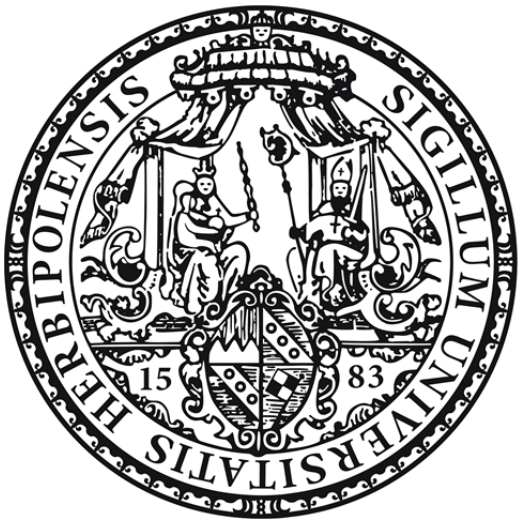

Nucleotide Excision Repair:
From Recognition to Incision of damaged DNA

Nukleotid-Exzisions-Reparatur:
Vom Erkennen zum Schneiden der geschädigten DNA

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

submitted by
Heide Marie Roth

from Marktoberdorf

Würzburg 2011

Submitted on: …….

Office stamp

Members of the *Promotionskomitee*:

Chairperson:	**Ulrike Holzgrabe**
Primary Supervisor:	**Caroline Kisker**
Supervisor (Second):	**Bennett Van Houten**
Supervisor (Third):	**Thomas Müller**
Supervisor (Forth):	**Helga Stopper**

Date of Public Defense: ……...

Date of Receipt of Certificates: …….

The scientist does not study nature because it is useful;
he studies it because he delights in it,
and he delights in it because it is beautiful.
If nature were not beautiful, it would not be worth knowing,
and if nature were not worth knowing,
life would not be worth living.

Jules Henri Poincaré (1854-1912)
French mathematician

Summary

The Nucleotide Excision Repair (NER) pathway is able to remove a vast diversity of structurally unrelated DNA lesions and is the only repair mechanism in humans responsible for the excision of UV-induced DNA damages. The NER mechanism raises two fundamental questions: 1) How is DNA damage recognition achieved discriminating damaged from non-damaged DNA? 2) How is DNA incision regulated preventing endonucleases to cleave DNA non-specifically but induce and ensure dual incision of damaged DNA? Thus, the aim of this work was to investigate the mechanisms leading from recognition to incision of damaged DNA.

To decipher the underlying process of damage recognition in a prokaryotic model system, the intention of the first part of this work was to co-crystallize the helicase UvrB form *Bacillus caldotenax* together with a DNA substrate comprising a fluorescein-adducted thymine as an NER substrate. Incision assays were performed to address the question whether UvrB in complex with the endonuclease UvrC is able to specifically incise damaged DNA employing DNA substrates with unpaired regions at different positions with respect to the DNA lesion. The results presented here indicate that the formation of a specific pre-incision complex is independent of the damage sensor UvrA. The preference for 5' bubble substrate suggests that UvrB is able to slide along the DNA favorably in a 5' $\rightarrow$ 3' direction until it directly encounters a DNA damage on the translocating strand to then recruit the endonuclease UvrC.

In the second part of this work, the novel endonuclease Bax1 from *Thermoplasma acidophilum* was characterized. Due to its close association to archaeal XPB, a potential involvement of Bax1 in archaeal NER has been postulated. Bax1 was shown to be a Mg^{2+}-dependent, structure-specific endonuclease incising 3' overhang substrates in the single-stranded region close to the ssDNA/dsDNA junction. Site directed mutagenesis of conserved amino acids was employed to identify putative active site residues of Bax1. In complex with the helicase XPB, however, incision activity of Bax1 is altered regarding substrate specificity. The presence of two distinct XPB/Bax1 complexes with different endonuclease activities indicates that XPB regulates Bax1 incision activity providing insights into the physical and functional interactions of XPB and Bax1.

Zusammenfassung

Die Nukleotid-Exzisions-Reparatur (NER) ist in der Lage, eine Vielfalt an strukturell unterschiedlichen DNA Schädigungen zu entfernen, und ist überdies der einzige DNA-Reparaturmechanismus im Menschen, der UV-induzierte DNA-Schädigungen entfernen kann. Der NER Mechanismus impliziert zwei grundlegende Fragen: 1) Wie wird geschädigte DNA erkannt und worauf gründet sich die Unterscheidung zwischen geschädigter und nicht-geschädigter DNA? 2) Wie wird das Schneiden der DNA reguliert? Wie wird unspezifisches Schneiden verhindert und sichergestellt, dass die geschädigte DNA auf beiden Seiten der Schädigung herausgeschnitten wird? Das Ziel dieser Arbeit war es daher, die Mechanismen zu untersuchen, die vom Erkennen zum Herausschneiden geschädigter DNA führen.

Um im bakteriellen Modelsystem den zugrundeliegenden Prozess der Schadenserkennung zu entschlüsseln, sollte im ersten Teil dieser Arbeit die Helikase UvrB aus *Bacillus caldotenax* zusammen mit einem geschädigten DNA Substrat kristallisiert werden. Als Schädigung wurde ein Fluorescein-Molekül genutzt, das an eine Thymin-Base gekoppelt wurde. Biochemische Experimente wurden durchgeführt um herauszufinden, ob UvrB im Komplex mit der Endonuklease UvrC spezifisch geschädigte DNA schneiden kann. Dafür wurden DNA-Substrate eingesetzt, die ungepaarte Basen an verschiedenen Stellen bezüglich der DNA-Schädigung enthielten. Die hier gezeigten Ergebnisse deuten darauf hin, dass ein spezifischer Komplex gebildet werden kann, der auch unabhängig von dem Schadenssensor UvrA zum Schneiden der DNA befähigt ist. Die Schnitt-Präferenz für die 5‘ ungepaarte Region lässt vermuten, dass UvrB bevorzugt in 5‘→3‘ Richtung an der DNA entlanggleiten kann. Sobald UvrB auf eine Schädigung auf diesem DNA Strang trifft, wird die Endonuklease UvrC rekrutiert.

Im zweiten Teil dieser Arbeit wurde die neuartige Endonuklease Bax1 aus *Thermoplasma acidophilum* charakterisiert. Aufgrund der engen Assoziation zu archaischem XPB wurde eine Beteiligung an der archaischen NER postuliert. Es konnte gezeigt werden, dass Bax1 eine Mg^{2+}-abhängige, strukturspezifische Endonuklease ist, die 3‘-Überhang Substrate im Einzelstrangbereich nahe des Einzelstrang/Doppelstrang-Überganges schneidet. Konservierte Aminosäuren wurden gezielt verändert, um diejenigen Reste zu identifizieren, die

möglicherweise das aktive Zentrum bilden. Im Komplex mit der Helikase XPB veränderte sich jedoch das Schneideverhalten im Hinblick auf die Substratspezifizität. Die Existenz von zwei verschiedenen XPB/Bax1 Komplexen mit unterschiedlicher Aktivität bezüglich des Schnittverhaltens könnte darauf hinweisen, dass XPB Bax1 reguliert. Diese Beobachtung erlaubt zugleich Einblicke in die Interaktion von XPB und Bax1 auf physikalischer und funktioneller Ebene.

Table of contents

1. Introduction

1.1. Nucleotide Excision Repair as a fundamental concept

DNA repair mechanisms have evolved in all three kingdoms of life to maintain their genomic information. These mechanisms are strongly required because DNA damages occur about 200,000 times per cell per day caused by endogenous and exogenous sources (Lindahl, 2000). Endogenous damaging agents are for example reactive oxygen species resulting from cellular reactions involved in metabolism and respiration (Sedelnikova et al, 2010). Examples for exogenous agents are UV-light and polycyclic aromatic hydrocarbons (PAH). PAHs are found in cigarette smoke and exhaust emissions, for example, and their metabolites form large DNA adducts such as benzo[α]pyrene-diol-epoxides (BPDEs) (Hecht, 1999). The problem of DNA damages caused by UV-light is becoming more and more important due to environmental pollution resulting in a further diminishing of the global ozone layer (de Gruijl, 1999; Diffey, 2004; van der Leun, 2004). It is interesting to note, that a combination of two different damaging agents does not only additively induce DNA lesions but also seems to multiply the damaging effects on DNA (Routledge et al, 2001). Several specialized DNA repair pathways have been identified in mammals which are responsible for the removal of numerous and diverse damages: mismatch repair, DNA damage reversal, homologous recombination, non-homologous end joining, base excision repair and nucleotide excision repair (Friedberg et al, 2006).

The nucleotide excision repair (NER) pathway is one of the most versatile DNA repair pathways as it is responsible for the removal of a vast diversity of structurally unrelated DNA lesions (Sancar, 1994; Sancar, 1996). In humans, it is the only DNA repair mechanism being able to remove UV lesions such as cyclobutane pyrimidine dimers (CPDs) and pyrimidine (6-4) pyrimidone photoproducts (6-4PP) (Gillet & Scharer, 2006).

Mutations in genes encoding NER proteins lead to the disease xeroderma pigmentosum (XP) which is characterized by a high sensitivity towards sunlight and thus results in an increased skin cancer rate. Other diseases caused by defective NER proteins are Cockayne's Syndrome (CS) and trichothiodystrophy

(TTD) (Bergmann & Egly, 2001; de Boer & Hoeijmakers, 2000; van Gool et al, 1997; Vermeulen et al, 1997).

1.2. Prokaryotic Nucleotide Excision Repair

The prokaryotic nucleotide excision repair pathway is investigated to decipher the mechanism leading to the discrimination of damaged and non-damaged DNA. To find a DNA lesion within a plethora of non-damaged DNA is both an essential task and a major challenge in DNA repair.

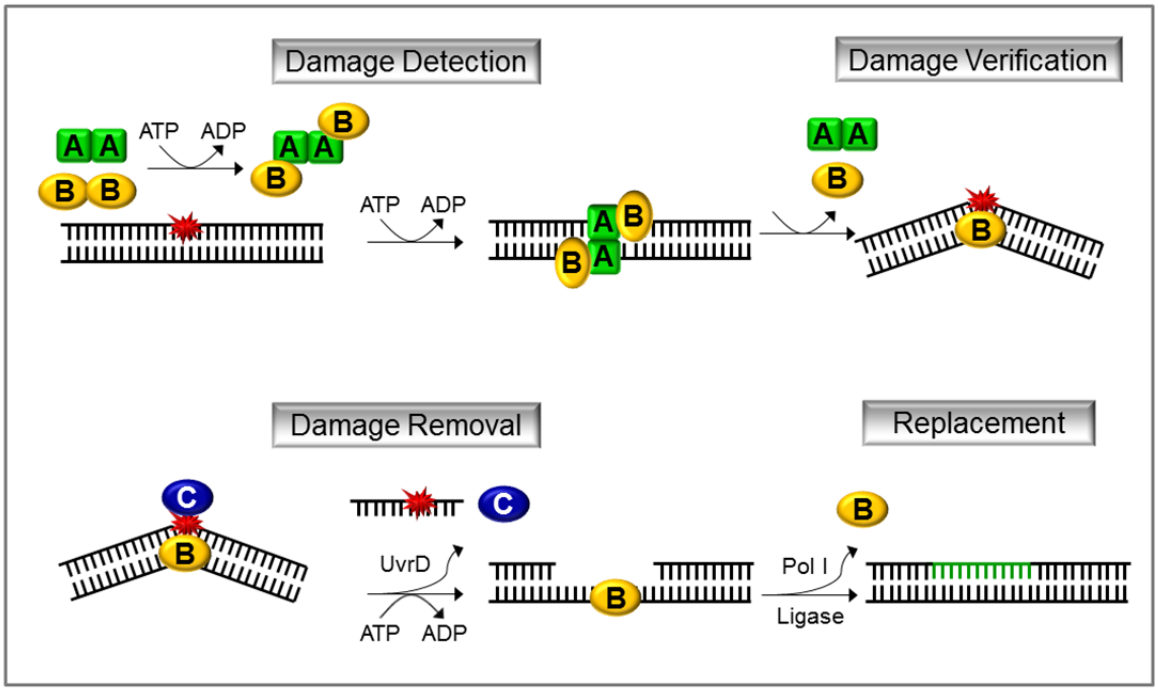

Figure 1.1 The prokaryotic nucleotide excision repair pathway [modified from (Theis et al, 2000)]. UvrAB form a DNA damage recognition complex which probes the DNA for lesions. Encountering a potential DNA damage, UvrB is loaded onto the DNA by UvrA which subsequently dissociates from the UvrB-DNA complex. UvrB verifies the presence of a NER substrate and thus establishes a stable pre-incision complex. UvrC is recruited to perform the dual incision 3' and 5' to the damaged nucleotide. UvrD removes the incised oligonucleotide together with UvrC to allow DNA Polymerase I and DNA ligase to accomplish the final DNA synthesis and ligation.

The NER pathway, in principal, can be divided into four major steps, namely DNA damage detection, verification, excision and replacement. These different steps are accomplished by the UvrABC proteins. UvrA is responsible for the detection of a helical distortion in the DNA double helix (DellaVecchia et al, 2004; Van Houten et al, 2005). Although being discussed controversially in the literature,

it is assumed that UvrA together with UvrB forms a heterotetrameric $UvrA_2UvrB_2$ (UvrAB) complex sliding along the DNA to initially detect DNA lesions (Kad et al, 2010; Orren & Sancar, 1990; Verhoeven et al, 2002; Wang et al, 2006). If the UvrAB complex encounters a DNA damage, UvrA loads UvrB onto the DNA and subsequently dissociates from the UvrB-DNA complex. UvrB verifies the NER substrate and forms a stable pre-incision complex (DellaVecchia et al, 2004). The endonuclease UvrC is then recruited and performs both the 3' and the 5' incision reaction with respect to the DNA lesion (Skorvaga et al, 2002; Verhoeven et al, 2000). The helicase UvrD releases UvrB, UvrC and the excised oligonucleotide to allow DNA polymerase I to carry out repair synthesis (Sancar, 1996). DNA ligase is finally responsible to seal the newly synthesized oligonucleotide to the parental DNA (Sancar, 1996). This mechanism is schematically depicted in **Figure 1.1**.

1.2.1. The role of UvrA in prokaryotic NER

UvrA fulfills a major role in prokaryotic NER probing the DNA for potential DNA damages (Truglio et al, 2006a). UvrA is an ATPase comprising two ATP-binding domains each containing a Walker A, a Walker B and an ABC signature motif (Pakotiprapha et al, 2008; Truglio et al, 2006a) (**Figure 1.2**). The UvrB-binding domain mediates the interactions with domain 2 of UvrB (Pakotiprapha et al, 2008; Pakotiprapha et al, 2009; Truglio et al, 2004). The most important contacts to DNA seem to be engaged by a patch of conserved residues within the Signature II domain close to the C-terminal Zn-finger (Croteau et al, 2008; Jaciuk et al, 2011). Two additional Zn-finger motifs have been observed in the Signature I domain, however, only the C-terminal Zn-finger was reported to be involved in DNA damage detection (Croteau et al, 2006; Jaciuk et al, 2011). The structure of UvrA in complex with damaged DNA suggests that UvrA does not directly interact with the DNA lesion, but senses helical distortions and the deformability of the DNA double helix (Jaciuk et al, 2011). This indirect readout mechanism is accomplished by the dimeric form of UvrA thus enabling the protein to probe the DNA substrate from both sides (Jaciuk et al, 2011). For DNA damage verification the DNA is then subsequently handed off to UvrB.

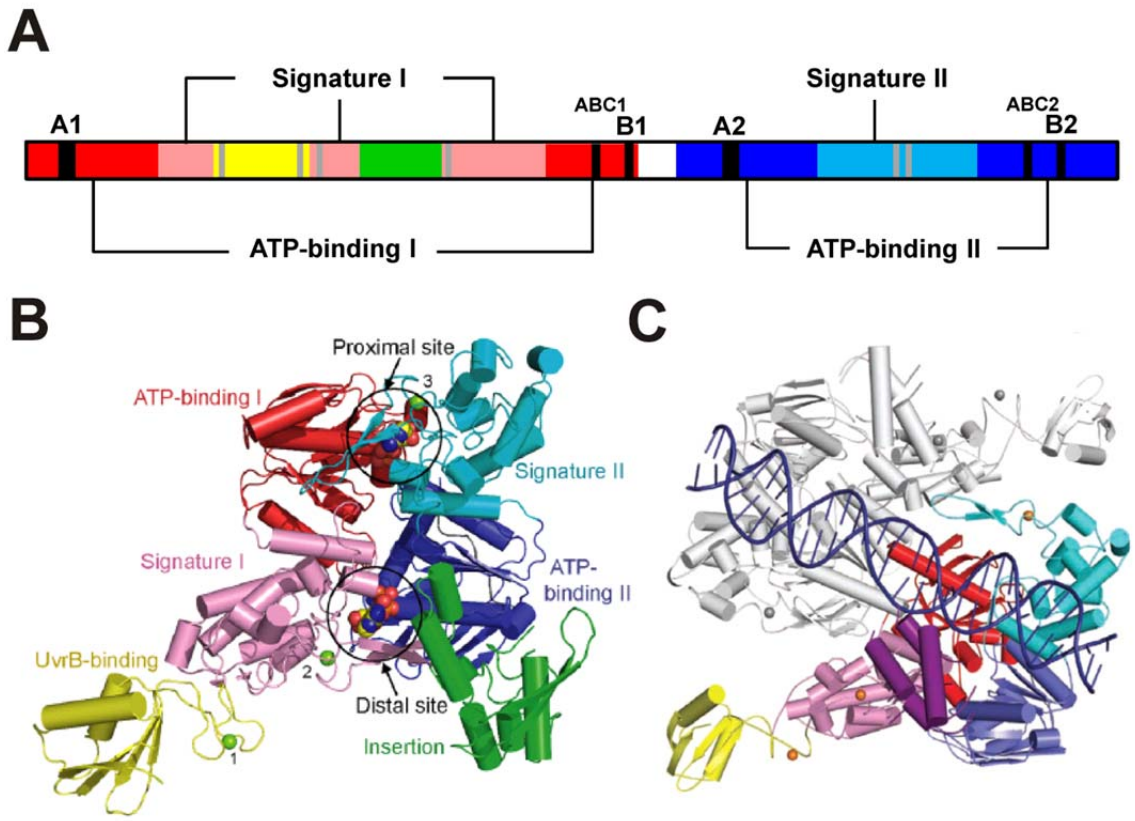

Figure 1.2 Domain architecture of UvrA. (A) UvrA consists of two domains separated by a linker region (white). Both domains comprise each a Walker A motif (A1 and A2), a Walker B motif (B1 and B2), and an ABC-signature sequence (ABC1 and ABC2) within the ATP-binding domains (depicted in red and blue). The conserved ATPase motifs are depicted in black. One signature domain is inserted in each ATP-binding domain (pink and cyan). The UvrB-binding domain (yellow) and the DNA-binding domain (green) are integrated in the signature I domain. For the coordination of each Zn^{2+}-ion, four residues are involved and are shown in gray (each gray bar representing two residues); thus, three Zn-fingers are present in UvrA each with one Zn^{2+}-ion coordinated by four residues. (B) Crystal structure of *Bacillus stearothermophilus* UvrA (PDB entry: 2R6F) (Pakotiprapha et al, 2008). The color coding is the same as described for panel A. (C) Crystal structure of *Thermotoga maritima* UvrA in complex with damaged DNA (PDB: 3PIH) (Jaciuk et al, 2011). The UvrA dimer (one monomer is colored as in panel A despite the DNA-binding domain here colored in purple, the second monomer is depicted in gray) probes the DNA from both sides for helical distortions indicating an indirect readout mechanism.

1.2.2. The structure and function of UvrB in prokaryotic NER

After having detected a potential DNA lesion, UvrA hands over the DNA to UvrB. UvrB plays an essential role in the NER cascade as it guides the DNA from the initial damage detection to the subsequent incision reaction.

UvrB is a superfamily 2 (SF2) helicase comprising 5 domains, 1a, 1b, 2, 3 and 4 (**Figure 1.3**). The two RecA-like helicase domains (domain 1a and 3) consist of seven helicase motifs providing an ATP binding site at the domain interface to translate energy from ATP hydrolysis into motion. (Theis et al, 1999; Theis et al, 2000). Domains 1a, 1b together with the protruding β-hairpin, and domain 3 mediate the interactions with DNA (Theis et al, 2000).

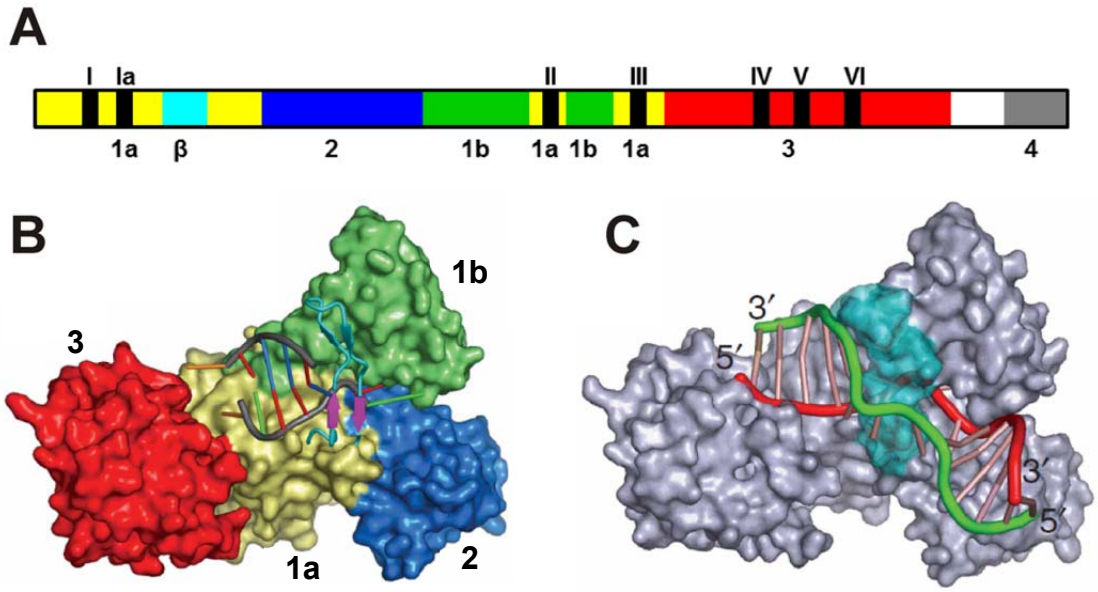

Figure 1.3 Domain architecture and crystal structure of UvrB in complex with DNA [modified from: (Truglio et al, 2006b)] (A) Domain architecture of UvrB. Domains 1a, 1b, 2, 3 and 4 are colored in yellow, blue, green, red and gray, respectively. The β-hairpin is shown in cyan. The seven helicase domains are depicted in black with their corresponding numbers. (B) The structure of *Bacillus caldotenax* UvrB was solved in complex with partially duplexed DNA (PDB: 2FDC). Please note: domain 4 is flexible and therefore not visible in the crystal structure. (C) The DNA binding model derived from the crystal structure indicates that the β-hairpin inserts between the two strands of DNA. The inner strand (depicted in red) is clamped between the β-hairpin and domain 1b.

The crystal structure of UvrB in complex with DNA (PDB: 2FDC) proposed that the β-hairpin inserts between the two strands of DNA supporting the crucial role of the β-hairpin in DNA damage recognition as previously suggested by mutagenesis studies (Moolenaar et al, 2001; Skorvaga et al, 2002; Truglio et al, 2006b) (**Figure 1.3**). Domain 2 is responsible for UvrA binding, whereas domain 4 interacts with both UvrA and UvrC (Theis et al, 2000). Intriguingly, truncation of domain 4 led to an increase in DNA binding and ATPase activity of the protein, suggesting that this part of the protein acts as an auto-inhibitory domain (Hsu et al, 1995; Wang et al, 2006).

1.2.3. The functions of UvrC in prokaryotic NER

The verification of a DNA damage and subsequent formation of a stable pre-incision complex consisting of UvrB and damaged DNA leads to the recruitment of the endonuclease UvrC. UvrC consists of two domains with different function (**Figure 1.4**).

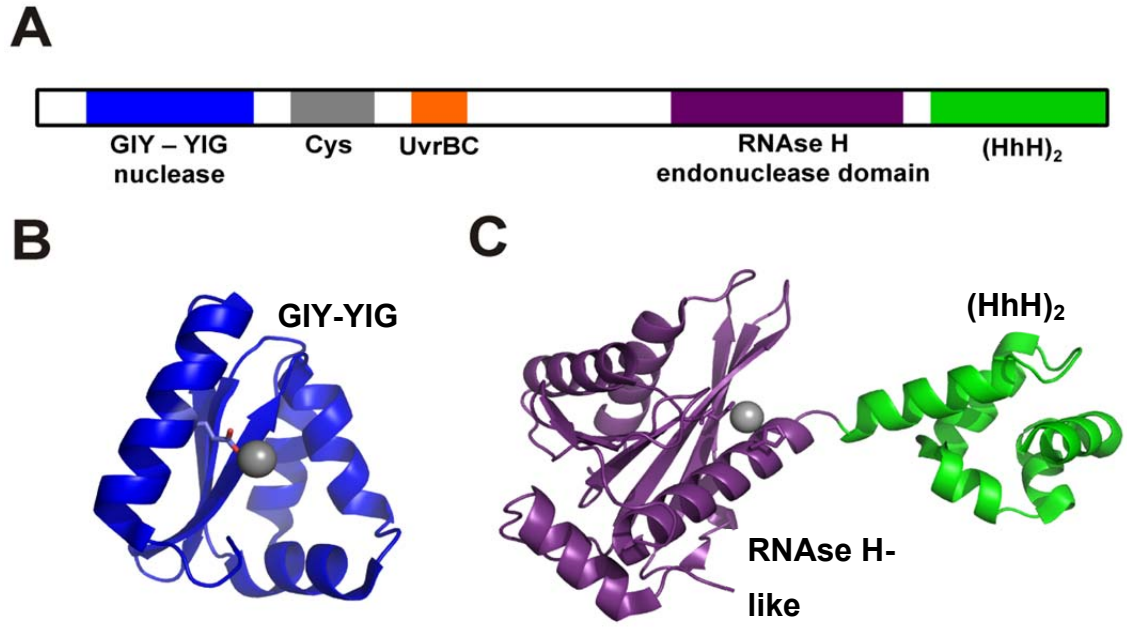

Figure 1.4 Domain architecture and crystal structure of UvrC. (A) UvrC consists of two endonuclease domains, particularly an N-terminal GIY-YIG nuclease domain (blue) and a C-terminal RNAse H-like endonuclease domain (purple). UvrC also comprises a Cys-rich region (gray), an UvrB-interacting domain (orange) and a C-terminal tandem Helix-hairpin-helix motif (green). (B) Crystal structure of the N-terminal GIY-YIG endonuclease domain of *T. maritima* UvrC (PDB: 1YD1) (Truglio et al, 2005). The residue E76 coordinating the Mg^{2+}-ion (gray sphere) is shown as stick model. (C) Crystal structure of the C-terminal RNAse H-like endonuclease domain (purple) and the $(HhH)_2$-domain (green) of *T. maritima* UvrC (PDB entry: 2NRZ) (Karakas et al, 2007). The catalytic triade DDH consisting of residues D367, D429 and H488 which coordinates one Mn^{2+}-ion (gray sphere) is represented as stick model.

The N-terminal GIY-YIG nuclease domain is responsible for the 3' incision reaction whereas the C-terminal RNAse H-like endonuclease domain performs the 5' incision reaction (Lin & Sancar, 1992; Truglio et al, 2006a; Van Houten et al, 2005; Verhoeven et al, 2000). The activity of both domains ensures dual incision and therefore complete removal of the damaged oligonucleotide. The structures of both UvrC nuclease domains were solved separately by X-ray crystallography (Karakas et al, 2007; Truglio et al, 2005) but the structure of the full-length protein

is not known. The structural studies on UvrC in addition to biochemical assays provide detailed insights into both incision mechanisms utilizing divalent cations to catalyze the cleavage of a phosphodiesterbond.

In addition to the two nuclease domains and the $(HhH)_2$-domain, a Cys-rich region and a UvrB-interacting domain are present in UvrC (Aravind et al, 1999). A model for the UvrB-UvrC interaction was derived from the crystal structure of UvrB's C-terminus (Sohi et al, 2000), however, the role of the Cys-rich region remains elusive (Truglio et al, 2006a).

1.2.4. From recognition to incision

The discovery of DNA repair immediately raised the fundamental question how proteins are able to discriminate damaged from non-damaged DNA. Since then, many puzzle pieces have been assembled to draw a conclusive picture about NER and damage recognition.

UvrA is the first protein in the NER cascade initially detecting distortions in the DNA double helix. The crystal structure of a UvrA dimer in complex with damaged DNA provides detailed information about the interactions between the two molecules which allows UvrA to distinguish damaged from non-damaged DNA (Jaciuk et al, 2011). In order to find a potential DNA lesion within a sea of non-damaged DNA, UvrA has recently been reported to perform a three-dimensional search which is characterized by only short-term binding events and jumping between different DNA molecules (Kad et al, 2010). In the presence of UvrB, however, the UvrAB complex applies a one-dimensional search by sliding along the DNA (Kad et al, 2010). The directed motion of UvrAB requires energy provided by ATP hydrolysis (Kad et al, 2010). UvrA hydrolyses ATP and GTP independent from the presence of DNA. UvrB, however, exerts only limited ATPase activity which is highly DNA-dependent and is also stimulated by UvrA (Goosen & Moolenaar, 2001; Moolenaar et al, 2000a). In contrast to processive helicases, UvrB in complex with UvrA rather translocates or slides on DNA but does not unwind large stretches of DNA (Kad et al, 2010; Oh & Grossman, 1989; Theis et al, 2000). It is interesting to note that the UvrAB complex moves on DNA in a 5' → 3' direction whereas UvrB was reported to employ its helicase activity to

locally unwind DNA in the opposite direction (Ahn & Grossman, 1996; Goosen, 2010; Oh & Grossman, 1987; Oh & Grossman, 1989; Truglio et al, 2006b; Van Houten, 1990).

Based on pioneering studies, it was suggested that UvrB is loaded by UvrA catalytically but not stoichiometrically onto damaged DNA (Orren & Sancar, 1989). Succeeding studies indicated that UvrB is positioned in close proximity to the DNA lesion and UvrB only locally separates the DNA for a few bases to verify the presence of a DNA damage (DellaVecchia et al, 2004; Orren & Sancar, 1990; Theis et al, 2000). Using partially single-stranded DNA such as bubble substrates, it has been shown that UvrB can bind to these DNA substrates without the help of UvrA (Moolenaar et al, 2000b; Zou & Van Houten, 1999). Subsequent incision assays revealed that UvrC can be recruited to perform specific incision reactions indicating that UvrB is able to form specific pre-incision complexes on damaged DNA independent from UvrA (Moolenaar et al, 2000b; Zou & Van Houten, 1999; Zou et al, 1997).

The β-hairpin in UvrB was demonstrated to be involved in DNA binding and DNA damage verification by the analysis of point mutants as well as deletion mutants of UvrB in biochemical approaches (Moolenaar et al, 2001; Skorvaga et al, 2002). The crystal structure of UvrB in complex with partially duplexed DNA suggested that UvrB inserts its protruding β-hairpin between the two strands of the DNA supporting the crucial role of the β-hairpin in damage recognition (Truglio et al, 2006b) (**Figure 1.3**). In addition to defects in DNA binding and damage verification, an increased ATPase activity of the UvrAB complex on damaged DNA was observed using these UvrB variants (Moolenaar et al, 2001; Skorvaga et al, 2002). This observation might be caused by the defective UvrB protein which could not be positioned on DNA by UvrA. Simultaneously, UvrB aimed to verify the presence of a DNA damage thereby continuously consuming ATP (Moolenaar et al, 2001; Skorvaga et al, 2002).

The hydrolysis of ATP induces helicase activity of UvrB which is proposed to be concomitantly translated into domain motion of the RecA-like domains 1 and 3 (Theis et al, 1999). During the process leading from recognition to incision of DNA, this movement seems to be required for correct positioning of the DNA, including base flipping of the base adjacent to the DNA lesion and bending of the DNA

(Malta et al, 2006; Skorvaga et al, 2004). In addition to these types of DNA deformations, DNA wrapping by UvrB was implicated to facilitate DNA damage recognition (Verhoeven et al, 2001; Wang et al, 2009).

It was suggested that bending of DNA leads to conformational changes in the pre-incision complex which then triggers the incision reaction by the endonuclease UvrC (Moolenaar et al, 2000a; Skorvaga et al, 2004). Interestingly, the *Eschericia coli* UvrB mutant D478A (corresponding to residue E478 in *B. caldotenax* located in domain 3) failed to bend DNA after damage detection and was thus not able to promote incision by UvrC (Lin et al, 1992). The presentation of the damaged DNA to UvrC seems to involve a regulatory mechanism to prevent UvrC from incising non-damaged DNA (Van Houten et al, 2005). Incision assays using pre-nicked DNA substrates suggested that formation of a specific UvrBC-DNA complex prior to the 3' incision is the rate limiting step (Moolenaar et al, 2000b; Zou et al, 1997). Moreover, it was demonstrated that the 5' incision is coupled to a preceding 3' incision (Lin et al, 1992).

These mechanisms ensure a concerted activity of UvrB and UvrC from recognition to incision of DNA in order to avoid unspecific cleavage of DNA by the endonuclease UvrC.

1.2.5. Objectives in prokaryotic NER

The ability of UvrB to verify a huge diversity of DNA lesions and to initiate dual incision by the endonuclease UvrC raises the question how this process is achieved and regulated on a molecular level. A plethora of biochemical and structural studies on the UvrABC proteins have been performed providing detailed insights into the prokaryotic NER system (Truglio et al, 2006a; Van Houten et al, 2005). Additionally, *in silico* methods such as molecular modeling and molecular dynamics simulations have been applied, for example, to model a DNA lesion at different positions on the inner and outer strand bound to UvrB (Jia et al, 2009) (**Figure 1.3**, panel C). Despite these comprehensive studies adumbrated in the previous paragraphs, the mechanism of DNA damage verification by UvrB remains elusive as well as the subsequent steps triggering dual incision by the endonuclease UvrC.

The term damage recognition includes the mechanism which allows UvrB to discriminate between damaged and non-damaged DNA. It also comprises the question how UvrB discriminates NER substrates from DNA damages repaired by other repair mechanisms such as the mismatch repair or base excision repair system. It is interesting to note that NER substrates vary among each other both in size and in their chemical properties. The only common denominator seems to be that these lesions distort and thus destabilize the DNA double helix (Chandrasekhar & Van Houten, 2000; Geacintov et al, 2002; Hoare et al, 2000; Truglio et al, 2006a). This huge diversity of different DNA lesions makes it difficult to envision how UvrB accomplishes DNA damage recognition.

To decipher the underlying process which enables UvrB to identify DNA damages, the intention of the first part of this work is to co-crystallize UvrB together with a hairpin-bubble DNA substrate comprising a fluorescein-adducted thymine as a NER substrate within an unpaired region. This substrate allows UvrB to bind the DNA in close proximity to the lesion without the help of UvrA. The crystal structure is expected to provide a snapshot of UvrB specifically interacting with the DNA damage.

The process of DNA damage verification ultimately raises another crucial question in NER: How does the endonuclease UvrC know, when and where to cut the DNA? After a DNA lesion has been verified by UvrB, conformational changes are postulated to trigger and regulate dual incision by UvrC. Thus, incision assays employing DNA substrates with unpaired regions at different positions with respect to the location of the DNA lesion will be performed to investigate the formation of a specific DNA incision complex of UvrBC on damaged DNA. Moreover, we intent to solve the crystal structure of UvrB in complex with UvrC as well as a damaged oligonucleotide in order to shed light onto the mechanism leading from recognition to incision of damaged DNA. Both the choice of a suitable DNA substrate as well as of a non-hydrolysable ATP-analog is supposed to ensure the formation of a stable protein-DNA complex prior to the first incision reaction. The structure of the quaternary complex consisting of UvrB, UvrC, damaged DNA and an AMPPNP is anticipated to reveal essential information on interactions and conformational changes crucial for the DNA incision reaction.

The second part of this thesis focuses on mechanisms leading from the damage recognition to its incision in eukaryotic NER. Archaeal homologs of the eukaryotic NER proteins thereby serve as a model to gain insights into their function using biochemical and biophysical approaches.

1.3. Eukaryotic Nucleotide Excision Repair

Similar to the prokaryotic nucleotide excision repair system, the eukaryotic pathway comprises basically four major steps. The DNA damage is first detected, then verified, removed and finally the excised oligonucleotide is replaced.

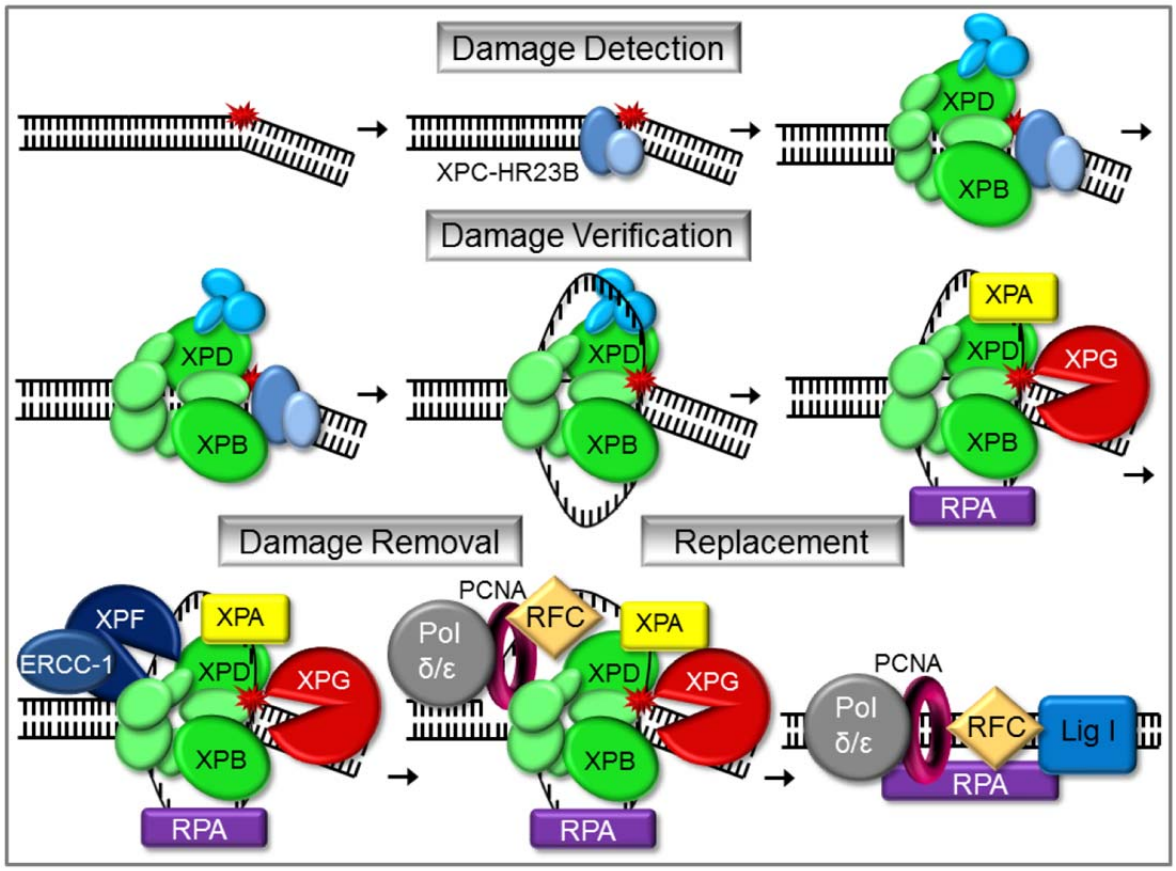

Figure 1.5 The eukaryotic NER pathway. [modified from: (Gillet & Scharer, 2006)] DNA lesions are initially detected by the XPC-HR23B complex. The ten subunit containing transcription factor TFIIH is recruited to locally unwind the DNA by the concerted activity of XPB and XPD. The two endonucleases XPG and XPF-ERCC1 ensure dual incision to remove the damaged oligonucleotide after the CAK complex dissociated from the TFIIH core complex. The non-damaged strand is used as a template for error-free DNA synthesis performed by DNA polymerases δ or ε in complex with the polymerase processivity factor PCNA, the replication factor RFC and ligase I.

The initial detection step in eukaryotic NER is differently achieved in actively transcribed DNA in contrast to transcriptionally silent regions (Bohr et al, 1985; Madhani et al, 1986). This phenomenon has also been observed and described for prokaryotic NER (Selby & Sancar, 1993; Selby & Sancar, 1994). In eukaryotic transcription-coupled repair (TCR), DNA damages are initially detected by RNA-polymerase II which is stalled during transcription at bulky DNA lesions (Brueckner et al, 2007). Cockayne's Syndrome complementation group A and B proteins (CSA and CSB) are then recruited as well as the transcription factor TFIIH to promote NER (Laine & Egly, 2006a; Laine & Egly, 2006b; van Hoffen et al, 1993).

In global genome repair (GGR), the initial detection of helical distortions is achieved by the XPC-HR23B complex (Sugasawa et al, 1998; Sugasawa et al, 2001). It was speculated that UV-DDB is also involved in sensing DNA lesions and in stimulating NER (Wakasugi et al, 2002; Wittschieben et al, 2005). The XPC-HR23B DNA detection complex then triggers the assembly of the repair complex consisting of TFIIH, both endonucleases XPG and XPF-ERCC1, and the DNA-binding proteins RPA and XPA (schematically represented in **Figure 1.5**) (Riedl et al, 2003; Wakasugi & Sancar, 1998; Yokoi et al, 2000).

The transcription factor TFIIH contains ten subunits. XPB, XPD, p62, p52, p44, p34 and p8/TTD-A form the core complex, whereas MAT1, cdk7 and cyclin H build the Cdk-activating kinases (CAK) complex (Chang & Kornberg, 2000; Coin et al, 2006; Giglia-Mari et al, 2004; Schultz et al, 2000). The DNA is locally unwound by the concerted activity of the ATPase XPB and the helicase XPD. Subsequently, the two endonucleases XPG and XPF-ERCC1 are recruited to perform both the 3' and 5' incision reactions after dissociation of the CAK complex from the TFIIH core complex. Polymerases δ or ε together with the polymerase processivity factor PCNA, the replication factor RFC and ligase I employ the non-damaged DNA strand as a template to newly synthesize a complementary DNA strand into the gap and to ligate the newly synthesized strand to the parental DNA strand (Gillet & Scharer, 2006).

1.3.1. Between recognition and incision – a central role of XPB in eukaryotic NER

Analogously to the prokaryotic NER system, the eukaryotic NER pathway employs a bipartite recognition mechanism using distinct steps for DNA damage detection and verification (Maillard et al, 2008). In GGR, the DNA damage is detected by UV-DDB or XPC-HR23B (Kulaksiz et al, 2005; Reardon & Sancar, 2003). The initial damage detection mechanism applied by XPC was investigated by structural as well as biochemical approaches (Clement et al, 2010; Min & Pavletich, 2007; Sugasawa et al, 2009). These studies provide insights into an indirect read-out mechanism to search for helical distortions caused by bulky DNA lesions.

To verify the presence of an NER substrate, the TFIIH complex is recruited by XPC-HR23B (Volker et al, 2001). Two central subunits of TFIIH are the helicases XPB and XPD. XPB and XPD are SF2-helicases of opposite polarity unwinding DNA duplexes in 3' → 5' and 5' → 3' direction, respectively (Hwang et al, 1996; Schaeffer et al, 1994; Schaeffer et al, 1993; Sung et al, 1993). Interestingly, the two helicases display different activities within the TFIIH complex depending on whether they cooperate in NER or in transcription initiation (Laine et al, 2006; Le May et al, 2010). Moreover, the sub-complexes XPB/p52 and XPD/p44 were shown to exert distinct roles in NER (Coin et al, 2007; Oksenych & Coin, 2010; Winkler et al, 2000). Interestingly, XPB's helicase activity was dispensable in NER, however its ATPase activity was essential for DNA opening in combination with the helicase activity of XPD/p44 (Coin et al, 2007; Oksenych et al, 2009). Thus, it was concluded that the ATPase XPB and the helicase XPD work together on opposite strands but in the same direction on the antiparallel DNA duplex to promote unwinding of the DNA (Maillard et al, 2008).

For the helicase XPD it was suggested that it plays and important role in the damage verification process (Naegeli et al, 1992; Oksenych & Coin, 2010). Despite the lack of a high overall sequence identity, XPD displays remarkable similarity to UvrB regarding the scaffold of the helicase domains providing evidence for a potential role of XPD in damage verification (Bienstock et al, 2003). Both biochemical and structural approaches were employed to shed light on XPD's contribution to NER (Fan et al, 2008; Liu et al, 2008; Mathieu et al, 2010;

Oksenych & Coin, 2010; Wolski et al, 2008; Wolski et al, 2010). Particularly, the helicase activity of XPD with regard to the iron sulfur cluster was investigated (Honda et al, 2009; Pugh et al, 2008a; Pugh et al, 2008b; Rudolf et al, 2006). Although controversially discussed, it was recently reported that XPD is stalled at a site of a DNA lesion indicating that XPD is able to verify the presence of a NER lesion (Mathieu et al, 2010; Rudolf et al, 2010).

Additionally, XPB has been investigated for its ability to unwind DNA and thus to possibly contribute to DNA damage recognition (Oksenych & Coin, 2010). Although the N-terminal domain of XPB was designated to act as a damage recognition domain, a potential ability in damage discrimination has not been proven by any other study (Fan et al, 2006; Oksenych & Coin, 2010). In contrast, XPB seems to exert different roles in NER: XPB was proposed to mainly mediate interaction with the XPC-HR23B DNA detection complex thus facilitating the recruitment of the TFIIH complex to damaged DNA (Bernardes de Jesus et al, 2008; Yokoi et al, 2000). Furthermore, XPB's ATPase activity is necessary to anchor the TFIIH complex to the DNA damage (Oksenych et al, 2009). Although XPA was suggested to be responsible to load and position XPF-ERCC1 onto the DNA 5' to the lesion (Tsodikov et al, 2007; Volker et al, 2001), there is some evidence that the recruitment of the endonuclease XPF-ERCC1 is mediated by XPB, likely in a concerted manner with XPA (Coin et al, 2004; Evans et al, 1997; Oh et al, 2007). In addition to its role in protein anchoring, XPB also seems to regulate XPF-ERCC1 since phosphorylation of XPB's S751 abolishes 5' incision by XPF-ERCC1 (Coin et al, 2004). This indicates that XPB is not directly involved in DNA damage discrimination, but might play an essential role in triggering the incision reaction after damage verification by XPD thereby bridging the process from recognition to incision (Oksenych & Coin, 2010).

The crystal structure of an archaeal XPB homolog (*Archeoglobus fulgidus* XPB) was solved providing information about XPB's architecture and function (**Figure 1.6**) (Fan et al, 2006). XPB consists of two helicase domains (HD1 and HD2) comprising seven helicase motifs in total including a Walker A motif as well as a conserved DExD/H-box motif (Fan et al, 2006). In the first helicase domain, the so-called RED motif is located and was reported to be essential for helicase

activity whereas in the second helicase domain a protruding thumb domain is supposed to be involved in DNA binding (Fan et al, 2006).

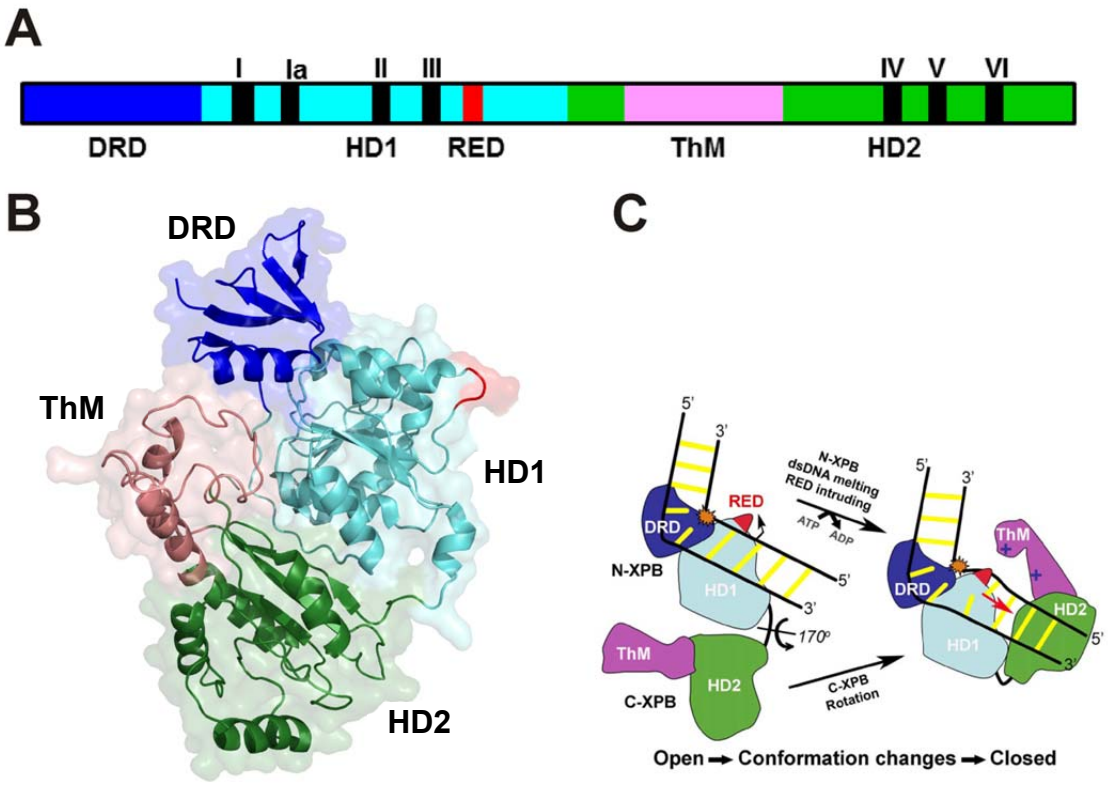

Figure 1.6 The domain architecture and crystal structure of *A. fulgidus* XPB. [modified from (Fan et al, 2006)] (A) Domain architecture of XPB. Two helicase domains HD1 and HD2 depicted in cyan and green, respectively, comprise in total seven helicase motifs. HD1 additionally contains the RED motif (depicted in red). The thumb domain (ThM; depicted in pink) is inserted in HD2. The damage recognition domain (DRD) is located at the N-terminus (blue). (B) The structure of XPB (PDB: 2FWR) reveals a 4 domain architecture comprising two helicase domains (cyan and green) a damage recognition domain (depicted in blue) and a thumb domain (pink). The RED motif within the first helicase domain is highlighted in red. (C) The proposed DNA binding and damage detection mechanism includes major conformational changes from an open to a closed conformation (Fan et al, 2006).

Moreover, an N-terminal damage recognition domain (DRD) was identified and suggested to be involved in discrimination of damaged from non-damaged DNA (Fan et al, 2006). Superposition of crystal structures of XPB with the helicase from hepatitis virus C (HCV NS3 helicase) was used to generate a potential DNA binding model (**Figure 1.6**, panel C) (Fan et al, 2006; Kim et al, 1998).

The crystal structure, however, provides only limited insights into XPB's role in NER, since XPB seems to undergo major conformational changes necessary for mediating DNA binding as well as the protein-protein interactions within the TFIIH

and XPF-ERCC1 complexes (Oh et al, 2007; Tapias et al, 2004). After assembly of all NER factors, XPF-ERCC1 was suggested to perform the 5' incision reaction. Although it has been reported that XPB is responsible for XPF-ERCC1 recruitment (Oh et al, 2007), it is not clear which signals are necessary for binding and engaging the endonuclease. Phosphorylation of XPB has been shown to inhibit XPF's endonuclease activity (Coin et al, 2004), indicating that dephosphorylation of this residue might contribute to the trigger of the first incision reaction. The resulting free ssDNA-overhang would then be displaced by DNA polymerase δ/ε which simultaneously synthesizes the first part of the missing oligonucleotide (Staresincic et al, 2009). XPG subsequently conducts the second incision reaction and allows repair synthesis to be completed (Staresincic et al, 2009). It was speculated that this concerted mechanism of incision and repair synthesis is not only initiated by the TFIIH components but also regulated by the DNA synthesis machinery (Staresincic et al, 2009) since it has been reported, that both endonucleases XPF-ERCC1 and XPG interact with PCNA (Chapados et al, 2004; Gary et al, 1999; Mocquet et al, 2008; Roberts et al, 2003a). Although controversially discussed in the literature (Mocquet et al, 2008), the "cut-patch-cut-patch" mechanism would avoid the presence of large stretches of ssDNA which are particularly prone to DNA strand breaks. The concerted activity would ensure both dual incision and complete DNA repair synthesis to take place (Staresincic et al, 2009).

It is a fundamental question in eukaryotic NER how dual incision of DNA is regulated. This regulation mechanism includes both the recruitment of the endonucleases as well as the induction and triggering of the incision reaction at a defined position and a defined time. In order to decipher the molecular mechanism from DNA unwinding and damage recognition to dual incision, archaeal homologs of the eukaryotic NER proteins are often used in both biochemical *in vitro* assays and in X-ray crystallography experiments.

1.3.2. Archaeal homologs serve as a model for eukaryotic NER

Besides prokaryotes and eukaryotes, archaea were identified as the third kingdom of life (Woese & Fox, 1977). Based on the phylogenetic analysis of

ribosomes, archaea are subdivided into two groups, the euryarchaea and the crenarchaea (Cox et al, 2008). Especially for hyperthermophilic archaea, the question arises which mechanisms are used to maintain genomic stability (Grogan, 2000; Grogan, 2004). Analyses of archaeal genomes revealed the presence of several conserved eukaryotic homologs, among them the XP-proteins responsible for NER in eukaryotes (Aravind et al, 1999; Ogrunc et al, 1998; White, 2003). The majority of archaea comprise homologs for the helicases XPB and XPD. In contrast to their human counterparts, archaeal XPB lacks the N-terminal domain and approximately 100 residues at the C-terminus (Fan et al, 2006), whereas archaeal XPD is only slightly shorter at the C-terminus compared to human XPD (Wolski et al, 2008). The core domains of both helicases, however, are conserved between human and archaeal proteins. In addition to the archaeal homologs of the helicases XPB and XPD, most archaea contain genes encoding FEN-1, a homolog of the eukaryotic endonuclease XPG (Kelman & White, 2005). The archaeal homolog of XPF seems to be present in two forms: the euryarchaeal long form comprises an N-terminal helicase domain and a C-terminal nuclease domain and is also known as Hef endonuclease, whereas the crenarchaeal short form only encodes the nuclease domain (Rouillon & White, 2011). In eukaryotic NER, the structure-specific endonucleases XPF and XPG perform dual incision 5' and 3' to the lesion, respectively (O'Donovan et al, 1994; Sijbers et al, 1996). Similar incision specificities of the endonucleases towards their substrates have been observed for their archaeal counterparts: Hef was shown to incise 3' flap model DNA substrates whereas Fen-1 incised 5' flap substrates consistent with the 5' and 3' incision reactions, respectively (Hutton et al, 2008; Newman et al, 2005; Nishino et al, 2006; Roberts et al, 2003a). So far, it has not been confirmed whether XP-proteins are employed in archaea for NER at all or if they exhibit their activity in different processes such as DNA replication (Kelman & White, 2005; Rouillon & White, 2011).

Several observations contradict a potential role for the XP-proteins in archaeal NER: Firstly, some archaea additionally comprise genes encoding the prokaryotic UvrABC system (Rouillon & White, 2011). *In vivo* studies on *Halobacterium sp.* NRC-1 revealed an elevated sensitivity towards UV-light upon depletion of the UvrABC system indicating that this organism primarily employs the

prokaryotic NER system rather than the eukaryotic system for the repair of UV-induced DNA damages (Crowley et al, 2006). Secondly, it has been published recently that knock-out of *xpb* and *xpd* genes in the hyperthermophilic euryarchaeon *Thermococcus kodakaraensis* results only in a slight increase in sensitivity towards UV-irradiation compared to the wildtype strain supporting that this archaeon does not use the XP-homologs for the removal of UV-lesions (Fujikane et al, 2010). Thirdly, some of the essential proteins involved in eukaryotic NER such as XPA and XPC are not present in any archaeal organism (Kelman & White, 2005; White, 2003). Lastly, not all archaeal organisms contain the same set of XP-homologs (Kelman & White, 2005; Rouillon & White, 2011). Interestingly, the archaeon *Methanopyrus kandleri* lacking the prokaryotic UvrABC system does not contain homologs of XPB or XPD (Rouillon & White, 2011).

Primarily, these studies reflect a huge diversity within the third kingdom of life displaying an almost random distribution of DNA repair homologs in archaea. Notably, this does not necessarily imply that none of the archaea utilize XP-homologs for NER. A UV-induced up-regulation of homologs for XPB, XPF and XPG in the crenarchaeon *Sulfolobus solfataricus* has been reported but contradictory results have been published as well (Frols et al, 2007; Gotz et al, 2007; Salerno et al, 2003; Wood et al, 1997). These discrepancies might be explained by a study where up-regulation of NER proteins in human keratinocytes was observed only at a defined UV-dose, whereas at higher doses, NER factors were down-regulated and apoptosis was induced (Maeda et al, 2001).

Although archaeal NER is still not well understood, the archaeal homologs of the eukaryotic NER proteins are widely employed for biochemical as well as structural studies as a model for the eukaryotic NER proteins since archaeal and eukaryotic NER proteins share significant sequence similarity. The core domains of human and *Thermoplasma acidophilum* XPB, for example, share a sequence identity of 26 %. Several *in vitro* studies on archaeal proteins provided detailed insights into NER on a molecular level (Fan et al, 2006; Liu et al, 2008; Ma et al, 2011; Mathieu et al, 2010; Nishino et al, 2003; Richards et al, 2008; Roberts et al, 2003a; Wolski et al, 2008; Wolski et al, 2010). Notably, the molecular consequences of mutations leading to XP, CS and TTD could be explained by the

crystal structure of an archaeal XPD homolog from *T. acidophilum* (Lehmann, 2008; Wolski et al, 2008).

Proteins from the euryarchaeon *T. acidophilum* were chosen as a model to investigate eukaryotic NER homologs since this organism does not comprise genes encoding for the prokaryotic UvrABC proteins which might be employed for NER instead (Rouillon & White, 2011). Besides, the complete genome sequence of *T. acidophilum* is available (Ruepp et al, 2000). Since *T. acidophilum* grows at a temperature of ~60 °C, the proteins are thermophilic and thus more amenable to biochemical and structural approaches than their human counterparts.

1.3.3. Bax1 – a novel nuclease in archaeal NER?

Nucleases are enzymes cleaving phosphodiester bonds in the DNA backbone and are basically classified by numerous different properties towards their substrates and their cleavage mechanism (Mishra, 2002). Nucleases are, for example, characterized as DNases or RNases, they act as exonucleases or endonucleases and their substrate is recognized in a sequence-specific or structure-specific manner. In DNA repair, structure-specific nucleases play a crucial role in DNA processing. However, the nuclease activity has to be tightly regulated on the one hand to avoid pre-mature or non-specific incision of DNA but on the other hand to ensure incision of verified DNA lesions.

Recently, genome analysis revealed the presence of a gene locus in close proximity to an archaeal XPB (Richards et al, 2008). It has been confirmed that the protein encoded by this gene is able to form a stable complex with archaeal XPB and was therefore named Bax1 (binds archaeal XPB). The association with XPB implied a role in archaeal NER although its function was not known (Richards et al, 2008). Bax1 does not seem to be present in eukaryotes but exists in many archaea and it appears to be abundant in some prokaryotes, particularly in cyanobacteria (http://www.expasy.ch/tools/blast/). Interestingly, Bax1 was predicted to contain a DUF790-domain and was therefore suggested to act as an endonuclease (Kinch et al, 2005). Thus, the question arises if Bax1 is an active endonuclease and which role Bax1 might play in archaeal NER.

1.3.4. Objectives in archaeal Nucleotide Excision Repair

In order to address these questions, this work will investigate Bax1's putative endonuclease activity and its biochemical properties both alone and in complex with archaeal XPB. Firstly, the physical interaction of XPB and Bax1 will be confirmed for *T. acidophilum* proteins. In the next step, the endonuclease activity of Bax1 as well as of the XPB/Bax1 complex will be characterized. Incision assays using different DNA substrates will be applied to disclose the specificity of Bax1. Since nucleases often employ divalent cations for cleaving phosphodiester bonds (Dupureur, 2008), incision assays in the presence and absence of different cations will uncover Bax1's requirement for metal ions in catalysis. To identify Bax1's active site residues mediating DNA cleavage, site-directed mutagenesis of conserved amino acids will be performed. The results obtained from these experiments might help to assess which role Bax1 plays in archaeal NER.

In addition, the activity of the XPB/Bax1 complex compared to Bax1 might contribute to shed light onto the regulation of nucleases and the processes from recognition to incision of damaged DNA. Incision assays as well as biophysical approaches such as analytical ultracentrifugation and atomic force microscopy will be used to investigate complex formation of XPB and Bax1. Since XPB's role within the TFIIH complex was postulated to include the recruitment of the endonuclease XPF-ERCC1 to the repair complex and subsequent initiation of the first incision reaction (Oh et al, 2007), these insights might help to understand the interactions and conformational changes of XPB within the TFIIH complex necessary to promote the primary incision by XPF-ERCC1.

2. Materials

2.1. DNA substrates

All DNA substrates used in this work were purchased from Biomers, IDT, Sigma Genosys or Metabion.

Table 2.1 DNA substrates used in this work

Name	Sequence
F_{26}50	5‘ GACTACGTACTGTTACGGCTCCATC[FldT]CTACCGCAAT CAGGCCAGATCTGC 3’
NDT	5‘ GACTACGTACTGTTACGGCTCCATCTCTACCGCAATCAG GCCAGATCTGC 3’
NDT30	5’ CCATCTCTACCGCAATCAGGCCAGATCTGC 3’
NDB	5’ GCAGATCTGGCCTGATTGCGGTAGAGATGGAGCCGTAA CAGTACGTAGTC 3’
NDB22	5’ GGAGCCGTAACAGTACGTAGTC 3’
NDB26	5’ AGATGGAGCCGTAACAGTACGTAGTC 3’
NDB30	5’ GTAGAGATGGAGCCGTAACAGTACGTAGTC 3’
NDBr15	5’ GCAGATCTGGCCTGA 3`
NDBr20	5’ GCAGATCTGGCCTGATTGCG 3’
NDBr25	5’ GCAGATCTGGCCTGATTGCGGTAGA 3’
NDBr30	5’ GCAGATCTGGCCTGATTGCGGTAGAGATGG 3’
NDBB16	5’ GCAGATCTGGCCTGATCTTCCGTTTTTACCTTCCGTAA CAGTACGTAGTC 3’
100mer_bottomF	5’ GGGGGTACGAGACTCGAGGCATGCGGTCGACTCT AGAGGATCAGATC[FldT]GGAACCTCTAGACTCGAGG CATGCACCTCTAGACTCGAGGCATGCGCATGGG 3’
100mer_top_NB	5’ CCCATGCGCATGCCTCGAGTCTAGAGGTGCATGCCTC GAGTCTAGAGGTTCCAGATCTGATCCTCTAGAGTCGACC GCATGCCTCGAGTCTCGTACCCCC 3’
100mer_top_CB	5’ CCCATGCGCATGCCTCGAGTCTAGAGGTGCATGCCTC GAGTCTAGAGGTCTTTCTTCTGATCCTCTAGAGTCGACC GCATGCCTCGAGTCTCGTACCCCC 3’

100mer_top_5B	5' CCCATGCGCATGCTCTCCCTCTAGAGGTGCATGCCTC GAGTCTAGAGGTTCCAGATCTGATCCTCTAGAGTCGACC GCATGCCTCGAGTCTCGTACCCCC 3'
100mer_top_3B	5' CCCATGCGCATGCCTCGAGTCTAGAGGTGCATGCCTC GAGTCTAGAGGTTCCAGATCTGATCCTCTAGAGTCGACC GCATGCCTCCTCCTCCGTACCCCC 3'
A_F24	5' TACGGCCTCATC[FldT]CTAACGCAATC 3'
A_comp24	5' GATTGCGTCTTTTTTTGAGGCCGTA 3'
E	5' CGACTTTTTTCCTTTGGATC[FldT]CTATCG 3'

2.2. Bacterial strains and plasmids

Table 2.2 Bacterial strains used in this work

Strain	Genotype	Reference/Supplier
BL21-CodonPlus®(DE3)-RIL	*E. coli* B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte [*argU ileYleuW* Cam^r]	Stratagene
DH5α	F- φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-	Invitrogen
XL1-blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*^{q}ZΔ*M15* Tn*10* (Tet^r)]	Stratagene

Table 2.3 Plasmids used in this work

Plasmid	Tag	Resistance	Reference/Supplier
pETM-11	N-terminal His_6-Tag	Kanamycin	EMBL
pBADM-11	N-terminal His_6-Tag	Kanamycin	EMBL
pTXB1	C-terminal Intein-Tag	Ampicillin	New England Biolabs
pTYB1	C-terminal Intein-Tag	Ampicillin	New England Biolabs

2.3. Media and antibiotics

Table 2.4 Media used in this work

Medium	Reference/Supplier
LB-Medium	Roth, Karlsruhe
LB-Agar	Roth, Karlsruhe
ZYM-5052 Medium	(Studier, 2005)

Table 2.5 Antibiotics used in this work

Antibiotic	Final concentration	Supplier
Ampicillin	100 µg/ml	Roth, Karlsruhe
Chloramphenicol	34 µg/ml	Roth, Karlsruhe
Kanamycin	50 µg/ml	Roth, Karlsruhe

2.4. Enzymes

Table 2.6 Enzymes used in this work

Enzyme	EC number	Reference
*Hind*III	EC 3.1.21.4	New England Biolabs
*Nco*I	EC 3.1.21.4	New England Biolabs
*Pci*I	EC 3.1.21.4	New England Biolabs
*Xho*I	EC 3.1.21.4	New England Biolabs
Nt.*Alw*I	EC 3.1.21.4	New England Biolabs
Nt.*Bst*NBI	EC 3.1.21.4	New England Biolabs
Trypsin	EC 3.4.21.4	Roth, Karlsruhe
Tobacco Etch Virus (TEV) protease	EC 3.4.22	EMBL Hamburg
Lysozyme	EC 3.2.1.17	Roth, Karlsruhe
Phusion DNA polymerase	EC 2.7.7.7	Finnzymes
Pfu DNA polymerase	EC 2.7.7.7	Fermentas
T4 DNA Ligase	EC 6.5.1.1	Invitrogen
T4 polynucleotide kinase	EC 2.7.1.78	Invitrogen
Desoxyribonuclease I	EC 3.1.21.1	Invitrogen

2.5. Chemicals

If not noted differently, all chemicals and solutions were purchases from Carl Roth (Karlsruhe), Sigma Aldrich (Seelze), Fluka (Neu-Ulm), Hampton Research (Laguna Hills, USA), or Applichem (Darmstadt). All chemical were of analytical grade or better, the chemicals used for crystallization were of the highest available purity.

2.6. Relevant buffers and solutions

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

Table 2.7 Buffers for the purification of UvrB and UvrC

Buffer	Components	pH
UvrB-AC/GF	20 mM TrisHCl, 500 mM NaCl	8.0
UvrB cleavage	20 mM TrisHCl, 500 mM NaCl, 100 mM DTT	8.0
UvrB dialysis	20 mM TrisHCl, 150 mM NaCl	8.0
UvrC-AC	20 mM TrisHCl, 500 mM NaCl, 0.1 mM EDTA	9.0
UvrC cleavage	20 mM TrisHCl, 500 mM NaCl, 0.1 mM EDTA, 100 mM DTT	9.0
UvrC-GF	20 mM TrisHCl, 500 mM NaCl	9.0
UvrBC-GF	20 mM TrisHCl, 500 mM NaCl, 0.1 mM EDTA	8.0

Table 2.8 Buffers for the purification of XPB, Bax1, XPB/Bax1

Buffer	Components	pH
IMAC-A	50 mM TrisHCl, 500 mM NaCl, 50 mM Imidazole	7.5
IMAC-B	50 mM TrisHCl, 500 mM NaCl, 500 mM Imidazole	7.5
IEX-A	50 mM TrisHCl, 50 mM NaCl	7.5
IEX-B	50 mM TrisHCl, 1 M NaCl	7.5
GF-1	20 mM TrisHCl, 500 mM NaCl	7.5
GF-2	20 mM TrisHCl, 200 mM NaCl	7.5

Table 2.9 Buffers for biochemical assays

Buffer	Components	pH
ABC	50 mM TrisHCl, 10 mM $MgCl_2$, 50 mM KCl, 1 mM ATP, 5 mM DTT	7.5
Bax1	20 mM MES, 150 mM NaCl, 1 mM DTT, 0.1 mg/ml BSA	6.5
TE	10 mM TrisHCl, 1 mM EDTA	7.5
Annealing-buffer	0.1 x TE, 100 mM KCl	7.5

Table 2.10 Buffers and solutions for SDS-PAGE, native and urea gels

Buffer	Components
Staining solution	50 % (*v/v*) Methanol, 10 % (*v/v*) Acetic acid, 0.1 % (*w/v*) Coomassie brilliant blue
Destaining solution	10 % (*v/v*) Methanol, 5 % (*v/v*) Acetic acid
Running buffer	192 mM Glycin, 0.1 % (*w/v*) SDS, 25 mM Tris
Loading buffer	50 mM Tris pH 6.8, 100 mM DTT, 2 % (*w/v*) SDS, 0.1 % (*w/v*) Bromphenol blue, 10 % (*w/v*) Glycerol
15 % resolving gel	15 % (*v/v*) Acrylamide/Bisacrylamide (37.5:1), 375 mM TrisHCl, pH 8.8, 0.1 % (*w/v*) SDS, 0.25 % (*w/v*) APS, 0.05 % TEMED
5 % stacking gel	5 % (*v/v*) Acrylamide/Bisacrylamide (37.5:1), 125 mM TrisHCl, pH 6.8, 0.1 % (*w/v*) SDS, 0.25 % (*w/v*) APS, 0.15 % (*v/v*) TEMED
TBE-buffer	89 mM Tris Base, 89 mM Boric acid, 2 mM EDTA, pH 8.3
6 % native gel	6 % (*v/v*) Acrylamide/Bisacrylamide (29:1), 0.5 x TBE, 0.2 % (*v/v*) APS, 0.1 % (*v/v*) TEMED
6 x Native sample buffer	60 mM TrisHCl, pH 7,5, 30 mM Na-acetat, 12 mM EDTA, 60 % (*w/v*) Glycerol, 0,36 % (*w/v*) Orange G
15 % urea gel	15 % (*v/v*) Acrylamide/Bisacrylamide (19:1), 1 x TBE, 4 M urea,0.05 % (*v/v*) APS, 0.005 % (*v/v*) TEMED
2 x Urea sample buffer	8 M Urea, 5 mM TrisHCl, pH 7.5, 0.5 % (*w/v*) Orange G

2.7. Crystallization Screens

Table 2.11 Crystallization Screens used in this work

Screen	Reference/ Supplier
Crystal Screen 1 and 2	Hampton Research
Index Screen HT	Hampton Research
Wizard Screen I and II	Emerald Biosystems
OptiMix-1-5	Topaz
The Nucleix Suite	Qiagen
The Protein Complex Suite	Qiagen
The PEGs Suite	Qiagen
Easy Xtal JCSG	Qiagen

2.8. Equipment and instrumentation

Table 2.12 Important equipment used in this work

Device	Name	Company
Bio Layer Interferometry	Octet RED	Forté BIO
Chromatography columns	Chitin Beads HiLoad™16/60Superdex™200pg HiLoad™26/60Superdex™200pg HiTrap SP HP (1 ml) Ni-MAC (1 ml) Superdex™75 10/300 GL Superdex™200 10/300 GL	NEB GE Healthcare GE Healthcare GE Healthcare Novagen GE Healthcare GE Healthcare
CD-Spectropolarimeter	J-810	Jasco
Dynamic Light Scattering	DynaPro Titan	Wyatt Technology Corporation
FPLC systems	ÄKTA purifier 10 ÄKTA xpress	GE Healthcare GE Healthcare
Liquid Handling System	Honeybee 963 Lissy	Zinsser Analytic Zinsser Analytic
Molecular Imager	Molecular Imager Pharos (FX) System	BioRad
Spectrophotometer	NanoDrop ND 1000	Peqlab
X-ray detector	Raxis HTC	Rigaku
X-ray generator	MicroMax 007 HF	Rigaku

2.9. Software and databases

Table 2.13 List of software and databases used in this work

	Software/Database	Author/ Reference
3D viewer	PyMOL	(DeLano, 2003)
Chromatography	Unicorn 5	GE Healthcare
CD-Spectroscopy	Spectra Manager	Jasco
BioLayer Interferometry	Octet RED Version 6.3	FortéBIO
Diverse protein analyses	ExPASy Proteomics Server	www.expasy.ch
Dynamic Light Scattering	Dynamics	Wyatt Technology Corporation
Molecular Imager	Quantity One	BioRad
Protein Data Bank	PDB	www.pdb.org
Secondary Structure Assigment	DSSP	(Kabsch & Sander, 1983)
Structure determination and refinement	CNS	(Adams et al, 1997)
	COOT	(Emsley & Cowtan, 2004)
	The CCP4 suite	(1994)
	Deformable elastic network (DEN) approach	(Schroder et al, 2010)
	Matthews	(Matthews, 1968)
	MolProbity Server	http://molprobity.biochem.duke.edu/
	MOSFLM	(Leslie, 1990)
	PHASER	(McCoy et al, 2007)
	PHENIX	(Adams et al, 2002)
	POINTLESS	(Evans, 2006; Grosse-Kunstleve et al, 2002)
	PROCHECK	(Laskowski et al, 1993; Laskowski et al, 1996)
	REFMAC5	(Murshudov et al, 1997)
	SCALA	(Kabsch, 1988)
	TLSMD server	(Painter & Merritt, 2006)

3. Methods

3.1. Molecular Biology techniques

3.1.1. Polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) was used to amplify DNA sequences encoding *T. acidophilum* XPB and Bax1. DNA polymerases are able to extent short DNA stretches, so-called primers, by adding deoxynucleotidetriphosphates (dNTPs) complementarily to the single stranded template DNA. Primers were designed to include recognition sites for the restriction enzymes *Pci*I and *Hind*III or *Nco*I and *Xho*I (Table 3.1; underlined) to enable cloning of the respective genes into the multiple cloning sites of the expression vectors. The forward primer in addition contains a start codon, the reverse primer comprises a stop codon (Table 3.1; both highlighted in bold). The following primer sequences were used (Table 3.1) in a PCR reaction containing the following agents (Table 3.2):

Table 3.1 Primer sequences used

Name	Sequence
XPB_*Pci*I_for	5‘ ATGGTCTCAC**ATG**TCCGATATCGTCTATTCTGGA 3‘
XPB_*Hind*III_rev	5‘ ACCGTCCTGAAGC**TTA**CGTACCACCATCAGT 3‘
Bax1_*Nco*I_for	5‘ AGGAGAAACC**ATG**GTCCCGGCAGAACTGAT 3‘
Bax1_*Xho*I_ rev	5‘ ATGGACTCGAG**TCA**GGTGGAACGGACTATCT 3‘

Table 3.2 PCR reaction set-up

	Stock concentration	Final concentration	amounts for 50 µl set-up
10 x *Pfu* Buffer with $MgSO_4$	10 x	1 x	5 µl
Template	200 ng/µl	200 ng	1 µl
Primer for	20 ng/µl	20 ng	1 µl
Primer rev	20 ng/µl	20 ng	1 µl
dNTPs	2 mM	60 µM	1.5 µl
Pfu DNA Pol	2 U/µl	1 U	0.5 µl
ddH_20			40 µl

The following PCR program was used (Table 3.3):

Table 3.3 PCR program for gene amplification

Step	Temperature	Time	Cycles
Initial Denaturation	95 °C	2 min	
Denaturation	95 °C	1 min	30 cycles
Annealing	55 °C	1 min	
Extension	72 °C	3 min	
Final Extension	72 °C	15 min	
cooling	4 °C	forever	

A 5 µl sample of the PCR reactions was analyzed on a 1 % (*w/v*) agarose gel. The PCR product was then purified using the PCR purification kit (Macherey-Nagel) according to the manufacturer's protocol.

3.1.2. Restriction digestion

PCR products for XPB and Bax1 were digested at 37 °C for 3 h with the restriction enzymes *Pci*I and *Hind*III or *Nco*I and *Xho*I, respectively, and destination vector pETM-11 with the combination of *Nco*I and *Hind*III or *Nco*I and *Xho*I. Restriction enzymes *Pci*I and *Nco*I produce compatible sticky ends and can thus be interchanged. The endonucleases were applied in NEB-buffer #2 according to the manufacturer's instructions. The destination vectors were additionally treated with alkaline phosphatase for 1 h at 37 °C to prevent reannealing of the vector.

3.1.3. Ligation

Sticky ends produced by the restriction enzymes allow the correct positioning of the insert into the destination vector in terms of direction and number. To subsequently generate phosphodiester bonds between vector and insert, Quick T4 DNA Ligase (Fermentas) was used in 2 x Ligase buffer and incubated for 5 min at room temperature (RT). Ligation reactions are then transformed into an *E. coli* host keeping strain such as XL1-blue or DH5α.

3.1.4. Transformation

To transform *E. coli* cells, 50-100 µl chemically competent cells were incubated with 1-2 µl plasmid DNA (miniprep scale; generally 50-100 ng/µl), 25 µl ligation reaction or 25 µl site-directed mutagenesis reactions for 30 min on ice. A heat shock (90 sec, 42 °C) was performed to perforate the cell wall and allow the DNA to enter the cytoplasm. 600 µl LB-medium were added and cells were grown for 45 min (37 °C, 600 rpm). The cells were centrifuged at 1,000 *g* for 2 min and resuspended in 100 µl supernatant. This suspension was then applied to LB-agar plates containing the appropriate antibiotics to select for cells which contain the plasmid DNA. The agar plates were incubated at 37 °C over night.

3.1.5. Plasmid isolation

A single colony grown on a LB-agar plate was picked and transferred to a 5 ml liquid culture containing LB and the appropriate antibiotics. After the cells were grown at 37 °C, 200 rpm over night to form a dense culture, the suspension was centrifuged at 4,000 rpm for 10 min to pellet the cells. The cells were then lysed and DNA was extracted using the Plasmid Kit (Macherey-Nagel) according to the manufacturer's instructions.

3.1.6. Colony PCR

In order to screen for positive clones carrying the desired gene, colony PCR was applied using the same primer sequences as for cloning. Plasmid DNA from a single colony was isolated and transferred into a PCR reaction as described above (Table 3.2, Table 3.3).

3.1.7. Site-directed mutagenesis

To introduce an exchange of amino acids at a specific site, site-directed mutagenesis on the DNA level was applied. Two complementary primers were designed both carrying the mutated sequence. A two-step protocol was carried out with eight PCR cycles containing just one primer each to avoid the formation of

primer-dimers (Wang & Malcolm, 1999). The single reactions including forward and reverse primers were pooled and 20 additional PCR cycles were performed. The reactions were composed of the following reagents (Table 3.4):

Table 3.4 Site-directed mutagenesis reaction

	Stock concentration	Final concentration	Amounts for 12.5 µl reaction
Quick change reaction buffer	10 x	1 x	1.25 µl
Template	40 ng/µl	20 ng	0.5 µl
Primer	5 µM	0.5 µM	1.25 µl
dNTPs	10 mM	400 µM	0.5 µl
Pfu Polymerase	2 U/µl	1 U	0.5 µl
ddH_2O			8.5 µl

The following PCR program was used (Table 3.5):

Table 3.5 PCR program for site-directed mutagenesis

Step	Temperature	Time	Cycles
Initial Denaturation	95 °C	30 sec	
Denaturation	95 °C	30 sec	8 cycles for each primer reaction, 20 cycles for combined reaction
Annealing	55 °C	1 min	
Extension	68 °C	6 min	
Final Extension	68 °C	10 min	
cooling	4 °C	forever	

Parental plasmids were then digested using *Dpn*I specifically recognizing methylated DNA. The reaction was then transformed into the *E. coli* host keeping strain DH5α. DNA was extracted from single colonies and checked for the introduced mutation using DNA sequence analysis at MWG-Biotech AG (Ebersberg) or Seqlab - Sequence Laboratories Göttingen GmbH (Göttingen).

3.1.8. Labeling and duplexing of single-stranded DNA

For ^{32}P-labeling, ssDNA was 5' end labeled for 10 min at 37 °C in a 25 µl reaction containing 200 nM ssDNA, 1 µCi/µl [γ ^{32}P]-ATP (Hartmann Analytic), 1 x

forward buffer and 10 U T4 Polynucleotide Kinase (Invitrogen). Alternatively, ssDNA was 3' end labeled for 1 h at 37 °C in a 50 µl reaction comprising 200 nM ssDNA, 25 µCi [α ^{32}P]-ATP (Hartmann Analytic), 5 x TdT buffer and 10 U Terminal transferase (Invitrogen). ssDNA was purified using MicroSpin™ G 25 columns (GE Healthcare) to remove unincorporated [γ ^{32}P]-ATP or [α ^{32}P]-ATP and subsequently annealed in the presence of 0.1 x TE buffer containing 100 mM KCl by heating to 80 °C and subsequent cooling slowly to RT. Fluorescently labeled ssDNA was purchased from Metabion or IDT and annealed in 0.1 x TE-buffer containing 100 mM KCl.

3.1.9. Generation of DNA substrates suitable for Atomic Force Microscopy (AFM)

For AFM, a 522 basepair (bp) DNA fragment (*uvrb* gene from *B. caldotenax*) was amplified by PCR using the primers UvrB_AFM__for (5' TCCGGACGATTTTTTGATC 3') and UvrB_rev5 (5' GACCAATACATCGTATTTGCC 3'). To generate a 32 nucleotide 3' overhang, the PCR product was incised using the nickase *Nt.BstNBI* (New England Biolabs) according to the manufacturers' protocol. A 30 nucleotide 5' overhang was produced by cutting the PCR product with the nickase *Nt.AlwI* (New England Biolabs). To remove the incised oligonucleotide from the DNA substrate 5 cycles of heating to 65 °C and spinning in a centrifuge concentrator (with a molecular weight cutoff of 50 kDa) were performed.

3.2. Protein expression

E. coli cells were transformed with an expression plasmid comprising the desired gene under control of the T7-promotor. Thus, protein expression can be induced by adding lactose or the allolactose analogue IPTG. IPTG is not consumed by bacterial cells and therefore only small amounts are required to induce protein expression. However, in contrast to IPTG, lactose and its metabolite allolactose does not stress the cells and thus enhances both bacterial growth and protein expression (Studier, 2005). In the so-called auto-induction

medium, glucose prevents the uptake of lactose by *E. coli* cells by catabolite repression. This mechanism allows the cells to grow to a certain density by consuming glucose. The lack of glucose then leads to the uptake of lactose and thereby induction of protein expression.

Since the different expression protocols are optimized for each protein, the detailed procedures are included in the Results section.

3.3. Protein purification

3.3.1. Cell lysis

UvrB expressing cells, UvrC expressing cells, and both XPB and Bax1 expressing cells were resuspended in UvrB-AC/GF, UvrC-AC and IMAC-A buffer (5 ml per g cells), respectively, and supplemented with DNaseI (1 µl per 50 ml suspension). The suspension was then lysed at 1.8 kbar in a cell disruptor (constant cell system). The lysate was centrifuged at 50,000 *g*, 4 °C, for 30 min. The supernatant was then applied to the first chromatography step.

3.3.2. Chromatography

Proteins were purified using a series of chromatography steps starting with an affinity chromatography as a first capture step. Finally, a polishing step is performed to not only remove contaminating proteins but also to separate different oligomeric states of the protein.

3.3.2.1. Affinity chromatography

The affinity chromatography is usually employed as a first capture step. Naturally occurring or tag-mediated affinities of proteins can be exploited to specifically and selectively bind the desired protein to the matrix of chromatography columns. Protein elution can be achieved by the addition of competitors, ligands, high concentration of salt or by cleaving the tag from the protein.

Immobilized metal affinity chromatography (IMAC)

Hexahistidin (His_6)-tags exert a strong affinity to Ni^{2+}-ions chelated by iminodiacetic acid (IDA) groups. Proteins containing a His_6-tag thus can be bound to the column and eluted by the addition of the competitor imidazole which is chemically and structurally similar to histidine residues.

The supernatant of the cleared cell lysate was loaded on a pre-equilibrated Ni-MAC column (CV = 1 ml; Novagen) with a flow rate of 1 ml/min. Both the protein sample and the binding buffer IMAC-A were supplemented with 50 mM imidazole to prevent unspecific binding. The column was washed with approximately 20 CV binding buffer to remove unbound sample until a stable baseline of the UV-absorption was achieved. The desired protein was then eluted by a gradient up to 500 mM imidazole.

Chitin-Intein system

The IMPACT^{TM} system (New England Biolabs) provides expression vectors to generate fusion proteins consisting of the desired protein and an intein-tag including a chitin binding domain. This chitin binding domain specifically binds to chitin beads. Thiols such as DTT induce the autocatalytic activity of inteins to cleave themselves from the target protein, which is thereby eluted from the chromatography column whereas the intein-chitin-tag stays bound on the chitin beads (Chong et al, 1997).

The supernatant of the cleared cell lysate was loaded onto a pre-equilbrated column (CV = 20 ml; New England Biolabs) and incubated for 2 h at 4 °C in order to improve binding. The column was then washed extensively with approximately 50 CV buffer UvrB-AC or UvrC-AC to remove unbound proteins. The cleavage buffer containing 100 mM DTT was applied to the chitin beads and incubated for 48 h at RT to ensure complete cleavage of the intein. The target protein was eluted by rinsing the column with buffer UvrB-AC or UvrC-AC.

3.3.2.2. Ion exchange chromatography

The net charge of a protein at certain pHs is exploited in ion exchange chromatography. At basic pH, proteins with a positive net charge ionically interact with cation exchange chromatography columns. The protein can be eluted by changing the pH or by increasing the salt concentration.

Protein purified by affinity chromatography was diluted with binding buffer to achieve a salt concentration of 60 mM NaCl and subsequently loaded onto a HiTrap SP HP cation exchange column (CV = 1 ml; GE Healthcare) with a flow rate of 1 ml/min. The column was washed with approximately 20 CV binding buffer to remove unbound proteins. The target protein was then eluted by a gradient up to 1 M NaCl.

3.3.2.3. Size-exclusion chromatography

As the final purification step, typically size-exclusion chromatography was perfomed. The column consists of a spherical composite of cross-linked agarose and dextran matrix which allows the separation of molecules with a molecular weight of 10-600 kDa. Thus, aggregates as well as higher oligomers can be separated from monomeric protein.

The concentrated protein sample was applied to a pre-equilbrated HiLoad™ 16/60 Superdex™ 200 pg or HiLoad™ 26/60 Superdex™ 200 column pg (CV = 124ml and 330 ml, respectively; GE Healthcare) and isocratically eluted with buffer. The presence and purity of the target protein in the elution fractions was analyzed by SDS-PAGE. Protein concentrations were determined by UV-spectrophotometry at 280 nm. Protein samples were concentrated using Centricon® concentrators (Millipore) with suitable molecular cut-offs, split in 25-50 µl aliquots, frozen in liquid nitrogen and stored at -80 °C until use.

3.4. Spectroscopy methods

3.4.1. Determination of Protein and DNA concentrations by UV-spectrophotometry

The concentration of protein and DNA samples was measured using a NanoDrop System (Peqlab) exploiting the law of Lambert-Beer (Pfeiffer & Liebhafsky, 1951):

Equation 1

$$c = \frac{A}{\varepsilon \cdot d}$$

with *A* being the Absorption at a wavelength of 280 nm or 260 nm for protein or DNA, respectively, *c* the concentration in mol/l, *d* the length of the pathway in cm and ε being the extinction coefficient in $M^{-1} cm^{-1}$. For proteins, ε is calculated according to the following equation:

Equation 2

$$\varepsilon_{280} = n_{Tyr} \cdot \varepsilon_{Tyr} + n_{Trp} \cdot \varepsilon_{Trp} + n_{Cys} \cdot \varepsilon_{Cys}$$

with ε_{280}(Tyr) = 1.490, ε_{280}(Trp) = 5.500, ε_{280}(Cys) = 125 (*ProtParam*, www.ExPASy.org). For DNA, ε was determined by the nearest neighbor method (Cantor et al, 1970; Fasman, 1975).

3.4.2. Dynamic Light Scattering

Dynamic Light Scattering (DLS) was used to determine the particle size in solution. The oligomeric state of a protein is often in an equilibrium which is reestablished again after size-exclusion chromatography and after concentrating the protein. To investigate whether the protein is monomeric or adopts different oligomeric states, the distribution of particle size in solution was analyzed by DLS. Laser light is scattered from dissolved macromolecules or suspended particles undergoing random Brownian motion resulting in fluctuations in light intensity (Murphy, 1997). The scattering fluctuations within a period of time are directly related to the translational diffusion coefficient of the scattering particles (Philo, 2006). The direct correlation of diffusion and size of a particle is used to determine

the hydrodynamic radius. Assuming that globular proteins are spherical in shape, the Stokes-Einstein relation can be applied to deduce the apparent hydrodynamic radius R_h(app) (Wilson, 2003):

Equation 3

$$R_h(\text{app}) = \frac{kT}{6\pi\eta D(c)}$$

with the concentration-dependent translational diffusion coefficient D(c), k is the Boltzmann constant, T is the absolute temperature in Kelvin and η the viscosity of the solvent.

Protein samples were extensively centrifuged for 30 min at 16,000 *g*, 20 °C, immediately filled into a quartz cuvette and measured by dynamic light scattering (DynaPro, Wyatt Technology) over 10 measurements at 20 °C. For crystallization purposes, the protein solution ideally should exert a polydispersity lower than 20 %.

3.4.3. Circular Dichroism (CD) Spectroscopy

CD-spectroscopy was applied to study the folding of proteins (Kelly et al, 2005). Circular polarized light is absorbed by chiral substances in solution and the resulting ellipticity is plotted against the wavelength. Tertiary structure elements such as α-helices and β-strands lead to characteristic CD-spectra. CD-spectra can thus be deconvoluted to estimate the content of structure elements or unfolded parts in a protein. In this study, CD-spectroscopy was utilized to verify that point mutations introduced into the Bax1 protein do not affect the overall structure of the protein. Moreover, the two identified XPB/Bax1 complexes were compared.

Measurements were performed at 20 °C in a 0.1 cm quartz cuvette using a J-810 CD-spectropolarimeter (Jasco). The protein was diluted in a buffer containing 20 mM sodium-phosphate, pH 8.0, to a final concentration of 0.1 mg/ml. Spectra were registered at wavelengths from 260 nm to 190 nm and five spectra were accumulated to optimize the signal to noise ratio. The band width was adjusted to 2.0 nm and a sensitivity of 100 mdeg was used. The scan speed was 20 nm min^{-1}, the time response 1 s and the data pitch 0.1 nm.

3.5. Biochemical Assays

3.5.1. Electrophoretic mobility shift assays (EMSA)

Electrophoretic mobility shift assays (EMSAs) were performed to study protein-DNA interactions. This method employs the different mobility of free DNA in comparison to DNA which is bound to a protein in native polyacrylamide gels (**Figure 3.1**). The DNA can be either detected by ^{32}P radio-isotope or fluorescent labeling.

DNA was incubated with protein in an appropriate reaction buffer with a total volume of 10 µl. The reaction was supplemented with sample buffer containing glycerol and color dyes prior to gel-electrophoresis on native polyacrylamide gels. Gels were exposed to a phosphorimager screen over night, visualized by a Personal Molecular Imager (PMI) System (BioRad) and quantified employing the software Quantity One (BioRad). Using DNA labeled with fluorescein, the polyacrylamide gels were directly visualized by the PharosFX™ imager system.

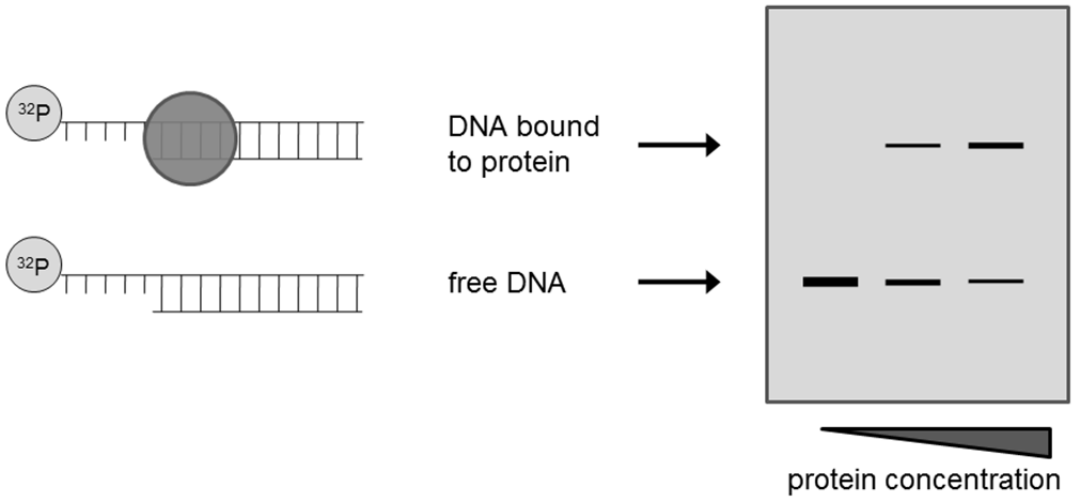

Figure 3.1 Schematic representation of an electrophoretic mobility shift assay. Free DNA in comparison to DNA which is bound to a protein (depicted as a filled circle) exhibits a higher mobility in native polyacrylamide gels. With increasing protein concentration, the percentage of protein-bound DNA will rise if a complex is formed.

3.5.2. Bio-Layer interferometry

Bio-Layer interferometry was performed with the Octet Red System (Forté Bio) using streptavidin coated biosensors. The DNA top strand NDT (Table 2.1) was

5’ biotinylated and annealed to form either 3’ overhang or double-stranded DNA (dsDNA) substrates. A 3’ biotinylated version of the DNA substrate NDT was used to form 5’ overhang substrates. DNA substrates and proteins were diluted in a buffer containing 20 mM MES, pH 6.5, 150 mM NaCl, 10 mM DTT, 1 mg/ml BSA and 10 mM $CaCl_2$.

Streptavidin biosensors were coated with the biotinylated DNA substrate by incubation for 10 min in a 100 nM DNA solution followed by a 2 min washing step in buffer. The biosensors were then transferred into protein solution for 10 min to allow association and subsequently conveyed into buffer to detect protein dissociation. All steps were performed at 37 °C and 1000 rpm in a 96-well plate containing 200 µl solution in each well. From changes in light interference over time upon binding and dissociation events, k_{on}, k_{off}, and the equilibrium dissociation constant K_D were derived using a 1:1 binding model (ForteBio OctetRED Evaluation software 6.1).

3.5.3. Incision assays

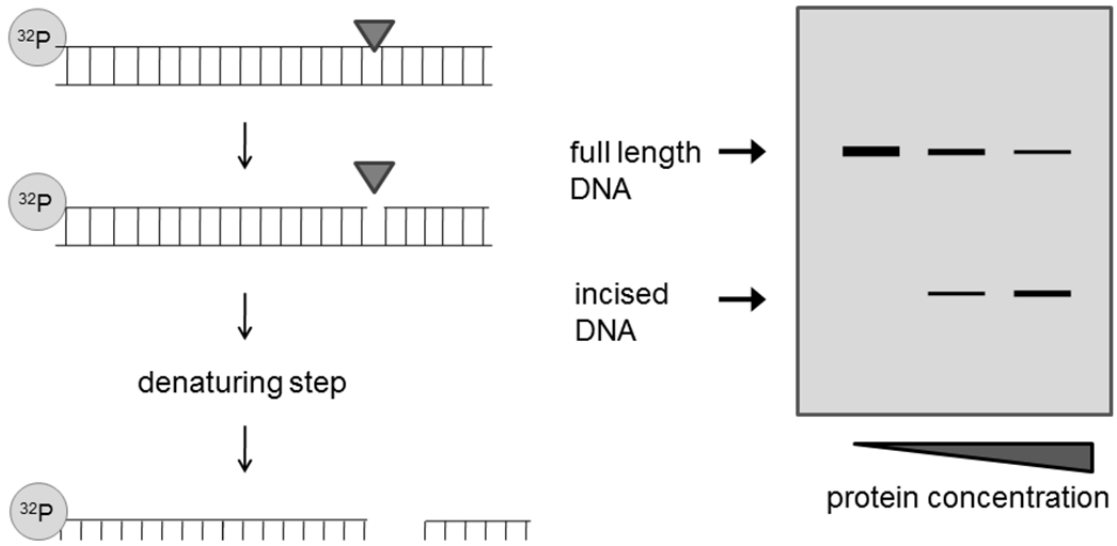

Figure 3.2 Schematic representation of an incision assay. Endonucleases (depicted as a triangle) cleave phosphodiesterbonds within the DNA substrate. A denaturing step is necessary to separate the duplex DNA. On the denaturing polyacrylamide gel, full length and incised DNA are separated by their different running velocity.

To investigate DNA endonuclease activity, incision assays were performed by incubation of protein and DNA substrates in an appropriate reaction buffer as schematically depicted in **Figure 3.2**. Urea sample buffer was added and samples were heated to 90 °C for 10 min prior to electrophoresis at 300 V on a denaturing 15 % polyacrylamide gel. Gels containing ^{32}P-labeled DNA were exposed to a phosphorimager screen over night, fluorescently labeled DNA comprising gels were directly visualized by the PharosFX^{TM} Imager System (BioRad) and quantified employing the software Quantity One (BioRad).

3.5.4. Helicase assays

To investigate DNA unwinding activity, helicase assays were performed by incubation of protein and DNA substrates in an appropriate reaction buffer as shown in **Figure 3.3**.

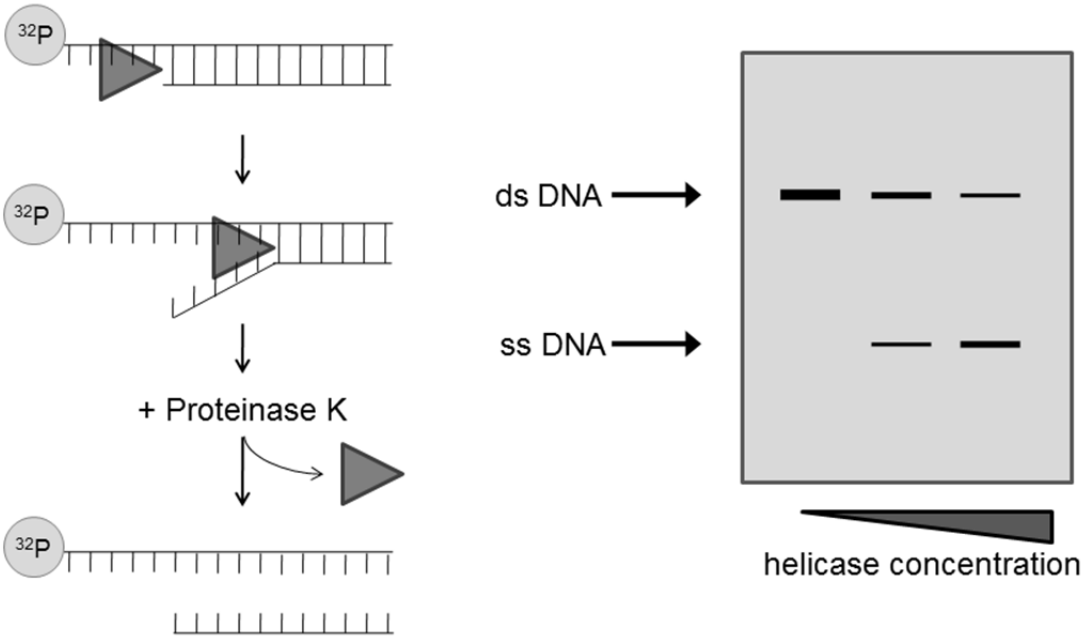

Figure 3.3 Schematic representation of a helicase assay. Helicases (shown as triangle) unwind dsDNA to form ssDNA. The helicase is digested by proteinase K to allow analyses of protein-free ssDNA and dsDNA on native polyacrylamide gels.

Stop buffer containing 10 mM TrisHCl, pH 8.0, 5 mM EDTA, 0.5 % SDS, 1 mg/ml Proteinase K and 5 µM competitor DNA was added and incubated for 15 min at RT to digest the helicase and to prevent the unwound DNA from

re-annealing. Then, 6 x sample buffer (Invitrogen) was added and samples were separated at 120 V on an 8 % native polyacrylamide gel. Gels containing ^{32}P-labeled DNA were exposed to a phosphorimager screen over night, fluorescently labeled DNA comprising gels were directly visualized by the PharosFX^{TM} Imager System (BioRad) and quantified employing the software Quantity One (BioRad).

3.5.5. ATPase assays

Hydrolysis of ATP provides energy for many molecular processes. Helicases, for example, consume ATP to be able to unwind DNA (Singleton et al, 2007). ATPase activity was determined using a coupled enzyme assay (Panuska & Goldthwait, 1980). Phosphoenolpyruvate is dephosphorylated to pyruvate by pyruvate kinase transferring the phosphate group to ADP. Pyruvate is further processed to *L*-Lactate by the *L*-Lactate dehydrogenase oxidizing NADH to NAD$^+$. This step can be monitored in an UV-spectrophotometer as NADH in contrast to NAD$^+$ absorbs light at a wavelength of 340 nm. Since no ADP is added to the reaction, this process is running only in the presence of an ATPase consuming ATP and simultaneously releasing ADP which can then serve as a substrate for the pyruvate kinase (schematically depicted in **Figure 3.4**). In this assay, the rate-limiting step needs to be restricted to the ATPase to obtain turn-over rates.

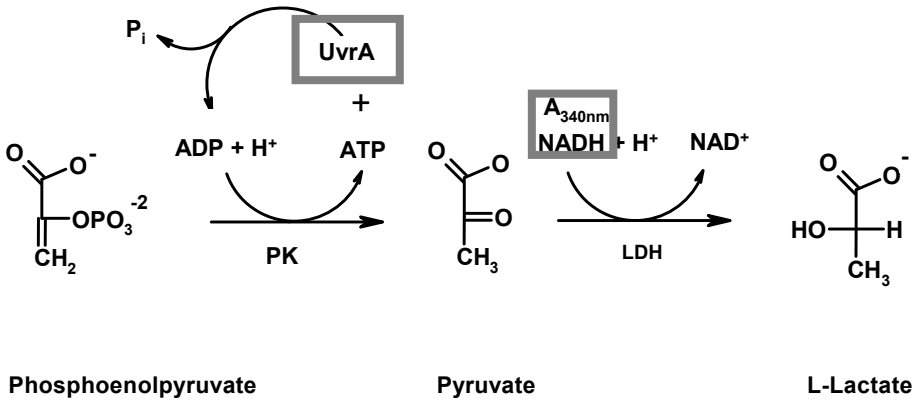

Figure 3.4 Schematic reaction scheme of the ATPase assay. Only in the presence of an active ATPase, ADP is provided for the reaction mechanism which consumes NADH. The oxidation of NADH can be monitored in spectrophotometrically as NADH in contrast to NAD$^+$ absorbs light at 340 nm.

3.5.6. Limited proteolysis

Limited proteolysis was used to monitor protein folding since partially unfolded proteins are more prone to proteolytic digestion compared to folded proteins (Heiring & Muller, 2001) This method also allows identifying stable protein fragments, flexible linker regions or even different conformations of a protein.

To study the stability of Bax1 and to determine conformational differences between XPB/Bax1 complexes, protein samples (4 µM) were subjected to limited proteolysis using 100 nM trypsin. Experiments were carried out in buffer containing 20 mM TrisHCl, 500 mM NaCl, pH 7.5 at room temperature for the times indicated. Samples were denatured in 5 x SDS sample buffer and subsequently analyzed by SDS-PAGE.

3.6. Biophysical approaches

3.6.1. Analytical gelfitration

To investigate protein-protein interactions, analytical size exclusion chromatography was performed using a Superdex^{TM} 200 10/300GL or Superdex^{TM} 75 10/300GL column (GE Healthcare).

Standard proteins were used to calibrate the size-exclusion columns (Gel filtration calibration Kit, GE Healthcare). The gel phase distribution coefficient K_{av} was determined for each standard protein using the following equation:

Equation 4

$$K_{av} = \frac{V_e - V_0}{V_t - V_0}$$

with V_e being the elution volume, V_0 was the void volume and V_t the total bed volume of the column. K_{av} was then plotted against the logarithm of the molecular weight log(MW) providing a linear calibration curve. For the Superdex^{TM} 200 10/300GL column, the calibration curve is described by the following equation:

Equation 5

$$K_{av} = -0.319 \times \log(MW) + 1.969$$

Calibration of the Superdex^{TM} 75 10/300GL column results in Equation 6:

Equation 6

$$K_{av} = -0.352 \times \log(MW) + 0.748$$

The molecular weight could subsequently be determined by linear regression. All proteins were eluted isocratically with 20 mM TrisHCl, 500 mM NaCl, pH 7.5 and collected in 0.5 ml fractions which were then analyzed by SDS PAGE.

3.6.2. Analytical ultracentrifugation

Analytical ultracentrifugation (AUC) was employed to analyze protein complexes in solution with regard to their molecular size and shape. The particle size is reflected by the sedimentation coefficient, whereas the shape of particles is indicated by the frictional ratio. Therefore, the stoichiometry and conformation of multi-protein complexes can be examined.

Sedimentation Velocity (SV) Analytical Ultracentrifugation was performed on purified Bax1, XPB and XPB/Bax1 complexes at a concentration of 1 µM each. Experiments were conducted with a Beckman Optima XL-I analytical ultracentrifuge (Beckman Coulter) using an eight-hole An-50 Ti rotor at 40,000 rpm and 20 °C, with 400 µl samples in standard double-sector charcoal-filled Epon centerpieces equipped with sapphire windows. Data were collected in continuous mode at a step-size off 0.003 cm using absorption optical detection at a wavelength of 280 nm. Data were analyzed using the software SEDFIT to determine continuous distributions for solutions to the Lamm equation *c*(*s*), as described (Schuck, 2000). Analysis was performed with regularization at confidence levels of 0.68 and floating frictional ratio (*f/fo*), time-independent noise, baseline, and meniscus position, to RMSD values of 0.0038 to 0.016.

3.6.3. Atomic Force Microscopy

Atomic Force Microscopy (AFM) was applied to study single molecules such as protein-protein and protein-DNA assemblies (Fotiadis et al, 2002; Yang et al, 2003). In AFM, a biological sample surface is mechanically raster-scanned and

translated into a 3D-image. The resulting surface topography reveals qualitative and quantitative information about stoichiometry, conformation and structure of biomolecular assemblies (Hirano et al, 2008; Yang et al, 2003).

To study the interaction of proteins and protein complexes on DNA on individual molecules, different DNA substrates and different incubation conditions were employed. DNA substrates were heated to 65 °C for 10 minutes and slowly cooled down to room temperature to remove any salt crystals formed during storage. Proteins (0.5 and 1 µM) and DNA (100 nM for linear DNA fragments, 20 nM for circular DNA) were incubated in AFM incubation buffer (20 mM TrisHCl pH 7.5, 150 mM NaCl, 10 mM DTT, 10 mM $CaCl_2$) for 30 minutes at 45 °C. Linear DNA samples were diluted 75-fold, circular DNA samples 20-fold in AFM deposition buffer (25 mM HEPES pH 7.5, 25 mM Na-acetate, 10 mM Mg-acetate) after incubation, immediately deposited onto freshly cleaved mica, rinsed with deionized water and dried in a gentle stream of nitrogen. All images were collected on an MFP-3D-BIO atomic force microscope (Asylum Research) in oscillating mode using Olympus OMCL-AC240 silicon probes with spring constants of ~2 N/m and resonance frequencies of ~70 kHz. Images were captured at a scan size of 2 × 2 μm^2, a scan rate of 0.5 Hz and a resolution of 1024 × 1024 pixels.

For analysis, AFM images were flattened to 1st order. Peak heights and AFM volumes were measured and calculated using Gwyddion and ImageSXM software. AFM volumes can be translated into protein molecular weights by calibrating the instrument with proteins of known molecular weight to derive a standard linear relationship. For our set-up, the calibration curve is described by the following equation:

Equation 7

$$V = 1.2 \times MW - 5.9$$

where V is the AFM volume and MW is the molecular weight. To determine the approximate molecular weights of DNA-bound protein complexes, we subtracted the DNA volume that is covered by a protein complex from the total complex volume. The DNA volume was calculated using the same length as that covered by the bound protein complex, the measured DNA height, and a width of 2 nm, and was on average (72 ± 6) nm^3 for the XPB/Bax1 complexes. Positions of

protein peaks on the DNA were measured using the NIH ImageJ software. The position distribution histograms are presented as the occurrence probability P_i versus position

Equation 8

$$P_i = \frac{n_i}{N_{bp,\,bins} \times \sum n_i}$$

where i is the position on the DNA, n_i is the number of complexes observed at position i, $\Sigma n_i = n$ is the total number of binding occurrences observed within the position range, and $N_{bp,bins}$ is the number of DNA base pairs in each position bin. From maxima in the distribution of relative occupancies, specificities for particular DNA strand internal positions can be determined. The specificity S for different DNA fragment end types was calculated from the ratio of relative occupancies of DNA ends and DNA strand internal positions (X) by the following equation:

Equation 9

$$S = N \times X + 1$$

(Yang et al, 2005), where N is the number of binding sites on the DNA (here N = 522 bp for the blunt end DNA substrate and N = 490 bp for the 3' and 5' overhang DNA substrates). For analysis of protein coverage of circular plasmid DNA, protein-DNA complexes were visually inspected and counted using the AFM software (Asylum Research on Igor Pro).

3.7. Crystallization

A single crystal is the pre-condition to solve a protein structure with X-ray crystallography. In vapor diffusion experiments, highly pure protein is mixed with reservoir solution and opposed to a much higher volume of reservoir to allow water to diffuse from the protein solution to attempt dilution of the reservoir or vice versa. The diffusion events result in a higher concentration of protein and precipitant in the crystallization drop thus facilitating nucleation and crystal growth. Crystallization screens are used to probe for many different conditions to finally identify a suitable condition to grow protein crystals.

Crystallization screens were set-up in a 96-well format with the crystallization robot (Honeybee, Cartesian Systems). Usually, 300 nl protein were mixed with 300 nl reservoir solution on a sitting drop next to 40 µl reservoir solution. For manual optimization, hanging drop vapor diffusion experiments were conducted by mixing 1 µl protein solution with 1 µl reservoir solution on a cover slide which is deposited upside down onto a 1 ml reservoir solution containing well. The crystallization conditions were checked on a regular basis for crystal growth.

3.8. X-ray crystallography

3.8.1. Data collection and structure determination

To determine the structure of a macromolecule by X-ray crystallography, the intensities of waves scattered from Bragg planes (referred to as the Miller indices *hkl*) in the crystal are measured. The square root of the detected intensities $[|I(hkl)|^{\frac{1}{2}}]$ is proportional to the structure factor amplitude |*F*(*hkl*)| of the reflection (*hkl*). To determine the electron density *ρ* at a position (*xyz*) in the unit cell of a crystal, the Fourier transformation is calculated by summation over all *hkl* planes (Taylor, 2003):

Equation 10

$$p(xyz) = \frac{1}{V} \sum_{hkl} |F(hkl)|\, e^{i\,\alpha_{hkl}} e^{[-2\pi i\,(hx+ky+lz)]}$$

with *V* being the volume of the unit cell, *i* is the imaginary number, and α_{hkl} is the phase angle associated with the structure-factor amplitude |*F*(*hkl*)|. Although the amplitudes can be determined, the phases are lost during the experiment. This fact is termed the phase problem in crystallography. To solve the phase problem, different approaches are available. Molecular replacement is used if a homologous search model is available. The isomorphous replacement method employs the additional information derived from heavy-atom scattering introduced by soaking or co-crystallization, whereas the anomalous diffraction method uses anomalous scattering elements such as selenium with are incorporated in the protein. In this study, the molecular replacement approach was applied.

The diffraction data set was collected from a single UvrBΔ4-DNA co-crystal at the synchrotron beamline BL14-1 the Berlin Electron Storage Ring Society for Synchrotron Radiation (BESSY). Data were indexed, integrated and scaled using the programs MOSFLM, POINTLESS and SCALA (Evans, 2006; Grosse-Kunstleve et al, 2002; Kabsch, 1988; Leslie, 1990). The structure was solved by molecular replacement with PHASER (McCoy et al, 2007) using a previously solved crystal structure of the UvrB-DNA complex (PDB: 2FDC, chain B) as a search model.

3.8.2. Refinement

Refinement consists of alternating cycles of manual model building and automated refinement of coordinates and *B*-factors according to the maximum likelihood approach. Model building in real space was performed employing COOT (Emsley & Cowtan, 2004) by fitting the structure model into the electron density map. Automated refinement was done in the reciprocal space to optimize the correlation of observed $|F_o(hkl)|$ and calculated structure-factor amplitudes $|F_c(hkl)|$. A qualitative measure of this correlation is reflected by the R_{work} and R_{free} factors.

Equation 11

$$R_{work,free} = \frac{\sum_{hkl}\big||F_o(hkl)| - |F_c(hkl)|\big|}{\sum_{hkl}|F_o(hkl)|}$$

For the calculation of the R_{free} factor, 5-10 % of the data are randomly selected and not included in the refinement procedure to avoid bias and over-interpreting the structure model (Brunger, 1997; Kleywegt & Jones, 1997).

For refinement of the UvrBΔ4-DNA complex, REFMAC5 (Murshudov et al, 1997), a combination of CNS (Adams et al, 1997) and the deformable elastic network (DEN) approach (Schroder et al, 2010), and PHENIX (Adams et al, 2002) was employed.

The DEN approach was established to improve the low-resolution data. In contrast to rigid body refinement, this method accounts for deformations of a high-resolution reference model (Schroder et al, 2010). The DEN potential can be optimized by varying the parameters w_{DEN} and γ as defined by following equations:

Equation 12

$$E_{total} = E_{geometric} + w_a E_{ML} + w_{DEN} E_{DEN}(\gamma)$$

The total energy function E_{total} consists of a weighted sum of a stereochemical energy function $E_{geometric}$, a maximum likelihood target function E_{ML} and the DEN potential E_{DEN} with its relative weights w_a and w_{DEN}. The elastic network energy term E_{DEN} is defined by Equation 13 referring to the sum of distance deviations to the power *p* over all selected atom pairs *i,j,*

Equation 13

$$E_{DEN}(\gamma, n) = \sum_{\text{N pairs } i,j} (d_{ij} - d_{ij}^0(\gamma, n))^p$$

where d_{ij} is the distance between atom pair *i* and *j* in the current atomic model and $d_{ij}^0(\gamma,n)$ is the corresponding equilibrium distance after the DEN update step *n* (Schroder et al, 2010).

Additionally, the weighting term w_{xc} between X-ray diffraction data and stereochemical restraints was optimized for refinement with CNS and used for both programs CNS and PHENIX (Adams et al, 2002; Adams et al, 1997).

Using PHENIX, TLS refinement was additionally applied and the model was split into 8 TLS mobility groups (comprising residues 3-119, 120-251, 252-274, 275-294, 295-415, 416-445, 446-512 and 513-596) as determined by the TLSMD server (Painter & Merritt, 2006).

3.8.3. Data evaluation

To evaluate the quality of the raw crystallographic data, the reliability factors R_{sym}, R_{meas} and $R_{p.i.m.}$ were calculated (Diederichs & Karplus, 1997; Weiss, 2001; Weiss & Hilgenfeld, 1997).

Equation 14

$$R_{sym} = \frac{\sum_{hkl} \sum_j \left| I_{hkl,j} - \langle I_{hkl} \rangle \right|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

Equation 15

$$R_{meas} = \frac{\sum_{hkl} \sqrt{\frac{n}{n-1}} \sum_{j=1}^{n} \left| I_{hkl,j} - \langle I_{hkl} \rangle \right|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

Equation 16

$$R_{p.i.m} = \frac{\sum_{hkl} \sqrt{\frac{1}{n-1}} \sum_{j=1}^{n} \left| I_{hkl,j} - \langle I_{hkl} \rangle \right|}{\sum_{hkl} \sum_{j} I_{hkl,j}}$$

After refining the structure to reasonable *R*-factors, the structure model was analyzed for its geometry by the MolProbity Server and PROCHECK (Chen et al, 2010; Laskowski et al, 1993; Laskowski et al, 1996). MolProbity and PROCHECK involve Ramachandran plot analysis to identify unusual or disallowed combinations of the main chain dihedral angles φ and ψ (Lovell et al, 2003). Deviations (given as root mean square deviations) from the ideal bond length and bond angle are reported as well.

All figures representing crystal structures were generated with PyMOL (DeLano, 2003). For correct structure representation, secondary structure elements were assigned using DSSP (Kabsch & Sander, 1983).

4. Results

4.1. Prokaryotic Nucleotide Excision Repair – Structural Analysis of the Damage Recognition and the Pre-Incision Complex

4.1.1. Expression and purification of UvrB and UvrBΔ4 from *B. caldotenax*

According to an already established protocol for the wildtype protein as well as the UvrBΔ4 variant (Skorvaga et al, 2002), *E. coli* BL21-CodonPlus®(DE3)-RIL cells were transformed with the pTYBI::uvrB or pTYBI::uvrBΔ4 plasmid providing an N-terminal intein-tag (Truglio et al, 2006b). The UvrBΔ4 variant corresponds to a truncated UvrB protein lacking domain 4 (residues Pro612 to Gly658). A single colony was used to inoculate a 100 ml $LB^{Amp/Cam}$ pre-culture. Cells were transferred into 2 l $LB^{Amp/Cam}$ and grown at 37 °C and 200 rpm until the optical density (oD_{600}) reached 0.5. Expression was induced by the addition of 0.3 mM IPTG and continued for 20 h at 15 °C. Cells were harvested by centrifugation and lysed in UvrB-AC/GF buffer by the use of a cell disruptor system. The lysate was cleared by centrifugation and the supernatant was then incubated with pre-equilibrated chitin beads in a batch approach at 4 °C for 2 h to improve binding. The beads were washed with buffer to remove unbound protein. UvrB (UvrBΔ4) was cleaved from its intein-chitin-tag by incubation with DTT for 48 h at RT. The protein was then eluted with buffer, concentrated in centrifuge concentrators (MW cutoff: 50,000 Da) and subjected to size-exclusion chromatography using a HiLoad™ 26/60 Superdex™ 200 pg column with UvrB-AC/GF buffer (**Figure 4.1**). Peak fractions were analyzed by SDS-PAGE for the presence and purity of UvrB (UvrBΔ4) and concentrated in centrifuge concentrators up to 50 mg/ml. The protein concentration was determined spectrophotometrically (UvrB: MW = 75.5 kDa, ε_{280} = 38740 M^{-1} cm^{-1}; UvrBΔ4: MW = 70.0 kDa, ε_{280} = 38740 M^{-1} cm^{-1}). 50 µl aliquots were flash frozen with liquid nitrogen and stored at -80 °C.

A representative size-exclusion chromatogram and corresponding gels are depicted in **Figure 4.1**. UvrBΔ4 elutes in a single peak at an elution volume of

~200 ml. On the SDS-gels, peak fractions show pure bands corresponding to UvrBΔ4 at approximately 70 kDa.

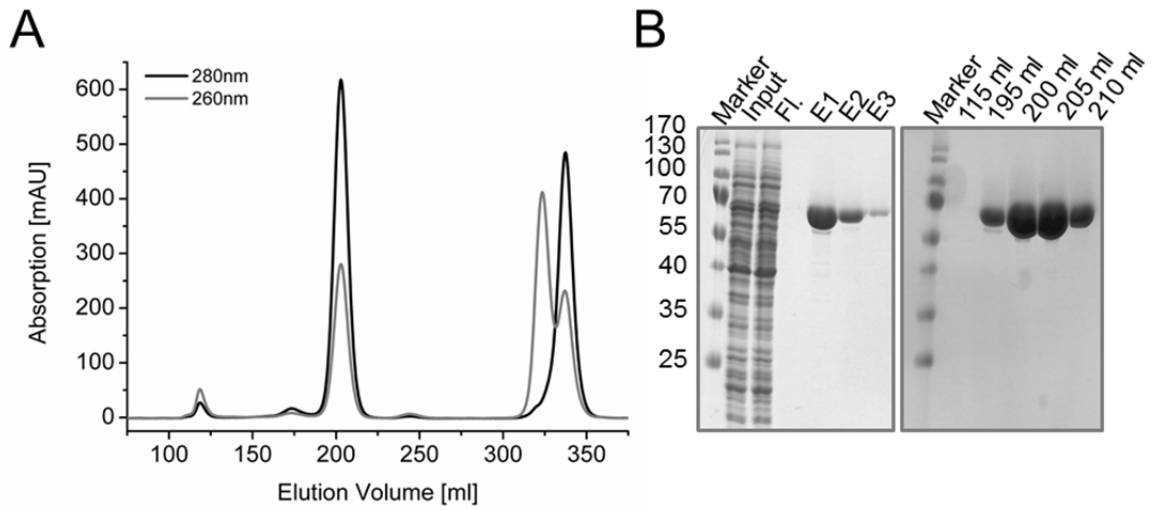

Figure 4.1 Chromatogram of the final size-exclusion chromatography of UvrBΔ4 and SDS-PAGE of the elution fractions. (A) UvrBΔ4 elutes from the HiLoad™ 26/60 Superdex™ 200 pg column at an elution volume of ~200 ml forming one defined peak. (B) The SDS-gel (left) shows the input, flow through (Fl.) and elution fractions (E1-E3) from the chitin beads. The SDS-gel in the right panel represents peak fractions from panel A revealing pure bands corresponding to UvrBΔ4 at approximately 70 kDa.

4.1.2. Expression and purification of UvrC from *B. caldotenax*

The $UvrC^{Bca}$ gene has been cloned in the expression vector pTXBI providing a C-terminal intein tag (Skorvaga et al, 2002). *E. coli* BL21-CodonPlus®(DE3)-RIL cells were transformed with pTXBI::$uvrC^{Bca}$ and grown on $LB^{Amp/Cam}$-agar plates. A 2 l ZYM-$5052^{Amp/Cam}$ culture was inoculated and grown at 37 °C and 200 rpm until oD_{600} = 0.3. The temperature was lowered to 15 °C and expression was continued for ~20 h. Cells were harvested by centrifugation and subsequently lysed in UvrC-AC buffer using the cell disruptor system. The cleared cell lysate was incubated with pre-equlibrated chitin beads at 4 °C for 2 h to improve binding of the fusion protein to the chitin beads in a batch set-up. The beads were washed extensively with approximately 50 CV buffer UvrC-AC prior to the addition of cleavage buffer containing 100 mM DTT. The solution was then incubated for 48 h at RT to cleave UvrC from the intein-chitin-tag. Protein was eluted, concentrated in centrifuge concentrators (MW cutoff: 30,000 Da) and subjected to size-exclusion

chromatography using a HiLoad™ 26/60 Superdex™ 200 pg column equilibrated in UvrC-GF buffer (**Figure 4.2**). Peak fractions were analyzed by SDS-PAGE and concentrated up to 50 mg/ml. The protein concentration was determined spectrophotometrically (UvrC: MW = 67.9 kDa, ε_{280} = 44240 M^{-1} cm^{-1}). 50 µl aliquots were flash-frozen with liquid nitrogen and stored at -80 °C.

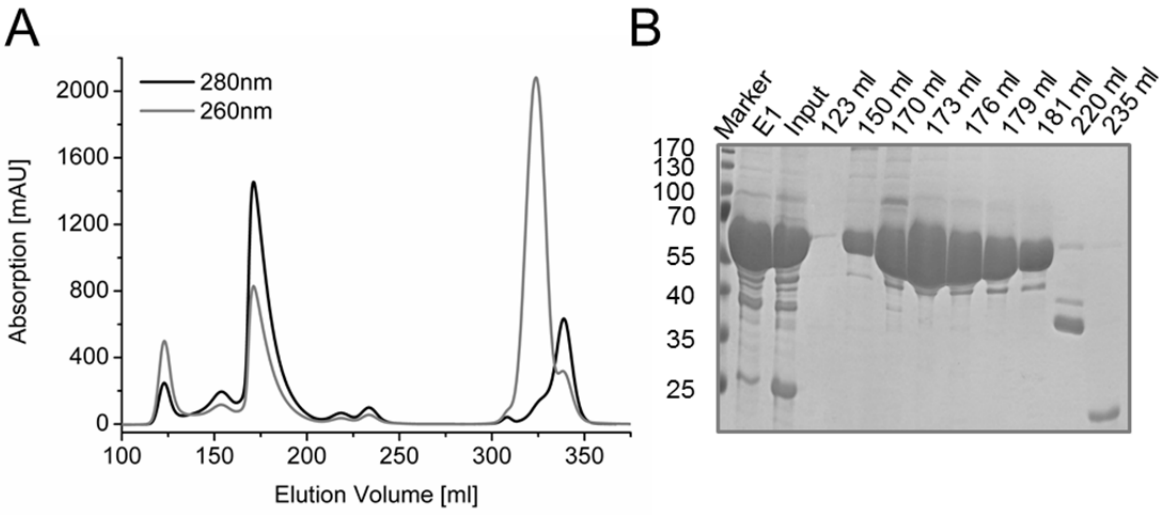

Figure 4.2 Chromatogram of the final size-exclusion chromatography of UvrC and SDS-PAGE of the peak fractions. (A) UvrC elutes in the size-exclusion chromatography (HiLoad™ 26/60 Superdex™ 200 pg column) in two peaks corresponding to monomeric and dimeric UvrC (172 ml and 150 ml, respectively). (B) The SDS-gel shows minor impurities eluting with the 68 kDa-band referring to UvrC. In addition to the peak fractions from the size-exclusion chromatography, elution fraction E1 from the chitin beads as well as the gel-filtration input sample is depicted.

The representative chromatogram (**Figure 4.2**) and the corresponding SDS-gel show that UvrC elutes in two peaks corresponding to monomeric and dimeric UvrC (172 ml and 150 ml, respectively). In addition to the UvrC band in the SDS-gel (~68 kDa) only minor impurities are visible.

4.1.3. Crystallization, data collection and crystal structure determination of the UvrB-DNA complex

To decipher the process of DNA damage recognition, UvrBΔ4 was co-crystallized with the DNA-substrate E, which contains a fluorescein adduct within a hairpin-bubble structure mimicking a NER-substrate to which UvrB is able to bind without the help of UvrA (Zou & Van Houten, 1999). UvrBΔ4 was used

instead of wildtype UvrB because it lacks its auto-inhibitory domain (domain 4; consistent with residues 612-658) thus exhibiting a higher affinity towards DNA (Wang et al, 2006). Protein (280 µM UvrBΔ4) and DNA were mixed in a 1:1.2 molar ratio and dialyzed against UvrB-dialysis buffer pre-heated to 45 °C. The dialysis set-up was slowly cooled to RT to allow the protein to bind its DNA substrate. Crystallization was performed immediately by hanging-drop vapor diffusion experiments at 20 °C. Diffracting crystals were obtained by mixing 1 µl crystallization condition containing 5 mM $ZnCl_2$, 100 mM Bicine pH 9.0, 9.4 % PEG 20,000 and 1 µl of UvrBΔ4-DNA solution. The drop was then opposed to a reservoir comprising 100 mM TrisHCl, pH 8.5, 500 mM NaCl and 20 % PEG 6,000 (**Figure 4.3**).

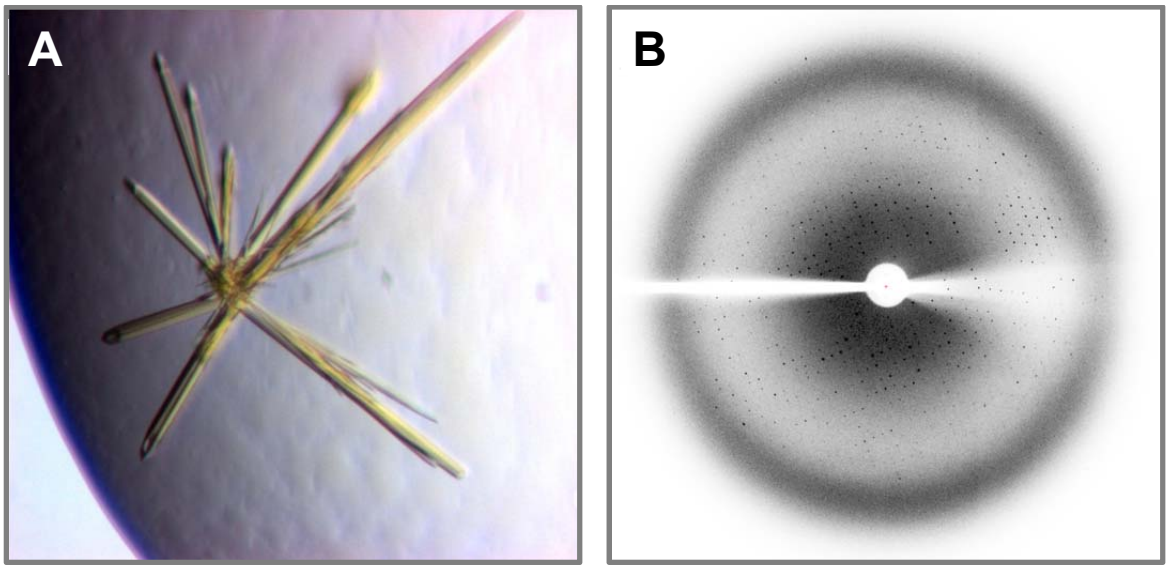

Figure 4.3 Crystallization and data collection of UvrBΔ4-DNA crystals. (A) Crystals of the UvrBΔ4-DNA complex. The yellow color indicates that fluorescein containing DNA is present in the crystal. (B) Diffraction pattern of the UvrBΔ4-DNA crystal.

The UvrBΔ4-DNA crystal was fished with a nylon loop, soaked in mother liquor supplemented with the cryoprotectant glycerol and flash-frozen in liquid nitrogen. Data were collected at the synchrotron beamline BL14.1 at BESSY and diffracted to a minimum of 3.2 Å resolution.

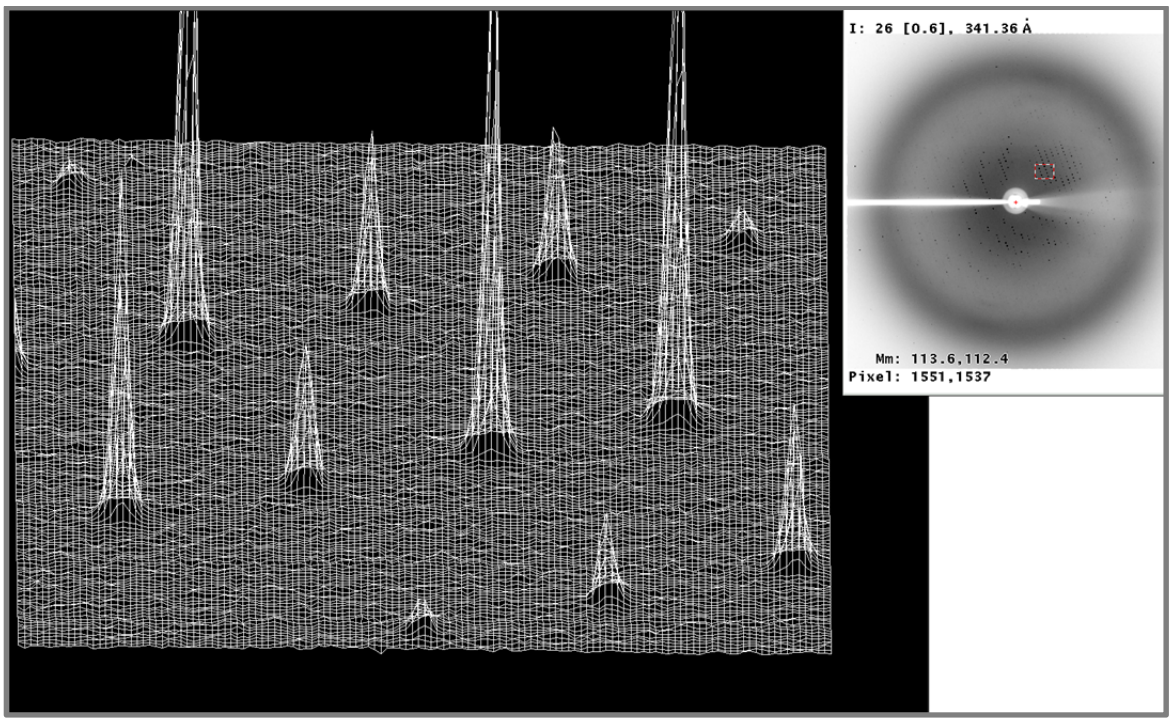

Figure 4.4 Diffraction spots of the UvrBΔ4-DNA crystal. Diffraction spots are enlarged and represented in a 3D-view with the spot intensities translated into height information.

Diffraction spots were indexed, integrated and scaled with MOSFLM, POINTLESS and SCALA (Kabsch, 1988; Leslie, 1990). The H test suggested that the data were twinned with a twinning fraction of 5 % described by the twinning operator h+k, -k, -l (Yeates, 1988). The data were processed to a resolution of 3.5 Å and the crystal structure was solved by molecular replacement with PHASER (McCoy et al, 2007) using the structure of the UvrB-DNA complex (PDB: 2FDC, chain B) as a search model.

The structure was refined by multiple cycles of model building and automated refinement using REFMAC5 (Murshudov et al, 1997), CNS/DEN (Adams et al, 1997) (Schroder et al, 2010) and PHENIX/TLS (Adams et al, 2002) as described in 3.8.2 and to reasonable *R*-factors (R_{work} = 25.8 %, R_{free} = 29.2 %). Further refinement steps in order to fix both bad rotamers and to improve the Ramachandran statistics for most favored dihedral φ/ψ angles did result in an over-fitting of the model indicated by a deviation between R_{work} and R_{free} of considerably more than 5 %.

The corresponding crystallographic data and refinement statistics are listed in Table 4.1.

Table 4.1 Crystallographic data collection and refinement statistics

Data collection	
X-ray source	BESSY-MX, BL14.1
Detector	Marmosaic 225 mm X-ray detector
Wavelength (Å)	0.95372
Resolution (Å)	50.89-3.50 (3.69-3.50) [a]
Space group	P 3_1 2 1
Cell parameters	$a = b$ = 152.27 Å, c = 80.04 Å $\alpha = \beta$ = 90°, γ = 120°
Number of molecules per ASU	1
Solvent content (%)	67.88
Matthews coefficient	3.3
Mean $I/\sigma(I)$	5.7 (1.9)
R_{sym} (%)	14.8 (71.2)
R_{meas} (%)	17.5 (84.3)
R_{pim} (%)	9.3 (44.7)
No. of reflections	47,351 (6,812)
No. of unique reflections	13,743 (1,966)
Redundancy	3.4
Completeness (%)	99.9 (100)
Average mosaicity	0.75
Wilson B-value(Å^2)	99.59

Refinement statistics (50.89-3.50 Å)	
No. of TLS groups used in refinement	8
R_{work} / R_{free} / R_{total} (%) [b]	25.84 / 29.22 / 26.00
No. of atoms	4443
No. of residues	594
Ligands	DNA (73 atoms, 5 residues)
Overall mean B-value (Å^2)	137.01
mean B-value protein(Å^2)	135.01
mean B-value DNA (Å^2)	258.6

R.M.S.D from ideal geometry	
bond length (Å)	0.005
bond angles (°)	0.821
Ramachandran statistics [c]	77.4 / 22.6 / 0.0 / 0.0

[a] Values given in parentheses in this column are for the highest resolution shell
[b] *R*-factors are defined in 3.8. The R_{free} was calculated from 4.91 % of the data that were randomly removed before the refinement was started.
[c] The Ramachandran statistics were obtained with PROCHECK. Reported is the percentage of residues in the most favored, additional allowed, generously allowed and disallowed areas of the Ramachandran plot.

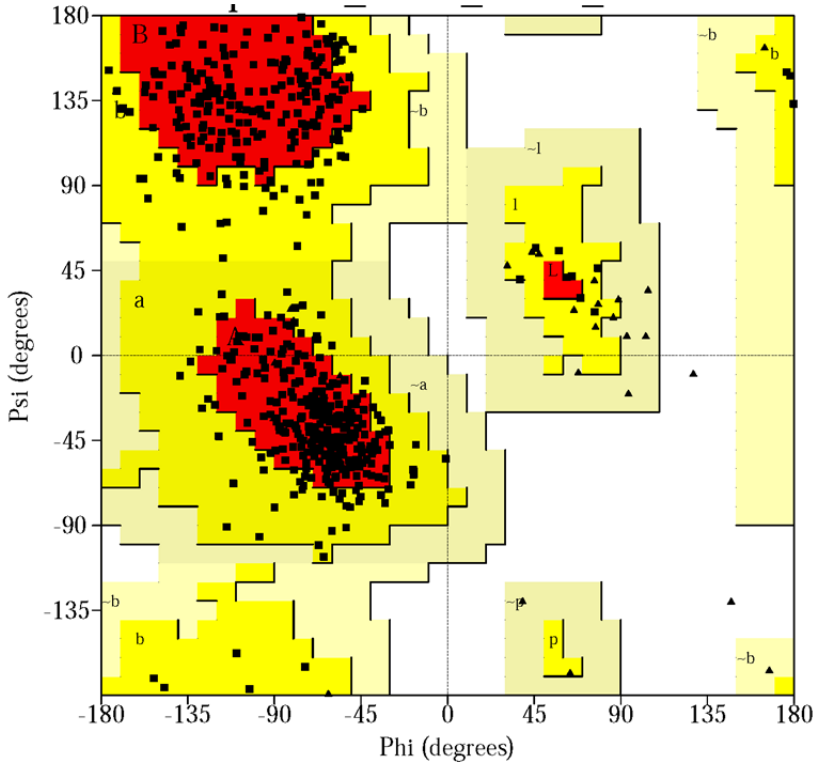

Figure 4.5 Ramachandran Plot of dihedral angles φ and ψ. For each residue, the corresponding φ and ψ are plotted and depicted as squares (triangles are related to Gly residues). Red areas delineate the most favored regions, yellow areas define the allowed regions and the light yellow area corresponds to the generously allowed region.

The CCP4 program suite was used to determine *B*-factors as well as the Matthews coefficient, solvent content and the number of molecules per asymmetric unit (ASU) (Matthews, 1968). It should be noted that, consistent with the high solvent content of 67.9 % and the elevated Wilson *B*-value of 99.6 Å^2, the resulting *B*-values of the model are quite high. Increased *B*-values are observed especially for the DNA bases where only residual density is present, either as a measure of high mobility or a very low occupancy in this region. On account of the

low resolution of the structure, the structural restraints thus were kept very tight during refinement. Consequently, the r.m.s.d. values for bond lengths and bond angles finally resulted in 0.005 Å and 0.821°, respectively. In general, r.m.s.d. values of ≤ 0.02 Å and ≤ 1.90° are tolerated.

4.1.4. Crystal structure of the UvrBΔ4-DNA complex

The crystal structure of UvrBΔ4 shows a 4-domain architecture comprising two RecA-like helicase domains (domains 1a and 3; colored in yellow and red in **Figure 4.6**, respectively) and a protruding β-hairpin (highlighted in cyan in **Figure 4.6**) inserted in domain 1a.

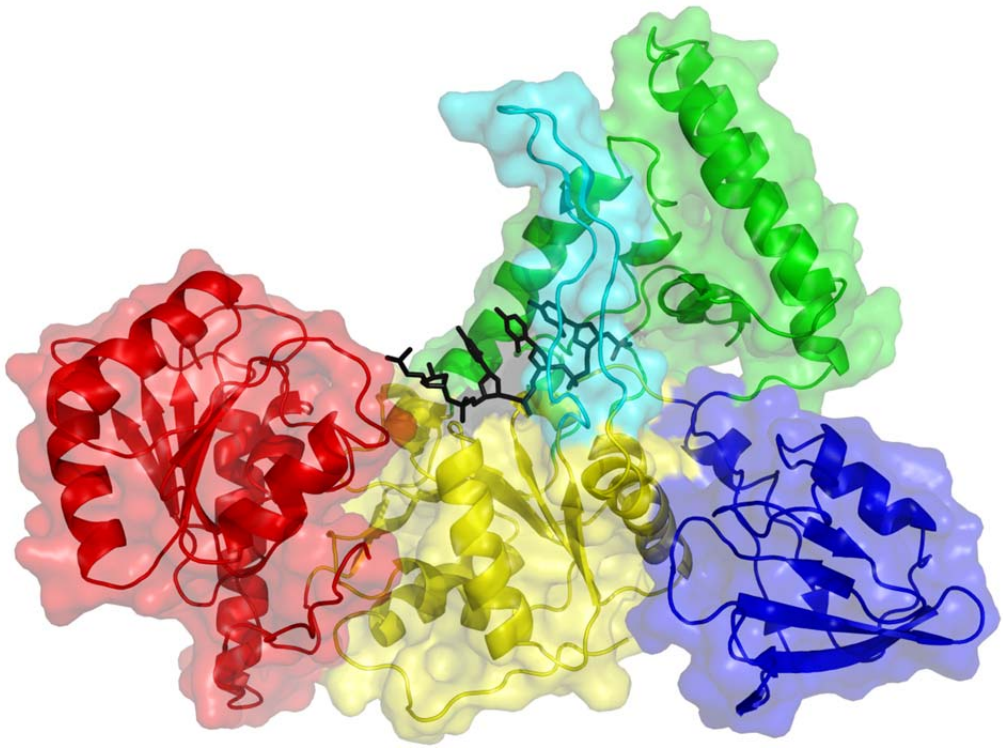

Figure 4.6 Crystal structure of the UvrBΔ4-DNA complex. The overall structure of the UvrBΔ4-DNA complex is depicted in colored ribbon and transparent surface representation, the DNA (black) is shown as stick model. The colors represent the domain architecture of UvrBΔ4 (yellow: domain 1a, green: domain 1b, blue: domain 2, red: domain 3, cyan: β-hairpin).

Additionally, a short stretch of DNA is bound to UvrBΔ4. Although UvrBΔ4 was co-crystallized with a hairpin-bubble substrate containing a fluorescein-adducted thymine as NER substrate, only three complete bases are visible in the crystal structure (**Figure 4.7**). At the 5' end, electron density can be observed to position

one ribose and one phosphate, whereas, at the 3' end, the electron density is only sufficient for one phosphate-group.

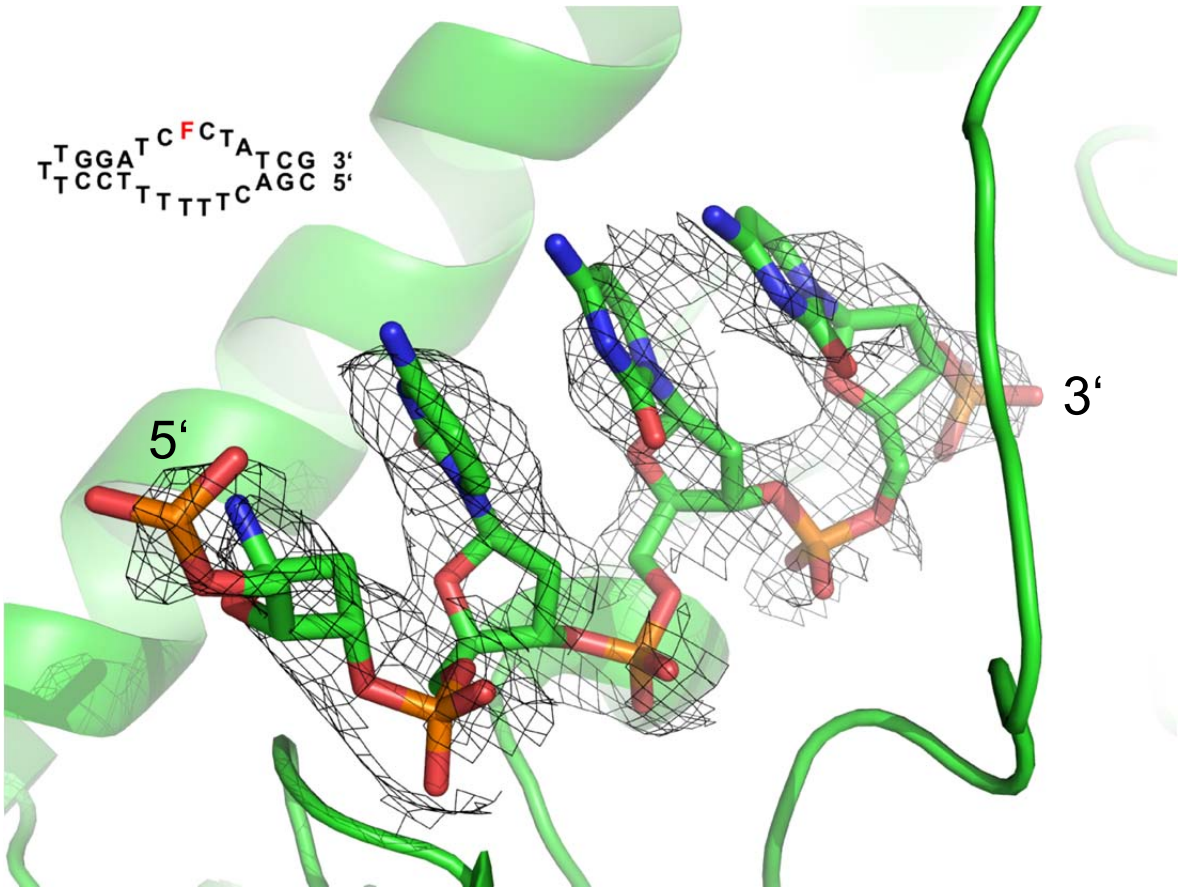

Figure 4.7 Electron density map covering the DNA. UvrBΔ4 is shown in green, the DNA is depicted as stick model, the black mesh represents the $2F_o$-F_c electron density map covering the DNA contoured at 1σ. The sequence and structure of the DNA substrate used for co-crystallization is shown above with F representing a fluorescein-adducted thymine.

The $2F_o$-F_c electron density map contoured at 1σ is not sufficiently resolved to allow the identification of the DNA sequence in the crystal structure, although the electron density does not seem to be suitable in size and shape to harbor purines. Since the DNA substrate used for co-crystallization contains several pyrimidines in a row within the bubble structure, the DNA sequence remains unclear. Moreover, it is an open question whether the thymine containing the fluorescein is part of the DNA observed in the density or the DNA stretch opposite the DNA adduct. The fluorescein seems to be too flexible to be visible in the electron density.

4.1.5. The UvrBC-complex is able to perform specific incision reactions

Although the mechanism of damage verification remains to be deciphered, it is known that UvrB binds to DNA bubble substrates without the help of UvrA (Zou & Van Houten, 1999) and is able to confirm the presence of a NER lesion in DNA even in the absence of UvrA (Moolenaar et al, 2000b). Therefore, it is assumed that UvrB jumps onto the DNA at the site of the bubble and moves to the DNA damage which is recognized and verified. UvrC would then be recruited to specifically incise the DNA.

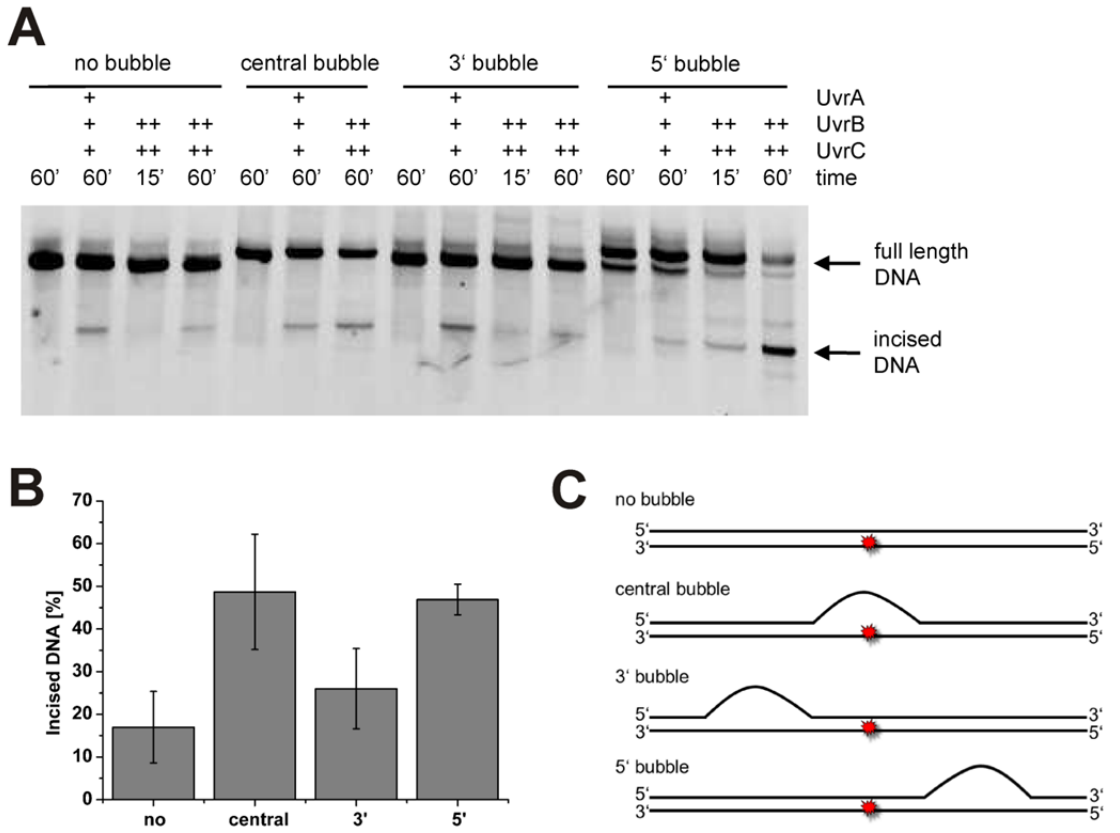

Figure 4.8 Incision assays of different bubble sites with regard to a DNA damage. (A) Incision assays were performed with 20 nM UvrA, 100 nM UvrB, 50 nM UvrC (+) or alternatively 2 µM UvrB and 2 µM UvrC (++). DNA substrates were generated by annealing the DNA substrate 100mer_bottomF containing a central fluorescein with either 100mer_topNB, 100mer_top_CB, 100mer_top5B or 100mer_top_3B to yield dsDNA, DNA with a 6 nts central bubble, DNA with a 6 nt-bubble 3’ to the fluorescein or DNA with a 6 nt-bubble 5’ to the lesion, respectively. Incision assay with different locations for a 6 nt unpaired region leads to an identical incision product consistent with the incision site 8 bp 5’ to the fluorescein adducted thymine both for UvrABC and for UvrBC. (B) Quantification of the incision products reveal that DNA substrates with either central or 5’ bubble are preferred over DNA substrates containing no bubble or a bubble located 3’ to the lesion. (C) Schematic representation of DNA substrates used for this incision assay.

To test this hypothesis, incision assays were performed employing dsDNA as well as DNA substrates containing a 6 nts bubble at different positions with regard to the central DNA damage (**Figure 4.8**, panel C). The reaction was carried out in 1 x ABC-buffer using 20 nM UvrA, 100 nM UvrB, 50 nM UvrC or alternatively 2 µM UvrB and 2 µM UvrC.

Quantification of three independently conducted experiments revealed that the substrates containing a central bubble and a 5' bubble are incised more efficiently compared to a dsDNA or a 3' bubble substrate (**Figure 4.8**). It was also observed that incision by UvrBC is performed even in the absence of a bubble but not as effective as in the presence of either UvrA or a bubble. UvrB therefore is able to bind to dsDNA but with much lower affinity than to bubble substrates. The denaturing gel, however, shows that UvrABC and UvrBC incision leads to the same incision product independent from the position of the bubble (**Figure 4.8**). Thus, we conclude that UvrBC forms a specific complex incising DNA specifically at the same site as UvrABC.

4.1.6. Towards the crystallization of the UvrBC-DNA complex

After UvrB has verified the presence of a DNA damage, UvrC is recruited to perform the dual incision 3' and 5' to the DNA lesion. This crucial step in NER is thought to include major conformational changes to hand over the DNA from UvrB to UvrC to then enable the endonuclease to incise the DNA. Having demonstrated that UvrBC in the absence of UvrA is able to incise DNA specifically, the UvrBC complex on damaged DNA was sought to be crystallized prior to incision. The non-hydrolysable ATP-analog AMPPNP was used to stall the UvrBC complex on the DNA substrate without being able to perform the incision reaction which would lead to the dissociation of the proteins from the DNA product.

A stable UvrBC-DNA complex was obtained by mixing the proteins UvrB, UvrC and the fluorescein-adduct containing DNA substrate A_F24/A_comp24 in a 1:1:1.2 molar ratio together with AMPPNP in 1 x ABC-buffer. The mixture was incubated at RT for 30 min and subsequently purified by size-exclusion chromatography in buffer UvrBC-GF using a Superdex™ 200 10/300GL column (**Figure 4.9**).

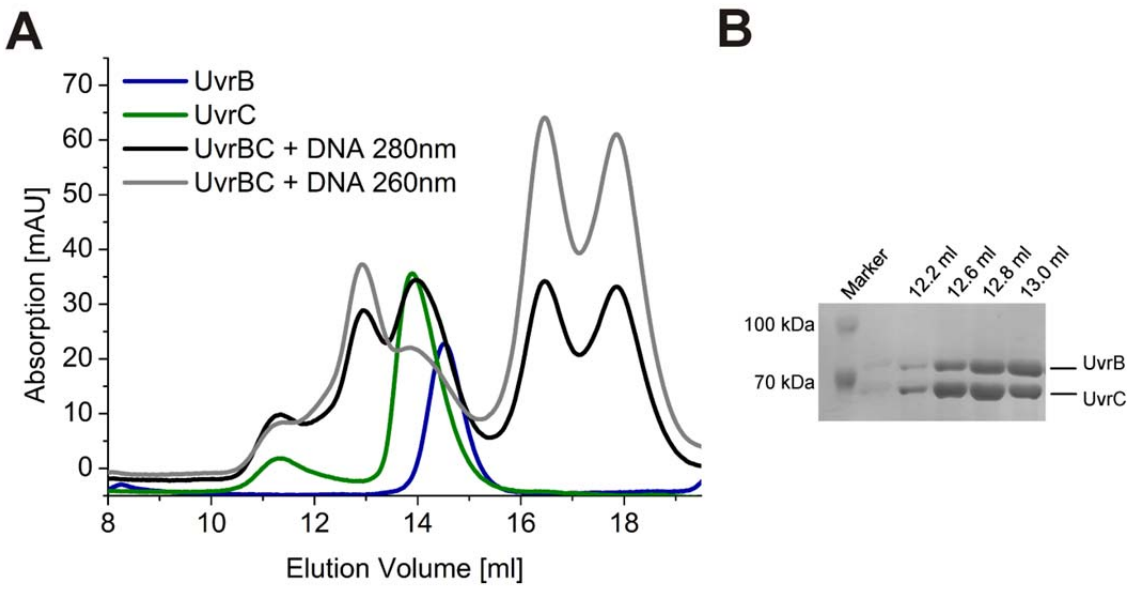

Figure 4.9 Chromatogram and SDS-PAGE of the UvrBC-DNA complex applied to analytical size-exclusion chromatography. (A) For reference, elution profiles of UvrB and UvrC are depicted in blue and green, respectively. The elution profile for the UvrBC-DNA complex is shown with the UV-absorption at 280 nm and 260 nm in black and gray, respectively. The UvrBC-DNA complex elutes at ~12.8 ml on the Superdex™ 200 10/300GL column. The higher elution volume compared to the single proteins indicates an increase of molecular weight and the elevated absorption at 260 nm indicates the presence of DNA. In addition, peaks are observed for the single proteins (~13.9 and 14.5 ml) and for free DNA (16.3 ml and 17.8 ml). (B) The presence of both proteins UvrB and UvrC in this peak is confirmed by SDS-PAGE.

The chromatogram resulting from the UvrBC-DNA complex (black and gray curves in **Figure 4.9**) reveals a peak eluting at ~ 12.8 ml. The increased elution volume compared to the single proteins indicates a complex of higher molecular weight (**Figure 4.9**). Moreover, the higher absorption at 260 nm in the newly formed peak compared to the single proteins suggests that DNA and AMPPNP are present in the complex. In the corresponding SDS-gel, the presence of both proteins, UvrB and UvrC, was verified. The peak fractions were concentrated using centrifuge concentrators (MW cutoff: 50,000 Da) and immediately used for crystallization employing vapor diffusion approaches. Table A1 (Appendix) lists all conditions in which the UvrBC-DNA complex was set-up for crystallization. However, none of these crystallization conditions was suitable to form crystals.

Possible alternatives for crystallizing the UvrBC-DNA complex would be to try different DNA substrates and to exchange AMPPNP to ADP, or to form the complex even in the presence of ATP but with an incision deficient UvrC variant. Due to time limitations, these experiments could not be pursued.

4.2. Functional and structural characterization of the archaeal endonuclease Bax1 and its interaction with the helicase XPB

4.2.1. Cloning of *T. acidophilum* XPB and Bax1

Genes for *T. acidophilum* XPB and Bax1 were amplified by PCR as described. PCR products were analyzed on a 1 %-agarose gel (**Figure 4.10**, panel A and B).

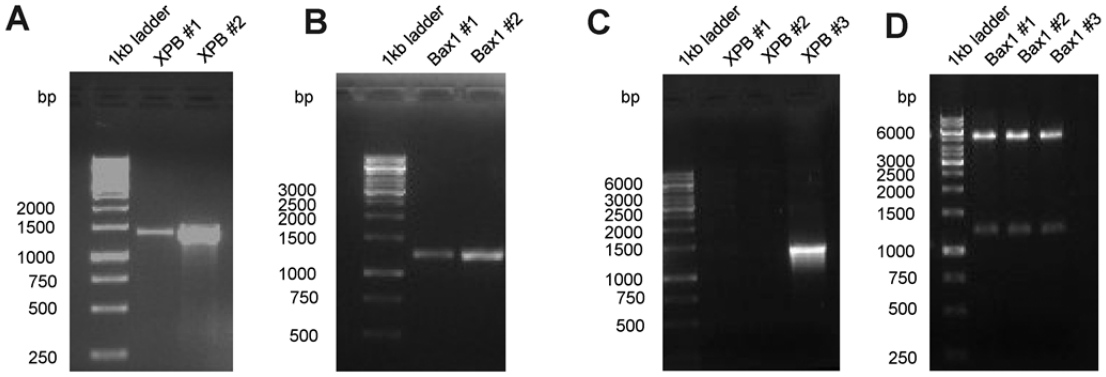

Figure 4.10 Cloning of *T. acidophilum* XPB and Bax1. (A) PCR product for the XPB gene (expected size: 1362 bp) on a 1 % agarose gel. (B) PCR product for the Bax1 gene (expected size: 1224 bp) on a 1 % agarose gel. (C) Colony PCR on a 1 % agarose gel with three different clones. Clone #3 carries the XPB gene within the expression vector pETM-11. (D) Test digestion on a 1 % agarose gel with three different clones. All three tested clones carry the Bax1 gene within the expression vector pETM-11.

The PCR products XPB #2 and Bax1 #2 were purified as described (3.1.1). The destination vectors as well as the PCR products were digested to allow for cloning via sticky ends. XPB was cloned using the restriction enzymes *Pci*I and *Hind*III, Bax1 was cloned employing *Nco*I and *Xho*I. Ligation was performed using a 3-fold molar excess of insert over vector DNA. The ligation reaction was then transformed into DH5α-cells and selected on LB^{Kan}–Agar plates for a circular plasmid. The presence of the gene was confirmed via colony-PCR and digestion of the plasmid. **Figure 4.10**, panel C, shows that XPB #3-clone carries the XPB gene which was amplified by PCR and runs at the expected size of 1362 bp in the agarose gel. **Figure 4.10**, panel D, depicts the digestion pattern of isolated

plasmids on an agarose gel. In addition to the linearized pETM-11 vector (~5400 bp) the Bax1 gene with an expected size of 1224 bp can be observed. Plasmids from clones XPB #3 and Bax1 #1 were isolated and the correct sequence was determined by DNA sequencing (Seqlab Göttingen).

4.2.2. Expression and purification of *T. acidophilum* Bax1

The Bax1 gene cloned into the pETM-11 expression vector (*vide supra*) was transformed into *E. coli* BL21-CodonPlus®(DE3)-RIL cells. A $LB^{Cam/Kan}$ pre-culture was inoculated with a single colony and used for inoculation of a 2 l $ZYM\text{-}5052^{Cam/Kan}$ culture. The cells were grown at 37 °C to an oD600 = 0.3. The temperature was lowered to 15 °C and expression was continued for additional 14 h at 200 rpm. Cells were harvested by centrifugation and lysed in buffer IMAC-A supplemented with DNaseI using a cell disruptor system. The cleared lysate was loaded onto a 1 ml Ni-MAC column. Protein was then eluted using a 2-step imidazole gradient up to 500 mM imidazole (**Figure 4.11**).

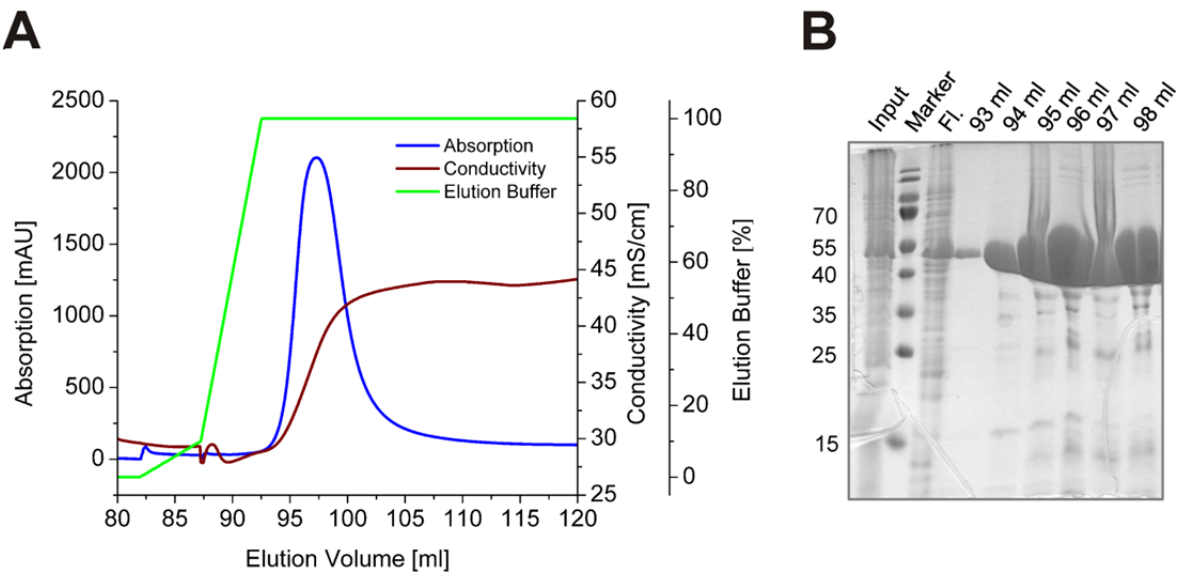

Figure 4.11 Chromatogram and SDS-PAGE of a representative affinity chromatography of Bax1. (A) Bax1 was eluted from 1 ml Ni-MAC column by a 2-step gradient (0-10 %, 10-100 % elution buffer IMAC-B) to a final concentration of 500 mM imidazole. (B) Input, flow through (Fl.) and elution fractions were analyzed by SDS-PAGE for amount and purity of Bax1 (46 kDa).

Fractions containing Bax1 were pooled, concentrated and subsequently applied to a HiLoad™ 26/60 Superdex™200 pg column (**Figure 4.12**). Bax1 elutes in two peaks at ~180 ml and ~210 ml corresponding to dimeric and monomeric Bax1, respectively. Protein samples from both peaks show a prominent band on the SDS-gel at ~46 kDa which refers to Bax1. The different oligomeric states were pooled, concentrated up to 40 mg/ml and stored separately. The protein concentration of Bax1 was determined spectrophotometrically (MW = 46.7 kDa, ε_{280} = 49,975 M^{-1} cm^{-1}).

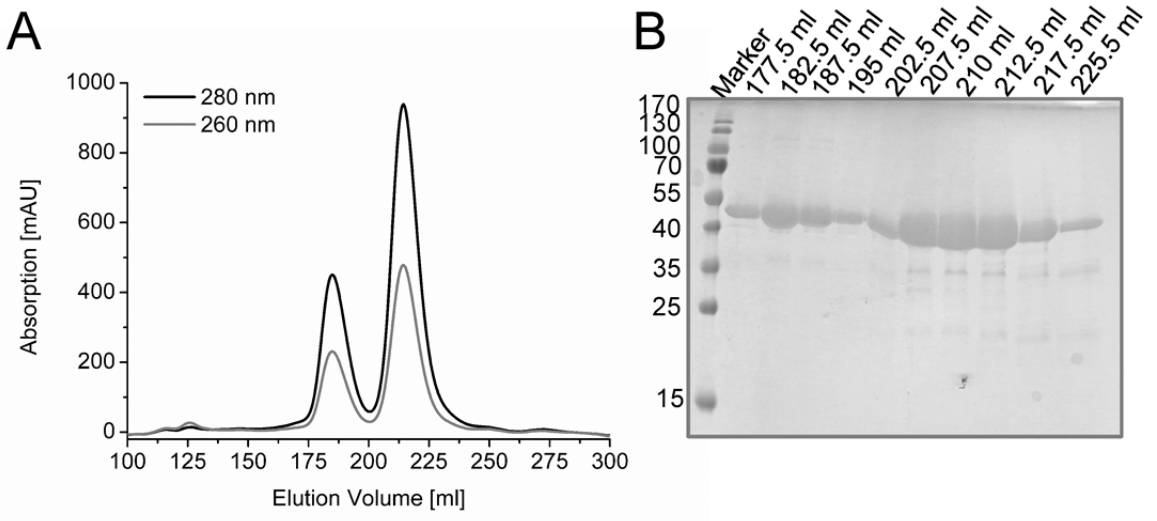

Figure 4.12 Chromatogram of the final size-exclusion chromatography and SDS-PAGE of peak fractions. (A) The elution profile of Bax1 on a HiLoad™ 26/60 Superdex™200 pg column shows two peaks at ~180 ml and ~218 ml corresponding to dimeric and monomeric Bax1, respectively. (B) Peak fractions were analyzed by SDS-PAGE with Bax1 running at the expected size of ~46 kDa.

All Bax1 variants were purified by an identical approach.

4.2.3. Expression and purification of *T. acidophilum* XPB

XPB was expressed and purified using the same protocol as for Bax1, namely an affinity chromatography employing the N-terminal His_6-tag (**Figure 4.13**) and a final size-exclusion chromatography using a HiLoad™ 16/60 Superdex™200 pg column (**Figure 4.14**).

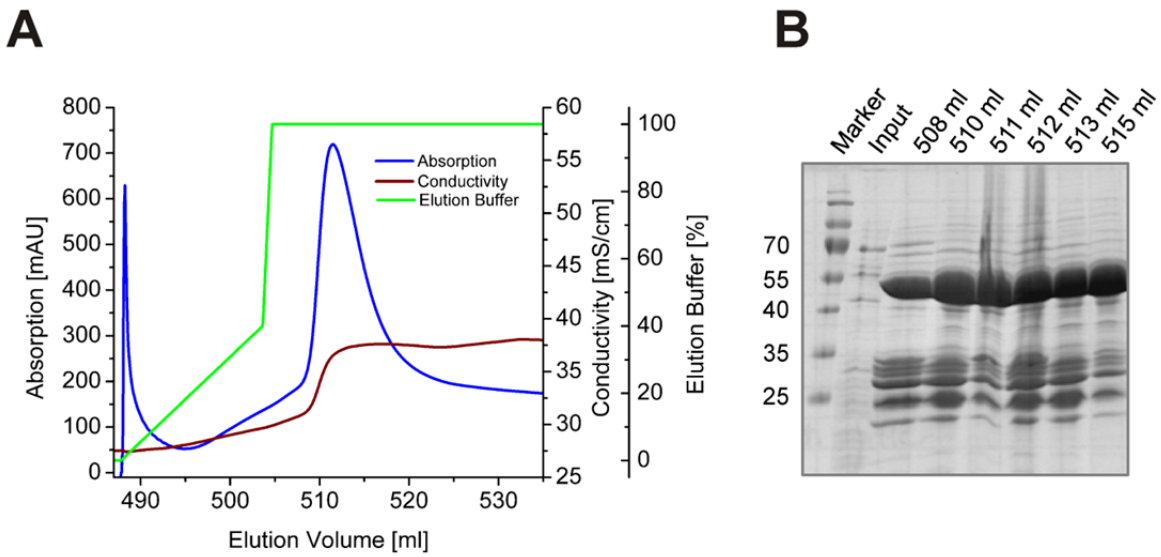

Figure 4.13 Chromatogram and SDS-PAGE of a representative affinity chromatography of XPB. (A) XPB was eluted from the 1 ml Ni-MAC column by a 2-step gradient (0-40 %, 40-100 % elution buffer IMAC-B) up to a final concentration of 500 mM imidazole. (B) Input and elution fractions were analyzed by SDS-PAGE for the presence and purity of XPB (52 kDa).

XPB was eluted in one defined peak by a 2-step imidazole gradient (**Figure 4.13**). The corresponding gel shows XPB running at approximately 55 kDa and some impurities of lower size. XPB could be highly enriched on the Ni-MAC column as indicated by the poor over-expression (**Figure 4.13**, panel B, Input) compared to the elution fractions (**Figure 4.13**, panel B).

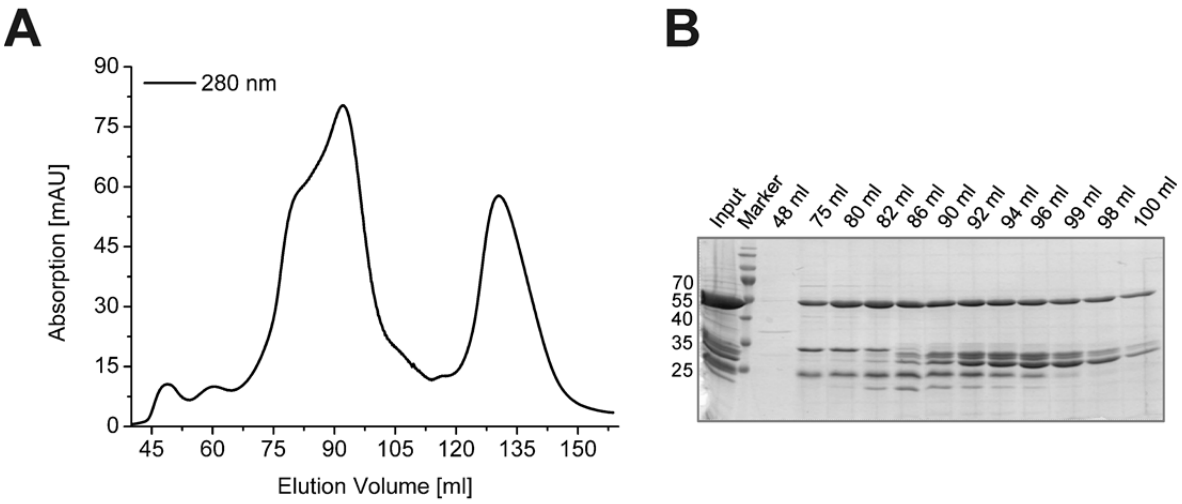

Figure 4.14 Chromatogram of the final size-exclusion chromatography and SDS-PAGE of peak fractions (A) The elution profile of XPB on a HiLoad™ 16/60 Superdex™200 pg column shows one major peak at an elution volume of ~90 ml with a shoulder containing XPB. (B) Peak fractions were analyzed by SDS-PAGE.

XPB elutes in one major peak at an elution volume of ~90 ml with a shoulder to a higher elution volume. Both the peak and the shoulder contain XPB with impurities of lower molecular weight as observed from the SDS-gel. The contaminating bands could not be removed by additional purification approaches such as ion exchange chromatography. The protein concentration of XPB was determined spectrophotometrically (MW = 51.6 kDa, ε_{280} = 58,790 M^{-1} cm^{-1}).

4.2.4. Expression and purification of the *T. acidophilum* XPB/Bax1 complex

The XPB/Bax1 complex was formed by mixing the cell pellets containing XPB and Bax1 prior to cell lysis. Cells were lysed and subjected to affinity chromatography as described for the single proteins and eluted by a gradient up to 500 mM imidazole (**Figure 4.15**).

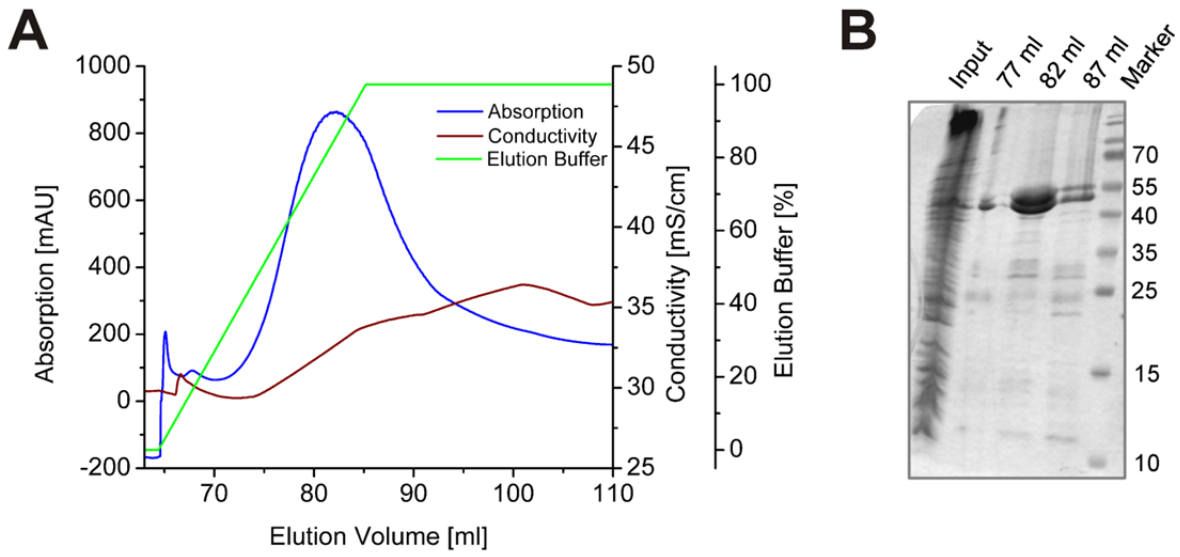

Figure 4.15 Chromatogram and SDS-PAGE of a representative affinity chromatography of XPB/Bax1. (A) XPB/Bax1 was eluted by a gradient (0-100 % elution buffer IMAC-B) up to 500 mM imidazole in one single peak. (B) Input and peak fractions were analyzed by SDS-PAGE for the presence and purity of both proteins XPB (52 kDa) and Bax1 (46 kDa).

The protein solution was diluted with buffer IEX-A buffer to adjust the concentration of NaCl to 50 mM. The diluted protein mixture was then loaded onto a cation-exchange chromatography HiTrap SP HP column and eluted by a gradient up to 1 M NaCl (**Figure 4.16**).

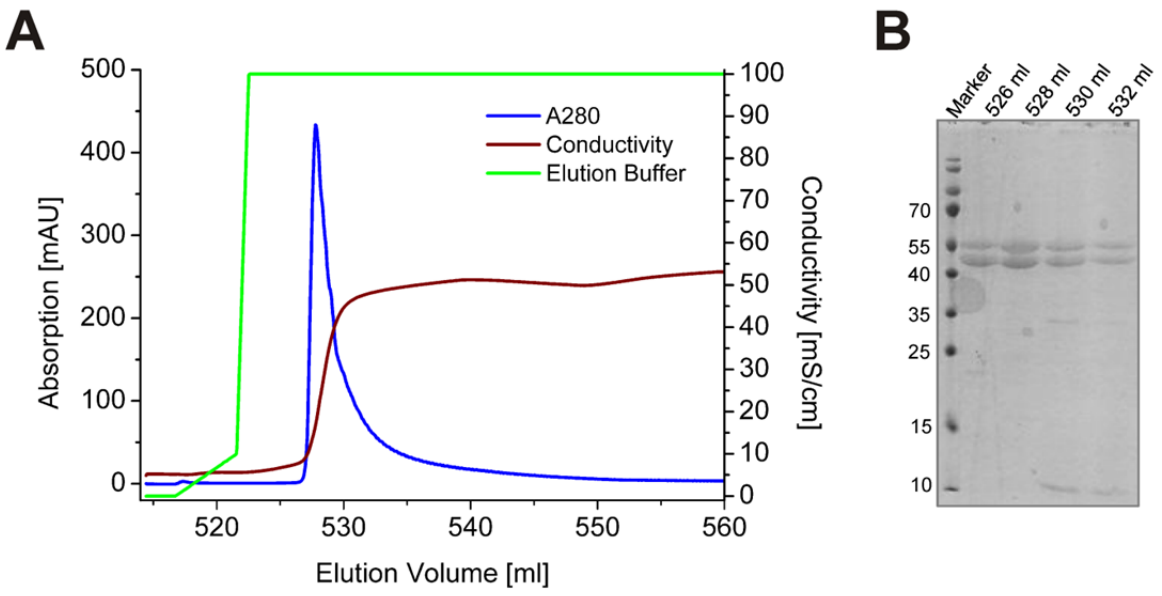

Figure 4.16 Chromatogram and SDS-PAGE of a representative cation-exchange chromatography of XPB/Bax1. (A) XPB/Bax1 was eluted from the 1 ml HiTrap SP HP column by a 2-step gradient (0-10 %, 10-100 % elution buffer IEX-B) up to 1 M NaCl in one single peak. (B) Peak fractions were analyzed for the presence and purity of the XPB/Bax1 complex by SDS-PAGE.

XPB and Bax1 were eluted together in one defined peak employing a steep NaCl gradient. Peak fractions were concentrated and subsequently subjected to a size-exclusion chromatography using a HiLoad™ 16/60 Superdex™200 pg column (**Figure 4.17**). Peak fractions were analyzed by SDS-PAGE regarding the protein content and purity.

XPB/Bax1 elutes from the size-exclusion chromatography in a peak at ~60 ml with a shoulder to higher molecular weight and in a second peak at ~73 ml (**Figure 4.17**). The SDS-gel shows the presence of both proteins XPB and Bax1 in the peaks eluting at 60 ml and 73 ml, whereas Bax1 by itself is present in the peak eluting at 84 ml. The putative XPB/Bax1 complexes as well as Bax1 are highly pure after purification employing the described procedure. Peak fractions (XPB/Bax1 (1) 55-65 ml; XPB/Bax1 (2) 70-75 ml) were pooled, concentrated up to 4 mg/ml, flash-frozen with liquid nitrogen and stored at -80 °C. The concentration of the protein XPB/Bax1 was determined spectrophotometrically (MW = 98.3 kDa, ε_{280} = 108,765 M^{-1} cm^{-1}).

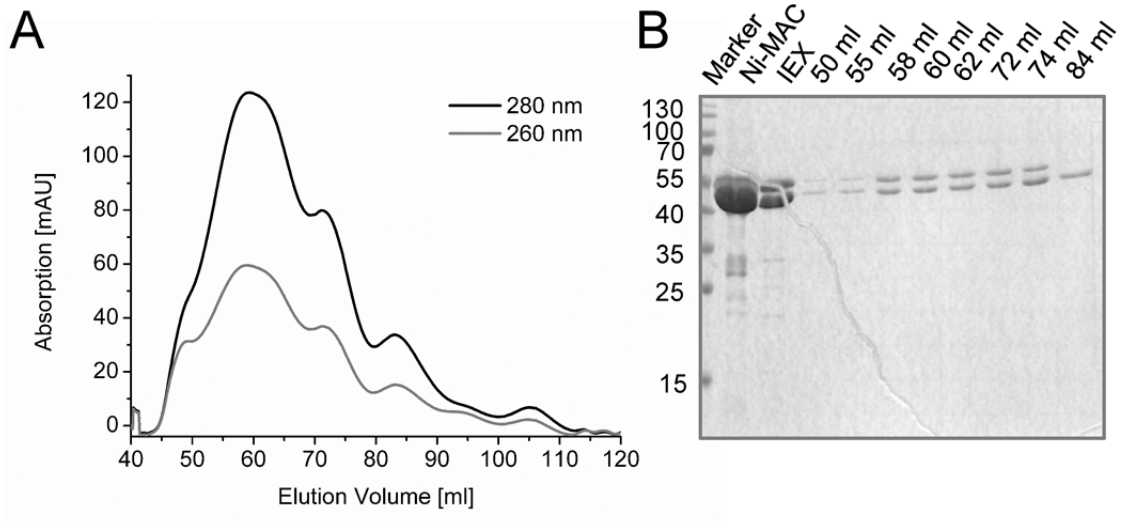

Figure 4.17 Chromatogram of the final size-exclusion chromatography and SDS-PAGE of peak fractions. (A) Elution profile of the XPB/Bax1 complex using a HiLoad™ 16/60 Superdex™200 pg column shows several peaks. (B) SDS-PAGE analysis revealed that XPB/Bax1 is present in both peaks eluting at ~60 ml with a shoulder to higher molecular weight and at ~73 ml. Free Bax1 elutes at ~84 ml.

4.2.5. Towards the crystallization of Bax1

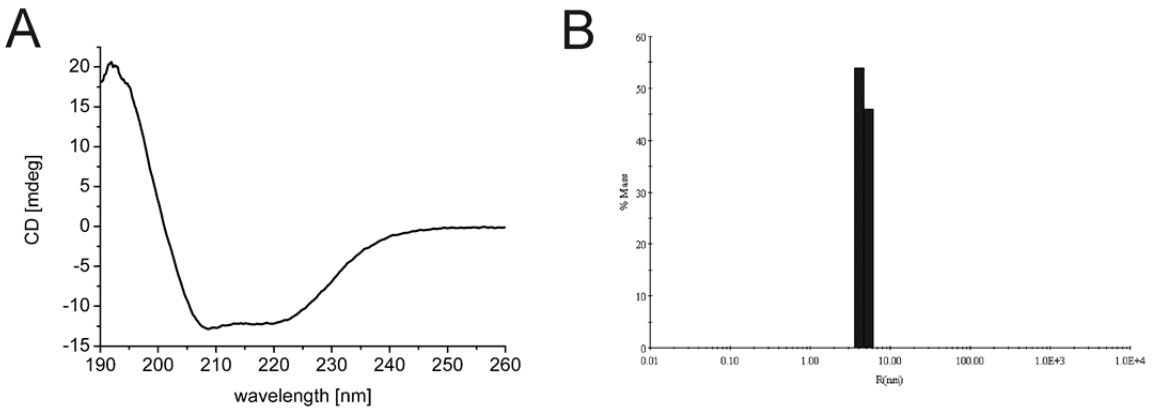

Figure 4.18 CD-spectroscopy and dynamic light scattering (DLS) experiments with Bax1 (A) The CD-spectrum of Bax1 indicates a mixed α/β-structure. (B) DLS measurements reveal that Bax1 was mono-disperse in solution at a concentration of 28 mg/ml. The apparent hydrodynamic radius R_h(app) for monomeric Bax1 was 4.8 nm.

For the crystallization of proteins, highly pure protein samples are essential. Additional criteria are the degree of aggregation, polydispersity and folding of the protein. These parameters were determined using dynamic light scattering and CD-spectroscopy (**Figure 4.18**). Both analyses were performed at 20 °C as described above (3.4.2 and 3.4.3).

As predicted by the secondary structure prediction server, CD-spectroscopy revealed that Bax1 was properly folded in solution adopting a mixed α/β-structure (**Figure 4.18**, panel A) indicated by the two characteristic minima at 208 and 222 nm. The transition from negative to positive ellipticity values at a wavelength of approximately 200 nm is consistent with the absence of disordered regions. Since mis-folded proteins are prone to aggregation and precipitation, the poly-dispersity of 13.4 % observed in DLS measurements (**Figure 4.18**, panel B) indicates that Bax1 does not tend to aggregate in solution, and thus supports the finding derived from CD-spectroscopy that Bax1 is properly folded in solution.

In addition, a renewed equilibrium between both oligomeric states could not be detected after separation of monomeric and dimeric Bax1 in size-exclusion chromatography. Proteins should meet these requirements to enhance the probability to form protein crystals.

Bax1 was employed in diverse crystallization set-ups. Table A2 in the appendix lists the conditions which were used to screen for crystal growth of Bax1 and XPB/Bax1. However, the crystallization set-ups did not yield diffracting crystals under these conditions.

4.2.6. Bax1 forms a stable complex with XPB

Bax1 has been recently identified as a novel binding partner of the helicase XPB in *S. solfataricus* (Richards et al, 2008). Several archaeal genomes, among them *T. acidophilum*, comprise the genes coding for XPB and Bax1 in close proximity. To investigate the complex formation of the two *T. acidophilum* proteins, Bax1 was subjected to both analytical ultracentrifugation and analytical size exclusion chromatography individually and also in the presence of XPB (**Figure 4.19**). The resulting peak fractions from analytical size exclusion chromatography were analyzed by SDS-PAGE to confirm their protein content. The peak including both proteins was clearly shifted compared to the peak containing Bax1 only, indicating stable complex formation in solution. According to the calibration curve obtained for the Superdex™ 200 10/300GL column (3.6.1), the elution volume of the complex peak translated into a molecular weight of

approximately 105 kDa, suggesting a 1:1 stoichiometry of XPB and Bax1. The sedimentation curves derived from analytical ultracentrifugation support this finding. Bax1 displayed a monomer-dimer equilibrium with sedimentation coefficients of approximately 3 S for the monomer and 5 S for the Bax1 dimer. The purified XPB/Bax1 complex showed a sedimentation coefficient of 5 S. Optimized fits for the data were obtained with frictional ratios of 1.65 and 1.35 for Bax1 and the XPB/Bax1 complex, respectively. The different frictional ratios indicate different geometries for the particles in the two samples, while the similar sedimentation coefficients confirm similar sizes for the Bax1 dimer and the XPB/Bax1 complex, consistent with a heterodimeric XPB/Bax1 complex and hence a 1:1 stoichiometry.

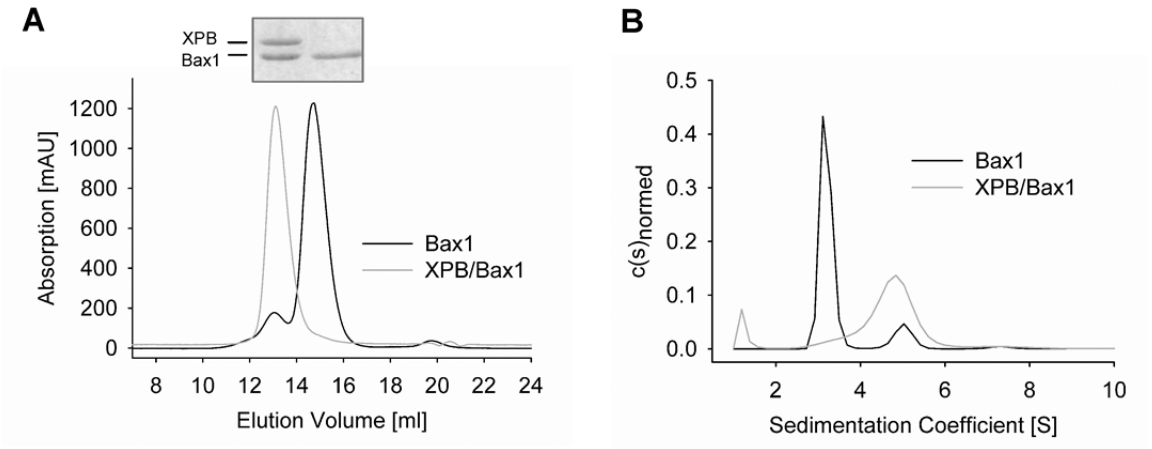

Figure 4.19 ***T. acidophilum*** **Bax1 forms a complex with XPB.** (A) Size exclusion chromatography using a Superdex™ 200 10/300GL column was performed to verify the presence of a stable XPB/Bax1 complex. Bax1 was first analyzed individually (shown in black), followed by a 1:1 stoichiometric mixture of Bax1 and XPB (shown in gray). Peak fractions were analyzed by SDS-PAGE to confirm complex formation (see inset). (B) Both Bax1 and the purified XPB/Bax1 complex were subjected to analytical ultracentrifugation in two separate experiments to support the finding that XPB and Bax1 interact in solution (same color coding as in A).

4.2.7. Bax1 is a structure-specific endonuclease

Despite the finding that Bax1 binds to XPB, there is no known function annotated to Bax1 until now. Database research, however, revealed that Bax1 encloses a DUF790 domain which includes a restriction endonuclease-like fold (Kinch et al, 2005). Therefore, incision assays were conducted and a specific incision product was formed in a concentration dependent manner (**Figure 4.20**,

panel A). It could be proven that Bax1 acts preferentially as a monomer (**Figure 4.20**, panel B)

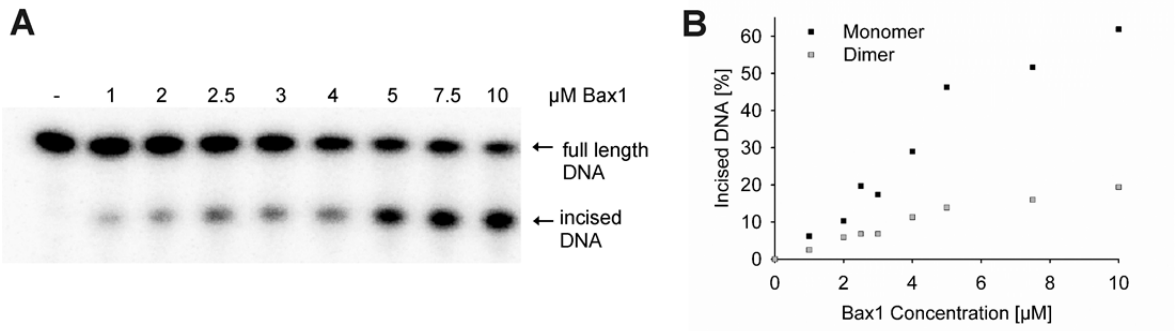

Figure 4.20 ***T. acidophilum*** **Bax1 is an active endonuclease.** (A) The DNA substrate NDT/NDB22 was labeled at the 5' end of the upper strand and used at a concentration of 10 nM. Arrows indicate full-length DNA as well as the incision product. Increasing concentrations of Bax1 led to increasing amounts of the incision product. (B) Quantification of incision assays performed with both monomeric and dimeric Bax1 (DNA substrate NDT/NDB22, 10 nM) revealed that the Bax1 monomer is more active than the dimer.

To verify that the observed incision activity is not due to contamination but is initiated by Bax1, the protein was purified as described above including size-exclusion chromatography as a final purification step (**Figure 4.21**, panel A).

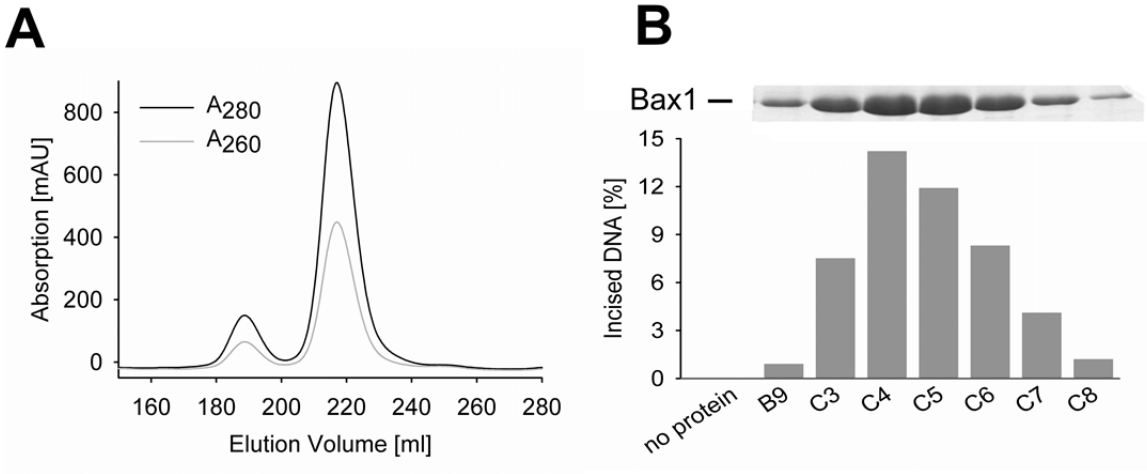

Figure 4.21 **The endonuclease activity is specific to Bax1.** (A) Bax1 elution profile after size exclusion chromatography. Absorptions at 260 nm and 280 nm are shown in gray and black, respectively. The SDS-gel indicates the Bax1 content of the different fractions. Both peaks contained Bax1 and revealed a monomer-dimer equilibrium according to a calibration of the column. (B) Samples from 4 ml fractions of the dimer peak (D9, corresponding to an elution volume of 185 ml) and of the monomer peak (C3-C8, corresponding to elution volumes from 210 ml to 230 ml) were subjected to both SDS-PAGE and incision assays. The incised product was quantified and visualized in a bar chart confirming a direct correlation of the amount of Bax1 and the formation of the DNA fragment.

The peak fractions were analyzed by SDS-PAGE with respect to their protein content and were then subjected to incision assays (**Figure 4.21**, panel B). Comparing the resulting incision product to the employed amount of Bax1, a strong correlation can be observed confirming that the endonuclease activity is specific to Bax1.

Incision assays employing different DNA substrates with respect to the length of the 3' overhang indicated that Bax1 cuts 4 to 6 nts to the ssDNA/dsDNA junction in the scissile 3' overhang (**Figure 4.22**, panel A). Moreover, Bax1 only incised DNA substrates comprising a free 3' overhang, whereas it did not cut dsDNA, DNA substrates forming a bubble with a maximum of 16 unpaired bases or DNA substrates including a 5' overhang (**Figure 4.22**, panel B). This finding indicates that Bax1 is a structure-specific endonuclease.

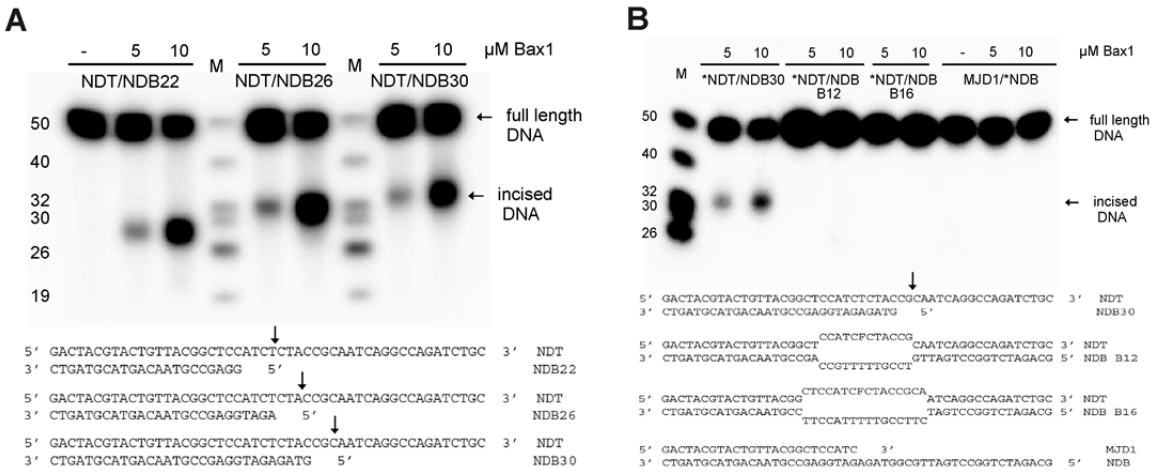

Figure 4.22 ***T. acidophilum*** **Bax1 is a structure-specific endonuclease.** (A) Bax1 incised 4 to 6 nts away from a dsDNA/ssDNA junction. Incision assays employing diverse DNA substrates, which differed with respect to the length of the 3' overhang, indicated that Bax1 cuts specifically 4 to 6 nts to the dsDNA/ssDNA junction in the ssDNA region. The DNA ladder M was generated by mixing each 10 nM of 5' labeled ssDNAs of known length, in this case 50 nts, 40 nts, 32 nts, 30 nts, 26 nts and 19 nts. The different substrates used for the incision assays are presented below. Arrows depict where Bax1 cuts the DNA substrates. (B) Incision assays were performed as described using 10 nM of different DNA substrates (depicted below). Bax1 incised DNA containing a 3' overhang, whereas no incision product could be observed employing DNA comprising a 5' overhang or unpaired bases forming a bubble. The asterisks mark the labeling site within the DNA substrate with ^{32}P.

4.2.8. Bax1 mutants show impaired nuclease activity revealing essential residues for DNA incision

Bax1 mutants were generated to further scrutinize Bax1's endonuclease activity. Since there is no structural information available for Bax1, multiple sequence alignments of homologous proteins from diverse archaeal organisms were performed and revealed a patch of highly conserved residues (**Figure 4.23**, panel A). Moreover, similarities with the restriction endonuclease *Tsp*RI were identified in a BLAST search (**Figure 4.23**, panel B). Although the overall sequence identity of Bax1 and *Tsp*RI is only 14 %, a small patch of 14 amino acids shares 60 % identical residues. Interestingly, the two patches identified by multiple sequence alignment and BLAST search overlap, indicating a crucial role for the conserved sites with regard to DNA binding or catalysis. To confirm the importance of individual residues, site directed mutagenesis was performed. After successful expression and purification, the mutants were first analyzed by CD-spectroscopy. The CD-spectra were almost identical for Bax1 WT and all Bax1 mutant proteins, indicating that the point mutations do not affect the overall structure of the protein (**Figure 4.23**, panel D).

Quantification of the incision assays revealed three different groups of mutants (**Figure 4.23**, panel C): The first group was only slightly impaired in the incision assays as it exerted more than 70 % of WT Bax1 activity, i.e. F116A, E125A and N153A. The second group showed impaired incision activity in a range of 25 to 35 % of WT activity, i.e. E124A/E125A, Y128A, D130A, E134A and L137E. Mutants of the third group displayed a dramatically reduced activity of less than 20 % of the WT activity, i.e. E124A, D132A, Y152A and N153A/E155A. These different levels of incision activity strongly indicate varying roles for the mutated amino acids either in DNA binding or catalysis.

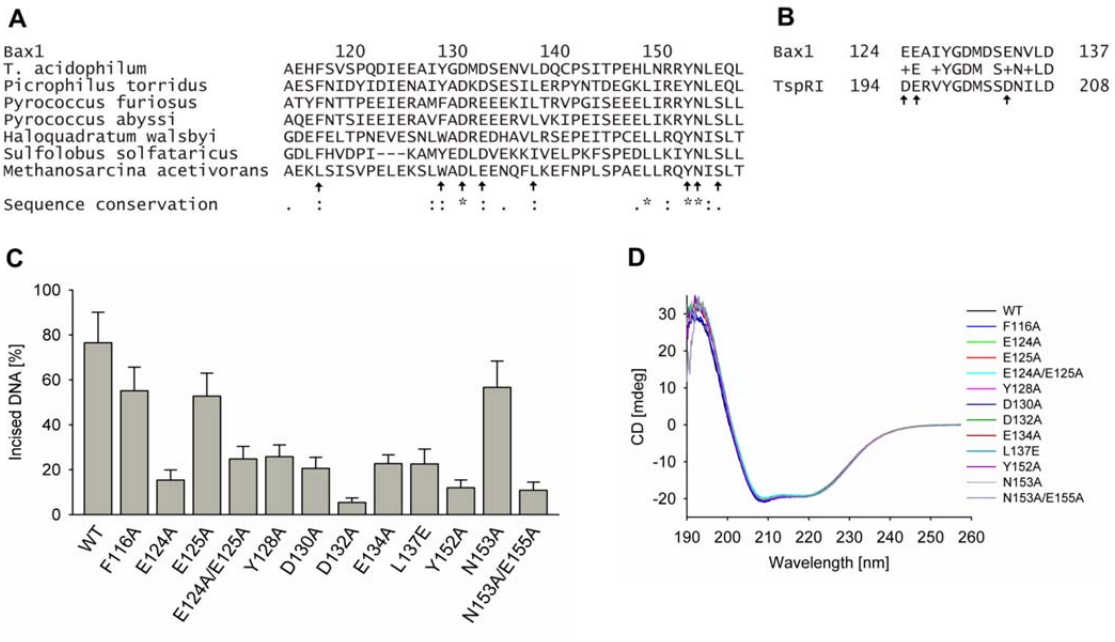

Figure 4.23 Identification of active site residues (A) Conservation of Bax1 among archaea. Multiple sequence alignment of Bax1 (*T. acidophilum*) and related proteins from *Picrophilus torridus*, *Pyrococcus furiosus*, *Pyrococcus abyssi*, *Haloquadratum walsbyi*, *S. solfataricus* and *Methanosarcina acetivorans* depict the level of sequence conservation among different archaea. The arrows point to sites where point mutations were introduced, dots correspond to similar residues, colons to conserved residues and asterisks to identical residues, respectively. Numbering above the alignment corresponds to the *T. acidophilum* sequence. (B) Conserved residues of Bax1 and *Tsp*RI endonucleases. The alignment of Bax1 and *Tsp*RI endonuclease reveals a high local sequence similarity indicating conserved and thus probably important residues for Bax1 endonuclease activity. (C) Incision activity of Bax1 mutants. Quantification of three independently conducted incision assays with DNA substrate NDT/NDB26 (2 nM) revealed different cutting efficiencies for diverse Bax1 variants (8 µM each). Error bars represent the standard deviation of three independent measurements. (D) Both the CD-spectrum depicting the WT protein and the mutant spectra show the same pattern thus verifying that all variants adopt the same mixed α/β-fold as observed for the WT protein.

Table 4.2 K_Ds for Bax1 wildtype and Bax1 variants

Bax1 variant	Dissociation constant K_D [µM]
wildtype	1.34 ± 0.57
E124A	0.92 ± 0.08
D132A	0.50 ± 0.37
Y152A	0.54 ± 0.40
N153A/E155A	0.93 ± 0.34

Employing bio-layer interferometry, the affinities of incision deficient Bax1 variants towards their DNA substrate were determined (Table 4.2).

Measurements were performed using three different protein concentrations. Data were considered for the table below only if the fitting accuracy of observed and calculated data was $R^2 \geq 0.95$. Table 4.2 shows that DNA binding of Bax1 variants was not reduced compared to wildtype Bax1. Contrarily, they seemed to bind to DNA even with slightly higher affinity.

In order to investigate the stability and to identify defined domains, Bax1 was subjected to limited proteolysis experiments using the unspecific protease trypsin. Trypsin cleaves peptide bonds preferentially after the basic residues arginine and lysine at locally disordered parts of proteins such as loops and flexible linker regions. These parts of a protein are more accessible and thus more susceptible for proteolytic digest.

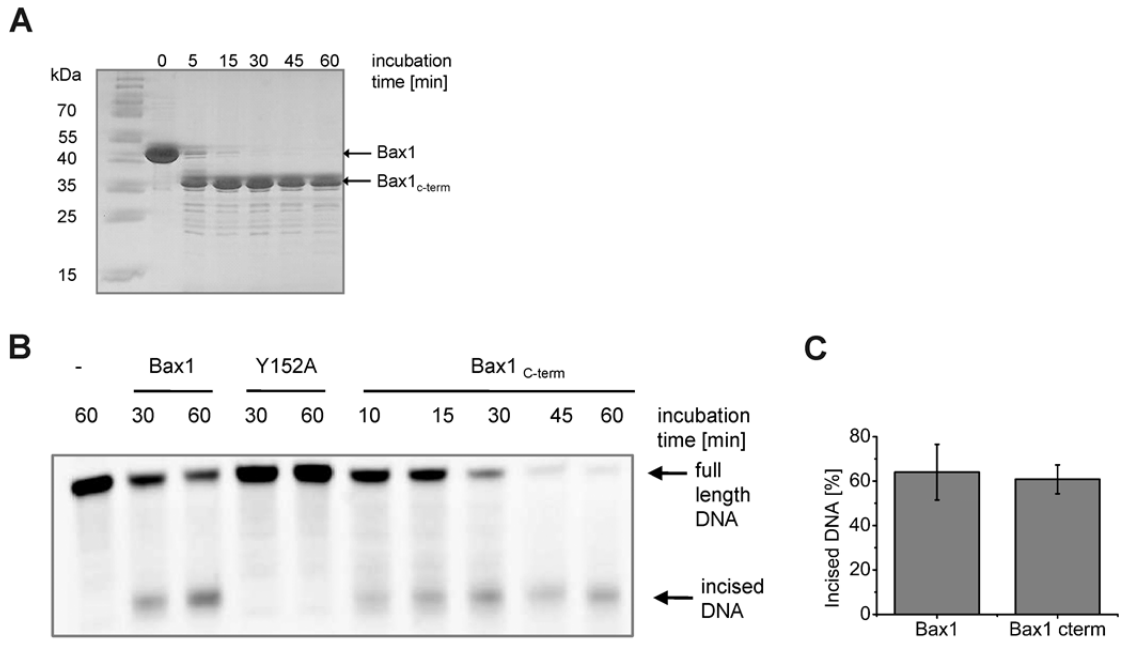

Figure 4.24 The C-terminal domain of Bax1 harbors endonuclease activity. (A) Limited proteolysis of Bax1 using trypsin protease uncovered a stable fragment at an apparent size of 35 kDa. Mass spectrometry analysis verified that this fragment is consistent with the C-terminal part of Bax1 comprising amino acids 90 to 403. (B) Both $Bax1_{c\text{-}term}$ and Bax1 were subjected to an incision assay. (C) Incision products from 3 independently performed experiments were quantified (5 µM Bax1, 20 nM DNA substrate NDT/NDB22).

Limited proteolysis of Bax1 revealed the formation of a stable fragment with a molecular weight of about 35 kDa (**Figure 4.24**, panel A). Mass spectrometry

analysis confirmed that the stable fragment produced by tryptic digest is consistent with the C-terminal part of Bax1 comprising amino acids 90 to 403 (referred to $Bax1_{c\text{-}term}$). Consistent with the location of the active site residues, incision assays verified that $Bax1_{c\text{-}term}$ comprising residues 90-403 is still active as an endonuclease yielding an incision product of the same size as full-length Bax1 (**Figure 4.24**, panel B and C).

So far, this work was focussing on the endonuclease activity of Bax1. Now, the activity of Bax1 is investigated mainly with regard to the interaction with its binding partner XPB.

4.2.9. XPB and Bax1 form two different complexes with different activity

Co-purification of XPB and Bax1 using a HiLoad™ 16/60 Superdex™ 200 column results in an elution profile displaying three major peaks (**Figure 4.17**). The peak eluting at ~85 ml contained free Bax1, as observed in the SDS-PAGE gel shown in the inset. The fractions eluting at ~60 ml and at ~75 ml contained both proteins XPB and Bax1. The fractions from the two differently eluting peaks containing XPB and Bax1 were re-applied onto an analytical size-exclusion chromatography Superdex™ 75 10/300 GL column. The analytical size-exclusion profiles revealed that the two different isolated complex forms did not convert into each other. The fractions from peak 1 eluted at 8.5 ml (XPB/Bax1 (1), gray curve) and those from peak 2 eluted at 9.5 ml (XPB/Bax1 (2), black curve) (**Figure 4.25**). Both peaks contained XPB as well as Bax1 (SDS-gel inset in **Figure 4.25**). This indicated that XPB and Bax1 can form two different, distinct types of complexes which did not seem to convert into each other.

Surprisingly, incision assays revealed that these two XPB/Bax1 complexes exert different DNA incision activities: the XPB/Bax1 (1) complex eluting at 8.5 ml is able to incise 5' and 3' ssDNA overhangs, whereas the XPB/Bax1 (2) complex eluting at 9.5 ml does not cut either of the DNA substrates under these conditions (**Figure 4.25**, panel B). From here on, we will refer to the XPB/Bax1 (1) and the XPB/Bax1 (2) complex as the incision-competent ($XPB/Bax1_{comp}$) and the incision-incompetent ($XPB/Bax1_{incomp}$) complex, respectively.

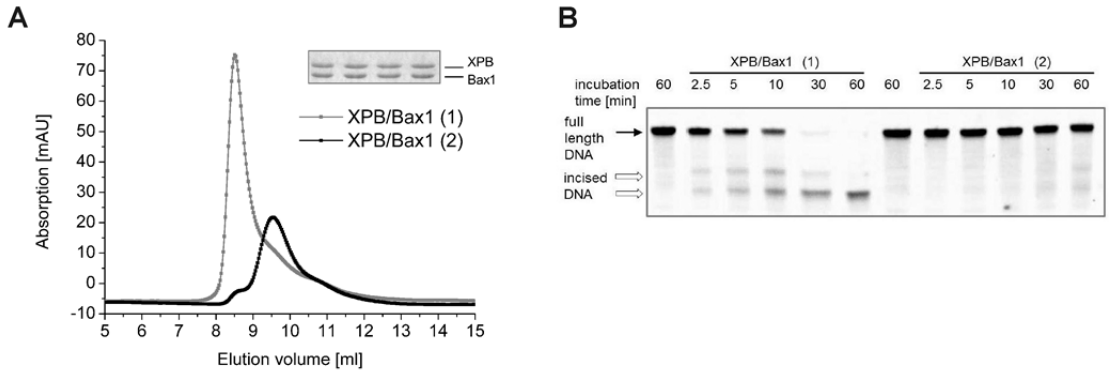

Figure 4.25 XPB and Bax1 form two different complexes with different activity. (A) The presence of two different XPB/Bax1 complexes was detected by size-exclusion chromatography using a Superdex™ 75 10/300GL column. (B) The XPB/Bax1 complex 1 (shown in gray in panel A) was able to cut DNA in contrast to the XPB/Bax1 complex 2 (shown in black in panel A), referring to as XPB/Bax1 incision competent (XPB/Bax1$_{comp}$) and XPB/Bax1 incision incompetent (XPB/Bax1$_{incomp}$), respectively. Both protein complexes were applied at a final concentration of 5 µM, the concentration of the DNA substrate NDT/NDB22 was 20 nM.

4.2.10. XPB/Bax1$_{comp}$ and Bax1 exert different incision activities

To study the effect of XPB binding on Bax1 activity, we directly compared DNA incision by Bax1 and by the XPB/Bax1$_{comp}$ complex, both from *T. acidophilum*. Incision assays showed that Bax1 cuts 3' overhang substrates leading to one distinct incision product, while XPB/Bax1$_{comp}$ produces a pattern of different, multiple incision products (**Figure 4.26**, panel A). In the Figure, the different positions and intensities of incision for Bax1 and for the XPB/Bax1$_{comp}$ complex are indicated by white and gray arrows on the schematic representations of the employed DNA substrates, respectively. The different incision behavior of Bax1 and XPB/Bax1$_{comp}$ was even more pronounced for 5' overhang DNA substrates: Bax1 by itself did not incise any of the 5' overhang substrates, whereas the XPB/Bax1$_{comp}$ complex produced clearly visible incision products with these substrates (**Figure 4.26**, panel B). This suggests that XPB is required to load and position Bax1 properly onto the DNA enabling Bax1 to increase its substrate range and to achieve exertion of the correct DNA incisions by controlled positioning and possibly orientation of Bax1 on the DNA. Furthermore, the presence of the XPB/Bax1$_{incomp}$ complex in addition to the XPB/Bax1$_{comp}$ complex suggests a regulation mechanism for Bax1.

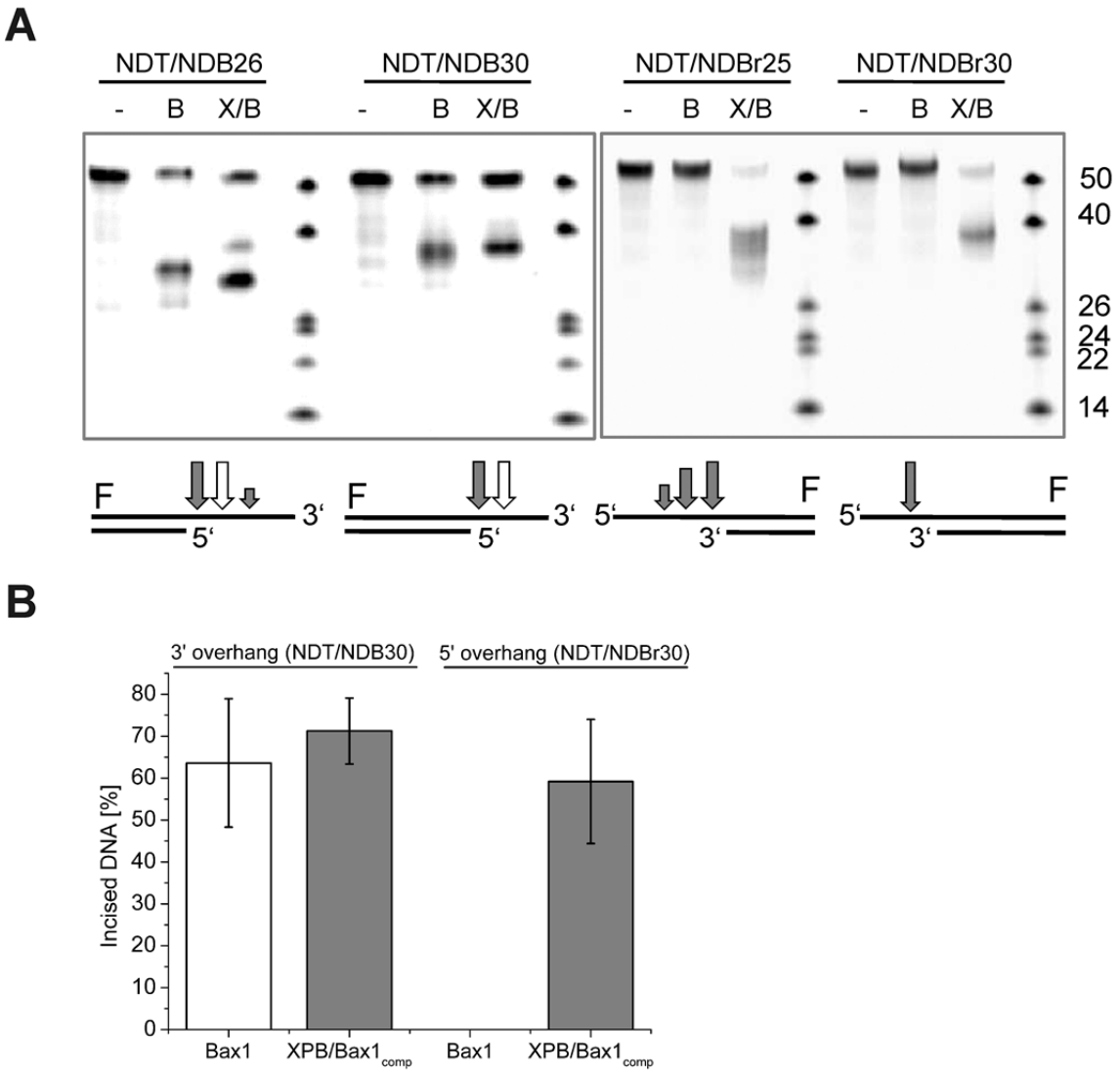

Figure 4.26 XPB/Bax1$_{comp}$ and Bax1 exert different incision activities. (A) Incision assays conducted with XPB/Bax1$_{comp}$ and Bax1 show a different pattern of incision products for 3’ overhang substrates (schematically represented by gray and white arrows, respectively). In contrast, 5’ overhang substrates were only cut by XPB/Bax1$_{comp}$. A DNA ladder (50, 40, 26, 22, 20, 14 nucleotides) was generated to serve as a reference. (B) Incision products from 3 independently performed experiments were quantified (5 µM Bax1 or XPB/Bax1$_{comp}$, 20 nM DNA substrate).

The presence of different incision sites within the ssDNA region of an overhang substrate suggests that XPB/Bax1$_{comp}$ incised ssDNA non-specifically. Therefore, differently structured DNA substrates carrying the fluorophore at either the 5’ or the 3’ end were tested (**Figure 4.27**).

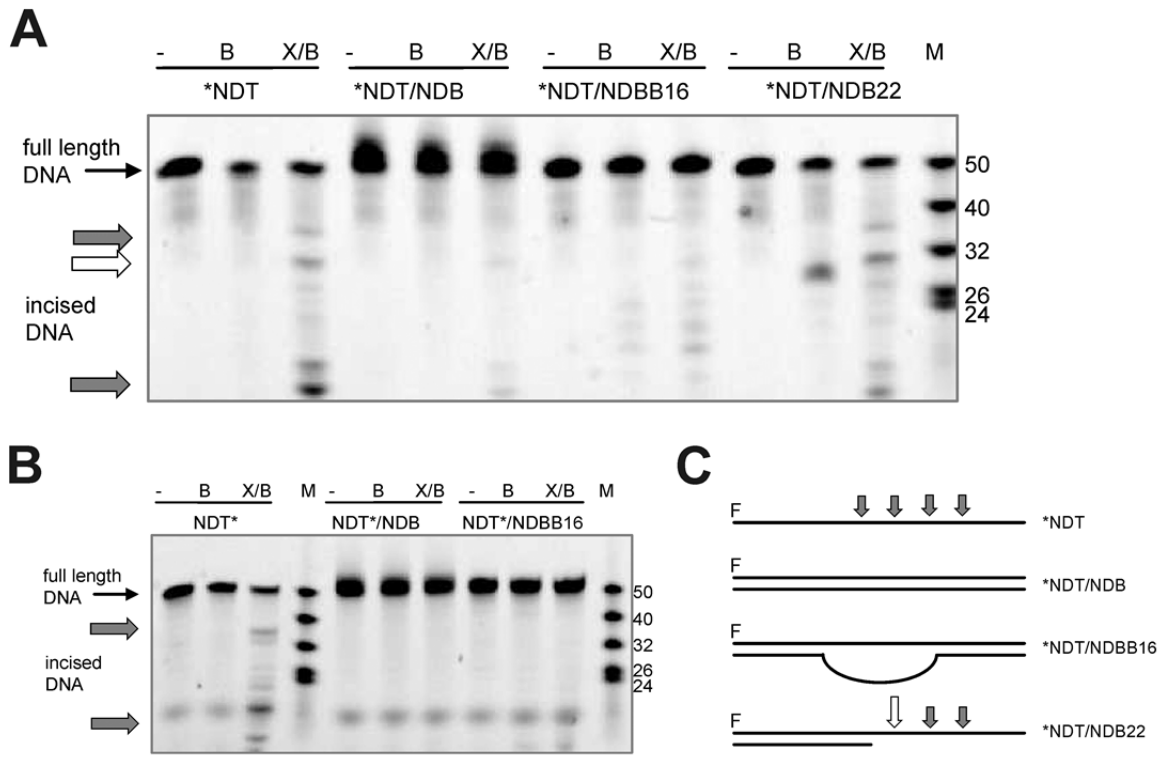

Figure 4.27 XPB/Bax1$_{comp}$ cuts ssDNA non-specifically. (A) 5' fluorescein labeled DNA was used as ssDNA or annealed with oligos to form dsDNA, 16 nt bubble substrate or a 3' overhang substrate at a final concentration of 20 nM each. XPB/Bax1 in contrast to Bax1 (both 5 µM) also incised ssDNA at different positions. (B) 3' fluorescein labeled DNA was used analogously demonstrating that XPB/Bax1 did not require a distinct free ssDNA end. (C) Schematic representation of the DNA substrates used for this study. Sites of incision by Bax1 and XPB/Bax1 are indicated by white and gray arrows, respectively.

The incision assays using ssDNA demonstrated that Bax1 in contrast to XPB/Bax1$_{comp}$ was not able to incise ssDNA (**Figure 4.27**). Both Bax1 and XPB/Bax1$_{comp}$ incised 3' overhang DNA substrates but no dsDNA or DNA containing a 16 nt-bubble.

4.2.11. Bax1 and XPB/Bax1$_{comp}$ require different divalent cations for incision activity

Several endonucleases require divalent cations. The prokaryotic endonuclease involved in NER, UvrC, requires one divalent cation for its N-terminal active site and most likely two divalent cations for the C-terminal endonuclease site (Karakas et al, 2007; Truglio et al, 2005). To investigate whether divalent cations are also essential for the incision activity of Bax1, assays were conducted in the presence

of different divalent cations: Mg^{2+}, Ca^{2+}, Mn^{2+} or the chelating reagent EDTA (**Figure 4.28**, panel A). A specific incision product could only be observed in the presence of Mg^{2+}, whereas an unspecific smear could be detected when Mn^{2+} was added. Both Ca^{2+} and EDTA completely abolished Bax1 incision activity.

Intriguingly, we observed, that the XPB/Bax1$_{comp}$ complex incised both 3' and 5' ssDNA-overhang substrates in the presence of 10 mM $MgCl_2$, $CaCl_2$, $MnCl_2$ or EDTA (**Figure 4.28**, panel B and C).

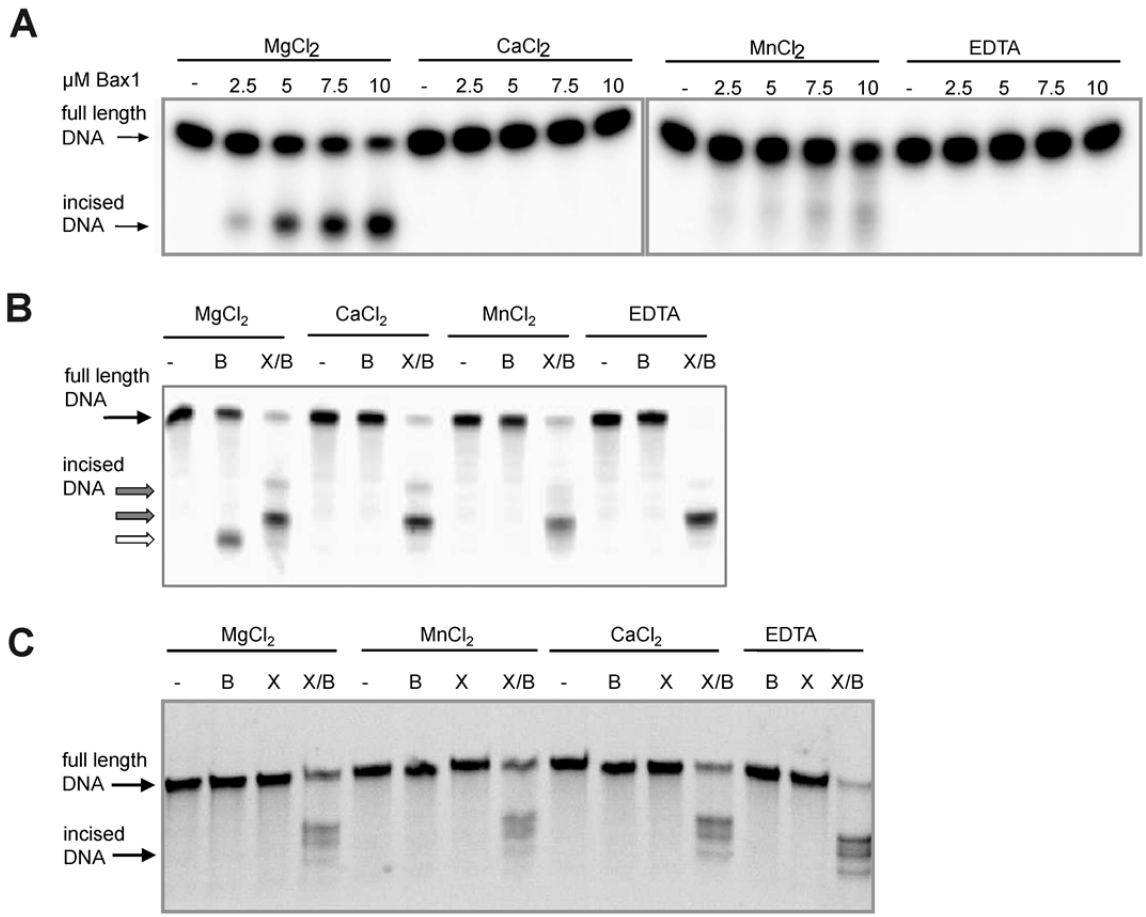

Figure 4.28 The addition of Mg^{2+} is crucial for Bax1 incision activity but not for XPB/Bax1. (A) Incision assays with different Bax1 concentrations were performed in the presence of different divalent cations or the chelating reagent EDTA (10 mM each) using the DNA substrate NDT/NDB22 (10 nM). Only in the presence of $MgCl_2$ a specific incision product was formed by Bax1. (B) XPB/Bax1 (X/B) incision of 3' overhang substrate, in contrast, was not affected by the addition of $MgCl_2$, $CaCl_2$, $MnCl_2$ or EDTA. (C) XPB/Bax1 (X/B) additionally incised 5' overhang substrates independent from divalent cations, whereas Bax1 (B) did not incise this DNA substrate. XPB (X) was applied as well to serve as control. In panels B and C, both XPB/Bax1 and Bax1 were applied at a concentration of 5 µM, the DNA concentration of the substrates F_{26}50/NDB22 and F_{26}50/NDBr25 was 20 nM, divalent cations and EDTA were used at 10 mM final.

4.2.12. XPB's helicase activity is not required for XPB/Bax1 incision

Neither Bax1 nor XPB/Bax1 incised bubble substrates comprising a region of 16 unpaired nucleotides (**Figure 4.29**, panel A). The addition of nucleotides such as ATP, the non-hydrolysable ATP-analog AMPPNP, or ADP which could have initiated XPB's helicase activity, did not induce incision of these DNA substrates.

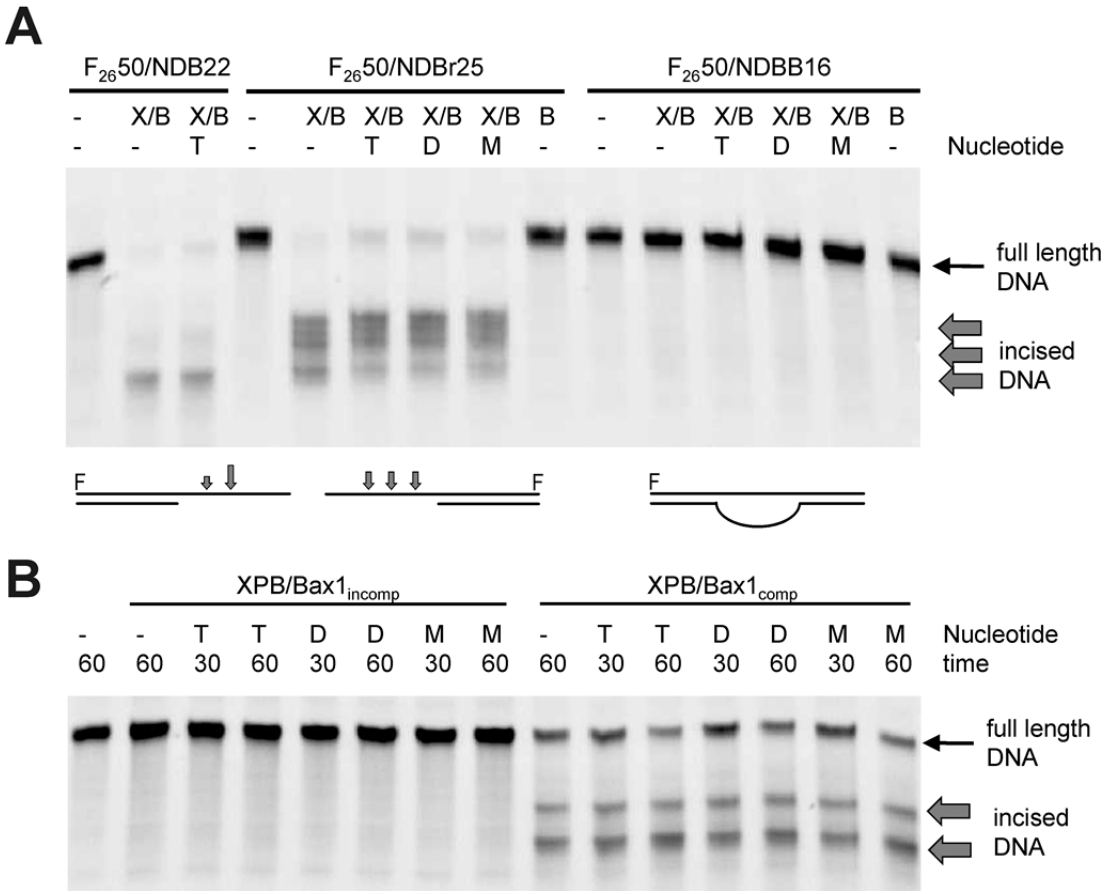

Figure 4.29 XPB/Bax1 incision occurs independently from nucleotides. (A) XPB/Bax1$_{comp}$ (X/B) was able to cut 3' and 5' overhang substrates but not bubble substrates. DNA substrates are schematically shown and the sites of incision are indicated by arrows. Incision of bubble substrates by XPB/Bax1$_{comp}$ could not be induced by the addition of nucleotides. (B) XPB/Bax1$_{incomp}$ incision of a 3' overhang substrate (NDT/NDB22) could not be triggered by adding nucleotides (T = ATP, D = ADP, M = AMPPNP, 1 mM each). XPB/Bax1 comp incision, however, was not altered by the addition of nucleotides. Proteins were applied at a final concentration of 5 µM, DNA concentration for each substrate was 20 nM.

Employing the same conditions for helicase and ATPase assays as described for incision assays, neither helicase nor ATPase activity could be detected for XPB. Since also the incision pattern obtained by XPB/Bax1 on 3' and 5' overhang substrates did not change with the addition of nucleotides, it is concluded that the helicase activity is not required for the loading of Bax1 onto DNA. Moreover,

nucleotides are not able to induce incision activity of the XPB/Bax1$_{incomp}$ complex (**Figure 4.29**, panel B).

4.2.13. XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ have the same stoichiometry but different conformations

Since Bax1 was shown to be a Mg^{2+}-dependent endonuclease (4.2.11), the impact of divalent cations on XPB/Bax1 complex formation by the addition of divalent cations to the cell lysate was investigated. The effect of depletion of divalent cations was tested by extensive dialysis of the XPB/Bax1 complex against the chelating agent EDTA prior to size-exclusion chromatography.

After addition of different cations to the cell lysate, the elution profiles contained three major peaks (**Figure 4.30**, panel A, gray curve), which were consistent with the XPB/Bax1$_{comp}$ complex at an elution volume of ~60 ml (gray arrow in **Figure 4.30**, panel A), the XPB/Bax1$_{incomp}$ complex at an elution volume of ~75 ml (black arrow in **Figure 4.30**, panel A) and Bax1 eluting at ~85 ml (white arrow in **Figure 4.30**, panel A). After dialysis against EDTA only two peaks were observed in the elution spectrum (**Figure 4.30**, panel A, black curve), consistent with the XPB/Bax1$_{incomp}$ complex and Bax1. EDTA induced Mg^{2+}-depletion thus appears to prevent the formation of the XPB/Bax1$_{comp}$ complex.

Potential structural differences in the two different XPB/Bax1 complexes, XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$, were examined by analytical ultracentrifugation (AUC), limited proteolysis, CD-spectroscopy and AFM. The two complex types showed similar behavior in all of these experiments. AUC analysis of two to four different experiments yielded sedimentation coefficients of 4.8 for both types of XPB/Bax1 complexes, with standard deviations of 0.2 and 0.05 for XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$, respectively (**Figure 4.30,** panel B). The frictional ratios which provided the best fit to the data (RMSDs between 0.0038 and 0.016) were 1.5 ± 0.02 for XPB/Bax1$_{comp}$ and 1.7 ± 0.2 for XPB/Bax1$_{incomp}$. These parameters translated into approximate molecular weights of 100 ± 15 kDa and 115 ± 20 kDa for XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$, respectively. Frictional ratios > 1.4 indicate elongated protein structures. Here, especially the incision incompetent complex showed frictional ratios consistent with deviation from a globular protein complex

structure. However, stronger elongation of the incision-incompetent complex compared to the incision-competent form is in contrast to the higher elution volume for the incision incompetent complex in analytical size-exclusion chromatography (**Figure 4.25**, panel A). Smaller elution volumes in size-exclusion chromatography are suggestive of either a larger molecular weight or stronger structural elongation possibly leading to enhanced interactions with the column matrix. The longer retention time on the chromatography column and therefore larger elution volume of XPB/Bax1$_{incomp}$ together with its larger frictional ratio in AUC sedimentation experiments may result from local structural changes in this complex leading to the exposure of protein parts with enhanced affinity to the column matrix (Michael Fried [University of Kentucky], personal communication). However, both analytical size-exclusion chromatography and AUC suggest maximal differences in molecular weight for the two complex types that are too low to allow for different complex stoichiometries. Calculated from the amino acid sequence, the theoretical molecular weight for a heterodimeric XPB/Bax1 complex is 98 kDa. Calibration of the size-exclusion chromatography column Superdex75 10/300GL provided elution volumes which indicate a molecular weight of 119 ± 13 kDa and 79 ± 11 kDa for the XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ complex, respectively.

In addition, AFM images showed similar AFM volumes for the two different complex types, which correspond to 118 kDa for XPB/Bax1$_{comp}$ and 115 kDa for XPB/Bax1$_{incomp}$ (**Figure 4.30,** panel C). The two complex types showed slight differences in the amount of dissociated Bax1 and XPB monomers in the samples (peaks at approximately 50 nm^3 in the distributions in **Figure 4.30**, panel C), which may be interpreted as differences in complex stability (**Figure 4.30**, panel C). Interestingly, a different deposition behavior between the two complex types was observed: XPB/Bax1$_{incomp}$ showed consistently higher affinity to the mica substrate surface, resulting in higher surface coverage at the same protein concentration (Ingrid Tessmer, personal communication). Different affinities to the AFM substrate surface for XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ are consistent with local differences in protein surface structure.

From these data obtained by size-exclusion chromatography, AUC, and AFM, it can be concluded that there are no stoichiometric differences between the two complex types XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$. However, differences in elution,

sedimentation, and AFM deposition behavior suggest that there are small, local conformational differences between the two types of XPB/Bax1 complexes.

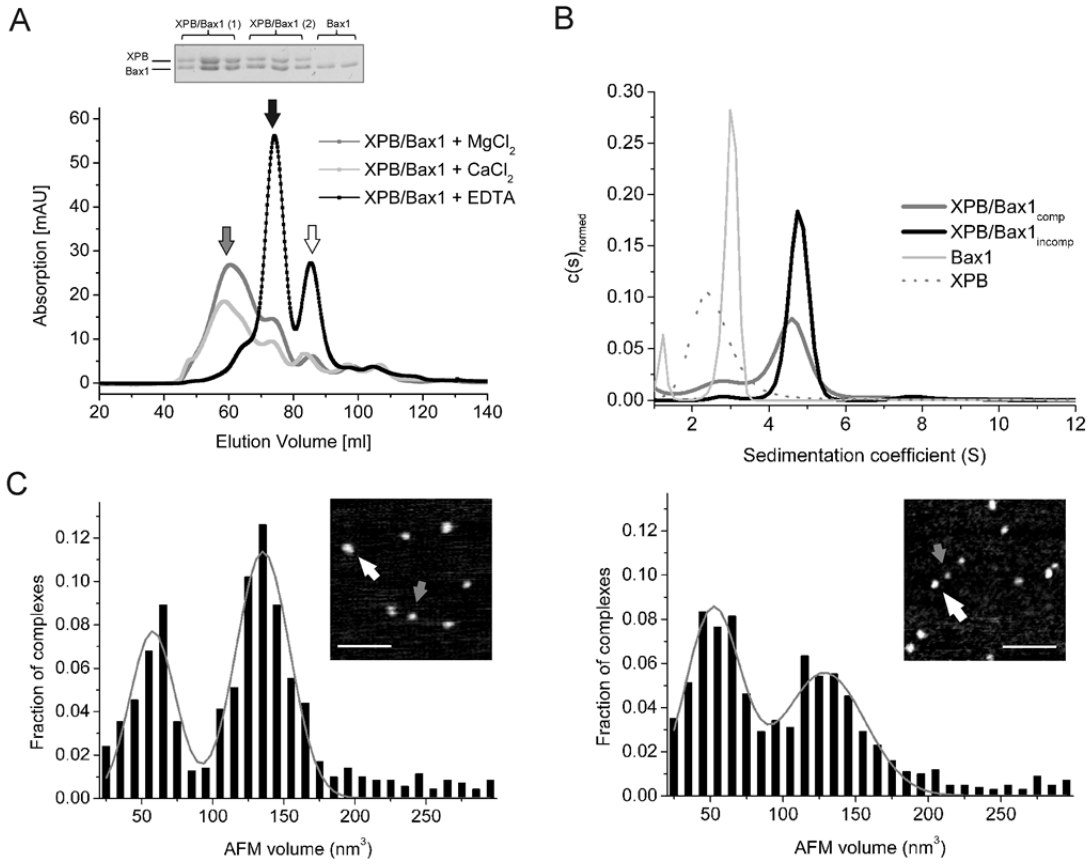

Figure 4.30 XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ share the same stoichiometry. (A) The elution profile in size-exlusion chromatography (HiLoad™ 16/60 Superdex™200 pg column) of XPB/Bax1 could be triggered by the addition of divalent cations or their removal by dialysis against the chelating agent EDTA. Gray, black and white arrows point to XPB/Bax1$_{comp}$, XPB/Bax1$_{incomp}$ and Bax1 peaks, respectively. The SDS-gel inset represents fractions from the size-exclusion chromatography of XPB/Bax1 with $MgCl_2$ (B) AUC sedimentation velocity experiments showed comparable sedimentation coefficients (4.8 S) for XPB/Bax1$_{comp}$ (gray line) and XPB/Bax1$_{incomp}$ (black line). For reference, the sedimentation coefficient profiles of Bax1 (light gray line) and XPB (gray dotted line) alone are also shown. (C) AFM volume analysis of XPB/Bax1$_{comp}$ (left, n=705) and XPB/Bax1$_{incomp}$ (right, n=994) complexes revealed the existence of a heterodimeric state for both complex types. Double-Gaussian fits to the AFM volume distributions (with $R^2 > 0.91$) revealed volume states of 135 nm^3 for XPB/Bax1$_{comp}$ and 130 nm^3 for XPB/Bax1$_{incomp}$. These AFM volumes translated into molecular weights of 118 kDa and 115 kDa, respectively, consistent with a heterodimeric state. Both complex types also showed dissociation into the individual protein components at the low concentrations of AFM sample deposition (nM-range). XPB and Bax1 individual volumes yielded maxima in both distributions at approximately 50 nm^3, corresponding to 50 ± 2 kDa. The insets represent exemplary zooms into AFM images for each of the two complex types, with scale bars of 100 nm. White and gray arrows indicate heterodimeric and monomeric particles, respectively.

In principle, conformational changes with regard to secondary structure elements can be detected by CD-spectroscopy. Thus, both XPB/Bax1 complexes were employed at identical conditions in order to register CD-spectra. These measurements resulted in identical curves (**Figure 4.31**, panel A), indicating that the differences in incision activity observed for the XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ complex are not caused by unfolding or major conformational changes in secondary structure elements.

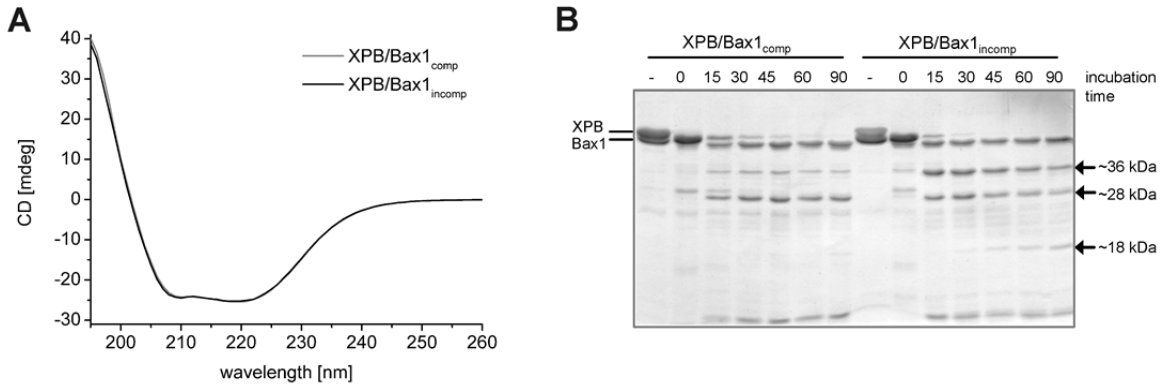

Figure 4.31 XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ adopt slightly different conformations. (A) CD-spectroscopy analysis of both XPB/Bax1 complexes were conducted at 20 °C using 5 µM protein each in a buffer containing 20 mM Na phosphate, pH 8.0. 5 spectra were accumulated to optimize the signal to noise ratio. Both spectra show the same pattern thus indicating that the two different XPB/Bax1 complexes adopt the same mixed α/β-fold. Different incision activities observed for XPB/Bax1$_{comp}$ (gray) and XPB/Bax1$_{incomp}$ (black) thus do not result from unfolding. (B) Limited proteolysis experiments of XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ using trypsin result in different digestion patterns, indicating different conformations for the two complex forms (XPB = 52 kDa, Bax1 = 46 kDa). Arrows point to prominently different digestion products and their estimated molecular weights are shown.

As different protein conformations result in different digestion patterns (Fontana et al, 2004; Heiring & Muller, 2001), potential conformational differences in the XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ complexes may be detectable by limited proteolysis. The presence of equal amounts of XPB and Bax1 in both XPB/Bax1 complexes was confirmed by SDS-gel electrophoresis (**Figure 4.31**, panel B). Upon addition of the protease trypsin both complexes degraded over time. Interestingly, for the two different types of complexes, different digestion patterns could be observed (indicated by arrows in **Figure 4.31**). The combination of CD-spectroscopy and limited proteolysis suggests that there are only slight conformational differences or differences caused by domain movement between

the two XPB/Bax1 complexes which are sufficient to change the accessibility of protease cleavage sites but do not induce different CD-spectra.

These experiments strongly suggest that the two XPB/Bax1 complexes identified via size-exclusion chromatography slightly differ in their conformations whereas they still comprise XPB and Bax1 at the same stoichiometry.

4.2.14. Bax1 and XPB/Bax1 are visualized on DNA substrates by AFM imaging

AFM imaging was used to directly visualize binding of the two complex types, XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp,}$ as well as Bax1 alone to different linear DNA substrates (**Figure 4.32**, panel A). AFM analyses revealed that Bax1 binds DNA as a monomer, consistent with previous results demonstrating that the monomeric form of the protein is functionally more active (4.2.7): a molecular weight of 53 ± 9 kDa was calculated from the AFM volumes of DNA-bound Bax1 complexes (after subtraction of the covered DNA volume, see Material and Methods).

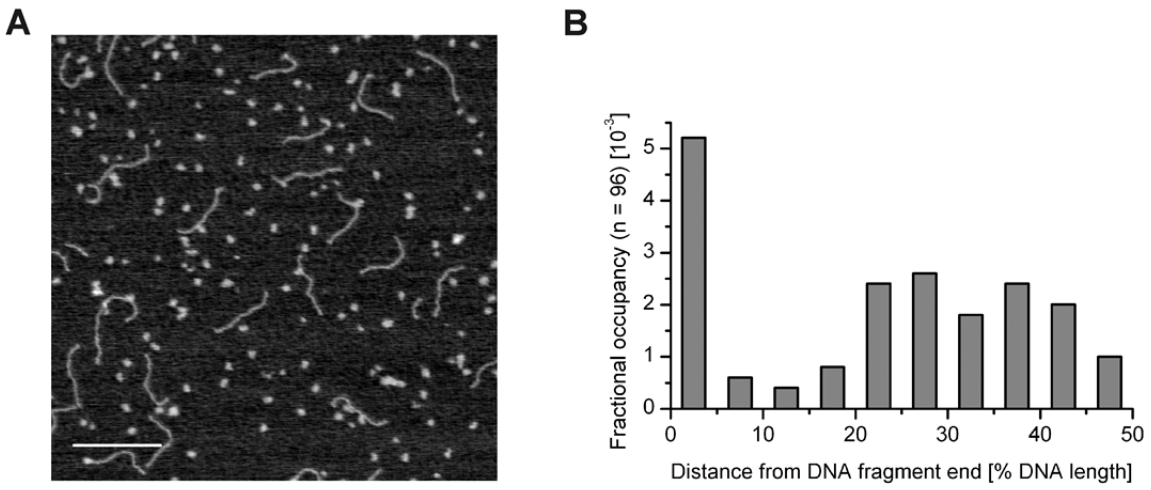

Figure 4.32 XPB/Bax1 binds to DNA non-specifically and as a heterodimer. (A) AFM image of XPB/Bax1$_{incomp}$ complexes bound to a 3' ssDNA overhang substrate. Volume analyses of representative DNA-bound complexes revealed molecular weights of approximately 110 kDa for both XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$. The scale bar is 200 nm, the surface area is 1 x 1 µm^2. (B) The position distribution of XPB/Bax1 complexes on DNA shows preferential occupation of DNA fragment ends and no specificity for any DNA strand internal sequence or base pair composition.

Both XPB/Bax1 complexes bind to DNA as a heterodimer: from the AFM volumes of DNA-bound complexes, approximate molecular weights of 112 ± 31 kDa and 107 ± 35 kDa for XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ were calculated, respectively (after subtraction of the estimated DNA volume). Under the applied incubation conditions using 1 µM protein and 100 nM DNA, non-saturating low coverage of the DNA with protein complexes (0.26 ± 0.15 / DNA fragment) was observed. AFM results for XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$ were indistinguishable. Distributions of XPB/Bax1 complexes on the DNA fragments suggest a specificity of approximately 200-fold (± 40) for DNA fragment ends (**Figure 4.32**, panel B). The position distributions further demonstrate an absence of any preference for DNA sequences or base pair composition within the DNA fragments (**Figure 4.32**, panel B).

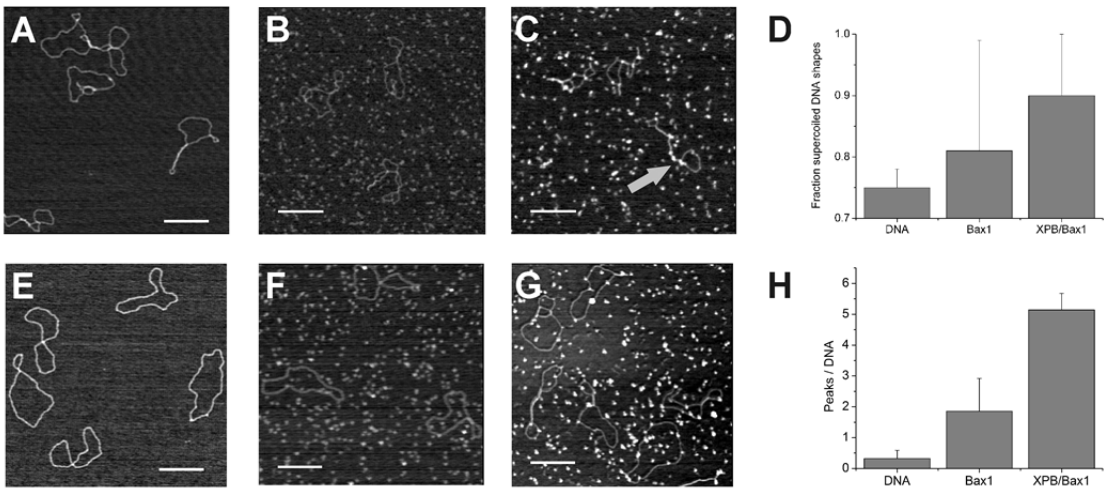

Figure 4.33 AFM imaging of Bax1 and XPB/Bax1 binding to circular DNA. Top row: AFM images of pUC19 supercoiled DNA (A) in the absence of protein, (B) in the presence of 1 µM Bax1, and (C) in the presence of 1 µM XPB/Bax1. (D) Quantification of the percentage of supercoiled DNA structures (versus open relaxed circles such as seen in (E)) upon protein binding reflects an effect on DNA superstructure (n = 82 and 116 for Bax1 and XPB/Bax1, respectively). We observe that both Bax1 and XPB/Bax1 bind to supercoiled DNA conformations, possibly supported by protein-protein interactions on the DNA (gray arrow in C). Large variations in the graph reflect ambiguity in DNA superstructure. In contrast, protein complexes on relaxed DNA (E-G) are more defined. Bottom row: AFM images of nicked pUC19 DNA (E) in the absence of protein, (F) in the presence of 0.5 µM Bax1, and (G) in the presence of 0.5 µM XPB/Bax1. (H) Quantitative analysis of the number of protein complexes bound to DNA (n = 218 and 389 for Bax1 and XPB/Bax1, respectively) shows increased binding of XPB/Bax1 to DNA compared to Bax1 alone, for protein concentrations of 0.5 -1 µM.

To investigate physiologically relevant DNA binding by Bax1 and XPB/Bax1 complexes, we imaged complex formation with circular plasmid DNA (pUC19) by AFM. The images suggest a correlation between protein-DNA complex formation and DNA superstructure (**Figure 4.33**, panel A-D). Because quantification of protein-DNA complexes on the plectonemic, supercoiled DNA structures was difficult, Bax1 and XPB/Bax1 were compared by binding to relaxed DNA circles (instead of negatively supercoiled DNA). Panels E-H in **Figure 4.33** show enhanced protein complex formation on DNA for the XPB/Bax1 complex compared to Bax1 in the absence of XPB (5.1 ± 0.5 versus 1.9 ± 1.1, respectively), suggesting that XPB may function in loading Bax1 onto circular dsDNA.

Loading of Bax1 onto DNA by XPB is supported by equilibrium dissociation constants, K_D, obtained by bio-layer interferometry (**Figure 4.34**), which revealed that Bax1 by itself binds to DNA with much lower affinity than XPB or the XPB/Bax1 complexes.

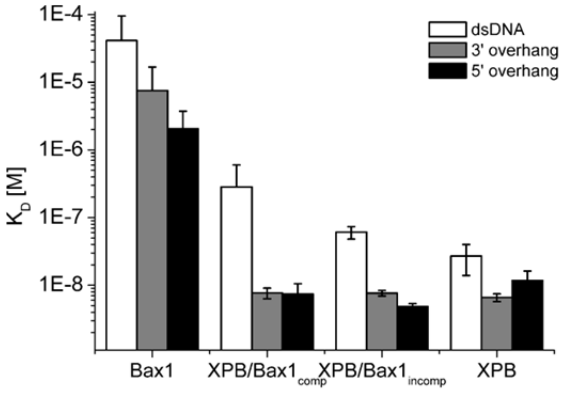

Figure 4.34 DNA binding of Bax1, XPB/Bax1 complexes and XPB towards different DNA substrates. DNA substrates NDT/NDB (dsDNA), NDT/NDB22 (3' overhang) and NDT/NDBr25 (5' overhang) were used in bio-layer interferometry measurements to obtain K_D-values for Bax1, both XPB/Bax1 complexes and XPB.

For Bax1, equilibrium dissociation constants in the low µM-range were obtained, almost independent of the nature of the DNA substrate. For the XPB/Bax1 complexes, however, a binding preference of about one order of magnitude were observed for substrates containing ssDNA-overhangs, as compared to dsDNA (**Figure 4.34**). For example, XPB/Bax1$_{incomp}$ showed equilibrium dissociation

constants of 8.0 ± 0.9 nM and 54 ± 5 nM for ssDNA overhang and dsDNA substrates, respectively. These results are consistent with the AFM data (**Figure 4.33**) and indicate that XPB mediates Bax1-binding onto DNA.

5. Discussion

5.1. Insights into the prokaryotic nucleotide excision repair

Nucleotide excision repair in prokaryotes is achieved by the UvrABC machinery (Truglio et al, 2006a; Van Houten et al, 2005). UvrA detects helical distortions in the DNA and loads UvrB onto the DNA which then verifies the presence of a NER lesion forming a tight UvrB-DNA pre-incision complex. Finally, UvrC is recruited to conduct both the 3' and 5' incision reaction to excise the damaged oligonucleotide. It has been demonstrated previously that UvrB is able to bind to DNA bubble substrates and also is able to detect DNA damages even in the absence of UvrA (Moolenaar et al, 2000b; Zou & Van Houten, 1999). Although many biochemical studies have been published unraveling the mysteries of the NER pathway, the mechanism of DNA damage verification, however, still remains unclear (Moolenaar et al, 2001; Truglio et al, 2006b). In addition, structural insights into the recruitment and positioning of UvrC to the pre-incision complex as well as the mechanism leading to the dual incision reactions are elusive. Therefore, the first part of this work seeked to investigate the DNA damage recognition and incision mechanism conducted by the proteins UvrB and UvrC.

5.1.1. UvrBC forms specific complexes on damaged DNA

In the present work, the question whether the UvrBC complex is able to specifically incise damaged DNA even in the absence of UvrA was addressed by incision assays employing double-stranded DNA and DNA substrates with 6 nt bubbles at different positions with regard to a DNA lesion (4.1.5). Both set-ups containing either UvrABC or UvrBC yielded an incision product of the same size on a denaturing polyacrylamide gel independent of the position of the bubble. This finding shows that UvrA does not seem to be essential for DNA damage verification and also indicates that the formation of a specific pre-incision complex is independent of UvrA as already suggested by previous studies (Zou & Van Houten, 1999). In the present study, additional DNA substrates were employed which contain a 6 nts unpaired region at different positions which do not correlate

with the position of the DNA lesion. Quantification of the incision assays showed that UvrB preferentially bound to a bubble, but, with lower affinity, is also able to bind to dsDNA. From the observation that the 5' bubble substrate is preferred over the 3' bubble substrate, we deduced that UvrB is able to slide along the DNA favorably in a 5' → 3' direction until it encounters a DNA damage to be detected and verified. UvrC is then recruited to the pre-incision complex to perform the incision reaction. This observation also suggests that UvrB binds to the damaged oligonucleotide to verify the presence of the damage. DNA damages on the opposite strand are ignored and UvrB does not seem to transfer from the leading strand to the lagging strand or vice versa. Thus, UvrB would directly interact with the DNA lesion in contrast to XPC which employs three β-hairpins to probe the non-damaged DNA-strand for locally unpaired nucleotides (Min & Pavletich, 2007). Recently, also UvrA was reported to employ an indirect readout mechanism to detect DNA lesions through the deformability of the DNA helix (Jaciuk et al, 2011).

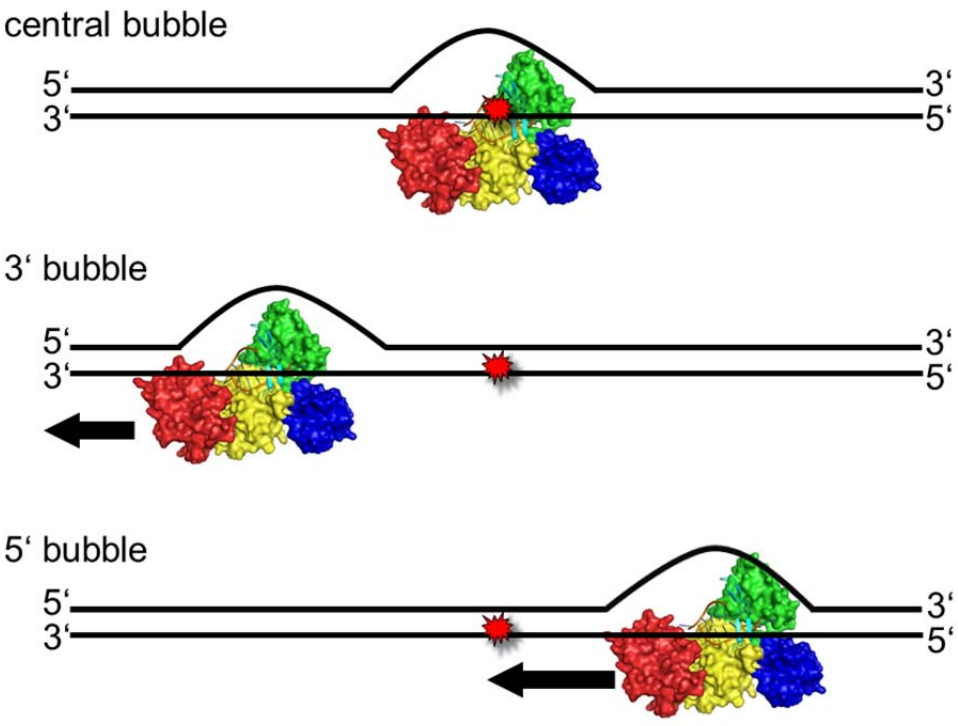

Figure 5.1 Schematic presentation of the sliding properties of UvrB prior to incision by UvrC. UvrB binds to DNA at the site of a bubble to then slide along the DNA preferentially in 5' → 3' direction. Encountering a damage, UvrB recruits UvrC which incises DNA 3' and 5' to the DNA lesion.

It has been shown recently, that UvrB in complex with UvrA is able to slide along DNA, whereas UvrA by itself only binds transiently to DNA (Kad et al, 2010). UvrB thus would reduce UvrA's three-dimensional search for damaged DNA to a one-dimensional search (Kad et al, 2010). Our data provide a first hint that UvrB is able to slide on DNA either alone or in complex with UvrC. The proposed sliding mechanism based on the incision assays presented in this work is schematically illustrated in **Figure 5.1**.

5.1.2. Structural analysis of the UvrB-DNA complex

The biochemical data presented here suggest that UvrB directly binds and sticks to the DNA lesion but does not probe the non-damaged DNA for unpaired regions as it has been shown for the XPC homolog Rad4 (Min & Pavletich, 2007). However, the molecular interactions leading to the discrimination of damaged from non-damaged DNA remain elusive. Therefore, we sought to solve the crystal structure of UvrB in complex with damaged DNA to decipher the process of damage verification. In 2006, a crystal structure of UvrB in complex with a ssDNA containing a fluorescein-adduced thymine has been published (PDB: 2FDC) (Truglio et al, 2006b). Since the DNA lesion was only partially ordered in the crystal structure, no information about the damage recognition process could be deduced. However, part of the ssDNA was clamped behind the protruding β-hairpin, which was previously shown to be important for DNA damage detection (Moolenaar et al, 2001; Skorvaga et al, 2002). Originating from the position and orientation of the DNA bound to UvrB, a model was generated suggesting that the β-hairpin inserts between the two strands of the DNA (Truglio et al, 2006b). Based on this model, DNA substrates containing a fluorescein-adducted thymine within an unpaired region in dsDNA were created to allow specific binding of UvrB in the absence of UvrA (Hong Wang and Bennett Van Houten, personal communication). To stabilize the DNA bubble structure, a self-complementary DNA substrate including a 3 nts turn and thus adopting a secondary structure referred to as hairpin bubble was chosen to obtain the UvrB-DNA complex.

In the present study, a hairpin bubble substrate comprising 27 nucleotides was used for co-crystallization with UvrBΔ4 (see also **Figure 4.7**). UvrBΔ4 was thought

to bind to the DNA substrate by inserting its β-hairpin into the 6 nts bubble. In the crystal structure of the UvrBΔ4-DNA complex (present work; **Figure 4.6**; **Figure 4.7**; **Figure 5.4**, depicted in green), however, only a small single-stranded part of the DNA is visible.

The crystal structure of the UvrBΔ4-DNA complex was superimposed in COOT (Krissinel & Henrick, 2004) with two previously published crystal structures of UvrB, namely apo-UvrB (PDB: 1D9X) and the UvrB-ATP complex (PDB: 1D9Z) depicted in blue and gray, respectively (**Figure 5.2**). The structures presented in **Figure 5.2** superimpose precisely with an overall r.m.s.d. of 0.99 Å and 0.95 Å, respectively.

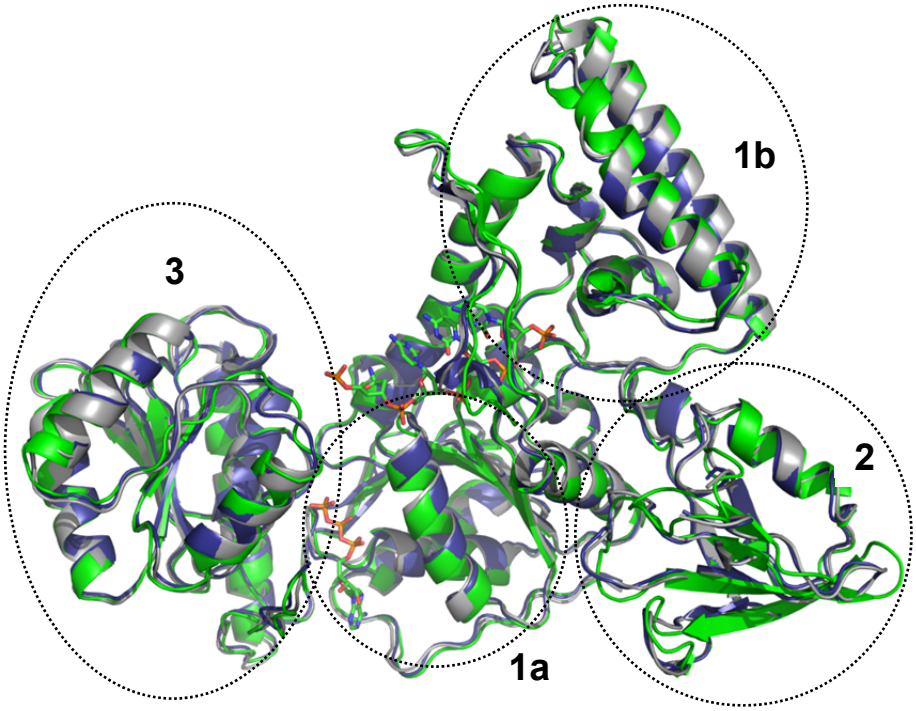

Figure 5.2 Superposition of apo-UvrB, the UvrB-ATP complex and the UvrBΔ4-DNA complex. The structures of apo-UvrB (PDB: 1D9X, blue), the UvrB-ATP complex (PDB: 1D9Z, gray) and the UvrBΔ4-DNA complex (green) are superposed. The DNA as well as the ATP is represented as a colored stick model. The domain architecture is highlighted by dotted circles. The numbers are related to the domains nomenclature 1a, 1b, 2 and 3.

The most striking difference between apo-UvrB (PDB: 1D9X) and the UvrBΔ4-DNA complex is observed with regard to the DNA binding site at the β-hairpin. In the apo-UvrB structure, a short α-helix is located at the base of the β-hairpin which points into the DNA binding site (**Figure 5.3**, highlighted by a black arrow). Upon DNA binding, the α-helix undergoes a conformational change. In the

UvrBΔ4-DNA structure, this part appears as random coil (**Figure 5.3**, depicted in green).

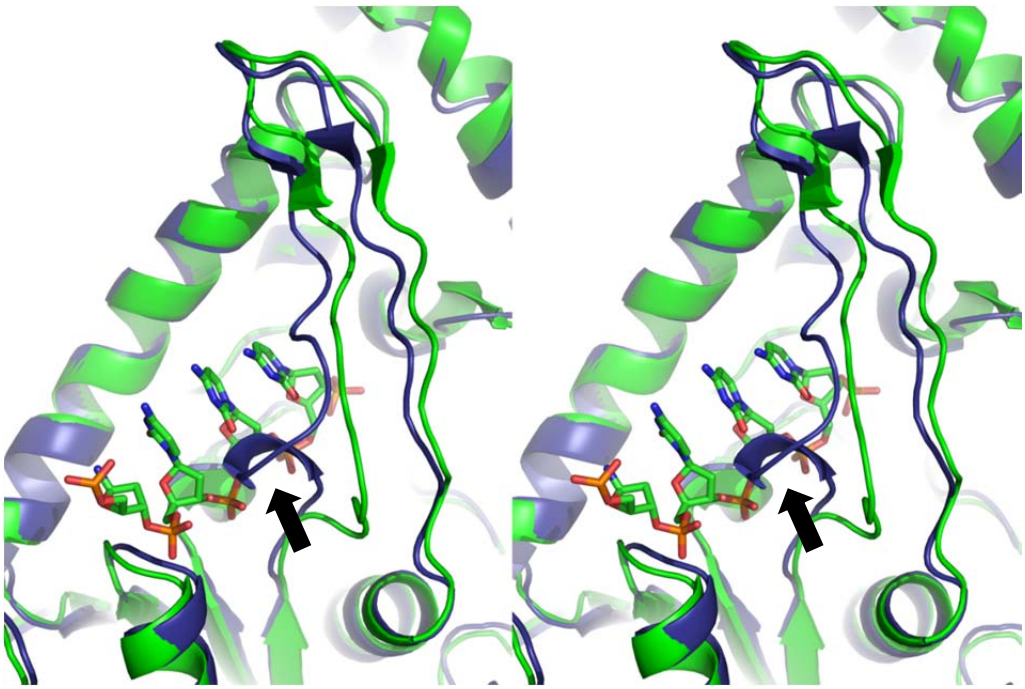

Figure 5.3 Stereo view of the UvrBΔ4-DNA complex superposed with apo-UvrB. The UvrBΔ4 structure (green ribbon) with bound DNA (colored stick model) is superimposed with the apo-UvrB structure (blue ribbon). The black arrow highlights a short α-helix in the apo-UvrB structure which undergoes a conformational change upon DNA binding.

Contrarily, a 2-stranded β-sheet is formed at the bottom of the β-hairpin in the UvrB-DNA structure (PDB: 2FDC) (Truglio et al, 2006b). Superposition of the UvrBΔ4-DNA complex with the published UvrB-DNA complex (**Figure 5.4**, PDB: 2FDC, depicted in gray) visualizes the differences of the β-hairpin conformations. Despite the observed discrepancies, the superposition reveals that the protein adopts almost the same overall conformation with an r.m.s.d of 1.28 Å. Also the partially visible DNA is bound at the same binding site in the same orientation.

The random coiled structure of the β-hairpin in the UvrBΔ4-DNA structure might be an artifact caused by the poor electron density. Several side chains are only partially defined by the electron density map, whereas the main chain is discernable as a rather bulky tube but not clearly defined. The poor electron density does not allow the discrimination of the actual conformation from other possible conformations. Therefore, the current model represents only one potential

topology slightly differing with regard to their atom position, bond lengths, bond angles as well as dihedral angles. To assign a particular secondary structure element, the values for the main chain torsion angles φ, ψ must not deviate from a defined range given for ideal β-sheets or α-helices. Otherwise, these parts are recognized as random coils.

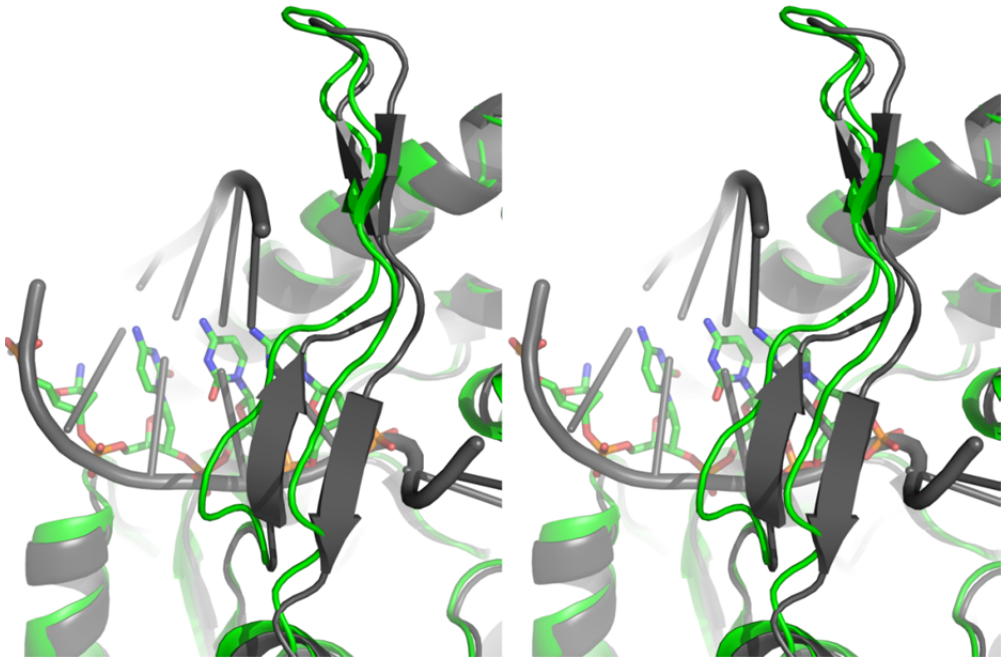

Figure 5.4 DNA binding site in the UvrBΔ4-DNA complex and the UvrB-DNA complex (stereo view). UvrBΔ4 (green) and UvrB (PDB: 2FDC, gray) adopt almost the same conformation. The DNA substrates bound to the proteins (represented as colored stick model and as gray ribbon, respectively) are located at the same binding site in the same orientation. The β-hairpins in UvrBΔ4-DNA (green) and UvrB-DNA (gray) adopt a different secondary structure. Black arrows point to the β-sheet in the UvrB-DNA complex which changes into a random coil structure in the UvrBΔ4-DNA complex.

However, the observed difference of the UvrB-DNA and the UvrBΔ4-DNA complex might be caused by the different DNA molecules bound in the UvrB structures (**Figure 5.4**). Binding of duplex DNA in contrast to single-stranded DNA might result in additional conformational changes and might also stabilize secondary structure elements in the DNA binding site by protein-DNA interactions.

Furthermore, it needs to be considered that although domain 4 is not visible in any *B. caldotenax* UvrB structure, the UvrBΔ4-DNA complex structure presented in this work contains a truncation of domain 4 by introducing a pre-mature stop-codon. Since the auto-inhibitory effect of domain 4 might be allosteric, as

suggested for UvrB's ATPase activity (Wang et al, 2006), the increased affinity of UvrBΔ4 for DNA could be explained by a more flexible β-hairpin.

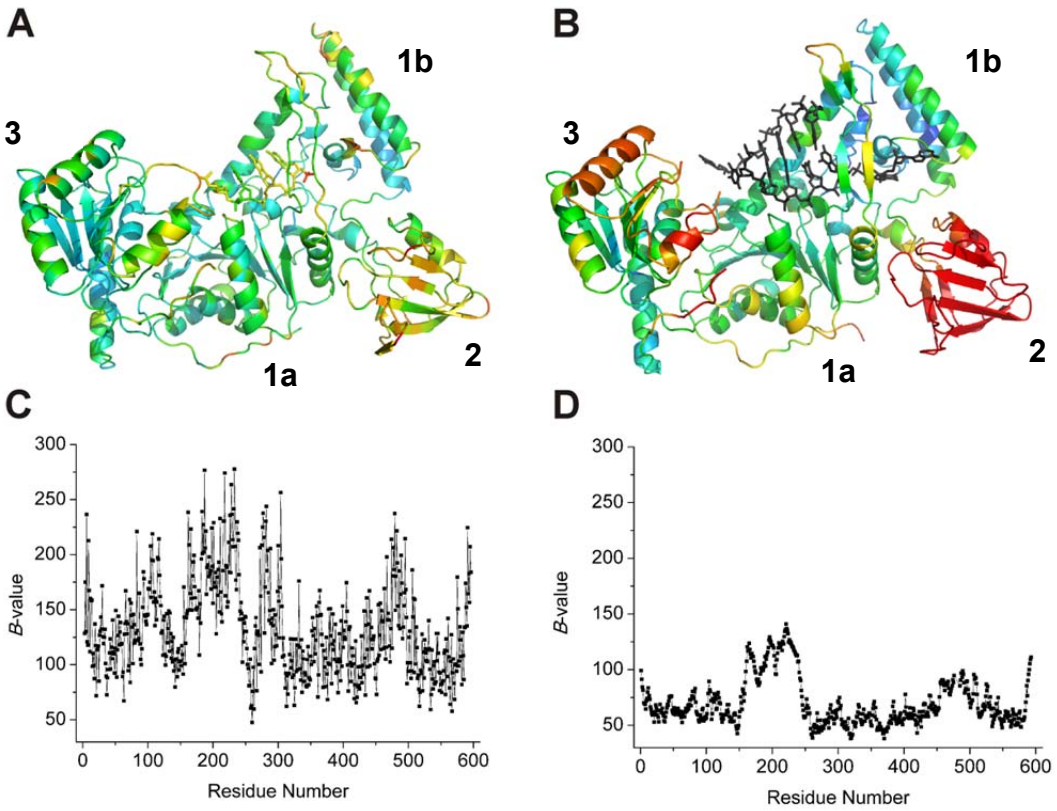

Figure 5.5 B-values of the UvrBΔ4-DNA and UvrB-DNA structure. (A) The UvrBΔ4-DNA crystal structure and (B) the UvrB-DNA structure (PDB: 2FDC) are colored with regard to the relative B-factors. Red is related to high B-values whereas blue corresponds to low B-values. The DNA is depicted as black stick model. The absolute B-values observed for the UvrBΔ4-DNA structure (C) and the UvrB-DNA structure (D) (PDB: 2FDC) are plotted against the corresponding residue number. Please note that the 3D-representations in (A) and (B) show relative B-values, whereas the plots in (C) and (D) provides absolute *B*-values.

The crystal structure of *B. subtilis* UvrB (PDB: 2D7D) (Eryilmaz et al, 2006; Waters et al, 2006) is assumed to comprise a C-terminally truncated protein corresponding to the UvrBΔ4 variant described in the present work. The *B. subtilis* UvrB was co-crystallized with ADP and with either a tri-nucleotide (Eryilmaz et al, 2006) or in complex with a penta-nucleotide containing a fluorescein moiety attached via a flexible linker (Waters et al, 2006). Both structures (PDB: 2D7D and PDB: 2NMV) exhibit the 2-stranded β-sheet at the bottom of the hairpin as also observed for UvrB in complex with the duplexed DNA. Thus, β-sheet formation

seems to be linked to DNA binding to either single-stranded or double-stranded DNA. The absence of this β-sheet in the UvrBΔ4-DNA structure (present work) therefore is maybe due to an artifact caused by the low resolution. As outlined above, the poorly defined electron density map facilitates slight deviations from the ideal φ, ψ angles defined for a β-strand conformation. Previously published UvrB structures were solved at a similar resolution in the range of 2.1 Å to 3.3 Å, however, the Wilson *B*-values and the overall *B*-values were lower indicating a more defined crystal structure. The UvrB-DNA structure (PDB: 2FDC), for example, displays an overall *B*-factor of 74.2 $Å^2$ in contrast to 137.0 $Å^2$ for the UvrBΔ4-DNA structure presented here despite a very similar solvent content of 62 % and 68 %, respectively. The distribution of *B*-values within the crystal structures is visualized in **Figure 5.5**. The 3D-representation of the UvrBΔ4-DNA structure (present work) and the UvrB-DNA structure (PDB: 2FDC) compares relative *B*-values, whereas the plot provides absolute *B*-values (**Figure 5.5**). Interestingly, the *B*-values for domain 2 (residues 151-251), one part of domain 3 (residues 450-510) and for the C-terminus are elevated in both UvrB structures, indicating a high flexibility for these regions.

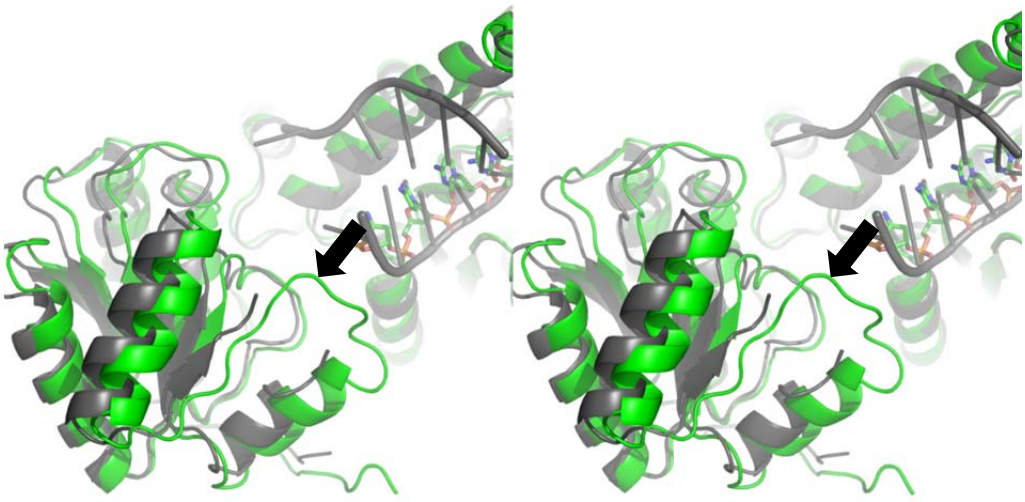

Figure 5.6 Different conformations of a loop within domain 3. The black arrow points to a disordered loop in the UvrB-DNA structure (the last and first residue visible are His476 and Ile485, respectively). In the UvrBΔ4-DNA structure, this loop was ordered possibly by a stabilization and elongation of the successive α-helix.

The increased *B*-values in domain 3 and the correlated flexibility becomes apparent in the structural discrepancy between the UvrBΔ4-DNA structure

(present work) and the previously published UvrB structure (PDB: 2FDC) as depicted in **Figure 5.6**.

The loop region in domain 3 comprised by residues His476 to Ile485, which is disordered in the UvrB-DNA structure, became visible in the UvrBΔ4-DNA structure (see black arrow in **Figure 5.6**). The subsequent helix could also be extended and starts already at residue Glu483 instead of Ile488 in the UvrB-DNA structure. This loop segment seems to be very flexible which is also indicated by the increased *B*-values (see also **Figure 5.5**) and is thus not detectable in the UvrB-DNA structure (**Figure 5.6**).

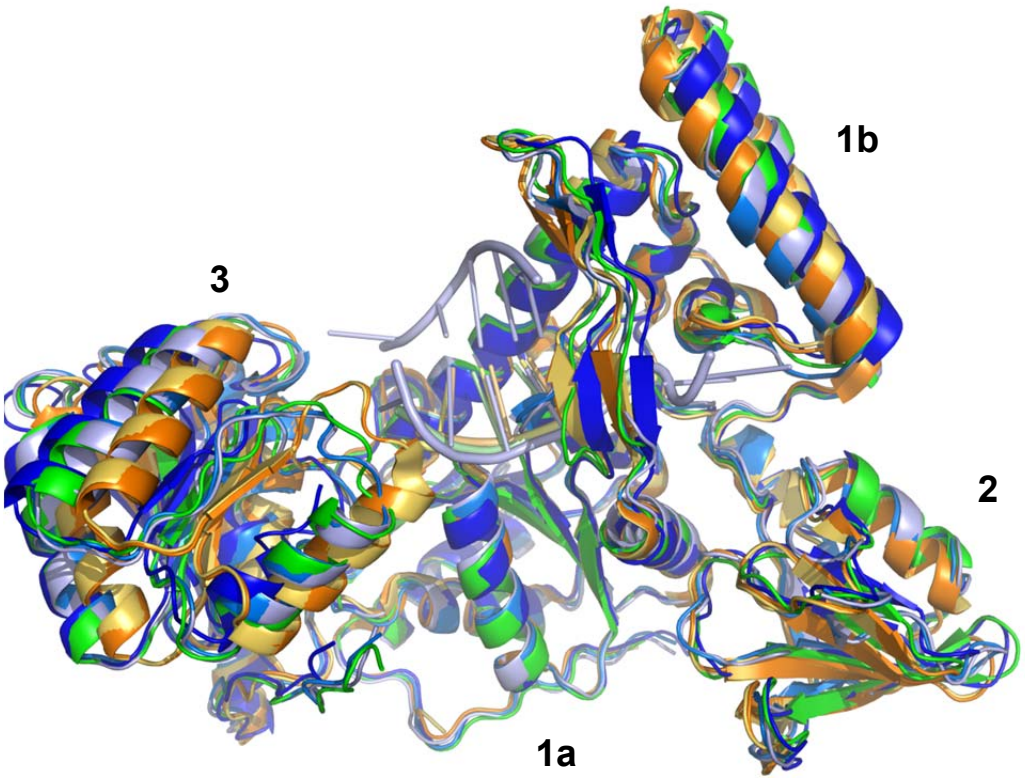

Figure 5.7 Superposition of *B. subtilis* and *B. caldotenax* UvrB structures. The UvrBΔ4-DNA structure (shown in green) is superimposed with UvrB structures (PDB codes: 2FDC, 1D9X, 19DZ, 2D7D and 2NMV are depicted in dark blue, blue, light blue, orange and yellow, respectively; the DNA substrates bound are depicted in the corresponding faint color). Different conformations of domain 3 and of the β-hairpin suggest flexibility for these parts of the protein.

Superposition of UvrB crystal structures reveals that the β-hairpin as well as domain 2 and domain 3 are rather flexible whereas domain 1a seems to adopt the same conformation among the six superimposed structures (**Figure 5.7**). This is also reflected by the corresponding *B*-values and the function of these domains: the β-hairpin binds to DNA and was shown to be crucial for DNA damage

detection (Moolenaar et al, 2001; Skorvaga et al, 2002; Truglio et al, 2006b); domain 2 is mainly responsible for the interaction with UvrA (Truglio et al, 2004) whereas domain 3 is one of the two RecA-like helicase domains translating ATP-hydrolysis into motion against the second helicase domain (domain 1a) (Theis et al, 2000).

The crystal structure of the UvrBΔ4-DNA complex did not lead to additional insights into the DNA damage verification step. However, incision assays in the absence of UvrA indicate that UvrB is able to translocate on DNA preferably in 5' → 3' direction to recognize and verify the presence of a NER substrate. Our data also suggest that UvrB in contrast to UvrA directly probes the DNA and interacts with the damaged nucleotide. The DNA is then specifically incised by the endonuclease UvrC to remove DNA lesions from damaged DNA.

5.2. Characterization of the novel endonuclease Bax1 in context of archaeal nucleotide excision repair

In the first part of this thesis, insights into the DNA damage verification and incision mechanism in prokaryotic NER were discussed. From here on, *in vitro* studies on archaeal proteins are summarized and implications for eukaryotic NER are deduced. Although no eukaryotic homolog for Bax1 has been identified so far, analyses of archaeal NER proteins shed light onto the NER mechanism in general and also with regard to the evolution of DNA repair pathways.

Bax1 was initially identified as a novel binding partner for the archaeal helicase XPB (Richards et al, 2008). Analysis of the genomic context of *S. solfataricus* revealed that one of two adjacent *xpb* genes is closely associated with a gene encoding an unknown protein, later on named Bax1. The protein was expressed recombinantly and used for interaction studies confirming that Bax1 binds to XPB as proposed by the genomic proximity (Richards et al, 2008).

This work provides the first biochemical and biophysical characterization of Bax1. Additionally, the activity of Bax1 was investigated in the presence of XPB to shed light onto Bax1's role in archaeal NER.

5.2.1. Bax1 is a novel, Mg^{2+}-dependent, structure-specific endonuclease

Bax1 was predicted to possess a DUF790 endonuclease-like domain. The DUF790 family, which comprises hypothetical proteins from archaea and cyanobacteria, was identified recently using a combination of diverse bioinformatic tools since the family members share very low sequence similarity with previously characterized nucleases (Kinch et al, 2005). A potential nuclease activity was also indicated by a probable membership ($p = 0.96$) within the PD-(D/E)XK-nuclease superfamily (Laganeckas et al, 2010).

We identified Bax1 as an active Mg^{2+}-dependent endonuclease which acts preferentially as a monomer (see 4.2.7). Incision studies provided initial insights to a possible mechanism how Bax1 incises DNA and how it may play an important role in NER. Bax1 is a structure-specific endonuclease cutting DNA in the

single-stranded region 4 to 6 nts next to the dsDNA/ssDNA junction. Other substrates such as dsDNA, a 5' overhang or a substrate containing mismatches are not recognized. The endonuclease activity of Bax1 requires at least one divalent cation, preferably Mg^{2+}.

5.2.2. Structural insights into the endonuclease Bax1

Although Bax1's endonuclease activity has been characterized by biochemical approaches, no structural information on Bax1 or other DUF790 endonucleases is available. Proteins, of which crystal structures are deposited in the PDB databank, do not share more than 14 % sequence identity. Proteins with less than 20 % sequence identity usually display large structural differences (Chothia & Lesk, 1986). Thus, the resulting structures from diverse homology modeling approaches such as provided by 3D JIGSAW or PHYRE are not reliable (Bates et al, 2001; Kelley & Sternberg, 2009). Moreover, no similarity to other single-strand-specific nucleases (Desai & Shankar, 2003) or structure-specific nucleases (Nishino et al, 2006) was observed although the activity of Bax1 indicates a functional relationship. The functional relationship, however, provides information about the potential active site of Bax1.

The crystal structure of UvrC's C-terminal endonuclease domain from *T. maritima*, visualizes how a metal ion is coordinated by the catalytic triad, consisting of residues D367, D429 and H488, either directly or indirectly through water molecules (Karakas et al, 2007). Similar catalytic triads, i.e. DDH, DDD and DDE, are present in *Pyrococcus furiosus* Argonaute (Song et al, 2004), *Bacillus halodurans* RNase H (Nowotny et al, 2005) and *E. coli* Tn5 transposase (Lovell et al, 2002), respectively. The latter two were shown to bind two divalent cations in their active site. The first metal ion facilitates the formation of a hydroxide ion which in turn performs a nucleophilic attack of the scissile phosphate. The second metal ion fulfills the role of a Lewis acid stabilizing the reaction intermediate via interaction with the 3' oxygen (Karakas et al, 2007; Steitz & Steitz, 1993). UvrC and Argonaute most likely also conduct a two-metal-ion mechanism as suggested by the strong homology in the active site (Karakas et al, 2007; Song et al, 2004; Steitz & Steitz, 1993). Analyzing the crystal structure of UvrC's N-terminal domain

reveals how the residue E76 directly coordinates one Mg^{2+} ion, whereas residues Y29 and N88 indirectly coordinate the metal ion through water molecules (Truglio et al, 2005). In this case, the Mg^{2+} ion acts as a Lewis acid, and it was speculated that Y29 facilitates the formation of a nucleophilic hydroxide catalyzing the cleavage of the phosphodiester bond.

In stark contrast to the structure-specific endonucleases, restriction endonucleases are known to be sequence-specific. Nevertheless, a BLAST search revealed a small patch of 14 amino acids in Bax1, which shares 60 % sequence identity between the type II restriction enzyme *Tsp*RI (Roberts et al, 2003b) and Bax1. The residues within this patch might be part of the putative active site of Bax1 where one or two divalent cations could be coordinated to catalyze the cleavage of a phosphodiester bond. This finding together with a multiple sequence alignment of Bax1 homologs from other archaea uncovered conserved residues which might play an essential role in DNA binding as well as in catalysis. To verify the importance of a residue, point mutants were generated and analyzed in incision assays.

The residues E124, D132, Y152 and E155 are highly conserved among archaea and provide the putative functional groups required for binding metal ions. Mutation of these amino acids to alanine strongly reduced incision efficiency. A reduced incision activity compared to WT Bax1 indicates that these residues are crucial for Bax1 to act effectively on its DNA substrate either resulting from a decreased DNA binding affinity, from the inability to coordinate Mg^{2+} or from a lower capability to stabilize reaction intermediates. We tested the Bax1 variants E124A, D132A, Y152A and N153A/E155A with regard to their DNA binding properties employing bio-layer interferometry. Since the mutants bound DNA in the same affinity range, the residues do not seem to be essential for binding the DNA but for mediating incision. Therefore, it is hypothesized that the above mentioned residues contribute to the active site and could also be involved in coordinating Mg^{2+} ions thus catalyzing the incision reaction.

Interestingly, Bax1 from *S. solfataricus* lacks the residue which corresponds to E124 in *T. acidophilum* Bax1 (**Figure 4.23**). To ensure a functional *S. solfataricus* Bax1, another acidic residue would have to take over the role of E124 in *T. acidophilum* Bax1. Indeed, Malcolm White's group postulated that residue D301

is crucial for incision activity of *S. solfataricus* Bax1 (Rouillon & White, 2010). Interestingly, this residue is located in a patch of conserved residues consistent with the sequence motif for DUF790 domains. This sequence motif, however, is only poorly conserved in the archaeal order of *Thermoplasmatales* suggesting that different active sites might have evolved in *T. acidophilum* and *S. solfataricus* Bax1.

Residues Y128, D130, E134 are also suitable for binding metal ions regarding their functional groups, but mutation to alanine had a milder effect on their incision activity. Thus amino acids Y128, D130, E134 might support coordination of the Mg^{2+} ion or play an important role with respect to the formation of a stable active site but are not essential for catalysis. H488 in the C-terminal domain of UvrC seems to fulfill a similar role (Karakas et al, 2007).

Moreover, we performed limited proteolysis experiments using the protease trypsin, which revealed the formation of a stable fragment with a molecular weight of about 35 kDa (**Figure 4.24**). Mass spectrometry analysis confirmed that the stable fragment is consistent with the C-terminal part of Bax1 comprising amino acids 90 to 403 (referred to $Bax1_{c\text{-}term}$). Incision assays showed that both full length Bax1 and $Bax1_{c\text{-}term}$, incised the DNA substrate and led to an incision product of the same size. Moreover, the activity seemed not to be changed as reflected by the bar chart diagram presented in **Figure 4.24**. The C-terminal domain of Bax1 therefore corresponds to the catalytically active endonuclease domain. These findings are consistent with the localization of the active site residues E124, D132, Y152 and E155 in the $Bax1_{c\text{-}term}$ protein.

Figure 5.8 Domain architecture of Bax1. Bax1 consists of two domains, an N-terminal domain and a C-terminal endonuclease domain.

From these experiments, it was deduced that Bax1 probably consists of two domains connected by a protease-sensitive linker, an N-terminal domain and a catalytically active C-terminal endonuclease domain (**Figure 5.8**). Interaction studies using analytical size-exclusion chromatography hint to a potential

involvement in binding XPB (Grundler, 2010), however, the role of the N-terminal domain has not been fully assessed so far.

5.2.3. XPB regulates Bax1 in the absence of helicase activity

In this work, it was possible to locate Bax1's endonuclease activity in its C-terminal domain and specifically assigned to the active site residues E124, D132, Y152 and E155. Moreover, Bax1 was shown to act as an endonuclease without being supported by any other protein. However, the physical interaction of Bax1 with XPB raises the question how the two proteins affect or regulate each other.

Malcolm White's group postulated that *S. solfataricus* XPB and Bax1 work together as a helicase-nuclease machine (Rouillon & White, 2010). In contrast to that study, *T. acidophilum* Bax1 and XPB were shown to form two different XPB/Bax1 complexes which were identified by size-exclusion chromatography (**Figure 4.25**). Interestingly, only one of the two complexes possesses DNA incision activity (XPB/Bax1$_{comp}$) while the other is completely inactive in terms of DNA incision (XPB/Bax1$_{incomp}$). The two XPB/Bax1 complex forms were found to be indistinguishable in their stoichiometry as well as DNA binding properties, using a series of biochemical and biophysical characterization methods (**Figure 4.30**). Both complex types bind to DNA as a heterodimer with preference for ssDNA overhang structures and similar dissociation constants in the nM-range (**Figure 4.34**). Loss of incision by XPB/Bax1$_{incomp}$ does therefore not originate from a decrease in DNA affinity. However, protease digestion patterns indicated differences in local conformation for the two XPB/Bax1 complexes (**Figure 4.30**). A small conformational change in the protein complex may be responsible for the different incision activities of XPB/Bax1$_{comp}$ and XPB/Bax1$_{incomp}$.

While Bax1 required Mg^{2+} for catalysis, the XPB/Bax1$_{comp}$ complex incised DNA in the presence of $MgCl_2$, $CaCl_2$, $MnCl_2$, or EDTA (**Figure 4.28**). These observations suggest that a pre-bound Mg^{2+}-ion in the incision competent complex XPB/Bax1$_{comp}$ (but not in XPB/Bax1$_{incomp}$) may be responsible for its incision ability. Subsequently, the Mg^{2+}-binding site would be inaccessible to the surrounding solution once the protein complex is formed. Extensive dialysis

against the chelating agent EDTA prior to size-exclusion chromatography favors the formation of the XPB/Bax1$_{incomp}$ complex suggesting that the presence or absence of divalent cations might play a role in XPB/Bax1 complex formation. Further structural analyses of the two different complexes will be necessary to decipher the underlying mechanism which leads to an incision-competent and an incision-incompetent XPB/Bax1 complex. X-ray crystallography as well as NMR spectroscopy are suitable techniques to address these questions. The presence of the XPB/Bax1$_{incomp}$ complex in addition to the XPB/Bax1$_{comp}$ complex may point towards a regulation mechanism for Bax1. This notion is also consistent with the presence of two different XPB forms in the organism *S. solfatarius and S. tokodaii*, only one of which is able to promote DNA incisions in complex with Bax1 (Ma et al, 2011; Richards et al, 2008).

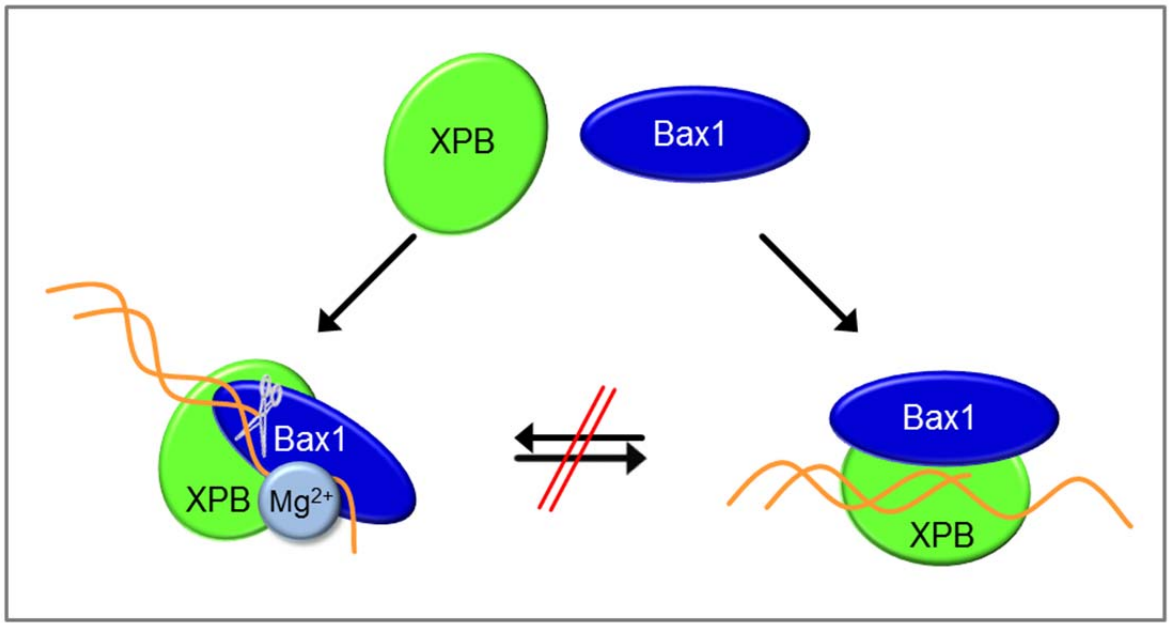

Figure 5.9 XPB and Bax1 are able to form two conformationally different complexes. One conformation promotes the incision activity of Bax1 whereas the other conformation prevents Bax1 from incising the DNA substrate. It is speculated that the XPB/Bax1$_{comp}$ complex might use a pre-bound Mg^{2+} ion for catalysis.

Comparing incision activities of Bax1 and of the XPB/Bax1$_{comp}$ complex, surprising differences were observed: First, Bax1 and the XPB/Bax1$_{comp}$ complex displayed different incision patterns for substrates with 3' overhang ends. Furthermore, while Bax1 alone did not cut 5' overhang substrates, the XPB/Bax1$_{comp}$ complex was able to incise both types of ssDNA overhang

substrates (**Figure 4.26**). These data support a role of XPB in loading and positioning Bax1 on the DNA to increase its substrate range. The DNA binding studies provided further support for a possible loading mechanism of Bax1 by XPB. Specifically, the XPB/Bax1 complexes bound DNA with a lower K_D compared to Bax1 (**Figure 4.34**), consistent with an increased number of DNA bound complexes on circular plasmid DNA for XPB/Bax1 complexes compared to Bax1 in the absence of XPB (**Figure 4.33**). Interestingly, neither DNA dissociation constants nor incision properties of the XPB/Bax1 complexes were affected by the presence or absence of nucleotides such as ATP, the non-hydrolysable ATP-analogue AMPPNP or ADP (**Figure 4.29**). This independency of nucleotide binding and/or hydrolysis suggests that XPB's helicase activity was not essential for loading and positioning Bax1 onto the DNA.

Incubation of 3' and 5' overhang DNA substrates with XPB/Bax1$_{comp}$ resulted in the occurrence of multiple incision sites. Bax1, in contrast, produced one specific incision product with 3' overhangs but did not cut 5' overhang substrates at all. XPB/Bax1$_{comp}$ also was able to cut ssDNA, whereas Bax1 alone did not incise this substrate (**Figure 4.27**). Taken together, these findings support a mechanism of non-specific ssDNA incisions for the XPB/Bax1$_{comp}$ complex.

Lack of specificity of Bax1 within the XPB/Bax1$_{comp}$ complex raises the question, if additional proteins may be required to regulate correct DNA incisions by controlled positioning and orientation of Bax1 on the DNA. Notably, discrimination between damaged and non-damaged DNA by the archaeal NER helicase XPD has been reported (Mathieu et al, 2010). Studies on eukaryotic cells suggest a concerted activity of the ATPase XPB and the helicase XPD to unwind DNA and to verify the presence of an NER substrate (Coin et al, 2007).

5.2.4. Bax1 is loaded and positioned onto DNA by XPB

The interaction of XPB and Bax1 was not only investigated with regard to DNA incision activity but also in terms of DNA binding. In a single-molecule approach, AFM was applied to obtain qualitative information in addition to the quantitative data attained by bio-layer interferometry. Employing AFM imaging, considerable amounts of XPB/Bax1 complexes were not only bound to the single-stranded part

of the DNA substrate, which provided a model target for both Bax1 and XPB proteins (Biswas et al, 2009; Fan et al, 2006; Rouillon & White, 2010), but surprisingly also in the double-stranded DNA regions (**Figure 4.32**). The position distributions of protein complexes on non-damaged DNA substrates revealed no specific, DNA strand-internal target site for the XPB/Bax1 complexes such as more flexible AT-rich regions. These experiments were carried out in the absence of ATP to prevent helicase activity of XPB and to support observation of the complexes in the original DNA-binding state. In the images, more than 50 % of the protein-DNA complexes were bound at DNA strand internal sites. Considering the vast excess of non-specific strand-internal binding sites over binding sites at DNA fragment ends, the data were nevertheless consistent with an approximately 200-fold specificity for DNA fragment ends. Preferential location of protein complexes at DNA fragment ends can be interpreted as a consequence of either of two scenarios: Firstly, the protein complexes preferentially bind to DNA ends which might also locally melt to appear partially single-stranded or secondly, the fragment ends provide “road blocks” to a sliding movement of the complexes on the DNA.

To distinguish between preferential DNA end binding and a blocking of sliding protein complexes at DNA fragment ends, AFM was used to visualize complexes of Bax1 and XPB/Bax1 with relaxed circular plasmid DNA (B-form DNA containing no ends). The presence of more XPB/Bax1 complexes per DNA base pair (bp) for the circular DNA compared to the linear DNA fragments (5.1 ± 0.5 complexes per DNA molecule for the circular plasmid, 2729 bp, versus 0.26 ± 0.15 complexes per linear DNA fragment, 500 bp) argued against a preferential binding to fragment ends and supported a DNA sliding mechanism for XPB/Bax1. Recently, evidence has amounted that sliding could be a common mechanism for DNA repair proteins to locate their target sites (Friedman & Stivers, 2010; Kad et al, 2010; Szczelkun et al, 2010). Sliding leads to a reduction of a three-dimensional to a one-dimensional target site search on the DNA substrate, therefore speeding up the search process. Since these experiments were carried out in the absence of ATP, it was concluded that, if sliding was employed here, then it was independent of ATP binding and/or hydrolysis by XPB.

In summary, the direct comparison of Bax1 with the XPB/Bax1$_{comp}$ complex from *T. acidophilum* suggests that XPB, even in the absence of helicase activity, is able to load and to position Bax1 onto the scissile DNA substrate.

5.2.5. Implications for other endonucleases in archaeal NER

In this study, the novel endonuclease Bax1 has been characterized. The physical and functional interaction with XPB suggests a role for Bax1 in archaeal NER. However, the question arises why another nuclease, Bax1, is present in the archaeal organism *T. acidophilum*. In eukaryotes, the presence of XPF and XPG, two endonucleases of opposite polarity, ensures the complete removal of the damaged oligonucleotide (Araujo et al, 2000).The archaeal homologs of XPG and XPF were confirmed to act as structure-specific endonucleases as well (Chapados et al, 2004; de Laat et al, 1998; Harrington & Lieber, 1994). In contrast to Bax1, which cuts DNA in the single-stranded region, both XPG and XPF incise DNA in the double-stranded region of a DNA substrate close to the dsDNA/ssDNA junction. XPG has been shown to incise 3' to a damaged site and remove a 5' overhang (O'Donovan et al, 1994), whereas XPF-ERCC1 makes the 5' incision and also cleaves 3' overhangs (de Laat et al, 1998). Interestingly, *T. acidophilum* seems to lack an XPF homolog in contrast to many other archaeal organisms (Kelman & White, 2005). Moreover, genome analysis of *T. acidophilum* did not uncover an additional NER mechanism, for example the UvrABC system, which is present not only in prokaryotes, but also in several archaea (Ogrunc et al, 1998). However, a database search revealed that *T. acidophilum* contains a Hef protein (helicase-associated endonuclease for fork-structured DNA). Hef usually consists of two domains, an N-terminal helicase domain and a C-terminal nuclease domain, which exhibits high similarity to the XPF/Rad1/Mus81 nuclease family (Nishino et al, 2003) with respect to their sequences and to their incision properties. Intriguingly, *T. acidophilum* Hef comprises only the helicase domain but no nuclease domain (Uniprot entry: Q9HI46). Thus, Bax1 would be a suitable candidate to replace the missing XPF protein in *T. acidophilum* as this endonuclease is able to remove 3' overhangs and thus could mediate the 5' incision (**Figure 5.10**).

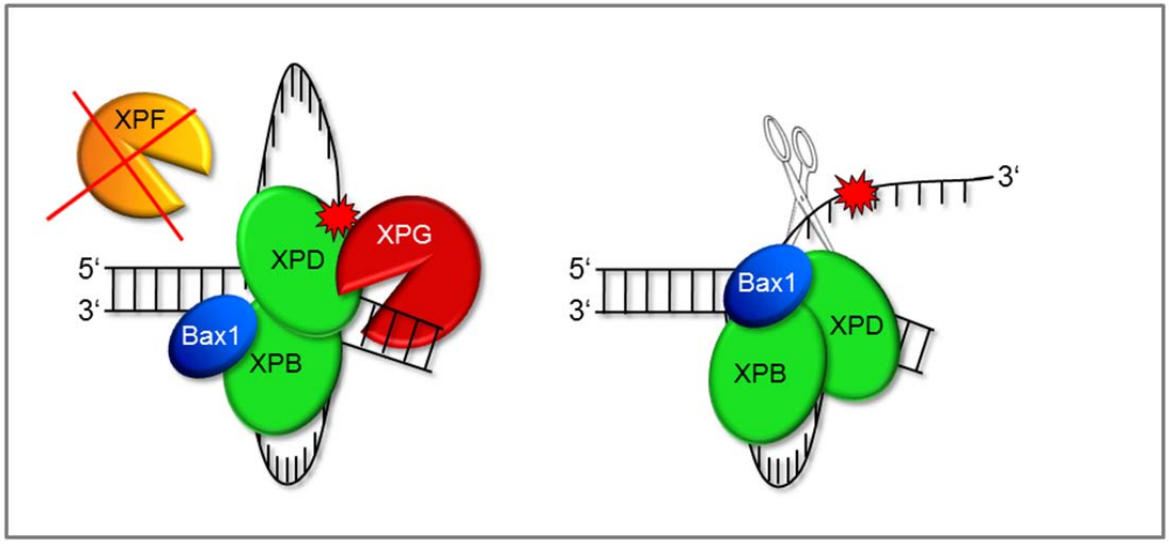

Figure 5.10 Proposed mechanism for Bax1 incision in archaeal NER. Due to their close genomic proximity it is likely that XPB and Bax1 are transcribed and translated concomitantly (Richards et al, 2008). XPB and Bax1 are proposed to form a complex which subsequently binds to a DNA substrate. Unwinding of the DNA would be accomplished by the concerted activity of XPB and the 5’ - 3’ helicase XPD. XPD was suggested to perform the DNA damage verification step (Coin et al, 2007), although recent studies on archaeal XPD controversially discuss whether XPD is stalled at a DNA damage (Mathieu et al, 2010; Rudolf et al, 2010). However, after the presence of an NER substrate has been confirmed, the damage has to be removed by a structure-specific endonuclease, which specifically recognizes the 3’ overhang produced by XPG’s endonuclease activity. The incision properties such as the polarity are suggestive for the endonuclease Bax1 to be a suitable candidate to perform the second incision reaction instead of the missing XPF in *T. acidophilum*.

However, in consideration of several other archaea, the question arises why in addition to XPF another endonuclease seems to be involved in archaeal NER as Bax1 is associated to the helicase XPB. This situation resembles the prokaryotic NER system in *E. coli*, which contains the UvrC homolog Cho in addition to UvrC performing both 3’ and 5’ incisions (Moolenaar et al, 2002). Cho conducts the 3’ incision at DNA lesions which are poorly incised by UvrC and thus leads to a broader substrate range for NER (Moolenaar et al, 2002). A similar mechanism could have evolved also in archaea employing Bax1 as a complementary endonuclease.

To scrutinize the role of Bax1 in archaeal NER, *in vivo* studies employing *bax1*-knockout strains would be necessary. Genetic tools as well as expertise and equipment for growing archaea in laboratories were not available to perform these studies in this present work. However, *in vivo* studies would provide detailed

insights into archaeal DNA repair mechanisms complementing many studies on recombinant NER homologs (Fan et al, 2006; Fan et al, 2008; Mathieu et al, 2010; Nishino et al, 2003; Richards et al, 2008; Rouillon & White, 2010; Wolski et al, 2008).

5.2.6. Nucleotide Excision Repair among diverse archaeal species – discrepancies between *T. acidophilum* and *S. solfataricus* Bax1

In vivo studies would also be required to shed light onto the genetic diversity among archaea which is reflected by the evolutionary subdivision into two separate groups within the archaeal kingdom: the euryarchaea and the crenarchaea (Cox et al, 2008). The point of separation in evolution is still under discussion (Cox et al, 2008). Focusing on DNA repair, it is interesting to note that the presence and number of different NER homologs seems to be almost randomly distributed among different archaeal species (Grogan, 2004; Kelman & White, 2005). Some archaea also contain the prokaryotic UvrABC proteins in addition to eukaryotic NER homologs probably introduced by lateral gene transfer (Grogan, 2004; Kelman & White, 2005).

Another important point contributing to the archaeal diversity is the DNA superstructure: in all three domains of life, eukaryotes, prokaryotes and archaea, mechanisms have evolved to pack and compact DNA (White & Bell, 2002). Hyperthermophilic archaea, such as *S. solfataricus*, contain a reverse gyrase which positively supercoils DNA to protect DNA from denaturation at high growth temperatures (Collin, 1988; Lopez-Garcia & Forterre, 1997; Musgrave et al, 1991; Perugino et al, 2009). The genome of *T. acidophilum*, however, comprises one circular negatively supercoiled chromosome (Collin, 1988; Kawashima et al, 2000). Using AFM, we were able to visualize *T. acidophilum* XPB/Bax1 binding to negatively supercoiled plasmid DNA, mimicking physiological conditions. We speculate that DNA binding proteins may have adapted to the superstructure of the DNA present in the organism. The different sign of supercoiling in *T. acidophilum* and *S. solfataricus* (negative and positive, respectively) is only one example underlining the evolutionary divergence between the two organisms. Moreover, XPB proteins from the two organisms share in total 39 % sequence

identity. In comparison, the helicase core domains of *T. acidophilum* and *S. solfataricus* share 26 % and 27 % sequence identity to human XPB, respectively (Biswas et al, 2009; Huang, 1991). *T. acidophilum* and *S. solfataricus* Bax1 are only 22 % identical in sequence. Thus, the diversity within the archaeal kingdom has to be considered when comparing the activity of archaeal proteins from different species.

To decipher DNA repair mechanisms in a specific organism, *in vivo* studies would be necessary. To date, only very few *in vivo* studies on archaeal organisms are published (Crowley et al, 2006; Fujikane et al, 2010; Salerno et al, 2003), in contrast to several studies on recombinant archaeal NER homologs (Fan et al, 2006; Fan et al, 2008; Mathieu et al, 2010; Nishino et al, 2003; Richards et al, 2008; Rouillon & White, 2010; Wolski et al, 2008). It is even not known if the NER homologs in archaea are involved in NER at all or if a different DNA repair mechanism has evolved. So far, NER homologs have been identified through sequence conservation to their human counterparts but it has yet not been shown that these proteins exhibit a role in DNA repair. This reflects the urgent need for these studies to gain insights into archaeal NER especially considering the huge genetic diversity among archaea.

The presence of the prokaryotic UvrABC system as well as photolyases in some archaea might suggest that these organisms do not employ the eukaryotic XP homologs to remove UV-lesions. This has been supported by a study revealing that a *uvrabc* knock-out strain responded more sensitively to UV-light than the wildtype archaeon (Crowley et al, 2006). This finding, however, raises the question how archaea lacking the prokaryotic UvrABC system remove UV-lesions from their genome. The hypothesis that archaea could have established unique DNA repair mechanisms different from prokaryotic and eukaryotic pathways was discussed previously (Grogan, 2004).

A recent publication presents the knock-out of *xpb*, *xpd* and *hef* genes in *T. kodakaraensis* and concludes that XPB and XPD seem not to be involved in NER in this organism whereas Hef is proposed to be involved in several different DNA repair pathways (Fujikane et al, 2010). These results were obtained from sensitivity tests of the knock-out strains towards various damaging agents such as UV-light and mitomycin C which are responsible for the formation of pyrimidine

dimers and interstrand crosslinks, respectively. However, considering that even the wildtype strain of *T. kodakaraensis* displays a higher sensitivity towards UV-light than other archaea and prokaryotes, the presented study might not be directly transferable to other archaea (Fujikane et al, 2010).

Despite the fragmentary knowledge of archaeal DNA repair, archaeal proteins are often used to gain insights into their eukaryotic homologs (Fan et al, 2006; Fan et al, 2008; Liu et al, 2008; Mathieu et al, 2010; Nishino et al, 2003; Richards et al, 2008; Wolski et al, 2008). The conserved sequence of archaeal and human proteins proposes a similar 3D-structure and thus comparable properties with regard to their function (Aravind et al, 1999; Grogan, 2000; Kelman & White, 2005; Ogrunc et al, 1998; White, 2003). The structure of archaeal XPD, for example, could actually explain in part the phenotype of *xpd*-patients suffering from XP, XP/CS or TTD (Fan et al, 2008; Lehmann, 2008; Liu et al, 2008; Wolski et al, 2008). Moreover, it has been demonstrated that archaeal XPD is able to distinguish between damaged and non-damaged DNA in *in vitro* assays (Mathieu et al, 2010). Thus, archaeal homologs nevertheless seem to be a valid model, at least to a certain extent, to gain insights of their human counterparts.

5.3. Final conclusion

The present studies provided insights into the DNA damage recognition and incision reaction in NER obtained by structural and biochemical approaches on recombinant prokaryotic and archaeal proteins.

In the first part of this thesis, the prokaryotic mechanisms of DNA damage verification and subsequent incision reactions were addressed. Although the co-crystal structure of UvrBΔ4 and damaged DNA could not decipher the damage verification process, incision assays in the absence of UvrA proposed that UvrB translocates on DNA predominantly in 5' → 3' direction. Encountering a potential DNA damage, UvrB verifies the NER lesion to subsequently recruit the endonuclease UvrC. UvrC then specifically incises the DNA to remove the damaged oligonucleotide.

The second part of this thesis focused on archaeal NER, especially on the characterization of the novel endonuclease Bax1. Incision studies provided initial

insights to a possible mechanism how Bax1 incises DNA and how it may play an important role in NER. Bax1 was shown to act as a Mg^{2+} dependent, structure specific endonuclease incising DNA in the single-stranded region 4 to 6 nts next to the dsDNA/ssDNA junction. Other substrates such as dsDNA, a 5' overhang or a substrate containing mismatches are not incised.

The presence of two conformationally different XPB/Bax1 complexes with diverse incision activity points towards a regulation mechanism for Bax1. The direct comparison of Bax1 with the XPB/$Bax1_{comp}$ complex from *T. acidophilum* suggests that XPB, even in the absence of helicase activity, is able to load and to position Bax1 onto the scissile DNA substrate. Thus, XPB might hand over the DNA to Bax1 within the XPB/Bax1 complex and thereby adopt a different conformation to enable Bax1 to incise the DNA. This proposed mechanism would ensure that the sequential process of damage recognition, followed by verification, is completed prior to incision. Therefore, XPB and Bax1 would not only interact physically but also functionally.

Insights into archaeal NER provided by the *in vitro* studies presented here require further complemented by continuative *in vivo* studies in order to understand the underlying mechanisms of DNA repair.

Appendix

Literature

Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58:** 1948-1954

Adams PD, Pannu NS, Read RJ, Brunger AT (1997) Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. *Proceedings of the National Academy of Sciences of the United States of America* **94:** 5018-5023

Ahn B, Grossman L (1996) The binding of UvrAB proteins to bubble and loop regions in duplex DNA. *The Journal of biological chemistry* **271:** 21462-21470

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

Aravind L, Walker DR, Koonin EV (1999) Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic acids research* **27:** 1223-1242

Bates PA, Kelley LA, MacCallum RM, Sternberg MJ (2001) Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins* **Suppl 5:** 39-46

Bergmann E, Egly JM (2001) Trichothiodystrophy, a transcription syndrome. *Trends Genet* **17:** 279-286

Bernardes de Jesus BM, Bjoras M, Coin F, Egly JM (2008) Dissection of the molecular defects caused by pathogenic mutations in the DNA repair factor XPC. *Mol Cell Biol* **28:** 7225-7235

Bienstock RJ, Skorvaga M, Mandavilli BS, Van Houten B (2003) Structural and functional characterization of the human DNA repair helicase XPD by comparative molecular modeling and site-directed mutagenesis of the bacterial repair protein UvrB. *The Journal of biological chemistry* **278:** 5309-5316

Biswas T, Pero JM, Joseph CG, Tsodikov OV (2009) DNA-dependent ATPase activity of bacterial XPB helicases. *Biochemistry* **48:** 2839-2848

Bohr VA, Smith CA, Okumoto DS, Hanawalt PC (1985) DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. *Cell* **40:** 359-369

Brueckner F, Hennecke U, Carell T, Cramer P (2007) CPD damage recognition by transcribing RNA polymerase II. *Science (New York, NY* **315:** 859-862

Brunger AT (1997) Free R value: cross-validation in crystallography. *Methods Enzymol* **277:** 366-396

Cantor CR, Warshaw MM, Shapiro H (1970) Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers* **9:** 1059-1077

Chandrasekhar D, Van Houten B (2000) In vivo formation and repair of cyclobutane pyrimidine dimers and 6-4 photoproducts measured at the gene and nucleotide level in Escherichia coli. *Mutat Res* **450:** 19-40

Chang WH, Kornberg RD (2000) Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH. *Cell* **102:** 609-613

Chapados BR, Hosfield DJ, Han S, Qiu J, Yelent B, Shen B, Tainer JA (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell* **116:** 39-50

Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66:** 12-21

Chong S, Mersha FB, Comb DG, Scott ME, Landry D, Vence LM, Perler FB, Benner J, Kucera RB, Hirvonen CA, Pelletier JJ, Paulus H, Xu MQ (1997) Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **192:** 271-281

Chothia C, Lesk AM (1986) The relation between the divergence of sequence and structure in proteins. *The EMBO journal* **5:** 823-826

Clement FC, Camenisch U, Fei J, Kaczmarek N, Mathieu N, Naegeli H (2010) Dynamic two-stage mechanism of versatile DNA damage recognition by xeroderma pigmentosum group C protein. *Mutat Res* **685:** 21-28

Coin F, Auriol J, Tapias A, Clivio P, Vermeulen W, Egly JM (2004) Phosphorylation of XPB helicase regulates TFIIH nucleotide excision repair activity. *The EMBO journal* **23:** 4835-4846

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

Coin F, Proietti De Santis L, Nardo T, Zlobinskaya O, Stefanini M, Egly JM (2006) p8/TTD-A as a repair-specific TFIIH subunit. *Molecular cell* **21:** 215-226

Collin RG, Morgan, H. W., Musgrave, D.R., Daniel, R.M. (1988) Distribution of reverse gyrase in representative species of eubacteria and archaebacteria. *FEMS Microbiology Letters* **55:** 235-240

Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM (2008) The archaebacterial origin of eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 20356-20361

Croteau DL, DellaVecchia MJ, Perera L, Van Houten B (2008) Cooperative damage recognition by UvrA and UvrB: identification of UvrA residues that mediate DNA binding. *DNA Repair (Amst)* **7:** 392-404

Croteau DL, DellaVecchia MJ, Wang H, Bienstock RJ, Melton MA, Van Houten B (2006) The C-terminal zinc finger of UvrA does not bind DNA directly but regulates damage-specific DNA binding. *The Journal of biological chemistry* **281:** 26370-26381

Crowley DJ, Boubriak I, Berquist BR, Clark M, Richard E, Sullivan L, DasSarma S, McCready S (2006) The uvrA, uvrB and uvrC genes are required for repair of ultraviolet light induced DNA photoproducts in Halobacterium sp. NRC-1. *Saline Systems* **2:** 11

de Boer J, Hoeijmakers JH (2000) Nucleotide excision repair and human syndromes. *Carcinogenesis* **21:** 453-460

de Gruijl FR (1999) Skin cancer and solar UV radiation. *Eur J Cancer* **35:** 2003-2009

de Laat WL, Appeldoorn E, Jaspers NG, Hoeijmakers JH (1998) DNA structural elements required for ERCC1-XPF endonuclease activity. *The Journal of biological chemistry* **273:** 7835-7842

DeLano W (2003) The PyMOL Molecular Graphics System.

DellaVecchia MJ, Croteau DL, Skorvaga M, Dezhurov SV, Lavrik OI, Van Houten B (2004) Analyzing the handoff of DNA from UvrA to UvrB utilizing DNA-protein photoaffinity labeling. *The Journal of biological chemistry* **279:** 45245-45256

Desai NA, Shankar V (2003) Single-strand-specific nucleases. *FEMS microbiology reviews* **26:** 457-491

Diederichs K, Karplus PA (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nature structural biology* **4:** 269-275

Diffey B (2004) Climate change, ozone depletion and the impact on ultraviolet exposure of human skin. *Phys Med Biol* **49:** R1-11

Dupureur CM (2008) Roles of metal ions in nucleases. *Curr Opin Chem Biol* **12:** 250-255

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126-2132

Eryilmaz J, Ceschini S, Ryan J, Geddes S, Waters TR, Barrett TE (2006) Structural insights into the cryptic DNA-dependent ATPase activity of UvrB. *Journal of molecular biology* **357:** 62-72

Evans E, Moggs JG, Hwang JR, Egly JM, Wood RD (1997) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *The EMBO journal* **16:** 6559-6573

Evans PR (2006) Scaling & assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62:** 82-82

Fan L, Arvai AS, Cooper PK, Iwai S, Hanaoka F, Tainer JA (2006) Conserved XPB Core Structure and Motifs for DNA Unwinding: Implications for Pathway Selection of Transcription or Excision Repair. *Molecular cell* **22:** 27-37

Fan L, Fuss JO, Cheng QJ, Arvai AS, Hammel M, Roberts VA, Cooper PK, Tainer JA (2008) XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell* **133:** 789-800

Fasman GD (ed) (1975) *Handbook of Biochemistry and Molecular Biology*: CRC Press, 589-pp

Fontana A, de Laureto PP, Spolaore B, Frare E, Picotti P, Zambonin M (2004) Probing protein structure by limited proteolysis. *Acta Biochim Pol* **51:** 299-321

Fotiadis D, Scheuring S, Muller SA, Engel A, Muller DJ (2002) Imaging and manipulation of biological structures with the AFM. *Micron* **33:** 385-397

Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T (2006) *DNA Repair and Mutagenesis*, 2nd edn. Washington, DC: ASM Press.

Friedman JI, Stivers JT (2010) Detection of damaged DNA bases by DNA glycosylase enzymes. *Biochemistry* **49:** 4957-4967

Frols S, Gordon PM, Panlilio MA, Duggin IG, Bell SD, Sensen CW, Schleper C (2007) Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV damage. *Journal of bacteriology* **189:** 8708-8718

Fujikane R, Ishino S, Ishino Y, Forterre P (2010) Genetic analysis of DNA repair in the hyperthermophilic archaeon, Thermococcus kodakaraensis. *Genes & genetic systems* **85:** 243-257

Gary R, Kim K, Cornelius HL, Park MS, Matsumoto Y (1999) Proliferating cell nuclear antigen facilitates excision in long-patch base excision repair. *The Journal of biological chemistry* **274:** 4354-4363

Geacintov NE, Broyde S, Buterin T, Naegeli H, Wu M, Yan S, Patel DJ (2002) Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers* **65:** 202-210

Giglia-Mari G, Coin F, Ranish JA, Hoogstraten D, Theil A, Wijgers N, Jaspers NG, Raams A, Argentini M, van der Spek PJ, Botta E, Stefanini M, Egly JM, Aebersold R, Hoeijmakers JH, Vermeulen W (2004) A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat Genet* **36:** 714-719

Gillet LC, Scharer OD (2006) Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chemical reviews* **106:** 253-276

Goosen N (2010) Scanning the DNA for damage by the nucleotide excision repair machinery. *DNA Repair (Amst)* **9:** 593-596

Goosen N, Moolenaar GF (2001) Role of ATP hydrolysis by UvrA and UvrB during nucleotide excision repair. *Res Microbiol* **152:** 401-409

Gotz D, Paytubi S, Munro S, Lundgren M, Bernander R, White MF (2007) Responses of hyperthermophilic crenarchaea to UV irradiation. *Genome Biol* **8:** R220

Grogan DW (2000) The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol* **8:** 180-185

Grogan DW (2004) Stability and repair of DNA in hyperthermophilic Archaea. *Curr Issues Mol Biol* **6:** 137-144

Grosse-Kunstleve RW, Sauter NK, Moriarty NW, Adams PD (2002) TheComputational Crystallography Toolbox: crystallographic algorithms in are usable software framework. *J Appl Crystallogr* **35:** 126-136

Grundler V (2010) Interaktionen der Nukleotid-Exzisions DNA-Reparatur-Proteine XPB und Bax1 und Kristallisation von IGHMBP2. Diploma Thesis, Rudolf-Virchow-Zentrum für experimentelle Biomedizin, Strukturbiologie, Würzburg

Harrington JJ, Lieber MR (1994) The characterization of a mammalian DNA structure-specific endonuclease. *The EMBO journal* **13:** 1235-1246

Hecht SS (1999) Tobacco smoke carcinogens and lung cancer. *J Natl Cancer Inst* **91:** 1194-1210

Heiring C, Muller YA (2001) Folding screening assayed by proteolysis: application to various cystine deletion mutants of vascular endothelial growth factor. *Protein Eng* **14:** 183-188

Hirano Y, Takahashi H, Kumeta M, Hizume K, Hirai Y, Otsuka S, Yoshimura SH, Takeyasu K (2008) Nuclear architecture and chromatin dynamics revealed by atomic force microscopy in combination with biochemistry and cell biology. *Pflugers Arch* **456:** 139-153

Hoare S, Zou Y, Purohit V, Krishnasamy R, Skorvaga M, Van Houten B, Geacintov NE, Basu AK (2000) Differential incision of bulky carcinogen-DNA adducts by the UvrABC nuclease: comparison of incision rates and the interactions of Uvr subunits with lesions of different structures. *Biochemistry* **39:** 12252-12261

Honda M, Park J, Pugh RA, Ha T, Spies M (2009) Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase. *Molecular cell* **35:** 694-703

Hsu DS, Kim ST, Sun Q, Sancar A (1995) Structure and function of the UvrB protein. *The Journal of biological chemistry* **270:** 8319-8327

Huang X, Miller, M. (1991) A time-efficient, linear-space local similarity algorithm. *Adv Appl Math* **12:** 337-357

Hutton RD, Roberts JA, Penedo JC, White MF (2008) PCNA stimulates catalysis by structure-specific nucleases using two distinct mechanisms: substrate targeting and catalytic step. *Nucleic acids research* **36:** 6720-6727

Hwang JR, Moncollin V, Vermeulen W, Seroz T, van Vuuren H, Hoeijmakers JH, Egly JM (1996) A 3' --> 5' XPB helicase defect in repair/transcription factor TFIIH of xeroderma pigmentosum group B affects both DNA repair and transcription. *The Journal of biological chemistry* **271:** 15898-15904

Jaciuk M, Nowak E, Skowronek K, Tanska A, Nowotny M (2011) Structure of UvrA nucleotide excision repair protein in complex with modified DNA. *Nat Struct Mol Biol* **18:** 191-197

Jia L, Kropachev K, Ding S, Van Houten B, Geacintov NE, Broyde S (2009) Exploring damage recognition models in prokaryotic nucleotide excision repair with a benzo[a]pyrene-derived lesion in UvrB. *Biochemistry* **48:** 8948-8957

Kabsch W (1988) *J Appl Cryst* **21:** 916-924

Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22:** 2577-2637

Kad NM, Wang H, Kennedy GG, Warshaw DM, Van Houten B (2010) Collaborative dynamic DNA scanning by nucleotide excision repair proteins investigated by single-molecule imaging of quantum-dot-labeled proteins. *Molecular cell* **37:** 702-713

Karakas E, Truglio JJ, Croteau D, Rhau B, Wang L, Van Houten B, Kisker C (2007) Structure of the C-terminal half of UvrC reveals an RNase H endonuclease domain with an Argonaute-like catalytic triad. *The EMBO journal* **26:** 613-622

Kawashima T, Amano N, Koike H, Makino S, Higuchi S, Kawashima-Ohya Y, Watanabe K, Yamazaki M, Kanehori K, Kawamoto T, Nunoshiba T, Yamamoto Y, Aramaki H, Makino K, Suzuki M (2000) Archaeal adaptation to higher temperatures revealed by genomic sequence of Thermoplasma volcanium. *Proc Natl Acad Sci U S A* **97:** 14257-14262

Kelley LA, Sternberg MJ (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4:** 363-371

Kelly SM, Jess TJ, Price NC (2005) How to study proteins by circular dichroism. *Biochim Biophys Acta* **1751:** 119-139

Kelman Z, White MF (2005) Archaeal DNA replication and repair. *Current opinion in microbiology* **8:** 669-676

Kim JL, Morgenstern KA, Griffith JP, Dwyer MD, Thomson JA, Murcko MA, Lin C, Caron PR (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* **6:** 89-100

Kinch LN, Ginalski K, Rychlewski L, Grishin NV (2005) Identification of novel restriction endonuclease-like fold families among hypothetical proteins. *Nucleic acids research* **33:** 3598-3605

Kleywegt GJ, Jones TA (1997) Model building and refinement practice. *Methods Enzymol* **277:** 208-230

Krissinel E, Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60:** 2256-2268

Kulaksiz G, Reardon JT, Sancar A (2005) Xeroderma pigmentosum complementation group E protein (XPE/DDB2): purification of various complexes of XPE and analyses of their damaged DNA binding and putative DNA repair properties. *Mol Cell Biol* **25:** 9784-9792

Laganeckas M, Margelevicius M, Venclovas C (2010) Identification of new homologs of PD-(D/E)XK nucleases by support vector machines trained on data derived from profile-profile alignments. *Nucleic acids research*

Laine JP, Egly JM (2006a) Initiation of DNA repair mediated by a stalled RNA polymerase IIO. *The EMBO journal* **25:** 387-397

Laine JP, Egly JM (2006b) When transcription and repair meet: a complex system. *Trends Genet* **22:** 430-436

Laine JP, Mocquet V, Egly JM (2006) TFIIH enzymatic activities in transcription and nucleotide excision repair. *Methods Enzymol* **408:** 246-263

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26:** 283-291

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8:** 477-486

Le May N, Egly JM, Coin F (2010) True lies: the double life of the nucleotide excision repair factors in transcription and DNA repair. *J Nucleic Acids* **2010**

Lehmann AR (2008) XPD structure reveals its secrets. *DNA Repair (Amst)* **7:** 1912-1915

Leslie AGW (1990) *Crystallographic Computing*: Oxford University Press.

Lin JJ, Phillips AM, Hearst JE, Sancar A (1992) Active site of (A)BC excinuclease. II. Binding, bending, and catalysis mutants of UvrB reveal a direct role in 3' and an indirect role in 5' incision. *The Journal of biological chemistry* **267:** 17693-17700

Lin JJ, Sancar A (1992) Active site of (A)BC excinuclease. I. Evidence for 5' incision by UvrC through a catalytic site involving Asp399, Asp438, Asp466, and His538 residues. *The Journal of biological chemistry* **267:** 17688-17692

Lindahl T (2000) Suppression of spontaneous mutagenesis in human cells by DNA base excision-repair. *Mutat Res* **462:** 129-135

Liu H, Rudolf J, Johnson KA, McMahon SA, Oke M, Carter L, McRobbie AM, Brown SE, Naismith JH, White MF (2008) Structure of the DNA repair helicase XPD. *Cell* **133:** 801-812

Lopez-Garcia P, Forterre P (1997) DNA topology in hyperthermophilic archaea: reference states and their variation with growth phase, growth temperature, and temperature stresses. *Mol Microbiol* **23:** 1267-1279

Lovell S, Goryshin IY, Reznikoff WR, Rayment I (2002) Two-metal active site binding of a Tn5 transposase synaptic complex. *Nature structural biology* **9:** 278-281

Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50:** 437-450

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

Madhani HD, Bohr VA, Hanawalt PC (1986) Differential DNA repair in transcriptionally active and inactive proto-oncogenes: c-abl and c-mos. *Cell* **45:** 417-423

Maeda T, Chua PP, Chong MT, Sim AB, Nikaido O, Tron VA (2001) Nucleotide excision repair genes are upregulated by low-dose artificial ultraviolet B: evidence of a photoprotective SOS response? *J Invest Dermatol* **117:** 1490-1497

Maillard O, Camenisch U, Blagoev KB, Naegeli H (2008) Versatile protection from mutagenic DNA lesions conferred by bipartite recognition in nucleotide excision repair. *Mutat Res* **658:** 271-286

Malta E, Moolenaar GF, Goosen N (2006) Base flipping in nucleotide excision repair. *The Journal of biological chemistry* **281:** 2184-2194

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

Matthews BW (1968) Solvent content of protein crystals. *Journal of molecular biology* **33:** 491-497

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40:** 658-674

Min JH, Pavletich NP (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* **449:** 570-575

Mishra NC (2002) *Nucleases: Molecular Biology and Applications*: Wiley.

Mocquet V, Laine JP, Riedl T, Yajin Z, Lee MY, Egly JM (2008) Sequential recruitment of the repair factors during NER: the role of XPG in initiating the resynthesis step. *The EMBO journal* **27:** 155-167

Moolenaar GF, Herron MF, Monaco V, van der Marel GA, van Boom JH, Visse R, Goosen N (2000a) The role of ATP binding and hydrolysis by UvrB during nucleotide excision repair. *The Journal of biological chemistry* **275:** 8044-8050

Moolenaar GF, Hoglund L, Goosen N (2001) Clue to damage recognition by UvrB: residues in the beta-hairpin structure prevent binding to non-damaged DNA. *The EMBO journal* **20:** 6140-6149

Moolenaar GF, Monaco V, van der Marel GA, van Boom JH, Visse R, Goosen N (2000b) The effect of the DNA flanking the lesion on formation of the UvrB-DNA preincision complex. Mechanism for the UvrA-mediated loading of UvrB onto a DNA damaged site. *The Journal of biological chemistry* **275:** 8038-8043

Moolenaar GF, van Rossum-Fikkert S, van Kesteren M, Goosen N (2002) Cho, a second endonuclease involved in Escherichia coli nucleotide excision repair. *Proceedings of the National Academy of Sciences of the United States of America* **99:** 1467-1472

Murphy RM (1997) Static and dynamic light scattering of biological macromolecules: what can we learn? *Curr Opin Biotechnol* **8:** 25-30

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53:** 240-255

Musgrave DR, Sandman KM, Reeve JN (1991) DNA binding by the archaeal histone HMf results in positive supercoiling. *Proc Natl Acad Sci U S A* **88:** 10397-10401

Naegeli H, Bardwell L, Friedberg EC (1992) The DNA helicase and adenosine triphosphatase activities of yeast Rad3 protein are inhibited by DNA damage. A potential mechanism for damage-specific recognition. *The Journal of biological chemistry* **267:** 392-398

Newman M, Murray-Rust J, Lally J, Rudolf J, Fadden A, Knowles PP, White MF, McDonald NQ (2005) Structure of an XPF endonuclease with and without DNA suggests a model for substrate recognition. *The EMBO journal* **24:** 895-905

Nishino T, Ishino Y, Morikawa K (2006) Structure-specific DNA nucleases: structural basis for 3D-scissors. *Current opinion in structural biology* **16:** 60-67

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:** 445-457

Nowotny M, Gaidamakov SA, Crouch RJ, Yang W (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* **121:** 1005-1016

O'Donovan A, Davies AA, Moggs JG, West SC, Wood RD (1994) XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371:** 432-435

Ogrunc M, Becker DF, Ragsdale SW, Sancar A (1998) Nucleotide excision repair in the third kingdom. *Journal of bacteriology* **180:** 5796-5798

Oh EY, Grossman L (1987) Helicase properties of the Escherichia coli UvrAB protein complex. *Proceedings of the National Academy of Sciences of the United States of America* **84:** 3638-3642

Oh EY, Grossman L (1989) Characterization of the helicase activity of the Escherichia coli UvrAB protein complex. *The Journal of biological chemistry* **264:** 1336-1343

Oh KS, Imoto K, Boyle J, Khan SG, Kraemer KH (2007) Influence of XPB helicase on recruitment and redistribution of nucleotide excision repair proteins at sites of UV-induced DNA damage. *DNA Repair (Amst)* **6:** 1359-1370

Oksenych V, Coin F (2010) The long unwinding road: XPB and XPD helicases in damaged DNA opening. *Cell Cycle* **9:** 90-96

Oksenych V, de Jesus BB, Zhovmer A, Egly JM, Coin F (2009) Molecular insights into the recruitment of TFIIH to sites of DNA damage. *The EMBO journal* **28:** 2971-2980

Orren DK, Sancar A (1989) The (A)BC excinuclease of Escherichia coli has only the UvrB and UvrC subunits in the incision complex. *Proceedings of the National Academy of Sciences of the United States of America* **86:** 5237-5241

Orren DK, Sancar A (1990) Formation and enzymatic properties of the UvrB.DNA complex. *The Journal of biological chemistry* **265:** 15796-15803

Painter J, Merritt EA (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr* **62:** 439-450

Pakotiprapha D, Inuzuka Y, Bowman BR, Moolenaar GF, Goosen N, Jeruzalmi D, Verdine GL (2008) Crystal structure of Bacillus stearothermophilus UvrA provides insight into ATP-modulated dimerization, UvrB interaction, and DNA binding. *Molecular cell* **29:** 122-133

Pakotiprapha D, Liu Y, Verdine GL, Jeruzalmi D (2009) A structural model for the damage-sensing complex in bacterial nucleotide excision repair. *The Journal of biological chemistry* **284:** 12837-12844

Panuska JR, Goldthwait DA (1980) A DNA-dependent ATPase from T4-infected Escherichia coli. Purification and properties of a 63,000-dalton enzyme and its conversion to a 22,000-dalton form. *The Journal of biological chemistry* **255:** 5208-5214

Perugino G, Valenti A, D'Amaro A, Rossi M, Ciaramella M (2009) Reverse gyrase and genome stability in hyperthermophilic organisms. *Biochem Soc Trans* **37:** 69-73

Pfeiffer HG, Liebhafsky HA (1951) The origins of Beer's law. *Journal of Chemical Education* **28:** 123-125

Philo JS (2006) Is any measurement method optimal for all aggregate sizes and types? *AAPS J* **8:** E564-571

Pugh RA, Honda M, Leesley H, Thomas A, Lin Y, Nilges MJ, Cann IK, Spies M (2008a) The iron-containing domain is essential in Rad3 helicases for coupling of ATP hydrolysis to DNA translocation and for targeting the helicase to the single-stranded DNA-double-stranded DNA junction. *The Journal of biological chemistry* **283:** 1732-1743

Pugh RA, Lin Y, Eller C, Leesley H, Cann IK, Spies M (2008b) Ferroplasma acidarmanus RPA2 facilitates efficient unwinding of forked DNA substrates by monomers of FacXPD helicase. *Journal of molecular biology* **383:** 982-998

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

Richards JD, Cubeddu L, Roberts J, Liu H, White MF (2008) The archaeal XPB protein is a ssDNA-dependent ATPase with a novel partner. *Journal of molecular biology* **376:** 634-644

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

Roberts JA, Bell SD, White MF (2003a) An archaeal XPF repair endonuclease dependent on a heterotrimeric PCNA. *Mol Microbiol* **48:** 361-371

Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, Blumenthal RM, Degtyarev S, Dryden DT, Dybvig K, Firman K, Gromova ES, Gumport RI, Halford SE, Hattman S, Heitman J, Hornby DP, Janulaitis A, Jeltsch A, Josephsen J, Kiss A, Klaenhammer TR, Kobayashi I, Kong H, Kruger DH, Lacks S, Marinus MG, Miyahara M, Morgan RD, Murray NE, Nagaraja V, Piekarowicz A, Pingoud A, Raleigh E, Rao DN, Reich N, Repin VE, Selker EU, Shaw PC, Stein DC, Stoddard BL, Szybalski W, Trautner TA, Van Etten JL, Vitor JM, Wilson GG, Xu SY (2003b) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic acids research* **31:** 1805-1812

Rouillon C, White MF (2010) The XBP-Bax1 helicase-nuclease complex unwinds and cleaves DNA: implications for eukaryal and archaeal nucleotide excision repair. *The Journal of biological chemistry* **285:** 11013-11022

Rouillon C, White MF (2011) The evolution and mechanisms of nucleotide excision repair proteins. *Res Microbiol* **162:** 19-26

Routledge MN, McLuckie KI, Jones GD, Farmer PB, Martin EA (2001) Presence of benzo[a]pyrene diol epoxide adducts in target DNA leads to an increase in UV-induced DNA single strand breaks and supF gene mutations. *Carcinogenesis* **22:** 1231-1238

Rudolf J, Makrantoni V, Ingledew WJ, Stark MJ, White MF (2006) The DNA repair helicases XPD and FancJ have essential iron-sulfur domains. *Molecular cell* **23:** 801-808

Rudolf J, Rouillon C, Schwarz-Linek U, White MF (2010) The helicase XPD unwinds bubble structures and is not stalled by DNA lesions removed by the nucleotide excision repair pathway. *Nucleic acids research* **38:** 931-941

Ruepp A, Graml W, Santos-Martinez ML, Koretke KK, Volker C, Mewes HW, Frishman D, Stocker S, Lupas AN, Baumeister W (2000) The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature* **407:** 508-513

Salerno V, Napoli A, White MF, Rossi M, Ciaramella M (2003) Transcriptional response to DNA damage in the archaeon *Sulfolobus solfataricus*. *Nucleic acids research* **31:** 6127-6138

Sancar A (1994) Mechanisms of DNA excision repair. *Science (New York, NY* **266:** 1954-1956

Sancar A (1996) DNA excision repair. *Annual review of biochemistry* **65:** 43-81

Schaeffer L, Moncollin V, Roy R, Staub A, Mezzina M, Sarasin A, Weeda G, Hoeijmakers JH, Egly JM (1994) The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *The EMBO journal* **13:** 2388-2392

Schaeffer L, Roy R, Humbert S, Moncollin V, Vermeulen W, Hoeijmakers JH, Chambon P, Egly JM (1993) DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science (New York, NY* **260:** 58-63

Schroder GF, Levitt M, Brunger AT (2010) Super-resolution biomolecular crystallography with low-resolution data. *Nature* **464:** 1218-1222

Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophysical journal* **78:** 1606-1619

Schultz P, Fribourg S, Poterszman A, Mallouh V, Moras D, Egly JM (2000) Molecular structure of human TFIIH. *Cell* **102:** 599-607

Sedelnikova OA, Redon CE, Dickey JS, Nakamura AJ, Georgakilas AG, Bonner WM (2010) Role of oxidatively induced DNA lesions in human pathogenesis. *Mutat Res* **704:** 152-159

Selby CP, Sancar A (1993) Molecular mechanism of transcription-repair coupling. *Science (New York, NY* **260:** 53-58

Selby CP, Sancar A (1994) Mechanisms of transcription-repair coupling and mutation frequency decline. *Microbiol Rev* **58:** 317-329

Sijbers AM, de Laat WL, Ariza RR, Biggerstaff M, Wei YF, Moggs JG, Carter KC, Shell BK, Evans E, de Jong MC, Rademakers S, de Rooij J, Jaspers NG, Hoeijmakers JH, Wood RD (1996) Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86:** 811-822

Singleton MR, Dillingham MS, Wigley DB (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annual review of biochemistry* **76:** 23-50

Skorvaga M, DellaVecchia MJ, Croteau DL, Theis K, Truglio JJ, Mandavilli BS, Kisker C, Van Houten B (2004) Identification of residues within UvrB that are important for efficient DNA binding and damage processing. *The Journal of biological chemistry* **279:** 51574-51580

Skorvaga M, Theis K, Mandavilli BS, Kisker C, Van Houten B (2002) The beta -hairpin motif of UvrB is essential for DNA binding, damage processing, and UvrC-mediated incisions. *The Journal of biological chemistry* **277:** 1553-1559

Sohi M, Alexandrovich A, Moolenaar G, Visse R, Goosen N, Vernede X, Fontecilla-Camps JC, Champness J, Sanderson MR (2000) Crystal structure of Escherichia coli UvrB C-terminal domain, and a model for UvrB-uvrC interaction. *FEBS Lett* **465:** 161-164

Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science (New York, NY* **305:** 1434-1437

Staresincic L, Fagbemi AF, Enzlin JH, Gourdin AM, Wijgers N, Dunand-Sauthier I, Giglia-Mari G, Clarkson SG, Vermeulen W, Scharer OD (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. *The EMBO journal* **28:** 1111-1120

Steitz TA, Steitz JA (1993) A general two-metal-ion mechanism for catalytic RNA. *Proceedings of the National Academy of Sciences of the United States of America* **90:** 6498-6502

Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **41:** 207-234

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. *Molecular cell* **36:** 642-653

Sugasawa K, Ng JM, Masutani C, Iwai S, van der Spek PJ, Eker AP, Hanaoka F, Bootsma D, Hoeijmakers JH (1998) Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair. *Molecular cell* **2:** 223-232

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 & development* **15:** 507-521

Sung P, Bailly V, Weber C, Thompson LH, Prakash L, Prakash S (1993) Human xeroderma pigmentosum group D gene encodes a DNA helicase. *Nature* **365:** 852-855

Szczelkun MD, Friedhoff P, Seidel R (2010) Maintaining a sense of direction during long-range communication on DNA. *Biochem Soc Trans* **38:** 404-409

The CCP4 suite: programs for protein crystallography. (1994) *Acta Crystallogr D Biol Crystallogr* **50:** 760-763

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

Taylor G (2003) The phase problem. *Acta Crystallogr D Biol Crystallogr* **59:** 1881-1890

Theis K, Chen PJ, Skorvaga M, Van Houten B, Kisker C (1999) Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. *The EMBO journal* **18:** 6899-6907

Theis K, Skorvaga M, Machius M, Nakagawa N, Van Houten B, Kisker C (2000) The nucleotide excision repair protein UvrB, a helicase-like enzyme with a catch. *Mutat Res* **460:** 277-300

Truglio JJ, Croteau DL, Skorvaga M, DellaVecchia MJ, Theis K, Mandavilli BS, Van Houten B, Kisker C (2004) Interactions between UvrA and UvrB: the role of UvrB's domain 2 in nucleotide excision repair. *The EMBO journal* **23:** 2498-2509

Truglio JJ, Croteau DL, Van Houten B, Kisker C (2006a) Prokaryotic nucleotide excision repair: the UvrABC system. *Chemical reviews* **106:** 233-252

Truglio JJ, Karakas E, Rhau B, Wang H, DellaVecchia MJ, Van Houten B, Kisker C (2006b) Structural basis for DNA recognition and processing by UvrB. *Nat Struct Mol Biol* **13:** 360-364

Truglio JJ, Rhau B, Croteau DL, Wang L, Skorvaga M, Karakas E, DellaVecchia MJ, Wang H, Van Houten B, Kisker C (2005) Structural insights into the first incision reaction during nucleotide excision repair. *The EMBO journal* **24:** 885-894

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

van der Leun JC (2004) The ozone layer. *Photodermatol Photoimmunol Photomed* **20:** 159-162

van Gool AJ, van der Horst GT, Citterio E, Hoeijmakers JH (1997) Cockayne syndrome: defective repair of transcription? *The EMBO journal* **16:** 4155-4162

van Hoffen A, Natarajan AT, Mayne LV, van Zeeland AA, Mullenders LH, Venema J (1993) Deficient repair of the transcribed strand of active genes in Cockayne's syndrome cells. *Nucleic acids research* **21:** 5890-5895

Van Houten B (1990) Nucleotide excision repair in Escherichia coli. *Microbiol Rev* **54:** 18-51

Van Houten B, Croteau DL, DellaVecchia MJ, Wang H, Kisker C (2005) 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. *Mutat Res* **577:** 92-117

Verhoeven EE, van Kesteren M, Moolenaar GF, Visse R, Goosen N (2000) Catalytic sites for 3' and 5' incision of Escherichia coli nucleotide excision repair are both located in UvrC. *The Journal of biological chemistry* **275:** 5120-5123

Verhoeven EE, Wyman C, Moolenaar GF, Goosen N (2002) The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands. *The EMBO journal* **21:** 4196-4205

Verhoeven EE, Wyman C, Moolenaar GF, Hoeijmakers JH, Goosen N (2001) Architecture of nucleotide excision repair complexes: DNA is wrapped by UvrB before and after damage recognition. *The EMBO journal* **20:** 601-611

Vermeulen W, de Boer J, Citterio E, van Gool AJ, van der Horst GT, Jaspers NG, de Laat WL, Sijbers AM, van der Spek PJ, Sugasawa K, Weeda G, Winkler GS, Bootsma D, Egly JM, Hoeijmakers JH (1997) Mammalian nucleotide excision repair and syndromes. *Biochem Soc Trans* **25:** 309-315

Volker M, Mone MJ, Karmakar P, van Hoffen A, Schul W, Vermeulen W, Hoeijmakers JH, van Driel R, van Zeeland AA, Mullenders LH (2001) Sequential assembly of the nucleotide excision repair factors in vivo. *Molecular cell* **8:** 213-224

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. *The Journal of biological chemistry* **277:** 1637-1640

Wakasugi M, Sancar A (1998) Assembly, subunit composition, and footprint of human DNA repair excision nuclease. *Proceedings of the National Academy of Sciences of the United States of America* **95:** 6669-6674

Wang H, DellaVecchia MJ, Skorvaga M, Croteau DL, Erie DA, Van Houten B (2006) UvrB domain 4, an autoinhibitory gate for regulation of DNA binding and ATPase activity. *The Journal of biological chemistry* **281:** 15227-15237

Wang H, Lu M, Tang MS, Van Houten B, Ross JB, Weinfeld M, Le XC (2009) DNA wrapping is required for DNA damage recognition in the Escherichia coli DNA nucleotide

excision repair pathway. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 12849-12854

Wang W, Malcolm BA (1999) Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. *BioTechniques* **26:** 680-682

Waters TR, Eryilmaz J, Geddes S, Barrett TE (2006) Damage detection by the UvrABC pathway: crystal structure of UvrB bound to fluorescein-adducted DNA. *FEBS Lett* **580:** 6423-6427

Weiss MS (2001) Global indicators of X-ray data quality. *J Appl Crystallogr* **34:** 130-135

Weiss MS, Hilgenfeld R (1997) On the use of the merging R-factor as a quality indicator for X-ray data. *J Appl Crystallogr* **30:** 203-205

White MF (2003) Archaeal DNA repair: paradigms and puzzles. *Biochem Soc Trans* **31:** 690-693

White MF, Bell SD (2002) Holding it together: chromatin in the Archaea. *Trends Genet* **18:** 621-626

Wilson WW (2003) Light scattering as a diagnostic for protein crystal growth--a practical approach. *J Struct Biol* **142:** 56-65

Winkler GS, Araujo SJ, Fiedler U, Vermeulen W, Coin F, Egly JM, Hoeijmakers JH, Wood RD, Timmers HT, Weeda G (2000) TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair. *The Journal of biological chemistry* **275:** 4258-4266

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

Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences of the United States of America* **74:** 5088-5090

Wolski SC, Kuper J, Hanzelmann P, Truglio JJ, Croteau DL, Van Houten B, Kisker C (2008) Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS biology* **6:** e149

Wolski SC, Kuper J, Kisker C (2010) The XPD helicase: XPanDing archaeal XPD structures to get a grip on human DNA repair. *Biol Chem* **391:** 761-765

Wood ER, Ghane F, Grogan DW (1997) Genetic responses of the thermophilic archaeon Sulfolobus acidocaldarius to short-wavelength UV light. *Journal of bacteriology* **179:** 5693-5698

Yang Y, Sass LE, Du C, Hsieh P, Erie DA (2005) Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions. *Nucleic acids research* **33:** 4322-4334

Yang Y, Wang H, Erie DA (2003) Quantitative characterization of biomolecular assemblies and interactions using atomic force microscopy. *Methods* **29:** 175-187

Yeates TO (1988) Simple statistics for intensity data from twinned specimens. *Acta Crystallogr A* **44 (Pt 2):** 142-144

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. *The Journal of biological chemistry* **275:** 9870-9875

Zou Y, Van Houten B (1999) Strand opening by the UvrA(2)B complex allows dynamic recognition of DNA damage. *The EMBO journal* **18:** 4889-4901

Zou Y, Walker R, Bassett H, Geacintov NE, Van Houten B (1997) Formation of DNA repair intermediates and incision by the ATP-dependent UvrB-UvrC endonuclease. *The Journal of biological chemistry* **272:** 4820-4827

Tables

Table A1: Crystallization conditions for UvrBC-DNA complex

Protein concentration (A_{280})	Drop Size (Protein:Reservoir)	Hanging/sitting drop	Temperature	Screen
4.0	0.25 µl + 0.25 µl	sitting drop	20 °C	• Crystal Screen I+II • Nextal PEGs • Easy Xtal JCSG • The Nucleix Suite
6.3	0.4 µl + 0.2 µl	sitting drop	20 °C	• Crystal Screen I+II • Easy Xtal JCSG • The Nucleix Suite
9.4	1 µl + 1 µl 1.5 µl + 0.75 µl 1.5 µl + 0.75 µl 0.75 µl + 1.5 µl	hanging drop	20 °C 20 °C 4 °C 20 °C	sub-screen: 50 mM Na-cacodylate pH 6.5, 2-2.5 mM Spermine, 30 mM $MgCl_2$, 0-1 mM $CoCl_2$, 3-10 % PEG400
15.1	1.5 µl + 0.75 µl	hanging drop + streak seeding	20 °C	sub-screen: 50 mM Na-cacodylate pH 5.5-pH 7.0, 2-2.5 mM Spermine, 10-30 mM $MgCl_2$, 2-7 % PEG400
8.4	0.3 µl + 0.3 µl	sitting drop	4 °C, 10 °C, 20 °C, 37 °C	• Crystal Screen I+II • Index Screen HT • The Nucleix Suite • OptiMix-1-5 • Wizard Screen I+II
8.4	1.5 µl + 0.75 µl	hanging drop	10 °C	sub-screen: 50 mM Na-cacodylate pH 5.5-pH 7.0, 2-2.5 mM Spermine, 10-30 mM $MgCl_2$, 2.5-15 % PEG400 or 5-20 % tert-Butanol or 2-12 % MPD or 5-20 % Isopropanol
6.4	1 µl + 1 µl	hanging drop	20 °C	sub-screen: 0.4-1 M K-thiocyanate 0.1 mM Na citrate
9.4	0.3 µl + 0.3 µl 0.4 µl + 0.2 µl 0.6 µl + 0.3 µl 0.4 µl + 0.3 µl 0.6 µl + 0.2 µl	sitting drop	20 °C	sub-screen: 50 mM HEPES pH 7.5, 1-2 mM Spermine, 10-30 mM $MgCl_2$, 0-15 % PEG400 or 0-20 % PEG8000 or 0-2.5 M $(NH_4)_2SO_4$ or 0-20 % Isopropanol

Table A2: Crystallization of Bax1, XPB and the XPB/Bax1 complex

Protein type and concentration	Buffer composition	Drop Size (Protein:Reservoir)	Hanging/sitting drop	Temperature	Screen	DNA	Additives
Bax1 monomer 28 mg/ml	IMAC-A (20 mM Imidazole)	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Wizard I+II • Protein Complex Suite • OptiMix 1 • Nextal pH • Nextal PEG		
Bax1 dimer 9.3 mg/ml	IMAC-A (20 mM Imidazole)	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Wizard I+II • Protein Complex Suite • OptiMix 1 • Nextal pH • Nextal PEG		
Bax1 12 mg/ml	20mM MES pH6.5, 500mM NaCl	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Wizard I+II • OptiMix 1 • Nextal pH • Nextal PEG		
XPB/Bax1$_{incomp}$ 2.7 mg/ml	20mM MES pH6.5, 500mM NaCl	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Protein Complex Suite		
Bax1 12 mg/ml	20mM MES pH6.5, 500mM NaCl	0.3µl + 0.3µl	sitting	20 °C	• Index Screen • OptiMix 3		

Bax1 10 mg/ml	20mM MES pH6.5, 500mM NaCl	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Index Screen • Wizard II+III • OptiMix 1 • Nextal pH • Nextal PEG		
Bax1 28 mg/ml	20mM MES pH6.5, 500mM NaCl	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Index Screen • Wizard II+III • OptiMix 1 • Nextal pH • Nextal PEG		
Bax1 250 µM	20mM MES pH6.5, 150mM NaCl, 10mM $MgCl_2$	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Index Screen • Wizard I+II • OptiMix 1 • OptiMix 3 • Nextal PEG	F_{15}26GAA11 (hairpin bubble)	
Bax1 250 µM	20mM MES pH6.5, 150mM NaCl, 10mM $MgCl_2$	0.3µl + 0.3µl 0.6µl + 0.3µl 0.3µl + 0.6µl	sitting	20 °C	• OptiMix 1 • OptiMix 3	F_{15}26GAA11 (hairpin bubble) HB47asF (hairpin bubble)	
Bax1 20 mg/ml	GF buffer	0.3µl + 0.3µl 0.6µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Index Screen • Nucleix Suite • Protein Complex Suite • OptiMix 1		seeds from amorphous non-diffracting quasi-crystals

Bax1 178 µM	GF buffer + 10 mM $CaCl_2$	0.4µl + 0.4µl	sitting	20 °C	• Crystal Screen I+II • Index Screen • Nucleix Suite • Protein Complex Suite • OptiMix 1 • OptiMix 3 • Wizard I+II	NDB22 (214 µM)	
Bax1 31 mg/ml	GF buffer	0.6µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Index Screen • Nucleix Suite • Protein Complex Suite • OptiMix 1 • Wizard I+II		Zeolith
XPB/$Bax1_{comp}$ 8.5 µM	GF buffer (250 mM NaCl) + 10 mM $MgCl_2$	0.3µl + 0.3µl 0.6µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Protein Complex Suite		
XPB/$Bax1_{comp}$ 21 µM	GF buffer (500 mM NaCl) + 10 mM $MgCl_2$	0.6µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Protein Complex Suite • OptiMix 1 • OptiMix 3 • Wizard I+II • Nextal PEG		
XPB/$Bax1_{comp}$ 15 µM	GF buffer (500 mM NaCl) + 10 mM $CaCl_2$	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Protein Complex Suite • Wizard I+II		
XPB/$Bax1_{comp}$ 12 µM	GF buffer (500 mM NaCl) + 10 mM $CaCl_2$	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Protein Complex Suite • Wizard I+II	DNA NDB30/NDBr 15 (14.4 µM)	5 mM EDTA

XPB/Bax1$_{comp}$ 12 µM	GF buffer (500 mM NaCl) + 10 mM $CaCl_2$	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Index Screen • Nucleix Suite • Protein Complex Suite • Wizard I+II • OptiMix 1	DNA NDB30/NDBr 15 (14.4 µM)	5 mM EDTA, crystal seeds
methylated Bax1 57 µM (2.6 mg/ml)	GF buffer (200 mM NaCl)	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II • Index Screen • Nucleix Suite • Protein Complex Suite • Wizard I+II • OptiMix 1 • Nextal PEG		
methylated Bax1 500 µM	GF buffer (200 mM NaCl)	0.3µl + 0.3µl	sitting	20 °C	• Crystal Screen I+II		
methylated Bax1 250 µM	GF buffer (100 mM NaCl)	0.3µl + 0.3µl	sitting	20 °C 10 °C 4 °C	• Crystal Screen I+II • Index Screen • Nucleix Suite • Protein Complex Suite • Wizard I+II • OptiMix 1 • Nextal PEG		

Affidavit

(Eidesstattliche Erklärung)

I hereby declare that my thesis entitled

“**Nucleotide Excision Repair: From Recognition to Incision of damaged DNA”**

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

Würzburg,……

Date Signature

Acknowledgments

In the first place, I would like to thank my supervisor Prof. Caroline Kisker for the opportunity to work on several exciting projects, for her continuous support and advice, for her guidance and for her encouragement. Special thanks also to my second supervisor Prof. Bennett Van Houten for many fruitful and vivid discussions, and for being a great mentor. Especially, I would like to thank him for the great time I had during my research stay in his laboratory.

I also would like to thank Prof. Thomas Müller for his support and many helpful scientific discussions and suggestions in my committee meetings. Thanks also to Prof. Helga Stopper for her contributions in my advisory committee.

I am grateful to many people in the lab for their contributions to this present work: Florian Rohleder cloned the genes encoding XPB and Bax1 into expression vectors. Atomic Force Microscopy experiments as well as analytical ultracentrifugation experiments were kindly performed by Dr. Ingrid Tessmer, Verena Grundler and Johannes Römer. For data collection and his help with solving the crystal structure of UvrB, I would like to acknowledge Dr. Jochen Kuper. ATPase assays were kindly carried out by Gudrun Michels. Thanks also to Dr. Hong Wang and Dr. Deborah Croteau for their advice and help with biochemical assays. And I would like to especially thank Dr. Ingrid Tessmer for many fruitful and enthusiastic discussions, for her scientific support and her contributions to our publications.

Personally, I would like to express my heartfelt gratitude to Dr. Marcus Resch for many helpful scientific discussions, for his help with the UvrB crystal structure, for critically reading and correcting the draft of this manuscript and of course for his constant support.

Finally, I would like to sincerely thank my beloved parents for supporting and encouraging me throughout my whole life.

Publications

Parts of the present work resulted in the following publications:

H. M. Roth, J. Römer, V. Grundler, B. Van Houten, C. Kisker and I. Tessmer, “The nucleotide excision repair helicase XPB regulates Bax1 incision”, *manuscript submitted*

H. M. Roth, I. Tessmer, B. Van Houten, and C. Kisker, "Bax1 is a novel endonuclease: implications for archaeal nucleotide excision repair," J Biol Chem **284** (47), 32272-32278 (2009).

Additional publications:

L. Polzien, A. Baljuls, **H. M. Roth**, J. Kuper, R. Benz, K. Schweimer, M. Hekman, U. R. Rapp, “Pore-forming activity of BAD is regulated by specific phosphorylation and structural transitions of the C-terminal part”, BBA - General Subjects (2010), *in press*

M. Resch, **H. M. Roth**, M. Kottmair, M. Sevvana, R. Bertram, F. Titgemeyer, and Y. A. Muller, "Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of YvoA from Bacillus subtilis," Acta Crystallogr Sect F Struct Biol Cryst Commun **65** (Pt 4), 410-414 (2009).

E. F. Pimentel-Schmitt, A. W. Thomae, J. Amon, M. A. Klieber, **H. M. Roth**, Y. A. Muller, K. Jahreis, A. Burkovski, and F. Titgemeyer, "A glucose kinase from Mycobacterium smegmatis," J Mol Microbiol Biotechnol **12** (1-2), 75-81 (2007).