figure 18). In the association step, DNA coated sensors were dipped in the protein solution to detect DNA binding activity of the protein (analyte). Upon immersion into ligand-free buffer solution, the protein molecules dissociate from the sensor. Association rate constants (k_{obs}) and dissociation (k_{diss}) were derived from fits to association and dissociation curves using a 1:1 binding model from the Octet Data Analysis software. K_D values were determined as the ratio of dissociation and association rate constants for the applied protein concentration c :

$$K_D = \frac{k_{diss}}{k_{ass}} = \frac{k_{diss}}{k_{obs}*(c^{-1})} \text{ (Equation 9)}$$

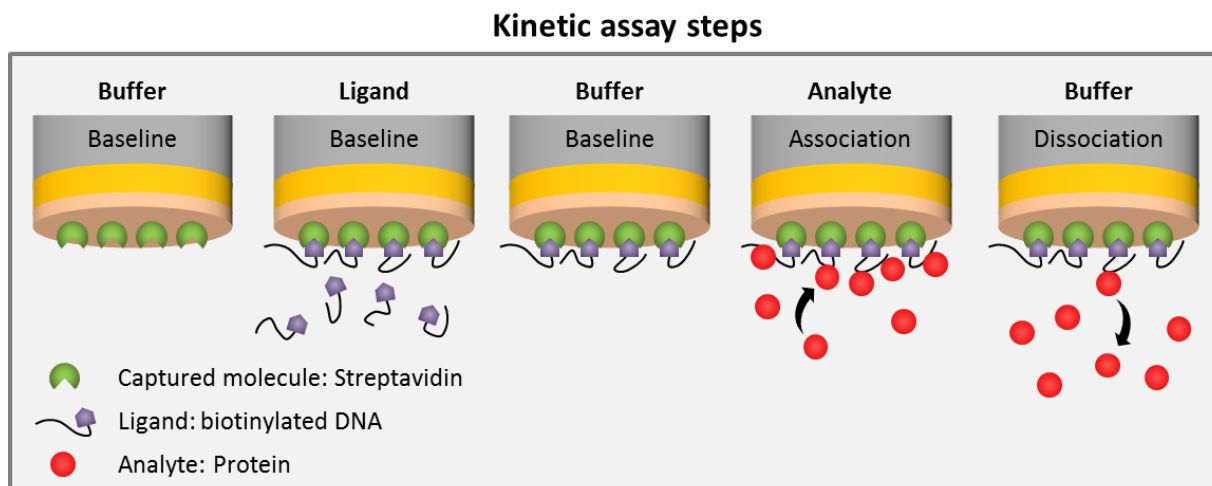


Figure 18: BLI binding assay.

The BLI binding assay consists of five distinct steps using a ‘Dip and Read™’ workflow. Prior to immobilization of the ligand, the baseline of the buffer is recorded. In the loading step, the ligand (biotinylated DNA) is bound to the capture molecules (streptavidin sensors). Subsequently, a second baseline is recorded after loading of the ligand. In the association step, binding of the analyte (protein) to the immobilized ligand molecules is detected. Finally, dissociation of the analyte from the ligand is determined by dipping of the sensor into ligand-free buffer solution (adapted from Octet RED brochure, Menlo Park, CA).

For better comparison, the DNA substrates for Octet were of the same DNA sequence as the 48 nt inserts used for AFM (Table 11). The bottom strand (Table 11, A) was purchased with a biotin group attached to the 3’ end and mixed at a ratio of 1:1 with the corresponding top strands (Table 11, substrates B, C, D, H, I and L) for annealing.

To investigate XPD-DNA interactions, 50 nM DNA and 100 – 300 nM XPD were applied for the DNA substrates containing ssDNA-regions, and 100 nM DNA and 400 – 500 nM XPD for fully dsDNA substrates. The binding reactions were performed at room temperature in XPD octet buffer (Table 9). BLI experiments for XPD were carried out at least in triplicates and with three different protein batches.

4.5.3 2-Aminopurine fluorescence based flipping assay

Nucleotide flipping activity of hTDG and hOgg1 was monitored using a 2-Aminopurine (2-AP) fluorescence based base flipping assay (Figure 19). 2-AP is a fluorescent base analog of adenine and can be incorporated into oligonucleotides as a sensitive probe for local conformational changes in the DNA structure (Millar, 1996). Due to its capability to form a Watson-Crick base pair with thymine, 2-AP maintains the overall topology of the DNA double helix (Sowers et al., 1986). The fluorescence intensity of 2-AP is highly sensitive to the surrounding environment (Ward et al., 1969) and is significantly quenched upon hybridization with a complementary DNA strand (Raney et al., 1994) (Figure 19, green star) and by stacking interactions with neighboring bases (Guest et al., 1991, Xu et al., 1994). Moreover, the fluorescence excitation and emission maxima of this nucleoside analog can be separated from those of tryptophan and tyrosine and 'normal' DNA bases. Thus, 2-AP is an ideal reporter for enzyme-induced conformational changes in the DNA duplex and has been used to study several different protein-DNA systems (Holz et al., 1998, Nakano et al., 2005, Bellamy et al., 2007). Upon destacking of the DNA duplex structure by protein-induced destabilization such as base flipping, the 2-AP fluorescence is expected to increase (Figure 19, red star).

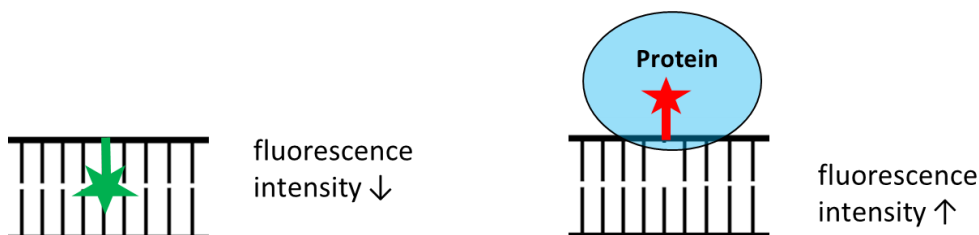


Figure 19: 2-AP fluorescence base flipping assay.

For details see main text.

2-AP was incorporated in short DNA substrates (22 bp) either within homoduplex DNA (, F/H) to follow flipping of nonspecific bases, or neighboring hTDG or hOGG1 target sites (G:U^F in a CpG context or 8oxoG:C,

E/H and G/I, respectively) to detect substrate specific base flipping. Due to the weak fluorescence signal of 2-AP, a DNA concentration of at least 170 nM was needed to reach sufficient fluorescence intensities as determined in concentration series prior to the experiments. To separate 2-AP fluorescence from intrinsic protein (tryptophan and tyrosine) and DNA fluorescence, an excitation wavelength of $\lambda_{ex} = 320$ nm was chosen for the fluorescence assays. Steady-state fluorescence emission spectra (340-400 nm) were acquired with a Fluoromax-4 fluorescence spectrometer (Horiba Jobin Yvon) at excitation and emission bandwidths of 4 and 6 nm, respectively. Fluorescence measurements were carried out at room temperature (25 °C) in hTDG or hOgg1 optimized 2-AP base flipping assay buffer. Increasing amounts of

concentrated protein were titrated to DNA substrate (final concentration of 170 nM) to reach final protein concentrations of 3 and 5 μM in the reactions. The samples were rapidly mixed and incubated for at least 5 min at room temperature to ensure binding was at equilibrium. Reference spectra were recorded for the hOgg1 and hTDG binding buffers in the absence of DNA and protein to correct the emission spectra for the fluorescence background of the buffers, respectively. Furthermore, control spectra were determined for the protein solutions (3 and 5 μM for hTDG and 5 μM for hOgg1) in the absence of DNA (in the same buffer solutions) that were subtracted from the emission spectra in the presence of 2-AP DNA and protein to account for contributions to the fluorescence spectra from the proteins. All experiments (except for the control spectra of the protein solutions in the absence of DNA) were performed in duplicates of individual titrations.

For quantification of the protein induced 2-AP fluorescence intensity change ($\Delta I_{2\text{-AP}}$), the maxima of the 2-AP fluorescence intensity at a wavelength of $\lambda_{\text{em}} = 373 \text{ nm}$ were determined for 2-AP DNA in the absence ($I_{2\text{-AP,DNA}}$) and in the presence of the maximum applied protein concentration ($I_{2\text{-AP,DNA+protein}}$) for hTDG and hOGG1. Hence, the relative 2-AP fluorescence change $\Delta I_{2\text{-AP}}$ induced by base flipping or destabilization is given by:

$$\Delta I_{2\text{-AP}} = I_{2\text{-AP,DNA+protein}} - I_{2\text{-AP,DNA}} \text{ (Equation 10)}$$

Fluorescence intensities were determined as averages from two different titration experiments with standard deviations derived from the variation between experiments. The consistently higher 2-AP fluorescence intensity changes observed for hOgg1 compared to hTDG may in part be due to the different chemical environments for the two proteins; hOGG1 contains 10 tryptophans while hTDG only has two.

5. Results and discussion

The two DNA excision repair pathways, NER and BER, differ significantly in their strategies employed for DNA lesion search and recognition. The broad DNA substrate versatility of NER is based on the bipartite substrate discrimination including an initial damage sensing step and an ATP-dependent damage verification step. In eukaryotic NER, these processes involve a coordinated cascade of different proteins and a multi-subunit damage verification complex. Contrarily, damage recognition and removal in BER is fulfilled by a variety of different DNA glycosylases—each of them highly specialized for a narrow subset of base lesions. Notably, DNA glycosylases neither require further protein factors nor the energy from ATP hydrolysis to accomplish these tasks. Specifically, I focus on the NER helicase XPD and on two representatives from the vast glycosylase family, hTDG and hOgg1, to compare those different approaches for damage search and recognition.

5.1 Substrate versatility of the NER helicase XPD in damage verification DNA

Efficient and correct damage verification is of vital importance for the cell. On the one hand, the GG-NER system has to be highly specific for damaged bases to avoid futile incisions of undamaged bases. On the other hand, the detection machinery has to be versatile enough to recognize a variety of structurally different lesions. In the proposed bipartite damage recognition model (see 2.3.2), a helix-distortion induced by a ‘base-pairing disruption’ is initially detected by XPC followed by damage verification by the TFIIH complex, which ensures the presence of a chemical modification (Naegeli et al., 1992, Hess et al., 1997, Sugawara, 2010). The TFIIH complex is comprised of two sub-complexes, the core-complex (7 subunits) and the CAK-complex (3 subunits), which is dispensable for NER. The two key players of the TFIIH are the ATPase XPB, which has been proposed to enlarge the repair bubble by its ATPase activity (Coin et al., 2007) and the 5’-3’ helicase XPD, which is implicated in unwinding the DNA around the damage by its helicase activity (Tapias et al., 2004, Compe and Egly, 2012). Apart from this well established function in NER, XPD has been suggested to be involved in damage verification and DNA translocation of XPD has been shown to be stalled at lesion sites (Naegeli et al., 1992, Hess et al., 1997, Sugawara, 2010). Sugawara and co-workers provided evidence that the orientation in which the TFIIH complex is loaded on DNA is essential for correct damage verification and processing (Sugawara et al., 2009). Currently, the exact processes of ‘kinetic proofreading’, however, are not completely understood and it is still a matter of debate, which TFIIH proteins are involved in this process. Although XPD appears to be the most obvious candidate, damage verification may also depend on other TFIIH subunits such as XPB or even on TFIIH non-members such as XPA, which is indicated to possess a binding preference for

single-stranded, chemically modified DNA (Camenisch et al., 2006, Sugawara et al., 2009). To focus solely on the function of XPD, I reconstituted a simplistic *in vitro* model for damage verification in NER using archaeal XPD from *T. acidophilum*, which requires only a few NER factors including XPB and the XPG homolog, FEN-1 (Rouillon and White, 2011). Due to the high sequence similarity to its human counterpart, the archaeal XPD homolog can be employed as an adequate model for eukaryotic NER as described above (2.3.3). In these studies, I compared interactions of taXPD with DNA substrates containing two structurally different NER lesions, a bulky fluorescein-adduct and a CPD damage, which causes only minor helix-distortions compared to other UV-photoproducts (Figure 9).

5.1.1 Purification of taXPD proteins

To obtain sufficient amounts of active taXPD protein for my studies, the protein was overexpressed in *E.coli* cells and purified in 2 steps by metal affinity and size exclusion chromatography based on an established protocol (Wolski et al., 2008, Kuper et al., 2012) (see 4.3).

At an absorption wavelength of 280 nm, the size exclusion chromatogram of taXPD wt (Figure 20A) shows five different peaks at elution volumes of ~115 mL (peak 1), ~120 mL (peak 2), ~215 mL (peak 3), ~245 mL (peak 4) and ~270 mL (peak 5). Subsequently, the fractions corresponding to these peaks were analyzed by SDS-PAGE (Figure 20B). The predominant peak 3 (at ~215 mL elution volume) contains taXPD as indicated by bands on the SDS gel corresponding to the molecular weight of the overexpressed protein (72 kDa) (red arrow). Moreover, the absorption of the FeS cluster at 410 nm overlaps with the major protein absorption peak at 280 nm (peak 3) in the chromatogram. Fractions running at an elution volume of 205-235 mL were pooled for further concentration and re-checked on a SDS-gel (Figure 20E).

The size exclusion chromatogram of the taXPD variant R567W (Figure 20C) shows three distinct peaks at an absorption wavelength of 280 nm. The predominant peak in protein absorption at an elution volume of ~86 mL corresponds to XPD as visible from the SDS gel (Figure 20D, red arrow) showing pooled fractions (elution volume 82-95 mL) of the purified protein R567W. The purity of the purified proteins was estimated from the intensities of the bands on the SDS gel. The purity of the variant R567W (~80 %) was lower compared to the purity of the wt (~95 %) (compare Figure 20D and E, red arrows). Impurities at a running position of ~25 kDa, most likely protein degradation products, were more pronounced for the mutant than the wt protein (compare Figure 20D and E, blue arrows).

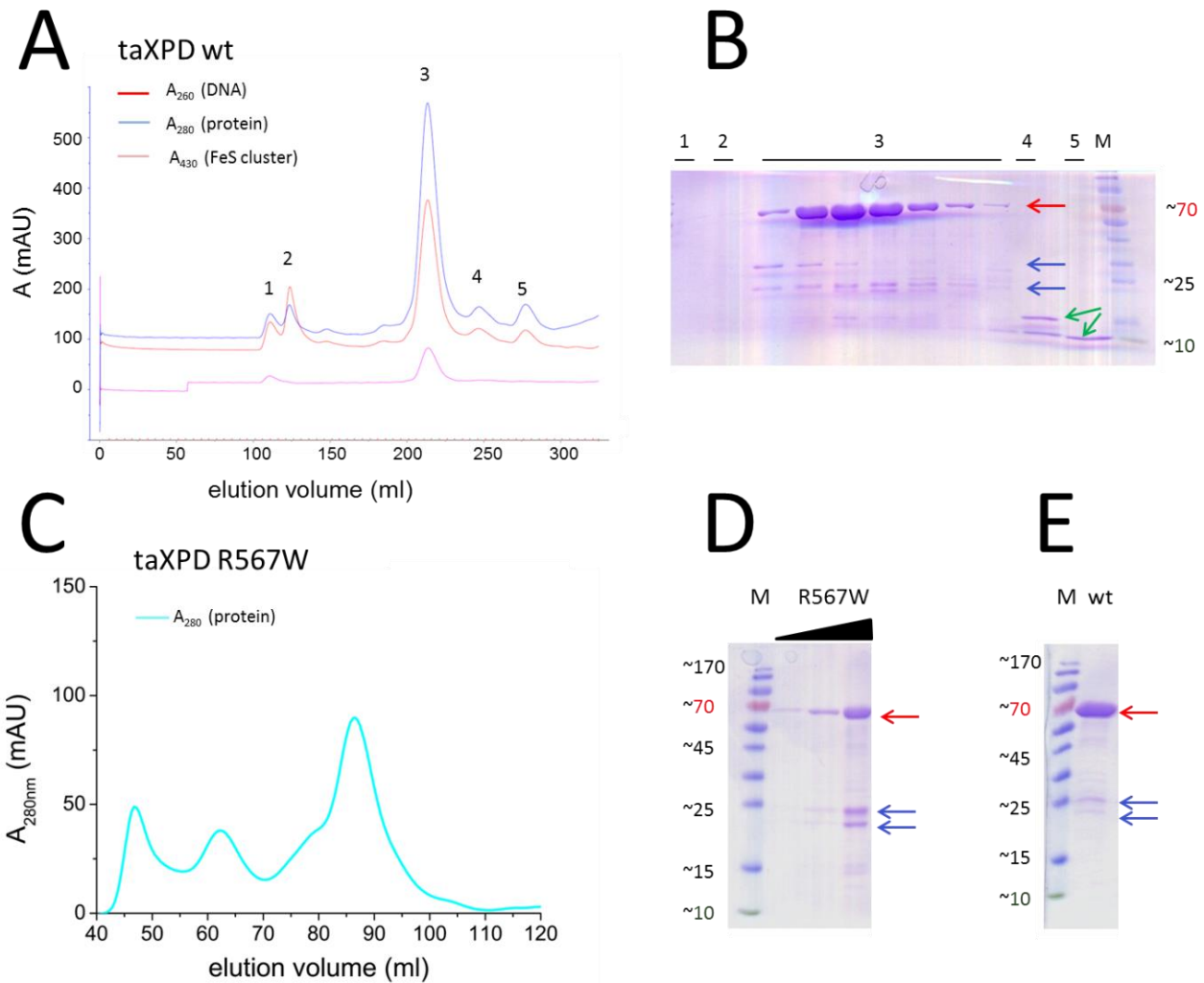


Figure 20: Purification of XPD wt from *T. acidophilum*.

A) Size exclusion chromatogram of taXPD wt. During SEC, absorption was recorded at three different wavelengths as indicated. B) SDS-PAGE of taXPD wt samples showing the elution fractions, which correspond to the five peaks in the elution profile in (A). C) Size exclusion chromatogram of taXPD R567W. For the variant, only protein absorption at a wavelength of 280 nm was detected. D) SDS-PAGE of pooled and concentrated samples of taXPD R567W (8.0 ng) and two different dilutions (0.9 ng and 0.2 ng). E) SDS-PAGE of pooled and concentrated samples of taXPD wt (4.9 ng). The running position of taXPD (~70 kDa) is indicated by red arrows in the SDS gel. Green and blue arrows point to impurities running at either 10-15 kDa or at 25-30 kDa.

In order to confirm proper folding of the taXPD wt and the variant R567W CD spectroscopy was applied. The spectra for XPD wt and the protein variant are similar (Figure 21), with minima observed at 208 nm and 222 nm, which are characteristic for correct protein folding in α -helices and β -sheets. The transition from negative to positive ellipticity values at a wavelength of approximately 200 nm indicates the absence of disordered regions in the protein structures.

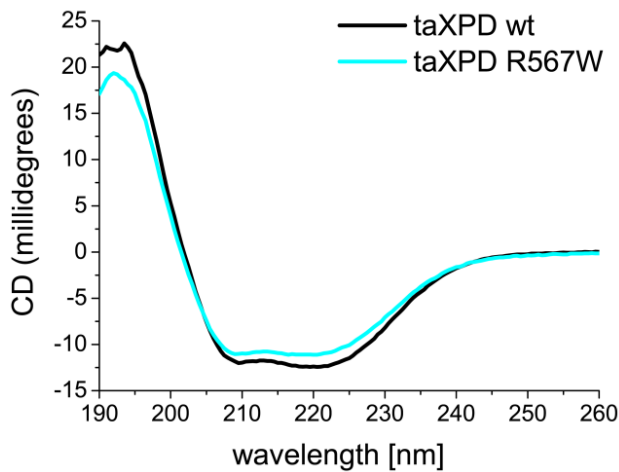


Figure 21: CD spectra of taXPD wt and variants.

The wt protein (black) and the protein variant R567W (cyan) showed similar CD profiles. The comparable spectra for both proteins indicate correct protein folding in α -helices and β -sheets and the absence of disordered regions in the protein structures.

Finally, DNA helicase activity of the purified taXPD wt protein was tested. The unwinding assay is based on different electrophoretic mobilities of dsDNA and ssDNA resulting in gel shifts. The applied DNA substrate ('open fork', see schematic in Figure 22) allowed loading of taXPD in 5'-3' direction. With rising concentration of XPD, the intensity of the bands corresponding to ssDNA slightly increased (Figure 22), which indicated weak 5'-3' helicase activity of taXPD. Although no dsDNA separation was expected in the negative control without ATP and in the negative dsDNA control without protein, we observed faint bands corresponding to ssDNA most likely due to incomplete annealing of the duplex.

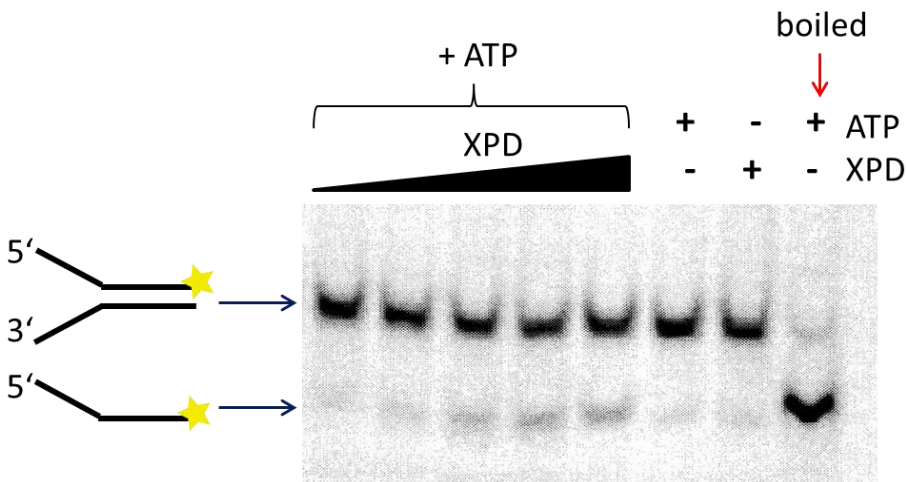


Figure 22: Gel-based helicase assay of taXPD.

The gel shows different samples of DNA and/or taXPD in the presence or absence of ATP as indicated. With rising taXPD concentration (0.1, 0.25, 0.5, 0.75 and 1 μ M), the helicase activity (separation of dsDNA is visible as a shift in electrophoretic mobility) increases. The structure of the applied DNA substrate is schematically depicted (Cy3 label = yellow star) with the running positions of dsDNA and ssDNA indicated by blue arrows.









After validation by SDS-PAGE analysis, CD spectroscopy and a gel-based helicase assay (for the wt protein), DNA interactions of the purified taXPD proteins were investigated in BLI and AFM experiments.

5.1.2 taXPD preferentially loads on unpaired regions within dsDNA

The current model of NER (Figure 6) suggests that the physiological target of XPD contains a small unpaired DNA region (bubble), which is introduced by the concerted action of XPC and XPB *in vivo* (Lindahl and Wood, 1999, Hoeijmakers, 2001, Friedberg, 2001, Friedberg, 2003, Oksenych et al., 2009). Consistent with that, XPD has been shown to possess a binding preference for ssDNA and ss/dsDNA junctions in several *in vitro* studies (Mathieu et al., 2010, Rudolf et al., 2010, Kuper et al., 2012).

To test for the proposed binding preference, DNA binding affinities of taXPD were determined for different DNA substrates using the label-free biosensor technique BLI. The DNA substrates tested included ssDNA, DNA bubble (8 nt) substrates and fully matched DNA duplexes either in the absence or in the presence of a fluorescein or CPD damage (Table 16). In order to concentrate exclusively on DNA binding effects, the binding experiments were carried out in the absence of ATP to prevent ATP-dependent DNA duplex unwinding and translocation on DNA.

Table 16: DNA binding affinities of taXPD for different DNA substrates.

DNA substrate		K_D [nM]	n	P
ssDNA		155 ± 61	13	-
dsDNA nsp		541 ± 234	12	6.4 x 10 ⁻⁶
- / bubble (8nt)		155 ± 75	16	0.48
F / bubble (8nt)		163 ± 69	14	0.36
F / -		499 ± 210	12	7.9 x 10 ⁻⁶
CPD / bubble (8nt)*		138 ± 52	8	0.28
CPD / -		340 ± 94	11	5.5 x 10 ⁻⁶

Equilibrium binding constants (K_D) for different DNA substrates (as depicted schematically) were determined by BLI. Binding affinities were obtained as means ± SD from n different experiments. Binding experiments were performed with three (or *two) different protein purification batches. Binding affinities of taXPD for ssDNA and dsDNA substrates containing single stranded DNA regions (-/bubble, F/bubble, CPD/bubble) are similar ($P > 0.28$) (highlighted in blue). The difference in binding affinity between ssDNA and fully matched dsDNA substrates (dsDNA nsp, F/-, CPD/-) is highly significant ($P < 10^{-5}$). One of the dsDNA substrates (CPD/-) is bound by taXPD with significantly higher affinity compared to completely matched nsp dsDNA ($P = 0.0095$). Statistical analysis of significances was performed with a two-tailed student t-test. All P-values given in the table were derived in comparison with the ssDNA substrate. DNA sequences of DNA substrates for BLI are indicated in Table 11. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

In the BLI experiments, taXPD bound to ssDNA and to dsDNA containing single stranded DNA regions with moderately higher affinity (~ 150 nM, Table 16, blue) compared to completely matched dsDNA substrates (~ 500 nM, Table 16, white). The observed binding preference for ssDNA and DNA bubble substrates over fully matched DNA duplexes suggests that taXPD may preferentially load onto DNA on ssDNA regions and is consistent with the proposed binding preference of XPD for ssDNA and ss/dsDNA junctions in previous studies (see above). The finding that, taXPD bound to ssDNA (155 ± 61 nM) and to DNA substrates containing an 8 nt bubble (155 ± 75 nM) with similar affinity further indicated that the size of the bubble (8 nt) is sufficient to allow loading of taXPD.

Furthermore, the K_D values obtained for DNA substrates containing either a CPD (138 ± 52 nM) or fluorescein damage (163 ± 69 nM) in a bubble context were comparable to those of damage-free bubble substrates (155 ± 75 nM). This implies that the DNA binding affinity of taXPD is rather dominated by the preference for the ssDNA region than by preference for the damage. Interestingly, the DNA binding affinity of XPD for the DNA substrate containing a CPD within fully matched DNA (340 ± 94 nM) was significantly higher compared to completely matched nsp dsDNA (541 ± 234 nM, $P = 0.0095$) and compared to dsDNA substrates containing fluorescein within completely matched DNA (499 ± 210 nM). In contrast to fluorescein, which is connected extra-helically to a thymine base with a short linker, the CPD damage is an intra-strand crosslink between two adjacent thymines and mildly destabilizes the DNA helix (Sugasawa et al., 2001, Reardon and Sancar, 2003). Thus, the CPD damage may represent a small access site for XPD and allow slightly enhanced loading of XPD to the site of the damage compared to fluorescein or nsp homoduplex DNA.

In accordance with BLI measurements, which revealed a moderate binding affinity of taXPD for nonspecific homoduplex DNA compared to ssDNA and DNA substrates containing unpaired DNA regions (Table 16), complex formation of taXPD with nonspecific dsDNA was also discovered in AFM experiments. Exemplarily, Figure 23A shows an AFM image of taXPD bound to circular non-damaged pUC19N DNA. This observed complex formation with nonspecific dsDNA represents a novel finding, which has not been proven experimentally before. To quantify DNA coverage with XPD in AFM experiments, AFM binding densities (number of peaks/bp, see Equation 1 in 4.1.3) were further determined in the absence and in the presence of ATP for different DNA substrates including circular plasmid DNA, linear undamaged DNA fragments and DNA substrates containing a loading site for XPD (8 nt bubble) either in the absence or in the presence of fluorescein (Figure 23B).

Consistent with the BLI experiments, ATP was generally not required for association of nonspecific taXPD-DNA complexes (grey bars in Figure 23B). Interestingly, however, the binding affinity for circular nonspecific DNA was significantly reduced compared to linear nonspecific DNA in the absence of ATP

(compare grey bars for circular and linear nsp DNA in Figure 23B, $P = 0.011$, mildly significant). This indicates that taXPD also may be capable of loading on DNA termini, which represent destabilized DNA structures and are a preferred binding site for many other DNA repair proteins (Yang et al., 2005, Tessmer et al., 2008, Wagner et al., 2009, Tessmer et al., 2012).

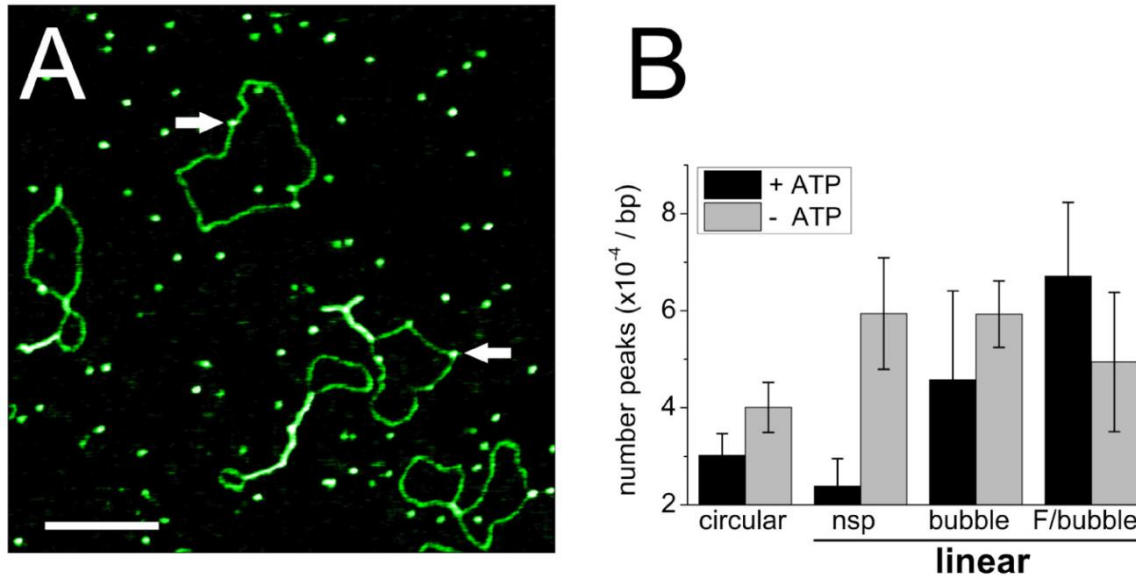


Figure 23: DNA binding of taXPD in AFM experiments.

A) AFM image of taXPD incubated with non-damaged plasmid pUC19N (2729 bp) in the presence of ATP. White arrows point to taXPD complexes bound to nonspecific circular DNA. The scalebar corresponds to 200 nm. B) AFM binding densities of taXPD on different DNA substrate in the presence (black) and absence of ATP (grey). For calculation of binding densities, the protein coverage is normalized to the number of DNA fragments (counted from the AFM images) and the number of bp (number of peaks/bp). DNA substrates: circular = pUC19N plasmid; linear: nsp = linearized full length pUC19N (2729 bp) and 1813 bp fragment, bubble = 916 bp fragment containing a 8 nt DNA bubble, F/bubble = 916 bp fragment containing a fluorescein (F) within the context of a DNA bubble. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

Upon addition of ATP, the binding density for linear undamaged DNA was significantly lowered compared to the absence of ATP (compare black and grey bars for linear nsp DNA in Figure 23B, $P = 0.00073$, highly significantly different) and only slightly lowered for circular DNA (compare black and grey bars for circular nsp DNA in Figure 23B, $P = 0.023$, mildly significantly different). As XPD is an ATP-dependent DNA translocase, this finding suggests that taXPD may translocate on nonspecific dsDNA in an ATP-dependent manner and slide off the DNA fragment ends.

To investigate whether the presence of a DNA damage may prevent that taXPD slides off or dissociates from the DNA fragments, a fluorescein adduct, which is known to be removed by the NER system (Hoeijmakers, 2009, Rudolf et al., 2010), was introduced into the DNA substrates for AFM (see 4.2.2) within the context of a small bubble (8 nt). As expected, the binding density of taXPD on DNA substrates containing a fluorescein damage within a bubble was significantly higher than for undamaged DNA

bubble substrates, when adding ATP (compare black bars for linear bubble and F/bubble DNA in Figure 23B, $P = 0.014$, mildly significant), which suggests that the ATP-dependent DNA translocation of XPD may be constrained by the presence of a lesion. However, in the absence of ATP and hence in the absence of ATP-dependent translocation on DNA, the number of XPD-DNA peaks was similar for both DNA substrates (compare grey bars for linear bubble and F/bubble DNA in Figure 23B, $P = 0.164$, not significant). Consistent with these findings, an NER lesion may serve as a mechanical road block for XPD translocation on DNA.

5.1.3 Lesion recognition and stalling at lesion sites requires ATP hydrolysis

Figure 24 shows a representative AFM image of taXPD incubated with DNA substrates containing a fluorescein lesion within a bubble context at $\sim 30\%$ of DNA fragment length in the presence of ATP.

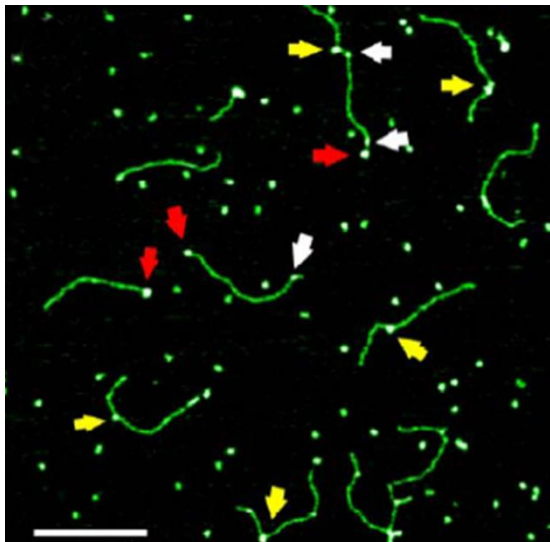







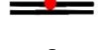




Figure 24: Lesion binding of taXPD in AFM.

Representative AFM image of taXPD incubated with lesion specific DNA substrates (916 bp) containing a fluorescein in the context of a 8 nt bubble (at $\sim 30\%$ of DNA fragment length) in the presence of ATP. To distinguish between specific (lesion-bound) and nonspecific XPD complexes, protein positions on DNA were determined. Yellow arrows point to XPD complexes specifically bound at the lesion site, white and red arrows point to nonspecific XPD-DNA complexes bound to undamaged dsDNA and DNA ends, respectively. The scale bar corresponds to 200 nm. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

Based on the finding that XPD showed enhanced binding density on DNA substrates containing a fluorescein damage in a bubble context compared to undamaged DNA bubble substrates in the presence of ATP (see above in 5.1.1 and Figure 23B), binding stability of XPD is assumed to be increased at an NER target lesion. To quantify this proposed enhancement, binding positions of XPD-DNA complexes were determined with respect to the closer DNA terminus (see Methods 4.1.3). As the position of the DNA target site on the DNA substrates is known ($\sim 30\%$ from the closer DNA terminus), specific XPD-

complexes at the damage (Figure 24, yellow arrows) and nonspecific protein complexes bound at DNA fragment ends (red arrows) or bound elsewhere on undamaged DNA (white arrows) can be distinguished. Using the ratio of areas under the Gaussian peak at the specific site and of the nonspecific background from the resulting position distribution histograms, the localization specificity (or binding specificity *S*) of XPD to a specific target site was calculated (for exact values see Table 17).

Table 17: AFM binding specificities of taXPD for different DNA substrates and under various conditions.

DNA substrate		+ ATP	+ ATPys	- ATP
- / bubble (8 nt)	1 	131 ± 51	n.d.	213 ± 21
F / bubble (8 nt)	2 	851 ± 280 655 ± 9	231 ± 67	169 ± 168 112
F / bubble (14 nt)	2b 	468 ± 157	n.d.	n.d.
F / 5' bubble (8 nt)	3 	445 ± 167	102 ± 60 101 ± 53	n.d.
F / 3' bubble (8 nt)	4 	95 ± 24	n.d.	n.d.
F / -	5 	113 ± 89	n.d.	n.d.
CPD / bubble (8 nt)	6 	522 ± 112	374 ± 105	66 ± 51 51
CPD / 5' bubble (8 nt)	7 	117 ± 39 109 ± 11	96 ± 48	n.d.
CPD / 3' bubble (8 nt)	8 	283 ± 99 452 ± 3	82 ± 70	n.d.
CPD / -	9 	137 ± 105		

AFM binding specificities were determined from position distribution histograms (e.g. from Figure 25) and reflect average values from 2-4 individual experiments with standard deviations (SD) derived from the variations between single experiments. In grey below, results obtained with a second protein batch for some of the DNA substrates indicate good reproducibility (F = fluorescein). This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

In the presence of ATP, taXPD showed a high localization preference ($S = 851 \pm 280$) to a bulky fluorescein damage within a central bubble (DNA substrate fluorescein/bubble) (Figure 25A, black bars and Table 17) over an excess of nonspecific dsDNA. Interestingly, in the absence of ATP, the binding

specificity for fluorescein in a bubble context was significantly ($P = 0.027$) reduced (Figure 25A, grey bars, $S = 169 \pm 168$). The residual specificity for the target site in the absence of ATP is likely due to XPD's binding preference for the unpaired ssDNA region, consistent with the higher affinity of XPD for the fluorescein/bubble DNA substrate and the -/bubble DNA substrate ($K_D = 163 \pm 69$ nM and $K_D = 155 \pm 75$ nM vs. $K_D = 541 \pm 234$ nM for nonspecific dsDNA) as determined in BLI binding experiments. To identify, whether the high localization specificity of XPD in AFM to the DNA substrate containing a fluorescein located within a DNA bubble was due to its preference for loading onto DNA bubble structures or due to its specificity for the target site, two control experiments were performed in AFM. The interaction of taXPD with DNA substrates containing an unpaired DNA region without a lesion (-/bubble) was compared with those containing a fluorescein damage within perfectly matched DNA (fluorescein/-).

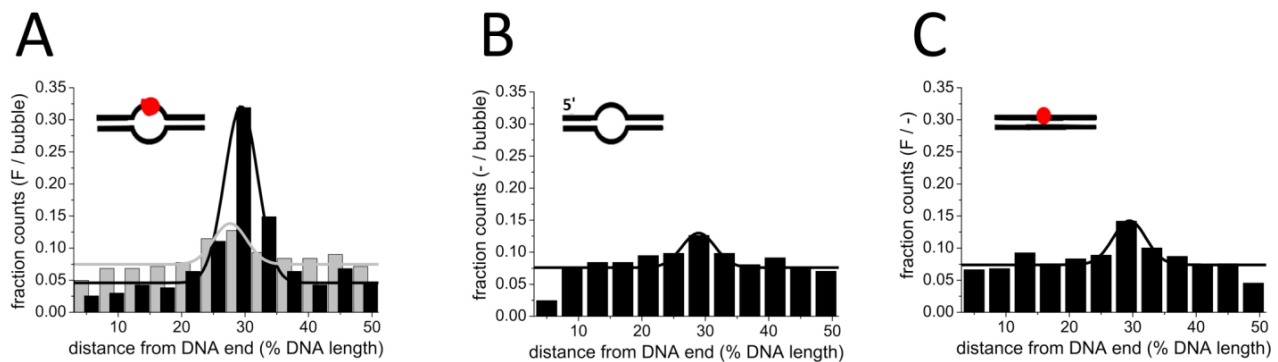


Figure 25: Specificity of taXPD for different DNA target sites.

XPD binding position distributions on three different DNA substrates as schematically indicated. The target sites are located at $\sim 30\%$ to the DNA fragment end. Fractional occupancies are plotted for ~ 33 bp long sections of 916 bp DNA from DNA fragment ends (0%) to 50% DNA length (at the DNA center). A) Complex position distribution on the fluorescein/bubble substrate showing high specificity for the target site in the presence of ATP (black, $n = 231$, $S = 851 \pm 280$) and considerably reduced binding specificity in the absence of ATP (grey, $n = 398$, $S = 169 \pm 168$). This finding indicates the relevance of ATP-dependent transitions for lesion recognition of XPD. B) The complex position distribution on a bubble substrate without a lesion site (bubble/-) ($n = 515$, $S = 131 \pm 51$) demonstrates a low binding specificity of taXPD for this target site in DNA most likely due to unhindered, ATP-dependent DNA translocation of XPD. C) For DNA substrates containing a fluorescein damage within perfectly matched DNA (fluorescein/-) ($n = 280$, $S = 113 \pm 89$), the observed low binding specificity is consistent with the lower binding affinity in BLI. All distributions represent pooled data from at least three individual AFM experiments. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

Firstly, the localization specificity to undamaged bubble sites over nonspecific DNA was poor ($S = 131 \pm 51$, Figure 25B) compared to the fluorescein/bubble target structure ($S \sim 850$), despite XPD's high binding affinity for unpaired DNA. Thus, the high specificity of taXPD for fluorescein in the bubble context can be clearly ascribed to its high specificity for the fluorescein lesion. The residual lesion specificity for the target site -/bubble in the presence of ATP is used to define the 'background level' of the localization specificity of taXPD in the AFM experiments.

Secondly, for the DNA substrate fluorescein/- (Figure 25C), a significantly reduced binding specificity of $S = 113 \pm 89$ was found in AFM compared to a fluorescein in the context of an 8 nt DNA bubble ($P = 0.00028$), which is consistent with the lower binding affinity for this DNA substrate in BLI ($K_D = 499 \pm 210$ nM) compared to ssDNA. These data indicate that loading of XPD onto DNA may be impaired in the absence of a ssDNA-region. Importantly, the lesion specificity of taXPD for fluorescein was not increased by enlarging the bubble size ($S = 468 \pm 157$ for fluorescein/bubble 14 nt compared to $S = 851 \pm 280$ for fluorescein/bubble 8 nt and Table 17), which suggests that an 8 nt bubble is sufficiently large to allow loading of XPD onto DNA.

To separate the effects of loading on a bubble from damage detection, we produced DNA substrates containing an 8 nt bubble in 27 bp distance 5' from the fluorescein damage (5' fluorescein/bubble). This distance between the entry site for loading of XPD onto DNA (bubble) and the target site is small enough to be overcome by XPD's 5'-3' helicase activity as has been previously shown (Mathieu et al., 2010). Although a distance of 27 bp (9 nm with 0.34 nm/bp) is too small to spatially resolve XPD complexes bound at the bubble or XPD complexes bound at the fluorescein damage in the AFM images, the resulting position distributions of XPD on these DNA substrates can reveal important information on the implications of separating loading and target recognition sites.

Indeed, I discovered that XPD could localize to the target site fluorescein (Table 17 and Figure 28A) with relatively high specificity of $S = 445 \pm 167$, when loaded at the bubble 5' to the damage in the presence of ATP. This finding suggests that the helicase XPD translocates on DNA in 5'-3' direction in an ATP-dependent manner and that its movement is stalled upon recognition of a fluorescein damage on the translocated DNA strand. In contrast, XPD hardly recognized the fluorescein damage, when located on the complementary, non-translocated strand (fluorescein/ 3' bubble $S = 95 \pm 24$, Figure 28B). Note that, XPD binding specificity for fluorescein on the translocated strand ($S = 445 \pm 167$, Table 17 and Figure 28B) was approximately 50% of that observed for fluorescein within the context of a central bubble ($S = 851 \pm 280$), consistent with approximately half of XPD translocating along DNA in 5' to 3' direction on either DNA strand. In the presence of the non- or only slowly-hydrolysable ATP analog ATP γ S, localization specificity of XPD to fluorescein was decreased to background levels (defined for the DNA substrate bubble/- in the presence of ATP), when loaded at a bubble 5' of the damage ($S = 102 \pm 60$, Table 17), which suggests that the 5'-3' translocation on DNA is coupled to ATP hydrolysis.

The finding that the processes leading to damage verification are ATP-dependent were confirmed by the effects of the taXPD variant R567W in AFM experiments. The variant R567W corresponds to R666W in the human protein and causes the disease XP, when mutated in XP patients (Wolski et al., 2008). As the

mutated residue Arg⁵⁶⁷ is located near the ATP-binding site in the crystal structure (Figure 7, between helicase domains 1 and 2, red and yellow), it has been proposed that the variant is deficient in ATP binding and hydrolysis (Wolski et al., 2008). In line with that, the localization specificity of R567W to the target site fluorescein/bubble in the presence of ATP ($S = 216 \pm 33$, Figure 26) was reduced 4-fold compared to XPD wt under the same conditions ($S = 851 \pm 280$, Table 17) and resembled the value obtained for the wt protein in the absence of ATP ($S = 169 \pm 168$, Table 17). In addition, DNA binding affinities of the variant R567W were determined for different DNA substrates by BLI in the absence of ATP (Figure 26). In the BLI experiments, taXPD R567W bound to ssDNA and to dsDNA containing single stranded DNA regions (F/bubble, -/bubble) with significantly lower affinity ($P < 1.9 \cdot 10^{-4}$, ~ 350 nM, Figure 26) compared to the wt enzyme (~ 150 nM, Figure 26 and Table 16, blue). The binding affinity of R567W determined for the nsp dsDNA substrate, however, was comparable to those measured for taXPD wt and dsDNA (624 ± 30 nM) and 541 ± 234 nM, respectively, $P = 0.26$, not significantly different, Figure 26). Thus, the reduced localization specificity of the taXPD variant to the target site fluorescein/bubble in AFM is consistent with this DNA binding phenotype observed for the variant R567W in the binding experiments (reduced binding affinity for ssDNA and DNA substrates containing ssDNA regions; similar binding affinity for dsDNA compared to wt).

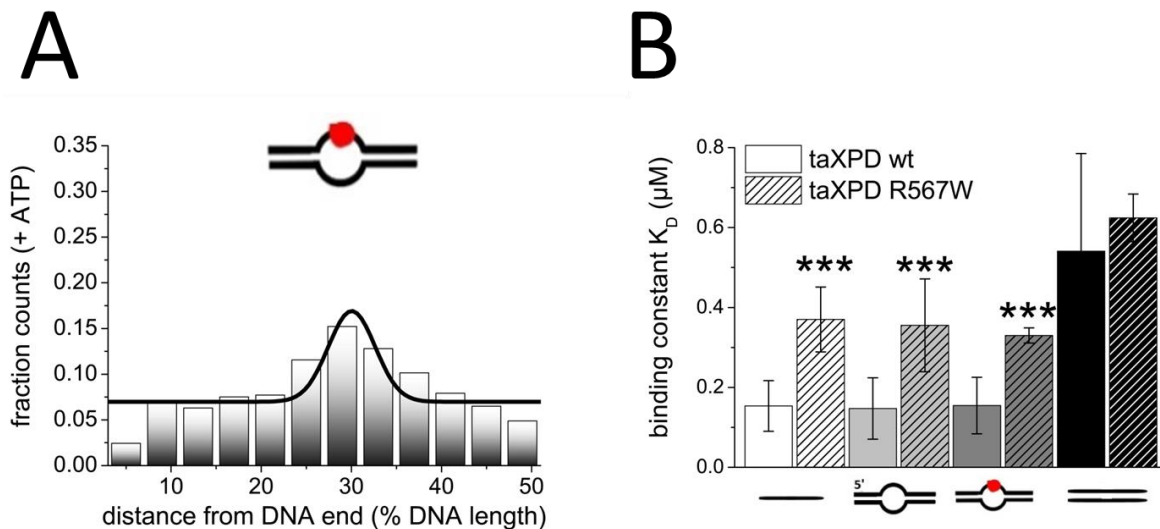


Figure 26: Localization specificity and binding affinities of taXPD R567W for different DNA substrates.

A) The complex position distribution of taXPD R567W on the fluorescein/bubble substrate shows low specificity ($S = 216 \pm 33$) for the target site in the presence of ATP ($n = 493$), which suggests that this variant is impaired in lesion recognition. The distribution represents pooled data from five individual AFM experiments. B) Equilibrium binding constants (K_D) of taXPD wt (blank bars) and taXPD R567W (hatched bars) for different DNA substrates (as depicted schematically) were determined by BLI in the absence of ATP. Binding affinities were obtained as means \pm SD from at least three different experiments. Differences in binding affinities between taXPD R567W for ssDNA and DNA substrates -/bubble and F/bubble were highly significant to those observed for the wt enzyme and the respective DNA substrates ($P < 1.9 \cdot 10^{-4}$). Statistical analysis of significances was performed with a two-tailed student t-test. All P-values given in the figure were derived in comparison between taXPD wt and the variant R567W for each DNA substrate. DNA sequences of DNA substrates for BLI are indicated in Table 11.

To further characterize the significance of proper lesion verification by XPD, AFM experiments were performed with a second mutant, taXPD K170A. It is likely that Arg¹⁷⁰ is involved in lesion recognition as the mutated residue is located in close vicinity to the FeS cluster (compare Figure 7), which has been implicated in damage verification (see 2.3.4).

In contrast to the variant R567W, which has been indicated to be impaired in ATP binding and hydrolysis, the variant K170A has recently been described to show helicase hyperactivity and elevated ATPase activity compared to the wt protein (Kuper et al., 2012). In the AFM experiments, taXPD K170A showed very low lesion specificity for a fluorescein damage located at a distance 3' of a bubble (DNA substrate fluorescein/5' bubble, $S = 67 \pm 14$, Figure 27). The finding that DNA translocation of this helicase mutant was not stalled by fluorescein is consistent with taXPD K170A 'overreading' the lesion sites due to its elevated helicase activity. Altogether, the data on the XPD variants R567W and K170A suggest that damage verification requires a) the ability to bind and hydrolyze ATP and b) 'normal' helicase activity and helicase stalling at lesion sites, respectively.

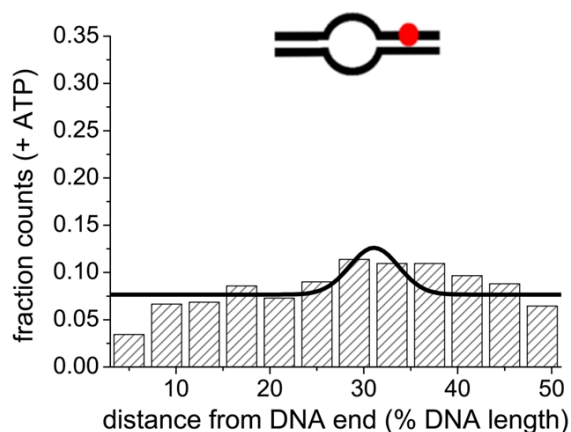


Figure 27: Specificity of taXPD K170A for the DNA substrate fluorescein/5' bubble.

The complex position distribution on the fluorescein/5' bubble substrate shows only a residual specificity ($S=74 \pm 20$ corresponding to background levels) for the target site in the presence of ATP ($n=562$) demonstrating that this variant is clearly impaired in lesion recognition. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

5.1.4 taXPD utilizes different mechanisms for verification of CPD and fluorescein lesions

The NER pathway recognizes and removes a variety of structurally diverse DNA targets embracing the class of helix-distorting lesions, which may disrupt normal base pairing. The fluorescein adduct, which has been used in these studies with taXPD so far, is an artificial substrate removed by the NER system and serves as representative for bulky DNA damages. In identical studies, the results for taXPD interactions with fluorescein will be compared with those for the less invasive UV-photoproduct CPD.

For a CPD damage located within the context of a central bubble (CPD/bubble), the lesion specificity ($S = 522 \pm 112$) was found to be comparable (or even slightly lower) to those for the fluorescein/bubble target structure ($S = 851 \pm 280$) ($P = 0.13$). Interestingly, however, for a DNA substrate with a bubble located 5' of a CPD damage (CPD/ 5' bubble), binding specificity was lowered to background levels ($S = 117 \pm 39$, Figure 28C) similar as observed for the DNA substrate -/bubble. This suggests that XPD is not able to detect the CPD lesion, when localized at a distance 3' from the loading site (bubble). This is in contrast to the observation for the fluorescein lesion, which was recognized with high specificity on the translocated strand, i.e. when loaded at a bubble 5' to the lesion (fluorescein/5' bubble substrate). However, XPD binding specificity for CPD on the non-translocated strand i.e. when loaded at a bubble 3' to the lesion (CPD/3'bubble, $S = 283 \pm 99$, Figure 28D) was significantly higher than for CPD placed on the translocated strand (CPD/5'bubble, $S = 117 \pm 39$, Figure 28C). Similar as observed for fluorescein, localization specificity to the target site CPD/3' bubble was approximately half of that to CPD within a bubble context (compare $S = 283 \pm 99$ and $S = 522 \pm 112$, respectively). AFM binding specificities of taXPD wt and variants K170A and R567W for different DNA substrates are summarized in Figure 28E.

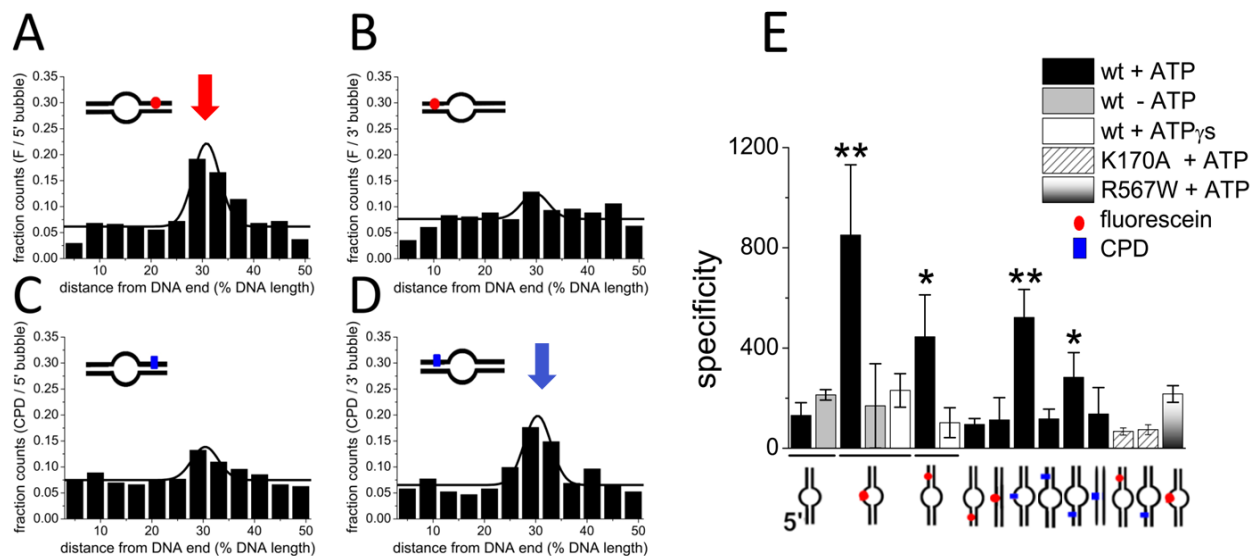


Figure 28: Different strand selectivity of taXPD for bulky fluorescein and CPD damages.

A) XPD displays a high specificity for the position of a fluorescein damage when loaded at bubble 5' to the damage (fluorescein/5' bubble substrate, $n = 537$ complexes, $S = 445 \pm 167$). B) When loaded on a bubble 3' to the damage, XPD poorly recognizes fluorescein (fluorescein/3' bubble substrate, $n = 272$ complexes, $S = 95 \pm 24$). C) XPD shows a low lesion preference for CPD, when loaded on a bubble 5' from the damage (CPD/5' bubble substrate, $n = 555$ complexes, $S = 117 \pm 39$). D) In contrast, XPD binding specificity for CPD is significantly enhanced, when XPD is loaded on a bubble 3' from the damage (CPD/3' bubble substrate, $n = 353$ complexes, $S = 283 \pm 99$). All distributions represent pooled data from at least three individual AFM experiments. E) Quantification of lesion specificities of taXPD wt and variants in AFM under different conditions. Lesion binding specificities were derived from the position distribution histograms shown here and in Figure 25, Figure 26 and Figure 27. Statistical analysis of significances was performed with one-tailed t-test relative to the bubble substrate (* < 0.05 , ** < 0.01 , *** < 0.005). This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

The observed difference in strand selectivity is most likely due to different structural prerequisites necessary for recognition of these structurally diverse lesions. It has been shown earlier that DNA repair efficiencies strongly depend on the degree of DNA helix destabilization by different lesions resulting in differences in NER mechanistic details for different types of lesions (Mu et al., 1996, Jain et al., 2013).

Under the conditions of the AFM experiments (pH 8), fluorescein occurs in the anionic and dianionic form (protolytic constant = $pK_a = 6.4$) (Sjöback et al., 1995). Due to its negative charge, fluorescein may not induce major DNA destabilization by intercalation into the negatively charged DNA. To load on DNA substrates containing a fluorescein lesion, XPD may hence require initial bubble opening, which is proposed to be performed by the concerted action of XPC and XPB *in vivo* (Lindahl and Wood, 1999, Hoeijmakers, 2001, Friedberg, 2001, Friedberg, 2003, Oksenych et al., 2009). taXPD's requirement for initial bubble opening is indicated by the only residual localization specificity to a fluorescein lesion within the context of perfectly matched DNA ($S = 113 \pm 89$ for the DNA substrate fluorescein/-) compared to the high specificity for fluorescein within a bubble context ($S = 851 \pm 280$ for the DNA substrate fluorescein/bubble). When loaded onto DNA at a bubble 5' from the fluorescein lesion, taXPDs movement on DNA in 5'-3' direction may be blocked mechanically by the protruding structure of the fluorescein damage placed on the translocated DNA strand that XPD 'holds on to'.

In case of CPD, the situation is more complex. Similar as fluorescein a CPD damage does not cause major DNA helix distortions compared to other UV-photolesions such as 6,4-photoproducts (Kim and Sancar, 1995, McAteer et al., 1998, Lee et al., 2004, Marteiijn et al., 2014). To characterize the two DNA lesions by AFM, intrinsic DNA bend angles were measured at the target site positions for fluorescein and CPD (30 % from DNA fragment end, Figure 29B and C). In principle, AFM-based measurements of intrinsic DNA bend angles in the absence of protein can provide information on the degree of DNA destabilization induced by the presence of lesions at these target sites. As a reference for intrinsic DNA bending at specific DNA sites, innate bending of nonspecific (undamaged) DNA sites (Figure 29A) was measured at regular intervals of 50 nm by moving a mask comparable to the size of the protein along DNA. The resulting distributions of DNA bend angles for lesion sites and undamaged sites were similar and can be fitted with a semi-Gaussian curve $\sim 0^\circ$, which is consistent with intrinsic bending of nonspecific DNA determined in previous studies (Wang et al., 2003). This suggests that neither fluorescein nor CPD exhibit innate DNA bending and hence may not induce major destabilization of the DNA helix.

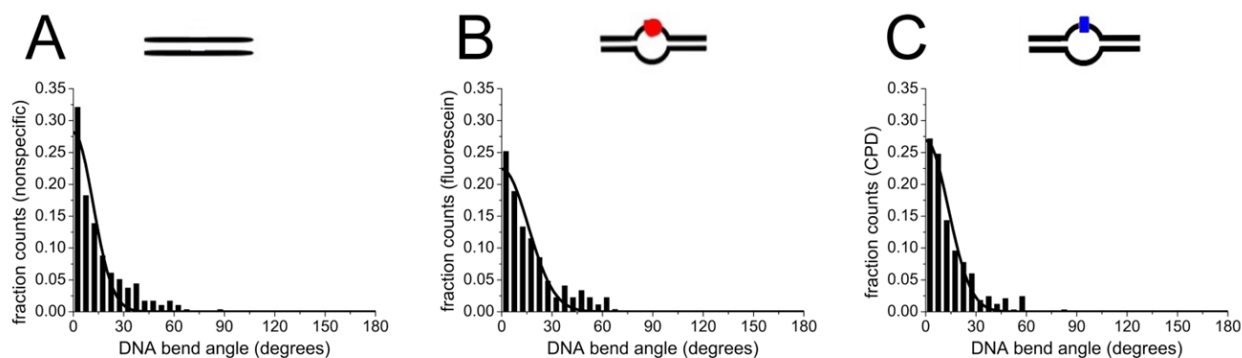


Figure 29: Intrinsic DNA bending of nonspecific and specific DNA substrates in AFM.

DNA bend angle measurements in the absence of protein on three different DNA substrates as schematically indicated. (A) Intrinsic bending of the nonspecific DNA substrate was determined by moving a mask comparable to the size of the protein at regular intervals of 50 nm along DNA ($n = 297$). (B) Fluorescein in the context of a DNA bubble ($n = 270$) (C) CPD in the context of a DNA bubble ($n = 335$). All distributions were fitted by a semi-Gaussian $\sim 0^\circ$ with fit qualities $R^2 > 0.91$ and with widths of 10-15°. The histograms represent pooled data from two individual experiments. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

However, compared to fluorescein, a CPD damage is not bulky and may not present a major mechanical barrier for translocation of XPD because CPD does not significantly stall movement of XPD, when loaded on a bubble 5' from the damage in the AFM studies (i.e. CPD on the translocated strand) (Figure 28). In contrast, when loaded on a bubble 3' from the damage (CPD on the displaced, non-translocated strand), CPD is recognized by XPD. Thus, for recognition of CPD, XPD may employ a different, so far unidentified damage verification mechanism. Based on the three-dimensional XPD-DNA model (see crystal structure in Figure 30), CPD is assumed to be in close vicinity to the FeS cluster, when placed on the complementary, non-translocated strand that XPD does not directly bind to. This observation would support a possible role of the FeS cluster in electrochemical sensing of CPD damages as described above (see 2.3.4 and Figure 8). Unfortunately, the exact location of the CPD damage on the displaced strand cannot be known with certainty as the path of the non-translocated DNA strand has not yet been resolved in the crystal structure of XPD (Kuper et al., 2012) and will require further investigation in crystallographic studies.

It is tempting to speculate that XPD may simultaneously use both detection strategies (mechanical detection of bulky damages and electrochemical sensing of CPD damages) during damage search. Depending on the type of lesion, the equilibrium may then be shifted towards one of the pathways.

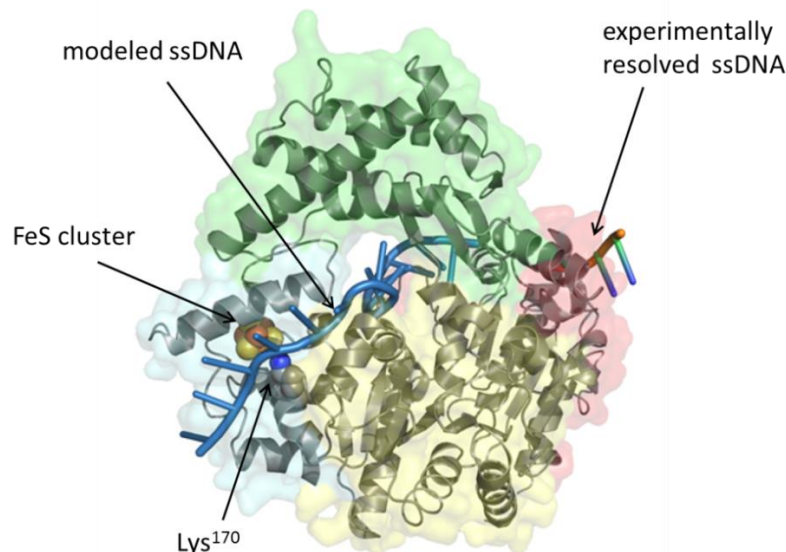


Figure 30: Model of the XPD-DNA complex.

The co-crystal structure of taXPD and a short stretch of ssDNA (experimentally resolved DNA in orange) (Kuper et al., 2012) was used to build a model of the path of ssDNA (modelled DNA in blue) through the enzyme. Color coding of the domains is similar to Figure 7 with the helicase domains 1 and 2 depicted in yellow and red, the FeS domain in cyan and the arch domain in green. The first three domains (helicase domain 1, FeS cluster and the arch domain) form a pore through which the ssDNA is threaded. The position of Lys170 (blue ball) mutated in the taXPD variant K170A and the FeS cluster (Fe atoms in red and S atoms in yellow) were depicted in a space-filling representation. Adapted from Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

Interestingly, however, the different strand selectivity observed for XPD from *T. acidophilum* is in contrast with a recent study performed by H.P. Naegeli's group with archaeal XPD from *Ferroplasma acidarmanus* (facXPD) (Mathieu et al., 2010). Based on electrophoretic competition analyses, facXPD recognized CPD both on the translocated and on the non-translocated DNA strand. Furthermore, facXPD protected a CPD lesion on the translocated strand from incision by the glycosylase T4 Endo V suggesting stable complex formation of XPD with CPD. Differences between the two archaeal species and the different experimental approaches (including the use of long DNA substrates in my AFM studies) may contribute to the different strand selectivity observed in the present study and the work of Naegeli (Mathieu et al., 2010, Mathieu et al., 2013). Although my data demonstrated that XPD preferentially recognizes CPD on the non-translocated strand, they do not exclude complex formation with CPD on the translocated strand.

5.1.5 Stable complex formation at the lesion involves ATP binding

AFM-based measurements of DNA bend angles induced at the site of the bound protein can reveal information on conformational properties of protein-DNA complexes. To characterize potentially different XPD-DNA conformations upon damage verification, DNA bend angles of nonspecific XPD complexes with undamaged DNA and specific complexes at the DNA target sites were analyzed

separately under different conditions and with different DNA substrates (Figure 31 and Table 18). Based on the position distributions, XPD-DNA complexes can be distinguished into specific complexes (bound at positions within two standard deviations from the center of the Gaussian fit) and nonspecific complexes (bound elsewhere on DNA). In addition to those measurements of protein-induced bend angles, intrinsic DNA bend angles were determined for nonspecific and specific DNA sites (at the fluorescein and at the CPD damage within the context of a bubble structure) as described above (see 5.1.4 and Figure 29). The DNA bend angle distributions in the absence of protein displayed narrow half Gaussian distributions (maximum at $\sim 0^\circ$) for both nonspecific and specific sites, which suggests that neither the presence of a bubble nor the presence of a damage induces innate DNA bending into the DNA substrates used for AFM.

In the absence of ATP, XPD showed a DNA bend angle of $\sim 50^\circ$ for both nonspecific and specific complexes at the fluorescein damage in the context of a bubble (Figure 31A, grey and black bars and Table 18, 2). However, in the presence of ATP, XPD showed a DNA bend angle of $\sim 50^\circ$ for nonspecific complexes (Figure 31B, grey bars and Table 18,2) and a bend angle of $\sim 65^\circ$ for specific complexes (Figure 31B, black bars and Table 18, 2). The observed transition in DNA bend angles from $\sim 50^\circ$ to $\sim 65^\circ$ at the site of the bound protein is highly significant ($P = 1.6 \cdot 10^{-12}$) and suggests that XPD may undergo ATP-dependent conformational changes at the target site. To identify whether the observed shift in DNA bend angles was due to ATP binding or ATP hydrolysis, XPD was incubated with the non-hydrolysable (or only very slowly hydrolysable) ATP-analog ATPys. Notably, the same bend angle change of $\sim 15^\circ$ between nonspecific and specific complexes could also be observed in the presence of ATPys (Figure 31C and Table 18, 2), indicating that ATP binding, not ATP hydrolysis, is sufficient to induce the shift in DNA bend angles.

The conditions leading to the potential conformational changes visible as a transition in DNA bend angles were further characterized by DNA bend angle analysis of XPD-DNA complexes on two DNA control substrates, fluorescein/- (Figure 31E and Table 18, 5) and -/bubble (Figure 31D and Table 18, 1). For both DNA substrates, the bend angle distributions for specific target sites (binding at the bubble position or at the damage) and nonspecific sites (binding at homoduplex DNA sites) were similar ($\sim 50^\circ$). The lack of the bend angle shift at the target sites is consistent with the protein position analysis on the same two DNA substrates (see 5.1.2 and Figure 25), which revealed a low localization specificity of taXPD to both target sites fluorescein/- and -/bubble compared to fluorescein/bubble. By excluding that the bend angle shift occurs in the two controls, the data indicate that XPD requires the combination of a bubble to support loading on DNA and the presence of a damage for stimulation of damage-specific conformational changes visible as a transition in bend angles.

Interestingly, a bend angle shift of 15° was noticed for stalled XPD-DNA complexes at the target site CPD/bubble (Figure 31F and Table 18, 6) similar as observed for the fluorescein/bubble target structure. This finding suggests a common verification mechanism for both lesions: Once a lesion has been recognized, the subsequent conformational changes are comparable, irrespective of the lesion type and the employed lesion detection strategy. The structural rearrangements might be a signal for the recruitment of further NER proteins including the endonucleases XPG and XPF.

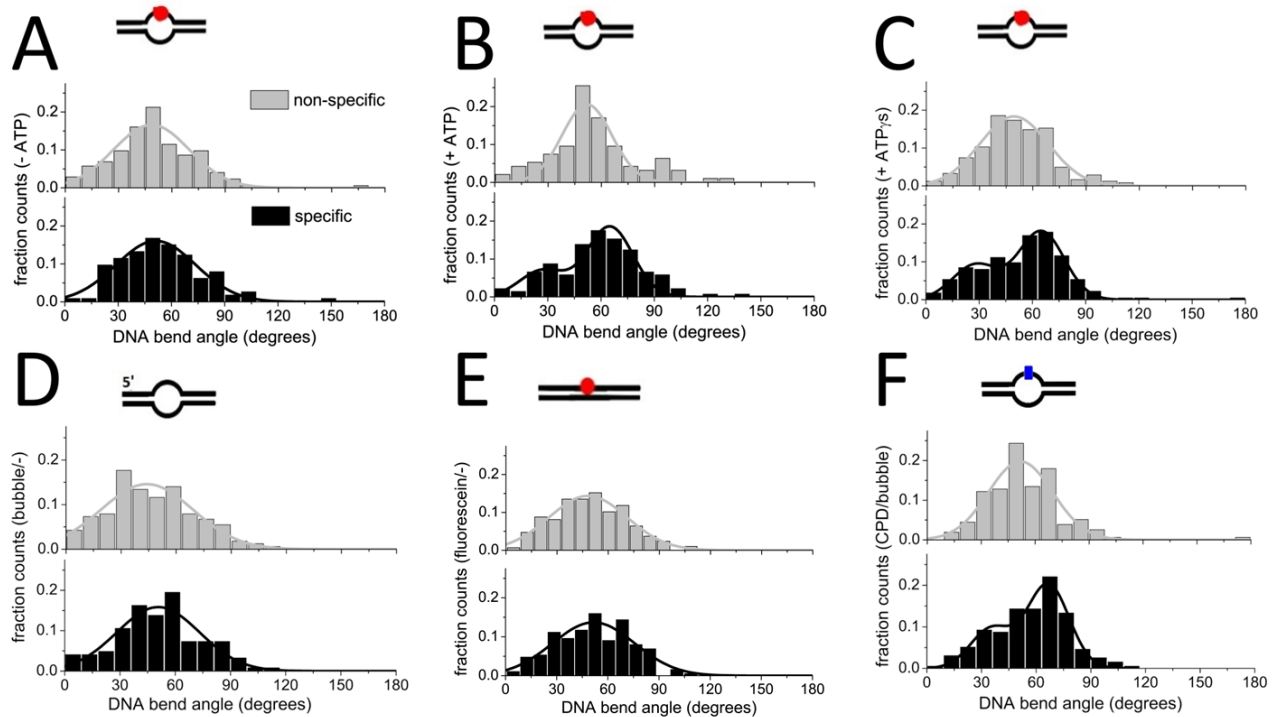











Figure 31: DNA bend angle distributions of taXPD with various DNA substrates.

A) In the absence of ATP, DNA bend angle distributions of XPD-DNA complexes are similar at nonspecific DNA sites (grey bars, $n = 174$) and specific DNA sites (black bars, $n = 113$). A Gaussian fit to the data centers at $\sim 50^\circ$ for both distributions. B) In the presence of ATP, XPD shows a DNA bend angle of $\sim 50^\circ$ for nonspecific complexes ($n = 166$) and a bend angle of $\sim 65^\circ$ for the specific complexes ($n = 119$). C) In the presence of the ATP-analog ATP γ s, bend angle distributions are different for nonspecific ($n = 242$) and specific complexes ($n = 224$). D)-F) AFM experiments with different DNA substrates in the presence of ATP. D) DNA bending for XPD complexes on DNA substrates containing a bubble without a lesion (-/bubble, Table 16, 1). Both distributions for specific ($n = 123$ complexes) and nonspecific sites ($n = 164$ complexes) are similar ($\sim 50^\circ$). E) DNA bending for XPD complexes on DNA substrates containing fluorescein within homoduplex DNA (fluorescein/-). Both distributions for specific ($n = 188$ complexes) and nonspecific sites ($n = 296$ complexes) are similar ($\sim 50^\circ$). F) A CPD damage within a bubble (Table 16, 6, ($n = 156$ for nsp sites and $n = 195$ for spec sites)) induced the same bend angle shift as observed for the fluorescein/bubble substrate. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

Table 18: DNA bend angle analysis of taXPD-DNA complexes under different conditions in AFM.

Bend angle (degrees)		+ ATP			+ ATPys		- ATP	
DNA substrate		nsp	spec	P-value	nsp	spec	nsp	spec
- / bubble (8nt)	 1	44 ± 25 n = 164	51 ± 22 n = 123	n.d.	n.d.	n.d.	47 ± 19 n = 246	48 ± 22 n = 196
F / bubble (8nt)	 2	52 ± 14 n = 94	27 ± 13 (27 %) 65 ± 13 (73 %) n = 137	1.6 x 10 ⁻¹²	49 ± 19 n = 242	28 ± 13 (26 %) 65 ± 13 (72 %) n = 224	49 ± 21 n = 174	51 ± 23 n = 113
F / bubble (14nt)	 2b	50 ± 13 n = 212	40 ± 9 (36 %) 66 ± 9 (64 %) n = 223	4.3 x 10 ⁻⁵¹	n.d.	n.d.	n.d.	n.d.
F / 5' bubble (8nt)	 3	53 ± 17 n = 184	33 ± 10 (37 %) 63 ± 10 (66 %) n = 188	9.8 x 10 ⁻¹²	51 ± 17 n = 201	27 ± 12 (35 %) 63 ± 12 (65 %) n = 166	n.d.	n.d.
F / 3' bubble (8nt)	 4	47 ± 19 n = 253	32 ± 13 (38 %) 68 ± 13 (62 %) n = 118	4.3 x 10 ⁻³¹	n.d.	n.d.	n.d.	n.d.
F / -	 5	48 ± 23 n = 296	52 ± 24 n = 188	n.d.	n.d.	n.d.	n.d.	n.d.
CPD / bubble (8nt)	 6	52 ± 18 n = 156	36 ± 12 (31 %) 66 ± 12 (69 %) n = 195	7.4 x 10 ⁻¹⁸	48 ± 19 n = 103	30 ± 12 (27 %) 65 ± 12 (73 %) n = 103	49 ± 20 n = 195	44 ± 22 n = 97
CPD / 5' bubble (8nt)	 7	52 ± 23 n = 313	26 ± 10 (34 %) 60 ± 10 (66 %) n = 212	7.4 x 10 ⁻⁸	n.d.	n.d.	n.d.	n.d.
CPD / 3' bubble (8nt)	 8	49 ± 18 n = 184	32 ± 11 (27 %) 66 ± 11 (73 %) n = 150	1.4 x 10 ⁻²³	n.d.	n.d.	n.d.	n.d.

DNA bend angles of XPD-DNA were measured at lesion-specific (spec) and nonspecific, undamaged DNA sites (nsp) for different DNA substrates (as schematically indicated) either in the presence or absence of ATP or ATPys. The values given in the table represent the maxima of single or double Gaussian fits to pooled bend angle distributions from at least 2 experiments. Sigma, the half-width of the Gaussian peak is given as the error. The total number of data points is indicated as n. The ratio between the bend angle states is given in brackets as % of total area derived from integration of the Gaussian peaks using the software Origin Pro. The significances of the shift in average bend angle between nonspecific and specific complexes (in the presence of ATP) were calculated by on-tailed student t-test. Significances of shifts in bend angles between the absence and presence of ATP (or ATPys) for specific complexes were $P = 2.2 \times 10^{-8}$ (4.8×10^{-9}) and 9.5×10^{-20} (5.3×10^{-16}) for DNA substrates 2 (F = fluorescein) and 6 (CPD), respectively. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

In addition to the DNA bend angle conformation with a maximum value of ~ 65°, a second DNA bend angle state of ~ 30° was discovered in all distributions of complexes at specific sites in the presence of

ATP and ATPys, when the distributions were fitted with a double Gaussian curve. Integration of the relative surface areas of the double Gaussians revealed that the predominant conformation is the larger bend angle state of $\sim 65^\circ$ ($\sim 70\%$ of all complexes), while the smaller bend angle of $\sim 30^\circ$ is populated to a lower degree ($\sim 30\%$ of all complexes). This slightly bent state may also be present in the broad distributions at nonspecific sites but may not be resolved by a double Gaussian fit because the centers of the Gaussians ($\sim 30^\circ$ and $\sim 50^\circ$) are too close to each other. A population of the $\sim 30^\circ$ bend angle is not detected in the distributions determined from intrinsic bend angles on nonspecific and specific DNA substrates in the absence of protein (Figure 29), which suggests that this state is not related to a pre-bent innate DNA structure XPD binds to. It could be envisioned that this state with a bend angle of $\sim 30^\circ$ is protein-induced and may reflect a conformation of XPD, which probes DNA upon initial encounter on its way to a specific damage verification complex. However, the transition in bend angles from $\sim 50^\circ$ to $\sim 65^\circ$ between nonspecific and specific is still considered as the relevant step for the lesion-dependent conformational changes.

DNA bend angle measurements of taXPD-K170A complexes at undamaged DNA and at lesion sites (for fluorescein and CPD, respectively) resulted in similar distributions for nonspecific and specific complexes (Figure 32, A and B) without the DNA bend angle shift of $\sim 15^\circ$ observed for the wt protein. The lack of lesion-dependent conformational changes confirms that this taXPD variant is incapable of damage recognition, which is consistent with the position distribution of K170A showing only residual localization specificity to a fluorescein lesion (Figure 27). The deficiency in damage recognition by taXPD K170A is most likely due to its elevated helicase activity, which confirms that stalling of the helicase activity is implicated in damage verification. In addition, the AFM data on K170A support the hypothesis that the FeS cluster may be involved in damage sensing using DNA-mediated CT (Figure 8) as the mutated Lysine residue is located in close proximity of the FeS cluster (Figure 30).

Consistent with the finding that the structural changes leading to damage verification by XPD wt require ATP-binding (Figure 31B and C), the taXPD variant R567W showed comparable bend angle distributions for nonspecific and fluorescein/bubble target sites without any bend angle shift (Figure 33) similar as observed for taXPD wt in the absence of ATP (Figure 31A). The absence of repair-competent conformational changes for R567W is in accordance with the position analysis data in Figure 26A, which indicate that this mutant is incapable of performing damage verification. This phenotype can be explained by a deficiency in ATP binding as it has been proposed here and in previous structural studies (Wolski et al., 2008), thus supporting the relevance of ATP binding for lesion-dependent conformational changes.

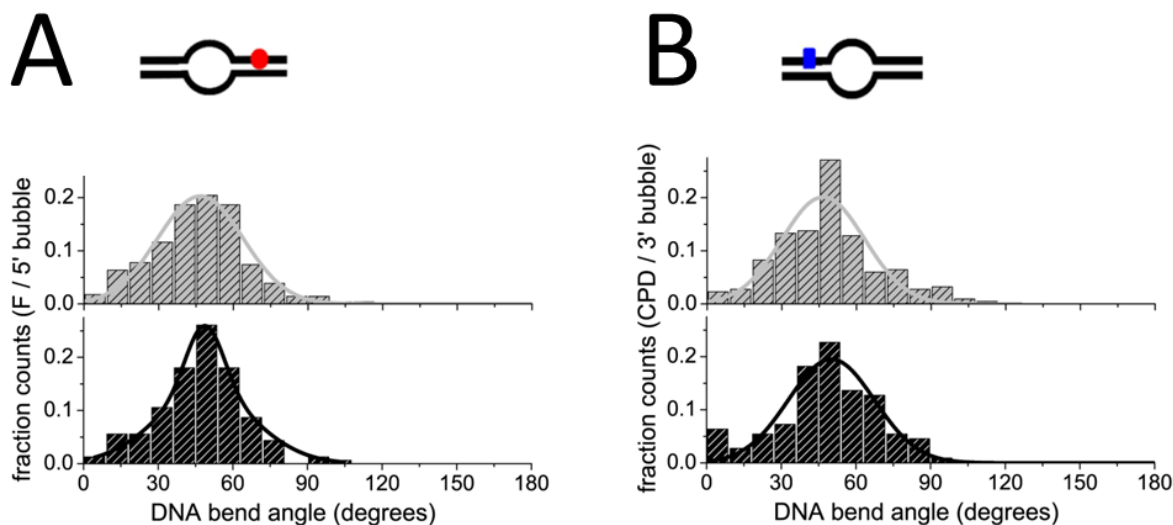


Figure 32: AFM experiments with taXPD variant K170A.

DNA bend angle measurements of taXPD-K170A complexes on both A) the DNA substrate fluorescein/ 5' bubble ($n_{\text{spec}} = 161$ $n_{\text{nsp}} = 284$) and B) on the DNA substrates CPD/ 3' bubble ($n_{\text{spec}} = 110$ $n_{\text{nsp}} = 218$) resulted in similar distributions for nonspecific complexes bound to non-damaged DNA sites (grey bars) and specific complexes bound to lesion sites (black bars). The missing DNA bend angle shift of $\sim 15^\circ$ observed for the wt protein suggests that K170A is deficient in lesion verification. All distributions represent pooled data n from three individual experiments. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

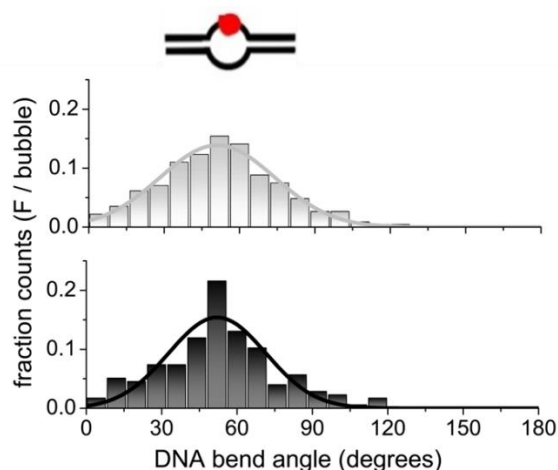


Figure 33: DNA bend angle distribution of the taXPD variant R567W.

Bend angle distributions for R567W in the presence of ATP are similar for nonspecific ($n = 227$) and specific complexes ($n = 176$) and are comparable to the bend angle distribution at nonspecific DNA sites for the wt protein. The distribution represents pooled data from five individual AFM experiments.

In principle, the requirement of ATP binding for the conformational changes closely resembles the prokaryotic NER pathway. Here, initial conformational changes of the prokaryotic XPD homolog UvrB via ATP hydrolysis lead to recruitment of UvrB to the lesion site. After successful damage verification, UvrB is converted to a pre-incision complex by ATP re-binding (Verhoeven et al., 2001) and prepared for the arrival of the endonuclease UvrC (Verhoeven et al., 2002). This similarity between the eukaryotic and the

prokaryotic NER mechanism supports the high level of conservation of NER across different species (Petit and Sancar, 1999, Batty and Wood, 2000, Kuper and Kisker, 2012).

5.1.6 Model of damage verification by taXPD

The model of XPD damage verification is based on the results of AFM and BLI experiments. XPD can load at bubble structures as a nonspecific complex (green circles, DNA bend angle of $\sim 50^\circ$) to translocate on DNA in 5'-3' direction (Figure 34). Upon damage verification, the nonspecific XPD-DNA complex is converted to a specific complex (red circles, DNA bend angle of $\sim 65^\circ$). Depending on the type of DNA lesion, fluorescein or CPD, two different scenarios can be distinguished. I) Movement of nonspecific XPD on DNA (green circle) is stalled when a bulky fluorescein damage (red ball) is located on the translocated DNA strand (5'-3') (5' bubble). The bulky damage may hence serve as a mechanical road block obstructing the pore for passage of ssDNA visualized in the crystal structure in Figure 30. XPD translocation is not blocked when fluorescein is located on the non-translocated DNA strand (3' bubble). II) In contrast, an intrastrand DNA crosslink (CPD damage, blue rectangle) is not recognized on the translocated DNA strand, but on the non-translocated DNA strand. It can be envisioned that taXPD employs a different mechanism for detecting CPD possibly involving the FeS cluster domain. Depending on the lesion type XPD may simultaneously use the two damage sensing mechanisms during damage search. However, once taXPD has verified a damage, either CPD or fluorescein, DNA bend angle analyses from AFM measurements indicate that the protein undergoes the same structural rearrangements, which may lead to recruitment of further NER factors (Buechner et al., 2014).

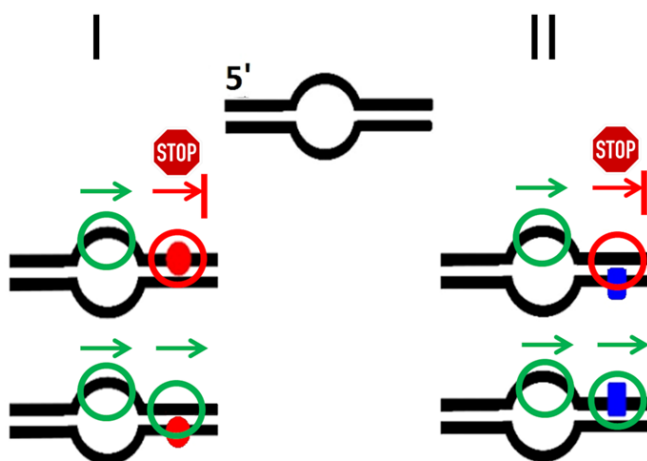


Figure 34: Model for lesion recognition by taXPD.

The green circle indicates taXPD moving along DNA. A bulky fluorescein damage (red ball) is preferentially recognized on the translocated strand, whereas a CPD damage (blue rectangle) is preferentially detected on the non-translocated strand. For details see main text. This research was originally published in Buechner et al. (2014) © the American Society for Biochemistry and Molecular Biology.

5.2 Lesion recognition strategies of DNA glycosylases in base excision repair

The ability of DNA glycosylases to detect and remove base lesions rapidly and efficiently from the genome is essential in base excision repair (BER). How can DNA glycosylases go about finding the rare target sites among the huge excess of undamaged bases? Here, this key question in BER was addressed by using single molecule AFM imaging supported by a 2-aminopurine fluorescence base flipping assay to study damage search and recognition by hTDG, which initiates BER of mutagenic and cytotoxic G:T and G:U wobble pairs. The findings for hTDG are summarized as a model in the context of general implications for DNA lesion search by DNA glycosylases (5.2.6). Similar studies were performed with hOgg1, a second representative of the glycosylase family targeting structurally different 8oxoG:C pairs, to test the generalizability of the proposed model.

5.2.1 hTDG recognizes G:U and G:T mismatch sites with moderate specificity

For the AFM studies, a single target site for hTDG, either a G:T or a G:U base mismatch in a CpG context, was introduced into long DNA fragments (549 bp) at 46% of DNA fragment length. Notably, the long DNA fragments used for the AFM experiments represent better mimics of the naturally occurring DNA substrates than the short oligonucleotides typically used in biochemical ensemble methods (Buechner and Tessmer, 2013, Buechner et al., 2014). To study specific DNA interactions of hTDG in its lesion recognition state, my AFM studies focused on the catalytically inactive hTDG variant N140A, which has been shown to possess similar DNA binding affinity as the wt enzyme, but hardly detectable base excision activity (Hardeland et al., 2000, Maiti et al., 2009). Figure 35A shows a representative AFM image of the hTDG variant N140A incubated with DNA substrates containing a single G:U mismatch (compare

for DNA sequence). Position analyses of specific hTDG complexes at mismatches (Figure 35B and D, black bars) can be fitted with a Gaussian curve (grey line) with maximum at ~ 45 % of DNA full length consistent with the position of the mismatch. Enhanced binding at the mismatch position hence indicates a preference of TDG for this target site binding over binding to nonspecific homoduplex DNA. From the Gaussian fits, moderate specificities $S = 163 \pm 47$ of hTDG-N140A for the G:U and $S = 305 \pm 95$ for the G:T mismatch were calculated (see Table 19 and Methods 4.6.1).

Due to local helix destabilization at DNA fragment ends, hTDG also displays a strong DNA end binding preference (compare bar at 0 %, with ~ 0.18 fraction of counts with bar at 45 %, with ~ 0.13 fraction of counts in Figure 35B) as has previously been observed for many other DNA repair proteins (Tessmer et al., 2008, Wagner et al., 2009, Tessmer et al., 2012, Buechner et al., 2014).

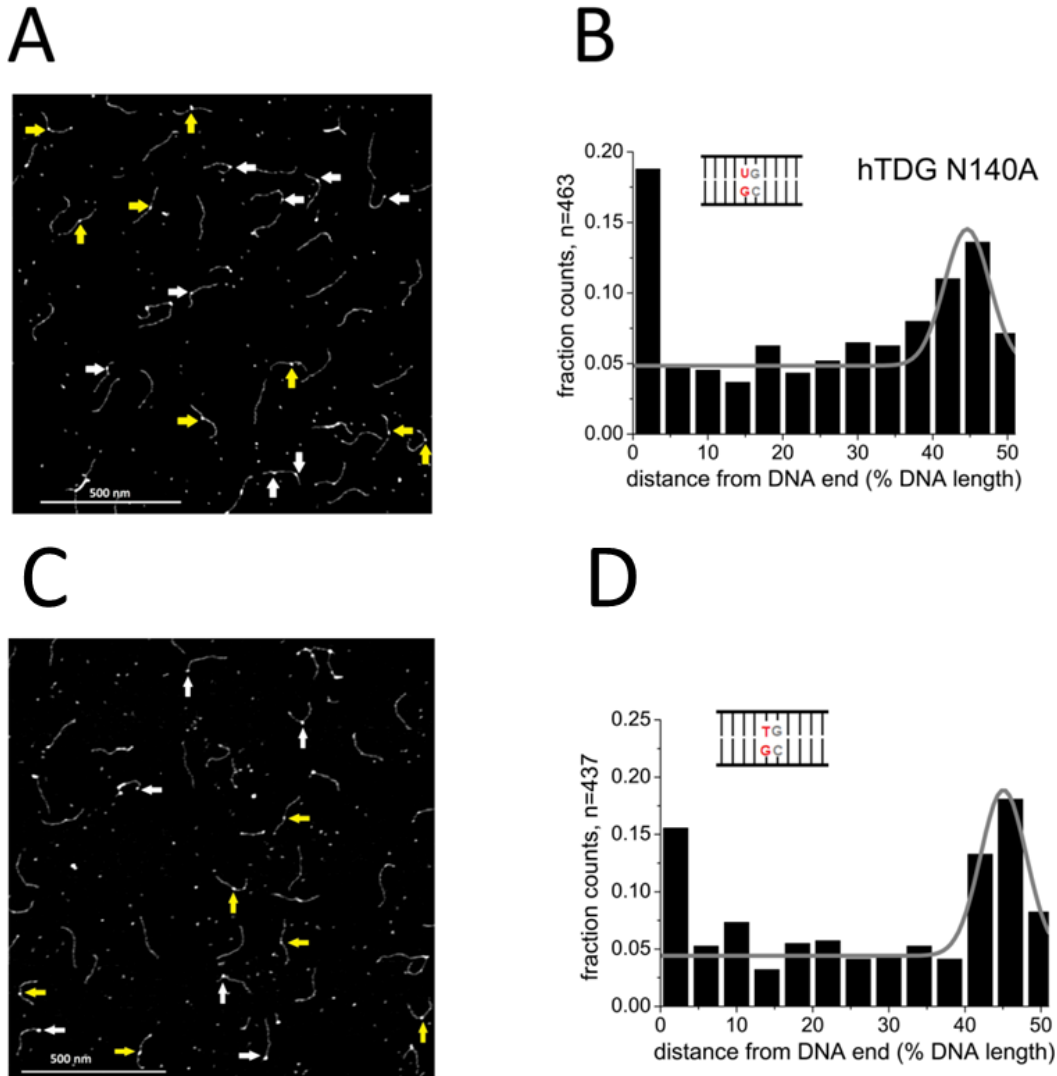


Figure 35: Lesion specificity of hTDG N140A at G:U and G:T mismatches

A,C. AFM images of hTDG N140A binding to DNA substrates (549 bp) containing a A) G:U mismatch and a C) G:T mismatch in a CpG context at 46 % of DNA fragment length. Arrows point to nonspecific (white) and specific hTDG N140A-DNA complexes bound to the target site (yellow). B,D. hTDG N140A binding position distribution on the DNA demonstrates a moderate binding preference for B) the G:U and D) the G:T mismatch (enhanced occupancy at ~ 46 % DNA length) over nonspecific DNA. Fractional occupancies are plotted for ~ 22 bp long sections of 549 bp DNA from DNA fragment ends (0 %) to DNA center (50 %). Results were pooled from three individual experiments (n = total data points).

Table 19: AFM binding specificities of hTDG variants to different DNA substrates.

Protein	DNA substrate	AFM binding specificity	n
hTDG N140A	dsDNA G:U	163 ± 47	371
hTDG N140A	dsDNA G:T	305 ± 95	364
hTDG R275A	dsDNA G:U ^F	358 ± 101	375

AFM binding specificities of hTDG wt and variants for different DNA substrates were determined as means ± SD from three position distribution histograms derived from three individual experiments. The number of datapoints n is indicated.

The moderate localization specificity of TDG to its target sites observed from the AFM experiments ($S \sim 150\text{-}300$, see Table 19) is likely due to nonspecific background binding at the applied protein concentration (50 nM). In addition to the specific target site located within hundreds of nonspecific bases, the DNA substrates used for AFM also contain several additional CpG sites (14 % of total bases) spread over the entire DNA sequence. hTDG has been reported to bind to CpG sites in DNA with only 4-fold lower affinity and to random homoduplex DNA sequences with only 16-fold reduced affinity compared to its G:T target sites (Morgan et al., 2011). Importantly, however, the AFM data show a distinct preference of TDG for its target lesions (G:T, G:U) over an excess of undamaged DNA.

5.2.2 DNA bending and base flipping by hTDG upon lesion encounter

Specific and nonspecific hTDG-DNA complexes can be distinguished based on their position on the DNA substrates (specific site at 46 % of the DNA length) and were treated separately for further DNA bend angle analyses (see 4.1.3).

Measurements of DNA bend angles for hTDG N140A bound to G:U (lesion encounter complexes, Figure 36A and Table 20) resulted in a distribution with two distinct populations. A double Gaussian fit to the distribution centered at $(29 \pm 10)^\circ$ and $(68 \pm 10)^\circ$ ($R^2 = 0.88$). Integration of the peak areas revealed that 33 % of the complexes are in the less bent state and 67 % are in the stronger bent state.

Interestingly, the less bent state with bend angle of $\sim 30^\circ$ is also clearly found at the mismatch site in the absence of protein (36 % of complexes) in addition to the straight conformation (semi-Gaussian peak at $\sim 0^\circ$) (Figure 36B), which is typically observed at homoduplex sites (Wang et al., 2003). This finding suggests that the G:U mismatch induces an intrinsic bend of $\sim 30^\circ$ in the DNA backbone (under the conditions of these AFM studies), which also prevails in part of the specific, mismatch-bound complexes in equilibrium with the stronger bent state (bend angle of $\sim 70^\circ$).

The distributions of DNA bend angles induced by hTDG-N140A at G:T sites (Figure 36C) also shows the two populations of $\sim 30^\circ$ and $\sim 70^\circ$ as observed for hTDG-N140A and G:U and were fitted by a double Gaussian. Although the two bend angle states induced by hTDG at G:T sites (Figure 36C) appear to be not as distinct as those observed for hTDG at G:U sites (Figure 36A), the ratios between the smaller and the larger bend angle are comparable ($(26 \pm 15)^\circ$ (42 %) and $(65 \pm 15)^\circ$ (58 %) for hTDG on G:T sites vs. $(29 \pm 10)^\circ$ (33 %) and $(68 \pm 10)^\circ$ (67 %) for hTDG on G:U sites, $P = 0.14$, not significantly different). Measurements of intrinsic DNA bend angles at the specific site in the absence of protein revealed intrinsic bending of the DNA at the site of the G:T mismatch by $(31 \pm 9)^\circ$ (32 %) similar as observed for the G:U mismatch (in addition to the straight conformation with bend angle of $\sim 0^\circ$ (68 %)).

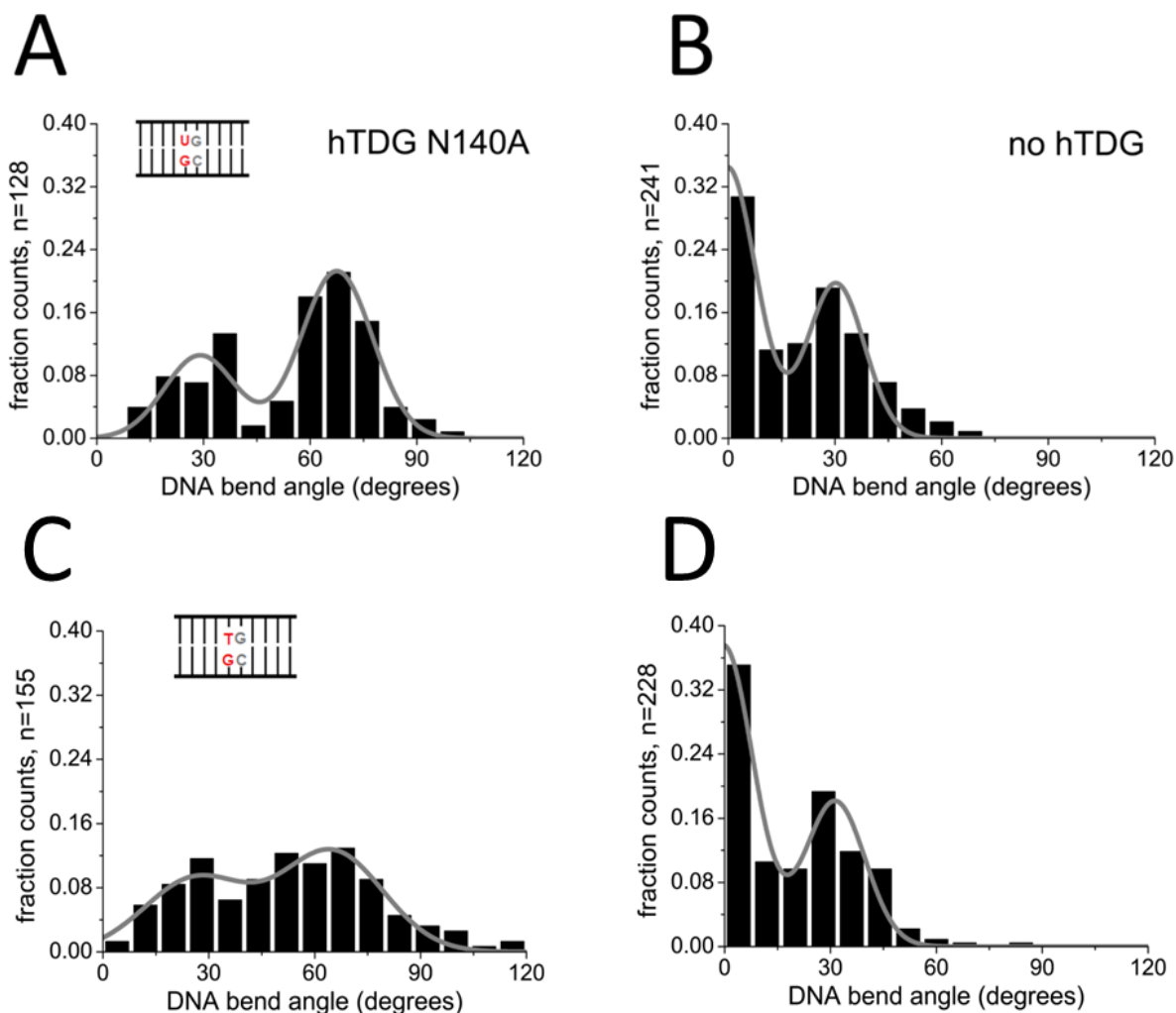


Figure 36: DNA bending by hTDG N140A at G:U and G:T mismatches

A) The distributions of DNA bend angles induced by hTDG-N140A at G:U sites shows two distinct populations. A double Gaussian fit ($R^2 = 0.88$) to the data centers at $(29 \pm 10)^\circ$ (33 %) and $(68 \pm 10)^\circ$ (67 %). B) Measurements of intrinsic DNA bend angles at the site of the G:U mismatch in the absence of protein reveal a similar less bent state with bend angle of $\sim 30^\circ$ (36 %) as observed in A) in the presence of protein. However, the predominant state (64 %) is the straight conformation (semi-Gaussian peak at $\sim 0^\circ$). C) Distribution of DNA bend angles induced by hTDG-N140A at G:T sites. A double Gaussian fit ($R^2 = 0.87$) centered at $(26 \pm 15)^\circ$ (42 %) and $(65 \pm 15)^\circ$ (58 %). Distribution of intrinsic DNA bend angles at the G:T mismatch in the absence of protein revealed similar DNA bending by $(31 \pm 9)^\circ$ (32 %) as in (B) in addition to a predominant straight DNA state ($0 \pm 9)^\circ$ (68 %). The ratio between the two states was derived from the areas of the Gaussian peaks. The bend angle distributions in A) and C) were pooled from three individual experiments and those in B) and D) from two individual AFM experiments. The total number of data points is indicated as n.

Furthermore, DNA binding and bending was studied by the hTDG variant R275A (Figure 37). Crystal structures of DNA-bound hTDG (to the abasic site analog THF and to the non-cleavable G:U analogue G:U^F) revealed that the strictly conserved residue Arg²⁷⁵ is located at the tip of an arginine finger (Maiti et al., 2008, Maiti et al., 2012). Based on these structural findings and further pre-steady-state kinetics experiments, the arginine finger had been implicated in promotion and/or stabilization of nucleotide

flipping in TDG by penetrating the DNA minor groove and filling the void created by base extrusion (Maiti et al., 2009). Complex position analysis from the AFM images revealed a lesion specificity of hTDG-R275A for G:U^F comparable to (or even slightly higher than) the specificity of the catalytically inactive hTDG variant N140A for G:U ($S = 358 \pm 101$ and $S = 163 \pm 47$, respectively, Figure 37B and Figure 35B, Table 19).

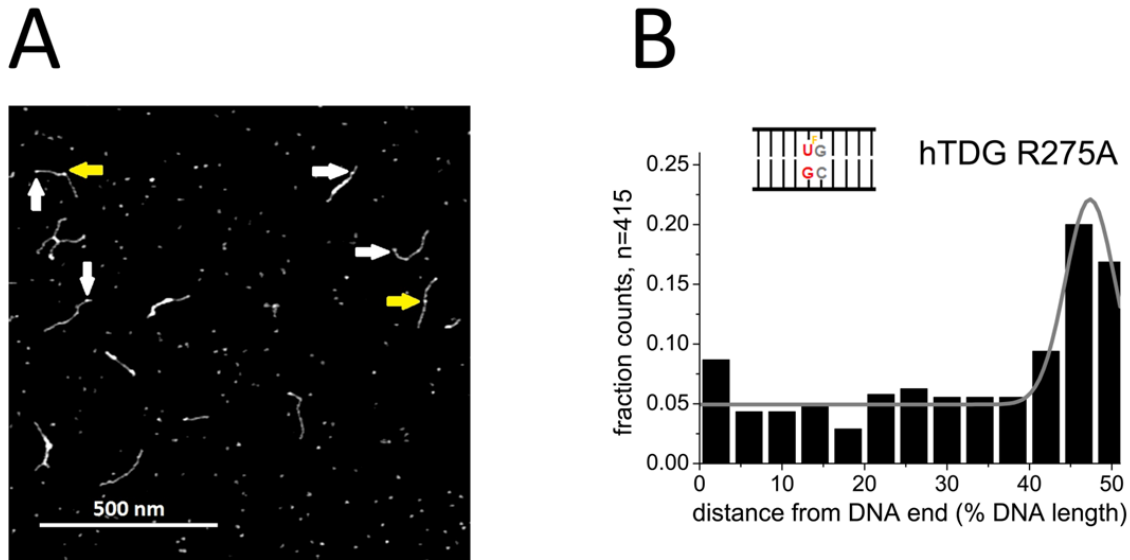


Figure 37: Lesion specificity of hTDG R275A at G:U^F mismatches.

A) AFM image showing hTDG-R275A incubated with DNA substrates (549 bp) containing a non-cleavable G:U^F mismatch in a CpG context at 46 % of DNA fragment length. Yellow arrows point to complexes specifically bound at the target site, white arrows indicate nonspecific hTDG-DNA complexes bound to non-damaged DNA sites including DNA ends. B) hTDG R275A binding position distributions on G:U^F containing DNA substrates revealed a lesion specificity comparable to the specificity of hTDG N140A for G:U. The position distribution in B) is pooled from three individual experiments. The total number of data points is indicated as n.

AFM experiments in the absence of protein showed intrinsic bending of the DNA at the site of the G:U^F mismatch by $(33 \pm 12)^\circ$ (Figure 38B), similar as observed for the G:U mismatch (see above and Table 20). Gaussian fits to the DNA bend angle distribution from AFM images of hTDG-R275A lesion encounter complexes bound to the G:U^F target site (Figure 38A, Table 20) showed an equilibrium between three states. In striking contrast to N140A, the percentage of kinked protein complexes (with a bend angle of $\sim 70^\circ$) was significantly reduced for R275A (67 % for N140A at G:U versus 15 % for R275A at G:U^F, respectively, $P = 0.0005$, Figure 38A and Figure 36A and Table 20).

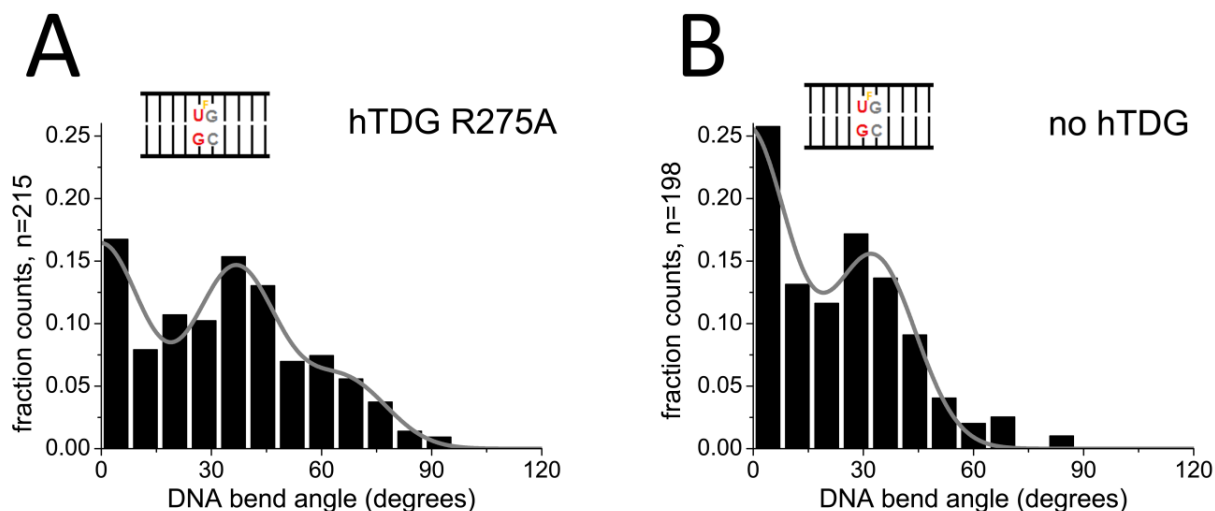


Figure 38: DNA bending of hTDG R275A at G:U^F mismatches

A) In contrast to N140A bound to G:U, the distribution of DNA bend angles induced by hTDG-R275A at G:U^F sites showed an equilibrium between three states. A triple Gaussian fit ($R^2 = 0.89$) to the data centered at $(0 \pm 12)^\circ$ (45%), $(37 \pm 12)^\circ$ (40%) and $(66 \pm 12)^\circ$ (15%). B) Measurements of intrinsic DNA bend angles at the specific site in the absence of protein revealed intrinsic bending of the DNA at the site of the G:U^F mismatch by $(33 \pm 12)^\circ$ (37%) similar as observed for the G:U mismatch in Figure 36A. However, the predominant population (63%) is the straight conformation with bend angle of $\sim 0^\circ$. The bend angle distributions in A) were pooled from three individual experiments and those in B) from two different experiments. The total number of data points is indicated as n. The ratios between the different states were derived from the areas of the Gaussian peaks.

In addition, 2-Aminopurine (2-AP) was placed as a fluorescent probe next to G:U^F to investigate specific DNA target site interactions of hTDG in a fluorescence based base flipping assay. As described above (4.5.3) and in previous studies (Stivers, 1998, Bellamy and Baldwin, 2001, Bellamy et al., 2007), the fluorescence signal of 2-AP can be also used to monitor flipping of the uracil nucleotide from the DNA, when located next to 2-AP. The 2-AP fluorescence is highly quenched within the DNA duplex structure and is expected to increase upon destacking of the DNA helix by protein-induced destabilization or base flipping. In titrations with hTDG wt, the 2-AP fluorescence intensity increased with increasing [TDG], consistent with the expectation that uracil flips out of the duplex and into the TDG active site (Figure 39A and C). At the highest TDG concentration examined, we measured a significant increase in 2-AP fluorescence compared to the free DNA ($P = 0.032$). In contrast, binding of R275A-TDG to G:U^F DNA does not result in a substantial increase in 2-AP fluorescence ($P = 0.373$, Figure 39B) and induced significantly lower fluorescence than for the wt protein ($P = 0.047$, Figure 39A and C and Table 21). This result is in accordance with the decreased population of strongly bent R275A complexes observed in the AFM data.

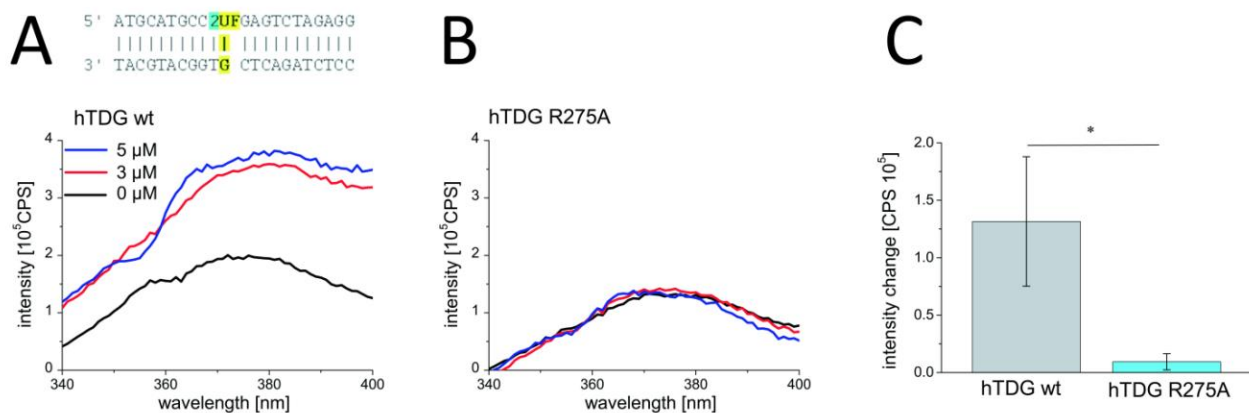


Figure 39: 2-AP fluorescence based base flipping assay of hTDG at G:U^F mismatches.

Increasing concentrations of A) hTDG wt and B) hTDG R275A were titrated to 170 nM 2-AP DNA substrates. Steady-state fluorescence emission spectra (340 – 400 nm) were recorded at an excitation wavelength of $\lambda_{\text{ex}} = 320$ nm. The inset in A) shows a schematic of the 2-AP (cyan) DNA substrate containing G:U^F (yellow). Fluorescence spectra were corrected for the fluorescence background of the reaction buffer and/or the protein C) Quantification of the 2-AP fluorescence intensity change for 5 μ M of hTDG wt (grey bar) and R275A (cyan bar). Results were derived from two individual titrations. The error bars indicate the standard deviation between the different experiments. Significances are calculated with a one-tailed student's t-test and are classed as * P < 0.05.

5.2.3 Nonspecific complexes of hTDG on homoduplex DNA

After characterizing specific TDG complexes bound at target sites of the glycosylase, hTDG was incubated with undamaged DNA fragments to analyze nonspecific, lesion search complexes of the DNA repair protein (Figure 40A-C and Table 20). The catalytically inactive hTDG variant N140A mutant showed comparable DNA bending behavior as the wt ((31 ± 12) (39 %) and $(65 \pm 12)^\circ$ (61 %) for hTDG N140A and $(33 \pm 11)^\circ$ (46 %) and $(68 \pm 11)^\circ$ (54 %) for hTDG wt, P = 0.478 using a one tailed student t-test, Figure 40A and B). The DNA bend angle distributions of these nonspecific complexes of hTDG-N140A (Figure 40A) revealed the same bend angle states as observed for hTDG-N140A bound specifically at a G:U mismatch (Figure 36A) with a Gaussian fit to the bend angle distribution centered at $\sim 30^\circ$ and $\sim 70^\circ$. DNA bend angle measurements for hTDG-R275A on nonspecific (non-damaged) DNA displayed a tri-phasic distribution as described above for specific, lesion-bound R275A, with a predominant peak at $\sim 0^\circ$, a large population at $\sim 30^\circ$, and an almost negligible population at $\sim 70^\circ$ (Figure 40C and Table 20). Interestingly, the percentage of complexes with an average bend angle of $\sim 30^\circ$ (36 % of complexes) was similar to that of hTDG-R275A lesion encounter complexes bound at a G:U mismatch (40 % with bend angle $\sim 30^\circ$, Table 20), although the non-damaged DNA itself showed no intrinsic bending in control experiments in the absence of protein (Figure 40D and Figure 41, see 4.1.3).

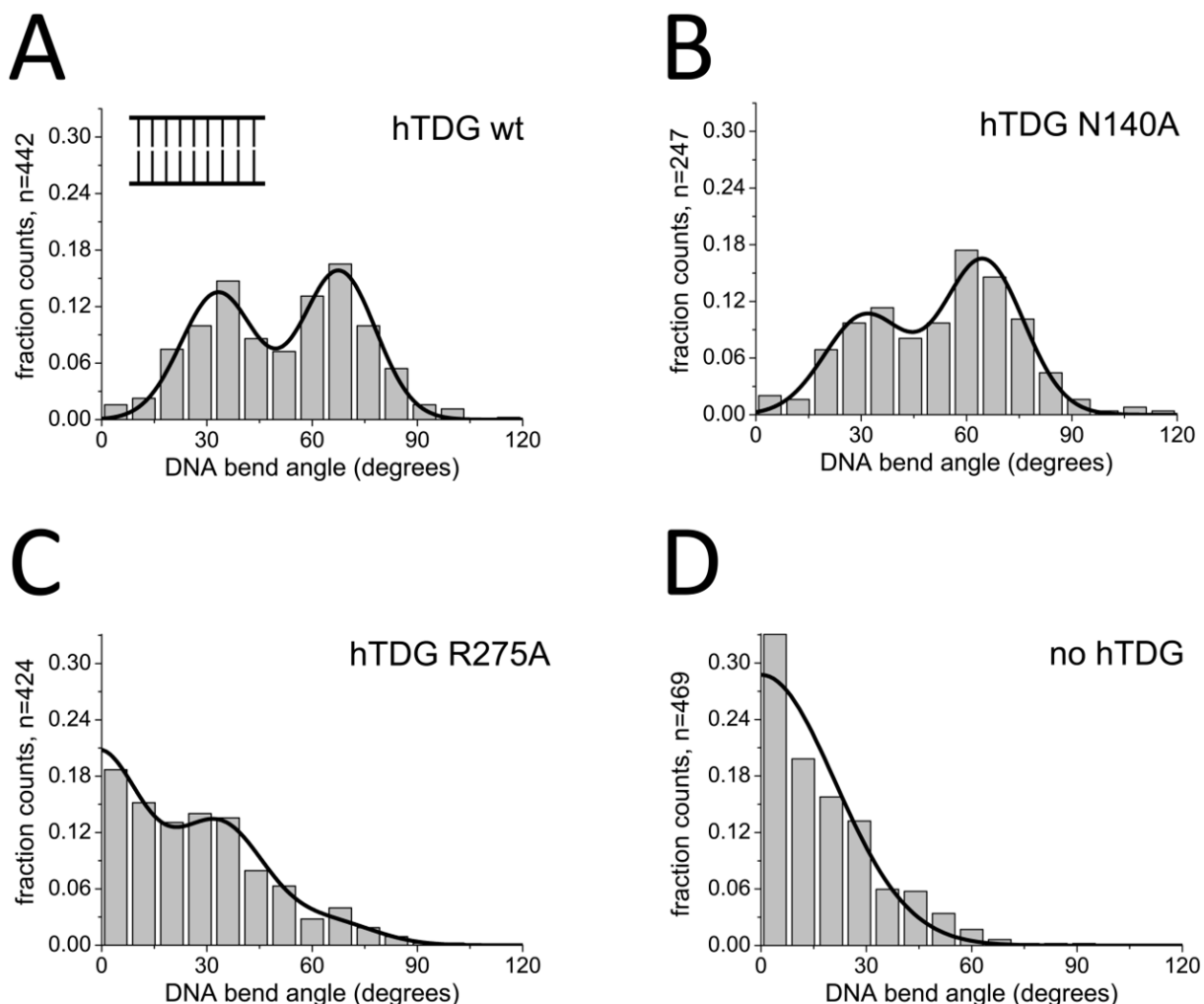


Figure 40: DNA bending of hTDG at nonspecific DNA sites.

A) The distributions of DNA bend angles induced by hTDG wt at nonspecific sites on 1813 bp DNA fragments shows two distinct populations. A double Gaussian fit ($R^2 = 0.96$) to the data centers at $(33 \pm 11)^\circ$ (46 %) and $(68 \pm 11)^\circ$ (54 %). B) DNA bend angle measurements from AFM experiments of hTDG N140A bound to nonspecific DNA fragments (549 bp) result in two distinct states. A double Gaussian fit ($R^2 = 0.94$) to the data centers at $(31 \pm 12)^\circ$ (39 %) and $(65 \pm 12)^\circ$ (61 %). C) DNA bend angle measurements from AFM experiments of hTDG R275A complexes bound to nonspecific DNA fragments (1813 bp) result in three different states. A triple Gaussian fit ($R^2 = 0.95$) to the data centers at $(-2 \pm 13)^\circ$ (57 %), $(34 \pm 13)^\circ$ (36 %) and $(65 \pm 13)^\circ$ (7 %). D) Intrinsic DNA bend angles measurements were pooled for two nonspecific DNA substrates (549 bp and 1813 bp). A semi-Gaussian fit to the data centers at $(0 \pm 21)^\circ$. The single distributions for the two DNA substrates (shown in Figure 41) are very comparable. All distributions represent pooled data from at least three individual experiments. The total number of data points is indicated as n.

Intrinsic bending was also measured for 549 bp undamaged DNA fragments (Figure 41B) at the same position as the target site mismatch (46 % from DNA fragment end, within identical sequence context) resulting in a similar distribution as observed for the longer (1813 bp) non-damaged DNA fragment (Figure 41D).

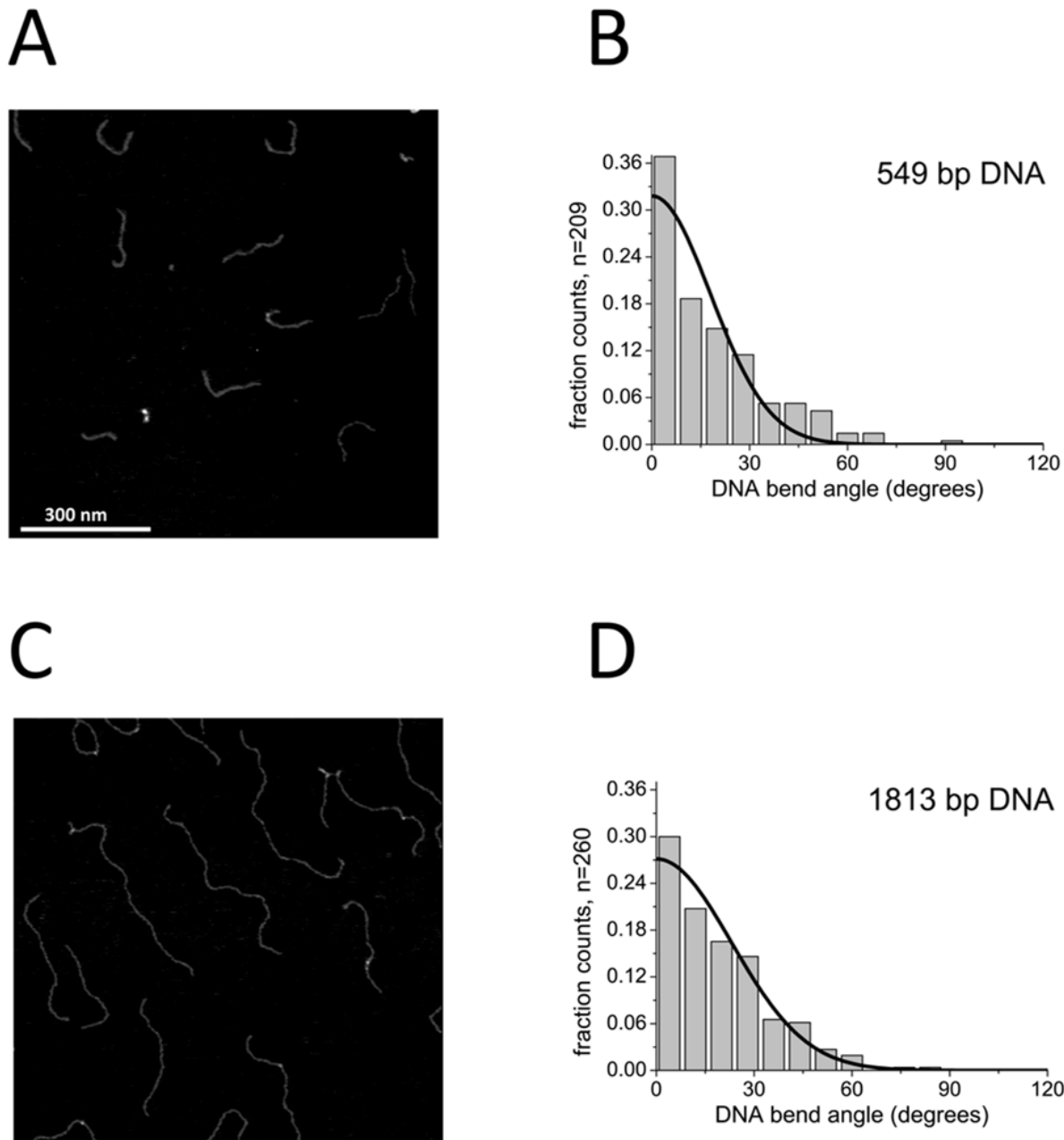


Figure 41: intrinsic bending of two nonspecific DNA substrates (549 bp and 1818 bp) in the absence of protein.

A, B) Using the same DNA substrate as in experiments with lesion containing DNA (549 bp), but without any base mismatches or lesions (nonspecific substrate), intrinsic bending was measured at 46 % of the DNA fragment length (target site position in the specific DNA substrates) in the absence of protein to exclude any influence of DNA sequence context on intrinsic DNA bending behavior. A semi-Gaussian fit centered at $(0 \pm 18)^\circ$ ($R^2 = 0.895$).

C,D) Intrinsic bending within a 1813 bp DNA fragment (that did not contain any base mismatches or lesions) was determined by moving a mask comparable to the size of TDG at regular intervals of 50 nm along the DNA. A semi-Gaussian fit centered at $(0 \pm 23)^\circ$ ($R^2 = 0.963$). Results were pooled from two individual experiments. The total number of data points is indicated as n. Both AFM images A) of the 549 bp fragments and C) of the 1813 bp fragments are $1 \mu\text{m} \times 1 \mu\text{m}$.

Table 20: DNA bend angle analysis for hTDG-DNA complexes and free DNA substrates.

protein	DNA substrate	straight	slightly bent	strongly bent	n
<u>protein induced bending (in degrees)</u>					
hTDG wt	Nsp DNA (1813 bp)	NA	33 ± 11 (46 %)	68 ± 11 (54 %)	442
hTDG N140A	G:U	NA	29 ± 10 (33 %)	68 ± 10 (67 %)	128
	G:T	NA	26 ± 15 (42 %)	65 ± 15 (58 %)	155
	Nsp DNA (549 bp)	NA	31 ± 12 (39 %)	65 ± 12 (61 %)	247
hTDG R275A	Nsp DNA (1813 bp)	-2 ± 13 (57 %)	34 ± 13 (36 %)	65 ± 13 (7 %)	424
	G:U ^F	0 ± 12 (45 %)	37 ± 12 (40 %)	66 ± 12 (15 %)	215
<u>intrinsic DNA bending (in degrees)</u>					
none	Nsp DNA (549 bp)	0 ± 18	NA	NA	209
none	Nsp DNA (1813 bp)	0 ± 23	NA	NA	260
none	Nsp DNA (549 bp and 1813 bp pooled)	0 ± 21	NA	NA	469
none	G:U mismatch	0 ± 8 (64 %)	30 ± 8 (36 %)	NA	241
none	G:U ^F mismatch	-2 ± 12 (63 %)	33 ± 12 (37 %)	NA	198
none	G:T mismatch	0 ± 9 (68 %)	31 ± 9 (32 %)	NA	228

Depending on DNA bending, ($\sim 0^\circ$, $\sim 30^\circ$ and $\sim 70^\circ$), complexes are classed as ‘straight’, ‘slightly bent’ or ‘strongly bent’. The given values represent the maxima of single, double or triple Gaussian fits to pooled bend angle distributions. The population of the different bend angle states is given in brackets as % of total area from integration of the Gaussian peaks. The total number of data points is indicated as n. NA = not applicable.

Results from 2-AP base flipping assays revealed a significant fluorescence shift for nonspecific 2-AP containing DNA in the presence of increasing concentrations of hTDG wt compared to the DNA control in the absence of protein ($P = 0.0002$, Figure 42A). This degree of base flipping activity of hTDG correlates extremely well with the percentage of strongly bent ($\sim 70^\circ$) complexes in the AFM experiments (Figure 40A and B); the ratios of fluorescence enhancement in the 2-AP base flipping assay and percentages of complexes in the stronger bent state in AFM are both approximately 3:2 between hTDG complexes bound at G:U sites and nonspecific hTDG complexes during damage search (at non-damaged sites). In contrast, and consistent with our AFM results, hTDG-R275A (Figure 42B and C) induced no significant increase in 2-AP fluorescence compared to DNA in the absence of protein ($P = 0.356$) and showed significantly reduced activity compared to the wt protein ($P = 0.003$, Figure 42C and Table 21).

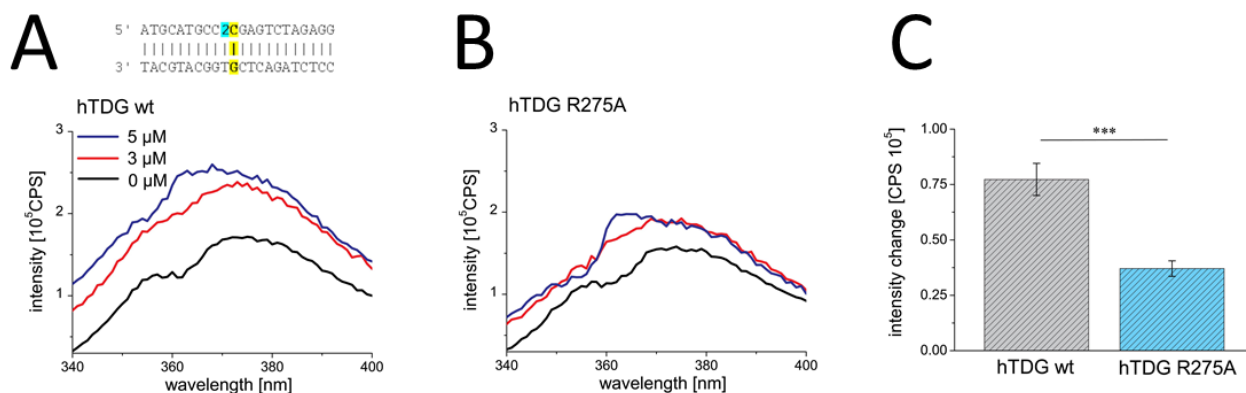


Figure 42: 2-AP fluorescence based base flipping assay of hTDG at nonspecific DNA.

Increasing concentrations of A) hTDG wt and B) hTDG-R275A were titrated to 170 nM 2-AP DNA substrates. Steady-state fluorescence emission spectra (340 – 400 nm) were recorded at an excitation wavelength of $\lambda_{\text{ex}} = 320$ nm. The inset in A) shows a schematic of the nonspecific DNA substrate used containing a 2-AP (cyan) neighboring a Watson-Crick G:C base pair (yellow). C) Quantification of the 2-AP fluorescence intensity change in A) and B) for protein concentrations of 5 μM compared to buffer control. Results were derived from two individual titrations. The error bars indicate the standard deviation. Significances are calculated with one-tailed student's t-test and are classed as *** $P < 0.005$.

Table 21: Quantification of 2-AP fluorescence based base flipping assay for hTDG proteins.

Protein	DNA substrate	rel. intensity change	P
hTDG wt	dsDNA 2-AP G:U ^F	131555 \pm 56391	0.047
hTDG R275A	dsDNA 2-AP G:U ^F	9370 \pm 7085	/
hTDG wt	dsDNA 2-AP nsp	77297 \pm 7239	0.003
hTDG R275A	dsDNA 2-AP nsp	37045 \pm 3514	/

2-AP fluorescence intensity changes for the indicated hTDG proteins (5 μM) and DNA substrates. Values represent means \pm SD derived from at least two individual experiments. Significances are calculated with a one-tailed student's t-test for the differences between hTDG variant R275A versus wt protein for lesion containing and nonspecific DNA substrate.

5.2.4 AFM volume analysis reveals the stoichiometry of free and DNA-bound hTDG

In two co-crystallographic studies, hTDG has been reported to form a dimer in complex with DNA (Maiti et al., 2008, Maiti et al., 2012) (see 2.4.3). In the X-ray structures, one hTDG subunit was bound at the specific site (abasic or uracil analog) and one subunit was bound to nonspecific DNA in direct proximity. However, the size of the hTDG dimer interface was small, which suggests only weak dimer interactions in the absence of DNA. Consistent with that, the free protein has been found to be monomeric in sedimentation velocity analysis even at high concentrations (Maiti et al., 2008). Although biochemical studies showed that hTDG, indeed, is able to bind different DNA substrates (unmodified and lesion-containing DNA) as a 2:1 complex at high protein concentrations ($K_{\text{D(dimerization)}} < 1.3 \mu\text{M}$), an hTDG

monomer has been indicated to be fully capable of repairing G:U and G:U mismatches (Maiti et al., 2008, Morgan et al., 2011).

To characterize the stoichiometry of hTDG in the absence and in the presence of DNA, AFM volumes of free (non-DNA-bound, Figure 43) and DNA-bound hTDG molecules (Figure 44) were determined. In principle, AFM volumes of proteins can be directly translated into molecular weight (see 4.1.3), thus providing information on their conformational and oligomeric properties.

Figure 43A shows an AFM image of hTDG in the absence of DNA. AFM volume analysis of free protein molecules resulted in a single Gaussian peak centered at $(48 \pm 20) \text{ nm}^3$ (Figure 43B) corresponding to a molecular weight of $\sim 45 \text{ kDa}$. This value is consistent with the size of an hTDG monomer (theoretical weight 46 kDa) and is in agreement with the previous studies (Maiti et al., 2008, Morgan et al., 2011).

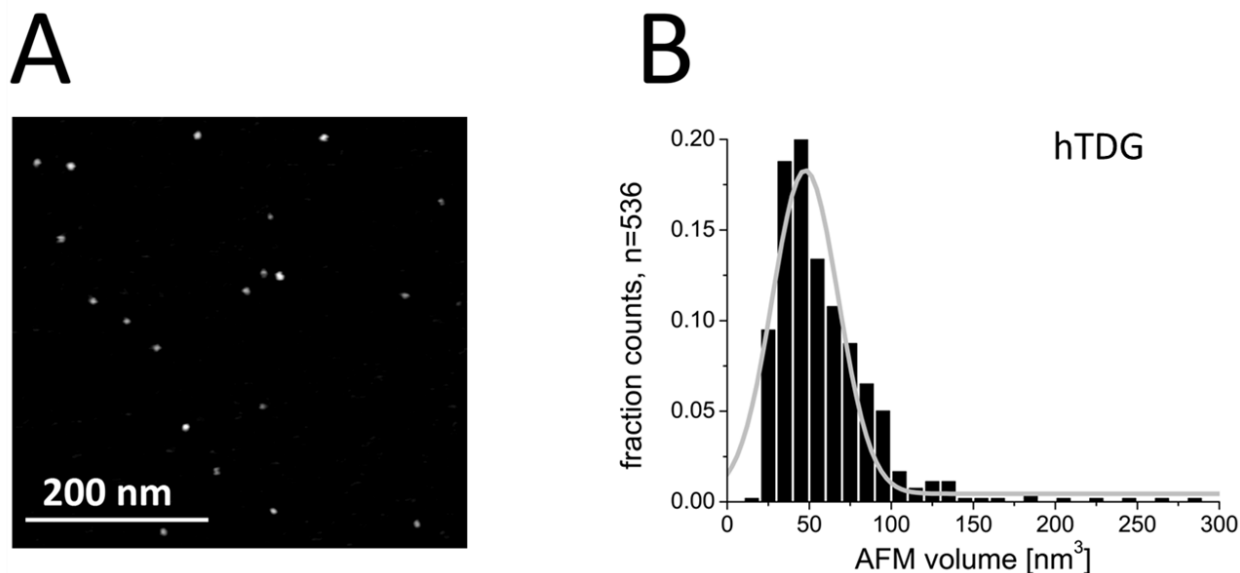


Figure 43: Stoichiometry of hTDG in AFM experiments.

A) AFM image of hTDG in the absence of DNA. B) AFM volume measurements of hTDG resulted in a distribution with a distinct peak. A Gaussian fit to the distribution centered at $(48 \pm 20) \text{ nm}^3$. This value corresponds to a molecular weight of $\sim 45 \text{ kDa}$ and is consistent with the size of a monomer. The distribution represents pooled data from three individual depositions. The number of datapoints is indicated.

To investigate conformational and/ or oligomeric states of hTDG on DNA, hTDG was incubated with specific DNA substrates containing a G:T mismatch. Incubations were performed at high protein concentrations ($2.5 \mu\text{M} > K_{\text{D(dimerization)}}$) (Figure 44) to allow potential dimer formation. For the volume analysis, specific hTDG complexes (bound at the mismatch, black bars) and nonspecific hTDG complexes with unmodified homoduplex DNA (grey bars) were treated separately.

For nonspecific hTDG complexes, two different volume states $(56 \pm 12) \text{ nm}^3$ (67 %) and $(82 \pm 12) \text{ nm}^3$ (33 %) can be observed from the resulting volume distribution, which correspond to molecular weights of $\sim 52 \text{ kDa}$ and $\sim 73 \text{ kDa}$, respectively. As these values reflect the size of the protein-DNA complexes

(the contribution of the DNA volume is estimated to be 5 - 10 nm³), these molecular weights are comparable with the size of an hTDG monomer (theoretical weight 46 kDa) and likely represent two different conformational states of hTDG formed on nonspecific DNA.

At lesion sites, three different populations of hTDG-DNA AFM volumes were discovered. A triple Gaussian fit to the data centers at (54 ± 13) nm³ (35 %), (83 ± 13) nm³ (44 %) and (116 ± 13) nm³ (21 %), which correspond to molecular weights of ~ 50 kDa, ~ 74 kDa and ~ 102 kDa, respectively. The two major states with AFM volumes of ~ 55 nm³ and ~ 80 nm³ likely represent two monomeric conformations similar as observed for nonspecific hTDG-DNA complexes. However, the equilibrium between these two conformational states is shifted towards the smaller volume state of ~ 55 nm³ for nonspecific complexes (compare 67 % for nsp complexes and 35 % for spec complexes) and towards the bigger volume state for specific complexes at the mismatch (compare 44 % for spec complexes and 33 % for nsp complexes). Only a minor population (21 %) of the largest volume (~ 115 nm³, corresponding to a molecular weight of ~102 kDa) has been found to consistent with the size of a dimer (92 kDa).

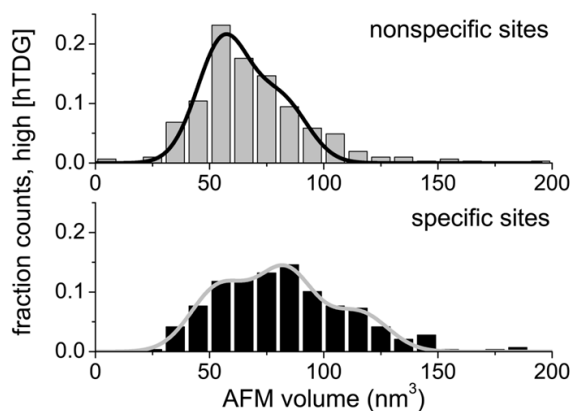


Figure 44: AFM volume analysis of hTDG-DNA complexes.

AFM volumes were determined at high hTDG concentration (2.5 μM) for nonspecific (grey bars) and specific hTDG-DNA complexes bound to G:T mismatches (black bars) separately. For nonspecific hTDG complexes (n = 307), two different states of AFM volumes were visible. A double Gaussian fit ($R^2 = 0.94$) centers at (56 ± 12) nm³ (67 %) and (82 ± 12) nm³ (33 %), corresponding to molecular weights of ~ 52 kDa and ~ 73 kDa. AFM volume measurements of specific hTDG-DNA complexes (n = 288) resulted in a triphasic distribution, with peaks at (54 ± 13) nm³ (35 %), (83 ± 13) nm³ (44 %) and (116 ± 13) nm³ (21 %), which correspond to molecular weights of ~ 50 kDa, ~ 74 kDa and ~ 102 kDa, respectively. Only the largest AFM volume (~ 115 nm³ corresponding to ~ 100 kDa) found at specific G:T sites is consistent with a dimeric form of hTDG. The distributions represent pooled data from three individual experiments.

5.2.5 hTDG complex size correlates with DNA bending

Based on the AFM volume analysis (Figure 44), hTDG-DNA complexes of smaller (35 - 70 nm³) and larger size (70 - 110 nm³), which likely include two different monomeric conformations, were separated for further DNA bend angle analysis. DNA bend angle measurements for smaller complexes (Figure 45A, cyan bars) resulted in a biphasic distribution with maxima at (36 ± 12)° (56 %) and at (68 ± 12)° (44 %)

respectively. For larger complexes, the distribution was also biphasic and a double Gaussian fit to the data centers at $(28 \pm 13)^\circ$ (32 %) and at $(62 \pm 13)^\circ$ (68 %) (Figure 45B, magenta bars). The two populations observed for both smaller and larger hTDG-DNA complexes with bend angles of $(28 - 36)^\circ$ and $(62 - 68)^\circ$ (Figure 45) closely resemble to the bend angle states of $\sim 30^\circ$ and $\sim 70^\circ$ measured for hTDG complexes, when nonspecific and specific DNA sites were separated (compare Figure 36 and Figure 40).

Notably, when separating the bend angles values according to their AFM volumes, the equilibrium between these two states with bend angles of $(28 - 36)^\circ$ and $(62 - 68)^\circ$ is shifted towards the smaller bend angle state for small complexes and towards the larger bend angle state for large complexes ($P = 0.006$, the difference between the relative areas of small and large complexes for the population with a bend angle of $\sim 65^\circ$ is highly significant). Hence, the stronger bent complexes seem to correlate with a larger complex size. Assuming that protein complexes in a flipped state appear larger in AFM (leading to larger AFM-volumes), the observed correlation between larger bend angles and larger AFM volumes may confirm a role of strong DNA bending (resulting in higher DNA flexibility) in facilitating nucleotide flipping. This finding is consistent with the AFM data on the hTDG variants N140A and R275A as described above (see 5.2.2 and 5.2.3) suggesting that hTDG may exploit the higher DNA flexibility at G:T and G:U mismatches to induce DNA kinking and extra-helical base interrogation (see model in Figure 46).

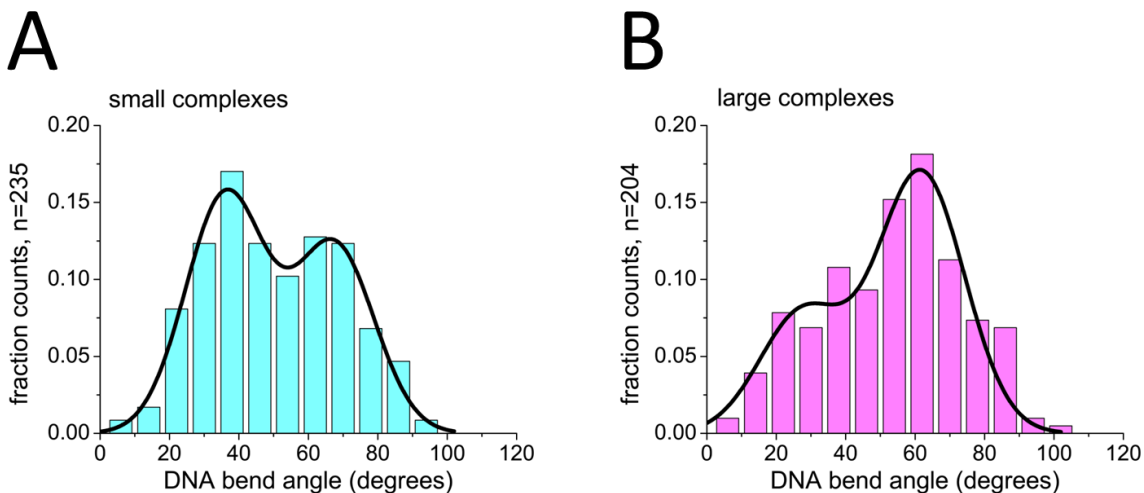


Figure 45: Correlation between DNA bending and AFM volume of hTDG-DNA complexes.

DNA bend angle measurements of hTDG-DNA complexes (at specific G:T sites and nonspecific DNA sites) were separated for A) smaller ($35 - 70 \text{ nm}^3$) and B) larger forms ($70 - 110 \text{ nm}^3$). A) For small complexes, two distinct populations of AFM volumes were apparent in a double Gaussian fit ($R^2 = 0.97$) with maxima at $(36 \pm 12)^\circ$ (56 %) and at $(68 \pm 12)^\circ$ (44 %). B) For large complexes, two bend angle populations were visible. A double Gaussian fit ($R^2 = 0.88$) to the data centers at $(28 \pm 13)^\circ$ (32 %) and at $(62 \pm 13)^\circ$ (68 %). The relative area (in %) of the second population with a DNA bend angle of $\sim 65^\circ$ is significantly increased for the larger complexes (**, $P = 0.006$). The significance was calculated by one-tailed student's t-test. The distributions represent pooled data from three individual experiments. The number of data points is indicated as n.

5.2.6 hTDG damage search and recognition model

TDG target sites in DNA (G:T, G:U, and its analog G:U^F) were characterized using AFM-based DNA bend angle analyses. In the absence of protein, the DNA was intrinsically bent by $\sim 30^\circ$ at these base mismatch sites (Figure 36B,D and Figure 38B). It is worth noting that in the AFM experiments, intrinsic bending at different, specific DNA sites may be somewhat promoted by the surface deposition process. However, DNA bending at mismatch sites in the AFM experiments still reflects the degree of destabilization at these sites because intrinsic bending is not observed for nonspecific DNA (Figure 41). In addition, the findings observed in AFM are in accordance with computational simulations, which show that G:T and G:U wobble pairs are kinetically less stable than canonical base pairs and are more prone to spontaneous base pair breathing, inducing minor DNA distortions (Fuxreiter et al., 2002, Seibert et al., 2002, Imhof and Zahran, 2013, Yin et al., 2014).

Glycosylases may hence have adapted to exploit these open (flipped) base configurations, stabilizing them to achieve a specific conformation required for lesion detection and removal. This scenario will be referred to as passive bending (model in Figure 46). Due to its higher flexibility, damaged DNA is also more prone to protein-induced DNA distortion or bending (active bending, Figure 46), which has been shown to reduce the energetic barrier for enzymatic base flipping (Sharma et al., 2013).

The same slightly bent state observed at DNA mismatches in the absence of protein (bend angle of $\sim 30^\circ$) is also observed in the protein complexes bound at DNA mismatch sites (Figure 36A,C). My data hence suggest that this bend angle state represents a complex of TDG adapted to the naturally pre-bent (or more flexible) DNA structure. Interestingly, a high fraction of these slightly bent hTDG search complexes (with a bend angle of $\sim 30^\circ$) is also present during damage search on nonspecific DNA (Figure 40A-C), which (in contrast to the mismatch sites) shows no intrinsic DNA bending (Figure 40D and Figure 41B,D). hTDG binding may hence impose this degree of DNA bending on undamaged DNA as part of its initial lesion sensing mechanism. This conformation (bend angle of $\sim 30^\circ$) will be termed as the lesion search complex (SC, green circles in model Figure 46) prior to damage recognition. Bending of nonspecific DNA would induce strain on the nucleotides, energetically facilitating their extrusion from the DNA double helix for interrogation by the glycosylase (active bending, Figure 46).

In contrast, at its target sites, hTDG would simply stabilize the pre-bent structure of the mismatches and the net energy consumption for bending would be zero (passive bending, Figure 46). The energetic cost of DNA bending may serve as an initial target site sensing mechanism by hTDG. My findings hence suggest the search conformation to be dictated by the intrinsic structure of the target substrate. The degree of bending observed for the SC of hTDG in the AFM experiments correlates well with the degree of DNA bending measured from crystal structures. The structures of hTDG bound to an abasic site analog

(product complex, (Maiti et al., 2008)) and of hTDG bound to G:U^F (lesion recognition complex, (Maiti et al., 2012)) contain a second subunit of hTDG that is bound to undamaged DNA and interacts with the specific complex via a weak dimer interface. The DNA bound by this nonspecific complex is bent by ~ 25° in the structures, consistent with the degree of bending observed in my AFM studies and with the idea that the bending exerted on nonspecific DNA substrates is dictated by a preformed DNA binding surface of hTDG.

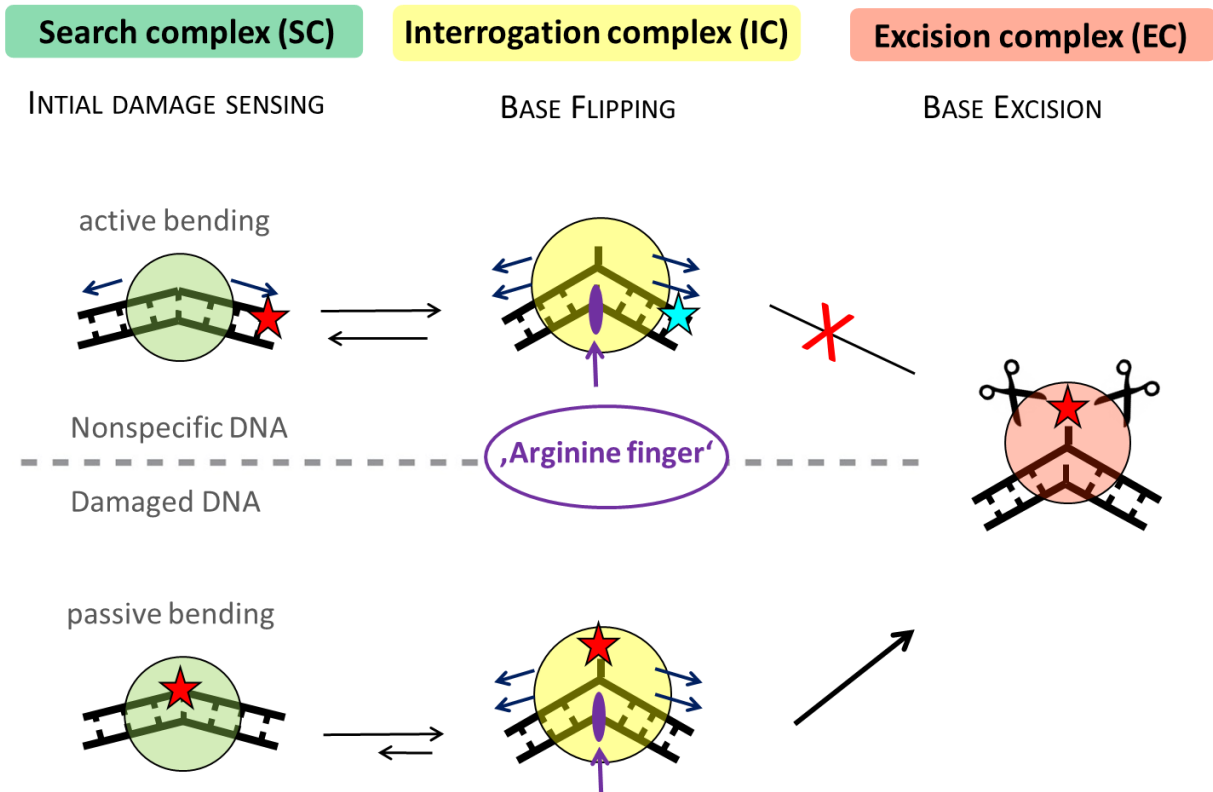


Figure 46: Model for lesion search and recognition by hTDG.

DNA lesion search and recognition by glycosylases involves three different types of TDG-DNA complexes: search complex (SC), interrogation complex (IC), and excision complex (EC). The glycosylase scans the DNA for intrinsically flexible sites as a SC, continuously converting to an IC state with the target base flipped out of the DNA double helix for better accessibility. While homoduplex DNA is actively bent by the glycosylase during the search process, G:T or G:U wobble base pairs (red star) display enhanced flexibility or intrinsic pre-bending that matches its conformation in the SC complex (passive bending). The energetic cost for DNA bending may hence serve as an initial damage sensing mechanism. Correct target bases are stabilized in the extrahelical state within the catalytic pocket of the glycosylase in the EC conformation, supported by interactions of protein residues with the DNA (for TDG e.g. R²⁷⁵, which corresponds to the 'arginine finger') and catalysis of base excision can occur.

The strictly conserved residues Arg²⁷⁵ and Asn¹⁴⁰ of hTDG have been previously proposed to be involved in nucleotide flipping and base excision, respectively (Maiti et al., 2009). In the AFM experiments, both hTDG variants N140A and R275A are able to recognize a G:U or G:T mismatch (Figure 35, Figure 37, Table 19) indicating that neither Asn¹⁴⁰ nor Arg²⁷⁵ are absolutely required for initial damage recognition by

hTDG. During lesion search on nonspecific DNA, the population of the slightly bent DNA state ($\sim 30^\circ$ bend angle) in both variants is also completely comparable to that in the wt protein. The data hence suggest a model (Figure 46), which includes an initial damage sensing step for TDG prior to base flipping and excision (represented by the slightly bent TDG-DNA structures), based on recognition of intrinsic structural properties of the target sites.

Final target identification by hTDG as well as base excision require extrahelical base flipping as visualized by crystal structures of the lesion recognition complex (Maiti et al., 2012, Hashimoto et al., 2013). In the structures, the target base is brought in close contact with catalytic residues of the active site by complete flipping out of the DNA duplex. The damaged base is thereby perfectly stabilized by active site residues, leading to an enhanced residence time of the correct (target) DNA substrate in the catalytic pocket, which results in a higher probability for base excision.

In my AFM analyses, a strongly bent DNA population (with a DNA bend angle of $\sim 70^\circ$) was observed in addition to the slightly bent one ($\sim 30^\circ$) as discussed above (Figure 36A,C and Figure 40A,B). The strongly bent DNA conformation is induced by hTDG because it was not displayed by negative controls of DNA (with or without a specific mismatch site) in the absence of protein. Thus, this more strongly bent state is interpreted as the conformation, in which a base is flipped into the active site pocket of TDG for interrogation and this conformation will be referred to as the lesion interrogation complex (IC, yellow circles Figure 46). Interestingly, the relative population of the strongly bent complex conformation in the experiments was slightly lower for hTDG bound to G:T versus G:U mismatch sites (58 % for G:T, 67% for G:U), consistent with previous findings that flipping of T is hindered relative to U flipping (Maiti et al., 2012).

Significantly reduced populations of the $\sim 70^\circ$ state observed in the AFM experiments (Figure 38A and Figure 40C) as well as a significant reduction of base destabilization activity in the fluorescence base flipping assays for the hTDG-R275A variant (Figure 39 and Figure 42) support this interpretation and are consistent with the proposed essential role of Arg²⁷⁵ in base flipping and stabilization of the “phosphate pinch” from crystal structures (Maiti et al., 2008, Maiti et al., 2009). It is worth noting that the DNA bending phenotype of hTDG-R275A in the AFM studies and its lack of base flipping activity observed in the 2-AP fluorescence assays is not due to impaired DNA substrate binding activity of the Arg variant as DNA binding affinities were found to be only slightly (≤ 3 -fold) reduced for hTDG-R275A versus hTDG wt (Maiti et al., 2009, Morgan et al., 2011).

Importantly, in my AFM studies, the strongly bent state ($\sim 70^\circ$ bend angle) was clearly present and similarly populated in nonspecific complexes of TDG with homoduplex DNA as in the specific TDG-lesion

complexes. Significant increase in 2-AP fluorescence in the base flipping assays further supports destabilization of nonspecific DNA bases by TDG during lesion search, likely by (partial or complete) base flipping into the protein's active site pocket (Figure 42). The finding that both SC and IC are present both at nonspecific and at specific DNA sites strongly suggests an active, constant DNA interrogation process by TDG involving base flipping as the protein scans the DNA in search of a target site.

5.2.7 Generalizability of hTDG lesion search and detection in comparison with hOgg1

To test whether the proposed damage search and recognition strategies for hTDG can be transferred to a general model for BER, the AFM and fluorescence analyses established for hTDG were extended to investigate hOgg1 (Figure 47), a different and also well-characterized DNA glycosylase (see 2.4.4). The two enzymes detect damaged bases with significantly different base pairing energies. hTDG targets intrinsically flexible G:U and G:T wobble pairs (Imhof and Zahran, 2013) that result in slightly bent structures in the AFM images (Figure 36B,D). In contrast, hOgg1 addresses oxoG, which forms stable Watson-Crick like base pairs with cytosine (Crenshaw et al., 2011) that lack intrinsic DNA bending and show straight conformations in AFM images in the absence of protein (Figure 47C and Table 22).

To investigate hOgg1 lesion search processes, the protein was incubated with long undamaged DNA fragments in the AFM experiments (see representative AFM image in Figure 47A). AFM-based DNA bend angle analysis of nonspecific hOgg1-DNA complexes reveals three different states of $(0 \pm 9)^\circ$ (44 %), $(33 \pm 9)^\circ$ (20 %), and $(70 \pm 9)^\circ$ (36 %) (Figure 47B and Table 22). The two major populations with bend angles of $\sim 0^\circ$ and $\sim 70^\circ$ are in agreement with previous AFM studies performed with hOgg1 (Chen et al., 2002). The straight DNA conformation can be interpreted as SC (with a bend angle of $\sim 0^\circ$ matching the DNA conformation of the target site) and the strongly bent conformation (with a bend angle of $\sim 70^\circ$), may represent the IC with the flipped out base. Similar as observed in the distribution of nonspecific hTDG complexes (Figure 40A,B), the bend angle distribution of hOgg1 suggests an equilibrium between the SC and IC during damage search likely involving an additional intermediate state (reflected by the small population of slightly bent hOgg1-DNA complexes with a bend angle of $\sim 30^\circ$). The presence of the kinked IC (with bend angle of $\sim 70^\circ$) in the AFM studies (Figure 47B and (Chen et al., 2002)) is consistent with the requirement for DNA flexibility (resulting in the strong DNA bending) for DNA repair activity by hOgg1 as discovered in high speed AFM experiments using DNA nanochips as scaffolds (Endo et al., 2010). In addition, the same DNA bend angle of $\sim 70^\circ$ was also observed in co-crystallographic studies of hOgg1 bound to its target site 8oxoG:C (Bruner et al., 2000) or bound to a nonspecific base with the base flipped into an exo-site in close proximity to the enzyme's active site pocket (Banerjee et al., 2005).

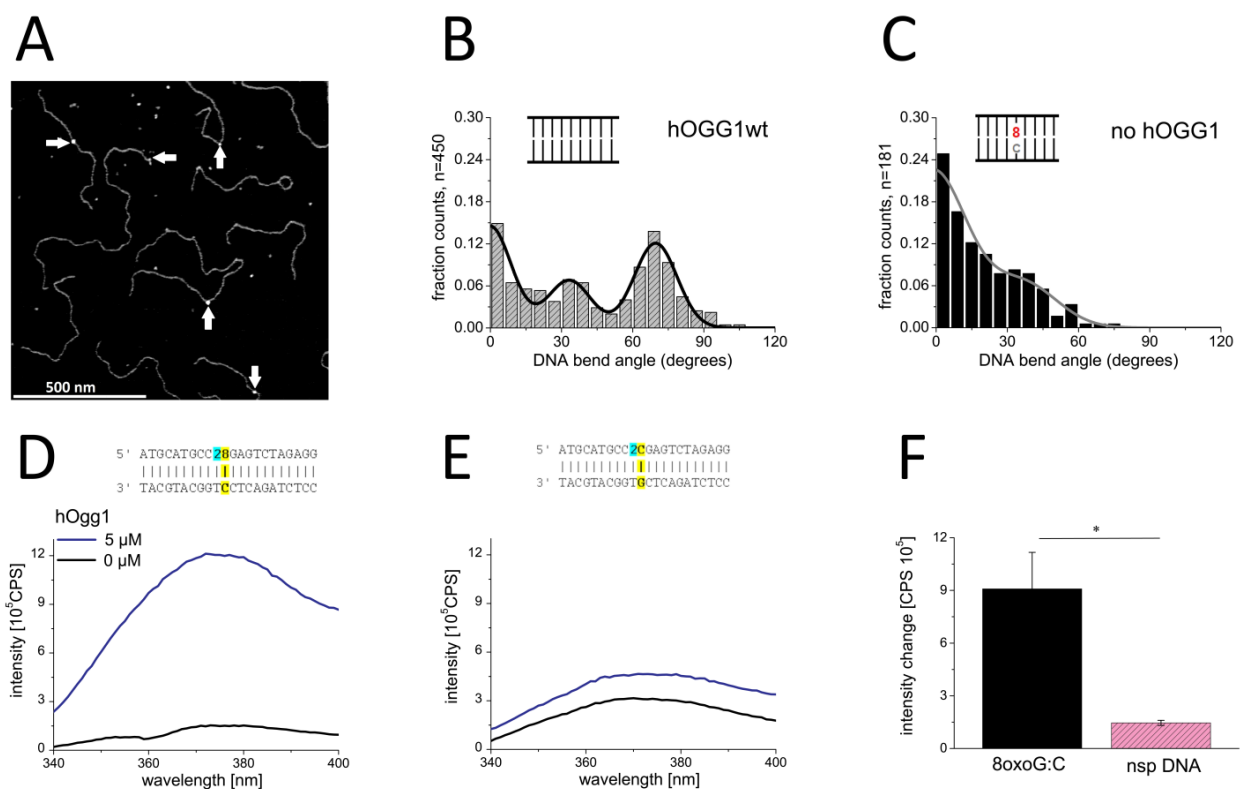


Figure 47: Characterization of hOgg1-DNA complexes by AFM and fluorescence experiments.

A) AFM image of hOgg1 binding to nsp DNA (1813 bp). Arrows point to hOgg1-DNA complexes. B) Distribution of DNA bend angles for hOgg1 bound to nonspecific DNA. A triple Gaussian fit ($R^2 = 0.87$) centered at $\sim 0^\circ$ (44 %), $\sim 30^\circ$ (20 %) and $\sim 70^\circ$ (36 %). C) Intrinsic DNA bend angles at an 8oxoG:C site (target site of hOgg1) in the absence of protein. A double Gaussian fit ($R^2 = 0.929$) revealed a dominant straight fraction with bend angle $\sim 0^\circ$ (78 %) and a small shoulder at $\sim 30^\circ$ (22 %). Bend angle results were pooled from at least two individual experiments. The total number of data points is indicated as n. D,E) hOgg1 was titrated to 2-AP DNA substrates containing D) 8oxoG:C or E) nonspecific 2-AP DNA substrates. The fluorescence intensities in the absence (black curve) and in the presence of hOgg1 (blue curve at 5 μM [protein]) are shown. Steady-state fluorescence emission spectra (340 – 400 nm) were recorded at $\lambda_{\text{ex}} = 320$ nm. The insets show schematics of the employed 2-AP (cyan) DNA substrates containing either an 8oxoG:C pair (D, yellow) or a stable Watson-Crick base pair (E, yellow). F) Quantification of the 2-AP fluorescence intensity changes in D) and E). Results are averages from two individual titrations with error bars indicating the standard deviations. Significances are calculated with a one-tailed student's t-test and are classed as * $P < 0.05$.

In the 2-AP fluorescence base flipping assays, large fluorescence increases can be observed for hOgg1 both with DNA substrates containing oxoG:C lesion sites and with non-damaged DNA (Figure 47D,E and Table 23). This finding suggests considerable stabilization of the flipped out base by hOgg1 both at its target sites and at nonspecific homoduplex DNA sites. The relative 2-AP fluorescence increase for hOgg1 is consistently higher than for hTDG, most likely due to the different chemical environments of the two proteins (see Methods 4.5.3). In addition, the increase in 2-AP fluorescence for the specific DNA substrate (containing 8oxoG:C) compared to nonspecific, homoduplex DNA is significant ($P = 0.018$). This is not the case for hTDG which induces a larger degree of 2-AP destabilization for nonspecific DNA relative to specific (G:U^F containing) DNA ($P = 0.085$, not significantly different).

Notably, in my AFM studies (Figure 36A,C, Figure 40A,B, Figure 47B), the bend angle conformations of the ICs of hTDG and hOgg1 closely resemble each other (DNA bend angle of $\sim 70^\circ$ for both proteins), which is also in line with a comparable degree of DNA bending in a number of crystal structures of different DNA glycosylases bound specifically to their flipped targets (Hollis et al., 2000, Parikh et al., 2000, Bruner et al., 2000, Zharkov et al., 2002, Fromme and Verdine, 2002, Bowman et al., 2010).

In contrast, the degree of DNA bending in the SC, which is difficult to capture by crystallographic studies due to its transient nature, differs between the two proteins: search complexes of hOgg1 are predominantly straight while those of hTDG are slightly bent consistent with the innate conformations of their target sites in the absence of protein (Figure 36B,D, Figure 38B and Figure 47C). The difference in DNA bending of the SC between the two glycosylases hence corroborates the hypothesis that DNA glycosylases have adapted to the native DNA structure of their target lesions.

Table 22: DNA bend angle analysis for hOgg1-DNA complexes and free DNA substrate 8oxoG:C.

protein	DNA substrate	straight	slightly bent	strongly bent	n
hOgg1 wt	Nsp DNA (1813 bp)	0 ± 9 (44 %)	33 ± 9 (20 %)	70 ± 9 (36 %)	450
none	8oxoG:C	-2 ± 15 (78 %)	36 ± 15 (22%)	NA	181

Depending on DNA bending (in degrees), ($\sim 0^\circ$, $\sim 30^\circ$ and $\sim 70^\circ$), complexes are classed as 'straight', 'slightly bent' or 'strongly bent'. The given values represent the maxima of single, double or triple Gaussian fits to pooled bend angle distributions. The different degree of population of the different bend angle states is given in brackets as % of total area from integration of the Gaussian peaks. The total number of data points is indicated as n. NA = not applicable.

Table 23: Quantification of 2-AP fluorescence based base flipping assay for hOgg1.

Protein	DNA substrate	rel. intensity change	P
hOgg1	dsDNA 2-AP nsp	145780 ± 14425	0.018
hOgg1	dsDNA 2-AP oxoG	908275 ± 208547	/

2-AP fluorescence intensity changes for hOgg1 (5 μ M) and the indicated DNA substrates. Values represent means \pm SD derived from two individual experiments. Significances are calculated with a one-tailed student's t-test for the differences between hOgg1 with specific 8oxoG:C DNA versus hOgg1 with nonspecific DNA substrate.

6. Final conclusions

Altogether, this thesis provided novel insights into damage search and recognition of two different DNA repair mechanisms. Detailed single molecule analysis from AFM imaging experiments complemented with biophysical ensemble techniques unraveled different lesion search and verification strategies of central players in NER and BER. In contrast to conventional ensemble methods, which represent an averaged sum of all molecules, single molecule analysis can resolve different populations and conformational states of individual molecules (Record et al., 1991). In addition, AFM can directly visualize protein-DNA interactions. Hence, AFM imaging has been exploited here to improve the understanding of fundamental steps in DNA repair.

More specifically, my AFM and BLI data on the archaeal homolog of the NER helicase XPD, taXPD, revealed different strand selectivity for two different types of lesions accounting for the high substrate versatility of NER. My data on taXPD resulted in a simplistic model for damage verification, which lays the foundation for understanding the role of XPD in eukaryotic NER in the context of the multi-subunit TFIIH complex and other NER factors. Recently, the genome of the thermophilic fungus *Chaetomium thermophilum* has been published (Amlacher et al., 2011), which can serve as a more adequate model for the human NER system in future AFM studies. Unlike archaea, the fungus contains all members of TFIIH. Proteins from *C. Thermophilum* are closely related to their human homologs, but more stable and hence more amenable to purification and *in vitro* studies (Johnson et al., 1985, Amlacher et al., 2011). The sequence of XPD from the fungus, for instance, shares a high homology of 74 % and an identity of 55 % (as defined by BLAST, NCBI) with the sequence of its human counterpart (Kuper et al., 2014).

The second part of the thesis demonstrates an essential role of base flipping in damage search and recognition by DNA glycosylases both by AFM imaging studies and 2-AP fluorescence based base flipping assays. hTDG serves as representative from the glycosylase family to investigate the poorly understood processes of initial damage sensing prior to nucleotide flipping in BER. The AFM data suggest that hTDG uses the enhanced intrinsic flexibility at G:T and G:U mismatches for initial damage sensing. Notably, the results from AFM and fluorescence analyses indicate a dynamic equilibrium between two different conformational states of hTDG-DNA complexes involving continuous base flipping during search. The findings for hTDG are summarized as a general model for damage search and recognition in BER, which is supported by the data on a second glycosylase, hOgg1, which recognizes structurally less flexible DNA lesions.

7. References

- AISHIMA, J., GITTI, R. K., NOAH, J. E., GAN, H. H., SCHLICK, T. & WOLBERGER, C. 2002. A Hoogsteen base pair embedded in undistorted B-DNA. *Nucleic Acids Research*, 30, 5244-5252.
- ALBRECHT, T. R., AKAMINE, S., CARBER, T. E. & QUATE, C. F. 1989. Microfabrication of cantilever styli for the atomic force microscope. *J. Vac. Sci. Technol.*, 8, 3386-3396.
- ALLEN, S., DAVIES, M. C., ROBERTS, C. J., TENDLER, S. J. B. & WILLIAMS, P. M. 1997. Atomic force microscopy in analytical biotechnology. *Trends in Biotechnology*, 15, 101-105.
- AMLACHER, S., SARGES, P., FLEMMING, D., VAN NOORT, V., KUNZE, R., DEVOS, D. P., ARUMUGAM, M., BORK, P. & HURT, E. 2011. Insight into structure and assembly of the nuclear pore complex by utilizing the genome of a eukaryotic thermophile. *Cell*, 146, 277-89.
- ANDO, T. 2013. High-speed atomic force microscopy. *Microscopy (Oxf)*, 62, 81-93.
- ANDO, T. 2014. High-speed AFM imaging. *Curr Opin Struct Biol*, 28C, 63-68.
- ANDO, T., UCHIHASHI, T. & KODERA, N. 2013. High-speed AFM and applications to biomolecular systems. *Annu Rev Biophys*, 42, 393-414.
- BALASHEV, K., JENSEN, T. R., KJAER, K. & BJORNHOLM, T. 2001. Novel methods for studying lipids and lipases and their mutual interaction at interfaces. Part I. Atomic force microscopy. *Biochimie*, 83, 387-97.
- BANERJEE, A., YANG, W., KARPLUS, M. & VERDINE, G. L. 2005. Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature*, 434, 612-8.
- BARTEK, J., BARTKOVA, J. & LUKAS, J. 2007. DNA damage signalling guards against activated oncogenes and tumour progression. *Oncogene*, 26, 7773-9.
- BARZILAY, G., MOL, C. D., ROBSON, C. N., WALKER, L. J., CUNNINGHAM, R. P., TAINER, J. A. & HICKSON, I. D. 1995. Identification of critical active-site residues in the multifunctional human DNA repair enzyme HAP1. *Nat Struct Biol*, 2, 561-8.
- BATTY, D. P. & WOOD, R. D. 2000. Damage recognition in nucleotide excision repair of DNA. *Gene*, 241, 193-204.
- BEARD, W. A. & WILSON, S. H. 2006. Structure and mechanism of DNA polymerase Beta. *Chem Rev*, 106, 361-82.
- BELLAMY, S. R. & BALDWIN, G. S. 2001. A kinetic analysis of substrate recognition by uracil-DNA glycosylase from herpes simplex virus type 1. *Nucleic Acids Res*, 29, 3857-63.
- BELLAMY, S. R., KRUSONG, K. & BALDWIN, G. S. 2007. A rapid reaction analysis of uracil DNA glycosylase indicates an active mechanism of base flipping. *Nucleic Acids Res*, 35, 1478-87.
- BENZANILLA, M., DRAKE, B., NUDLER, E., KASHLEV, M., HANSMA, P. K. & HANSMA, H. G. 1994. Motion and enzymatic degradation of DNA in the atomic force microscope. *Biophys. J.*, 67, 2454-2459.
- 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, 6929-48.
- BERNARDES DE JESUS, B. M., BJORAS, M., COIN, F. & EGLY, J. M. 2008. Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. *Mol Cell Biol*, 28, 7225-35.

- BILLINGSLEY, D. J., BONASS, W. A., CRAMPTON, N., KIRKHAM, J. & THOMSON, N. H. 2012. Single-molecule studies of DNA transcription using atomic force microscopy. *Phys Biol*, 9, 021001.
- BINNIG, G., QUATE, C. F. & GERBER, C. 1986. Atomic force microscope. *Phys Rev Lett*, 56, 930-933.
- BINNIG, G., ROHRER, H., GERBER, C. & WEIBEL, E. 1982. Tunneling through a Controllable Vacuum Gap. *Applied Physics Letters*, 40, 178-180.
- BJORAS, M., SEEBERG, E., LUNA, L., PEARL, L. H. & BARRETT, T. E. 2002. Reciprocal "flipping" underlies substrate recognition and catalytic activation by the human 8-oxo-guanine DNA glycosylase. *J Mol Biol*, 317, 171-7.
- BLACKLEY, H. K., SANDERS, G. H., DAVIES, M. C., ROBERTS, C. J., TENDLER, S. J. & WILKINSON, M. J. 2000. In-situ atomic force microscopy study of beta-amyloid fibrillization. *J Mol Biol*, 298, 833-40.
- BLAINEY, P. C., VAN OIJEN, A. M., BANERJEE, A., VERDINE, G. L. & XIE, X. S. 2006. A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc Natl Acad Sci U S A*, 103, 5752-7.
- BOWMAN, B. R., LEE, S., WANG, S. & VERDINE, G. L. 2010. Structure of Escherichia coli AlkA in complex with undamaged DNA. *J Biol Chem*, 285, 35783-91.
- BRUINSMA, R. F. 2002. Physics of protein-DNA interaction. *Physica a-Statistical Mechanics and Its Applications*, 313, 211-237.
- BRUNER, S. D., NORMAN, D. P. & VERDINE, G. L. 2000. Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA. *Nature*, 403, 859-66.
- BUECHNER, C. N., HEIL, K., MICHELS, G., CARELL, T., KISKER, C. & TESSMER, I. 2014. Strand-specific recognition of DNA damages by XPD provides insights into nucleotide excision repair substrate versatility. *J Biol Chem*, 289, 3613-24.
- BUECHNER, C. N. & TESSMER, I. 2013. DNA substrate preparation for atomic force microscopy studies of protein-DNA interactions. *J Mol Recognit*, 26, 605-17.
- BURROWS, C. J. & MULLER, J. G. 1998. Oxidative Nucleobase Modifications Leading to Strand Scission. *Chem Rev*, 98, 1109-1152.
- BUSTAMANTE, C., SMITH, S. B., LIPHARDT, J. & SMITH, D. 2000. Single-molecule studies of DNA mechanics. *Curr Opin Struct Biol*, 10, 279-85.
- BUSTAMANTE, C., VESENKA, J., TANG, C. L., REES, W., GUTHOLD, M. & KELLER, R. 1992. Circular DNA molecules imaged in air by scanning force microscopy. *Biochemistry*, 31, 22-6.
- CALDECOTT, K. W. 2003. XRCC1 and DNA strand break repair. *DNA Repair (Amst)*, 2, 955-69.
- CAMENISCH, U., DIP, R., SCHUMACHER, S. B., SCHULER, B. & NAEGELI, H. 2006. Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nat Struct Mol Biol*, 13, 278-84.
- CANTOR, C. R., WARSHAW, M. M. & SHAPIRO, H. 1970. Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers*, 9, 1059-77.
- CAO, C., JIANG, Y. L., KROSKY, D. J. & STIVERS, J. T. 2006. The catalytic power of uracil DNA glycosylase in the opening of thymine base pairs. *J Am Chem Soc*, 128, 13034-5.
- CAO, C., JIANG, Y. L., STIVERS, J. T. & SONG, F. 2004. Dynamic opening of DNA during the enzymatic search for a damaged base. *Nat Struct Mol Biol*, 11, 1230-6.

- CAREY, J. 1991. Gel retardation. *Methods Enzymol*, 208, 103-17.
- CAVALUZZI, M. J. & BORER, P. N. 2004. Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res*, 32, e13.
- CHEN, L., HAUSHALTER, K. A., LIEBER, C. M. & VERDINE, G. L. 2002. Direct visualization of a DNA glycosylase searching for damage. *Chem Biol*, 9, 345-50.
- CHU, G. & CHANG, E. 1988. Xeroderma pigmentosum group E cells lack a nuclear factor that binds to damaged DNA. *Science*, 242, 564-7.
- CLAESSON, P. M., HERDER, P., STENIUS, P., ERIKSSON, J. C. & PASHLEY, R. M. 1986. An Esca and Aes Study of Ion-Exchange on the Basal-Plane of Mica. *Journal of Colloid and Interface Science*, 109, 31-39.
- COIN, F., OKSENYCH, V. & EGLY, J. M. 2007. Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Mol Cell*, 26, 245-56.
- COIN, F., OKSENYCH, V., MOCQUET, V., GROH, S., BLATTNER, C. & EGLY, J. M. 2008. Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Mol Cell*, 31, 9-20.
- COLTON, R. J., BASELT, D. R., DUFRENE, Y. F., GREEN, J. B. & LEE, G. U. 1997. Scanning probe microscopy. *Curr Opin Chem Biol*, 1, 370-7.
- COMPE, E. & EGLY, J. M. 2012. TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol*, 13, 343-54.
- CORTAZAR, D., KUNZ, C., SAITO, Y., STEINACHER, R. & SCHAR, P. 2007. The enigmatic thymine DNA glycosylase. *DNA Repair (Amst)*, 6, 489-504.
- COX, C. J., FOSTER, P. G., HIRT, R. P., HARRIS, S. R. & EMBLEY, T. M. 2008. The archaeobacterial origin of eukaryotes. *Proc Natl Acad Sci U S A*, 105, 20356-61.
- CRAMPTON, N., BONASS, W. A., KIRKHAM, J., RIVETTI, C. & THOMSON, N. H. 2006. Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. *Nucleic Acids Res*, 34, 5416-25.
- CRAMPTON, N., YOKOKAWA, M., DRYDEN, D. T., EDWARDSON, J. M., RAO, D. N., TAKEYASU, K., YOSHIMURA, S. H. & HENDERSON, R. M. 2007. Fast-scan atomic force microscopy reveals that the type III restriction enzyme EcoP15I is capable of DNA translocation and looping. *Proc Natl Acad Sci U S A*, 104, 12755-60.
- CRENSHAW, C. M., NAM, K., OO, K., KUTCHUKIAN, P. S., BOWMAN, B. R., KARPLUS, M. & VERDINE, G. L. 2012. Enforced presentation of an extrahelical guanine to the lesion recognition pocket of human 8-oxoguanine glycosylase, hOGG1. *J Biol Chem*, 287, 24916-28.
- CRENSHAW, C. M., WADE, J. E., ARTHANARI, H., FRUEH, D., LANE, B. F. & NUNEZ, M. E. 2011. Hidden in plain sight: subtle effects of the 8-oxoguanine lesion on the structure, dynamics, and thermodynamics of a 15-base pair oligodeoxynucleotide duplex. *Biochemistry*, 50, 8463-77.
- DALHUS, B., FORSBRING, M., HELLE, I. H., VIK, E. S., FORSTROM, R. J., BACKE, P. H., ALSETH, I. & BJORAS, M. 2011. Separation-of-function mutants unravel the dual-reaction mode of human 8-oxoguanine DNA glycosylase. *Structure*, 19, 117-27.
- DE LAAT, W. L., APPELDOORN, E., SUGASAWA, K., WETERINGS, E., JASPERS, N. G. & HOEIJMAKERS, J. H. 1998. DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev*, 12, 2598-609.

- DEDON, P. C. 2011. Oxidation and deamination of DNA by endogenous sources. *Current Cancer Research: Chemical Carcinogenesis.*, 209–225.
- DIP, R., CAMENISCH, U. & NAEGELI, H. 2004. Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair. *DNA Repair (Amst)*, 3, 1409-23.
- DRAKE, B., PRATER, C. B., WEISENHORN, A. L., GOULD, S. A., ALBRECHT, T. R., QUATE, C. F., CANNELL, D. S., HANSMA, H. G. & HANSMA, P. K. 1989. Imaging crystals, polymers, and processes in water with the atomic force microscope. *Science*, 243, 1586-9.
- DUFRENE, Y. F. & LEE, G. U. 2000. Advances in the characterization of supported lipid films with the atomic force microscope. *Biochim Biophys Acta*, 1509, 14-41.
- DUNAND-SAUTHIER, I., HOHL, M., THOREL, F., JAQUIER-GUBLER, P., CLARKSON, S. G. & SCHARER, O. D. 2005. The spacer region of XPG mediates recruitment to nucleotide excision repair complexes and determines substrate specificity. *J Biol Chem*, 280, 7030-7.
- EGHIAIAN, F., RICO, F., COLOM, A., CASUSO, I. & SCHEURING, S. 2014. High-speed atomic force microscopy: imaging and force spectroscopy. *FEBS Lett*, 588, 3631-8.
- EGLY, J. M. & COIN, F. 2011. A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair (Amst)*, 10, 714-21.
- ELEY, D. D. & SPIVEY, D. I. 1962. Nucleic acid in the dry state. *Trans. Faraday Soc.*, 58, 411-415.
- ENDO, M., KATSUDA, Y., HIDAKA, K. & SUGIYAMA, H. 2010. A versatile DNA nanochip for direct analysis of DNA base-excision repair. *Angew Chem Int Ed Engl*, 49, 9412-6.
- ERZBERGER, J. P. & WILSON, D. M., 3RD 1999. The role of Mg²⁺ and specific amino acid residues in the catalytic reaction of the major human abasic endonuclease: new insights from EDTA-resistant incision of acyclic abasic site analogs and site-directed mutagenesis. *J Mol Biol*, 290, 447-57.
- EVANS, M. D., DIZDAROGLU, M. & COOKE, M. S. 2004. Oxidative DNA damage and disease: induction, repair and significance. *Mutat Res*, 567, 1-61.
- FAGBEMI, A. F., ORELLI, B. & SCHARER, O. D. 2011. Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair (Amst)*, 10, 722-9.
- FAN, L., ARVAI, A. S., COOPER, P. K., IWAI, S., HANAOKA, F. & TAINER, J. A. 2006. Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair. *Mol Cell*, 22, 27-37.
- FAN, L., FUSS, J. O., CHENG, Q. J., ARVAI, A. S., HAMMEL, M., ROBERTS, V. A., COOPER, P. K. & TAINER, J. A. 2008. XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell*, 133, 789-800.
- FASMAN, G. D. 1975. *Handbook of Biochemistry and Molecular Biology.*, CRC Press.
- FOTIADIS, D., HASLER, L., MULLER, D. J., STAHLBERG, H., KISTLER, J. & ENGEL, A. 2000. Surface tongue-and-groove contours on lens MIP facilitate cell-to-cell adherence. *J Mol Biol*, 300, 779-89.
- FOUSTERI, M., VERMEULEN, W., VAN ZEELAND, A. A. & MULLENDERS, L. H. 2006. Cockayne syndrome A and B proteins differentially regulate recruitment of chromatin remodeling and repair factors to stalled RNA polymerase II in vivo. *Mol Cell*, 23, 471-82.
- FRIEDBERG, E. C. 2001. How nucleotide excision repair protects against cancer. *Nat Rev Cancer*, 1, 22-33.
- FRIEDBERG, E. C. 2003. DNA damage and repair. *Nature*, 421, 436-40.

- FRIEDMAN, J. I., MAJUMDAR, A. & STIVERS, J. T. 2009. Nontarget DNA binding shapes the dynamic landscape for enzymatic recognition of DNA damage. *Nucleic Acids Res*, 37, 3493-500.
- FRIEDMAN, J. I. & STIVERS, J. T. 2010. Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry*, 49, 4957-67.
- FROMME, J. C. & VERDINE, G. L. 2002. Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM. *Nat Struct Biol*, 9, 544-52.
- FRONCZEK, D. N., QUAMMEN, C., WANG, H., KISKER, C., SUPERFINE, R., TAYLOR, R., ERIE, D. A. & TESSMER, I. 2011. High accuracy FIONA-AFM hybrid imaging. *Ultramicroscopy*, 111, 350-5.
- FUXREITER, M., LUO, N., JEDLOVSZKY, P., SIMON, I. & OSMAN, R. 2002. Role of base flipping in specific recognition of damaged DNA by repair enzymes. *J Mol Biol*, 323, 823-34.
- GAO, Y., KATYAL, S., LEE, Y., ZHAO, J., REHG, J. E., RUSSELL, H. R. & MCKINNON, P. J. 2011. DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. *Nature*, 471, 240-4.
- GILLET, L. C. & SCHARER, O. D. 2006. Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev*, 106, 253-76.
- GIRARD, P. M., GUIBOURT, N. & BOITEUX, S. 1997. The Ogg1 protein of *Saccharomyces cerevisiae*: a 7,8-dihydro-8-oxoguanine DNA glycosylase/AP lyase whose lysine 241 is a critical residue for catalytic activity. *Nucleic Acids Res*, 25, 3204-11.
- GRANDBOIS, M., DETTMANN, W., BENOIT, M. & GAUB, H. E. 2000. Affinity imaging of red blood cells using an atomic force microscope. *J Histochem Cytochem*, 48, 719-24.
- GROGAN, D. W. 2000. The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol*, 8, 180-5.
- GROGAN, D. W. 2004. Stability and repair of DNA in hyperthermophilic Archaea. *Curr Issues Mol Biol*, 6, 137-44.
- GROLLMAN, A. P. & MORIYA, M. 1993. Mutagenesis by 8-oxoguanine: an enemy within. *Trends Genet*, 9, 246-9.
- GUEST, C. R., HOCHSTRASSER, R. A., SOWERS, L. C. & MILLAR, D. P. 1991. Dynamics of mismatched base pairs in DNA. *Biochemistry*, 30, 3271-9.
- HALFORD, S. E. & MARKO, J. F. 2004. How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res*, 32, 3040-52.
- HALLETT, P., OFFER, G. & MILES, M. J. 1995. Atomic force microscopy of the myosin molecule. *Biophys J*, 68, 1604-6.
- HANAWALT, P. C. & SPIVAK, G. 2008. Transcription-coupled DNA repair: two decades of progress and surprises. *Nat Rev Mol Cell Biol*, 9, 958-70.
- HANSMA, H. G. & HOH, J. H. 1994. Biomolecular imaging with the atomic force microscope. *Annu Rev Biophys Biomol Struct*, 23, 115-39.
- HARDELAND, U., BENTELE, M., JIRICNY, J. & SCHAR, P. 2000. Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. *J Biol Chem*, 275, 33449-56.
- HASHIGUCHI, K., STUART, J. A., DE SOUZA-PINTO, N. C. & BOHR, V. A. 2004. The C-terminal alphaO helix of human Ogg1 is essential for 8-oxoguanine DNA glycosylase activity: the

- mitochondrial beta-Ogg1 lacks this domain and does not have glycosylase activity. *Nucleic Acids Res*, 32, 5596-608.
- HASHIMOTO, H., HONG, S., BHAGWAT, A. S., ZHANG, X. & CHENG, X. 2012. Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation. *Nucleic Acids Res*, 40, 10203-14.
- HASHIMOTO, H., ZHANG, X. & CHENG, X. 2013. Activity and crystal structure of human thymine DNA glycosylase mutant N140A with 5-carboxylcytosine DNA at low pH. *DNA Repair (Amst)*, 12, 535-40.
- HE, Y. F., LI, B. Z., LI, Z., LIU, P., WANG, Y., TANG, Q., DING, J., JIA, Y., CHEN, Z., LI, L., SUN, Y., LI, X., DAI, Q., SONG, C. X., ZHANG, K., HE, C. & XU, G. L. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, 333, 1303-7.
- HEGDE, M. L., HAZRA, T. K. & MITRA, S. 2008. Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res*, 18, 27-47.
- HEGNER, M., WAGNER, P. & SEMENZA, G. 1993. Immobilizing DNA on gold via thiol modification for atomic force microscopy imaging in buffer solutions. *FEBS Lett*, 336, 452-6.
- HERDER, P. C., CLAEISSON, P. M. & HERDER, C. E. 1987. Adsorption of Cationic Surfactants on Muscovite Mica as Quantified by Means of Esca. *Journal of Colloid and Interface Science*, 119, 155-167.
- HESS, M. T., SCHWITTER, U., PETRETTA, M., GIESE, B. & NAEGELI, H. 1997. Bipartite substrate discrimination by human nucleotide excision repair. *Proc Natl Acad Sci U S A*, 94, 6664-9.
- HOEIJMAKERS, J. H. 2001. Genome maintenance mechanisms for preventing cancer. *Nature*, 411, 366-74.
- HOEIJMAKERS, J. H. 2009. DNA damage, aging, and cancer. *N Engl J Med*, 361, 1475-85.
- HOLLIS, T., ICHIKAWA, Y. & ELLENBERGER, T. 2000. DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, *Escherichia coli* AlkA. *EMBO J*, 19, 758-66.
- HOLZ, B., KLIMASAUSKAS, S., SERVA, S. & WEINHOLD, E. 1998. 2-Aminopurine as a fluorescent probe for DNA base flipping by methyltransferases. *Nucleic Acids Res*, 26, 1076-83.
- HOOGSTEN, K. 1963. Crystal and Molecular Structure of a Hydrogen-Bonded Complex between 1-Methylthymine and 9-Methyladenine. *Acta Crystallographica*, 16, 907-+.
- HORTON, J. K., PRASAD, R., HOU, E. & WILSON, S. H. 2000. Protection against methylation-induced cytotoxicity by DNA polymerase beta-dependent long patch base excision repair. *J Biol Chem*, 275, 2211-8.
- HOU, E. W., PRASAD, R., ASAGOSHI, K., MASAOKA, A. & WILSON, S. H. 2007. Comparative assessment of plasmid and oligonucleotide DNA substrates in measurement of in vitro base excision repair activity. *Nucleic Acids Research*, 35.
- HU, T., GROSBERG, A. Y. & SHKLOVSKII, B. I. 2006. How proteins search for their specific sites on DNA: the role of DNA conformation. *Biophys J*, 90, 2731-44.
- HUANG, J. C., SVOBODA, D. L., REARDON, J. T. & SANCAR, A. 1992. Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc Natl Acad Sci U S A*, 89, 3664-8.

- IMHOF, P. & ZAHRAN, M. 2013. The effect of a G:T mispair on the dynamics of DNA. *PLoS One*, 8, e53305.
- ITO, S., KURAOKA, I., CHYMKOWITCH, P., COMPE, E., TAKEDACHI, A., ISHIGAMI, C., COIN, F., EGLY, J. M. & TANAKA, K. 2007. XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: implications for Cockayne syndrome in XP-G/CS patients. *Mol Cell*, 26, 231-43.
- JAIN, V., HILTON, B., LIN, B., PATNAIK, S., LIANG, F., DARIAN, E., ZOU, Y., MACKERELL, A. D., JR. & CHO, B. P. 2013. Unusual sequence effects on nucleotide excision repair of arylamine lesions: DNA bending/distortion as a primary recognition factor. *Nucleic Acids Res*, 41, 869-80.
- JANICIJEVIC, 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 (Amst)*, 2, 325-36.
- JIANG, Y. & MARSZALEK, P. E. 2011. Atomic force microscopy captures MutS tetramers initiating DNA mismatch repair. *EMBO J*, 30, 2881-93.
- JOHNSON, R. T., SQUIRES, S., ELLIOTT, G. C., KOCH, G. L. & RAINBOW, A. J. 1985. Xeroderma pigmentosum D-HeLa hybrids with low and high ultraviolet sensitivity associated with normal and diminished DNA repair ability, respectively. *J Cell Sci*, 76, 115-33.
- KATAN, A. J. & DEKKER, C. 2011. High-speed AFM reveals the dynamics of single biomolecules at the nanometer scale. *Cell*, 147, 979-82.
- KELLER, D. & BUSTAMANTE, C. 1993. Attaching molecules to surfaces for scanning probe microscopy. *Biophys J*, 64, 896-7.
- KELLY, S. M., JESS, T. J. & PRICE, N. C. 2005. How to study proteins by circular dichroism. *Biochim Biophys Acta*, 1751, 119-39.
- KELMAN, Z. & WHITE, M. F. 2005. Archaeal DNA replication and repair. *Curr Opin Microbiol*, 8, 669-76.
- KIM, S. T. & SANCAR, A. 1995. Photorepair of nonadjacent pyrimidine dimers by DNA photolyase. *Photochem Photobiol*, 61, 171-4.
- KUBOTA, Y., NASH, R. A., KLUNGLAND, A., SCHAR, P., BARNES, D. E. & LINDAHL, T. 1996. Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein. *EMBO J*, 15, 6662-70.
- KUPER, J., BRAUN, C., ELIAS, A., MICHELS, G., SAUER, F., SCHMITT, D. R., POTERSZMAN, A., EGLY, J. M. & KISKER, C. 2014. In TFIIH, XPD helicase is exclusively devoted to DNA repair. *PLoS Biol*, 12, e1001954.
- KUPER, J. & KISKER, C. 2012. Damage recognition in nucleotide excision DNA repair. *Curr Opin Struct Biol*, 22, 88-93.
- KUPER, J., WOLSKI, S. C., MICHELS, G. & KISKER, C. 2012. Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *EMBO J*, 31, 494-502.
- KYPR, J., KEJNOVSKA, I., RENCIOUK, D. & VORLICKOVA, M. 2009. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res*, 37, 1713-25.
- LAEMMLI, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680-5.
- LEE, J. H., PARK, C. J., SHIN, J. S., IKEGAMI, T., AKUTSU, H. & CHOI, B. S. 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, 2474-81.
- LEHMANN, A. R. 2001. The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases. *Genes Dev*, 15, 15-23.
- LEVIN, D. S., MCKENNA, A. E., MOTYCKA, T. A., MATSUMOTO, Y. & TOMKINSON, A. E. 2000. Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. *Curr Biol*, 10, 919-22.
- LINDAHL, T. 1993. Instability and decay of the primary structure of DNA. *Nature*, 362, 709-15.
- LINDAHL, T. & WOOD, R. D. 1999. Quality control by DNA repair. *Science*, 286, 1897-905.
- LIU, H., RUDOLF, J., JOHNSON, K. A., MCMAHON, S. A., OKE, M., CARTER, L., MCROBBIE, A. M., BROWN, S. E., NAISMITH, J. H. & WHITE, M. F. 2008. Structure of the DNA repair helicase XPD. *Cell*, 133, 801-12.
- LOHMAN, T. M. & BUJALOWSKI, W. 1991. Thermodynamic methods for model-independent determination of equilibrium binding isotherms for protein-DNA interactions: spectroscopic approaches to monitor binding. *Methods Enzymol*, 208, 258-90.
- LYUBCHENKO, Y. L. 2014. Nanoscale Nucleosome Dynamics Assessed with Time-lapse AFM. *Biophys Rev*, 6, 181-190.
- LYUBCHENKO, Y. L. & SHLYAKHTENKO, L. S. 2009. AFM for analysis of structure and dynamics of DNA and protein-DNA complexes. *Methods*, 47, 206-13.
- MA, X., HONG, Y., HAN, W., SHENG, D., NI, J., HOU, G. & SHEN, Y. 2011. Single-stranded DNA binding activity of XPBI, but not XPBII, from *Sulfolobus tokodaii* causes double-stranded DNA melting. *Extremophiles*, 15, 67-76.
- MAILLARD, O., CAMENISCH, U., CLEMENT, F. C., BLAGOEV, K. B. & NAEGELI, H. 2007. DNA repair triggered by sensors of helical dynamics. *Trends Biochem Sci*, 32, 494-9.
- MAITI, A. & DROHAT, A. C. 2011. Dependence of substrate binding and catalysis on pH, ionic strength, and temperature for thymine DNA glycosylase: Insights into recognition and processing of G.T mismatches. *DNA Repair (Amst)*, 10, 545-53.
- MAITI, A., MORGAN, M. T. & DROHAT, A. C. 2009. Role of two strictly conserved residues in nucleotide flipping and N-glycosylic bond cleavage by human thymine DNA glycosylase. *J Biol Chem*, 284, 36680-8.
- MAITI, A., MORGAN, M. T., POZHARSKI, E. & DROHAT, A. C. 2008. Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition. *Proc Natl Acad Sci U S A*, 105, 8890-5.
- MAITI, A., NOON, M. S., MACKERELL, A. D., JR., POZHARSKI, E. & DROHAT, A. C. 2012. Lesion processing by a repair enzyme is severely curtailed by residues needed to prevent aberrant activity on undamaged DNA. *Proc Natl Acad Sci U S A*, 109, 8091-6.
- MARTEIJN, J. A., LANS, H., VERMEULEN, W. & HOEIJMAKERS, J. H. 2014. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol*, 15, 465-81.
- MASUTANI, C., SUGASAWA, K., YANAGISAWA, J., SONOYAMA, T., UI, M., ENOMOTO, T., TAKIO, K., TANAKA, K., VAN DER SPEK, P. J., BOOTSMAN, D. & ET AL. 1994. Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *EMBO J*, 13, 1831-43.

- MATHIEU, N., KACZMAREK, N. & NAEGELI, H. 2010. Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proc Natl Acad Sci U S A*, 107, 17545-50.
- MATHIEU, N., KACZMAREK, N., RUTHEMANN, P., LUCH, A. & NAEGELI, H. 2013. DNA quality control by a lesion sensor pocket of the xeroderma pigmentosum group D helicase subunit of TFIIH. *Curr Biol*, 23, 204-12.
- MCATEER, K., JING, Y., KAO, J., TAYLOR, J. S. & KENNEDY, M. A. 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, 1013-32.
- MCCULLOUGH, A. K., DODSON, M. L. & LLOYD, R. S. 1999. Initiation of base excision repair: glycosylase mechanisms and structures. *Annu Rev Biochem*, 68, 255-85.
- MEIJER, M., KARIMI-BUSHERI, F., HUANG, T. Y., WEINFELD, M. & YOUNG, D. 2002. Pnk1, a DNA kinase/phosphatase required for normal response to DNA damage by gamma-radiation or camptothecin in *Schizosaccharomyces pombe*. *J Biol Chem*, 277, 4050-5.
- MICHAELS, M. L. & MILLER, J. H. 1992. The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J Bacteriol*, 174, 6321-5.
- MILLAR, D. P. 1996. Time-resolved fluorescence spectroscopy. *Curr Opin Struct Biol*, 6, 637-42.
- MIN, J. H. & PAVLETICH, N. P. 2007. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature*, 449, 570-5.
- MOGGS, J. G., YAREMA, K. J., ESSIGMANN, J. M. & WOOD, R. D. 1996. Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct. *J Biol Chem*, 271, 7177-86.
- MOLDOVAN, G. L., PFANDER, B. & JENTSCH, S. 2007. PCNA, the maestro of the replication fork. *Cell*, 129, 665-79.
- MOLLER, C., BULDT, G., DENCHER, N. A., ENGEL, A. & MULLER, D. J. 2000. Reversible loss of crystallinity on photobleaching purple membrane in the presence of hydroxylamine. *J Mol Biol*, 301, 869-79.
- MORGAN, M. T., BENNETT, M. T. & DROHAT, A. C. 2007. Excision of 5-halogenated uracils by human thymine DNA glycosylase. Robust activity for DNA contexts other than CpG. *J Biol Chem*, 282, 27578-86.
- MORGAN, M. T., MAITI, A., FITZGERALD, M. E. & DROHAT, A. C. 2011. Stoichiometry and affinity for thymine DNA glycosylase binding to specific and nonspecific DNA. *Nucleic Acids Res*, 39, 2319-29.
- MOSER, J., KOOL, H., GIAKZIDIS, I., CALDECOTT, K., MULLENDERS, L. H. & FOUSTERI, M. I. 2007. Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Mol Cell*, 27, 311-23.
- MOU, J., SHENG, S., HO, R. & SHAO, Z. 1996. Chaperonins GroEL and GroES: views from atomic force microscopy. *Biophys J*, 71, 2213-21.
- MU, D., HSU, D. S. & SANCAR, A. 1996. Reaction mechanism of human DNA repair excision nuclease. *J Biol Chem*, 271, 8285-94.
- MULLER, D. J., AMREIN, M. & ENGEL, A. 1997. Adsorption of biological molecules to a solid support for scanning probe microscopy. *J Struct Biol*, 119, 172-88.

- MYSZKA, D. G. 2000. Kinetic, equilibrium, and thermodynamic analysis of macromolecular interactions with BIACORE. *Methods Enzymol*, 323, 325-40.
- NABEL, C. S., MANNING, S. A. & KOHLI, R. M. 2012. The curious chemical biology of cytosine: deamination, methylation, and oxidation as modulators of genomic potential. *ACS Chem Biol*, 7, 20-30.
- NAEGELI, H., BARDWELL, L. & FRIEDBERG, E. C. 1992. 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, 392-8.
- NAKANO, S., UOTANI, Y., UENISHI, K., FUJII, M. & SUGIMOTO, N. 2005. DNA base flipping by a base pair-mimic nucleoside. *Nucleic Acids Res*, 33, 7111-9.
- NASH, H. M., BRUNER, S. D., SCHARER, O. D., KAWATE, T., ADDONA, T. A., SPOONER, E., LANE, W. S. & VERDINE, G. L. 1996. Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. *Curr Biol*, 6, 968-80.
- NASH, H. M., LU, R., LANE, W. S. & VERDINE, G. L. 1997. The critical active-site amine of the human 8-oxoguanine DNA glycosylase, hOgg1: direct identification, ablation and chemical reconstitution. *Chem Biol*, 4, 693-702.
- NELSON, S. R., DUNN, A. R., KATHE, S. D., WARSHAW, D. M. & WALLACE, S. S. 2014. Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases. *Proc Natl Acad Sci U S A*, 111, E2091-9.
- NIKOLOVA, E. N., KIM, E., WISE, A. A., O'BRIEN, P. J., ANDRICIOAEI, I. & AL-HASHIMI, H. M. 2011. Transient Hoogsteen base pairs in canonical duplex DNA. *Nature*, 470, 498-502.
- NISHI, R., OKUDA, Y., WATANABE, E., MORI, T., IWAI, S., MASUTANI, C., SUGASAWA, K. & HANAOKA, F. 2005. Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Mol Cell Biol*, 25, 5664-74.
- NISHINO, T., KOMORI, K., ISHINO, Y. & MORIKAWA, K. 2003. X-ray and biochemical anatomy of an archaeal XPF/Rad1/Mus81 family nuclease: similarity between its endonuclease domain and restriction enzymes. *Structure*, 11, 445-57.
- NORMAN, D. P., CHUNG, S. J. & VERDINE, G. L. 2003. Structural and biochemical exploration of a critical amino acid in human 8-oxoguanine glycosylase. *Biochemistry*, 42, 1564-72.
- ODA, M., FURUKAWA, K., OGATA, K., SARAI, A. & NAKAMURA, H. 1998. Thermodynamics of specific and non-specific DNA binding by the c-Myb DNA-binding domain. *J Mol Biol*, 276, 571-90.
- OGI, T., LIMSIRICHAIKUL, S., OVERMEER, R. M., VOLKER, M., TAKENAKA, K., CLONEY, R., NAKAZAWA, Y., NIIMI, A., MIKI, Y., JASPERS, N. G., MULLENDERS, L. H., YAMASHITA, S., FOUSTERI, M. I. & LEHMANN, A. R. 2010. Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Mol Cell*, 37, 714-27.
- OGRUNC, M., BECKER, D. F., RAGSDALE, S. W. & SANCAR, A. 1998. Nucleotide excision repair in the third kingdom. *J Bacteriol*, 180, 5796-8.
- OHNESORGE, F. M., HORBER, J. K., HABERLE, W., CZERNY, C. P., SMITH, D. P. & BINNIG, G. 1997. AFM review study on pox viruses and living cells. *Biophys J*, 73, 2183-94.
- OKSENYCH, V., BERNARDES DE JESUS, B., ZHOVMER, A., EGLY, J. M. & COIN, F. 2009. Molecular insights into the recruitment of TFIIH to sites of DNA damage. *EMBO J*, 28, 2971-80.
- PAPADOPOULOS, T. & SOYKAN, T. 2011. The role of collybistin in gephyrin clustering at inhibitory synapses: facts and open questions. *Front Cell Neurosci*, 5, 11.

- PARIKH, S. S., WALCHER, G., JONES, G. D., SLUPPHAUG, G., KROKAN, H. E., BLACKBURN, G. M. & TAINER, J. A. 2000. Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc Natl Acad Sci U S A*, 97, 5083-8.
- PARKER, J. B., BIANCHET, M. A., KROSKY, D. J., FRIEDMAN, J. I., AMZEL, L. M. & STIVERS, J. T. 2007. Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. *Nature*, 449, 433-7.
- PEARL, L. H. 2000. Structure and function in the uracil-DNA glycosylase superfamily. *Mutat Res*, 460, 165-81.
- PETIT, C. & SANCAR, A. 1999. Nucleotide excision repair: from E. coli to man. *Biochimie*, 81, 15-25.
- PFEIFFER, H. G. & LIEBHAFSKY, H. A. 1951. The origins of Beer's law. *Journal of Chemical Education*, 28, 123-125.
- PRASAD, R., DIANOV, G. L., BOHR, V. A. & WILSON, S. H. 2000. FEN1 stimulation of DNA polymerase beta mediates an excision step in mammalian long patch base excision repair. *J Biol Chem*, 275, 4460-6.
- PRASAD, R., SHOCK, D. D., BEARD, W. A. & WILSON, S. H. 2010. Substrate channeling in mammalian base excision repair pathways: passing the baton. *J Biol Chem*, 285, 40479-88.
- QI, Y., SPONG, M. C., NAM, K., BANERJEE, A., JIRALERSPONG, S., KARPLUS, M. & VERDINE, G. L. 2009. Encounter and extrusion of an intrahelical lesion by a DNA repair enzyme. *Nature*, 462, 762-6.
- RADOM, C. T., BANERJEE, A. & VERDINE, G. L. 2007. Structural characterization of human 8-oxoguanine DNA glycosylase variants bearing active site mutations. *J Biol Chem*, 282, 9182-94.
- RAMSTEIN, J. & LAVERY, R. 1988. Energetic coupling between DNA bending and base pair opening. *Proc Natl Acad Sci U S A*, 85, 7231-5.
- RANEY, K. D., SOWERS, L. C., MILLAR, D. P. & BENKOVIC, S. J. 1994. A fluorescence-based assay for monitoring helicase activity. *Proc Natl Acad Sci U S A*, 91, 6644-8.
- RATCLIFF, G. C. & ERIE, D. A. 2001. A novel single-molecule study to determine protein--protein association constants. *J Am Chem Soc*, 123, 5632-5.
- REARDON, J. T., GE, H., GIBBS, E., SANCAR, A., HURWITZ, J. & PAN, Z. Q. 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, 6482-7.
- REARDON, J. T. & SANCAR, A. 2003. Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes Dev*, 17, 2539-51.
- RECORD, M. T., JR., HA, J. H. & FISHER, M. A. 1991. Analysis of equilibrium and kinetic measurements to determine thermodynamic origins of stability and specificity and mechanism of formation of site-specific complexes between proteins and helical DNA. *Methods Enzymol*, 208, 291-343.
- RICHARDS, J. D., CUBEDDU, L., ROBERTS, J., LIU, H. & WHITE, M. F. 2008. The archaeal XPB protein is a ssDNA-dependent ATPase with a novel partner. *J Mol Biol*, 376, 634-44.
- RIEDL, T., HANAOKA, F. & EGLY, J. M. 2003. The comings and goings of nucleotide excision repair factors on damaged DNA. *EMBO J*, 22, 5293-303.

- ROBERTS, J. A., BELL, S. D. & WHITE, M. F. 2003. An archaeal XPF repair endonuclease dependent on a heterotrimeric PCNA. *Mol Microbiol*, 48, 361-71.
- ROMANO, C. A., SONTZ, P. A. & BARTON, J. K. 2011. Mutants of the base excision repair glycosylase, endonuclease III: DNA charge transport as a first step in lesion detection. *Biochemistry*, 50, 6133-45.
- ROTH, H. M., ROMER, J., GRUNDLER, V., VAN HOUTEN, B., KISKER, C. & TESSMER, I. 2012. XPB helicase regulates DNA incision by the *Thermoplasma acidophilum* endonuclease Bax1. *DNA Repair (Amst)*, 11, 286-93.
- ROUILLON, C. & WHITE, M. F. 2011. The evolution and mechanisms of nucleotide excision repair proteins. *Res Microbiol*, 162, 19-26.
- RUDOLF, J., ROUILLON, C., SCHWARZ-LINEK, U. & WHITE, M. F. 2010. The helicase XPD unwinds bubble structures and is not stalled by DNA lesions removed by the nucleotide excision repair pathway. *Nucleic Acids Res*, 38, 931-41.
- RUEPP, A., GRAML, W., SANTOS-MARTINEZ, M. L., KORETKE, K. K., VOLKER, C., MEWES, H. W., FRISHMAN, D., STOCKER, S., LUPAS, A. N. & BAUMEISTER, W. 2000. The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature*, 407, 508-13.
- SANTOS, N. C. & CASTANHO, M. A. 2004. An overview of the biophysical applications of atomic force microscopy. *Biophys Chem*, 107, 133-49.
- SCHERMERHORN, K. M. & DELANEY, S. 2014. A chemical and kinetic perspective on base excision repair of DNA. *Acc Chem Res*, 47, 1238-46.
- SCHEURING, S., FOTIADIS, D., MÖLLER, C., MÜLLER, S. A., ENGEL, A. & MÜLLER, D. J. 2001. Single Proteins Observed by Atomic Force Microscopy. *Single Mol.*, 2, 59-67.
- SCHEURING, S., RINGLER, P., BORGNIA, M., STAHLBERG, H., MULLER, D. J., AGRE, P. & ENGEL, A. 1999. High resolution AFM topographs of the *Escherichia coli* water channel aquaporin Z. *EMBO J*, 18, 4981-7.
- SCHUMACHER, B., GARINIS, G. A. & HOEIJMAKERS, J. H. 2008. Age to survive: DNA damage and aging. *Trends Genet*, 24, 77-85.
- SCHWERTMAN, P., LAGAROU, A., DEKKERS, D. H., RAAMS, A., VAN DER HOEK, A. C., LAFFEYER, C., HOEIJMAKERS, J. H., DEMMERS, J. A., FOSTER, M., VERMEULEN, W. & MARTEIJN, J. A. 2012. UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nat Genet*, 44, 598-602.
- SCRIMA, A., KONICKOVA, R., CZYZEWSKI, B. K., KAWASAKI, Y., JEFFREY, P. D., GROISMAN, R., NAKATANI, Y., IWAI, S., PAVLETICH, N. P. & THOMA, N. H. 2008. Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. *Cell*, 135, 1213-23.
- SCZEPANSKI, J. T., WONG, R. S., MCKNIGHT, J. N., BOWMAN, G. D. & GREENBERG, M. M. 2010. Rapid DNA-protein cross-linking and strand scission by an abasic site in a nucleosome core particle. *Proc Natl Acad Sci U S A*, 107, 22475-80.
- SEELERT, H., POETSCH, A., DENCHER, N. A., ENGEL, A., STAHLBERG, H. & MULLER, D. J. 2000. Structural biology. Proton-powered turbine of a plant motor. *Nature*, 405, 418-9.
- SEIBERT, E., ROSS, J. B. & OSMAN, R. 2002. Role of DNA flexibility in sequence-dependent activity of uracil DNA glycosylase. *Biochemistry*, 41, 10976-84.
- SEO, D. S. & LEE, J. K. 2008. AFM analysis of anisotropic dissolution in dense hydroxyapatite. *Ultramicroscopy*, 108, 1157-62.

- SEONG, G. H., YANAGIDA, Y., AIZAWA, M. & KOBATAKE, E. 2002. Atomic force microscopy identification of transcription factor NFkappaB bound to streptavidin-pin-holding DNA probe. *Anal Biochem*, 309, 241-7.
- SHARMA, M., PREDEUS, A. V., MUKHERJEE, S. & FEIG, M. 2013. DNA bending propensity in the presence of base mismatches: implications for DNA repair. *J Phys Chem B*, 117, 6194-205.
- SHIBUTANI, S., TAKESHITA, M. & GROLLMAN, A. P. 1991. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature*, 349, 431-4.
- SHLYAKHTENKO, L. S., LUSHNIKOV, A. Y., LI, M., LACKEY, L., HARRIS, R. S. & LYUBCHENKO, Y. L. 2011. Atomic force microscopy studies provide direct evidence for dimerization of the HIV restriction factor APOBEC3G. *J Biol Chem*, 286, 3387-95.
- SHLYAKHTENKO, L. S., LUSHNIKOV, A. Y., MIYAGI, A., LI, M., HARRIS, R. S. & LYUBCHENKO, Y. L. 2012a. Nanoscale structure and dynamics of ABOBEC3G complexes with single-stranded DNA. *Biochemistry*, 51, 6432-40.
- SHLYAKHTENKO, L. S., LUSHNIKOV, A. Y., MIYAGI, A. & LYUBCHENKO, Y. L. 2012b. Specificity of binding of single-stranded DNA-binding protein to its target. *Biochemistry*, 51, 1500-9.
- SIMSEK, D., FURDA, A., GAO, Y., ARTUS, J., BRUNET, E., HADJANTONAKIS, A. K., VAN HOUTEN, B., SHUMAN, S., MCKINNON, P. J. & JASIN, M. 2011. Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature*, 471, 245-8.
- SINGH, S., TURINA, P., BUSTAMANTE, C. J., KELLER, D. J. & CAPALDI, R. 1996. Topographical structure of membrane-bound Escherichia coli F1F0 ATP synthase in aqueous buffer. *FEBS Lett*, 397, 30-4.
- SINGHAL, R. K., PRASAD, R. & WILSON, S. H. 1995. DNA polymerase beta conducts the gap-filling step in uracil-initiated base excision repair in a bovine testis nuclear extract. *J Biol Chem*, 270, 949-57.
- SINGLETON, M. R., DILLINGHAM, M. S. & WIGLEY, D. B. 2007. Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem*, 76, 23-50.
- SJÖBACK, R., NYGREN, J. & KUBISTA, M. 1995. Absorption and fluorescence properties of fluorescein. *Spectrochimica Acta*, Part A 51, L7-L21.
- SONTZ, P. A., MUI, T. P., FUSS, J. O., TAINER, J. A. & BARTON, J. K. 2012. DNA charge transport as a first step in coordinating the detection of lesions by repair proteins. *Proc Natl Acad Sci U S A*, 109, 1856-61.
- SOWERS, L. C., FAZAKERLEY, G. V., ERITJA, R., KAPLAN, B. E. & GOODMAN, M. F. 1986. Base pairing and mutagenesis: observation of a protonated base pair between 2-aminopurine and cytosine in an oligonucleotide by proton NMR. *Proc Natl Acad Sci U S A*, 83, 5434-8.
- SOYKAN, T., SCHNEEBERGER, D., TRIA, G., BUECHNER, C., BADER, N., SVERGUN, D., TESSMER, I., POULOPOULOS, A., PAPADOPOULOS, T., VAROQUEAUX, F., SCHINDELIN, H. & BROSE, N. 2014. A conformational switch in collybistin determines the differentiation of inhibitory postsynapses. *EMBO J*, 33, 2113-33.
- STARK, M., MOLLER, C., MULLER, D. J. & GUCKENBERGER, R. 2001. From images to interactions: high-resolution phase imaging in tapping-mode atomic force microscopy. *Biophys J*, 80, 3009-18.

- STEENKEN, S. & JOVANOVIĆ, S. 1997. How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution. . *J. Am. Chem. Soc.*, 119, 617–618.
- STIVERS, J. T. 1998. 2-Aminopurine fluorescence studies of base stacking interactions at abasic sites in DNA: metal-ion and base sequence effects. *Nucleic Acids Res*, 26, 3837-44.
- STIVERS, J. T. & JIANG, Y. L. 2003. A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem Rev*, 103, 2729-59.
- SUGASAWA, K. 2010. Regulation of damage recognition in mammalian global genomic nucleotide excision repair. *Mutat Res*, 685, 29-37.
- SUGASAWA, K., AKAGI, J., NISHI, R., IWAI, S. & HANAOKA, F. 2009. Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning. *Mol Cell*, 36, 642-53.
- 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, 223-32.
- SUGASAWA, K., OKAMOTO, T., SHIMIZU, Y., MASUTANI, C., IWAI, S. & HANAOKA, F. 2001. A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes Dev*, 15, 507-21.
- SUNG, J. S., DEMOTT, M. S. & DEMPSEY, B. 2005. Long-patch base excision DNA repair of 2-deoxyribonolactone prevents the formation of DNA-protein cross-links with DNA polymerase beta. *J Biol Chem*, 280, 39095-103.
- SUZUKI, Y., HIGUCHI, Y., HIZUME, K., YOKOKAWA, M., YOSHIMURA, S. H., YOSHIKAWA, K. & TAKEYASU, K. 2010. Molecular dynamics of DNA and nucleosomes in solution studied by fast-scanning atomic force microscopy. *Ultramicroscopy*, 110, 682-8.
- SWENBERG, J. A., LU, K., MOELLER, B. C., GAO, L., UPTON, P. B., NAKAMURA, J. & STARR, T. B. 2011. Endogenous versus exogenous DNA adducts: their role in carcinogenesis, epidemiology, and risk assessment. *Toxicol Sci*, 120 Suppl 1, S130-45.
- TAPIAS, A., AURIOL, J., FORGET, D., ENZLIN, J. H., SCHARER, O. D., COIN, F., COULOMBE, B. & EGLY, J. M. 2004. Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *J Biol Chem*, 279, 19074-83.
- TAYLOR, E. M. & LEHMANN, A. R. 1998. Conservation of eukaryotic DNA repair mechanisms. *Int J Radiat Biol*, 74, 277-86.
- TAYLOR, M. R., CONRAD, J. A., WAHL, D. & O'BRIEN, P. J. 2011. Kinetic mechanism of human DNA ligase I reveals magnesium-dependent changes in the rate-limiting step that compromise ligation efficiency. *J Biol Chem*, 286, 23054-62.
- TESSMER, I., MELIKISHVILI, M. & FRIED, M. G. 2012. Cooperative cluster formation, DNA bending and base-flipping by O6-alkylguanine-DNA alkyltransferase. *Nucleic Acids Res*, 40, 8296-308.
- TESSMER, I., YANG, Y., ZHAI, J., DU, C., HSIEH, P., HINGORANI, M. M. & ERIE, D. A. 2008. Mechanism of MutS searching for DNA mismatches and signaling repair. *J Biol Chem*, 283, 36646-54.
- THAYER, M. M., AHERN, H., XING, D., CUNNINGHAM, R. P. & TAINER, J. A. 1995. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J*, 14, 4108-20.

- TIRODE, F., BUSSO, D., COIN, F. & EGLY, J. M. 1999. Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7. *Mol Cell*, 3, 87-95.
- TRUGLIO, J. J., CROTEAU, D. L., VAN HOUTEN, B. & KISKER, C. 2006. Prokaryotic nucleotide excision repair: the UvrABC system. *Chem Rev*, 106, 233-52.
- VAN HOUTEN, B., CROTEAU, D. L., DELLAVECCHIA, M. J., WANG, H. & KISKER, C. 2005. 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. *Mutat Res*, 577, 92-117.
- 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, 4196-205.
- 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, 601-11.
- VERMEULEN, W. & FOUSTERI, M. 2013. Mammalian transcription-coupled excision repair. *Cold Spring Harb Perspect Biol*, 5, a012625.
- VIDAL, A. E., HICKSON, I. D., BOITEUX, S. & RADICELLA, J. P. 2001. Mechanism of stimulation of the DNA glycosylase activity of hOGG1 by the major human AP endonuclease: bypass of the AP lyase activity step. *Nucleic Acids Res*, 29, 1285-92.
- 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, 213-24.
- WAGNER, K., MOOLENAAR, G., VAN NOORT, J. & GOOSEN, N. 2009. Single-molecule analysis reveals two separate DNA-binding domains in the Escherichia coli UvrA dimer. *Nucleic Acids Res*, 37, 1962-72.
- WAKASUGI, M., KAWASHIMA, A., MORIOKA, H., LINN, S., SANCAR, A., MORI, T., NIKAIDO, O. & MATSUNAGA, T. 2002. DDB accumulates at DNA damage sites immediately after UV irradiation and directly stimulates nucleotide excision repair. *J Biol Chem*, 277, 1637-40.
- WALKER, R. K., MCCULLOUGH, A. K. & LLOYD, R. S. 2006. Uncoupling of nucleotide flipping and DNA bending by the t4 pyrimidine dimer DNA glycosylase. *Biochemistry*, 45, 14192-200.
- WALLACE, S. S. 2002. Biological consequences of free radical-damaged DNA bases. *Free Radic Biol Med*, 33, 1-14.
- WALTERS, D. A., CLEVELAND, J. P., THOMSON, N. H., HANSMA, P. K., ENDMAN, M. A. W. & GURLEY, G. 1997. Short cantilevers for atomic force microscopy. *Rev. Sci. Instrum.*, 67, 3583-3590.
- WANG, H., YANG, Y., SCHOFIELD, M. J., DU, C., FRIDMAN, Y., LEE, S. D., LARSON, E. D., DRUMMOND, J. T., ALANI, E., HSIEH, P. & ERIE, D. A. 2003. DNA bending and unbending by MutS govern mismatch recognition and specificity. *Proc Natl Acad Sci U S A*, 100, 14822-7.
- WARD, D. C., REICH, E. & STRYER, L. 1969. Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J Biol Chem*, 244, 1228-37.
- WARSHAW, M. M. & TINOCO, I., JR. 1966. Optical properties of sixteen dinucleoside phosphates. *J Mol Biol*, 20, 29-38.

- WEISENHORN, A. L., DRAKE, B., PRATER, C. B., GOULD, S. A., HANSMA, P. K., OHNESORGE, F., EGGER, M., HEYN, S. P. & GAUB, H. E. 1990. Immobilized proteins in buffer imaged at molecular resolution by atomic force microscopy. *Biophys J*, 58, 1251-8.
- WINZOR, D. J. & WILLS, P. R. 1995. Thermodynamic nonideality of enzyme solutions supplemented with inert solutes: yeast hexokinase revisited. *Biophys Chem*, 57, 103-10.
- WOESE, C. R. & FOX, G. E. 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A*, 74, 5088-90.
- WOLSKI, S. C., KUPER, J., HANZELMANN, P., TRUGLIO, J. J., CROTEAU, D. L., VAN HOUTEN, B. & KISKER, C. 2008. Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS Biol*, 6, e149.
- WU, L., ZHOU, H., ZHANG, Q., ZHANG, J., NI, F., LIU, C. & QI, Y. 2010. DNA methylation mediated by a microRNA pathway. *Mol Cell*, 38, 465-75.
- XU, D., EVANS, K. O. & NORDLUND, T. M. 1994. Melting and premelting transitions of an oligomer measured by DNA base fluorescence and absorption. *Biochemistry*, 33, 9592-9.
- XU, S., BEVIS, B. & ARNSDORF, M. F. 2001. The assembly of amyloidogenic yeast sup35 as assessed by scanning (atomic) force microscopy: an analogy to linear colloidal aggregation? *Biophys J*, 81, 446-54.
- YANG, Y., SASS, L. E., DU, C., HSIEH, P. & ERIE, D. A. 2005. Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions. *Nucleic Acids Res*, 33, 4322-34.
- YIN, Y., YANG, L., ZHENG, G., GU, C., YI, C., HE, C., GAO, Y. Q. & ZHAO, X. S. 2014. Dynamics of spontaneous flipping of a mismatched base in DNA duplex. *Proc Natl Acad Sci U S A*, 111, 8043-8.
- YOKOI, M., MASUTANI, C., MAEKAWA, T., SUGASAWA, K., OHKUMA, Y. & HANAOKA, F. 2000. The xeroderma pigmentosum group C protein complex XPC-HR23B plays an important role in the recruitment of transcription factor IIH to damaged DNA. *J Biol Chem*, 275, 9870-5.
- YOKOKAWA, M., WADA, C., ANDO, T., SAKAI, N., YAGI, A., YOSHIMURA, S. H. & TAKEYASU, K. 2006. Fast-scanning atomic force microscopy reveals the ATP/ADP-dependent conformational changes of GroEL. *EMBO J*, 25, 4567-76.
- ZHARKOV, D. O., GOLAN, G., GILBOA, R., FERNANDES, A. S., GERCHMAN, S. E., KYCIA, J. H., RIEGER, R. A., GROLLMAN, A. P. & SHOHAM, G. 2002. Structural analysis of an Escherichia coli endonuclease VIII covalent reaction intermediate. *EMBO J*, 21, 789-800.
- ZHARKOV, D. O., MECHETIN, G. V. & NEVINSKY, G. A. 2010. Uracil-DNA glycosylase: Structural, thermodynamic and kinetic aspects of lesion search and recognition. *Mutat Res*, 685, 11-20.
- ZHARKOV, D. O., SHOHAM, G. & GROLLMAN, A. P. 2003. Structural characterization of the Fpg family of DNA glycosylases. *DNA Repair (Amst)*, 2, 839-62.
- ZOTTER, A., LUIJSTERBURG, M. S., WARMERDAM, D. O., IBRAHIM, S., NIGG, A., VAN CAPPELLEN, W. A., HOEIJMAKERS, J. H., VAN DRIEL, R., VERMEULEN, W. & HOUTSMULLER, A. B. 2006. Recruitment of the nucleotide excision repair endonuclease XPG to sites of UV-induced dna damage depends on functional TFIIH. *Mol Cell Biol*, 26, 8868-79.

8. Supplement

8.1 Abbreviations

2-AP	2-Aminopurine
8-oxoG	8-oxoguanine
(6-4)PP	(6-4) pyrimidine-pyrimidone photoproduct
°C	degree Celsius
μl	microliter
μg	microgram
A	Adenine
AA	Acrylamide
AFM	Atomic Force Microscopy
AlkA	3-methyladenine DNA glycosylase
Amp	Ampicillin
AP	apurinic site
APE1	AP endonuclease 1
APS	Ammonium peroxydisulfate
ATP	Adenosine triphosphate
ATPys	Adenosine 5'-[γ-thio]triphosphate
AUC	Analytical Ultracentrifugation
BAA	Bisacrylamide
BER	Base excision repair
BLI	Biolayer Interferometry
bp	base pair
BSA	Bovine serum albumin
CAK	Cdk-activating kinase
Cam	Chloramphenicol
CD	Circular dichroism
CETN2	centrin 2
CPD	Cyclobutane pyrimidine dimers
CS	Cockayne Syndrome
CSA	Cockayne Syndrome A
CSB	Cockayne Syndrome B
C	Cytosine
CV	Column volume
DDB	Damaged DNA-binding protein
DNA	deoxyribonucleic acid
DNase	desoxyribonuclease
dRP	5'-deoxyribose phosphate
dsDNA	Double stranded DNA
DTT	dithiothreitol
E.coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
EM	Electron microscopy
EMSA	electrophoretic mobility shift assays
ERCC1	Excision repair cross-complementation group 1

F	Fluorescein
FCS	Fluorescence correlation spectroscopy
FEN-1	5'-flap endonuclease 1
FeS	Iron-sulfur
G	Guanine
GDP	Guanosine 5'-diphosphate
GG-NER	Global genomic NER
GTP	Guanosine-5'-triphosphate
HhH	helix-hairpin-helix
hTDG	human Thymine DNA glycosylase
hOgg1	human 8-oxoguanine DNA glycosylase
HS-AFM	high-speed AFM
IC	Interrogation complex
IPTG	Isopropyl- β -thiogalactoside
ITC	Isothermal titration calorimetry
kb	kilobase
kDA	kilo Dalton
LB	Luria broth
Lig	DNA ligase
LP-BER	Long-patch BER
M	molar
mM	millimolar
mRNA	messenger RNA
MWCO	Molecular weight cut off
NEIL	Endonuclease VIII-like protein
NER	Nucleotide excision repair
Nsp	nonspecific
NTH1	Endonuclease III-like protein 1
OD	Optical density
P	Phosphate
PAGE	Polyacrylamide gel electrophoresis
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PEG	Polyethyleneglycol
PNK	polynucleotide kinase
Pol	Polymerase
PUA	3' phosphor- α,β -unsaturated aldehyde
RAD23B	RAD23 homologue B
RFC	replication factor-C
RMS	root mean square
RNA	ribonucleic acid
RPA	replication protein A
RT	Room temperature
S	specificity
SC	search complex
SDS	sodium dodecyl sulfate
SEC	Size exclusion chromatography
SF2	Superfamily 2
SN-BER	Single-nucleotide BER

SP-BER	Short-patch BER
Spec	specific
SPM	Scanning probe microscopies
SPR	surface plasmon resonance
ssDNA	Single stranded DNA
STM	Scanning tunneling microscopy
taXPD	Thermoplasma acidophilum XPD
T	Thymine
THF	Tetrahydrofuran
TCR	Transcription coupled repair
TEMED	tetramethylethylenediamine
TFIIH	Transcriptionfactor II H
TTD	Trichothiodystrophy
U	Uracil
U ^F	2'-deoxy-2'-flouroarabinouridine
UDG	Uracil DNA glycosylase
USP7	ubiquitin-specific-processing protease 7
UV	ultraviolet light
UVSSA	UV-stimulated scaffold protein A
WT	wild type
XP	Xeroderma Pigmentosum
XP(A-G)	Xeroderma pigmentosum complementation group (A-G)
XRCC1	X-ray repair cross-complementing protein 1

8.2. Figure index

Figure 1: Imaging modes in AFM.....	9
Figure 2: Typical setup of a tapping mode atomic force microscope.....	10
Figure 3: Conformational states of a neuro-protein visualized in AFM experiments.....	14
Figure 4: The twofold danger of DNA damages.....	15
Figure 5: Sources and types of DNA damages.....	16
Figure 6: Schematic of the eukaryotic NER mechanism.....	21
Figure 7: Primary and crystal structure of taXPD showing mutated residues as used in this work.....	25
Figure 8: Model for DNA mediated charge transfer.....	26
Figure 9: DNA substrates for lesion recognition studies with taXPD.....	28
Figure 10: Schematics of BER subpathways.....	30
Figure 11: Primary structure and co-crystal structure of hTDG bound to a uracil analog.....	33
Figure 12: Primary structure and co-crystal structure of hOgg1 with 8oxoG flipped in the active site.....	35
Figure 13: DNA substrates used for AFM studies with hTDG.....	36
Figure 14: Illustration of the position distribution of protein binding to a single target site in DNA.....	47
Figure 15: Schematics of DNA substrate preparation for AFM imaging experiments.....	51
Figure 16: Assaying DNA substrate preparation.....	53
Figure 17: Principle of Biolayer Interferometry.....	57
Figure 18: BLI binding assay.....	58
Figure 19: 2-AP fluorescence base flipping assay.....	59
Figure 20: Purification of XPD wt from <i>T. acidophilum</i>	63
Figure 21: CD spectra of taXPD wt and variants.....	64
Figure 22: Gel-based helicase assay of taXPD.....	64
Figure 23: DNA binding of taXPD in AFM experiments.....	67
Figure 24: Lesion binding of taXPD in AFM.....	68
Figure 25: Specificity of taXPD for different DNA target sites.....	70
Figure 26: Localization specificity and binding affinities of taXPD R567W for different DNA substrates.....	72
Figure 27: Specificity of taXPD K170A for the DNA substrate fluorescein/5' bubble.....	73
Figure 28: Different strand selectivity of taXPD for bulky fluorescein and CPD damages.....	74
Figure 29: Intrinsic DNA bending of nonspecific and specific DNA substrates in AFM.....	76
Figure 30: Model of the XPD-DNA complex.....	77
Figure 31: DNA bend angle distributions of taXPD with various DNA substrates.....	79

Figure 32: AFM experiments with taXPD variant K170A.....	82
Figure 33: DNA bend angle distribution of the taXPD variant R567W.....	82
Figure 34: Model for lesion recognition by taXPD.	83
Figure 35: Lesion specificity of hTDG N140A at G:U and G:T mismatches.....	85
Figure 36: DNA bending by hTDG N140A at G:U and G:T mismatches.....	87
Figure 37: Lesion specificity of hTDG R275A at G:U ^F mismatches.	88
Figure 38: DNA bending of hTDG R275A at G:U ^F mismatches.....	89
Figure 39: 2-AP fluorescence based base flipping assay of hTDG at G:U ^F mismatches.	90
Figure 40: DNA bending of hTDG at nonspecific DNA sites.	91
Figure 41: intrinsic bending of two nonspecific DNA substrates (549 bp and 1818 bp) in the absence of protein.....	92
Figure 42: 2-AP fluorescence based base flipping assay of hTDG at nonspecific DNA.	94
Figure 43: Stoichiometry of hTDG in AFM experiments.....	95
Figure 44: AFM volume analysis of hTDG-DNA complexes.	96
Figure 45: Correlation between DNA bending and AFM volume of hTDG-DNA complexes.....	97
Figure 46: Model for lesion search and recognition by hTDG.....	99
Figure 47: Characterization of hOgg1-DNA complexes by AFM and fluorescence experiments.....	102

8.3. Table index

Table 1: Feasibility comparison of AFM with other types of microscopes.	12
Table 2: Distribution of selected DNA repair proteins in sequenced archaeal genomes.	23
Table 3: Instruments.	37
Table 4: Disposable Materials.	38
Table 5: Kits.	39
Table 6: Computer software.....	39
Table 7: Media for bacterial cell culture.	39
Table 8: Media additives.	40
Table 9: Buffers and solutions.	40
Table 10: Enzymes.	41
Table 11: DNA substrates for AFM and BLI experiments with taXPD.	42
Table 12: DNA substrates for hOgg1 and hTDG.	43
Table 13: Bacterial strains.	43
Table 14: Plasmids.	43
Table 15: Nearest neighbor values for $\epsilon 260$ of dNTPs.	52
Table 16: DNA binding affinities of taXPD for different DNA substrates.	65
Table 17: AFM binding specificities of taXPD for different DNA substrates and under various conditions. .	69
Table 18: DNA bend angle analysis of taXPD-DNA complexes under different conditions in AFM.....	80
Table 19: AFM binding specificities of hTDG variants to different DNA substrates.	85
Table 20: DNA bend angle analysis for hTDG-DNA complexes and free DNA substrates.....	93
Table 21: Quantification of 2-AP fluorescence based base flipping assay for hTDG proteins.	94
Table 22: DNA bend angle analysis for hOgg1-DNA complexes and free DNA substrate 8oxoG:C.	103
Table 23: Quantification of 2-AP fluorescence based base flipping assay for hOgg1.	103

8.4. Affidavit

I hereby confirm that my thesis entitled 'single molecule studies of DNA repair processes' is the result of my own work. I did not receive any help or support from commercial consultants. All sources and / or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Place, Date Signature

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation ‚Einzelmolekül-Studien von DNA Reparatur Prozessen‘ eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ort, Datum Unterschrift

8.5. Publications

Claudia N. Buechner, Atanu Maiti, Alexander Drohat, Ingrid Tessmer. Base flipping in lesion search and recognition by the human thymine DNA glycosylase revealed by single molecule imaging (*submitted to Nucleic Acids Research*).

Soykan T, Schneeberger D, Tria G, Buechner C, Bader N, Svergun D, Tessmer I, Pouloupoulos A, Papadopoulos T, Varoqueaux F, Schindelin H, Brose N. A conformational switch in collybistin determines the differentiation of inhibitory postsynapses. *EMBO J.* 33(18):2113-33 (2014)

Buechner, C. N., Heil K., Michels, G., Carell, T., Kisker, C., Tessmer, I. Strand-specific recognition of DNA damages by XPD provides insights into nucleotide excision repair substrate versatility. *J Biol Chem.* 289(6) 3613-24 (2014)

Buechner, C.N., Tessmer, I. DNA substrate preparation for atomic force microscopy studies of protein-DNA interactions. *J Mol Recognit.* 26(12):605-17 (2013)

Göhler, A., Büchner, C., André, S., Doose, S., Kaltner, H., Gabius, H.J. Analysis of homodimeric avian and human galectins by two methods based on fluorescence spectroscopy: different structural alterations upon oxidation and ligand binding. *Biochimie.* 94(12) 2649-55. (2012)

Göhler, A., Büchner, C., André, S., Doose, S., Kaltner, H., Gabius, H.J. Sensing ligand binding to a clinically relevant lectin by tryptophan fluorescence anisotropy. *Analyst.* 136(24):5270-6. (2011)

8.6. Oral presentations and posters

- 02/2014 4th Euro AFM Forum in Göttingen
Talk: 'Atomic force microscopy studies of XPD–DNA interactions provide insights into Nucleotide Excision Repair substrate versatility.'
- 10/2013 Annual Summer Retreat of the Rudolf-Virchow-Center, University of Wuerzburg
Poster: 'Target site search of the human Thymine DNA Glycosylase'
- 10/2013 8th International Student Symposium of the Graduate School of Life Sciences, SCI, Scientific Crosstalk
Poster: 'Towards an understanding of damage recognition and verification in the DNA repair mechanism of nucleotide excision repair.'
- 10/2012 7th International Student Symposium of the Graduate School of Life Sciences, EPOS, Everything's Part Of Science
Poster: 'Mechanistic AFM-studies provide new insights in DNA damage recognition strategies of XPD; implications of disease-relevant mutations.'
- 10/2012 Annual Summer Retreat of the Rudolf-Virchow-Center, University of Wuerzburg
Talk: 'Mechanistic AFM-studies provide new insights in DNA damage recognition strategies of XPD.'
- 08/2012 73rd Harden Conference 'Machines on genes' in Oxford
Poster: 'Mechanistic AFM-studies provide new insights in DNA damage recognition strategies of XPD; implications of disease-relevant mutations.'
- 09/2011 Annual Summer Retreat of the Rudolf-Virchow-Center, University of Wuerzburg
Poster: 'How does Nucleotide Excision Repair work at the level of single molecules?'

8.7. Curriculum Vitae

Claudia Nadine Büchner

Date of birth: 11th May 1987

Place of birth: Schweinfurt

07/2011 - present

Doctoral research

Julius-Maximilians-University, Wuerzburg (Germany),
Rudolf Virchow Center, Department of Structural Biology

Title: Single molecule studies of
DNA lesion search and recognition strategies in DNA repair.

Supervisory Committee:

Ingrid Teßmer (primary supervisor)
Caroline Kisker
Markus Sauer

10/2006 – 03/2011

Study program Biology, Diploma

Julius-Maximilians-University, Wuerzburg (Germany)
Major field of studies: Biotechnology and Biophysics

Diploma thesis: Characterisation of galectins using fluorescence spectroscopy

Primary supervisor: Sören Doose
Secondary supervisor: Markus Sauer

09/1997 – 06/2006

Grammar School, school leaving certificate “Abitur”,
Jack-Steinberger-Gymnasium Bad Kissingen

Claudia Büchner
Würzburg 2014

8.8. Acknowledgments

Mein erster Dank gebührt meiner Betreuerin Dr. Ingrid Tessmer für die erfolgreiche Zusammenarbeit und für die vielen anregenden Ideen, die sehr zum Gelingen dieser Arbeit beigetragen haben! Außerdem danke ich Frau Prof. Dr. Caroline Kisker und Herrn Prof. Dr. Markus Sauer, die mich als weitere Mitglieder in meinem Thesis-Komitee mit Rat und Tat in der Diskussion von meinen Experimenten unterstützt haben. Herrn Prof. Dr. Markus Sauer danke ich zusätzlich für die Übernahme des Zweitgutachtens. Des Weiteren möchte ich mich bei Herrn Prof. Dr. Thomas Dandekar bedanken, der den Prüfungsvorsitz bei meiner Verteidigung übernimmt!

Besonderer Dank gilt allen Co-Autoren für die Zusammenarbeit an der XPD Publikation, mit deren Hilfe eine spannende Geschichte entstehen konnte! Vielen Dank auch an unsere Kollaborationspartner Prof. Dr. Alexander Drohat und Dr. Atanu Maiti im TDG-Projekt, die mich freundlicherweise sowohl mit aufgereinigten Proteinen als auch mit wertvollen Ratschlägen und Anregungen unterstützt haben! Danke auch an Dr. Daniela Schneeberger und Prof. Dr. Herrmann Schindelin für die erfolgreiche Zusammenarbeit am Collybistin-Projekt!

Zudem möchte ich mich sehr bei Steffi Wolski für die Hilfe bei der Aufreinigung von XPD bedanken und für viele hilfreiche Tipps! Besonderer Dank geht hierbei auch an Gudrun Michels – natürlich immer begleitet von Gesprächen über das Laufen. Vielen Dank an Nico Wirth, die Zusammenarbeit hat wirklich Spaß gemacht! Danke auch an alle ehemaligen und aktuellen Mitglieder der Strukturbiologie, die mich im Labor mit praktischer Hilfe und Ratschlägen unterstützt haben und für anregende Diskussionen!

Ganz herzlich möchte ich mich bei Petra Baron für ihre fortwährende und hervorragende Unterstützung danken, die mir vieles erleichtert hat! Großer Dank gilt auch meinen Büro-Kollegen: Agnes Elias, Monika Kuhn, Wolfgang Kölmel, Lisa Krebes und Lisa Schönwetter! Ihr seid einfach super und habt mich in allem sehr unterstützt! Und natürlich auch vielen Dank an meine ehemaligen Büro-Mitglieder, vor allem an die große Lisa, Anja, Annica, Bruno und Joachim!

Vielen lieben Dank an meinen Freund David Weber für die fantastische Hilfe und für das Korrekturlesen meiner Arbeit! Danke auch an all meine Freunde, die mich auf dem Weg zur Doktorarbeit so gut unterstützt haben! Besonders an Ise, Svetlana, Mia, Christina, Felix, Sophie, Maria, Ela, Flo und an die kleine und große Laura! Zuletzt bedanke ich mich bei meinen Eltern, meiner Schwester Nicole, bei Sammy, meiner Tante Renate und meiner Großcousine Vera, die mich in jeglicher Weise auf diesem Weg begleitet haben!