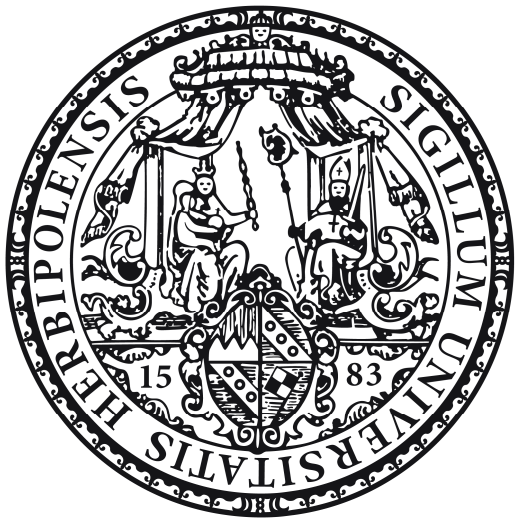


Structural and functional characterization of nucleotide excision repair proteins

Strukturelle und funktionelle Charakterisierung von Nucleotid-Exzisions-Reparatur Proteinen

Doctoral thesis for a doctoral degree

at the Graduate School of Life Sciences,

Julius-Maximilians-Universität Würzburg,

Section Biomedicine

submitted by

Stefanie Carola Wolski

from Vaihingen/Enz

Würzburg, September 2011

Submitted on: ……..

Office stamp

Members of the *Promotionskomitee*:

Chairperson: Ulrike Holzgrabe

Primary Supervisor: Caroline Kisker

Supervisor (Second): Thomas Müller

Supervisor (Third): Yves Muller

Supervisor (Forth): Bennett Van Houten

Date of Public Defense: ……

Date of Receipt of Certificates: ……

Table of Contents

Summary

XPD is a 5‘-3‘ helicase of the superfamily 2. As part of the transcription factor IIH it functions in transcription initiation and nucleotide excision repair. This work focus on the role of XPD in nucleotide excision repair. NER is a DNA repair pathway unique for its broad substrate range. In placental mammals NER is the only repair mechanism able to remove lesions induced by UV-light. NER can be divided into four different steps that are conserved between pro- and eukaryotes. Step 1 consists of the initial damage recognition, during step 2 the putative damage is verified, in step 3 the verified damage is excised and in the 4th and final step the resulting gap in the DNA is refilled. XPD was shown to be involved in the damage verification step.

It was possible to solve the first apo XPD structure by a MAD approach using only the endogenous iron from the iron sulfur cluster. Based on the apo XPD structure several questions arise: where is DNA bound? Where is DNA separated? How is damage verification achieved? What is the role of the FeS cluster? These questions were addressed in this work. Hypothesis driven structure based functional mutagenesis was employed and combined with detailed biochemical characterization of the variants. The variants were analyzed by thermal unfolding studies to exclude the possibility that the overall stability could be affected by the point mutation. DNA binding assays, ATPase assays and helicase assays were performed to delineate amino acid residues important for DNA binding, helicase activity and damage recognition. A structure of XPD containing a four base pair DNA fragment was solved by molecular replacement. This structure displays the polarity of the translocated strand with respect to the helicase framework. Moreover the properties of the FeS cluster were studied by electron paramagnetic resonance to get insights into the role of the FeS cluster. Furthermore XPD from *Ferroplasma acidarmanus* was investigated since it was shown that it is stalled at CPD containing lesions. The data provide the first detailed insight into the translocation mechanism of a SF2B helicase and reveal how polarity is achieved. This provides a basis for further anlayses understanding the combined action of the helicase and the 4Fe4S cluster to accomplish damage verification within the NERcascade.

Zusammenfassung

XPD ist eine 5‘-3‘ Helicase der Superfamilie 2. Als Untereinheit des Transkriptionsfaktors IIH ist XPD in Transkriptionsinitiation und Nucleotid-Exzisions-Reparatur involviert. Diese Arbeit fokusiert auf die Rolle von XPD in der NER. NER ist ein DNA Reparatur Weg der bekannt ist für seine breite Substratspezifität. In Säugetieren ist NER der einzige Reparaturmechanismus, der fähig ist Läsionen zu reparieren, die durch UV Strahlung induziert werden. NER kann man in vier unterschiedliche Schritte aufteilen die zwischen Pro- und Eukaryoten konserviert sind. Schritt 1 besteht aus der initialen Schadenserkennung, während des zweiten Schrittes wird der mögliche Schaden verifiziert, im dritten Schritt wird der verifizierte Schaden ausgeschnitten und im vierten und letzten Schritt wird die resultierende Lücke in der DNA geschlossen. Es wurde gezeigt, dass XPD in die Schadensverifizierung involviert ist.

Ein MAD Versuch, bei dem nur das endogene Eisen des Eisen-Schwefel-Clusters verwendet wurde ermöglichte die Strukturlösung der ersten apo XPD Struktur. Basierend auf der Struktur ergeben sich verschiedene Fragen: wo wird DNA gebunden? Wo wird DNA aufgetrennt? Wie wird Schadenserkennung ermöglicht? Was ist die Rolle des Eisen-Schwefel-Clusters? Diese Fragen werden in dieser Arbeit angesprochen. Strukturbasierte funktionelle Mutagenesestudien, die auf Hypothesen basiert sind, wurden angewendet und mit einer detailierten biochemischen Charakterizierung der Varianten kombiniert. Die Varianten wurden mittels thermischen Entfaltungsstudien analysiert, um die Möglichkeit auszuschliessen, dass die Stabilität durch die Punktmutation betroffen ist. DNA-Bindungs- Assays, ATPase Assays und Helikase Assays wurden durchgeführt um Aminosäurereste zu identifizieren, die für DNA Bindung, Helikase Aktivität und Schadenserkennung wichtig sind. Eine Struktur von XPD, die ein DNA Fragment mit vier Basen enthält, wurde mittels Molekularem Ersatz gelöst. Diese Struktur zeigt die Polarität des translozierenden DNA- Stranges im Verhältnis zu der Helikasestruktur auf. Desweiteren wurden die Eigenschaften des FeS Clusters mittels paramagnetischen Elektronenresonanz Studien untersucht, um Einblicke in die Rolle des FeS Clusters zu bekommen. Ausserdem wurde XPD aus *Ferroplasma acidarmanus* erforscht, da gezeigt wurde, dass es an CPD enthaltenden Läsionen hängen bleibt. Diese Daten stellen die ersten detailierten Einblicke in den Translokationsmechanismus einer SF2B Helikase dar und zeigen

wie Polarität erzielt wird. Das ist eine Basis für weitere Analysen, um die kombinierte Aktion von Helikase und dem 4Fe4S Cluster zu verstehen, die zur Schadenserkennung in der NER Kaskade führt.

1 Introduction

The DNA of each cell is endangered by a variety of modifications arising spontaneously by replication errors, endogenously by reactive metabolites, or exogenously after exposure to environmental mutagens or ultraviolet (UV) light. Damages occur quite frequently at a rate of about 200,000 damages per cell per day. If lesions in DNA cannot be eliminated, a cell may undergo apoptosis or will accumulate mutations and transform into a potentially cancerous cell that might proliferate into a tumor(Sarasin 2003).

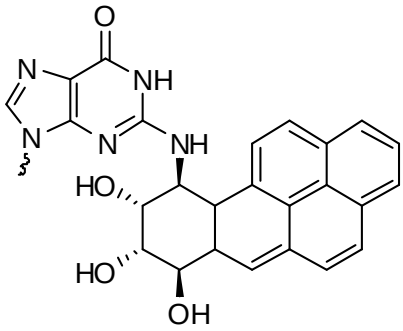
To deal with all the different types of DNA lesions different DNA repair pathways evolved in all kingdoms of life. DNA repair is commonly dived into five major pathways: direct damage reversal, base excision repair (BER), mismatch repair, double strand break repair and nucleotide excision repair (NER). NER is one of the most versatile DNA repair systems. It can be subdivided into two distinct subpathways: global genome repair (GGR), which can detect and remove lesions throughout the genome and relies on a dedicated set of proteins, and transcription-coupled repair (TCR), which ensures faster repair of many lesions when located on the transcribed strand of actively transcribed genes (Figure 1.2)(Guo et al. 2010). TCR is initiated by a RNA polymerase stalled at a lesion (Svejstrup 2002), (Brueckner et al. 2007), (Hanawalt and Spivak 2008). A key player in eukaryotic NER is transcription factor IIH (TFIIH) which functions in NER and transcription(Bradsher et al. 2000).

1.1 Lesions addressed by nucleotide excision repair

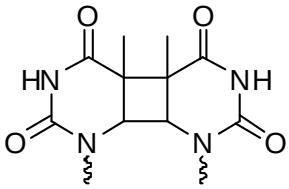
NER is a particularly intriguing DNA repair pathway because of its extraordinarily wide substrate specificity; it has the ability to recognize and repair a large number of structurally unrelated lesions(Sancar 1994), (Sancar et al. 1996), such as DNA damage formed upon exposure to the UV radiation from sunlight and numerous bulky DNA adducts induced by mutagenic chemicals from the environment or by cytotoxic drugs used in chemotherapy as reviewed in(Gillet and Scharer 2006)and (Nouspikel 2009).

NER is also involved in the repair of crosslinking agents. These chemicals can form two distinct bonds with DNA, either on the same strand (intrastrand

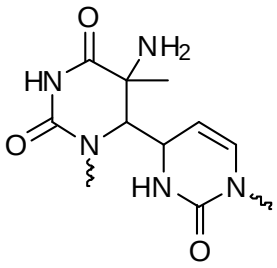
crosslinks) or across both strands (interstrand crosslinks)(Trimmer and Essigmann 1999). For the repair of interstrand crosslinks a combination of NER with translesion synthesis and/or homologous recombination is necessary, because NER requires an intact strand to serve as template to repair the other strand (Kuraoka et al. 2000). Classical examples for crosslinking agents are cis-dichloro-diaminoplatin (cisplatin) (Figure 1.1), nitrogen mustards and psoralens (Trimmer and Essigmann 1999), (Hearst et al. 1984).



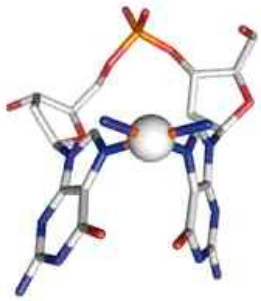
benzopyrene-diol-epoxide adduct



cyclobutane pyrimidine dimer



6-4 photoproduct



GG cisplatin lesion

Figure 1.1: DNA lesions repaired by NER(GG cisplatin lesion reprinted by permission from Macmillan Publishers Ltd: Nature Structural and Molecular Biology(Damsma et al. 2007), ©2007

Other chemicals handled by NER are for example benzo[a]pyrene (cigarette smoke) (Figure 1), aromatic amines, aflatoxin and nitrosamines (Wogan et al. 2004). These chemicals share a common property, they are present in the environment in a relatively harmless form. These so-called proximate carcinogens are absorbed with food (nitrosamines, aflatoxin) or air (benzo[a]pyrenes) and then activated by cellular metabolism to highly reactive species. These 'ultimate carcinogens' are highly reactive and have a strong tendency to form bulky adducts on DNA (Wogan et al. 2004).

Another class of lesions which is repaired by NER is UV-induced lesions. In placental mammals NER is the only error-free DNA repair mechanism which can repair these lesions (Hoeijmakers 2009), (Friedberg et al. 2006), (Reardon and Sancar 2003). Short-wavelength UV light can cause the formation of covalent bonds between two adjacent pyrimidines on the same strand. Cyclobutane pyrimidine dimers (CPDs) and (6-4)pyrimidine-pyrimidone photoproducts [(6-4)PPs] (Figure 1.1) are the two main UV-induced lesions (Pfeifer 1997), (Nouspikel 2009).

Oxidized bases are generally repaired by base excision repair; however several types of oxidative damage are recognized by NER. For example cyclopurines (Brooks 2007) and malondialdehyde formed DNA adducts (Johnson et al. 1997).

1.2 Molecular mechanism of eukaryotic NER

The mechanism of NER has been reconstitued in vitro (Aboussekhra et al. 1995), (Mu et al. 1995) and is now better understood. NER consists of several consecutive steps: first lesion sensing, second opening of a bubble, third incision of the damaged strand, as a fourth step displacement of the lesion-containing oligonucleotide, finally gap filling and ligation (Figure 1.2)(Nouspikel 2009).

There are two requirements that NER takes place(Sugasawa 2001). First there must be a distortion in the structure of the double helix and second there must be a chemical modification in the DNA. In mammalian GGR the distortion-sensing complex consists of the subunits XPC, HR23B and centrin 2 (Araki et al. 2001). The XPC/HR23B heterodimer opens the DNA locally (Min and Pavletich 2007). Upon DNA damage XPC is polyubiquitinated which increases XPC's affinity for DNA (Sugasawa et al. 2005).Centrin 2 stabilizes the complex and improves NER activity, but its presence is not strictly required for NER (Araki et al. 2001). Not all lesions cause an equal distortion of the double-helix. (6-4)PPs for example lead to a strong kink in the DNA making them very good NER substrates, whereas CPDs

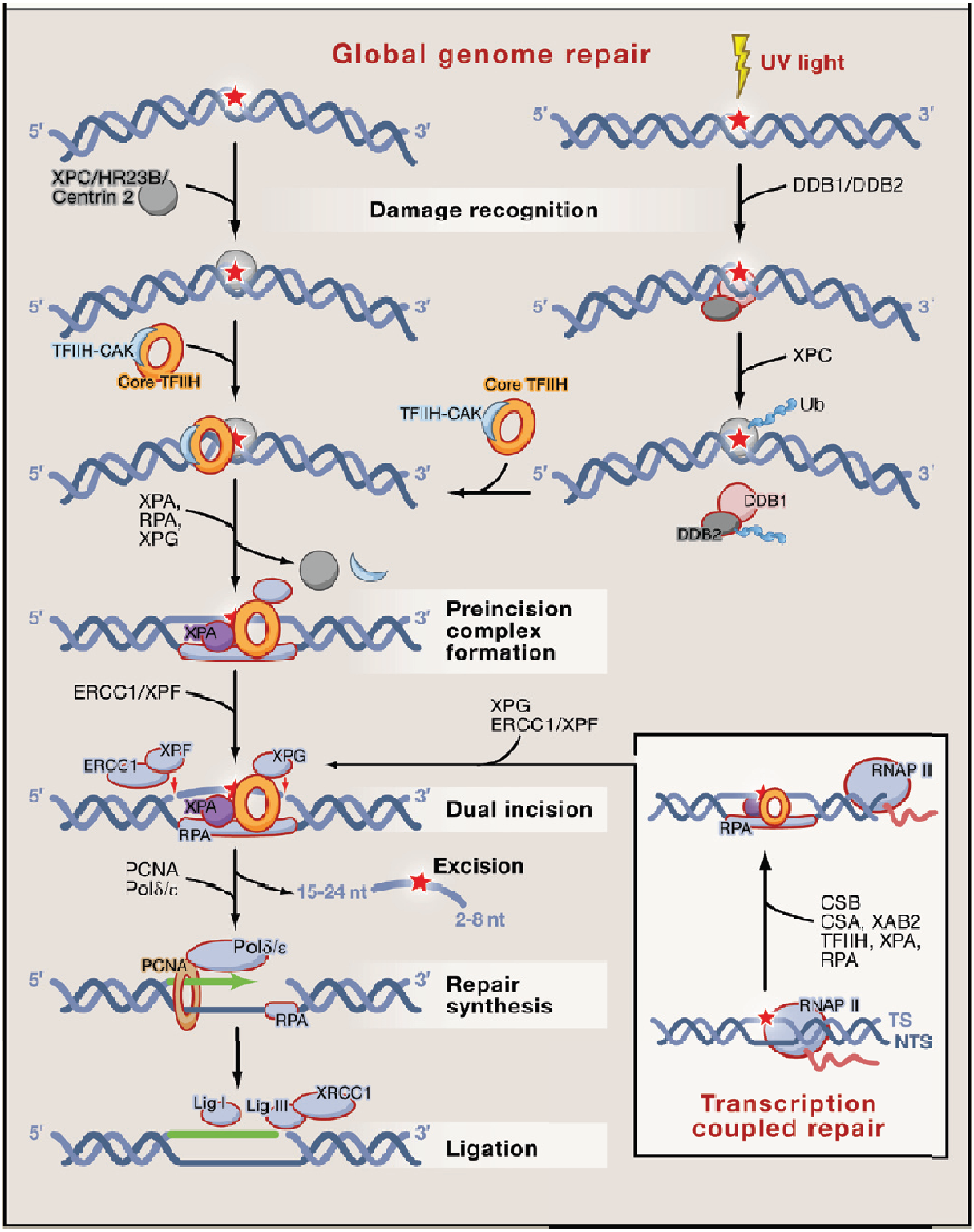


Figure 1.2: Model of the eukaryotic NER pathway.(Reprinted by permission from Elsevier, Cell(Guo et al., 2010), ©2010) DNA lesions (indicated by a red star) are initially detected by the XPC-HR23B complex, UV-lesions can alternatively be detected by the DDB-complex which recruits TFIIH. The ten subunit 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. Finally DNA synthesis takes place with the combined activities of DNA polymerases δ and ε or PCNA. Ligase III seals the nick. In TCR RNA-

Polymerase II is stalled at a lesion which leads to the recruitment of CSA, CSB and TFIIHthe following steps are the same as for GGR (Guo et al., 2010).

just lead to a modest distortion and are therefore poor substrates for NER (Kim et al., 1995). (6-4)PPs are recognized by the XPC complex alone, but for efficient repair of CPDs the DDB complex is needed. The DDB complex is a damaged DNA-binding heterodimer consisting of DDB1 and DDB2 (XPE). This complex has a high affinity for DNA damage, such as CPDs and (6-4)PPs (Wittschieben et al. 2005), (Payne and Chu 1994). It was proposed that DDB induces a kink in the DNA upon binding to a lesion which allows the recruitment of the XPC complex (Tang and Chu 2002), (Nouspikel 2009).

In transcription-coupled repair the initial detection step is achieved differently. RNA polymerase II is stalled at a lesion during transcription which is a strong signal for apoptosis(Brueckner et al. 2007), (Hanawalt and Spivak 2008). CSA and CSB (Cockayne's syndrome complementation group A and B) are then recruited as well as TFIIH to perform NER (Laine and Egly 2006), (Laine et al. 2006), (van Hoffen et al. 1993).

The next step in NER is the opening of a bubble around the lesion by the transcription factor TFIIH. TFIIH is a 10 subunit containing complex which functions in NER and in transcription(Bradsher et al. 2000). TFIIH contains the two helicases XPB and XPD, which are of opposite polarity. XPD, a 5'-3' helicase, plays a critical role in NER by opening a bubble around the lesion (Winkler et al. 2000). XPB, a 3'-5'helicase, is also required for this function, but in the repair process it is XPB's ATPase activity which is necessary, not its helicase activity (Coin et al. 2007). XPB hook up to the complex whereas XPD unwinds the DNA(Winkler et al. 2000), (Coin et al. 2007). XPB, XPD and five other subunits (p62, p52, p44, p34 and TTD-A) form a ring-shaped structure with a central cavity, which could accommodate double-stranded DNA(Schultz et al. 2000). This is the core of TFIIH to which the cyclin-activated kinase (CAK) complex (cyclin H, cdk7, MAT1) is attached (Chang and Kornberg 2000), (Schultz et al. 2000). MAT1 mediates the interaction between the CAK complex and the core of TFIIH (Busso et al. 2000). The CAK complex works as a phosphorylation cascade: cyclin H is phosphorylated by external kinases and phosphorylates Cdk7, which (when it is

part of TFIIH) phosphorylates the large subunit of RNA polymerase II(Nouspikel 2009). This allows RNA polymerase II to enter the elongation mode.

After TFIIH the XPA protein complex is recruited (Volker et al. 2001) which removes the XPC/HR23Bcomplex in this process (Hey et al. 2002) and catalyzes the detachment of CAK from the core (Coin et al. 2008). XPA binds DNA with a slight preference for damaged DNA (Jones and Wood 1993)and is associated with the three subunits of the RPA heterotrimer(Matsuda et al. 1995). However, the exact role of the XPA complex is still not clear. A possible role for the XPA complex could be the identification of the strand that carries the lesion (Sugasawa et al. 1998). It is also possible that this discrimination function is accomplished by TFIIH (Wood 1999). It was shown that translocation of the yeast homologue Rad3 is stalled at sites of damage *in vivo*(Naegeli et al. 1993).In addition it was shown that XPD from *Ferroplasma acidarmanus*(faXPD) is stalled at lesions containing a CPD (Mathieu et al. 2010) indicating a role for XPD in damage recognition.

The next step in NER is the incision of the damaged strand at both sides of the bubble after the dissociation of the CAK complex from the TFIIH core. The XPG endonuclease carries out the 3'cut (O'Donovan et al. 1994), whereas the 5'cut is performed by the XPF-ERCC1 heterodimer (Mu et al. 1996) which is recruited to the NER incision complex by XPA(Park and Sancar 1994), (Li et al. 1994), (Li et al. 1995),(Saijo et al. 1996). The presence of XPG is necessary for XPF-ERCC1 to incise the other end of the bubble, but the incision by XPG does not essentially come first (Wakasugi et al. 1997), (Gillet and Scharer 2006). As for the XPF-ERCC1 complex, it is XPF that harbors the endonuclease activity (Enzlin and Scharer 2002). The role of ERCC1 in the complex is not clear, aside from the fact that it stabilizes XPF (Houtsmuller et al. 1999).

The last steps in NER are gap filling and ligation. The fragment excised by mammalian NER is about 24-32 nucleotides in length, depending on the lesion (Moggs et al. 1996), (Matsunaga et al. 1995), (Svoboda et al. 1993). The excised lesion is located closer to the 3'side of the removed DNA as to the 5' side indicating a asymmetric pattern for incision(Moggs et al. 1996). The resulting gap is filled by either of the replicative DNA polymerases delta or epsilon (Popanda et al., 1992), associated with the 'sliding clamp' PCNA (Shivji et al. 1992), (Nichols and Sancar 1992). Recent data indicate that it is mostly ligase III, together with its

partner XRCC1, which seals the nick (Moser et al. 2007) and not as previously thought ligase I (Nocentini 1999).

1.3 Insights into the mechanism and functions of the XPD protein

XPD is an ATP-dependent 5'-3'helicase of the Superfamily 2 (SF2). Helicases are characterized by seven helicase motifs (walker motif I, Ia, II, III, IV, V and VI) constituted of conserved amino-acid sequences.

Most archaea encode a homologue of XPD. We have chosen to analyze *Thermoplasma acidophilum* XPD (taXPD) as a homologue of human XPD. However, the function of XPD in archaea is still unclear (Kelman and White 2005). The archaeal protein from *Sulfolobus acidocaldarius* was the first to be characterized following heterologous expression in *E.coli*(Rudolf et al. 2006). During purification it was revealed that XPD contains an FeS cluster indicated by a yellow-brownish colour of the protein (Rudolf et al. 2006)(Figure 1.3). There are four cysteine residues located towards the N-terminus of the protein which are conserved also in eukaryotic XPDs which indicates the presence of a FeS cluster. Moreover the four conserved cysteine residues were found in other superfamily 2 (SF2) helicases such as bacterial DinG (damage-inducible G) and the eukaryotic XPD paralogs FancJ (Fanconi's anemia complementation group J), RTEL (regular of telomere length) and Chl1 (chromatid cohesion in yeast) making XPD a founding member of a family of related helicases(Rudolf et al. 2006), (White 2009).

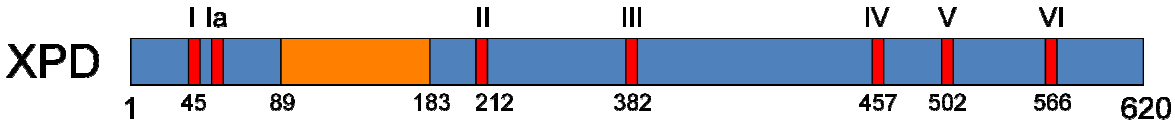


Figure 1.3:Diagram representing the primary structure of taXPD. The helicase motifs are coloured in red and the FeS cluster in orange.

It was shown that the iron-sulfur cluster domain is not essential for the stability or the enzyme's ability to bind to ssDNA or its ATPase activity, but it is essential for the helicase activity of XPD (Rudolf et al., 2006). However, the exact function of the FeS cluster remains unclear at present. Since XPD has been predicted to play a role in the detection of DNA damage during NER it could be by analogy with the DNA glycosylases that the FeS cluster plays an active role in damage detection

(Boal et al. 2005). Alternatively the FeS cluster might function in redox sensing (White 2009). Finally the FeS cluster could have a purely structural role, stabilizing the FeS domain where the DNA duplex is separated during helicase function (Pugh et al. 2008). It was furthermore shown for XPD from *Ferroplasma acidarmanus* that the single-stranded DNA binding protein RPA2 stimulates DNA unwinding by the faXPD helicase (Pugh et al. 2008). It was suggested that duplex melting by a cognate ssDNA-binding protein coordinated with translocation by a helicase may represent a common strategy for duplex unwinding by the Rad3 family of helicases (Pugh et al. 2008). Distance-dependent fluorescence quenching by the iron-sulfur cluster of archaeal XPD revealed that the two cognate ssDNA-binding proteins differentially affect XPD translocation. RPA1 competes with XPD for ssDNA access whereas RPA2 does not interfere with XPD-ssDNA binding but markedly slowed down XPD translocation (Honda et al. 2009). It was proposed that translocation on a protein-coated DNA lattice may represent a generalized feature used by Rad3 family helicases. There are two possible mechanism for these helicases, XPD in particular, to bypass ssDNA-binding proteins: XPD can translocate over the bound ssDNA-binding protein or the SSB protein dissociates ahead of the translocating XPD (Honda et al. 2009), (Spies and Ha 2010).

In eukaryotes XPD functions as part of the TFIIH complex whereas in archaea XPD is a monomer in solution and so far no stable interactions with other protein partners have been reported (Rouillon and White 2011). XPD is an essential structural component of TFIIH and bridges the TFIIH core subunits with the CAK complex (Reardon et al. 1996) by its interaction with p44 and MAT1. Once TFIIH is bound to the damaged DNA, the interaction between p44 and XPD leads to the stimulation of XPDs helicase activity (Coin et al. 1998), (Dubaele et al. 2003), (Egly and Coin 2011). The stimulation of XPD helicase activity by p44 allows the unwinding of the DNA around the damage, a prerequisite for the dual incision executed by XPG and ERCC1-XPF. The helicase activity of XPD is essential for NER, but dispensable for transcription initiation (Tirode et al. 1999). Mutations in the XPD gene lead to three severe diseases: xeroderma pigmentosum, Cockayne Syndrome and trichothiodystrophy, indicating the importance of a fully functional XPD (Lehmann 2001). The mutations found in the 3'-end of XPD represent more than 80 % of the mutations found in XP-D patients and hinder the XPD-p44

interaction (Coin et al. 1998), (Dubaele et al. 2003), (Egly and Coin 2011). Recently it was discovered that XPD also exists in TFIIH-free complexes and plays roles in other processes than transcription and repair. As part of the MMXD complex XPD is required for proper chromosome segregation (Ito et al. 2010). XPD can also form a complex with CAK which functions as a control for cell cycle progression (Cameroni et al. 2010).

A highly debated question is whether or not XPD is responsible for the damage verification in the NER pathway and if it acts alone or with other factors such as XPA. It was shown for the yeast homolog Rad3 that translocation is stalled at sites of damage *in vivo*(Naegeli et al. 1993). Moreover the enzymatic activity of XPD promotes its own anchoring to damaged DNA in living cells, thus supporting the conclusion that XPD is directly responsible for the predicted lesion verification step (Mathieu et al. 2010). However it was shown recently that XPD from *Sulfolobus acidocaldarius* (saXPD) is not stalled in the presence of DNA lesions which are removed by the NER pathway (Rudolf et al. 2010),whereas XPD from *Ferroplasma acidarmanus*(faXPD) is stalled at a CPD lesion during 5'-3'movement and forms a tight complex with DNA (Mathieu et al. 2010).Furthermore faXPD's ATPase activity is stimulated in the presence of a CPD containing oligonucleotide (Mathieu et al. 2010).

1.4 Diseases caused by mutations in XPD

There are several genetic diseases linked to mutations in genes involved in GGR and TCR, namely xeroderma pigmentosum (XP), Cockayne's syndrome (CS) and trichthiodystrophy (TTD)(Lehmann 2001)(Figure 1.4).

XP is a rare autosomal recessive genetic disorder and was the first DNA-repair disorder to be identified (Von Hebra 1874). XP is characterized by skin atrophy, pigmentation abnormalities, a more than 2000-fold increased incidence of skin cancer and neurological abnormalities caused by premature neuronal death (Kraemer et al. 1987), (Rapin et al. 2000). XP results from a deficiency in NER at the global genomic level (Cleaver 1968), (Setlow et al. 1969). XPD-mutations leading to XP are located at the C-terminus of the protein, i.e. R683W and R683Q, or in the helicase motifs, for example G47R, T76A and D234N (Figure 4). However

XP can also be caused by mutations in other genes. There are8 complementation groups in XP: XP-A through XP-G plus a variant group XP-V (Cleaver 1968), (Setlow et al. 1969), (Masutani et al. 1999).

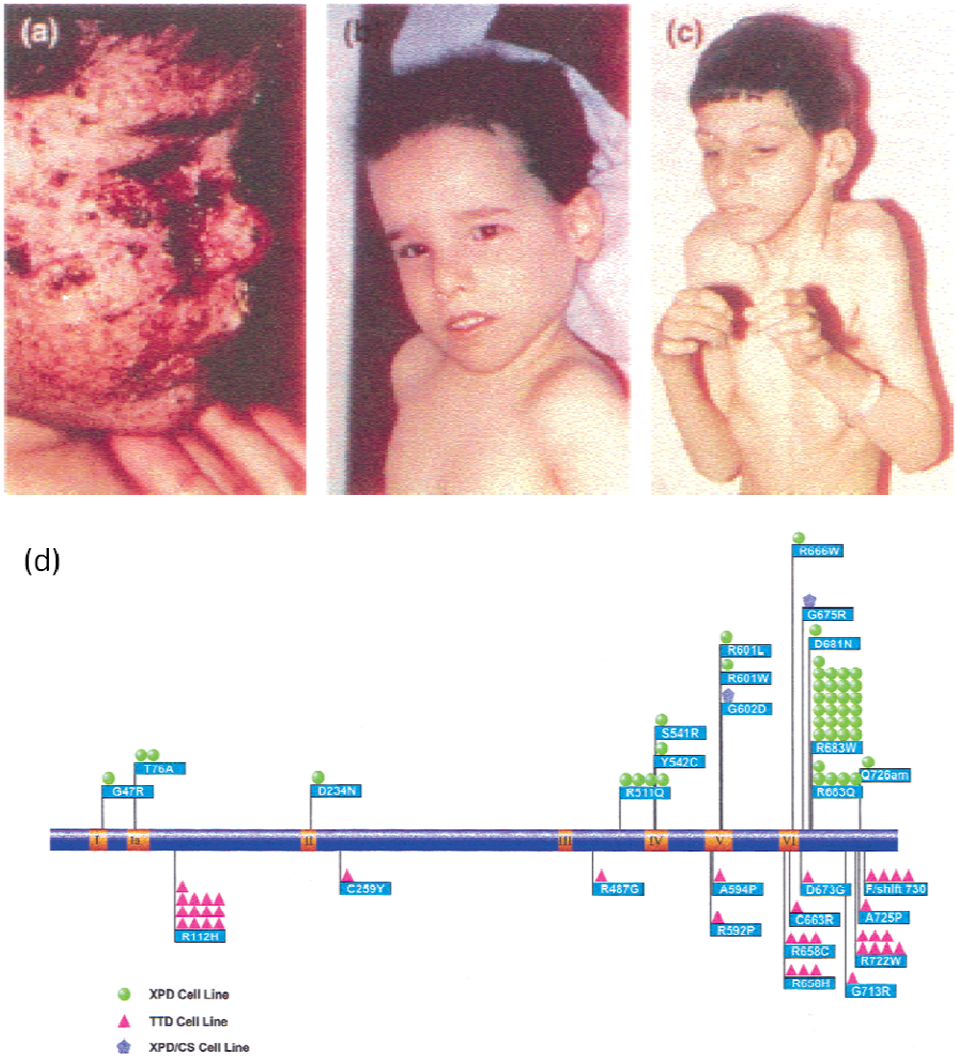


Figure 1.4: Diseases caused by mutations in XPD. Patients with XP (a), TTD (b) and CS (c). (d) Distribution of mutations in XP-D patients with different disorders. The XPD protein is shown as a bar with the seven helicase domains indicated in orange. The flags denote the amino acid changes in different patients, with the number of symbols indicating the number of times that the particular allele has been identified in patients (i.e., two symbols for homozygotes, one for compound heterozygotes). (Green circles) XP, (magenta triangles) TTD, (purple pentagons) XP/CS.Reprinted by permission fromCSHL Press, Genes & Development(Lehmann, 2001) ©2001.

CS was reported by Edward Alfred Cockayne in 1936 (Cockayne 1936). CS is a multisystem disorder characterized by mental retardation, microcephaly, retinal and skeletal abnormalities and severe cachectic dwarfism (Nance and Berry 1992), (Rapin et al. 2000). Since XP patients do not display several of the

symptoms observed in CS, the CS phenotype cannot be explained by a defect in TCR alone. However, it has been suggested that CS cells have a mild impairment in transcription on a more global level (Gillet and Scharer 2006) , (van Gool et al. 1997). Mutations in XPD which cause XP/CS are found in the C-terminal part of the protein, for example G575R and G602D (Figure 1.4).

TTD is characterized by sulfur-deficient brittle hair, caused by a greatly reduced content of cysteine-rich matrix proteins in the hairshafts (Gillespie and Marshall 1983). Other symptoms are mental retardation, unusual facies, ichthyotic skin and a reduced stature (Itin et al. 2001), (Itin and Pittelkow 1990). Many, but not all patients are sensitive to sunlight. However there are no reports on unusual pigmentation changes or cancer in TTD patients (Lehmann 2001). TTD is now known to correlate with mutations in genes involved in NER, namely *XPB*, *XPD* and *TTDA* which are subunits of TFIIH (Weeda et al. 1997), (Stefanini et al. 1986), (Giglia-Mari et al. 2004). Given that TFIIH operates in transcription as well as in NER it was suggested that TTD may result from a subtle defect in transcription, rather than from a deficiency in NER(Bootsma and Hoeijmakers 1993) , (Vermeulen et al. 1994). Mutations leading to TTD are mostly found at the C-Terminus of XPD with the exception of R112H and C259Y which are located in the N-terminal part (Figure 1.4).

1.5 Iron-sulfur clusters

In contemporary organisms, iron-sulfur (Fe-S) clusters are perhaps the most abundant and the most diversely employed enzymatic cofactors (Imlay 2006). The simplest FeS centre is comprised of a single iron atom liganded within a polypeptide by four cysteine residues. The more common FeS clusters have two, three or four iron atoms coordinated to polypeptide residues and bridged by inorganic sulfide (Figure 1.5)(Imlay 2006).

These clusters can fulfill different roles, for example iron storage, structural roles, electron transport(Boal et al. 2005), substrate binding (Lauble et al. 1992) and regulation of gene expression (Kiley and Beinert 2003), (Flint and Allen 1996), (Imlay 2006).

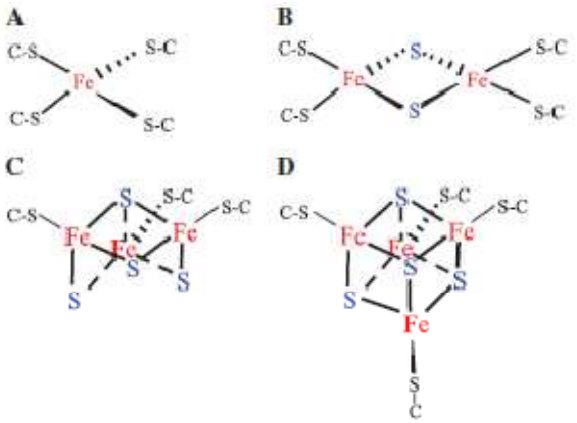


Figure 1.5: Structure of common iron-sulfur centers and clusters. **(A)** Rubredoxin-style iron centre. **(B)** [2Fe-2S] cluster. **(C)** [3Fe-4S] cluster. **(D)** [4Fe-4S] cluster. Reprinted by permission from John Wiley & Sons, Inc.: Molecular Microbiology(Imlay, 2006) ©2006

Recently, in several DNA binding proteins FeS clusters were identified, for example in eukaryotic Primase (Vaithiyalingam et al. 2010), (Weiner et al. 2007), (Klinge et al. 2007)in the helicases DinG (Ren et al. 2009), XPD and FancJ (Rudolf et al. 2006) as well as in DNA-glycosylases (Hinks et al. 2002) and in archaeal RNA polymerase (Hirata et al. 2008), which raises the question of the role of the iron sulfur cluster in these proteins.

MutY is a DNA-glycosylase which contains a 4Fe-4S cluster and functions in base excision repair. The data from intensive studies indicate that the region supported by the cluster is intimately involved in DNA binding (Guan et al. 1998), and that such binding interactions impact catalysis of base removal (Porello et al. 1998), (Chepanoske et al. 2000). Recent evidence has also established that binding of MutY to DNA facilitates oxidation of the $[4Fe\text{-}4S]^{2+}$ cluster to the $[4Fe\text{-}4S]^{3+}$ form. Based on this DNA mediated redox behavior, it has been suggested that this property may be used to enhance the activity of these enzymes by facilitating damaged DNA location (Lukianova and David 2005), (Boon et al. 2003).

1.6 Archaeal homologs: a model for eukaryotic NER

Archaea were, besides prokaryotes and eukaryotes, identified as the third kingdom of life (Woese and Fox 1977). Based on a ribosomal phylogenetic analysis, archaea are subdivided into two groups, the euryarchaea and the crenarchaea (Cox et al. 2008). Because of the advantages archaeal proteins have and because of the high sequence homology to human XPD we used the archaea

Thermoplasma acidophilum as a model organism. Until its genome was sequenced Thermoplasma was thought to be a eukaryotic organism (Ruepp et al. 2000).

Table 1.1: Distribution of selected DNA repair proteins in sequenced archaeal genomes.

Species	XPF	XPB	XPD	Fen1	Dpo4	Photolyase	UvrABC	MutS/MutL	Temp (°C)
Aeropyrum pernix	1	1	1	1					95
Pyrobaculum aerophilum	1	2	1	1					95
Sulfolobus solfataricus	1	2	1	1	1	1			80
Sulfolobus tokodaii	1	2	2	1	1	1			80
Sulfolobus acidocaldarius	1	2	1	1	1	1			75
Methanopyrus kandleri	1			1					100
Pyrococcus furiosus	1	2	1	1					100
Pyrococcus horikoshii	1	2	1	1					100
Pyrococcus abyssi	1	2	1	1					100
Thermococcus kodakarensis	1	1	1	1					85
Nanoarchaeum equitans	1	1	1	1					85
Methanococcus jannaschii	1		1	1					85
Archaeoglobus fulgidus	1	1		1					85
Methanothermobacter thermautotrophicus	1			1		1	•		70
Picrophilus torridus		1	1	1	1	1			60
Thermoplasma acidophilum		1	1	1					60
Thermoplasma volcanium		1	1	1					60
Methanosarcina acetivorans	1	1	1	1	1	1	•	•	<40
Methanosarcina mazei	1	1	1	1	1	1	•	•	<40
Haloarcula marismortui	1	2		1	1	3	•	•	<40

Crenarchaea are at the top, followed by euryarchaeal species, ranked by growth temperature. The numbers show the presence and number of each repair gene. A circle indicates that bacterial repair genes are present in archaeal genomes (Reprinted by permission fromElsevier: Current Opinion in Microbiology(Kelman & White, 2005), ©2005.

Especially for hyperthermophilic archaea the question arises which mechanisms are used to maintain genomic stability (Grogan 2000), (Grogan 2004). It is still an open question whether or not there is NER in archaea. The bacterial UvrABC machinery is only found in certain mesophilic archaea (Table 1.1), these instances probably reflect a lateral gene transfer event from a bacterial donor (Grogan 2000). For the majority of archaea a different pathway must be assumed. Most archaea contain several homologues of the eukaryotic NER nucleases XPF/ERCC1 and Fen1 and the helicases XPB and XPD (Table 1.1), raising the possibility that a eukaryal-type NER pathway is functional in archaea (Kelman and White 2005). However the NER proteins XPA and XPC are not present. Furthermore several archaeal genomes lack one or more of the eukaryal-type NER enzymes, for example *Methanopyrus kandleri* and *Methanothermobacterthermoautotrophicus* lack XPB and XPD. The presence of different NER proteins does not guarantee NER functionality of the present enzymes (Rouillon and White 2011). In addition all of these proteins can perform alternative functions *in vivo*.

Furthermore there is no published evidence for a NER-type patch repair pathway in archaea other than in *Methanothermobacterthermautotrophicus*(Ogrunc et al. 1998), which contains a complete set of UvrABC genes. *Thermoplasma acidophilum* lacks the UvrABC genes as well as XPF (Table 1.1), however it contains a clear homologue of XPD. XPD from *Sulfolobus acidocaldarius* was shown to be a monomer in solution (Rudolf et al. 2006), but lacking a known interaction partner the function of archaeal XPD remains unclear so far(Kelman and White 2005).Other NER proteins also fulfill different roles as in the NER pathway. XPF for example moonlights in several different repair pathways. In mammals XPF is involved in single strand annealing and gene conversion (Al-Minawi et al. 2008). Furthermore XPF limits L1 retrotransposition (Gasior et al. 2008). Fen1 on the other hand is involved in long patch BER and Okazaki fragment processing (Kelman and White 2005). These observations demonstrate multi-cellular functions for NER proteins indicating that these proteins not necessary fulfill NER roles.

2 Materials

2.1 Chemicals

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

2.2 Equipment, instrumentation and expendable materials

Table 2.1: Instruments used in this work

Equipment	
Autoclave Systec V-150	Systec
Bacteria shaker ISF-1-W	Kühner
Balances XS 105 Dual Range XS 6002S Dual Range	 Mettler Toledo
Bio-Photometer	Eppendorf
Camera AxioCam MRC	Zeiss
CD-Spectropolarimeter J-810	Jasco
Cell disruptor	Constant Systems
Centrifuges Centrifuge 5415 D Centrifuge 5810 R Avanti J-26 XP	 Eppendorf Eppendorf Beckmann Coulter
Chromatography-system Äkta purifier	GE Healthcare
Columns for affinitychromatography Econo-Column 1,5 x 15 cm Econo-Column 2,5 x 20 cm	 Biorad
Columnmaterial for affinitychromatography Ni-NTA Agarose	 Invitrogen
Cryo loop	Hampton Research
Flouromax4 fluorescence spectrometer	Horiba Jobin Yvon
Gelelectophoresis Mini-Protean 3-cell	Biorad
Gelfiltration HiLoad 26/60 Superdex 200 prep grade	GE Healthcare
HoneyBee 963 crystallization robot	Zinser Analytik
Illumination table	Roth

Incubator	Heraeus
Lightsource KL 2500 LCD	Zeiss
Lissy liquid handling robot	Zinser Analytik
Magnetic stirrer MR 3002	Heidolph
Microscope	Zeiss
Nanodrop ND 1000 Spectrophotometer	Peqlab
Octet RED System	Forté Bio
PCR Cycler	Eppendorf
PharosFX imager system	BioRad
pH-Electrode	Schott
pH-Meter	Schott
Pipettes	Rainin
Pippetboy easypet	Eppendorf
Power supply	Biorad
Real time PCR cycler (MX3005P)	Stratagene
Rotors JLA 25.50 JA 8.1000	Beckmann Coulter
Spectrophotometer	Agilent
Streptavidin tips	Forté Bio
Thermoblock Rotilabo-Block Heater 250	Roth
Thermomixer	Eppendorf
Expendable Materials	
24-well plate	Crystalgen
96-well plate	Greiner
Concentrator MWCO 30,000 Vivaspin15 Vivaspin 20	Sartorius Stedim Biotech
Cover slides	Jenabioscience
Filter 0,22 µm	Roth
Greinertube 15 ml, 50 ml	Greiner
Pipette tips 10 µl, 200 µl, 1000 µl, 5000 µl	Rainin
Reactiontube 1.5 ml	Eppendorf
UVette	Eppendorf

2.3 Bacterial cell culture

2.3.1 Bacterial strains

Table 2.2: Bacteria strains used in this work

Name	Genotype	Source
BL21-CodonPlus(DE3)-RIL	*E.coli* B F-*ompT hsdS*(r_B-m_B-) *dcm*+ Tet^r*gal* *I* (DE3) *endA* Hte [*argU ileY leuW* 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

2.3.2 Expression plasmids

Table 2.3: Expression plasmids used in this work

Name	Vector	Tag	Source
taXPD (T3)	pET16b	N-terminal His	J. Truglio
taXPD (T4)	pET16b	N-terminal His	J. Truglio
taXPD	pBadM11	N-terminal His	F. Rohleder
taXPD mutants	pET16b	N-terminal His	
faXPD	pET28a	N-terminal His	N. Mathieu

2.3.3 Media for bacterial cell culture

Table 2.4: Media for bacterial cell culture used in this work

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

2.3.4 Media additives

Table 2.5: Media additives used in this work

Substance	Stocksolution	Final concentration
Ampicillin	100 mg/ml in ddH_2O	100 µg/ml
Chloramphenicol	34 mg/ml in 70 % Ethanol	34 µg/ml
IPTG	1 M	0.1 mM
Arabinose	20 %	0.02 %

2.4 Enzymes

Table 2.6: Enzymes used in this work

Enzyme	Source
DpnI	New England Biolabs
Phusion DNA Polymerase	Finnzymes
T4 polynucleotide kinase	Invitrogen
Desoxyribonuclease I	Invitrogen
Proteinase K	Roth

2.5 Nucleotides and oligonucleotides

Table 2.7 Nucleotides and oligonucleotides used in this work

Nucleotide/ Oligonucleotide	Source	Purpose
ATP	Sigma-Aldrich	Helicase assay, ATPase assay, Crystallization
[γ^{32}P]-ATP	Hartmann Analytic	Helicase assay
AMPPNP	Sigma-Aldrich	Crystallization
8I-AMPPNP	Jena Bioscience	Crystallization
Br-AMPPNP	Jeno Bioscience	Crystallization
10 mer	CGAGCACTGC	Crystallization
22mer	CCCAGTACGACGGCCAGTGCGC	Crystallization
50mer	GCTCGAGTCTAGACTGCAGTTGAGAGCTTGCT AGGACGGATCCCTCGAGG	ATPase assay
BGAAF	ATTGCGGATTTTTTGGAGCCGAAGGCTCCATC**F** CTACCGCAAT	Crystallization
CPD_biotin	GACTACGTACTGTTACGGCTCCATC**X**CTACCGC AATCAGGCCAGATCTGC-biotinX=CPD	Octet

F1522	CGTACTGTTACGGC**F**CCATCTC	EPR
F2140	CGTACTGTTACGGCTCCATC**F**CTACCGCAATCAGGCCAGA	EPR
F2650_biotin	GACTACGTACTGTTACGGCTCCATC**F**CTACCGCAATCAGGCCAGATCTGC-biotin	Octet
K10T	GAGCTTAATTCCTTTTTTTTTT	Crystallization
CPD_22	GGATGGCTTAGAGC**X**AATTCC X=CPD	Crystallization
CPD_29	CCTTGGTCCATGGCTTAGAGC**X**AATTCC X=CPD	Crystallization
MJD1	Cy3-GACTACGTACTGTTACGGCTCCATC	Helicase assay
NDB	GCAGATCTGGCCTGATTGCGGTAGAGATGGAGCCGTAACAGTACGTAGTC	Helicase assay
NDB8	GCAGATCTGGCCTGATTGCGCCTTTTTCGGAGCCGTAACAGTACGTAGTC	Octet
NDB16	GCAGATCTGGCCTGATTGCGCCTTTTTACCTTCCGTAACAGTACGTAGTC	Octet
NDB24	TCTGGCCTGATTGCGGTAGAGATG	EPR
NDB40	TCTGGCCTGATTGCGGTAGAGATGGAGCCGTAACAGTACG	EPR
NDT_biotin	GACTACGTACTGTTACGGCTCCATCTCTACCGCAATCAGGCCAGATCTGC-biotin	Octet
U_K10T	TTTTTTTTTTGGAATTAAGCTC	Crystallization
XPD_bottom	Dabcyl-AGCTATGACCATGATTACGAATTGCTTGGAATCCTGACGAACTGTAG	Helicase assay
XPD_bottom_gegen18	GACCATGATTACGAATTG	Helicase assay
XPD DNA 1	CGAGCACTGC	Crystallization
XPD DNA 2	TTTTTTTGCAGTGCTCG	Crystallization
XPD_top	AGCTACCATGCCTGCACGAATTAAGCAATTCGTAATCATGGTCATAGCT-Cy3	Helicase assay

Annealing of F1522/NDB40, F2140/NDB24, MJD1/NDB and XPD DNA1/XPD DNA2 results in 5'-overhang substrates and annealing of NDT/NDB8, NDT/NDB16, CPD/NDB8, CPD/NDB16, F2650/NDB8 and F2650/NDB16 in bubble substrates, with 8 and 16 bases being mismatched, respectively. Annealing of XPD_top/XPD_bottomand K10T/U_K10T results in open fork substrates.

For sequences of mutagenesis primers see apendix.

2.6 Buffer and solutions

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

Cell lysis and protein purification

Lysis buffer	**Elution buffer**	**SEC buffer**
20 mM Tris pH 8	20 mM Tris pH 8	20 mM Tris pH 8
500 mM NaCl	500 mM NaCl	300 mM NaCl
5 mM Imidazole	500 mM Imidazole	10 mM $MgCl_2$

Reaction buffers

Octet Reaction buffer	**ATPase reaction buffer**	**Helicase reaction buffer**
20 mM Tris pH 8	60 mM HEPES pH 7.2	20 mM Tris pH 8.5
150 mM NaCl	10 mM KCl	10 mM NaCl
10 mM $MgCl_2$	20 mM $MgCl_2$	5 mM $MgCl_2$
1 mg/ml BSA	0.5 mM DTT	1 mM EDTA
	0.5 mM $CaCl_2$	2 mM DTT

Additional buffers

Staining solution	**Destaining solution**	**Running buffer (1x)**
50 % Methanol	10 % Methanol	192 mM Glycine
10 % Acetic acid	5 % Acetic acid	0.1 % SDS
0.1 % Coomassie brilliant blue		25 mM Tris

Loading buffer	**15 % Resolving Gel**	**5 % Stacking Gel**
50 mM Tris pH 6.8	15% AA/ BAA	15% AA/ BAA
100 mM DTT	375 mM Tris pH 8.8	125 mM Tris pH 6.8
2 % SDS	0.1 % SDS	0.1 % SDS
0.1 % Bromphenol blue	0.25 % APS	0.25 % APS
	0.05 % TEMED	0.15 % TEMED

1x TE buffer	**8% native Gel**	**1x Trisglycine buffer**
10mM Tris pH 8	8% AA/ BAA	25 mM Tris base
1mM EDTA	0.5x Trisglycine pH 8	192 mM Glycine
	0.2 % APS	
	0.1 % TEMED	

2.7 Oxidants and reducing agents

Table 2.8: Oxidants and reducing agents used in this work

Chemical	Source	Redox potential
Sodium dithionite	Roth	E_{ox} = -460 mV
$Ir(Cl)_6^{2-}$	Lisa Engstrom	E_{ox} = 892 mV
$K_3Fe(CN)_6$	Roth	E_{ox} = 424 mV
Potassiumsuperoxide	Acros Organics	
Peroxynitrite	Lisa Engstrom	
NO	Lisa Engstrom	
H_2O_2	Roth	E_{ox} = 692 mV
$Co(phen)_3^{3+}$	Lisa Engstrom	E_{ox} = 370 mV
$Co(5\text{-}Cl\text{-}phen)_3^{3+}$	Lisa Engstrom	E_{ox} = 430 mV

2.8 Kits

NucleoBond Plasmid Kit — Macherey & Nagel

2.9 Crystallization screens

Table 2.9: Crystallizationscreens used in this work

Name	Source
Additive Screen	Hampton Research
Crystal Screen I, II	Hampton Research
Index Screen HT	Hampton Research
JCSG Screen	Qiagen
Nextal PEG Suite	Qiagen
Nucleix Suite	Qiagen
Optimix 1 - 5	Topaz

Protein Complex Screen	Qiagen
Wizard Screen 1	Emerald BioSystems
Wizard Screen 2	Emerald BioSystems

2.10 Computersoftware

Table 2.10: Software used in this work

Software	Author/Reference
Arp/Warp	(Morris et al. 2003)
CCP4 program suite	(1994)
Coot	(Emsley and Cowtan 2004)
ESyPred3D web server	(Lambert et al. 2002)
ExPASY Proteomics Server	http://www.expasy.ch/
Mosflm	(Battye et al. 2011)
O	(Jones et al. 1991)
Octet RED Software Analysis 6.3	Forté Bio
OriginPro 8.5	Origin Lab Corporation
PDB	www.pdb.org
Phaser	(McCoy et al. 2007)
Phenix.refine	(Adams et al. 2010)
Pirate	(1994)
Pymol	DeLano 2002
Quantity One (Molecular Imager)	Biorad
Refmac5	(Murshudov et al. 1997)
Scala	(Evans 2006)
Sharp	(De La Fortelle and Bricogne 1997)
ShelxD	(Schneider and Sheldrick 2002)
Solomon	(Abrahams and Leslie 1996)

3 Methods

3.1 Molecular biology techniques

3.1.1 Mutagenesis

XPD mutants were generated by site-directed mutagenesis. Two complementary primers with a length between 20 and 40 bases 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 and Malcolm 1999). The single reactions including forward and reverse primers were combined and 25 additional PCR cycles were performed.

The reactions were composed of the following reagents:

Table 3.1: Contents of PCR reaction

	Stock concentration	Final concentration	Amounts for 25 µl reaction
Reaction buffer	5 x	1 x	5 µl
Template	100 ng/ µl	100 ng	1 µl
Primer	10 µM	400 nM	1 µl
dNTPs	10 mM	400 µM	1 µl
Phusion Polymerase			0.5 µl
ddH_2O			16.5 µl

The following PCR program was used:

Table 3.2: PCR program used in this work.

Step	Temperature	Time	Cycles
Initial Denaturation	98 ℃	30 sec	
Denaturation	98 ℃	10 sec	8 cycles for each primer reaction, 25 cycles for combined reaction
Annealing	65 ℃	30 sec	
Extension	72 ℃	90 sec	
Final Extension	72 ℃	5 min	
cooling	4 ℃	forever	

Afterwards, 10 U of *Dpn*I were added to each reaction tube and incubated for 1h at 37ºC. 10 µl of the digested DNA were transformed into DH5αcompetent cells, plated on an ampicillin *Luria Broth* (LB) agar plate and incubated overnight at 37ºC. DNA was extracted from single colonies and checked for the introduced

mutation using DNA sequence analysis at Seqlab – Sequence Laboratories Göttingen GmbH (Göttingen).

3.1.2 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℃, 200 rpm over night the suspension was centrifuged at 4,000 rpm for 10 min (Centrifuge 5810R, Eppendorf) to pellet the cells. The cells were then lysed and DNA was extracted using the NucleoBond Plasmid Kit (Macherey-Nagel) according to the manual.

3.1.3 Labeling and duplexing of single-stranded DNA

For ^{32}P-labeling, ssDNA was 5'-end labeled for 10 min at 37℃ in a 25 µl reaction solution containing 200 nM ssDNA, 1 µCi/µl [γ^{32}P]-ATP (Hartmann Analytic), 1x forward buffer and 10 U T4 Polynucleotide Kinase (NEB). SsDNA was purified using MicroSpin^{TM} G25 columns (GE Healthcare) to remove unincorporated [γ^{32}P]-ATP and subsequently annealed with the counter oligonucleotide in the presence of 0.1x TE buffer containing 100 mM KCl by heating the sample to 80℃ and subsequent cooling slowly to RT. Fluorescently labeled ssDNA was purchased from Biomers.

3.1.4 Transformation

1 µl of plasmid DNA (100 ng/µl) was added to an aliquot (100 µl) of chemo-competent *Escherichia coli* (*E.coli*) BL21-CodonPlus (DE3)-RIL cells (Stratagene). The cells wereincubated on ice for 30 minutes and after this heat-shocked at 42ºC for 90 sec. The cells were then put on ice for 2 min. After adding1 mL*Luria Broth* (LB) medium, the cells were incubated at 37ºC for 1h (600 rpm; Eppendorf Thermomixer). Afterwards the cellswere plated on a LB-agar plate with the appropriate antibiotics and incubated overnight at 37ºC.

The antibiotics ampicillin (Amp) and chloramphenicol (Cam) were used for LB-agar plates and liquid cultures in the concentrations 100 µg/mL and 34 µg/mL, respectively

3.2. Protein expression

The 200 mLLB-medium with the appropriate antibiotics Amp/Cam was inoculated with one colony of a fresh transformation and was incubated overnight (37℃, 200 rpm). For large scale expression 2 L *Luria Broth* (LB) Media (Amp/Cam) were inoculated with 20 mL of an overnight-culture. The culture was incubated (37℃, 200 rpm) to an OD_{600} of 0.5 to 0.6 and at this point protein expression was induced. For Induction 0.1 mM isopropyl-β-thiogalactoside (IPTG) (pet16b, pet28a) or 0.02% Arabinose (pBadM11) was added. Upon induction the temperature was decreased to 14℃ and after 18 to 20h the cells wer e harvested by Centrifugation (Rotor JLA 8.1000, 4℃, 5000g, 20 min; Beckmann Cou lter Avanti Centrifuge J-26 XP) and stored at -80℃.

3.3 Protein purification

3.3.1 Cell lysis

The cells were resuspended in Lysisbuffer (20 mM TrisHCl pH 8, 500 mM NaCl, 5 mM imidazole) and 5 µL of DNase I was added. The cells were lysed by passing the cell suspension twice through the cell disruptor at a pressure setting of 1.8 kbar. Afterwards the celllysate was centrifuged to separate soluble proteins from insoluble proteins and celldebris (Rotor JA 25.50, 4℃, 30000g, 1 h; Beckmann Coulter Avanti Centrifuge J-26 XP). The soluble expressed protein in the supernatant was used for protein purification.

Proteinpurification was performed in two steps. First affinity chromatography was applied followed bysize exclusion chromatography.

3.3.2 Affinity chromatography – Ni $^{2+}$ column

Affinity chromatography is based on properties of biological macromolecules to specifically and reversible bind other substances (ligands). Polyhistidine-tags exert a strong affinity to Ni^{2+}-ions. The matrix of the Ni^{2+}-column consists of Ni-NTA agarose (Invitrogen) that is charged with Ni^{2+}-Ions. Proteins containing a His_6-or His_{10}-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.

XPD contains as affinity-tag polyhistidine residues that bind nickel-ions with high affinity. The six (pET28a) or ten (pET16b) histidine residues are located at the N-terminus of the protein.

First the column was equilibrated with 5 column volumes (CV = 5 ml) of Lysisbuffer (20 mM TrisHCl pH 8, 500 mM NaCl, 5 mM imidazole). Afterwards the proteinsolution was applied to the column. To eliminate unbound protein the column was washed with 10 CV of Lysisbuffer. The bound protein was eluted with Elutionbuffer (20 mM TrisHCl pH 8, 500 mM NaCl, 500 mM imidazole). The different fractions were analyzed on a 15% SDS-Gel(Laemmli 1970). The elutionfractions containing the protein were combined and concentrated to 5 mL.

3.3.3 Size exclusion chromatography

Size exclusion chromatography (SEC) was performed on an ÄKTA-Purifier-Chromatography-System (GE Healthcare) equipped with a sample loading pump P960, a Monitor pH/C-900 for measurement of pH, conductivity, and temperature and a UV900 Monitor for UV-Vis detection of up to three wavelengths simultaneously. The fraction collector Frac-950 was connected to the Äkta system and was used for automated collection of the elution fractions.

Via SEC proteins are separated by size and form. The resin of the Superdex 200 XK 26/60 consists of dextran, which is covalently attached to agarose. The columnvolume of Superdex 200 XK 26/60 is 330 ml and proteins with a size between 10 kDa and 600 kDa can be separated. Large proteinmolecules cannot, in contrast to small proteins, enter the particles and therefore elute earlier.

The Superdex 200 XK 26/60 column was equilibrated with SEC buffer (20 mM TrisHCl pH 8, 300 mM NaCl, 10 mM $MgCl_2$). The proteinsolution was applied with a sample loop to the column and the SEC was performed with a flow of 1 ml/min for 1.2 CV. UV-detectors for various wavelengths were applied. Nucleic acids were detected at 260 nm, protein at 280 nm and the iron sulfur cluster of XPD at 410 nm. The 280 nm peak fractions were checked on a 15% SDS-Gel(Laemmli 1970).

3.4 Protein characterization

3.4.1 SDS-PAGE analysis

Via a sodium dodecyl sulfate (SDS) polyacrylamid gel electrophoresis (SDS PAGE) proteins are separated according totheir molecular weight under denaturing conditions. SDS is a strong anionic detergent thatbinds to the main chain of the proteins so that the charge of the SDS-protein complexes isabout proportional to the molecular weight of the protein. Additionally, a reducing agent is used to disruptpossible disulfide bonds. Proteins are thereby separated in an electrical field according to their size,independently of their natural charge or conformation.

After each purification step, an SDS PAGE (Laemmli 1970)was run, to identify the fractionscontaining the protein and to analyze the grade of purity. For the estimation of the molecularweight of the different protein bands, the PageRuler prestained protein ladder (Fermentas) thatshows bands between 10 and 170 kDa was loaded in one of the lanes of the gel. Beforeloading on the gel, 10 µL of the protein sample was mixed with 3 µL 5x protein samplebuffer, heated for 5 min at 95° C and was loaded onto the gel. The run wasperfo rmed for 45 min at 200 V. The gel was stained for 20 min in coomassie staining solution (see 2.6) and the unbound stain was removed afterwards with destaining solution (see 2.6) (Laemmli 1970).

3.4.2 Determination of the protein concentration

The concentration of the purified proteins was determined with a nanodrop UV/Vis spectrophotometer (Peqlab) using the spectrum mode from 175 nm to 875 nm. Via the Lambert-Beer law one can calculate the protein concentration c from the

extinction coefficient ε, the absorbance at 280 nm A and the thickness of the cuvette d (Pfeiffer and Liebhafsky 1951).

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

The extinction coefficient can be defined with the help of the number of the tryptophan-, tyrosine- and cysteine-Residues applying the following formula:

$$\varepsilon_{280} = (n_{Trp} \cdot 5690 + n_{Tyr} \cdot 1280 + n_{Cys} \cdot 120)$$

The *Thermoplasma acidophilum* XPD (taXPD) monomer containes two tryptophanes, 36 tyrosines and eight cysteines, the calculated extinction coefficient ε_{280} is 65,140 $M^{-1}cm^{-1}$.The molecular weight of taXPD is 72 kDa.

The *Ferroplasma acidarmanus* (faXPD) monomer containessixtryptophanes, 39 tyrosines and sevencysteines, the calculated extinctioncoefficient ε_{280} is 98,415 $M^{-1}cm^{-1}$. The molecular weight of faXPD is 72 kDa.

3.4.3 Circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarized light by chiral molecules. CD spectroscopy can be used to analyze the secondary structure of polypeptides and proteins (Kelly et al. 2005). Here, CD measurements were performed at 20℃ using a Jasco J-810 spectrop olarimeter. Measurements on taXPD and faXPD were carried out in 200 µl volume in 50 mM Na-phosphate pH 7.8 at a protein concentration of 0.1 mg/ml. Spectra were registered from 260 to 190nm and accumulated eight times with a scan speed of 20 nm/min,

3.4.4 ThermoFluor analysis

A ThermoFluor analysis is based on the use of fluorescence measured upon thermal unfolding of the analyzed protein. An environmentally sensitive fluorescent dye that interacts specifically with denatured protein is used as an indicator. A soluble protein contains a hydrophilic surface which makes it possible to interact with the solvent, however the core of the protein is usually hydrophobic. The dye is

quenched in aqueous environments only upon unfolding of the protein the dye can bind to the hydrophobic interior of the protein resulting in a large increase in quantum yield (Matulis et al. 2005).

The midpoint of the protein unfolding transition is defined as the melting temperature T_m at which folded and unfolded states of the protein are equally populated.

Here the ThermoFluor analysis was applied to determine whether the wild-type and XPD variants are folded in a similar way. For the thermal unfolding assays 5 µL of the protein sample in its respective buffer and 5 µL of the SYPRO® Orange dye (Sigma-Aldrich) in DMSO were added to 15 µL buffer in a concentration of 100 mM in a 96 PP-PCR-plate (Greiner Bio-One International AG). The experiment was performed in a Real-Time PCR system (Stratagene Mx3005P™). The program comprises 70 steps with the first step conducted at 25°C. With each subsequent step, the temperature is increased by 1°C and thus the last step's temperature is 95°C. The results analysis was carri ed out with an excel sheet provided by the Structural Genomics Consortium (SGC) in Oxford.

3.4.5 *In vitro* ATPase assay

XPD ATPase activity was measured with a *in vitro* ATPase assay, in which ATP consumption is coupled to the oxidation of NADH via pyruvate kinase and lactate dehydrogenase activity. Activities were measured at 37°C in 50 µL solution, containing 1.5 U pyruvate kinase, 1.9 U lactate dehydrogenase, 2 mM phosphoenolpyruvate and 0.15 mM NADH, 10 mM KCl, 20 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM $CaCl_2$ and 60 mM HEPES (pH 7.2). SsDNA was added at a final concentration of 0.5 µM. Under saturating concentrations of ATP (2 mM) measurements were carried out using XPD wild-type and variants at a concentration range of 200-2000 nM. For catalytic measurements, the mix of all reagents, with the exception of ATP, were pre-incubated at 37°C until a stable base line was achieved. Enzyme catalysis was initiated by the addition of ATP. The activity profiles were measured at 340 nm, using an Agilent 8453 diode array spectrophotometer. The initial velocities were recorded and using the molar

extinction coefficient of NADPH ATP consumption was determined. The measurements were carried out in triplicates using different protein batches.

3.4.6 Biolayer Interferometry

Real time binding assays between ssDNA and purified taXPD wild-type and variants were performed using biolayer interferometry on an Octet RED system (Fortebio, Menlo Park, CA). This system monitors interference of light reflected from the surface of a fiber optic sensor to measure the thickness of molecules bound to the sensor surface. 3'-biotinolyted DNA was obtained from Biomers and coupled to kinetics grade streptavidin biosensors at a concentration of 100 nM. Sensors coated with ssDNA were allowed to bind taXPD in a reaction buffer containing 20 mM Tris (pH8), 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT and 1 mg/ml BSA at different taXPD concentrations ranging from 0.5 to 5 µM. The measurements were carried out in triplicates using different protein batches. Binding kinetics were calculated using the Octet Data Analysis Software 6.3, with a 1:1 binding model to calculate the association rate constants. Binding affinities were calculated as the ratio of dissociation and association rate constants.

3.4.7 Gel-based helicase assay

To investigate DNA unwinding activity, helicase assays were performed by incubation of 500 nM protein, 50 nM DNA substrate and 1 µM competitor oligonucleotide in helicase reaction buffer for 5 min at RT (see 2.6).The reaction was started by addition of 5 mM ATP. Stop buffer containing 10 mM TrisHCl, pH 8.0, 5 mM EDTA, 0.5% SDS, 1mg/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 6x sample buffer (Invitrogen) was added and samples were separated at 100 V on a 8 % native polyacrylamide gel. Gels containing ^{32}P-labeled DNA were exposed to a phosphoimager screen over night, fluorescently labeled DNA comprising gels were directly visualized by the PharosFX^{TM} imager system (BioRad).

3.4.8. Fluorescence based helicase assay

Helicase activity was analyzed utilizing a fluorescence based helicase assay (Martinez-Senac and Webb 2005). An open fork substrate with a cy3 label at the 3'-end of the translocated strand oligonucleotide and a dabcyl modification on the 5'-end of the opposite strand was used. This results in a quenching of the cy3 fluorescence that is removed upon unwinding of the substrate. Assays were carried out in 20 mM Tris (pH 8.5), 10 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA and 2 mM DTT. 800 nM of protein were mixed with 40 nM open fork substrate and 800 nM capture oligonucleotide. The reaction was subsequently started with the addition of 500 µMATP. Kinetics were recorded with a Flouromax4 fluorescence spectrometer (Horiba Jobin Yvon) and monitored until the reaction was completed Fluorescence was detected at an excitation wavelength of 550 nm (slidwidth 2nm) and an emission wavelength of 570 nm (slidwidth 2 nm). Initial velocities were fitted with Origin8 and represent the averages of at least two different reactions and two independent protein batches.

3.4.9 Electron paramagnetic resonance (EPR)

EPR is a spectroscopic technique which detects transitions of unpaired electrons in an applied magnetic field. Like a proton, the electron has a spin, which gives it a magnetic property known as a magnetic moment. The magnetic moment makes the electron behave like a bar magnet.

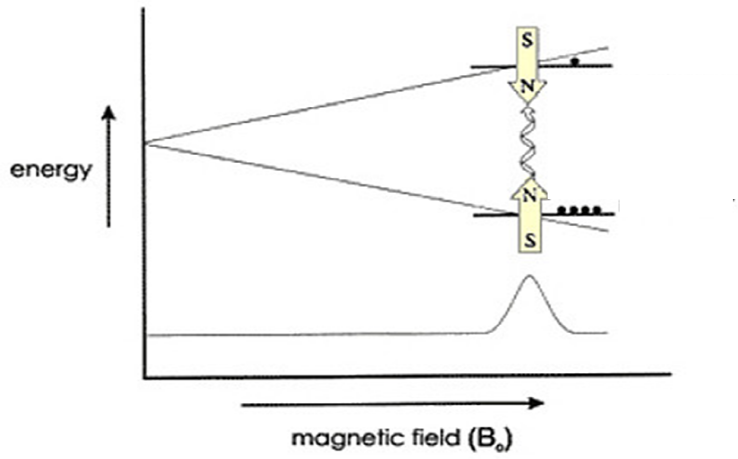


Figure 3.1: **Variation of the spin state energies as a function of the applied magnetic field**taken from (http://www.bruker-biospin.com)© Bruker Corp.

When an external magnetic field is supplied the paramagnetic electrons can either orient in a direction parallel or antiparallel to the direction of the magnetic field. This creates two distinct energy levels for the unpaired electrons and allows measuring them as they are driven between the two levels.

From quantum mechanics the most basic equation of EPR is obtained, where ΔE is the energy difference between the two levels of the electrons, h is the Planck's constant, v is the frequency of the radiation, g is the g-factor, μ_B is the moment of the electron and B_0 is the magnetic field.

$$\Delta E = h \cdot v = g \cdot \mu_B \cdot B_0$$

To measure EPR the sample is transfered into an EPR spectrometer. The spectrometer consists of a magnet, a cavity, a bridge, a signal channel, a field controller and a computer. The sample is placed in the microwave cavity, which is a metal box that helps to amplify weak signals from the sample. The microwave bridge contains the microwave source and the detector. The signal channel houses the required electronics for the phase sensitive detection, a technique used to enhance the sensitivity of the spectrometer. The magnetic field controller makes it possible to sweep the magnetic field in a controlled and precise manner for the EPR experiment.

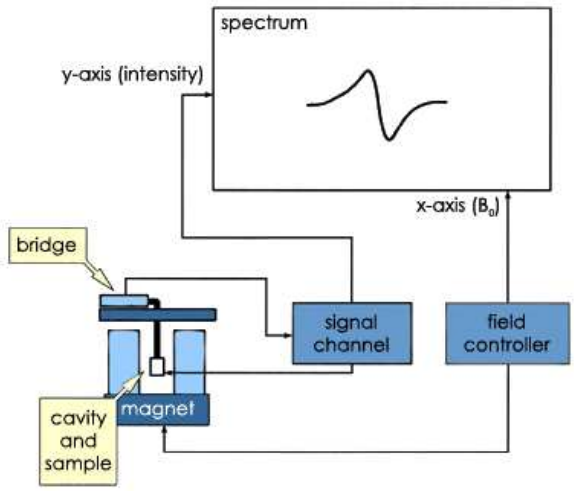


Figure 3.2: **Block diagram of an EPR Spectrometer**taken from (http://www.bruker-biospin.com)© Bruker Corp.

To acquire a spectrum the frequency of the electromagnetic radiation is changed and the amount of radiation which passes through the sample is measured with a

detector to observe the spectroscopic absorptions. Because spectra can be acquired at several different frequencies, the g-factor which is independent of the microwave frequency is a unique character for the identification of a compound. High values of g occur at low magnetic fields and vice versa.

$$g = h \cdot v / (\mu_B \cdot B_0)$$

In these experiments taXPD with or without DNA was incubated for 10 min at 37℃ in the presence of different reactive chemicals, transferred to an EPR tube, and carefully and slowly frozen in liquid nitrogen to allow the air to exhaust of the sample.

3.5 Crystallization

To determine the structure of a protein by x-ray diffraction, a well-ordered single protein crystal is required which is characterized by a consistent homogeneous crystal lattice. Therefore highly pure protein is needed (95% purity on a SDS-gel). Crystallization requires that molecules exchange from a supersaturation state into a solid state. Several techniques have been developed to achieve the supersaturation state: microbatch crystallization, dialysis and vapor diffusion for example. In this work only the vapor diffusion method was used.

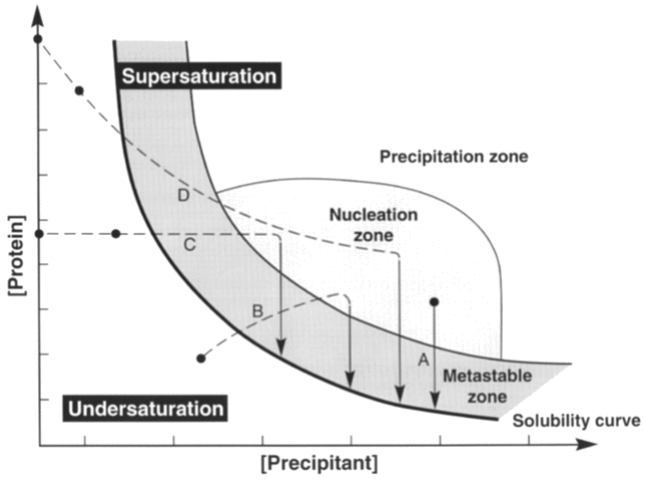


Figure 3.3: **A schematic drawing of a protein crystallization phase diagram.** This is based on two of the most commonly varied parameters, protein and precipitant concentrations. The diagram displays how supersaturation is achieved to trigger crystallization in vapor diffusion (B) and other crystallization methods. (Reprinted with permission by IUCr, Acta Crystallographica Section D: Biological Crystallography(Chayen 1998), ©1998)

For vapor diffusion (Figure 3.3, line B, (Chayen 1998)) the protein solution is combined with crystallization solution containing an adequate precipitant. The precipitant displaces at a certain concentration the protein molecules from the solution. In a sealed reservoir, the drop solution is equilibrated against a much larger reservoir of the crystallization solution. This technique can be performed in two different ways eitherby placing the drop on a shelf that is surrounded by the reservoir solution (sitting drop) or by hanging the drop from a cover slide towards the reservoir solution (hanging drop). In both cases, diffusion of H_2O from the drop to the reservoir slowly increases the precipitant and the protein concentration in the drop, the nucleation barrier is accomplished and small crystal nuclei form. Upon formation of nuclei, the protein concentration in the drop is decreased and the condition in the drop is pushed towards the metastable zone where growth of crystals at the nucleation sites occurs. The sitting drop method was used for initial screening for suitable conditions with the HoneyBee 963 crystallization robot (Zinser Analytik). The drop contained 0.3 µL protein solution and 0.3 µL crystallization solution and was equilibrated against 40 µL reservoir solution in a 96-well crystallization plate (Greiner Bio-One International AG). The hanging drop method was used for the follow-up screens where the drop (1 µL protein solution pipetted on 1 µL crystallization solution) was equilibrated against 1 mL reservoir solution.

3.6 X-ray crystallography

3.6.1 Data collection

An x-ray beam that encounters a crystal lattice of atoms or molecules induces oscillation of the electrons in the lattice at the x-ray frequency. Upon returning to the unexcited state, the electrons emit radiation at the original energy and wavelength in a random direction. The scattered waves can be described as reflections at the lattice planes of the crystal. In most directions the scattered waves add up to zero (destructive interference). When the x-ray beam hits the lattice plane under the glancing angle θ and Bragg's law is satisfied, constructive interference takes place and reflections appear at the detector. The intensity and the position of the reflections are measured in the diffraction experiment. From the position of the reflections the Miller indices h, k, l can be determined.

Bragg's law

$$n \cdot \lambda = 2 \cdot d \cdot \sin\theta$$

The electron density distribution ρ at every point in the unit cell (x,y,z) is calculated by the inverse Fourier transformation over all structure factors F(hkl).

inverse Fourier transformation

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F(hkl)e^{[-2\pi \cdot i(hx+ky+lz)]}$$

The measured intensity (I(hkl)) of a reflection is proportional to the square of the structure factor amplitude |F(hkl)|. To calculate the structure factor F(hkl) from the intensity, the phase φ of the wave is needed. The phase information cannot be determined during data collection and is lost. This fact is termed the phase problem of crystallography. Several methods exist to solve the phase problem. In this work multiwavelength anomalous dispersion and molecular replacement were used.

Calculation of structure factors

$$I(hkl) = |F(hkl)|^2$$

$$F(hkl) = |F(hkl)e^{i \cdot \varphi(hkl)}|$$

I(hkl): Intensity of the scattered waves, F(hkl): structure factor

For data collection, a single crystal was fished with a nylon loop (Hampton Research) out of the mother liquor, transferred into a cryo-protectant and frozen in liquid nitrogen. The cryo-protectant is used to prevent formation of ice crystals when the crystal is frozen and typically contains solvents like glycerol or polyethylene glycol together with the mother liquor of the protein crystal. As the XPD crystals grew in MPD and PEG 400 conditions, MPD and glycerol were chosen as cryo-protectant, respectively. The crystals were transferred from their mother liquor to a crystallization solution containing 5% glycerol. The concentration of the cryo-protectant was increased 5% stepwise up to a final concentration of 30 % glycerol by just dipping the crystal shortly into the particular solution. Crystals from MPD conditions were directly transferred to a cryo-solution

containing 30 % MPD. After one minute in the solution, the crystals were fished out with a cryo-loop and flash frozen in liquid nitrogen.

For data collection, a cryo-loop containing one single crystal was mounted in the beam path and simultaneously cooled by a cryogenic gas stream of nitrogen at 100 K.Diffraction data were collected at a synchrotron. In a synchrotron electrons and positrons are accelerated and then injected into a storage ring where they circulate near the speed of light, guided by bending magnets. Synchrotron radiation is emitted when the charged particles go through the bending magnets, or through wigglers or undulators, which are additional magnetic devices inserted into the straight sections of the ring. The radiation is captured by beam lines and guided to experimental stations that contain instrumentation for diffraction experiments.

3.6.2 Multiwavelength anomalous dispersion (MAD)

MAD utilizes anomalous scattering elements like endogenous iron and selenium that can be incorporated into proteins. When the energy of an incident X-Ray beam is varied across the absorption edge of an anomalous scatterer, there are substantial changes in the real and imaginary components (f' and f") of the X-ray scattering. Thus, qualitatively, multiwavelength anomalous dispersion experiments can be thought of as *in situ* isomorphous replacements generated by the variation in scattering strength that accompanies the change of the wavelength. The MAD technique requires the special properties of synchrotron radiation since data have to be collected at different wavelegths, but it has the advantages that isomorphism is perfect and all data can be measured from a single crystal.

In the case of small molecules, a direct solution of the phase problem is possible from statistical relations among the intensities. For these direct methods, a single native X-ray data set is sufficient, without the use of fragments of known structures or the need to prepare heavy-atom or selenomethionine derivatives, provided that the data are available at atomic resolution. Nowadays, direct methods are used routinely to locate the positions of heavy-atoms in isomorphous replacement or anomalous scatterers in MAD phasing of large proteins at moderate solutions.

In the present work, structure solution was achieved utilizing the anomalous signal of the endogenous Fe belonging to the 4Fe4S cluster at the Fe edge.

3.6.3 Molecular replacement

Molecular replacement is a method to solve the phase problem by using an already available structure with high homology to the unknown structure. This structure,here 2VSF was used as a search model, should have at least 25- 30 % sequence identity and has to be positioned correctly in the unit cell of the unknown structure. This is done by calculating a Patterson function of the measured data of the unknown structure and of the structure of the search model. The advantage of the Patterson function is that no phase information is needed for calculation. In the rotation function the correct orientation of the search model is calculated by rotation of the self-Patterson vectors of the search model over the Patterson map of the unknown structure. The correct orientation of the search model is indicated by a maximum correlation of the intra-atomic vectors (self-vectors) of the two structures. The translation function is used to find the correct position of the oriented search model in the unit cell. Analogous to the rotation function, the translation function consists of superimposing the cross-Patterson vectors from the search model on the Patterson map of the unknown structure until a maximum of the inter-atomic vectors (cross-vectors) indicates the exact position of the search model. Finally, an initial electron density map can be calculated using the structure factor amplitudes of the measured data and the phases of the correctly positioned search model. A six-dimensional search is computationally very intensive (three rotation variables and three translation variables), the six variables are therefore often determined in two steps with the rotation search first following the translation search.

With the known molecular weight of the protein (M_r), the volume of the unit cell (V_{EZ}) and the number of symmetry operators in the space group (z), the number of molecules (n) in the asymmetric unit can be estimated with the method of Matthews.

Calculation of Matthews coefficient

$$V_M = V_{EZ} / M_r \cdot z \cdot n$$

V_M is the packing parameter that describes the crystal volume per protein mass. With V_M the solvent content of the protein crystal can be calculated which is typically between 30 % and 70%.

3.6.4 Model building and refinement

After structure solution, usually an incomplete model is obtained that contains many errors. In order to produce an accurate model, the structure is manually modified in real space fitting the amino acids to the electron density and adding molecules like water, cofactors and DNA. For model building, the programs Coot(Emsley and Cowtan 2004) and O (Jones et al. 1991) were used which are molecular graphics applications with the primary focus on crystallographic macromolecular model building and manipulation. During crystallographic refinement, the model parameters like coordinates and B-factors are processed in the reciprocal space to improve the fit of observed and calculated structure-factor amplitudes. In order that the system is not under-determined, it is important that the number of parameters that define the structure do not exceed the number of observed reflections. This can especially be an issue when working with huge structures or at low resolution. Fortunately, it is possible to improve the data to parameter ratio by adding additional information to the measured intensities and the derived structure factor amplitudes. Restraints are conditions that apply within a certain standard deviation for the model. Information about the molecular geometry (bond lengths and bond angles), planarity of peptide bonds and phenyl groups and atomic distances can be restrained. If more than one molecule is present in the asymmetric unit, they can be related by non-crystallographic symmetry (NCS) and the geometry can be restrained to be similar (NCS-restraints). Restraints increase the amount of data available for refinement. Constraints are conditions that apply exactly for the structure, reducing the number of independent parameters to be refined. The occupancy of one atom is usually

constrained to be exactly 1.0. Displacement parameters can be constrained or the bond lengths and bond angles of hydrogen atoms.

Upon refinement, the phases become more accurate and an improved electron density can be calculated. The program Refmac5 (Murshudov et al. 1997) was used for maximum likelihood refinement. Alternating rounds of model building and automated refinement were carried out to successively improve the atomic model. The success of refinement is evaluated by the R-factor that measures the agreement between the crystallographic model and the measured data. A decreasing R-factor indicates that the model has improved.

Calculation of the R-factor

$$R = \frac{\sum_{hkl}\left||F_o(hkl)| - |F_c(hkl)|\right|}{\sum_{hkl}|F_o(hkl)|}$$

As the R-factor is related to data that is used during refinement, the obtained value might be biased towards the model. To prevent over-fitting of the data a more objective quality indicator, the free R-factor (R_{free}) was introduced (Brunger 1997), (Kleywegt and Jones 1997). For calculation of the R_{free}, a fraction of 5-10% of the data are randomly chosen and not used for the refinement process. After each refinement cycle, the R_{free} is calculated from these data. If after refinement only the R-factors decrease and the R_{free} stays the same or even increases, the data might have been over fitted and the structure has to be carefully reanalyzed. In general, the R-factor is an indicator of the precision of the model while the R_{free} displays more the accuracy of the model.

After interpretation of all features of the electron density map and when the R-factors indicate an appropriate value and convergence, refinement can be considered as complete. To evaluate the quality of the final model parameters like appropriate geometry, deviations of the bond lengths and angles from ideal values and unusual rotamers of amino acids are inspected. However, these parameters are often heavily restrained during refinement or closely monitored during rebuilding and are therefore not independent criteria of the model-quality. An other criterion for model validation is the Ramachandran plot. In a Ramachandran plot, the dihedral angles φ and ψ of the amino acid residues in a protein structure are plotted against each other and show the possible conformations of φ and ψ angles

for a polypeptide. The diagram differentiates between the favored regions, the additional allowed regions and the disallowed regions while exceptions apply only for proline and glycine. If an amino acid is found to lie in the disallowed region (outlier) there are two possibilities: either the conformation indeed is wrong or the residue actually hadopts an unusual conformation. Amino acids with unusual conformations should have well defined electron density and are often found in important regions of the protein structure like the active site.

4 Results

Archaeal organisms often exist in a more hostile environment as compared to eukaryotic organisms. This includes conditions like high temperature, high UV exposure, arid conditions or a very saline environment. Therefore archaeal proteins are often more robust as compared to their eukaryotic homologs. These characteristicsrender them very amenable for cloning and subsequent heterologous expression in *E.coli*. Because of the above mentioned advantages and the strong evolutionary relationship of archaea to eukaryotes we used the archaeon*Thermoplasma acidophilum* as a model organism for proteins involved in eukaryotic NER in particular the XPD helicase.

4.1 Protein expression and purification

4.1.1 Proteinexpression and Purification of XPD WT and variants from *Thermoplasma acidophilum*

The genes encoding XPDs from *Thermoplasma acidophilum*(taXPD) were cloned with variable N-termini (residues 1-622 and 23-622) into the pET16b vector using NdeI and XhoI restriction sites. This resulted in the constructs taXPDT3 and taXPDT4. PET16b::taXPD (either T3 or T4) was transformed into *E.coli* BL21(DE3) RIL cells as described in 3.1.4. One colony was used for the inoculation of the 200 mL overnight culture. 20 mL of the overnight culture were used to inoculate a 2 L cultures for large scale expression. To obtain sufficient amount of the protein large scale expressions of XPD between 16 L and 24 L cell culture were used. The bacteria were grown at 37℃ to an OD_{600} of 0.5 to 0.6 and then induced with 0.1 mM IPTG. Upon induction the temperature was reduced to 14℃ and the cultures were incubated overnight at 2 00 rpm as described in 3.2. The cells were harvested and 1 L of cell culture yielded approximately 3 g wet cell paste. The cell pellets from 16 or 24 L cell culture were resuspended in 200 or 250 mL lysisbuffer. The cells were lysed as described in 3.3.1 with a cell disruptor system and the soluble expressed protein was used for chromatographic purification.

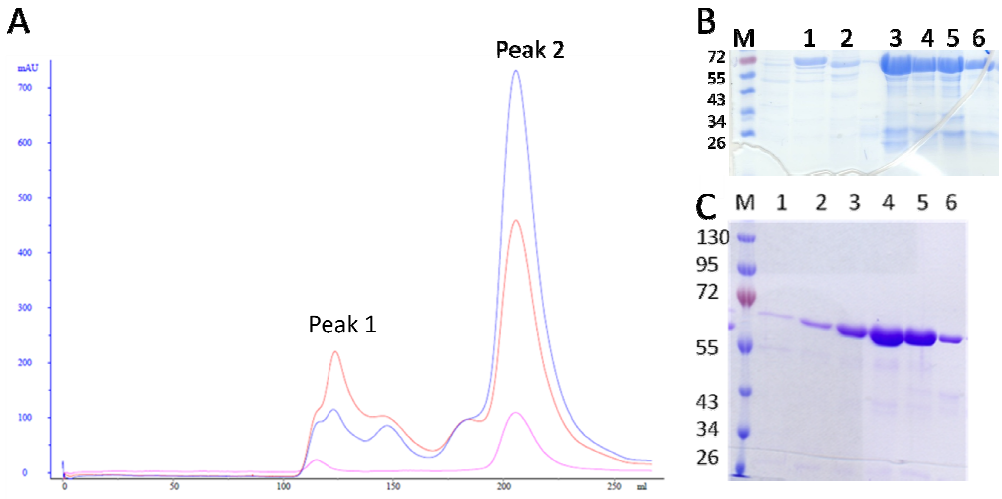


Figure 4.1: Purification of taXPD.(A) Chromatogramm of a size exclusion chromatography run of taXPD with the absorption at 280 nm colored in blue, 260 nm colored in red and 410 nm in pink. **(B)** SDS-PAGE ofNi-column. M indicates the marker, 1 and 2 the wash fractions and 3 to 6 the elution fractions. **(C)** SDS-PAGE after the size exclusion chromatography. With M indicating the marker and 1-6 peak elution fractions of (A). Lane 1 and 2 correspond to peak 1 and 3 to 6 to peak 2.

First the protein was purified via affinitychromatography with a Ni^{2+}-column as indicated in 3.3.2. The column was equilibrated with lysis buffer and afterwards the protein solution was applied. The column was washed with 50 mL of lysis buffer then the protein was eluted with elutionbuffer containing 500 mM imidazole in four fractions each of 5 mL. The input, wash and elution fractions were analyzed on a 15% SDS-PAGE as described in 3.4.1 (Figure 4.1, B). The elution fractions containing the taXPD protein fractions were combined and concentrated with a Vivaspin 20 (MWCO 30,000) to 5 mL.

As a second purification step size exclusion chromatography was performed as described in 3.3.3. The protein was filtered prior to application on the Superdex 200 26/60 column. 5 mL of the protein solution was loaded via a 5 mL loop on the column. The protein was eluted with SEC buffer for 1 CV with a flow rate of 1 mL/min. Afterwards the peak fractions were analyzed on a 15% SDS-PAGE (Figure 4.1, C). As shown in Figure 4.1, B the elution fractions of peak 2 (210 mL, fractions 3-6) contain pure taXPD protein. The chromatogram shows two major peaks at A410 (Figure 4.1, A) with peak one being an aggregation peak (lanes 1and 2, Figure 4.1, C) because it runs at the void volume of the column and peak 2 contains the monomeric taXPD protein (lanes 3-6, Figure 4.1, C).

Elution fractions 3-6 were combined and concentrated with a Vivaspin 20 (MWCO 30,000) to a final concentration of 50 µM. The concentration was determined as depicted in 3.4.2, 50 µL aliquots were flash frozen in liquid nitrogen and stored at -80℃.

All taXPD variants were expressed and purified analogously to taXPD WT.

4.1.2 Proteinexpression and purification of XPD from *Ferroplasma acidarmanus*

The construct for XPD from *Ferroplasma acidarmanus* (faXPD) was obtained from Nadine Mathieu a graduate student in Hanspeter Naegeli's laboratory, Zürich, Switzerland. PET28a::faXPD was transformed into BL21 cells as described in 3.1.4. A single colony was used to inoculate the 200 mL overnight culture. 25 mL of the overnight culture were used to inoculate 2 L cultures for large scale expression. 16 L of cell culture were grown at 37℃ to an OD_{600} of 0.5-0.6 and then induced with 0.1 mM IPTG as described in 3.2. The cells were incubated over night at 14℃ and afterwards har vested. The cell pellet from 16 Lcell culture (50 g) was resuspended in 200 mL of lysis buffer and lysed as depicted in 3.3.1 by a cell disruptor system. The soluble expressed protein was used for further purification.

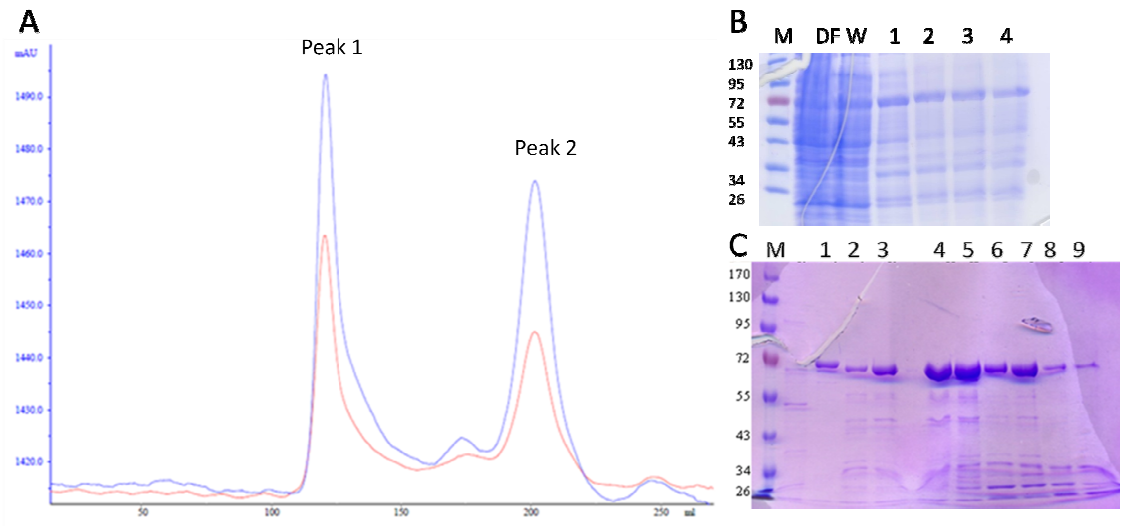


Figure 4.2: Purification of faXPD.(A) Size exclusion chromatogram of faXPD with the absorption at 260 nm colored in red and 280 nm colored in blue. **(B)** 15% SDS-PAGE with samples of the Ni-colum. M indicates the Marker, DF the flowthrough, W the wash fraction and 1 to 4 the elution

fractions. **(C)** 15% SDS-PAGE of the size exclusion chromatography run (A). M indicates the Marker and 1 to 3 the elution fractions of peak 1 and 4 to 9 the elution fractions of peak 2.

The first purification step was an affinitychromatography via a Ni^{2+}-column. The protein solution was applied to a Ni^{2+}-NTA column, washed with 50 mL lysisbuffer and eluted with elution buffer containing 500 mM imidazole in four 5 mL fractions. The wash and elution fractions were analyzed on a 15% SDS-PAGE (Figure 4.2, B). The fractions containing the faXPD protein were combined and concentrated to 5 mL.

Afterwards size exclusion chromatography was applied as described in 3.3.3. The protein solution was filtered and loaded via a 5 mL loop on a Superdex 200 XK 26/60 column. The protein was eluted with SEC buffer for 1 CV with a flow rate of 1 mL/min. The A280 peak fractions were analyzed on a 15% SDS-PAGE. The chromatogram (Figure 4.2, A) shows two major peaks at A280. Peak 1 is most likely an faXPD containing aggregate since the corresponding fractions lanes 1,2 and 3 in Figure 4.2, C contain protein at the size of about 72 kDa which corresponds to faXPD. The faXPD containing fractions 4-9 were pooled and concentrated with a Vivaspin 20 (MWCO 30,000) to a final concentration of 50 µM. The concentration was determined with a Nanodrop spectrophotometer as described in 3.4.2; 50 µL aliquots were flash frozen in liquid nitrogen and stored at -80℃.

4.2 CD spectroscopy reveals that taXPD and faXPD are properly folded

To check whether the two XPDs show an α/β-architecture indicating a proper folding of the proteins CD Spectroscopy was conducted as described in 3.4.3. The CD Spectra (Figure 4.3) reveal that taXPD and faXPD are folded properly in β-Sheets and α-helices indicted by the characteristic minima at 208 nm and 222 nm. The two spectra are quite similar indicating that a similar amount of α-helices and β-sheets are present in the two proteins. Furthermore the spectra show that there are no disordered regions in the proteins indicated by the transition from negative to positive ellipticity values at a wavelength of approximately 200 nm.

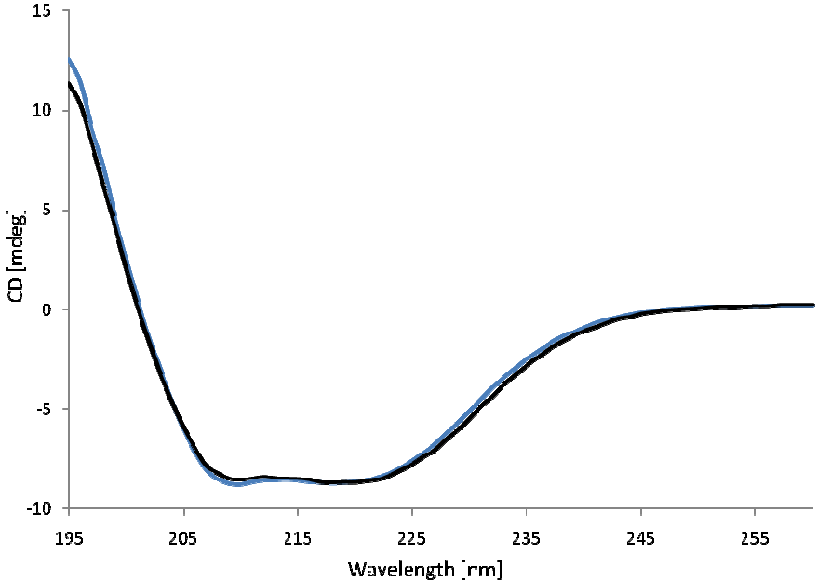


Figure 4.3: CD Spectra of taXPD and faXPD. TaXPD colored in black and faXPD colored in blue.

4.3 Crystallizationexperiments

4.3.1 Crystallization of taXPD

As shown in (Wolski 2007) screens were prepared as described in 3.5 and taXPD (construct taXPDT4) was used at a concentration of 50 µM. After two days in condition 23 of crystal screen I small crystals appeared (Figure 4.4, A). By varying the crystallization condition and by streak-seeding the crystals could be optimized (Figure 4.4, B).

TaXPD crystals were grown by vapor diffusion in hanging drops containing equal volumes of protein in 20 mM Tris/HCl (pH 8.0) and 500 mM NaCl at a concentration of 5 mg/mL, and a reservoir solution consisting of 200 mM $MgCl_2$, 100 mM Hepes (pH 8), and 5%–10% PEG 400 equilibrated against the reservoir solution. The optimized crystals were of hexagonal shape and of yellow-brownish color (Figure 4.4, D). After 7 days the crystal growth was completed and the crystals had a final size of 100 µM x 50 µM x 50 µM. In the initial screens crystals appeared also in crystal screen I in condition 49 (Figure 4.4, C), but optimization of these crystals wasn't successful. TaXPD was also co-crystallized in the presence of ATP and a non-hydrolysable analogue AMPPNP. It was not possible to obtain well diffracting crystals.

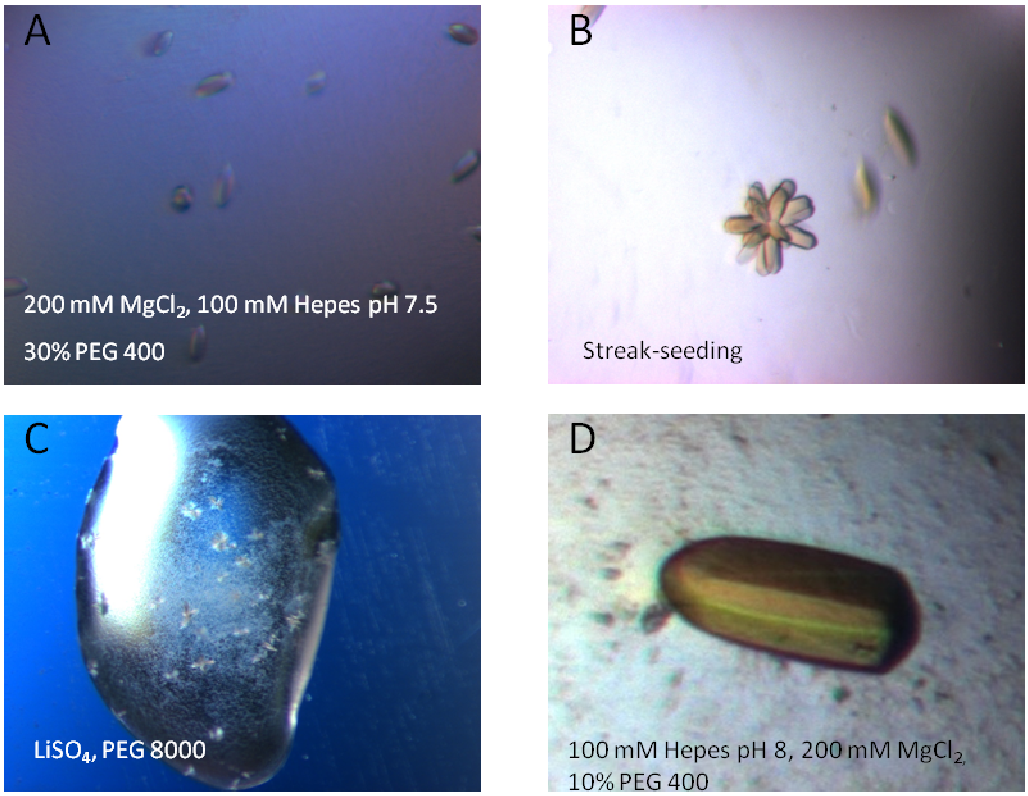


Figure 4.4: Crystals of taXPD.(A) initial Crystals of taXPD of Crystal Screen I&II. **(B)** with Streak-seeding optimized crystals. **(C)** initial Crystals of taXPD of Crystal Screen I. **(D)** optimized taXPD crystal.

Since there was no structure known which could be used for molecular replacement a multiwavelength anomalous diffraction approach was chosen to solve the phase problem and the XPD structure. To achieve this taXPD was co-crystallized in the presence of 8I-AMPPNP, 8Br-AMPPNP, K(Au)CN and furthermore the crystals were soaked with different heavy atom derivatives. These approaches did not result in high quality diffracting crystals with the additive bound to the protein.

4.3.2 Crystallization of taXPD with DNA

For the co-crystallization of taXPD (construct taXPDT3) with DNA different approaches were chosen as described in 3.5. For the first attempts taXPD protein and DNA were incubated at RT or at 37 ℃ for 1 h in a 1:1.2 molar ratio. Furthermore taXPD and DNA were dialyzed against the SEC buffer to form a stable XPD/DNA complex which could be used for crystallization. Moreover an analytical size exclusion chromatography was applied prior to crystallization which

should separate the unbound DNA from the protein-DNA complex. Those different approaches were performed in the presence and absence of ATP or AMPPNP and with different DNA substrates: ssDNA 10mer, ssDNA 22mer, B-GAAF bubble substrate, CPD containing 25mer and 30mer and a dsDNA K10T/KU10T (see 2.5). The best diffracting crystals were obtained by incubatinga ssDNA 22mer and taXPD in a stoichiometric ratio of 1.2:1 at RT for 30 min. Crystals were grown by vapor diffusion in hanging drops containing equal volumes of a protein-DNA complex in 20 mM Tris (pH 8), 200 mM NaCl and 10 mM $MgCl_2$ and a reservoir solution consisting of -50 mM MES pH 6.5, 10 mM $MgSO_4$, 5-30% MPD, 200 mM KCl equilibrated against the reservoir solution at 20℃.

4.3.3 Crystallization of faXPD

For XPD from *Ferroplasma acidarmanus* no crystallization conditions were available. Therefore different crystallization screens were prepared as depicted in 3.5 and faXPD was used in a concentration range of 50 µM to 100 µM. For a faXPD ssDNA co-crystal crystallization set ups were performed in the presence of a 22mer and a CPD-containing 25mer in presence and absence of ATP or AMPPNP. Initial needle shaped crystals of apo faXPD could be obtained from the Index Screen, condition 17 (Figure 4.5). However, the reproduction of those crystals was not successful. The ssDNA co-crystal setups did not result in faXPD crystals, only salt crystals were obtained.

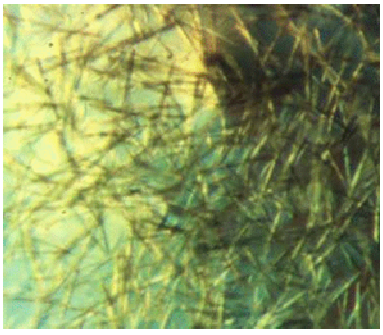


Figure 4.5: Crystals of faXPD. Crystals of faXPD from condition 17, Index Screen (100 mM Tris pH 8.5, 3 M NaCl).

4.4 TaXPD apo structure

4.4.1 Data collection and structure solution of apo taXPD

Prior to data collection, the crystals were cryoprotected by sequential transfer into mother liquor containing increasing amounts of glycerol in 5% steps to a final concentration of 30%.

Table 4.1: Data collection and refinement statistics

Dataset	Subparameter	Native	Fe Peak	Fe Inflection	Fe Remote
Resolution (Å)		2.9	3.6	3.6	3.6
Wavelength (Å)		1.0	1.7367	1.7419	1.7
Unique reflections		13.622	6.963	6.967	6.909
<I>/<σI>		38.3 (1.7)	23.1 (4.5)	19 (3.9)	32.6 (16.1)
Completness (%)		99.0 (99.4)	99.9 (100)	100 (100)	99.2 (100)
Redundancy		4	14.8	12.4	14.3
Rsym		0.07 (0.617)	0.14 (0.52)	0.16 (0.59)	0.07 (0.17)
Phasing to 4.0 Å	Number of sites			1	
Phasing power	Ano		1.505	0.758	1.915
	Iso (acentric/centric)		0.303/0.288	1.236/0.987	
Mean figure of merit (acentric/centric)				0.43/0.15	
Rcryst (Rfree)		0.209 (0.287)			
Rms deviations bond lengths (Å)		0.007			
Rms deviations bond angles (°)		0.92			
Mean B-factor (Å2)		54.2			
Ramachandran statistics		86.6/11.1/1.5/0.8			

The R-factoris defined in 3.6.4. <I>/<σI> indicates the average of the intensity divided by its average standard deviation. Numbers in parentheses refer to the respective highest resolution data shell in each dataset. Phasing power is the mean standard value of the heavy atom structure factor amplitude divided by the lack of closure for anomalous (ano) and isomorphous (iso) data. R_{free} is the same as R_{cryst}, but is calculated for 5% of the data randomly omitted from the refinement. Mean B-factor is the mean residual B-factor after TLS refinement.

The crystals were flash cooled in liquid nitrogen, and data collection was performed at 100 K. Data sets were collected at beamline BM14 (European Synchrotron Radiation Facility [ESRF]) at wavelengths of 1.0 Å, 1.7 Å, 1.7367 Å, and 1.7419 Å (Table 4.1). All data were indexed and processed using Moslfm and Scala. The taXPD protein crystallized in space group $P6_5$ with one XPD molecule in the asymmetric unit. The structure was solved by Dr. Jochen Kuper by multiwavelength anomalous diffraction (MAD) using the anomalous Fe signal of

the endogenous FeS cluster in the protein as described in 3.6.2. The Fe sites were located using ShelxD and with the program Sharp phase improvement was achieved. The sub-structure solution of the cluster and the refinement was carried out at 4 Å resolution and the 4Fe4S cluster was treated as one “super-Fe atom” for phasing. The programs Solomon and Pirate were used for solvent flattening and phase extension to 3.6 Å. Arp/WARP was used to autotrace the solvent-flattened maps by using the low-resolution quick-build option. The obtained model was further extended manually by Prof. Hermann Schindelin and Prof. Caroline Kisker using O and Coot. The model was exposed to phase-restrained simulated annealing and maximum likelihood refinement using phenix.refine after assigning the maximum number of residues possible. This model was then further refined by cycling between manual model building and refinement using Refmac5 with TLS refinement. The final structure was refined at 2.9 Å resolution to an R-factor of 0.209 and an R_{free} of 0.287. The model contains residues 23-507 and 515-615 (586 out of 602 residues) of the XPD construct with residues 20 to 22, 508 to 514, and 616 to 620 presumably disordered.

4.4.2 Overall taXPD structure

The taXPD structure consists of four distinct domains. Residues 23-87, 178-225, and 366-407 form domain 1, residues 88-177 form domain 2, residues 226-365 form domain 3 whereas residues 408-615 form domain 4 (Figure 4.6). Domains 1-3 together with α-helix α22 from domain 4 form a donut-shaped structure that contains a hole with a diameter of approximately 13 Å. The remainder of domain 4 is positioned in front of the ring. The overall dimensions of the protein can therefore be separated into the donut with a width and height of 65 Å and 75 Å and a thickness of 29 Å. At the location of domain 4, the width of the ring is increased to 45 Å. In the context of the overall structure, domain 1 can be viewed as the core domain surrounded by the three other domains. Domains 2 and 3 are insertions,which emerge from domain 1. Domain 2 is inserted between β-strands β3 and β4, while domain 3 is inserted between α-helices α11 and α17. Domain 4 is situated adjacent to domain 1 within the linear protein sequence.

The motor domains

Domains 1 and 4 are the two motor domains of the protein which convert the energy of ATP-hydrolysis into motion. The tandem occurrence of both domains that have the classical RecA-like fold is present in all helicases of superfamilies 1 and 2 (SF1 and SF2). The two domains can be superimposed with an rmsd of 2.4

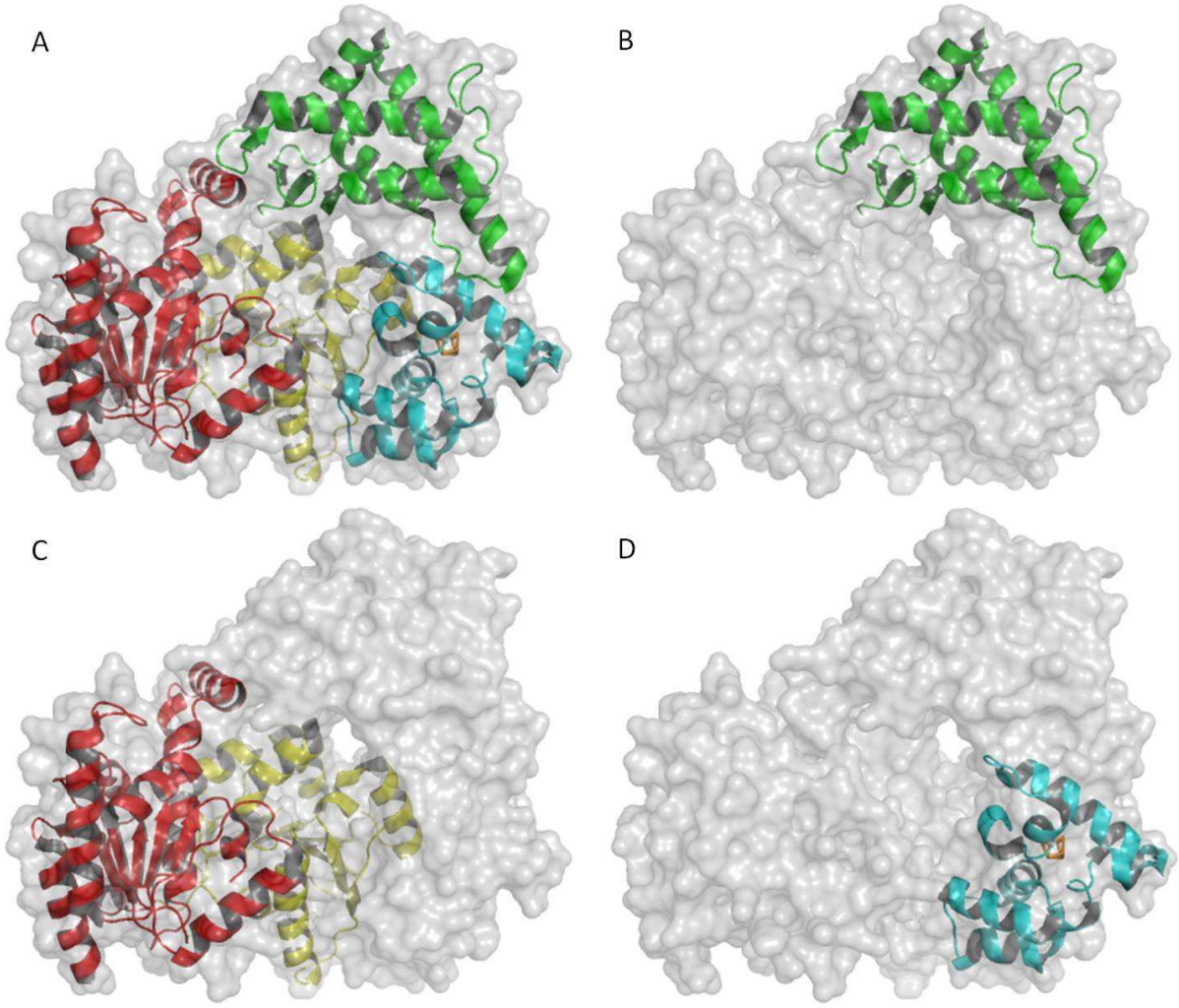


Figure 4.6: Overall structure of taXPD.(A) Front view of taXPD, with domain 1 in yellow, domain 2 in cyan, domain 3 in green, and domain 4 in red. The FeS cluster is shown in all-bonds representation. The surface representation of taXPD is shown in grey. **(B)** view of arch domain. **(C)** view of the two motor domains. **(D)** view of the FeS domain.

Å indicating their similarity. Motor domain 1 consists of seven β-sheets surrounded by seven α-helices. Motor domain 2 contains six β-sheets surrounded by seven α-helices and two 3_{10}-helices. The putative ATP-binding site is formed by the interface between domains 1 and 4. Domain 1 contains helicase motifs I, Ia, II and III, whereas domain 4 harbors helicase motifs IV, V and VI.

The arch domain

The antiparallel β-sheets 6, 7, 8 and 9 and the α-helices 12 to 16 form the arch domain and comprise a novel protein fold. The β-strands β6 to β9 form a β-bridge to domain 1. Furthermore the arch-domain contains two α-helical hairpins formed by α12 and α13 and α15 and α16, respectively. The interface of the two hairpins forms a hydrophobic core. SSM searches revealed no hit indicating the novelty of this protein fold. The interface of domains 2 and 3 is formed by 17 residues from each domain and buries 620 $Å^2$. The interaction between the two domains is mainly mediated by hydrophobic interactions and four salt bridges between residues Lys323/Asp99, Arg335/Glu103, Arg235/Glu103 and Glu315/Lys111.

The FeS cluster domain

The core of the FeS domain is a 4Fe4S cluster which is coordinated by the four cysteines Cys92, Cys113, Cys128 and Cys164 (Figure 4.7). Three of these cysteines are located in loops whereas Cys113 is positioned within α-helix 5. The cluster is surrounded by six α-helices and two 3_{10}-helices. The domain is stabilized by hydrophobic interactions. Hydrophobic residues like Arg88, Tyr166, Glu167, Lys117 and Phe131 shield the cluster with their long side chains from the solvent. They all point into a basic groove which lies between domains 1 and 2 and could be a putative DNA binding site. Residues Arg88 and Tyr166 form a small basic pocket in this groove. A SSM search revealed that the transcription factor c-myb is the closest structural homolog to the fold of the FeS cluster although c-myb does not contain an FeS cluster.

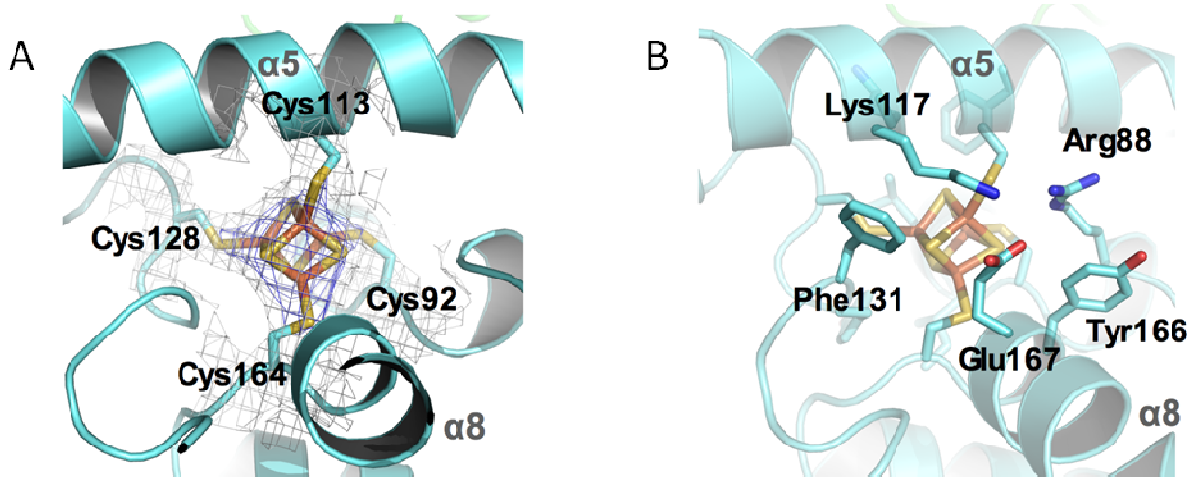


Figure 4.7: FeS cluster.(A) SIGMAA weighted 2Fo-Fc electron density map contoured at one times the standard deviation around the 4Fe4S cluster and the surrounding protein in grey, and at four times the standard deviation in blue. The cluster and the coordinating cysteines are shown in all-bonds representation. Secondary structure elements are labeled in grey. **(B)** Residues in close proximity to the 4Fe4S cluster. The cluster is shown as in (A), and residues Arg88, Tyr166, Lys117, Phe131, and Glu167 are shown in all-bonds representation, with secondary structure elements labeled in grey.(Wolski et al. 2008)

4.5 Model of faXPD based on taXPD

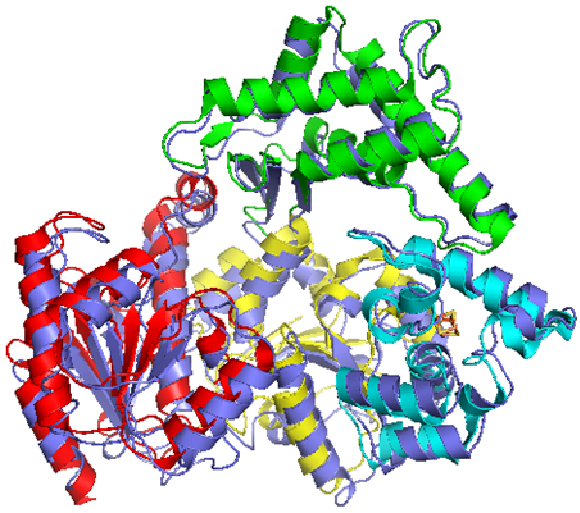


Figure 4.8: Model of faXPD. taXPD is colored with domain 1 in yellow, domain 2 in cyan, domain 3 in green and domain 4 in red. The FeS cluster is shown in all-bonds representation and colored orange. FaXPD is colored in blue.

Since there is no structure available for the *Ferroplasma acidarmanus* XPD a structure was generated through homology modeling. For automated homology

modeling the ESyPred3D web server was used with the XPD structures from *Thermoplasma* and *Sulfolobus* (2VSF and 3CRV) as templates. The superposition of the homology model and the taXPD structure reveals that the faXPD model has a quite similar overall fold as taXPD WT (Figure 4.8). The two structures can be superposed with an rms deviation of 2 Å indicating a similar overall fold which coincides with the CD spectroscopy result which revealed a similar secondary structure in both XPDs.

4.6 XPD–DNA binding model

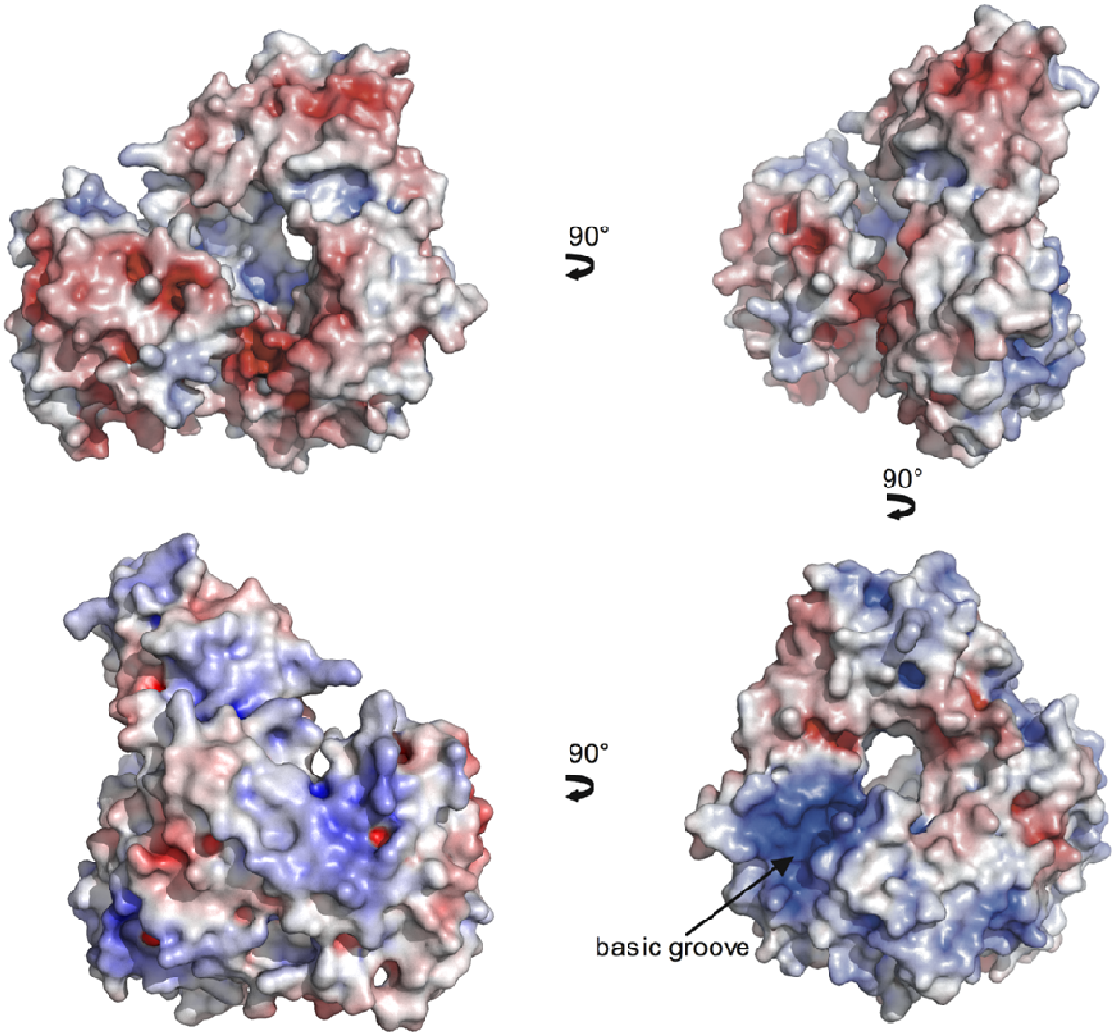


Figure 4.9: Electro-static surface potential of taXPD. Four different views of the electrostatic surface potential of taXPD. The surface potential has been calculated with Pymol/APBS at an ionic strength of 150 mM and is contoured at ±10kBT. The first view is the same as in Figure 4.6, A; the

arrows show the transition from one view to the other with the rotation indicated by the arrow. Putative DNA binding sites are marked by additional arrows.(Wolski et al. 2008)

To obtain insight into the DNA binding mode of XPD, the electrostatic surface potential of the protein was calculated with Pymol/APBS at an ionic strength of 150 mM (Figure 4.9). The surface potential indicates a positively charged path along domain 4, leading towards a highly conserved groove along domain 4 and domain 1. This basic path protrudes through a pore formed by domain 1, the arch domain and the FeS domain as described above into a basic pocket behind the pore.

Based on the calculations of the electrostatic surface potential (see Figure 4.9), mapping of highly conserved (yellow) and strictly conserved (green) residues on the XPD structure and the superposition with the structure of NS3 with ssDNA (see Apendix) a DNA-binding model can be proposed (Figure 4.10). The ssDNA passes through a conserved groove from domain 4 to the pore. The DNAprotrudes through the highly conserved hole into a basic groove in the FeS domain. This basic groove additionally harbors a narrow pocket formed by residues Arg88, Tyr166 and Tyr185 (Figure 4.10). This pocket is sufficient in size to accommodate purine and pyrimidine bases. This model is supported further byadditional peaks in the difference electron density maps which have a space of 6.5 Å between them. In DNA the space between the phosphates of the backbone is 6.4 Å (Figure 4.10). Since there is no phosphate present in the purification or crystallization conditions it could be possible that some DNA remained bound to the protein.

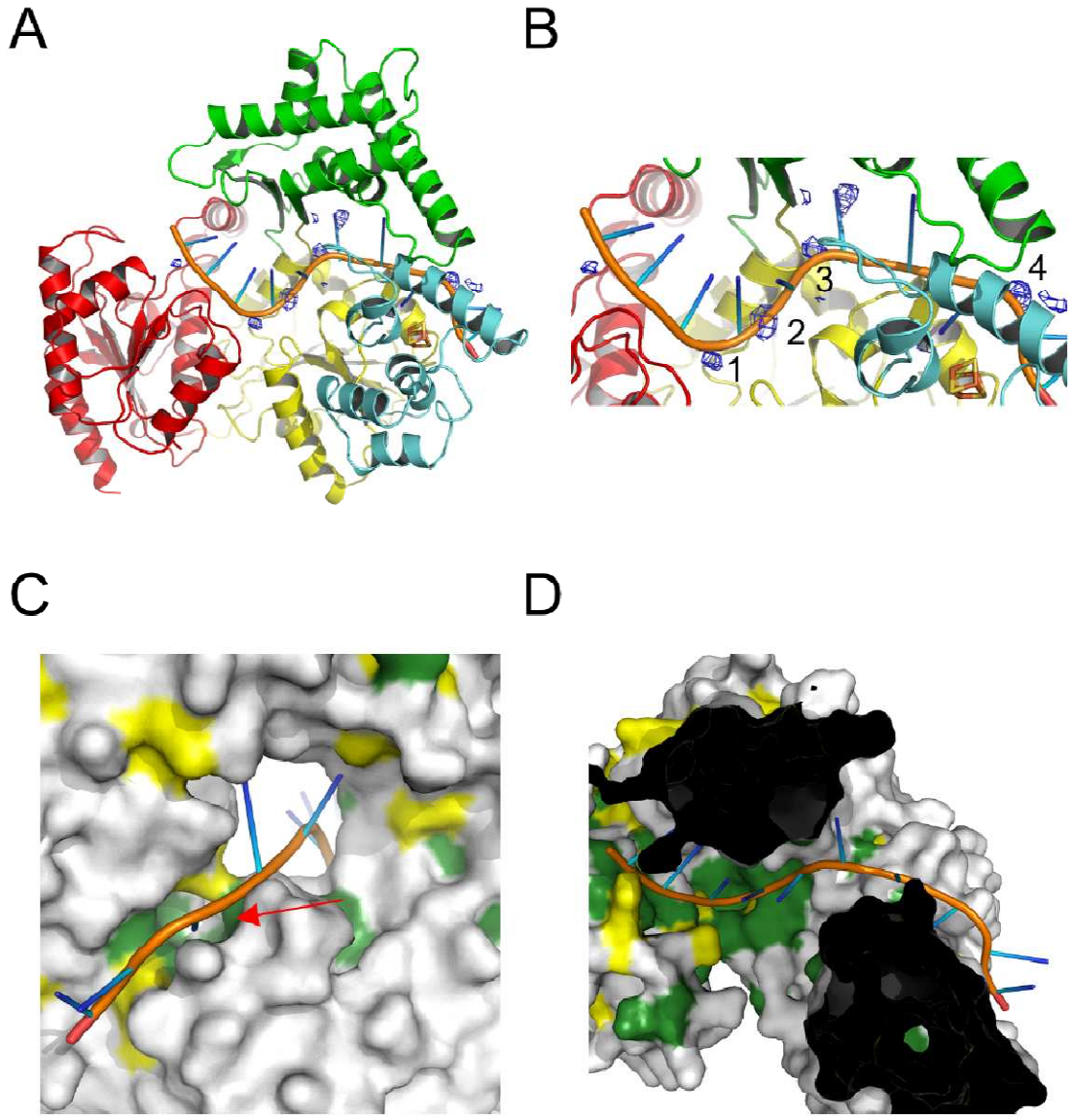


Figure 4.10: Hypothetical XPD-DNA model. Domain 1 is colored in yellow, domain 2 in cyan, domain 3 in green and domain 4 in red, respectively. **(A)** Overall view of XPD in complex with ssDNA. The model was obtained by superposition with the NS3 helicase in complex with ssDNA. Residual difference map peaks at 2.5 times the standard deviation are shown in blue, and peaks used for backbone phosphate positioning are numbered. Further extension of the ssDNA towards the hole was achieved by addition of B-form DNA. The DNA is shown with its backbone in orange, and the bases are shown in blue spokes. **(B)** Close up view of (A). **(C)** Surface representation of the XPD-DNA model. The DNA is shown as in (A), and the surface is colored according to sequence conservation, with green indicating strictly conserved, yellow highly conserved and gray residues that are not conserved. The DNA in this view emerges from the pore and fits nicely into the highly conserved pocket (indicated by the arrow), which could potentially play a role in damage recognition or which couples DNA binding to the FeS cluster. **(D)** Top view of the XPD-DNA model. For clarity, the top part of XPD has been removed to allow a view into the highly conserved groove, which leads the ssDNA towards the hole in the donut-shaped molecule.(Wolski et al. 2008)

4.7 taXPD-DNA complex

4.7.1 Data collection and structure solution of taXPD with DNA

Prior to data collection the crystals were cryoprotected by sequential transfer into mother liquor containing 30% MPD. The crystals were flash frozen in liquid nitrogen and data collection was performed at 100 K. The data set was collected at beamline BM14 (ESRF) at a wavelength of 0.972 Å and to 2.2 Å resolution. All data were indexed and processed using Mosflm and Scala. The crystal belong to space group P65 with unit cell dimensions of a=b= 79.1 Å and c=175.7 Å (Table

Table 4.2: Data Collection and refinement statistis

Space group	P65
Wavelength	0.972
Cell dimensions a, b, c α,β, y	 79.1, 79.1, 175.7 90, 90, 120
Resolution	44.5 – 2.2 (2.32 – 2.2)
Completness (%)	100 (100)
Rsym	0.085 (0.583)
Mean (I/σI)	9.4 (2.1)
Redundancy	3.8 (3.8)

Resolution	52 – 2.2
No.reflections	29,606
Rcryst/Rfree	0.195/0.251
No.atoms Protein/ions/DNA water	 4,933/16/96 138
B factors XPD DNA	 45.6 75.9
Bond lengths	0.012
Bond angles	1.736
Ramachandran statistics (%)	95.4/4.1/0.5

Values in parentheses refer to highest resolution shell. The R-factor is defined in 3.6.4. I/σI indicates the average of the intensity divided by its average standard deviation. R_{free} is the same as R for the 5% of the data randomly omitted from refinement. B-factor analysis was performed on the residual B-factors after TLS-refinement. Ramachandran statistics indicate the fraction of residues in the favoured, allowed, and outlier regions of the Ramachandran diagram.

4.6). The apo XPD structure (2VSF) was used as a model for molecular replacement which was performed with PHASER. To improve the model alternating rounds of manual model building in Coot and maximum likelihood refinement with Refmac5 were performed. The structure was refined at a resolution of 2.2 Å to an R-factor of 19.5 % and and R_{free} of 25.1 % with good stereochemistry as indicated by the Ramachandran statistics (Rampage).

4.7.2 XPD-DNA structure

For protein expression a construct (taXPDT3) with additional residues at the N-terminus was used, and therefore the electron density of this α-helix was observed. Furthermore four other significant electron density features were observed. The largest electron density feature could be identified as four DNA bases. From the density the sequence could be mapped as TACG, which is found at the 5'-end of the used DNA substrate. The DNA is located at domain 4. The 5'-end of the DNA points towards the solvent whereas the 3'-end points towards the conserved pass leading through the protein. The protein-DNA interface consists of an accessible surface area of 370 $Å^2$. This interface consists in equal parts of polar and non-polar residues. Trp549 forms a strong protein-DNA interaction by π-interactions with the adenine base. Arg584 also mediates strong interactions with DNA. Arg584 interacts via hydrogen bonds with the phosphate backbone of the adenine and cytosine bases. The Arg584-DNA interaction is further stabilized by interactions between Asp582 and Arg584.

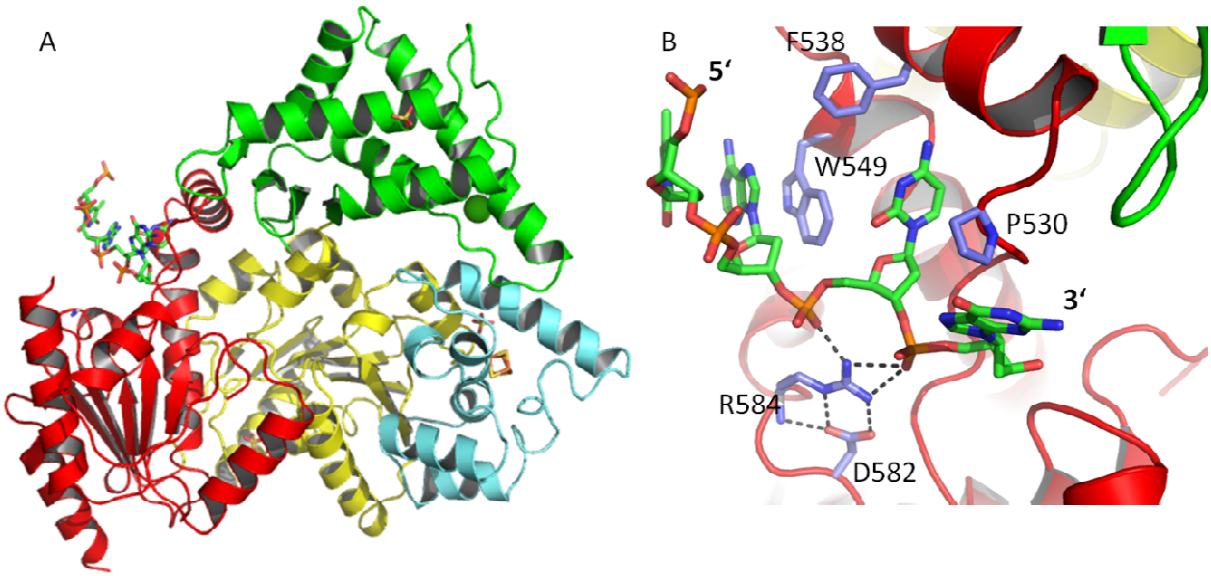


Figure 4.11: The taXPD-DNA complex. (A) Domain 1 is colored in yellow, domain 2 in cyan, domain 3 in green and domain 4 in red. The FeS cluster is shown in all-bonds representation. The DNA in green is shown in all-bonds representation. **(B)** Enlarged view showing the tetra-oligonucleotide visualized in the structure with its carbon atoms in green. Residues interacting with the DNA are shown with their carbon atoms in blue and hydrogen bonds are indicated by dashed black lines.

The three remaining features in the electron density could be identified as sulfate molecules which originate from the crystallization solution. The first sulfate site is

located in the arch-domain at α-helix α13. It is involved in crystal contacts. The second sulfate is located directly in a small pocket in the basic groove at the exit of the poreand forms hydrogen bonds with Tyr166, Lys170, Arg88 and His198. The third sulfate molecule is located in the ATP-binding pocket of the Walker A motif.

4.8 Sequence alignment of XPD from different organisms

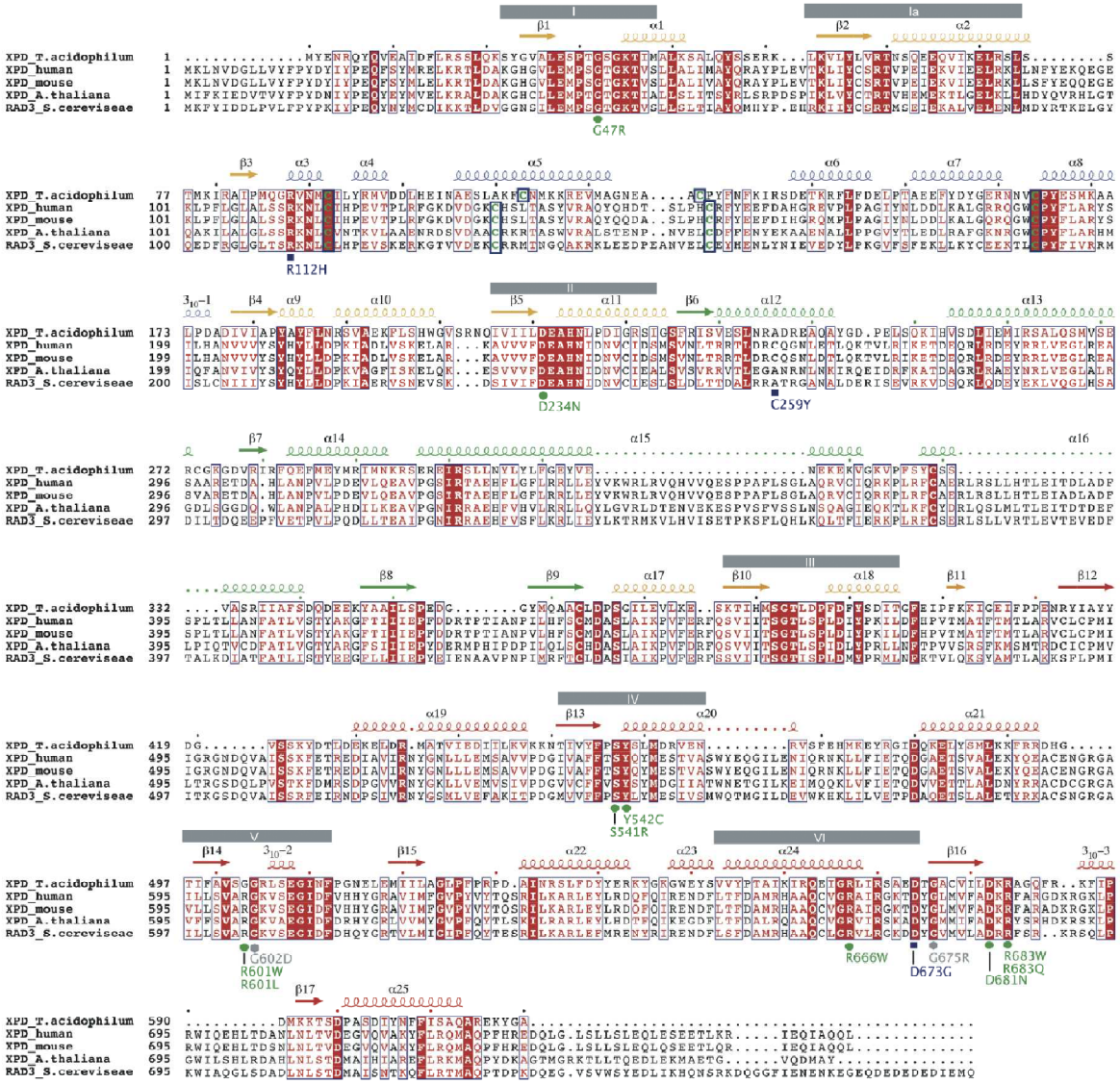


Figure 4.12: Sequence alignment. Sequencealignment of five different XPD proteins. From top to bottom: XPD from *T.acidophilum*, human, mouse, *Arabidopsis thaliana*, and *Saccharomyces cerevisiae*. Secondary structure elements are indicated above the sequence and are color coded according to their domains; arrows indicate β-strands and coils α-helices. Helicase motifs are shown as grey boxes. Mutations leading to xeroderma pigmentosum, Cockayne syndrome or trichothiodystrophy are indicated below the sequence and are colored in blue, green and grey, referring to their occurrence in trichothiodystrophy, xeroderma pigmentosum or xeroderma

pigmentosum/Cockayne syndrome patients, respectively. Conserved residues are boxed and strictly conserved residues are shown in white with a red background. The four cysteines that coordinate the 4Fe4S cluster are highlighted in green surrounded by thick blue boxes. (Wolski et al. 2008)

The sequence alignment in Figure 4.12 shows a comparison of XPDs from different organisms. The secondary structure elements and the helicase motifs are mapped on the alignment. This alignment reveals a high sequence homology in the helicase motifs. Furthermore, the Iron-sulfur cluster coordinating cysteines are conserved in the XPDs of all organisms shown in Figure 4.12 from *Thermoplasma acidophilum* to human. Most of the mutations which lead in humans to the diseases xeroderma pigmentosum, trichothiodystrophy and xeroderma pigmentosum/ Cockayne syndrome are strictly conserved residues in all organisms shown with exception of R601W, R601L and C259Y. This indicates that taXPD serves as a good model for human XPD due to the high sequence homology.

4.9 Site-directed mutagenesis

Based on the taXPD apo structure (Figure 4.6), the DNA binding model and the sequence alignment different XPD mutants were generated by site-directed mutagenesis of the taXPDT3 construct as described in 3.3.1. These mutants provide a basis to assess where the path of the translocated strand is located and identify functionally important regions of the protein that also apply to eukaryotic XPDs. The mutants can be divided into different groups: first there are the cluster mutants C92S, C113S, C128S and C164S where the FeS cluster coordinating cysteine is replaced by a serine; secondly there are the disease related mutations R88H, R567W, D582N and R584E which lead in humans to trichothiodystrophy and xeroderma pigmentosum, respectively. To find possible regions on both sides of the protein where double stranded DNA is separated a third group of mutants was created including M115R, R118A, F133A, F538S, Y545S, W549S and Y545S/W549S. Finally a fourth group of mutants was designed including mutants of conserved and strictly conserved residues: Y46A, R59A, R59E, E107A, Y166A, K170A, K170E, Y185A, Y166A/Y185A, E315A, F326A and Y425E.

Expression of the mutants was performed analogously to XPD WT. The expression and purification of M115R, C128S, C164S, E315A and Y166A/Y185A revealed no overexpression or the proteins were not stable during the purification procedure. Therefore they were not pursued further.

4.10 ThermoFluor of taXPD mutants

After the expression and purification of the taXPDvariants it had to be shown that the taXPD pointmutants have the same overall conformation and stability as the taXPD WT. To this end a ThermoFluor assay was performed as described in 3.4.4. As depicted in Figure 4.13 taXPD WT unfolds at 51℃. The shape of the unfolding curve indicates that the protein unfolds with one steep transition. The same melting temperature as taXPD WT was obtained for the mutants K170A, Y166A/Y185A, Y545S/W549S and W549S. The mutants Y46A, R59A, R59E, R88H, C92S, C113S, F133A, Y166A, Y185A, F326A, Y425E, F538S and R584E are more stable and unfold at slightly higher temperatures. The C113S, Y185A and the R59E differ in the unfolding transition. Their unfolding curves show a biphasic behavior indicating a slightly different fold as taXPD WT. Proteins with a different unfolding transition as taXPD WT were not used for further analysis with exception of C113S.

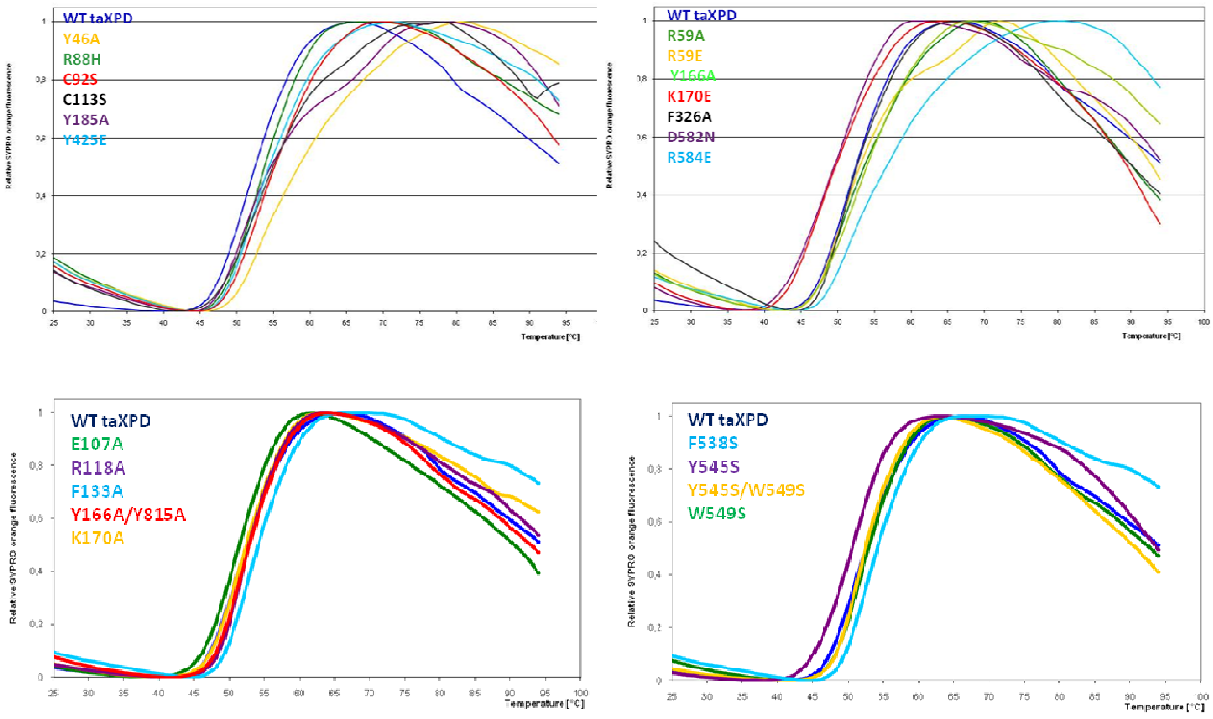


Figure 4.13:Unfolding Curves of taXPD WT (blue) and taXPD variants.

4.11 Biolayer Interferometry identifies residues which are involved in DNA binding

To identify residues which are important for ssDNA binding Biolayer Interferometry assays were performed in the Octet reaction buffer as described in 3.4.6. The measurements were performed at 30℃ with different protein concentrations and different protein batches. To compare the DNA binding properties of taXPD variants 100 nM of a ssDNA 50mer was immobilized at the sensor tip. The data was analyzed using the Octet RED data analysis software and the curves were fitted using the 1:1 binding model.

Comparable to XPD from *Sulfolobus acidocaldarius*(Fan et al. 2008) taXPD displays a binding constant (K_D) of 61 nm for the used substrate which indicates a high affinity for this DNA.

All the disease related mutations show reduced DNA binding ability, however they differ in the strength of the phenotype from 10-fold for D582N, 6-fold for R567W, 5-fold for R88H and a 37-fold reduction of the K_D for R584E. Y46A is the only mutant

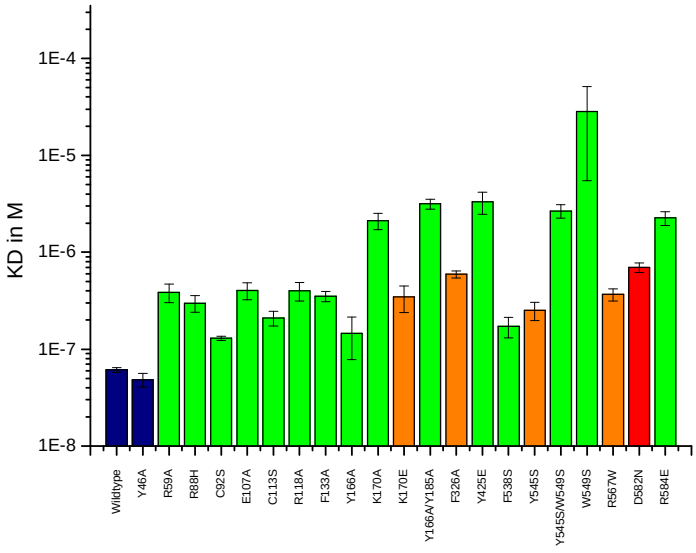


Figure 4.14: DNA binding properties of taXPD variants. The K_D values of different taXPD mutants are plotted in a logarithmic scale. The mutants are colored with respect to their K_D. Blue colored are K_Ds with 10^{-8}M, green with 10^{-7} M, orange with 10^{-6} M and red with 10^{-5} M. The error bars show the standard error.

Table 4.3: DNA binding properties of taXPD variants.

	K_D
Wildtype	61 ± 3.51 nM
Y46A	48.5 ± 7.79nM
R59A	385 ± 83.7nM
R88H	299 ± 57.9 nM
C92S	130 ± 6 nM
E107A	403 ± 79.9 nM
C113S	210 ± 36.8 nM
R118A	400 ± 87.5 nM
F133A	353 ± 43.1 nM
Y166A	146 ± 67.8 nM
K170A	2.12 ± 0.4 µM
K170E	345 ± 105 nM
Y166A/Y185A	3.16 ± 0.37 µM
F326A	594 ± 48.9 nM
Y425E	3.32 ± 0.85 µM
F538S	172 ± 41.2 nM
Y545S	251 ± 53.8 nM
Y545S/W549S	2.67 ± 0.42 µM
W549S	28.3 ± 22.8 µM
R567W	367 53.3 nM
D582N	697 ± 78.6 nM
R584E	2.26 ± 0.37 µM

The values are color coded with respect to the K_D. WT binding properties with 10^{-8}M are colored in blue, K_Ds with 10^{-7} M are colored in green, K_Ds with 10^{-6} M are colored orange and K_Ds with 10^{-5} M are colored red.

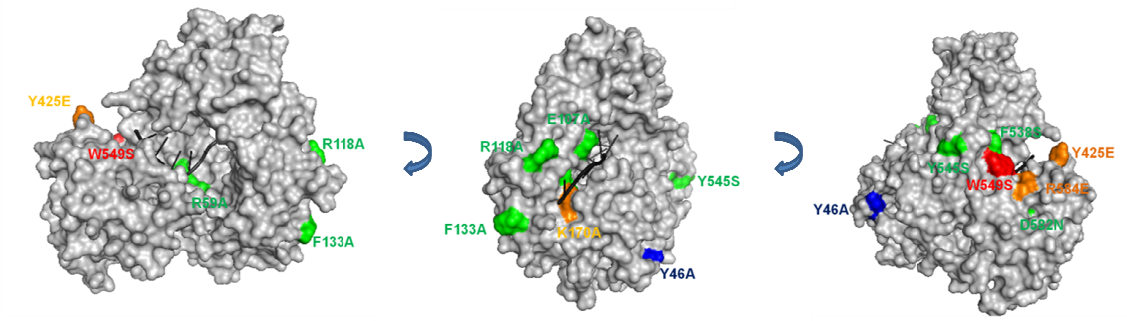


Figure 4.15: K_Ds of taXPD variants. Surface representation of taXPD. TaXPD variants are colored with respect to their K_D. 10^{-8} M is colored in blue, 10^{-7} M is colored in green, 10^{-6} M is colored in orange and 10^{-5} M is colored in red.

which shows with 48.5 nM a comparable K_D as taXPD WT (61 nM). The two cluster mutants C92S and C113S show a 2- to 3-fold reduction of the K_D. The

strongest effects on ssDNA-binding were observed in the second motor domain (domain 4) where Tyr425, Phe538, Tyr545, Trp549, Arg567, Gln582 and Arg584 are located. The W549S mutant even shows a 500-fold reduction in the DNA binding ability indicating an important role in protein-DNA interaction for Trp549. Interestingly the Y545S/W549S doublemutant has a lower $K_D$compared to W549S and therefore displays a better DNA binding ability. As displayed in the DNA binding model ssDNA continues from domain 4 between the two motor domains as indicated by the reduced affinity of the R59A (6-fold) variant. Arg59 is a strictly conserved residue located in helicase motif Ia. The side chain of Arg59 points into a cleft at the interface of the two motor domains which was proposed to accommodate ssDNA. This path protrudes through the pore formed by motor domain 1, the arch domain and the FeS domain as mentioned above. To verify if a DNA strand is passing through the pore, five variants were designed, within or in close proximity to the pore, namely R88H, E107A, Y166A, Y166A/Y185A and F326A. All five variants show a reduced binding affinity indicating that these residues are involved in protein-DNA interactions. These observations support the implication that the translocating DNA strand is passing through the pore. As suggested by the DNA binding model the translocated DNA strand most likely interacts with residues located in a basic groove that is formed by two α-helices (5 and 8) from the FeS domain and α-helix 10 from domain 1 since R118A, F133A and K170A show a decreased affinity. Taken together the location and binding affinities of the different variants suggest that a path of ssDNA leads from domain 4 along a groove formed by the two motor domains through the pore and thereby covering the complete length of the protein which could accommodate around 20 bases of DNA.

4.12 taXPDs ATPase activity is stimulated by ssDNA

To investigate which regions of taXPD are important for ATPase activity ATPase assays were kindly performed by Gudrun Michels. For XPD from *Sulfolobus acidocaldarius*it was demonstrated that its ATPase activity is stimulated by ssDNA

(Rudolf et al. 2006). Therefore taXPD WT and mutants were analyzed with respect to their ATPase activity stimulated by ssDNA. ATPase assays were performed as described in 3.4.5. In absence of ssDNA the ATPase activity is hardly detectable

(data not shown). Upon addition of ssDNA (0.5 µM) the specific activity of taXPD WT is 2.0 mol ATP*s-1*mol-1 XPD (Figure 4.16). This value is four times higher than that determined for XPD from *Sulfolobus acidocaldarius*(Fan et al. 2008). Based on the results of the ATPase assays the taXPD variants can be divided in two different groups. Group I consists of Y46A, R59A, Y425E, Y545S/W549S W549S, D582N and R584E whereas group II contains R88H, C92S, E107A, C113S, R118A, F133A, Y166A, Y166A/Y185A, K170A, K170E and F326A. Group I variants are located in the two motor domains and thereby in front of the pore with respect to the DNA binding model. Variants of group II are located in the arch domain and in the FeS domain and thus behind the pore.

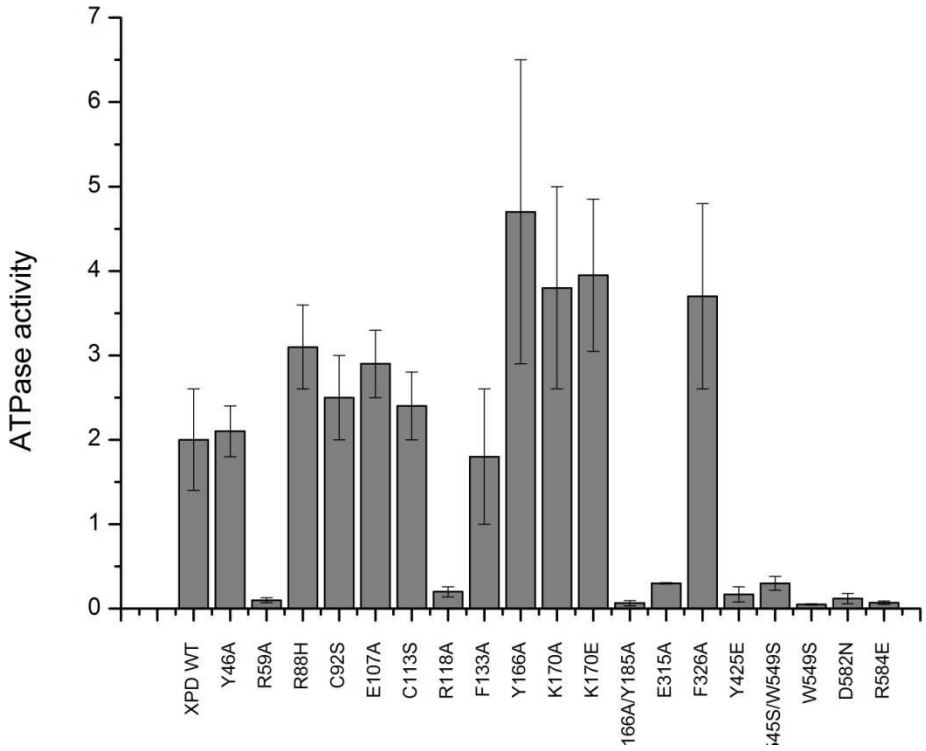


Figure 4.16: ATPase activity of taXPD WT and variants.

All group I variants display a significantly decreased DNA binding ability which comes along with an impaired ATPase activity. W549S displays the most severe phenotype with no measureable ATPase activity and a strong impaired DNA binding ability. Mutants which are located in front of the pore show at least a tenfold reduction in ATPase activity. Interestingly, mutants which are positioned within or behind the pore display a different phenotype. In these variants the ATPase activity remains high although ssDNA-binding is impaired. However, there

is an exception, R118A displays a reduced ATPase activity. The group II variant Y166A is located directly in the pore, it shows an increased activity of 4.7 mol ATP*s-1*mol-1 compared to 2.0 mol ATP*s-1*mol-1 for taXPD WT. K170A, F326A, R88H and E107A variants also display a significantly increased ATPase activity with 50% increase compared to taXPD WT. The activity of F133A is, however, comparable to taXPD WT.

4.13 taXPD is an ATP-dependent 5'-3'helicase and residues Y166, D582 and R584 are important for helicase activity

To demonstrate that taXPD is an active enzyme helicase assays were performed as described in 3.4.7. The results reveal that XPD unwinds 5'-overhang substrates in an ATP-dependent manner as shown in Figure 4.16.

To identify regions of the protein which are important for the activity of the enzyme, helicase assays with the taXPD variants were performed by Dr. Jochen Kuper and Gudrun Michels. Helicase assays were conducted as described in 3.4.8. All group I variants (R59A, Y425E, W549S, D582N and R584E) display impaired DNA unwinding ability which is in line with defects in DNA binding and ATPase activity. The two disease related variants D582N and R584E which lead in humans to XP display no helicase activity (Figure 4.18). The phenotype of the group II variants is

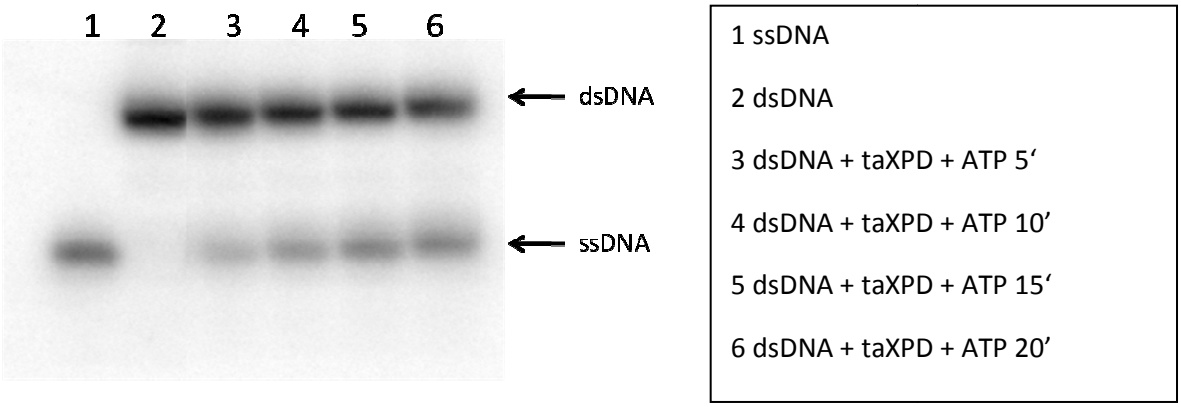


Figure 4.17: Helicase activity of taXPD WT.

more complex. Although they show defects in DNA binding and elevated ATPase activity the variants R88H and E107A display a helicase activity comparable to taXPD WT. The two variants F133A and K170A display a significantly increased helicase activity which is interesting since they are impaired in DNA binding.

Furthermore K170A displays with a 3-fold increase compared to taXPD WT the highest helicase activity. Since K170A also shows elevated ATPase activity this can result from the higher turnover rate for ATP. In contrast the Y166A and F326A variants also display an increased ATPase activity yet either show no (Y166A) or a highly reduced (F326A, 20-fold) helicase activity.

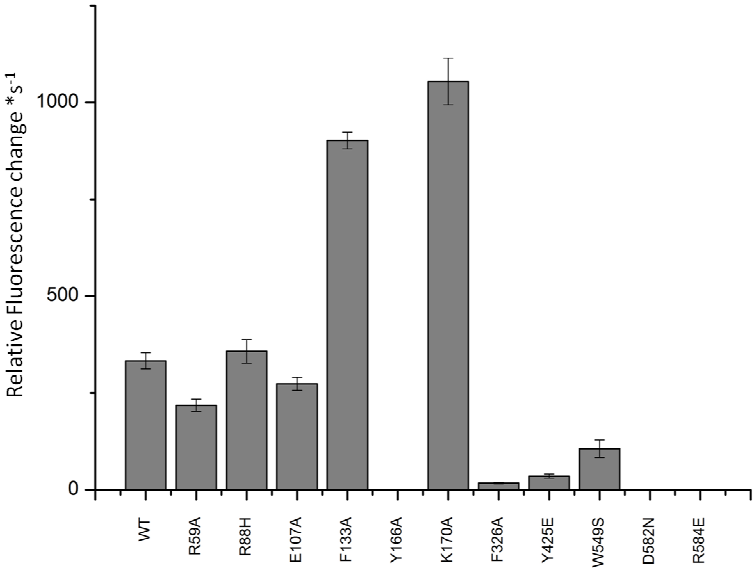


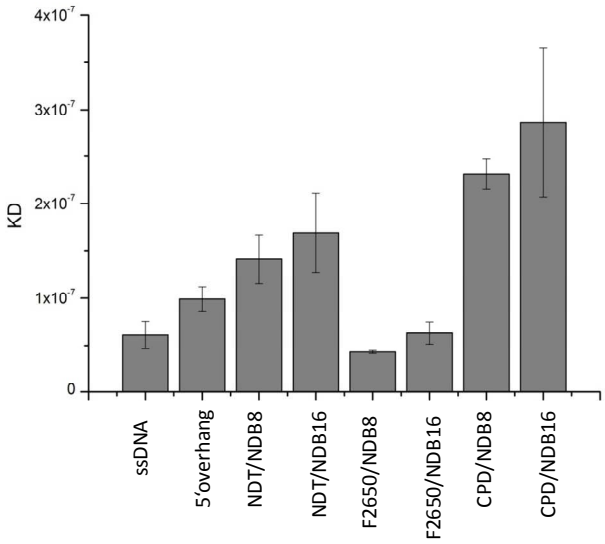
Figure 4.18: Helicase activity of taXPD variants.

4.14 taXPD and faXPD bind fluorescein-containing bubble substrates with a higher affinity as non-damaged bubble substrates

Different DNA substrates were tested in Biolayer Interferometry DNA binding assays with taXPD WT and faXPD to identify an optimal substrate for crystallization. 100nM of the particular substrate was immobilized at the sensor tip. The measurements were carried out as described in 3.4.6 in the Octet reaction buffer. The experiment was performed at 30℃ with different protein concentrations and different protein batches. The data was analyzed using the Octet RED data analysis software and the curves were fitted using the 1:1 binding model.

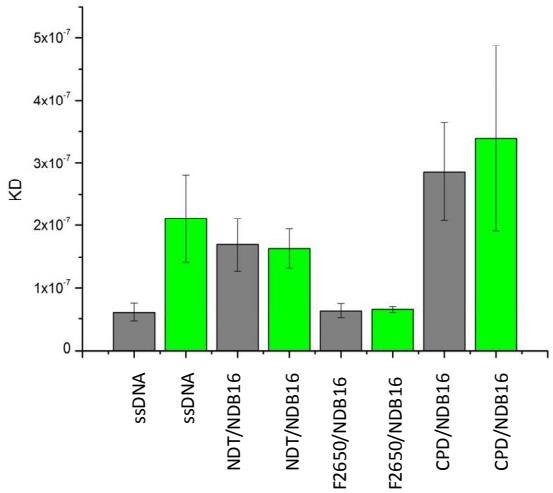
Different DNA lengths were tested from a 10mer up to a 50mer. DNA binding assays revealed that a ssDNA 50mer displays a K_D of 61 nM and is the best ssDNA substrate tested (Weller 2010). A ssDNA 50mer, a 5'-overhang substrate, different bubble substrates and different damages were compared subsequently

(Figure 4.19). TaXPD WT shows a slight preference for small 8 nucleotide bubbles compared to the larger 16 nucleotide bubbles. As a damage fluorescein and CPD were chosen since it was shown for faXPD that it is stalled by a CPD lesion (Mathieu et al. 2010). Interestingly the fluorescein containing bubble substrates have a 4-fold lower K_D as the CPD containing bubble substrates (Figure 4.19). taXPD WT also shows a preference for the 8 nucleotide fluorescein containing bubble substrate compared to the non-damaged and the 5'-overhang substrate. Interestingly, the CPD containing substrates display the worst affinity.



substrate	K_D
ssDNA	61 ± 14.5 nM
5' overhang	99.1 ± 12.8 nM
NDT/NDB8	141 ± 25.9 nM
NDT/NDB16	169 ± 41.8 nM
F2650/NDB8	42.8 ± 1.6 nM
F2650/NDB16	63 ± 11.6 nM
CPD/NDB8	231 ± 16.2 nM
CPD/NDB16	286 ± 79.3 nM

Figure 4.19: DNA binding behavior of taXPD WT to different DNA substrates.



substrate	K_D
taXPD ssDNA	61 ± 14.5 nM
faXPD ssDNA	211 ± 70.1 nM
taXPD NDT/NDB16	169 ± 41.8 nM
faXPD NDT/NDB16	163 ± 31.1 nM
taXPD F2650/NDB16	63 ± 11.6 nM
faXPD F2650/NDB16	65.6 ± 4.6 nM
taXPD CPD/NDB16	286 ± 79.2 nM
faXPD CPD/NDB16	339 ± 14.8 nM

Figure 4.20: Comparison of DNA binding properties of taXPD WT and faXPD.

Furthermore the DNA binding properties of taXPD WT and faXPD were compared (Figure 4.20). Therefore a ssDNA 50mer and 16 nucleotide bubble substrates with and without a damage were chosen. FaXPD shows as taXPD WT a preference for a fluorescein containing bubble substrate. TaXPD WT binds 4-fold better to the ssDNA 50mer as faXPD. FaXPD and taXPD WT bind better to non-damaged bubble substrates as to CPD-containing bubble substrates. Taken together the K_Ds for all substrates show that taXPD WT and faXPD bind with high affinity to the substrates. However both XPDs show a preference for the fluorescein-containing substrate.

The association and dissociation curves for taXPD WT and faXPD reveal for all substrates a similar k_{on} and k_{off}, respectively. The k_{on} is approximately 10^3 and the k_{off} approximately 10^{-5} for all the different substrates. This indicates that all substrates are bound in a similar way. Since for the bubble substrates there are two possible DNA binding sites it was also tried to fit with a 2:1 binding model, however only the 1:1 binding model revealed good fits for the association and dissociation curves.

4.15 Iron sulfur cluster analysis

4.15.1 Electron paramagnetic resonance studies

Iron-sulfur (FeS) clusters can have different redox states: $[4Fe4S]^{2+}$and the oxidized forms $[4Fe4S]^{3+}$and $[3Fe4S]^{1+}$. To further analyze the4Fe4S cluster of taXPD electron paramagnetic resonance studies were performed in collaboration with Lisa Engstrom a graduate student in Sheila David's laboratory, UC Davis, USA. A $[4Fe4S]^{2+}$ cluster shows no EPR signal, it is EPR silent, whereas the oxidized $[3Fe4S]^{1+}$ cluster shows a significant signal at 2.02 g. EPR analysis was performed as described in 3.4.9. First XPD was treated with different oxidants.

Table 4.4: Iron-sulfur cluster properties of taXPD

Oxidant	[oxidized cluster]	% oxidation
Co-$(phen)_3^{3+}$	13 µM	26 %
Co-$(5\text{-Cl-phen})_3^{3+}$	6.7 µM	13 %
Superoxide	33 µM	66 %

Untreated taXPD WT contains a $[4Fe4S]^{2+}$ cluster indicated by only 2% of the clusters oxidized. Only treatment with Co-$(phen)_3^{3+}$, Co-(5-Cl-$phen)_3^{3+}$ and superoxide resulted in a $[3Fe4S]^{1+}$ cluster signal indicating that taXPD is oxidized in the presence of these agents. Superoxide seems to be the best oxidantsince 66% of the clusters are oxidized compared to 26% and 13% for Co-$(phen)_3^{3+}$ and Co-(5-Cl-$phen)_3^{3+}$, respectively (Table 4.4, Figure 4.21). However,the FeS cluster of taXPD is resistant to oxidization by $Ir(Cl)_6^{2-}$, ferricyanide, H_2O_2,peroxynitrite and nitric oxide since taXPD is EPR silent in the presence of these oxidants (Figure 4.21). The FeS cluster could not be reduced to a $[4Fe4S]^{1+}$ cluster by sodium dithionite (spectra not shown) indicating that taXPD behaves like a HiPIP and not like Ferredoxin.

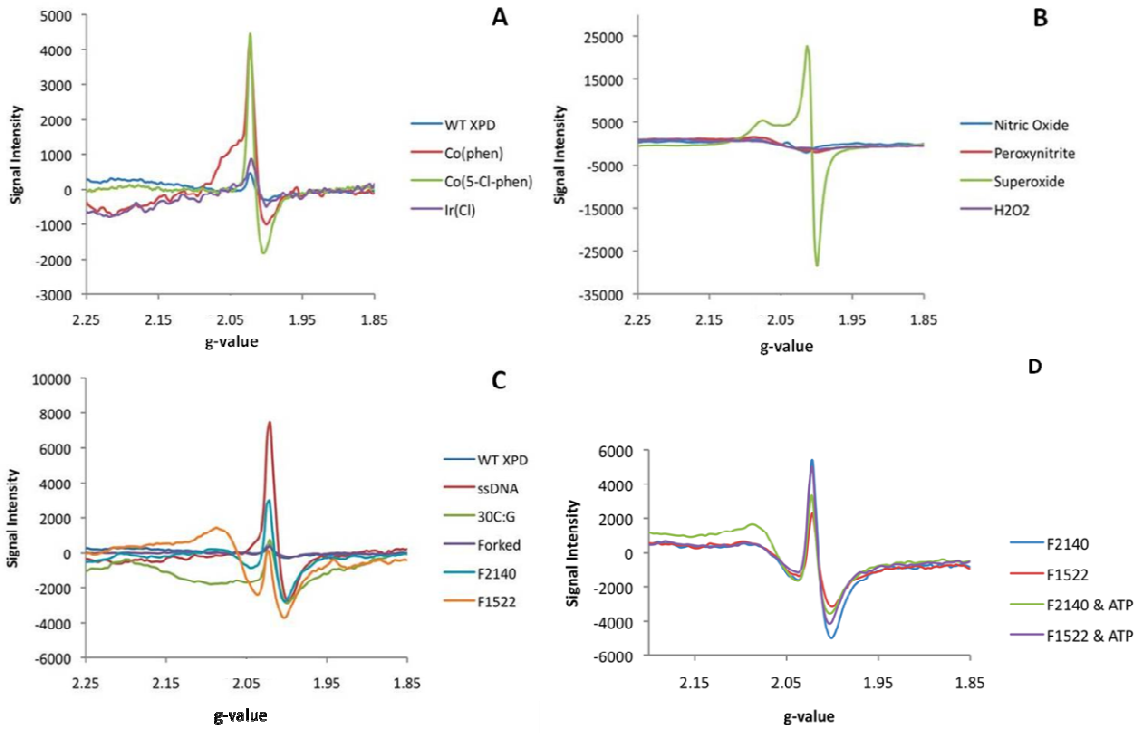


Figure 4.21: EPR spectra of taXPD WT. **(A)** Spectra of taXPD WT without oxidant in blue, with Co$(phen)_3^{3+}$ in red, with Co(5-Cl-$phen)_3^{3+}$ in green and with $Ir(Cl)_6^{2-}$ in purple. **(B)** Spectra of taXPD WT with Nitric Oxide (blue), Peroxynitrite (red), Superoxide (green) and H202 (purple). **(C)** Spectra of taXPD WT and different DNA substrates. Without DNA (blue), ssDNA (red), 30C:G (green), forked (purple), F2140 (cyan) and F1522 (orange). **(D)** Spectra of taXPD WT and different DNA substrates with and without ATP. F2140 colored in blue, F1522 colored in red, F2140 with ATP colored in green and F1522 with ATP colored in purple.

To further characterize the properties of the FeS cluster taXPD WT was incubated with different DNA substrates and then treated with Co-(5-Cl-phen)$_3^{3+}$ prior to EPR analysis. In the presence of ssDNA 7% of the clustersare oxidized compared to13% in the sample without DNA (Table 4.5, Figure 4.21). In the presence of dsDNA 13% of the clusters are oxidized. It was also tested what effect the presence of ATP and damage-containing substrates have on the cluster (Table 4.5).

In the presence of DNA and ATP and in the presence of Fluorescein containing DNA the cluster is oxidized more readily. In the presence of the 5'-overhang substrate mjd1/ndb the cluster is almost EPR silent, only 2% are oxidized. By addition of ATP the oxidation of the cluster increases to 20%. With a damage on the translocating strand (F2140) 42% of the cluster is oxidized and by ATP addition this increases to 50%. When the damage is located on the non-translocating strand (F1522) 25% of the cluster is oxidized. In the presence of ATP it is increased to 32%. It seems that the substrate shields the cluster from oxidizing agents and the addition of ATP leads to a conformational change which leads to a higher accessibility of the cluster for the oxidizing agent. In addition the fluorescein on either of the two strands leads to a structural change around the cluster which seems to make it more accessible for the oxidant.

Table 4.5: Iron-sulfur cluster properties of taXPD in presence of DNA

DNA substrate	[oxidized cluster]	% oxidation
ssDNA 40mer	3.7 µM	7 %
dsDNA 30mer	6.7 µM	13 %
Mjd1/ndb	1 µM	2 %
Mjd1/ndb & ATP	10 µM	20 %
F2140/NDB24	21 µM	42 %
F2140/NDB24 & ATP	25 µM	50 %
F1522/NDB40	13 µM	25 %
F1522/NDB40 & ATP	16 µM	32 %

4.15.2 Iron-sulfur cluster analysis of taXPD variants

To analyze the FeS cluster properties of selected taXPD variants the EPR was performed as described in 3.4.9. C92S and C113S were chosen because the cysteine residues coordinate the FeS cluster. Arg88 stabilizes the cluster and

interacts with DNA; furthermore the R88H variant (R112H in human XPD) causes trichothiodystrophy in humans.

Table 4.6: Iron-sulfur cluster properties of taXPD variants

mutant	[oxidized cluster]	% oxidation
R88H XPD	0 µM	0 %
R88H & Co(5-Cl-phen)$_3^{3+}$	31 µM	27 %
C92S	14 µM	27 %
C92S & Co(5-Cl-phen)$_3^{3+}$	38 µM	75 %
C113S	10 µM	90 %

In untreated taXPD WT only 2% of the FeS cluster is oxidized to a [3Fe4S]$^{1+}$ cluster, for the C92S variant it is 27% and for the C113S mutant it is 90% indicating that the cluster in C92S and C113S is not fully coordinated (Table 4.6, Figure 4.22). However the R88H mutant is similarto taXPD WT with 0% of the cluster oxidized. In the presence of Co-(5-Cl-phen)$_3^{3+}$ R88H is oxidized by 27% and C92S even by 75%. This indicates that the cluster in the variants C92S and C113S is not as stable as in taXPD WT and in C113S even only a [3Fe4S]$^{1+}$ cluster is present.

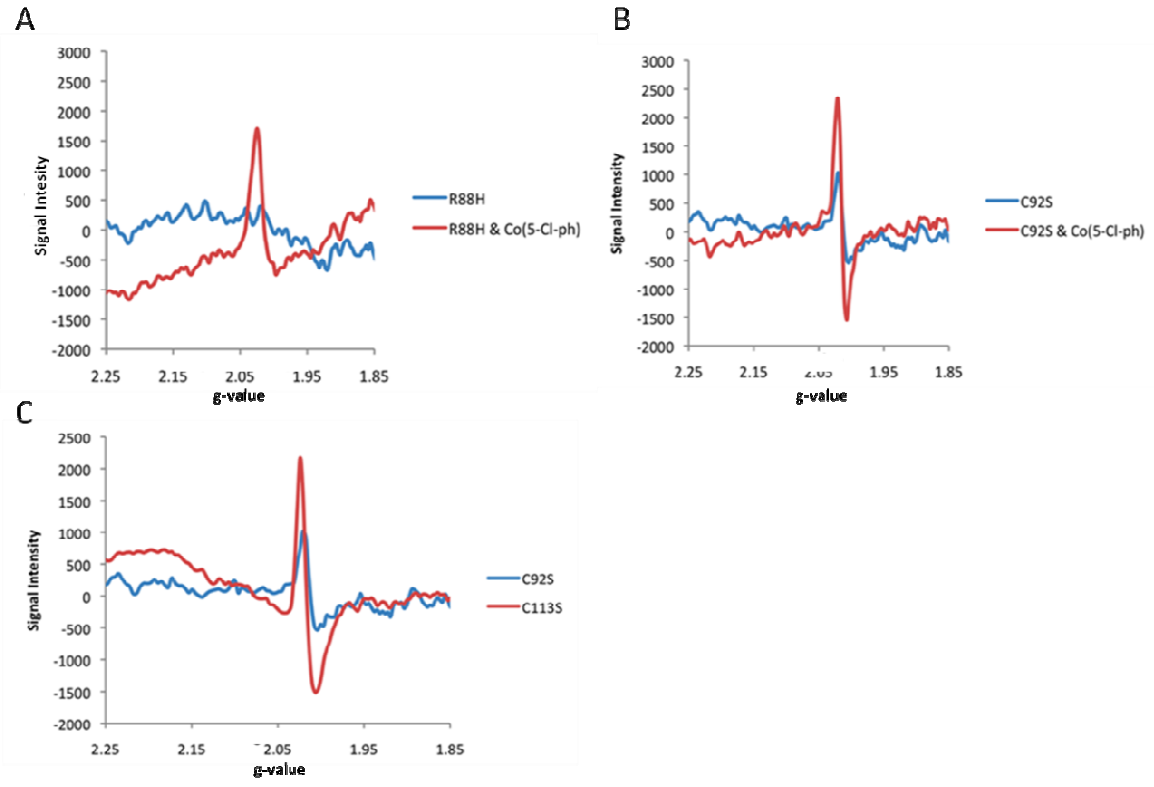


Figure 4.22: EPR Spectra of taXPD variants. **(A)** Spectrum of R88H colored in blue, R88H with Co-(5-Cl-phen)$_3^{3+}$ colored in red. **(B)** Spectrum of C92S colored in blue, C92S with Co-(5-Cl-phen)$_3^{3+}$ colored in red. **(C)** Spectrum of C92S colored in blue and spectrum of C113S colored in red.

4.15.3 taXPD WT shows reduced helicase activity in the presence of oxidizing and reducing agents

To investigate whether an intact 4Fe4S cluster is needed for helicase activity or if taXPD is also active when it contains a 3Fe4S cluster helicase assays were performed in the presence of oxidizing and reducing agents. The assays were conducted as described in 3.4.8 with slight modifications. The following buffer was used 20 mM Tris (pH 8.5), 10 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA and taXPD, DNA and oxidizing and reducing agents were used in the following concentrations 500 nM, 50 nM and 2.5 µM, respectively. The measurements were carried out at 37℃.

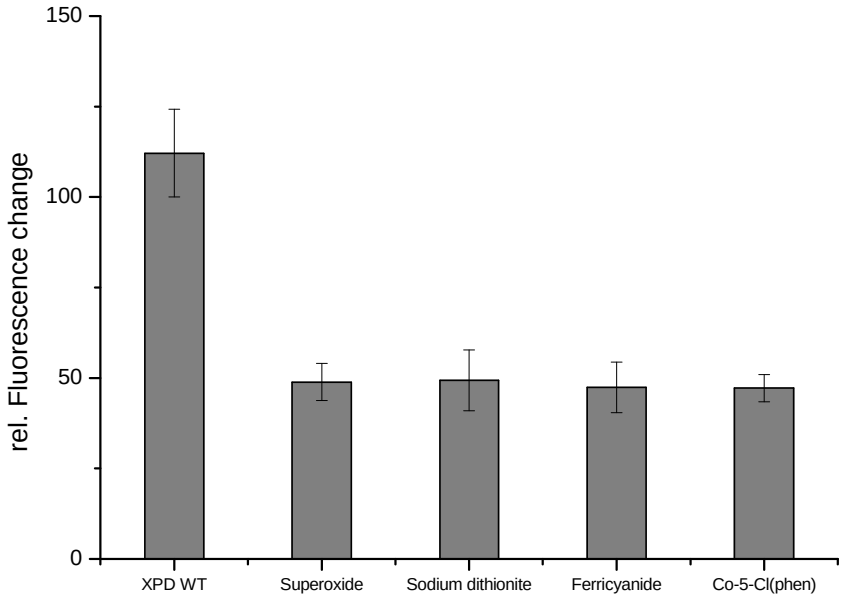


Figure 4.23: Helicase activity of taXPD WT in the presence of oxidizing and reducing agents.

Preliminary data revealed that treatment by superoxide, sodium dithionite, ferricyanide and Co-(5-Cl-phen)$_3^{3+}$ leads to a 50-fold decrease in helicase activity (Figure 4.23). However, treatment with sodium dithionite and ferricyanide didn't show an effect in EPR whereas treatment by superoxide and Co-(5-Cl-phen)$_3^{3+}$ resulted in an oxidation to a [3Fe4S]$^{1+}$ cluster. These helicase assays have to be further investigated and verified.

4.15.4 XPD is folded properly in the presence of different oxidants and reducing agents

To verify that the reduced helicase activity of taXPD WT in the presence of different agents is not due to a partially or completely unfolded protein CD spectroscopy was performed as depicted in 3.4.3. The CD spectra (Figure 4.24) show that untreated XPD is properly folded in an α/β-architecture. Treatment of XPD with superoxide, ferricyanide, sodium dithionite and Co-(5-Cl-phen)$_3^{3+}$ shows no effect on the overall fold of the protein indicating that these agents do not harm the protein which is in line with missing junk-iron peaks in EPR. However a conformational change without change of the secondary structure would not be detectable by CD spectroscopy.

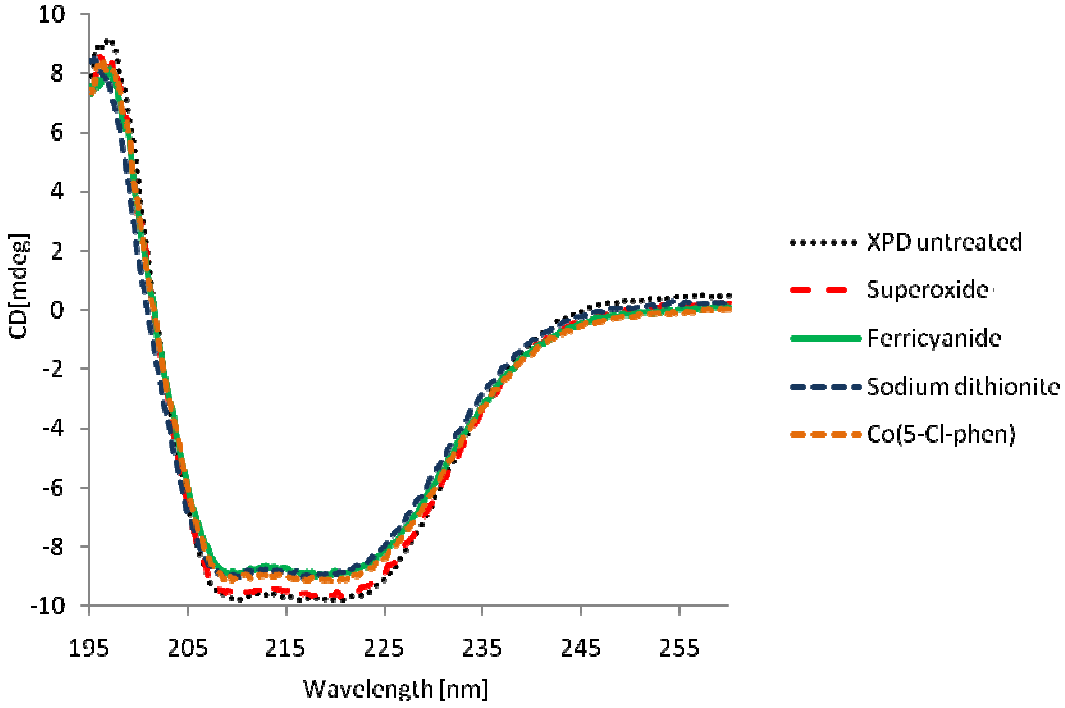


Figure 4.24: CD spectra of taXPD treated with different agents. Untreated taXPD colored in black. Superoxide, Ferricyanide, Sodium dithionite and Co-(5-Cl-phen)$_3^{3+}$ treated samples colored in red, green, blue and orange, respectively.

5 Discussion

5.1 The XPD crystal structure

XPD is a 5'-3'helicase belonging to the superfamily 2 (SF2)(Singleton et al. 2007). As a subunit of TFIIH XPD functions in nucleotide excision repair and transcription (Coin et al. 1998). Because of the high sequence homology to eukaryotes *Thermoplasma acidophilum* was chosen as a model organism to obtain insights into the human NER pathway and the XPD helicase in particular. The function of XPD in archaea still remains unclear.

In this work the structure of apo taXPD was solved, determined by multiwavelength anomalous diffraction using only the endogenous iron of the iron sulfur cluster. Size exclusion chromatography revealed that XPD is a monomer in solution as reported previously (Rudolf et al. 2006). As a subunit of TFIIH XPD interacts with p44, however p44 is not present in archaea. Therefore it is not remarkable that in contrast to human XPD the archaeal homologs lack p44 interacting residues at its C-Terminus. XPD contains four domains. The two RecA-like domains which are the motor domains of the protein and two insertions into motor domain 1, the arch-domain and the FeS domain. Domain 1 together with the two inserted domains form a donut-shaped structure with a pore in the middle.

Superposition of XPD from *Sulfolobus acidocaldarius* und taXPD revealed that both proteins display a four domain organization (Figure 5.1). They can be superposed with an rms deviation of 3Å indicating structural similarity, however the sequence identity is with 19% relatively low, but all disease related mutations are included in the 19%. The major difference between the two structures is the size of the pore. In taXPD the pore has a diameter of approximately 13 Å whereas the pore in saXPD only has a diameter of 8 Å. Therefore the pore in taXPD is sufficient in size to accommodate ssDNA whereas the pore size of saXPD is too small (Figure 5.1). The difference in pore size is achieved by a different positioning of the arch and theFeS domain relative to each other. This variation could be mediated by domain 4. Helix α22 from domain 4 would push the β-bridge of the arch-domain leading to a 10° tilt in α22. Furthermore saXPD displays a more closed conformation of the two motor domains. Therefore this mechanism could link ATP-hydrolysis to pore size dynamics. This is also supported by a normal mode analysis which reveals that movement of the two motor domains has an

effect on the pore. Interestingly, XPD has to open up when it acts on its natural substrate. This could be achieved by a transient opening between the arch and FeS domain which is in line with all three DNA binding models (Fan et al. 2008; Liu et al. 2008). Moreover single moleculestudies showed that during translocation

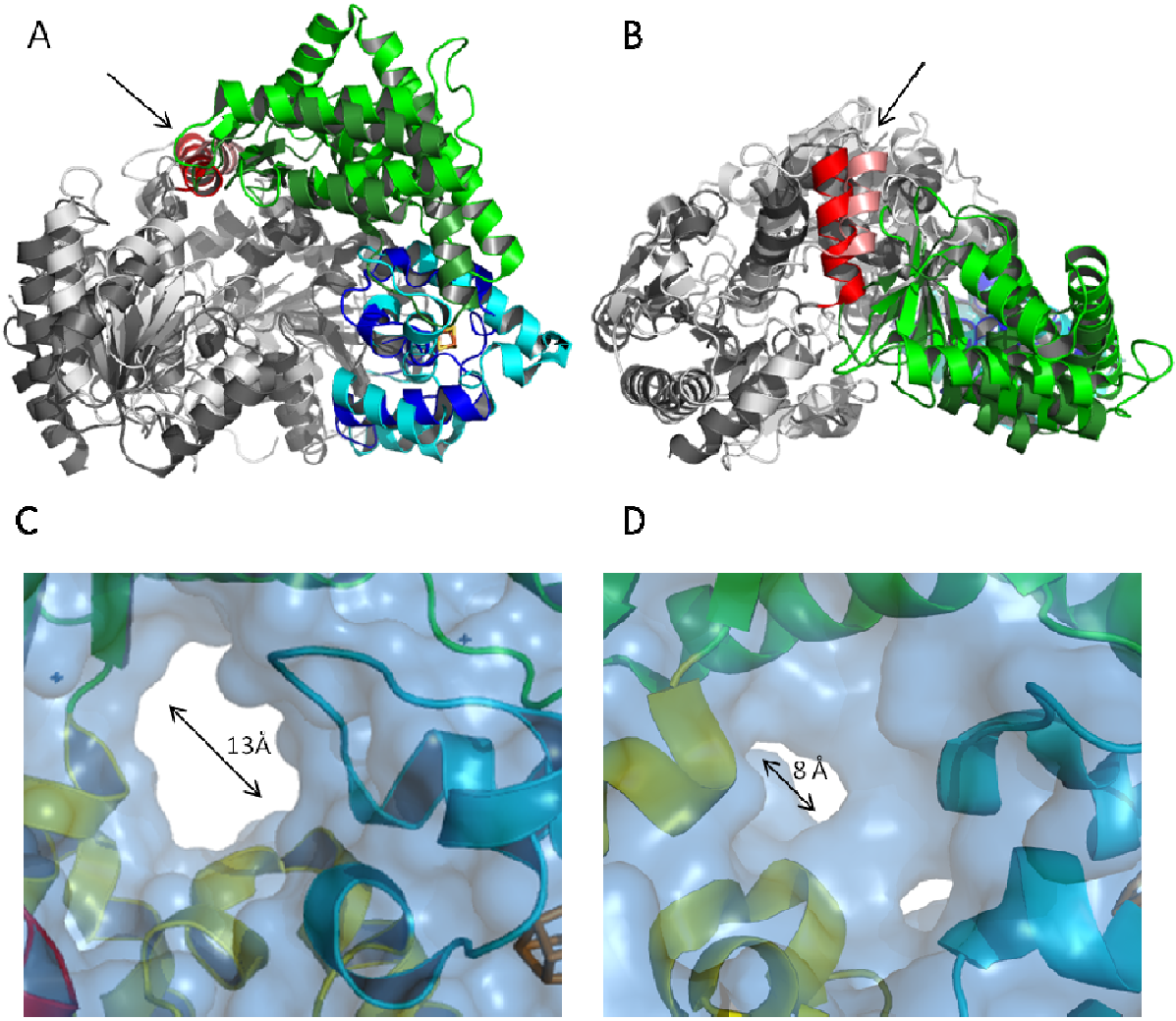


Figure 5.1: Superposition of taXPD and saXPD. The FeS and arch domain of taXPD are colored in cyan and green, respectively.The XPD helicaseHelix α22 of domain 4 is colored in red and highlighted by an arrow. The FeS andarch domain of saXPD are shown in blue and dark green, respectivley, and the helixcorresponding to helix α22 from taXPD is colored in salmon. The FeS cluster of taXPDis shown in orange. Domains 3 and 4 are colored in dark grey (taXPD) and grey(saXPD), respectively. **(A)** Front view, **(B)** top view. **(C)** Surface representations of thepore of taXPD and **(D)** the pore of saXPD.

XPD can bypass SSB proteins indicating a highly dynamic enzyme (Honda et al. 2009). Our biochemical data combined with the structure of the XPD-DNA complex specifies the path of and polarity of the translocated strand according to the helicase framework. The DNA is bound at domain 4 with the 5'-end of the DNA pointing towards the solvent and the 3'-end pointing towards domain 3.

Based on the XPD structures several questions arise: whereis the DNA bound? Where is DNA separated? Which residues are important for helicase functionality? Where are the disease causing mutations located in the tertiary structure? What is the role of the FeS cluster? How is damage verification achieved? These questions were addressed in the present work and are discussed here.

5.2 Disease causing mutations

XPD harbors phenotypes of three different diseases and is part of at least three different complexes TFIIH, CAK/XPD and MMXD. XPD is therefore involved in different cellular processes. TFIIH in humans is involved in nucleotide excision repair and transcription initiation as a general transcription factor for RNA polymerase II(Svejstrup 2002).XPD not only forms a complex with CAK in TFIIH but also with CAK alone(Ito et al. 2010). As part of a complex with CAK XPD may be involved in cell cycle regulation and as part of the MMXD(*M*MS19-*M*IP18-*XPD*)complex in chromosome segregation (Ito et al. 2010). XPD thereby has at least three different interaction partners. The interaction of p44 with XPD in TFIIH stimulates XPDs helicase activity (Coin et al. 1998). XPD interacts in TFIIH but also by itself with the CAK subunit MAT1 (Reardon et al. 1996), (Drapkin et al. 1996) and the MMXD subunit MMS19 (Ito et al. 2010). Mutations in XP-D can therefore lead to different effects. The first group of mutations can lead to an impaired interaction with p44 or MAT1 or MMS19 and thereby affecting XPDs activity in an indirect way. In the case of p44 and MAT1 this can also lead to a destabilization of the TFIIH core (without the CAK subunits) or the entire TFIIH complex. Furthermore the point mutants can affect the activity of XPD directly. Based on the structure the phenotype of mutations leading to xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy can be rationalized on a molecular level unparalleled so far. Mutations leading to xeroderma pigmentosum such as G47R, D234N and R666W (corresponding to G32R, D211N and R567W in taXPD) are located in helicase motifs I, II and VI (Figure 4.12), respectively and affect DNA and ATP binding thus inactivating the enzyme. This is also supported by biochemical data of the taXPD R567W variant which displays impaired DNA binding. However point mutations within other regions have quite distinct effects (Figure 5.2).

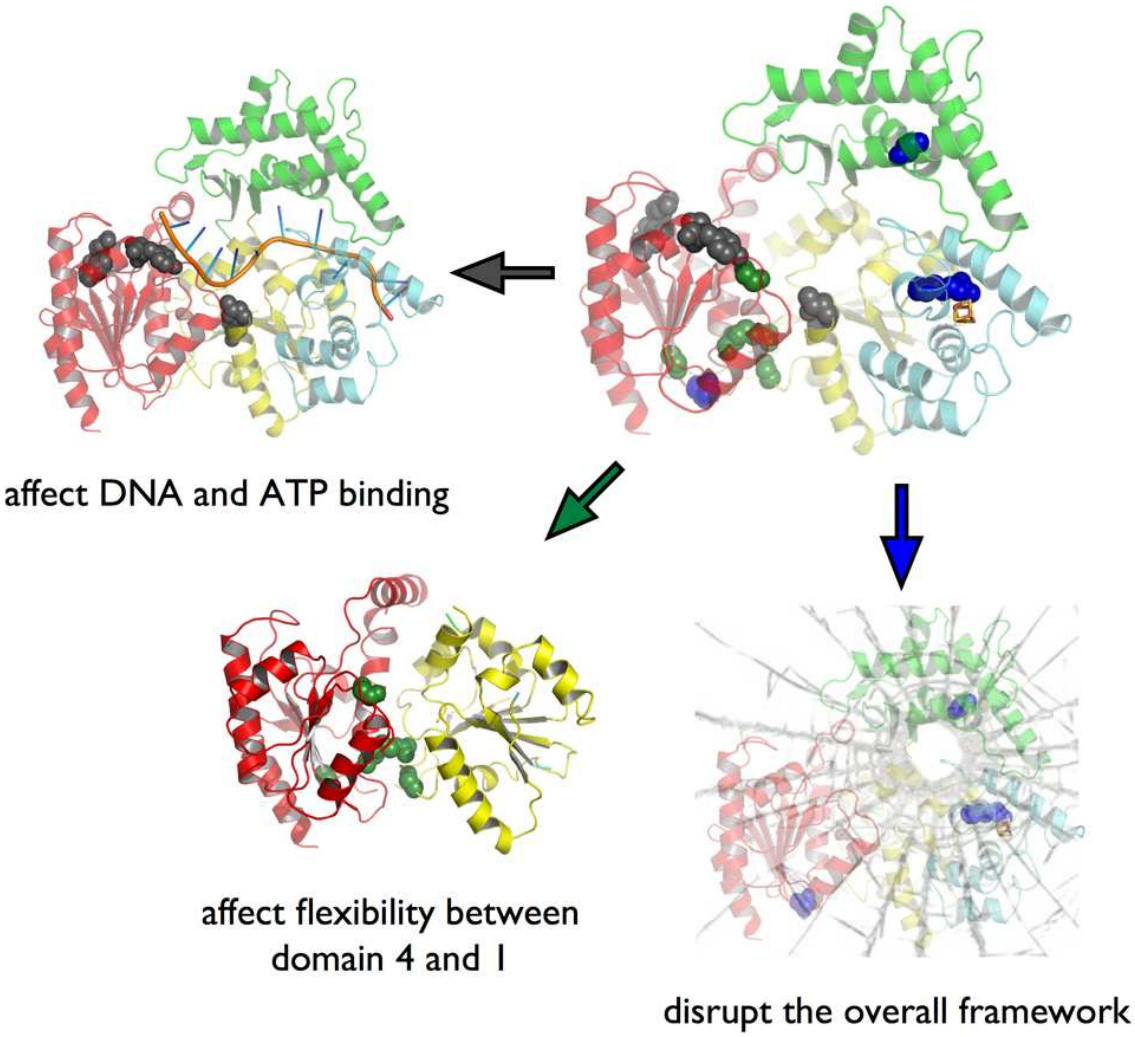


Figure 5.2: Mutants leading to xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy. XPD is shown in ribbon presentation and color coded as in Figure 4.6. Point mutations leading to either XP, XP/CS or TTD are shown in space-filling representation and are labeled according to the disease they cause in humans, black (XP), green (XP/CS) and blue (TTD).

In humans the replacement of Arg112 (Arg88 in *Thermoplasma acidophilum*) by an histidine leads to trichothiodystrophy (Lehmann 2001). Arg88 points into a pocket formed by the residues Tyr166 and Tyr185. This pocket can directly interact with DNA. Arg88 is in close neighborhood of Cys113 an Iron-sulfur-cluster coordinating residue. Arg88 stabilizes this cluster by a hydrogen bond. Furthermore with its long side chain it protects the cluster from solvent exposure. The mutation to a histidine at this position leads to steric clashes and thereby to a

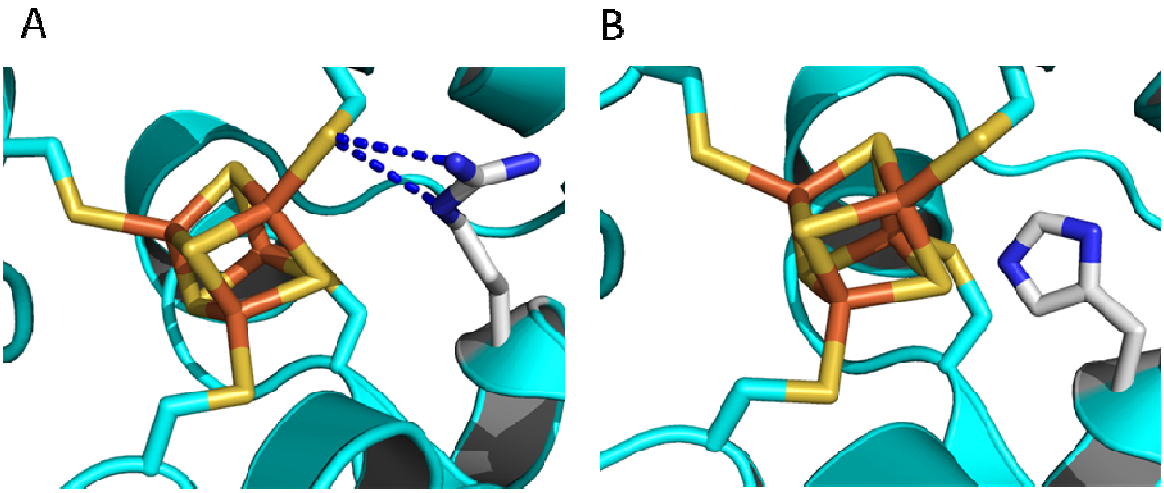


Figure 5.3: Zoom into the FeS cluster domain. **(A)**Arg88 is shown in sticks. H-Bonds are indicated by blue dashes. **(B)** Similar view as in panel **(A)** for the trichothiodystrophy-causing variant R88H.

destabilization of the protein. Moreover the R112H variant provides a shorter side chain which cannot donate a hydrogen bond to Cys113 leading to a destabilization. For the corresponding K84H variant from *Sulfolobus acidocaldarius* it was demonstrated that the mutation has no affect on the ATPase activity and DNA binding ability, however the helicase activity is affected (Rudolf et al. 2006), (Liu et al. 2008). The R112H variant in humans shows the same phenotype (Dubaele et al. 2003). The taXPD variant displays a slightly different behavior. There is also no effect on ATPase activity, the DNA binding ability is moderately impaired, however the R88H variant is still helicase active which might be due to a more stable FeS cluster in *Thermoplasma acidophilum*compared to the human and *Sulfolobus acidocaldarius*XPD protein. As discussed later on there might be an additional important role in damage verification related to this particular residue.

In humans the variant C259Y causes trichothiodystrophy (Lehmann 2001). In *Thermoplasma acidophilum* this residue corresponds to Ala236. Four α-helices in the arch domain are stabilized by a hydrophobic core. Ala236 points into this hydrophobic core. An exchange to a much larger tyrosine leads to steric clashes thereby destabilizing the entire domain.

The four mutations D673G, G675R, D681N and R683W/Q. correspond to D574G, G576R, D582N and R584E in *Thermoplasma acidophilum*. All these residues are

located in the C-terminal motor domain 2 and fulfill important structural roles or are involved in DNA interactions. Furthermore it was demonstrated that XPD interacts via its C-Terminus with p44 (Coin et al. 1998). G575R was analyzed with respect to its p44 interaction which revealed that this interaction is severely impaired (Coin et al. 1998). However in archaea there is no p44 homolog and therefore it is not surprising that XPD lacks residues at its C-Terminus which are important for p44 interaction. Asp574 interacts with Arg570 (Arg669 in human XPD) thereby stabilizing the transition between α-helix 24 and β-strand β16. A replacement of glycine in residue 576 by an arginine would lead to a destabilization of domain 4 since Gly576 points towards two hydrophobic residues and a much larger amino acid would increase the distance to these residues. Asp582 interacts with the strictly conserved residue Arg584 (corresponds to Arg683 in humans) which itself interacts with the two residues Asp426 and Phe527. However, our ThermoFluor data indicate that the D582N and the R584E variants have no effect on the stability of the protein and the exact role for the disease phenotype has to be elucidated in future studies.

It was shown that the function of MMXD in chromosome segregation is affected in XP, and CS cells but not in TTD cells (Ito et al. 2010). The interaction of MMS19, a subunit of the MMXD complex, and XPD is mediated by residues 438-637 indicated by pull down assays (Ito et al. 2010). However, this interaction can be affected by different mutations which are located in the region important for the interaction. Y542C and G602D (Tyr458 and Gly504 in *Thermoplasma acidophilum*) lead to XP and XP/CS, respectively. The two residues are very close to each other in the structure. Tyr458 is located in α-helix 20 of domain 4. It forms hydrophobic interactions with Val501 (Val599 in humans) which is a strictly conserved residue. The exchange from tyrosine to a cysteine would weaken the Tyr458/Val501 interaction. Gly504 is located between the two β-strands β14 and β15 in domain 4. A replacement of the glycine by an aspartic acid would lead to steric clashes with residue Tyr458. The mutations Y542C and G602D could therefore lead to an impairment of the MMS19-XPD interaction by affecting the interaction interface of the two proteins and therefore lead to an abnormality in chromosome segregation (Ito et al. 2010). However it was shown by Ito *et al.* that G602D did not affect the activity to bind MMS19(Ito et al. 2010), but it would be possible that the mutations affect somehow the enzymatic activity of XPD. These

findings suggest that abnormalities in the function of MMXD due to the XPD mutations lead to aneuploidy and /or cell death (Ito et al. 2010) and might be mediated by a difference in XPD's activity.

From this analysis it can be concluded that mutations leading to XP affect the DNA and ATP binding ability, whereas mutations leading to XP/CS affect the flexibility between the two motor domains and mutations leading to TTD affect the overall stability of XPD. Furthermore the mutations can affect the enzymatic activity of XPD by an impairment in the interaction of XPD with MMS19 and p44.

5.3 Conclusions from the mutagenesis studies

The XPD-DNA complex structure revealed that Trp549 interacts with DNA forming π-stacking interactions. Therefore it is not surprising that when mutated to a serine the DNA binding ability is severely impaired. Y545S and F538S show a similar phenotype with reduced DNA binding ability and impaired helicase and ATPase activity. All three residues are located in a region in domain 4 with several hydrophobic residues (Figure 5.4). Trp549 is not located in a strictly conserved region, Tyr545 and Phe538 however are conserved. Moreover neighboring amino acids are highly conserved and could fulfill in eukaryotes similar roles as Trp549. For example a phenylalanine (Phe652 in human XPD) which is highly conserved and located three residues C-terminally.

The two variants D582N and R584E (correspond to D681N and R683W/Q in human XPD) which cause XP in humans show severe effects in DNA binding, ATPase activity and in the ability to unwind dsDNA. The XPD-DNA complex revealed that Arg584 significantly interacts with the DNA backbone. Furthermore ThermoFluor analysis demonstrated that both variants are properly folded so that the disease phenotype cannot be caused by partially unfolded protein. The Y425E variant is also located in this wedge-like region and displays a moderate DNA binding phenotype but a strong reduction in ATPase and helicase activity. Since this residue is disordered in the XPD-DNA complex structure no conclusions with respect to a DNA-complex formation for this residue can be derived from this structure.

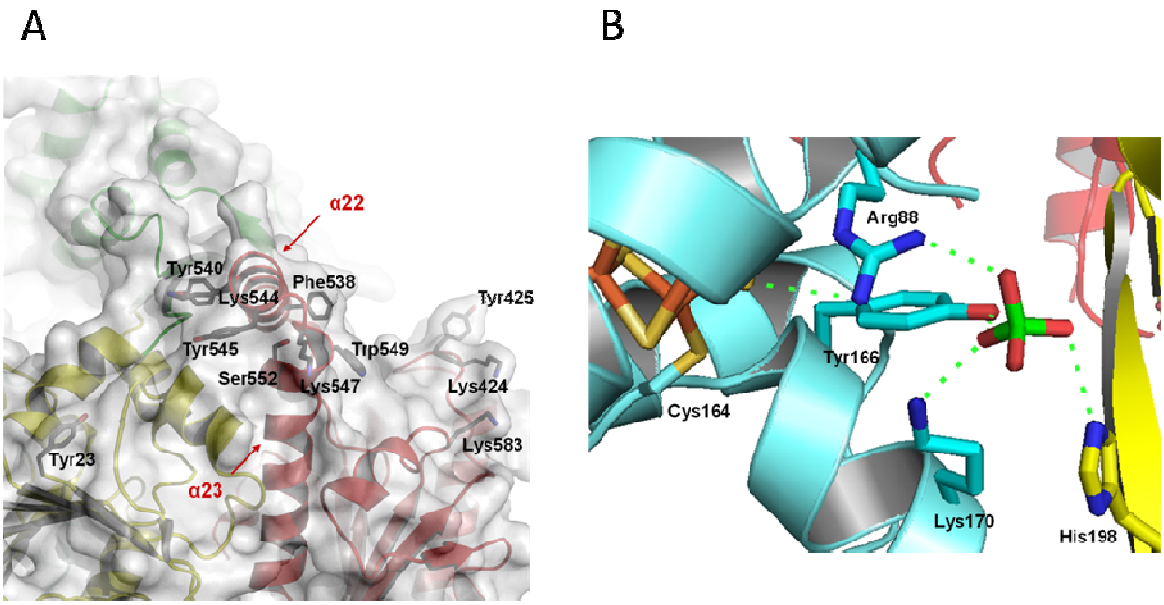


Figure 5.4: Important Residues for DNA binding. (A) A view towards helices α22 and α23 in domain 4. The color code from figure 5.1 was maintained, and a transparent surface was added to provide a view of the groove were one DNA strand binds. All residues that may play an important role in DNA binding are indicated and labeled. (Wolski et al. 2008)**(B)** Putative phosphate binding site. Color code from Figure 4.6 with the FeS cluster and important residues shown in all-bonds representation.

A bound sulfate between motor domain 1 and the arch domain indicates a putative phosphate binding site (Figure 5.4). Residues which interact with the sulfate were mutated and resulted in different phenotypes. The variants R88H, Y166A and K170A display reduced DNA binding and increased ATPase activity. The ability to unwind dsDNA however varies greatly between the three variants. R88H has a helicase activity comparable to taXPD WT, Y166A displays no helicase activity and K170A shows an elevated helicase activity and therefore seems to be a negative regulator for helicase activity. Interestingly, Y166A shows the smallest difference in DNA binding to WT in all three mutants. Tyr166 and Tyr185 form a hydrophobic pocket which is a putative DNA base binding site and sufficient in size to bind purine and pyrimidine bases. The Y166A/Y185A variant displays reduced DNA binding and ATPase activity, however the mutant did not over express well and was therefore not further pursued. F133A shows a similar phenotype as K170A with an increased helicase activity, however the ATPase activity is comparable to taXPD WT and the DNA binding ability is reduced. Arg118 is located near Phe133 on the surface of the FeS domain and displays a decreased

K_D and a therefore reduced ATPase activity. The E107A is located directly in the pore and shows a reduced DNA binding ability but interestingly the ATPase and helicase activity are comparable to taXPD WT.

The replacement of Arg112 (Arg88 in *Thermoplasma acidophilum*) to a histidine leads in humans to trichothiodystrophy. In contrast to the observation in *Sulfolobus acidocaldarius*(Rudolf et al. 2006) the R88H variant in *Thermoplasma acidophilum* displays no helicase phenotype. Furthermore R88H displays an elevated ATPase activity and a reduced DNA binding ability which is not surprising since Arg88 points into a basic pocket which is proposed to be a DNA binding site. Arg88 bridges the putative DNA binding site with the Cys113 cluster ligand. The exchange from arginine or lysine in *Sulfolobus acidocaldarius* to histidine shows a similar phenotype for the human protein and XPD from*Sulfolobus acidocaldarius* with an abolished helicase activity (Rudolf et al. 2006), (Dubaele et al. 2003). This observation can be explained by the loss of the FeS cluster indicating a more stable cluster in taXPD.

Phe326 is partly located in the interface of the arch domain and the FeS domain. F326A displays a decreased helicase activity and a moderate DNA binding phenotype combined with an elevated ATPase activity. The phenylring of Phe326 points towards the pore and could therefore be involved in π-stacking interactions with DNA. The location of F326A may affect the flexibility between the arch domain and FeS domain which is necessary for helicase activity.Our XPD-DNA structure combined with the biochemical data identified regions for DNA binding. However it is not clear were DNA separation takes place. Due to the polarity of the translocated strand the DNA has to be separated behind the pore. Possible regions could be on the FeS domain close to the residues Phe133 and Arg118 or the arch domain with Phe326 could act as a wedge to separate DNA duplexes.

5.4 The role of the FeS cluster

The importance of the FeS cluster of XPD for helicase activity was investigated by Rudolf et al. (Rudolf et al. 2006). They could demonstrate that the helicase activity of XPD is abolished when it loses the FeS cluster (Rudolf et al. 2006). MutY is an FeS cluster containing DNA-glycosylase and was studied structurally in the

presence of DNA (Fromme et al. 2004). Therefore MutY was compared with XPD. For MutY it was shown that the FeS cluster is required for DNA binding and enzymatic activity (Porello et al. 1998). The superposition of the FeS cluster domains of both proteins revealed that the orientation of two conserved arginine residues (Arg88 in taXPD and Arg153 in MutY) is similar (Figure 5.5). For MutY it was shown that the neighboring Arg149 is located ideally to interact with the DNA backbone (Guan et al. 1998). Arg88 has a similar location as it points into a conserved pocket and thereby could interact with DNA. This pocket is furthermore a potential position for damage recognition. MutY contains an FeS cluster loop motif (Lukianova and David 2005) and the FeS cluster plays an important role in arranging the residues of this motif. A loop motif(residues 160-178) with a high density of positively charged residues is also found in the FeS domain of XPD and the FeS cluster could have a similar function in the positioning of the residues as in MutY (Lukianova and David 2005).

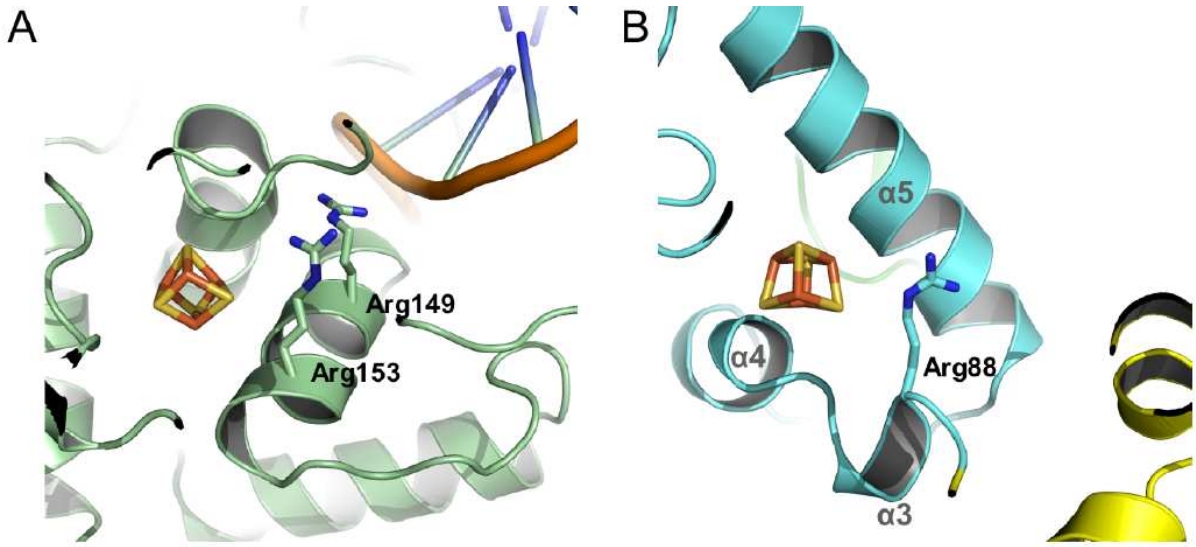


Figure 5.5: Structural similarity to MutY. Side-by-side presentation of the 4Fe4S clusters of MutY and XPD after superposition. **(A)** Tow arginines (Arg149 and Arg153) are located in close proximity to the 4Fe4S cluster and point towards the DNA in MutY. **(B)** In XPD, Arg88 occupies a similar position as observed for the arginines in MutY. The helices surrounding the cluster have been labeled (Wolski et al. 2008).

Boon et al. showed for MutY that DNA binding modulates the redox properties of the 4Fe4S cluster (Boon et al. 2003). Upon DNA binding the cluster is activated and thereby oxidized (Boal et al. 2009). It was proposed that via the two arginines

(Ar149 and Arg153) one electron could be transferred from the FeS cluster to the DNA (Boal and Barton 2005). This DNA-mediated charge transfer could accomplish rapid DNA scanning by several glycosylase molecules and they could thereby locate sites of damage (Boal et al. 2007). A similar mechanism seems unlikely in NER. However it is possible that DNA binding mediates the redox properties of the cluster and is therefore necessary for the damage verification step. This would suggest not only a structural but a functional role for the FeS cluster in XPD. Mutagenesis studies in *Sulfolobus acidocaldarius*(Rudolf et al. 2006)revealed the importance of the FeS cluster coordinating cysteines. Single mutations of three of the four cysteines to serine leads to a loss of the FeS cluster and helicase activity is abolished. However it seems that the cluster in *Thermoplasma acidophilum* is more stable since the C92S variant contains a $[4Fe4S]^{2+}$ cluster and C113S contains a stable $[3Fe4S]^{1+}$ cluster. Helicase activity of the *Thermoplasma acidophilum* variants C92S and C113S has to be further investigated to verify if the enzyme activity is depending on the FeS cluster redox state.

To obtain detailed insights into the FeS cluster properties they were further investigated by EPR analysis. Therefore XPD was treated with different oxidants and reducing agents. Interestingly, only Co-$(phen)_3^{3+}$, Co-$(5\text{-Cl-phen})_3^{3+}$ and superoxide showed an effect on the cluster indicated by an oxidation to a $[3Fe4S]^{1+}$ cluster. Whereas XPD was resistant to treatment by $Ir(Cl)_6^{2-}$, ferricyanide, sodium dithionite, nitric oxide, peroxynitrite and hydrogen peroxide indicating a stable $[4Fe4S]^{2+}$ cluster. The resistance to hydrogen peroxide suggests that XPD could be functional under oxidative stress conditions which corresponds to results for *E.coli* DinG (Ren et al. 2009). Furthermore EPR analysis of XPD in the presence of different DNA substrates and ATP was investigated. The FeS cluster is oxidized more readily in the presence of DNA which is further increased if ATP is added and the motor domains moveindicating that the cluster is more accessible for oxidants upon DNA binding and movement of the helicase domains. The importance of the cluster for helicase activity is supported by the our preliminary data that treatment of XPD with Co-$(phen)_3^{3+}$, Co-$(5\text{-Cl-phen})_3^{3+}$, superoxide and sodium dithionite leads to a 50-fold reduction in helicase activity indicating that the $[3Fe4S]^{1+}$cluster is not as capable for helicase activity as the $[4Fe4S]^{2+}$cluster. This reduction is not due to unfolded protein since EPR

displayed no junk iron peaks which indicate the loss of the FeS cluster and CD spectroscopy showed that in the presence of oxidizing and reducing agents XPD is folded as the non treated sample. However it is not clear whether or not a $[3Fe4S]^{1+}$ cluster can perform helicase activity whichhas to be further investigated by additional helicase and ATPase assays in the presence of different oxidizing agents.

In the last few years more and more of iron-sulfur cluster containing proteins that interact with DNA or RNA have been reported. With respect to their functions these proteins can be divided into two groups (Ren et al. 2009). Group I contains the transcription or translation regulators for example the redox transcription factor SoxR (Ding et al. 1996), (Gaudu and Weiss 1996) and the repressor IscR (Schwartz et al. 2001). In this group of proteins the iron sulfur cluster acts as a specific signal sensor and modulates the interaction between protein and DNA and RNA (Ren et al. 2009). Group II consist of the iron-sulfur enzymes that chemically modify RNA or DNA. Members of this group are for example endonuclease III (Cunningham et al. 1989), (Boal et al. 2005), MutY (Porello et al. 1998), XPD and *E.coli* DinG helicase (Ren et al. 2009). It was demonstrated for the DinG helicase that its $[4Fe4S]^{2+}$ cluster is redox-active and modulated by the second messenger molecule nitric oxide (Ren et al. 2009). NO is a physiological free radical involved in signal transduction in neuronal and cardiovascular systems (Ignarro 1999). Some studies revealed that chronic NO exposure is linked to carcinogenic processes and genomic instability (Gal and Wogan 1996), (Li et al. 2002). Mutations in the XPD gene lead to diseases which are associated with an increase of cancer incidence and aging phenotypes (Lehmann 2001), (Rudolf et al. 2006), (Fan et al. 2008), (Liu et al. 2008) it is therefore possible that chronic NO exposure may inactivate the FeS cluster of DNA repair enzymes like XPD and DinG. In this way this contributes to the initiation of the carcinogenic process and genomic instability (Gal and Wogan 1996), (Li et al. 2002), (Ren et al. 2009).However EPR studies revealed that XPD is resistant to oxidation by NO since no signal for a $[3Fe4S]^{1+}$ cluster was obtained, but it has not been investigated so far if NO affects XPDs helicase activity. Thus it is unclear if XPD fulfillsa role in the NO signaling process.

5.5 Mechanistic conclusions from the XPD structure

In eukaryotes the initial damage recognition is achieved by the XPC/Rad23B complex. Afterwards TFIIH is recruited and with the combined action of XPD and XPB the DNA is unwound. In this process only the ATPase activity of XPB is needed whereas the helicase activity of XPD is necessary. Naegeli et al. showed that Rad3 the yeast homolog of XPD is stalled at sites of damage during translocation *in vivo*(Naegeli et al. 1993). Furthermore it was demonstrated for XPD from *Ferroplasma acidarmanus* that it is stalled at lesions containing a CPD on the translocating strand *in vitro*(Mathieu et al. 2010). Furthermore XPD's ATPase activity is stimulated in the presence of a CPD lesion (Mathieu et al. 2010).

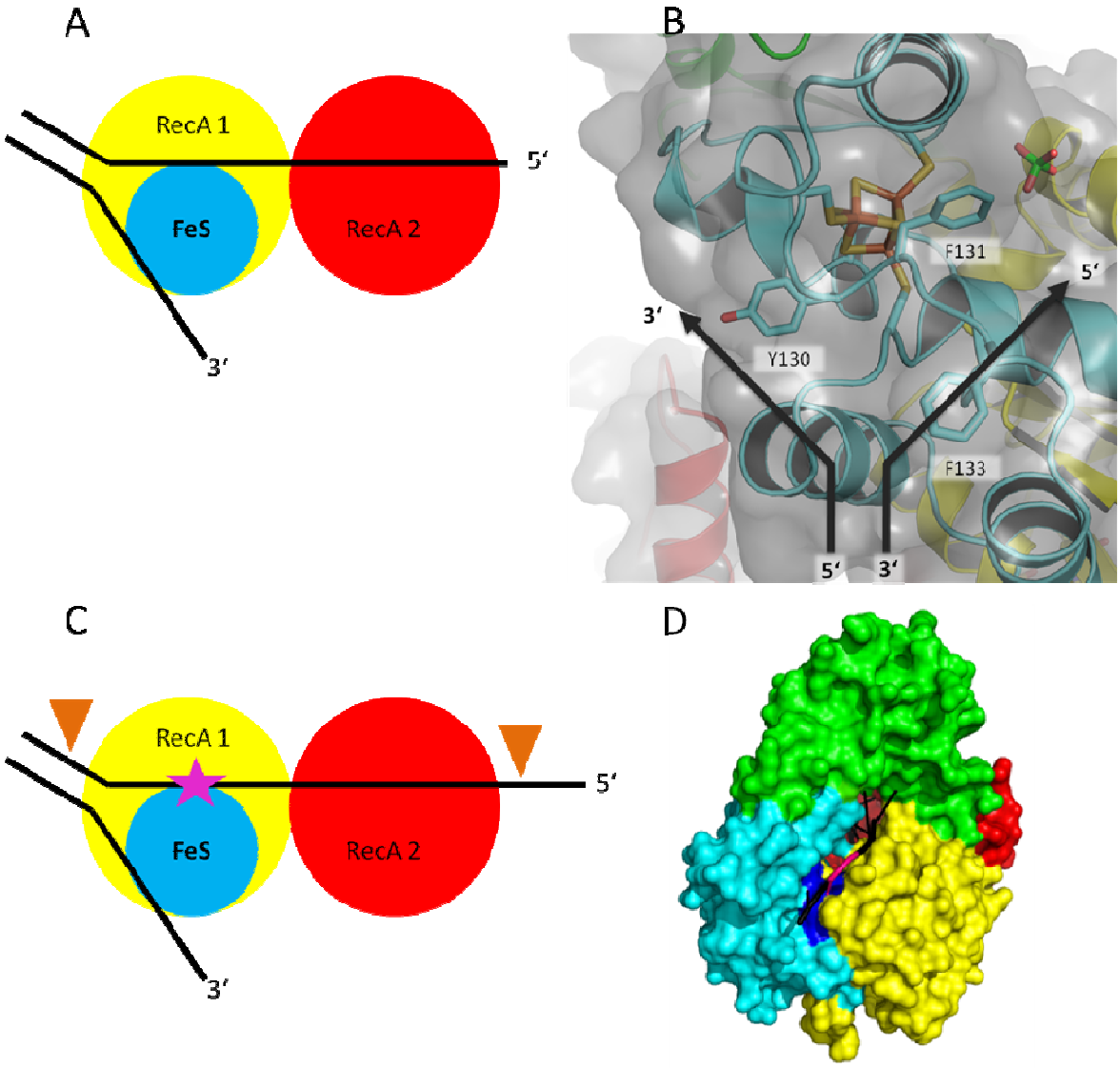


Figure 5.6: DNA separation and damage recognition. (A) A minimal model of XPD with the two motor domains colored in yellow and red and the FeS cluster domain colored in cyan. **(B)** Possible

DNA separation site with XPD color coded as in Figure 4.6. Important residues and the FeS cluster are displayed in all bonds representation. **(C)** A minimal model of XPD with the two motor domains color-coded as in (A). A damage in the DNA is indicated by a pink star and possible incision sites are indicated by orange triangles. **(D)** Surface representation of taXPD color coded as in (A) with the narrow pocketcolored in dark blue. The DNA is shown in black with a modified base depicted in pink.

UvrB is a homolog of XPD which was demonstrated to be involved in damage recognition. UvrB contains an β-hairpin which was shown to be essential for the damage recognition step (Truglio et al. 2006) and fulfills a similar function as XPD (Wang et al. 2006). Although XPD and UvrB display partially a high structural similarity XPD lacks a β-hairpin or similar feature. The question of damage verification in XPD therefore remains unsolved. The pore is with a diameter of 13 Å too large to be a trap for modified DNA. A good candidate for damage verification is the narrow pocket at the FeS domain formed by residues Arg88, Tyr166 and Tyr185. The two variants R88H and Y166A display a slightly reduced DNA binding ability in combination with an ATPase activity as taXPD WT (R88H) or an increase in ATPase activity (Y166A). However they differ in the helicase activity. R88H shows a helicase activity as taXPD WT whereas Y166A displays no helicase activity. This pocket is sufficient in size to accommodate non-modified purine and pyrimidine bases which would be held in place by van der Waals interactions with the residues mentioned above. However, modified bases would be excluded from the pocket. The location of the possible damage recognition site in close proximity to the pore is further supported by the results of restriction protection assays (Mathieu et al. 2010). Mathieu et al. demonstrated that the DNA was cut 16 bases prior (5'-3' direction) and 15 bases after the damage. DNA located close to the damage was protected from incision. Based on our structure and the polarity of the DNA it can be furthermore suggested that dsDNA unwinding takes place at the FeS domain and approximately 20-25 bases can be accommodated by XPD (Figure 5.6). This is further supported by biochemical data of the group of Maria Spies (Pugh et al. 2008) and in line with other DNA binding models (Liu et al. 2008). For eukaryotic NER it was shown that the excised fragment varies but the damage is always located 2-9 bases from the 3'-end (Wood 1999). Therefore the narrow pocket mentioned above would be an ideal position for damage verification (Figure 5.6). Furthermore Arg88 is located in close proximity to the FeS cluster and could in analogy to MutY bridge the distance

between the DNA and the FeS cluster (Fromme et al. 2004), (Boal and Barton 2005) thus involving the cluster in the damage verification process.

5.7 Final conclusions

In this work the first crystal structure of apo XPD from *Thermoplasma acidophilum* was solved. The structure revealed a four-domain organization with two motor domains and two additional domains. The two auxiliary domains, the arch and the FeS domain, emerge from motor domain 1. Motor domain 1 together with the arch and FeS domain form a central pore. A ssDNA could pass through this central pore in close proximity to the FeS cluster as suggested by a DNA binding model. The suggested path of the DNA is supported by our structure of XPD with 4 bound DNA bases. It was possible to identify the polarity of DNA with the 5'-end pointing towards the solvent of domain 4 and the 3'end pointing towards the conserved path through the enzyme. Due to the high sequence homology to eukaryotic XPDs human mutations leading to xeroderma pigmentosum, Cockayne syndrome and trichothiodystrophy can be explained. Combined with a detailed mutagenic analysis important functional regions of the protein could be identified and important functions of the helicase during NER can be explained. These findings provide a basis for further analyses to understand the damage verification process in the NER pathway.

6 References

(1994). "The CCP4 suite: programs for protein crystallography." Acta Crystallogr D Biol Crystallogr 50(Pt 5): 760-763.

Aboussekhra, A., M. Biggerstaff, M. K. Shivji, J. A. Vilpo, V. Moncollin, V. N. Podust, M. Protic, U. Hubscher, J. M. Egly and R. D. Wood (1995). "Mammalian DNA nucleotide excision repair reconstituted with purified protein components." Cell 80(6): 859-868.

Abrahams, J. P. and A. G. Leslie (1996). "Methods used in the structure determination of bovine mitochondrial F1 ATPase." Acta crystallographica. Section D, Biological crystallography 52(Pt 1): 30-42.

Adams, P. D., P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart (2010). "PHENIX: a comprehensive Python-based system for macromolecular structure solution." Acta crystallographica. Section D, Biological crystallography 66(Pt 2): 213-221.

Al-Minawi, A. Z., N. Saleh-Gohari and T. Helleday (2008). "The ERCC1/XPF endonuclease is required for efficient single-strand annealing and gene conversion in mammalian cells." Nucleic acids research 36(1): 1-9.

Araki, M., C. Masutani, M. Takemura, A. Uchida, K. Sugasawa, J. Kondoh, Y. Ohkuma and F. Hanaoka (2001). "Centrosome protein centrin 2/caltractin 1 is part of the xeroderma pigmentosum group C complex that initiates global genome nucleotide excision repair." J Biol Chem 276(22): 18665-18672.

Battye, T. G., L. Kontogiannis, O. Johnson, H. R. Powell and A. G. Leslie (2011). "iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM." Acta Crystallogr D Biol Crystallogr 67(Pt 4): 271-281.

Boal, A. K. and J. K. Barton (2005). "Electrochemical detection of lesions in DNA." Bioconjug Chem 16(2): 312-321.

Boal, A. K., J. C. Genereux, P. A. Sontz, J. A. Gralnick, D. K. Newman and J. K. Barton (2009). "Redox signaling between DNA repair proteins for efficient lesion detection." Proc Natl Acad Sci U S A 106(36): 15237-15242.

Boal, A. K., E. Yavin and J. K. Barton (2007). "DNA repair glycosylases with a [4Fe-4S] cluster: a redox cofactor for DNA-mediated charge transport?" J Inorg Biochem 101(11-12): 1913-1921.

Boal, A. K., E. Yavin, O. A. Lukianova, V. L. O'Shea, S. S. David and J. K. Barton (2005). "DNA-bound redox activity of DNA repair glycosylases containing [4Fe-4S] clusters." Biochemistry 44(23): 8397-8407.

Boon, E. M., A. L. Livingston, N. H. Chmiel, S. S. David and J. K. Barton (2003). "DNA-mediated charge transport for DNA repair." Proc Natl Acad Sci U S A 100(22): 12543-12547.

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

Bradsher, J., F. Coin and J. M. Egly (2000). "Distinct roles for the helicases of TFIIH in transcript initiation and promoter escape." J Biol Chem 275(4): 2532-2538.

Brooks, P. J. (2007). "The case for 8,5'-cyclopurine-2'-deoxynucleosides as endogenous DNA lesions that cause neurodegeneration in xeroderma pigmentosum." Neuroscience 145(4): 1407-1417.

Brueckner, F., U. Hennecke, T. Carell and P. Cramer (2007). "CPD damage recognition by transcribing RNA polymerase II." Science 315(5813): 859-862.

Brunger, A. T. (1997). "Free R value: cross-validation in crystallography." Methods Enzymol 277: 366-396.

Busso, D., A. Keriel, B. Sandrock, A. Poterszman, O. Gileadi and J. M. Egly (2000). "Distinct regions of MAT1 regulate cdk7 kinase and TFIIH transcription activities." J Biol Chem 275(30): 22815-22823.

Cameroni, E., K. Stettler and B. Suter (2010). "On the traces of XPD: cell cycle matters - untangling the genotype-phenotype relationship of XPD mutations." Cell Div 5: 24.

Chang, W. H. and R. D. Kornberg (2000). "Electron crystal structure of the transcription factor and DNA repair complex, core TFIIH." Cell 102(5): 609-613.

Chayen, N. E. (1998). "Comparative studies of protein crystallization by vapour-diffusion and microbatch techniques." Acta Crystallogr D Biol Crystallogr 54(Pt 1): 8-15.

Chepanoske, C. L., M. P. Golinelli, S. D. Williams and S. S. David (2000). "Positively charged residues within the iron-sulfur cluster loop of E. coli MutY participate in damage recognition and removal." Arch Biochem Biophys 380(1): 11-19.

Cleaver, J. E. (1968). "Defective repair replication of DNA in xeroderma pigmentosum." Nature 218(5142): 652-656.

Cockayne, E. A. (1936). "Dwarfism with retinal atrophy and deafness." Arch Dis Child 11(61): 1-8.

Coin, F., J. C. Marinoni and J. M. Egly (1998). "Mutations in XPD helicase prevent its interaction and regulation by p44, another subunit of TFIIH, resulting in Xeroderma pigmentosum (XP) and trichothiodystrophy (TTD) phenotypes." Pathol Biol (Paris) 46(9): 679-680.

Coin, F., J. C. Marinoni, C. Rodolfo, S. Fribourg, A. M. Pedrini and J. M. Egly (1998). "Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH." Nat Genet 20(2): 184-188.

Coin, F., V. Oksenych and J. M. Egly (2007). "Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair." Mol Cell 26(2): 245-256.

Coin, F., V. Oksenych, V. Mocquet, S. Groh, C. Blattner and J. M. Egly (2008). "Nucleotide excision repair driven by the dissociation of CAK from TFIIH." Mol Cell 31(1): 9-20.

Cox, C. J., P. G. Foster, R. P. Hirt, S. R. Harris and T. M. Embley (2008). "The archaebacterial origin of eukaryotes." Proc Natl Acad Sci U S A 105(51): 20356-20361.

Cunningham, R. P., H. Asahara, J. F. Bank, C. P. Scholes, J. C. Salerno, K. Surerus, E. Munck, J. McCracken, J. Peisach and M. H. Emptage (1989). "Endonuclease III is an iron-sulfur protein." Biochemistry 28(10): 4450-4455.

Damsma, G. E., A. Alt, F. Brueckner, T. Carell and P. Cramer (2007). "Mechanism of transcriptional stalling at cisplatin-damaged DNA." Nat Struct Mol Biol 14(12): 1127-1133.

De La Fortelle, E. and G. Bricogne (1997). "Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods." Methods in enzymology: macromolecular crystallography part A: 472-494.

Ding, H., E. Hidalgo and B. Demple (1996). "The redox state of the [2Fe-2S] clusters in SoxR protein regulates its activity as a transcription factor." J Biol Chem 271(52): 33173-33175.

Drapkin, R., G. Le Roy, H. Cho, S. Akoulitchev and D. Reinberg (1996). "Human cyclin-dependent kinase-activating kinase exists in three distinct complexes." Proc Natl Acad Sci U S A 93(13): 6488-6493.

Dubaele, S., L. Proietti De Santis, R. J. Bienstock, A. Keriel, M. Stefanini, B. Van Houten and J. M. Egly (2003). "Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients." Mol Cell 11(6): 1635-1646.

Egly, J. M. and F. Coin (2011). "A history of TFIIH: Two decades of molecular biology on a pivotal transcription/repair factor." DNA Repair (Amst).

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallogr D Biol Crystallogr 60(Pt 12 Pt 1): 2126-2132.

Enzlin, J. H. and O. D. Scharer (2002). "The active site of the DNA repair endonuclease XPF-ERCC1 forms a highly conserved nuclease motif." EMBO J 21(8): 2045-2053.

Evans, P. (2006). "Scaling and assessment of data quality." Acta Crystallogr D Biol Crystallogr 62(Pt 1): 72-82.

Fan, L., J. O. Fuss, Q. J. Cheng, A. S. Arvai, M. Hammel, V. A. Roberts, P. K. Cooper and J. A. Tainer (2008). "XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations." Cell 133(5): 789-800.

Flint, D. H. and R. M. Allen (1996). "Ironminus signSulfur Proteins with Nonredox Functions." Chem Rev 96(7): 2315-2334.

Friedberg, E. C., A. Aguilera, M. Gellert, P. C. Hanawalt, J. B. Hays, A. R. Lehmann, T. Lindahl, N. Lowndes, A. Sarasin and R. D. Wood (2006). "DNA repair: from molecular mechanism to human disease." DNA Repair (Amst) 5(8): 986-996.

Fromme, J. C., A. Banerjee, S. J. Huang and G. L. Verdine (2004). "Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase." Nature 427(6975): 652-656.

Gal, A. and G. N. Wogan (1996). "Mutagenesis associated with nitric oxide production in transgenic SJL mice." Proceedings of the National Academy of Sciences of the United States of America 93(26): 15102-15107.

Gasior, S. L., A. M. Roy-Engel and P. L. Deininger (2008). "ERCC1/XPF limits L1 retrotransposition." DNA repair 7(6): 983-989.

Gaudu, P. and B. Weiss (1996). "SoxR, a [2Fe-2S] transcription factor, is active only in its oxidized form." Proceedings of the National Academy of Sciences of the United States of America 93(19): 10094-10098.

Giglia-Mari, G., F. Coin, J. A. Ranish, D. Hoogstraten, A. Theil, N. Wijgers, N. G. Jaspers, A. Raams, M. Argentini, P. J. van der Spek, E. Botta, M. Stefanini, J. M. Egly, R. Aebersold, J. H. Hoeijmakers and W. Vermeulen (2004). "A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A." Nat Genet 36(7): 714-719.

Gillespie, J. M. and R. C. Marshall (1983). "A comparison of the proteins of normal and trichothiodystrophic human hair." J Invest Dermatol 80(3): 195-202.

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

Grogan, D. W. (2000). "The question of DNA repair in hyperthermophilic archaea." Trends Microbiol 8(4): 180-185.

Grogan, D. W. (2004). "Stability and repair of DNA in hyperthermophilic Archaea." Curr Issues Mol Biol 6(2): 137-144.

Guan, Y., R. C. Manuel, A. S. Arvai, S. S. Parikh, C. D. Mol, J. H. Miller, S. Lloyd and J. A. Tainer (1998). "MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily." Nat Struct Biol 5(12): 1058-1064.

Guo, C., T. S. Tang and E. C. Friedberg (2010). "SnapShot: nucleotide excision repair." Cell 140(5): 754-754 e751.

Hanawalt, P. C. and G. Spivak (2008). "Transcription-coupled DNA repair: two decades of progress and surprises." Nat Rev Mol Cell Biol 9(12): 958-970.

Hearst, J. E., S. T. Isaacs, D. Kanne, H. Rapoport and K. Straub (1984). "The reaction of the psoralens with deoxyribonucleic acid." Q Rev Biophys 17(1): 1-44.

Hey, T., G. Lipps, K. Sugasawa, S. Iwai, F. Hanaoka and G. Krauss (2002). "The XPC-HR23B complex displays high affinity and specificity for damaged DNA in a true-equilibrium fluorescence assay." Biochemistry 41(21): 6583-6587.

Hinks, J. A., M. C. Evans, Y. De Miguel, A. A. Sartori, J. Jiricny and L. H. Pearl (2002). "An iron-sulfur cluster in the family 4 uracil-DNA glycosylases." J Biol Chem 277(19): 16936-16940.

Hirata, A., B. J. Klein and K. S. Murakami (2008). "The X-ray crystal structure of RNA polymerase from Archaea." Nature 451(7180): 851-854.

Hoeijmakers, J. H. (2009). "DNA damage, aging, and cancer." N Engl J Med 361(15): 1475-1485.

Honda, M., J. Park, R. A. Pugh, T. Ha and M. Spies (2009). "Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase." Mol Cell 35(5): 694-703.

Houtsmuller, A. B., S. Rademakers, A. L. Nigg, D. Hoogstraten, J. H. Hoeijmakers and W. Vermeulen (1999). "Action of DNA repair endonuclease ERCC1/XPF in living cells." Science 284(5416): 958-961.

Ignarro, L. J. (1999). "Nitric oxide: a unique endogenous signaling molecule in vascular biology." Bioscience reports 19(2): 51-71.

Imlay, J. A. (2006). "Iron-sulphur clusters and the problem with oxygen." Mol Microbiol 59(4): 1073-1082.

Itin, P. H. and M. R. Pittelkow (1990). "Trichothiodystrophy: review of sulfur-deficient brittle hair syndromes and association with the ectodermal dysplasias." J Am Acad Dermatol 22(5 Pt 1): 705-717.

Itin, P. H., A. Sarasin and M. R. Pittelkow (2001). "Trichothiodystrophy: update on the sulfur-deficient brittle hair syndromes." J Am Acad Dermatol 44(6): 891-920; quiz 921-894.

Ito, S., L. J. Tan, D. Andoh, T. Narita, M. Seki, Y. Hirano, K. Narita, I. Kuraoka, Y. Hiraoka and K. Tanaka (2010). "MMXD, a TFIIH-independent XPD-MMS19 protein complex involved in chromosome segregation." Mol Cell 39(4): 632-640.

Johnson, K. A., S. P. Fink and L. J. Marnett (1997). "Repair of propanodeoxyguanosine by nucleotide excision repair in vivo and in vitro." J Biol Chem 272(17): 11434-11438.

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

Jones, T. A., J. Y. Zou, S. W. Cowan and M. Kjeldgaard (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models." Acta crystallographica. Section A, Foundations of crystallography 47 (Pt 2): 110-119.

Kelly, S. M., T. J. Jess and N. C. Price (2005). "How to study proteins by circular dichroism." Biochim Biophys Acta 1751(2): 119-139.

Kelman, Z. and M. F. White (2005). "Archaeal DNA replication and repair." Curr Opin Microbiol 8(6): 669-676.

Kiley, P. J. and H. Beinert (2003). "The role of Fe-S proteins in sensing and regulation in bacteria." Curr Opin Microbiol 6(2): 181-185.

Kleywegt, G. J. and T. A. Jones (1997). "Model building and refinement practice." Methods Enzymol 277: 208-230.

Klinge, S., J. Hirst, J. D. Maman, T. Krude and L. Pellegrini (2007). "An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis." Nat Struct Mol Biol 14(9): 875-877.

Kraemer, K. H., M. M. Lee and J. Scotto (1987). "Xeroderma pigmentosum. Cutaneous, ocular, and neurologic abnormalities in 830 published cases." Arch Dermatol 123(2): 241-250.

Kuraoka, I., W. R. Kobertz, R. R. Ariza, M. Biggerstaff, J. M. Essigmann and R. D. Wood (2000). "Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease." J Biol Chem 275(34): 26632-26636.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature 227(5259): 680-685.

Laine, J. P. and J. M. Egly (2006). "When transcription and repair meet: a complex system." Trends Genet 22(8): 430-436.

Laine, J. P., V. Mocquet and J. M. Egly (2006). "TFIIH enzymatic activities in transcription and nucleotide excision repair." Methods Enzymol 408: 246-263.

Lambert, C., N. Leonard, X. De Bolle and E. Depiereux (2002). "ESyPred3D: Prediction of proteins 3D structures." Bioinformatics 18(9): 1250-1256.

Lauble, H., M. C. Kennedy, H. Beinert and C. D. Stout (1992). "Crystal structures of aconitase with isocitrate and nitroisocitrate bound." Biochemistry 31(10): 2735-2748.

Lehmann, A. R. (2001). "The xeroderma pigmentosum group D (XPD) gene: one gene, two functions, three diseases." Genes Dev 15(1): 15-23.

Li, C. Q., L. J. Trudel and G. N. Wogan (2002). "Nitric oxide-induced genotoxicity, mitochondrial damage, and apoptosis in human lymphoblastoid cells expressing wild-type and mutant p53." Proceedings of the National Academy of Sciences of the United States of America 99(16): 10364-10369.

Li, L., S. J. Elledge, C. A. Peterson, E. S. Bales and R. J. Legerski (1994). "Specific association between the human DNA repair proteins XPA and ERCC1." Proc Natl Acad Sci U S A 91(11): 5012-5016.

Li, L., C. A. Peterson, X. Lu and R. J. Legerski (1995). "Mutations in XPA that prevent association with ERCC1 are defective in nucleotide excision repair." Mol Cell Biol 15(4): 1993-1998.

Liu, H., J. Rudolf, K. A. Johnson, S. A. McMahon, M. Oke, L. Carter, A. M. McRobbie, S. E. Brown, J. H. Naismith and M. F. White (2008). "Structure of the DNA repair helicase XPD." Cell 133(5): 801-812.

Lukianova, O. A. and S. S. David (2005). "A role for iron-sulfur clusters in DNA repair." Curr Opin Chem Biol 9(2): 145-151.

Martinez-Senac, M. M. and M. R. Webb (2005). "Mechanism of translocation and kinetics of DNA unwinding by the helicase RecG." Biochemistry 44(51): 16967-16976.

Masutani, C., R. Kusumoto, A. Yamada, N. Dohmae, M. Yokoi, M. Yuasa, M. Araki, S. Iwai, K. Takio and F. Hanaoka (1999). "The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase eta." Nature 399(6737): 700-704.

Mathieu, N., N. Kaczmarek and H. Naegeli (2010). "Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase." Proc Natl Acad Sci U S A 107(41): 17545-17550.

Matsuda, T., M. Saijo, I. Kuraoka, T. Kobayashi, Y. Nakatsu, A. Nagai, T. Enjoji, C. Masutani, K. Sugasawa, F. Hanaoka and et al. (1995). "DNA repair protein XPA binds replication protein A (RPA)." J Biol Chem 270(8): 4152-4157.

Matsunaga, T., D. Mu, C. H. Park, J. T. Reardon and A. Sancar (1995). "Human DNA repair excision nuclease. Analysis of the roles of the subunits involved in dual incisions by using anti-XPG and anti-ERCC1 antibodies." J Biol Chem 270(35): 20862-20869.

Matulis, D., J. K. Kranz, F. R. Salemme and M. J. Todd (2005). "Thermodynamic stability of carbonic anhydrase: measurements of binding affinity and stoichiometry using ThermoFluor." Biochemistry 44(13): 5258-5266.

McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read (2007). "Phaser crystallographic software." J Appl Crystallogr 40(Pt 4): 658-674.

Min, J. H. and N. P. Pavletich (2007). "Recognition of DNA damage by the Rad4 nucleotide excision repair protein." Nature 449(7162): 570-575.

Moggs, J. G., K. J. Yarema, J. M. Essigmann and R. D. Wood (1996). "Analysis of incision sites produced by human cell extracts and purified proteins during nucleotide excision repair of a 1,3-intrastrand d(GpTpG)-cisplatin adduct." J Biol Chem 271(12): 7177-7186.

Morris, R. J., A. Perrakis and V. S. Lamzin (2003). "ARP/wARP and automatic interpretation of protein electron density maps." Methods in enzymology 374: 229-244.

Moser, J., H. Kool, I. Giakzidis, K. Caldecott, L. H. Mullenders and M. I. Fousteri (2007). "Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner." Mol Cell 27(2): 311-323.

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

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

Murshudov, G. N., A. A. Vagin and E. J. Dodson (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallogr D Biol Crystallogr 53(Pt 3): 240-255.

Naegeli, H., L. Bardwell and E. C. Friedberg (1993). "Inhibition of Rad3 DNA helicase activity by DNA adducts and abasic sites: implications for the role of a DNA helicase in damage-specific incision of DNA." Biochemistry 32(2): 613-621.

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

Nichols, A. F. and A. Sancar (1992). "Purification of PCNA as a nucleotide excision repair protein." Nucleic Acids Res 20(13): 2441-2446.

Nocentini, S. (1999). "Rejoining kinetics of DNA single- and double-strand breaks in normal and DNA ligase-deficient cells after exposure to ultraviolet C and gamma radiation: an evaluation of ligating activities involved in different DNA repair processes." Radiat Res 151(4): 423-432.

Nouspikel, T. (2009). "DNA repair in mammalian cells : Nucleotide excision repair: variations on versatility." Cell Mol Life Sci 66(6): 994-1009.

O'Donovan, A., A. A. Davies, J. G. Moggs, S. C. West and R. D. Wood (1994). "XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair." Nature 371(6496): 432-435.

Ogrunc, M., D. F. Becker, S. W. Ragsdale and A. Sancar (1998). "Nucleotide excision repair in the third kingdom." J Bacteriol 180(21): 5796-5798.

Park, C. H. and A. Sancar (1994). "Formation of a ternary complex by human XPA, ERCC1, and ERCC4(XPF) excision repair proteins." Proc Natl Acad Sci U S A 91(11): 5017-5021.

Payne, A. and G. Chu (1994). "Xeroderma pigmentosum group E binding factor recognizes a broad spectrum of DNA damage." Mutat Res 310(1): 89-102.

Pfeifer, G. P. (1997). "Formation and processing of UV photoproducts: effects of DNA sequence and chromatin environment." Photochem Photobiol 65(2): 270-283.

Pfeiffer, H. and H. Liebhafsky (1951). "The origins of Beer's law." Journal of Chemical Education 28: 123-125.

Porello, S. L., M. J. Cannon and S. S. David (1998). "A substrate recognition role for the [4Fe-4S]2+ cluster of the DNA repair glycosylase MutY." Biochemistry 37(18): 6465-6475.

Pugh, R. A., M. Honda, H. Leesley, A. Thomas, Y. Lin, M. J. Nilges, I. K. Cann and M. Spies (2008). "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." J Biol Chem 283(3): 1732-1743.

Pugh, R. A., Y. Lin, C. Eller, H. Leesley, I. K. Cann and M. Spies (2008). "Ferroplasma acidarmanus RPA2 facilitates efficient unwinding of forked DNA substrates by monomers of FacXPD helicase." J Mol Biol 383(5): 982-998.

Rapin, I., Y. Lindenbaum, D. W. Dickson, K. H. Kraemer and J. H. Robbins (2000). "Cockayne syndrome and xeroderma pigmentosum." Neurology 55(10): 1442-1449.

Reardon, J. T., H. Ge, E. Gibbs, A. Sancar, J. Hurwitz and Z. Q. Pan (1996). "Isolation and characterization of two human transcription factor IIH (TFIIH)-related complexes: ERCC2/CAK and TFIIH." Proc Natl Acad Sci U S A 93(13): 6482-6487.

Reardon, J. T. and A. Sancar (2003). "Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease." Genes Dev 17(20): 2539-2551.

Ren, B., X. Duan and H. Ding (2009). "Redox control of the DNA damage-inducible protein DinG helicase activity via its iron-sulfur cluster." J Biol Chem 284(8): 4829-4835.

Rouillon, C. and M. F. White (2011). "The evolution and mechanisms of nucleotide excision repair proteins." Res Microbiol 162(1): 19-26.

Rudolf, J., V. Makrantoni, W. J. Ingledew, M. J. Stark and M. F. White (2006). "The DNA repair helicases XPD and FancJ have essential iron-sulfur domains." Mol Cell 23(6): 801-808.

Rudolf, J., C. Rouillon, U. Schwarz-Linek and M. F. White (2010). "The helicase XPD unwinds bubble structures and is not stalled by DNA lesions removed by the nucleotide excision repair pathway." Nucleic Acids Res 38(3): 931-941.

Ruepp, A., W. Graml, M. L. Santos-Martinez, K. K. Koretke, C. Volker, H. W. Mewes, D. Frishman, S. Stocker, A. N. Lupas and W. Baumeister (2000). "The genome sequence of the thermoacidophilic scavenger Thermoplasma acidophilum." Nature 407(6803): 508-513.

Saijo, M., I. Kuraoka, C. Masutani, F. Hanaoka and K. Tanaka (1996). "Sequential binding of DNA repair proteins RPA and ERCC1 to XPA in vitro." Nucleic Acids Res 24(23): 4719-4724.

Sancar, A. (1994). "Mechanisms of DNA excision repair." Science 266(5193): 1954-1956.

Sancar, G. B., W. Siede and A. A. van Zeeland (1996). "Repair and processing of DNA damage: a summary of recent progress." Mutat Res 362(1): 127-146.

Sarasin, A. (2003). "An overview of the mechanisms of mutagenesis and carcinogenesis." Mutat Res 544(2-3): 99-106.

Schneider, T. R. and G. M. Sheldrick (2002). "Substructure solution with SHELXD." Acta crystallographica. Section D, Biological crystallography 58(Pt 10 Pt 2): 1772-1779.

Schultz, P., S. Fribourg, A. Poterszman, V. Mallouh, D. Moras and J. M. Egly (2000). "Molecular structure of human TFIIH." Cell 102(5): 599-607.

Schwartz, C. J., J. L. Giel, T. Patschkowski, C. Luther, F. J. Ruzicka, H. Beinert and P. J. Kiley (2001). "IscR, an Fe-S cluster-containing transcription factor, represses expression of Escherichia coli genes encoding Fe-S cluster assembly proteins." Proc Natl Acad Sci U S A 98(26): 14895-14900.

Setlow, R. B., J. D. Regan, J. German and W. L. Carrier (1969). "Evidence that xeroderma pigmentosum cells do not perform the first step in the repair of ultraviolet damage to their DNA." Proc Natl Acad Sci U S A 64(3): 1035-1041.

Shivji, K. K., M. K. Kenny and R. D. Wood (1992). "Proliferating cell nuclear antigen is required for DNA excision repair." Cell 69(2): 367-374.

Singleton, M. R., M. S. Dillingham and D. B. Wigley (2007). "Structure and mechanism of helicases and nucleic acid translocases." Annu Rev Biochem 76: 23-50.

Spies, M. and T. Ha (2010). "Inching over hurdles: how DNA helicases move on crowded lattices." Cell Cycle 9(9): 1742-1749.

Stefanini, M., P. Lagomarsini, C. F. Arlett, S. Marinoni, C. Borrone, F. Crovato, G. Trevisan, G. Cordone and F. Nuzzo (1986). "Xeroderma pigmentosum (complementation group D) mutation is present in patients affected by trichothiodystrophy with photosensitivity." Hum Genet 74(2): 107-112.

Sugasawa, K. (2001). "[Molecular mechanism of mammalian nucleotide excision repair]." Tanpakushitsu Kakusan Koso 46(8 Suppl): 893-901.

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

Sugasawa, K., Y. Okuda, M. Saijo, R. Nishi, N. Matsuda, G. Chu, T. Mori, S. Iwai, K. Tanaka and F. Hanaoka (2005). "UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex." Cell 121(3): 387-400.

Svejstrup, J. Q. (2002). "Mechanisms of transcription-coupled DNA repair." Nat Rev Mol Cell Biol 3(1): 21-29.

Svoboda, D. L., J. S. Taylor, J. E. Hearst and A. Sancar (1993). "DNA repair by eukaryotic nucleotide excision nuclease. Removal of thymine dimer and psoralen monoadduct by HeLa cell-free extract and of thymine dimer by Xenopus laevis oocytes." J Biol Chem 268(3): 1931-1936.

Tang, J. and G. Chu (2002). "Xeroderma pigmentosum complementation group E and UV-damaged DNA-binding protein." DNA Repair (Amst) 1(8): 601-616.

Tirode, F., D. Busso, F. Coin and J. M. Egly (1999). "Reconstitution of the transcription factor TFIIH: assignment of functions for the three enzymatic subunits, XPB, XPD, and cdk7." Mol Cell 3(1): 87-95.

Trimmer, E. E. and J. M. Essigmann (1999). "Cisplatin." Essays Biochem 34: 191-211.

Truglio, J. J., D. L. Croteau, B. Van Houten and C. Kisker (2006). "Prokaryotic nucleotide excision repair: the UvrABC system." Chem Rev 106(2): 233-252.

Vaithiyalingam, S., E. M. Warren, B. F. Eichman and W. J. Chazin (2010). "Insights into eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase." Proc Natl Acad Sci U S A 107(31): 13684-13689.

van Gool, A. J., G. T. van der Horst, E. Citterio and J. H. Hoeijmakers (1997). "Cockayne syndrome: defective repair of transcription?" EMBO J 16(14): 4155-4162.

van Hoffen, A., A. T. Natarajan, L. V. Mayne, A. A. van Zeeland, L. H. Mullenders and J. Venema (1993). "Deficient repair of the transcribed strand of active genes in Cockayne's syndrome cells." Nucleic Acids Res 21(25): 5890-5895.

Vermeulen, W., R. J. Scott, S. Rodgers, H. J. Muller, J. Cole, C. F. Arlett, W. J. Kleijer, D. Bootsma, J. H. Hoeijmakers and G. Weeda (1994). "Clinical heterogeneity within xeroderma pigmentosum associated with mutations in the DNA repair and transcription gene ERCC3." Am J Hum Genet 54(2): 191-200.

Volker, M., M. J. Mone, P. Karmakar, A. van Hoffen, W. Schul, W. Vermeulen, J. H. Hoeijmakers, R. van Driel, A. A. van Zeeland and L. H. Mullenders (2001). "Sequential assembly of the nucleotide excision repair factors in vivo." Mol Cell 8(1): 213-224.

Von Hebra, F. K., M (1874). On Diseases of the Skin. Including the Exanthemata. 3.

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

Wang, H., M. J. DellaVecchia, M. Skorvaga, D. L. Croteau, D. A. Erie and B. Van Houten (2006). "UvrB domain 4, an autoinhibitory gate for regulation of DNA binding and ATPase activity." J Biol Chem 281(22): 15227-15237.

Wang, W. and B. A. Malcolm (1999). "Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis." Biotechniques 26(4): 680-682.

Weeda, G., E. Eveno, I. Donker, W. Vermeulen, O. Chevallier-Lagente, A. Taieb, A. Stary, J. H. Hoeijmakers, M. Mezzina and A. Sarasin (1997). "A mutation in the XPB/ERCC3 DNA repair transcription gene, associated with trichothiodystrophy." Am J Hum Genet 60(2): 320-329.

Weiner, B. E., H. Huang, B. M. Dattilo, M. J. Nilges, E. Fanning and W. J. Chazin (2007). "An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase." J Biol Chem 282(46): 33444-33451.

Weller, M.-C. (2010). "Functional Characterization of the Nucleotide Excision Repair Pathway."

White, M. F. (2009). "Structure, function and evolution of the XPD family of iron-sulfur-containing 5'-->3' DNA helicases." Biochem Soc Trans 37(Pt 3): 547-551.

Winkler, G. S., S. J. Araujo, U. Fiedler, W. Vermeulen, F. Coin, J. M. Egly, J. H. Hoeijmakers, R. D. Wood, H. T. Timmers and G. Weeda (2000). "TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair." J Biol Chem 275(6): 4258-4266.

Wittschieben, B. O., S. Iwai and R. D. Wood (2005). "DDB1-DDB2 (xeroderma pigmentosum group E) protein complex recognizes a cyclobutane pyrimidine dimer, mismatches, apurinic/apyrimidinic sites, and compound lesions in DNA." J Biol Chem 280(48): 39982-39989.

Woese, C. R. and G. E. Fox (1977). "Phylogenetic structure of the prokaryotic domain: the primary kingdoms." Proc Natl Acad Sci U S A 74(11): 5088-5090.

Wogan, G. N., S. S. Hecht, J. S. Felton, A. H. Conney and L. A. Loeb (2004). "Environmental and chemical carcinogenesis." Semin Cancer Biol 14(6): 473-486.

Wolski, S. C. (2007). "Strukturelle Charakterisierung des XPD-Proteins aus *Thermoplasma acidophilum* und der UvrB-Mutante Y95W aus *Bacillus caldotenax*.".

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

7 Apendix

7.1 Abreviations

(6-4)PP	(6-4) pyrimidine-pyrimidone photoproduct
AA	Acrylamide
Amp	Ampicillin
APS	Ammonium peroxydisulfate
BAA	Bisacrylamide
BER	Base excsion repair
BSA	Bovine serum albumin
CAK	Cdk-activiating kinase
Cam	Chloramphenicol
CD	Circular dichroism
Chl1	Chromatid cohesion in yeast
CPD	Cyclobutane pyrimidine dimers
CS	Cockayne Syndrome
CSA	Cockayne Syndrome A
CSB	Cockayne Syndrome B
CV	Column volume
Da	Dalton
DDB	Damaged DNA-binding protein
DTT	dithiothreitol
DinG	Damage-inducible G
DNA	Deoxyribonucleic acid
DNase	desoxyribonuclease
dsDNA	Double stranded DNA
E.coli	Escherichia coli
EDTA	ethylenediaminetetraacetate
EPR	Electron paramagnetic resonance
F	Fluorescein
FancJ	Fanconi's anemia complementation group J
faXPD	*Ferroplasma acidarmanus* XPD
FeS	Iron-sulfur
GGR	Global genome repair
h	hour
IPTG	Isopropyl-β-thiogalactoside
LB	Luria broth
M	molar
MAD	Multiwavelengt anomalous diffraction
min	minutes
mL	milliliter
MMXD	MMS19-MIP18-XPD
MPD	2-Methyl-2,4-pentanediol
MR	Molecular Replacement
MWCO	Molecular weight cut off
NER	Nucleotide excision repair
nm	nanometer
nM	nanomolar
OD	Optical density
PAGE	Polyacrylamide gel electophoresis
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PEG	Polyethylenglycol
RT	Room temperature
RTEL	Regular of telomere length

saXPD	*Sulfolobus acidocaldarius* XPD
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SF2	Superfamily 2
ssDNA	Single stranded DNA
taXPD	*Thermoplasma acidophilum* XPD
TCR	Transcription coupled repair
TEMED	tetramethylethylenediamine
TFIIH	Transcriptionfactor II H
TTD	Trichothiodystrophy
µL	microliter
µM	micromolar
UV	ultraviolet
WT	Wild type
XP	Xeroderma Pigmentosum
XP(B-G)	Xeroderma pigmentosum complementation group (B-G)

Aminoacids

Alanine	A
Arginine	R
Asparagine	N
Asparatic acid	D
Cysteine	C
Glutamic acid	E
Glutamine	Q
Glycine	G
Histidine	H
Isoleucine	I
Leucine	L
Lysine	K
Methionine	M
Phenylalanine	F
Proline	P
Serine	S
Threonine	T
Tryptophan	W
Tyrosine	Y
Valine	V

7.2 Table of table and figures

7.3 Superposition of XPD with UvrB, NS3 and Hel308

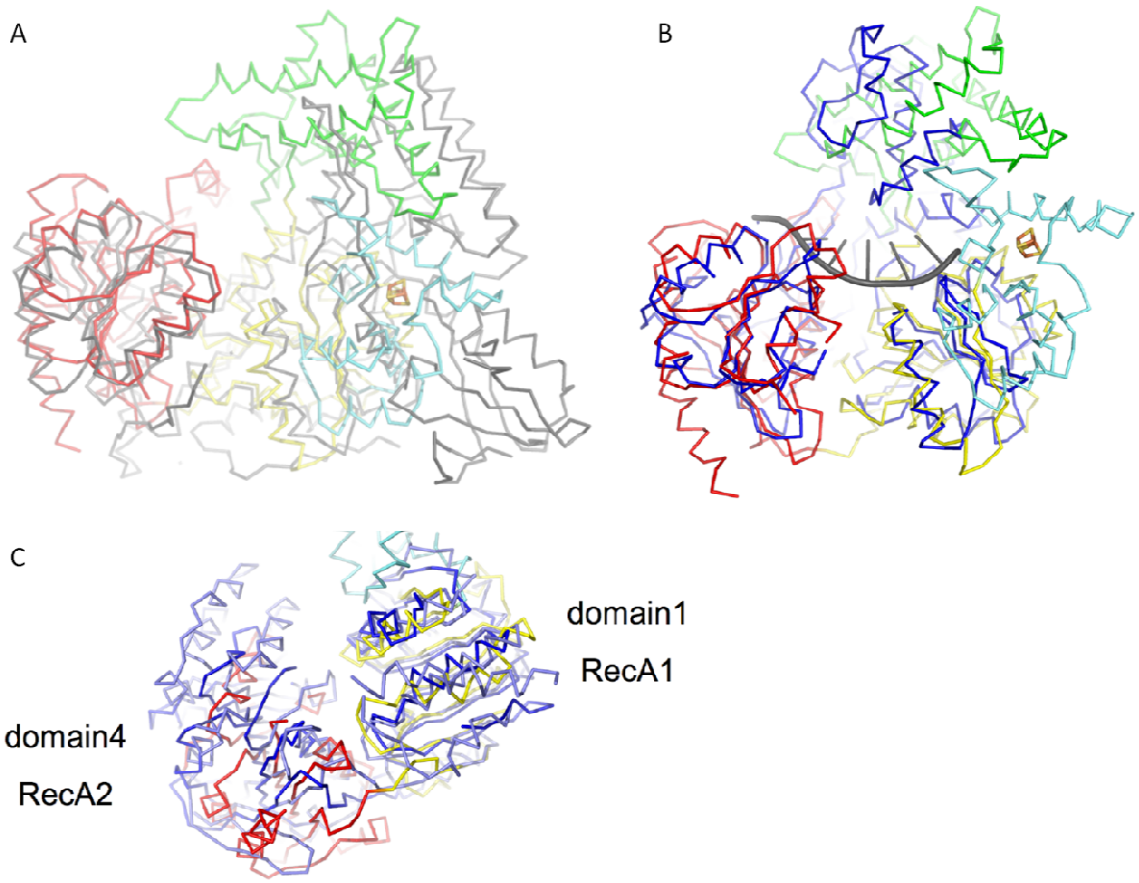


Figure 7.1: Superposition of XPD with UvrB, NS3 and Hel308. XPD iscolor-coded as in figure X, **(A)** XPD was superimposed with UvrB shown in grey. The two RecA-like domains (yellow and red) superimpose, whereas the remainder of the two protein models share no structural homology. **(B)** superposition of XPD and NS3 with NS3 colored in blue which lead to an rmsd of 3 Å. **(C)** superposition of the motor domains of XPD (yellow and red), NS3 (blue) and Hel308 (light blue) using the first RecA domain (domain 1) as the pivot point. (Wolski et al. 2008)

7.4 Sequences of constructs

XPD from *Ferroplasma acidarmanus*(pET28a derived from N. Mathieu)

MEKFTPREWQDALISTVSKNLEEKNKIAIEAPTGSGKTSFILYLAFITGKKLIYLTRTHNEFT
RLYEDNQKYFNIPALYLYGKSKLCPLKERWYDSEDDGQKNVCKGCPLKDKTIKLDLKNIR
SPESFLEDIEADSTELIRQDVESKALGRDTKGNLVTIKKENVKSYCPYYSVRSNMADAQII
AMTYNYLLNPSIRFNVFHGTEGEEMIDLNECYIVFDEAHNLDSVIENFGRTLQIKTVEKAID
MLVKDFPIETYNSYDRIEANLMLENLLSNMSSTRGTDSLKLFKFSNDFASRIANLESLRTF
SDEFEEKNENRPLNKRSRNYLENVYNFLYDYQRMENGSVYSTINNPEKKDLKLKLMYYD
TSGYLSFLRDNSVIFMSGTMPSVDHISKVWNMENVLYLKVDSIFKNAGGLKQYHIVNGYT
TLGRYRENLDDWAEMLRKYMQFAIDVYSKSDKSVLVAVPSYRILFGDTYTKTPGIGNYLP
ESMRANCVFENKKISYSYIEKKAKERKIIIFSVHGGKLLEGIQLVNRGKSLISDVIIAGLPLIPL
DDYRKDKIKYLEKVLKKNAYNLLYYEFALIKVKQAAGRSTRSPEDTSSIWLCDDRFDEPF
WQSNLLPNTK

ATGGAAAAATTTACCCCCAGGGAATGGCAGGATGCCCTGATAAGCACCGTATCAAAA
AACCTTGAGGAGAAAAACAAAATTGCCATAGAGGCACCTACCGGTTCCGGAAAGACC
AGCTTCATACTATATCTGGCTTTTATTACAGGGAAGAAATTAATTTACCTTACCAGAAC
ACATAATGAATTCACGAGGCTTTATGAGGATAACCAGAAATATTTCAACATACCCGCT
TTATATTTATATGGAAAAAGCAAATTATGCCCGCTTAAAGAGAGGTGGTACGACTCAG
AAGATGATGGGCAGAAAAATGTATGCAAAGGGTGCCCCTTGAAGGATAAAACCATAA
AGCTTGATCTGAAAAATATAAGGAGCCCGGAGAGTTTTCTGGAAGACATAGAAGCTG
ATTCTACGGAACTTATCAGGCAGGATGTGGAATCAAAGGCCCTCGGAAGGGATACAA
AGGGGAATCTTGTTACAATCAAGAAAGAAAATGTAAAATCCTATTGCCCATATTATTCT
GTGCGGTCCAACATGGCAGATGCACAGATAATAGCCATGACATATAACTATCTGTTAA
ATCCATCAATACGGTTCAATGTTTTTCATGGCACCGAAGGGGAAGAGATGATTGACC
TTAATGAATGCTATATAGTTTTTGATGAAGCACATAACCTGGATTCTGTAATAGAAAAT
TTTGGCCGGACGCTGCAGATAAAAACAGTAGAGAAGGCAATTGATATGCTTGTAAAG
GATTTTCCAATTGAAACTTATAATTCATATGATAGGATTGAAGCAAATTTGATGCTTGA
AAATCTTCTGTCCAACATGAGCTCAACAAGGGGAACCGATAGTCTCAAGCTATTCAAA
TTTTCTAATGATTTTGCATCAAGGATTGCAAACCTTGAATCACTCAGGACATTCAGTG
ATGAATTTGAAGAAAAAAACGAAAACAGGCCATTGAATAAGAGAAGCAGGAATTACCT
GGAAAACGTTTACAATTTTCTCTATGATTACCAGAGGATGGAAAATGGTTCAGTGTAC
TCCACTATCAATAATCCGGAAAAAAAAGATTTAAAACTTAAATTGATGTATTATGATAC
ATCCGGCTATCTTTCATTTTTACGGGATAACTCTGTTATTTTTATGTCCGGGACAATGC
CATCTGTAGACCATATCTCAAAGGTATGGAATATGGAAAATGTCCTGTACCTGAAAGT
TGACAGCATATTCAAAAACGCCGGCGGGCTAAAGCAATACCATATAGTTAACGGTTA
TACAACTCTGGGACGTTACCGCGAAAACCTGGATGATTGGGCTGAAATGCTCAGAAA
ATATATGCAGTTTGCCATTGATGTTTACTCAAAATCGGATAAAAGTGTACTTGTTGCTG
TCCCCTCATATCGCATACTTTTTGGCGACACCTATACAAAAACACCAGGCATAGGCAA
TTACCTGCCAGAAAGTATGAGGGCAAACTGTGTATTCGAAAACAAGAAAATTTCTTAT
TCCTATATAGAAAAGAAAGCGAAGGAGAGAAAAATAATTATTTTTTCTGTACACGGTG
GCAAACTTCTTGAGGGCATACAGCTGGTTAACCGGGGAAAAAGCCTGATTAGCGATG
TAATCATTGCAGGGCTTCCCCTGATACCCCTTGATGATTACAGAAAGGATAAAATAAA
ATATCTTGAAAAAGTGCTGAAAAAAAATGCATACAATCTGCTCTATTATGAGTTTGCG

CTTATCAAAGTAAAACAGGCAGCAGGCAGATCTACAAGGTCTCCAGAGGATACTTCC
AGCATATGGCTCTGTGATGACAGGTTTGATGAGCCTTTCTGGCAATCCAATTTGCTTC
CTAACACAAAATAACTCGAG

Restrictionenzymes used for cloning NdeI and Xho

XPD *Thermoplasma acidophilum*(pET16b, T4, residues 23-602)

Amino acid sequence:
MQKSYGVALESPTGSGKTIMALKSALQYSSERKLKVLYLVRTNSQEEQVIKELRSLSSTM
KIRAIPMQGRVNMCILYRMVDDLHEINAESLAKFCNMKKREVMAGNEAACPYFNFKIRSD
ETKRFLFDELTAEEFYDYGERNNVCPYESMKAALPDADIVIAPYAYFLNRSVAEKFLSHW
GVSRNQIVIILDEAHNLPDIGRSIGSFRISVESLNRADREAQAYGDPELSQKIHVSDLIEMIR
SALQSMVSERCGKGDVRIRFQEFMEYMRIMNKRSEREIRSLLNYLYLFGEYVENEKEKV
GKVPFSYCSSVASRIIAFSDQDEEKYAAILSPEDGGYMQAACLDPSGILEVLKESKTIHMS
GTLDPFDFYSDITGFEIPFKKIGEIFPPENRYIAYYDGVSSKYDTLDEKELDRMATVIEDIILK
VKKNTIVYFPSYSLMDRVENRVSFEHMKEYRGIDQKELYSMLKKFRRDHGTIFAVSGGRL
SEGINFPGNELEMIILAGLPFPRPDAINRSLFDYYERKYGKGWEYSVVYPTAIKIRQEIGRLI
RSAEDTGACVILDKRAGQFRKFIPDMKKTSDPASDIYNFFISAQAREKYGA

XPD *Thermoplasma acidophilum* (pET16b, T3; residues 1-602)

Amino acid sequence:
MYENRQYQVEAIDFLRSSLQKSYGVALESPTGSGKTIMALKSALQYSSERKLKVLYLVRT
NSQEEQVIKELRSLSSTMKIRAIPMQGRVNMCILYRMVDDLHEINAESLAKFCNMKKREV
MAGNEAACPYFNFKIRSDETKRFLFDELPTAEEFYDYGERNNVCPYESMKAALPDADIVI
APYAYFLNRSVAEKFLSHWGVSRNQIVIILDEAHNLPDIGRSIGSFRISVESLNRADREAQ
AYGDPELSQKIHVSDLIEMIRSALQSMVSERCGKGDVRIRFQEFMEYMRIMNKRSEREIR
SLLNYLYLFGEYVENEKEKVGKVPFSYCSSVASRIIAFSDQDEEKYAAILSPEDGGYMQA
ACLDPSGILEVLKESKTIHMSGTLDPFDFYSDITGFEIPFKKIGEIFPPENRYIAYYDGVSSK
YDTLDEKELDRMATVIEDIILKVKKNTIVYFPSYSLMDRVENRVSFEHMKEYRGIDQKELY
SMLKKFRRDHGTIFAVSGGRLSEGINFPGNELEMIILAGLPFPRPDAINRSLFDYYERKYG
KGWEYSVVYPTAIKIRQEIGRLIRSAEDTGACVILDKRAGQFRKFIPDMKKTSDPASDIYNF
FISAQAREKYGA
atgtacgagaacaggcagtaccaggtggaggccatcgattttctcaggagttctttgcagaaaagctacggcgtggcacttga
atctccaaccggttccggcaagaccatcatggcattgaaatctgccctgcagtattcaagtgaaaggaagctcaaggttctgt
accttgtccgcaccaattcgcaggaggaacaggttataaaggaattgagatctctttcatccacgatgaagatccgcgcaata
ccaatgcagggcagggtcaatatgtgcatactctacaggatggttgacgatctgcacgagataaatgcagaatctcttgcaaa
attctgcaacatgaagaagcgtgaggtcatggccggaaatgaggctgcatgcccgtacttcaacttcaagataagatccgat
gaaacgaagaggtttcttttcgacgagctgccgactgcggaggaattttacgattatggggaaaggaacaacgtgtgcccttа
tgaaagcatgaaggcagcactgccagatgcagacattgttatagcaccttatgcatattttctcaacaggtccgtggcagaga
agtttctcagccactggggcgtctcaagaaaccagatagtgataatactggatgaagcgcacaacctgccggatataggca
gatccataggatccttcaggatatctgtggaatcgctgaacagagcagacagggaggctcaggcatacggtgatccagagc
tgtcgcagaagatccatgtttccgatctgatagagatgatcagaagtgctttgcagagcatggtcagcgagagatgcgggaa
gggcgacgtcaggataaggttccaggaattcatggaatatatgaggataatgaacaagaggagcgagagagaaataaga
tcgctgctgaactatctctatctcttcggcgaatacgtggaaaacgagaaggaaaaggtgggcaaggtacctttcagttattgct
cgtccgttgccagcaggatcatcgcattctctgaccaggatgaggagaagtacgcggccatactttcgccggaagacggcg
gatatatgcaggcggcctgccttgatccgtctggaatactggaggtgctgaaggaatccaagacgatacacatgtccggaac
gcttgatcctttcgatttctattccgacatcaccggtttcgagattcctttcaagaagataggcgaaatatttcctcctgagaacag

atacattgcgtattacgatggggtgtcatcaaaatacgacacgctggatgaaaaggaactggacagaatggcgacggtgat
agaggatataatactgaaggtgaagaagaacaccatagtgtactttccatcatattcgctcatggacagggttgagaacagg
gtatcattcgaacacatgaaggagtacaggggcatagaccagaaggaactgtactccatgcttaagaaattcaggagggat
catggtacaattttcgcagtatctggcggaaggctttctgaaggcataaattttccgggaaacgaactggagatgataatactc
gcgggcctgcctttcccgaggccggacgccatcaacagatcgctgtttgactactatgagaggaaatacggcaagggctgg
gaatacagcgtcgtgtatccaacggccataaagataaggcaggagatagggaggctcataaggagtgcggaagatacag
gcgcctgcgtgatcctggacaagagggccggccagttcaggaaattcatacctgatatgaagaagacgtcggatccggcat
cggatatatacaatttttcatatctgcgcaggcacgcgaaaaatatggggcctga

XPD *Thermoplasma acidophilum*(pBadM11, residues 1-602; derived from Florian Rohleder)

MNENRQYQVEAIDFLRSSLQKSYGVALESPTGSGKTIMALKSALQYSSERKLKVLYLVRT
NSQEEQVIKELRSLSSTMKIRAIPMQGRVNMCILYRMVDDLHEINAESLAKFCNMKKREV
MAGNEAACPYFNFKIRSDETKRFLFDELPTAEEFYDYGERNNVCPYESMKAALPDADIVI
APYAYFLNRSVAEKFLSHWGVSRNQIVIILDEAHNLPDIGRSIGSFRISVESLNRADREAQ
AYGDPELSQKIHVSDLIEMIRSALQSMVSERCGKGDVRIRFQEFMEYMRIMNKRSEREIR
SLLNYLYLFGEYVENEKEKVGKVPFSYCSSVASRIIAFSDQDEEKYAAILSPEDGGYMQA
ACLDPSGILEVLKESKTIHMSGTLDPFDFYSDITGFEIPFKKIGEIFPPENRYIAYYDGVSSK
YDTLDEKELDRMATVIEDIILKVKKNTIVYFPSYSLMDRVENRVSFEHMKEYRGIDQKELY
SMLKKFRRDHGTIFAVSGGRLSEGINFPGNELEMIILAGLPFPRPDAINRSLFDYYERKYG
KGWEYSVVYPTAIKIRQEIGRLIRSAEDTGACVILDKRAGQFRKFIPDMKKTSDPASDIYNF
FISAQAREKYGA

ctgttggacagaaaaaaaggtttccatccgcttctaagtggaactatgaggattgctaatattcattaaccgggtcagcatctttc
acagaggatggacgagaacaggcagtaccaggtggaggccatcgattttctcaggagttctttgcagaaaagctacggcgt
ggcacttgaatctccaaccggttccggcaagaccatcatggcattgaaatctgccctgcagtattcaagtgaaaggaagctca
aggttctgtaccttgtccgcaccaattcgcaggaggaacaggttataaaggaattgagatctctttcatccacgatgaagatcc
gcgcaataccaatgcagggcagggtcaatatgtgcatactctacaggatggttgacgatctgcacgagataaatgcagaatc
tcttgcaaaattctgcaacatgaagaagcgtgaggtcatggccggaaatgaggctgcatgcccgtacttcaacttcaagataa
gatccgatgaaacgaagaggtttcttttcgacgagctgccgactgcggaggaattttacgattatggggaaaggaacaacgt
gtgcccttatgaaagcatgaaggcagcactgccagatgcagacattgttatagcaccttatgcatattttctcaacaggtccgtg
gcagagaagtttctcagccactggggcgtctcaagaaaccagatagtgataatactggatgaagcgcacaacctgccggat
ataggcagatccataggatccttcaggatatctgtggaatcgctgaacagagcagacagggaggctcaggcatacggtgat
ccagagctgtcgcagaagatccatgtttccgatctgatagagatgatcagaagtgctttgcagagcatggtcagcgagagat
gcgggaagggcgacgtcaggataaggttccaggaattcatggaatatatgaggataatgaacaagaggagcgagagag
aaataagatcgctgctgaactatctctatctcttcggcgaatacgtggaaaacgagaaggaaaaggtgggcaaggtacctttt
cagttattgctcgtccgttgccagcaggatcatcgcattctctgaccaggatgaggagaagtacgcggccatactttcgccgga
agacggcggatatatgcaggcggcctgccttgatccgtctggaatactggaggtgctgaaggaatccaagacgatacacat
gtccggaacgcttgatcctttcgatttctattccgacatcaccggtttcgagattcctttcaagaagataggcgaaatatttcctcct
gagaacagatacattgcgtattacgatggggtgtcatcaaaatacgacacgctggatgaaaaggaactggacagaatggc
gacggtgatagaggatataatactgaaggtgaagaagaacaccatagtgtactttccatcatattcgctcatggacagggttg
agaacagggtatcattcgaacacatgaaggagtacaggggcatagaccagaaggaactgtactccatgcttaagaaattc
aggagggatcatggtacaattttcgcagtatctggcggaaggctttctgaaggcataaattttccgggaaacgaactggagat
gataatactcgcgggcctgcctttcccgaggccggacgccatcaacagatcgctgtttgactactatgagaggaaatacggc
aagggctgggaatacagcgtcgtgtatccaacggccataaagataaggcaggagatagggaggctcataaggagtgcgg
aagatacaggcgcctgcgtgatcctggacaagagggccggccagttcaggaaattcatacctgatatgaagaagacgtcg
gatccggcatcggatatatacaatttttcatatctgcgcaggcacgcgaaaaatatggggcctgaaccctgcgtggaaaacg
ttttatgcctgttttgatt

Restrictionenyzmes used for cloning NcoI and HindIII

7.5 Primer for mutagenesis

R567W_fwd: ggc agg aga tag ggt ggc tca taa gga gtg c
R567W_rev: gca ctc ctt atg agc cac cct atc tcc tgc c

XPD_R118A_fwd; gca aca tga aga agg ctg agg tca tgg

XPD_R118A_rev; cca tga cct cag cct tct tca tgt tgc

XPD_R584A_fwd;gat ctt gga caa ggc ggc cgg cca gtt c

XPD_R584A_rev; gaa ctg gcc ggc cgc ctt gtc cag gat c

XPD_Y545S_fwd; gag agg aaa tcc ggc aag ggc tgg g

XPD_Y545S_rev; cca gcc ctt gcc gga ttt cct ctc

XPD_R567A_fwd; gga gat agg ggc gct cat aag gag tgc

XPD_R567A_rev; gca ctc ctt atg agc gcc cct atc tcc

XPD_Y185V_fwd; gca cct tat gca gtg ttt ctc aac agg

XPD_Y185V_rev; cct gtt gag aaa cac tgc ata agg tgc

XPD_Y182V_fwd; gtt ata gca cct gtg gca tat ttt ctc

XPD_Y182V_rev; gag aaa ata tgc cac agg tgc tat aac

XPD_F133A_fwd; ccg tac ttc aac gcc aag ata aga tcc g

XPD_F133A_rev; cgg atc tta tct tgg cgt tga agt acg g

XPD_R136A_fwd; ctt caa gat agc atc cga tga aac g

XPD_R136A_rev; cgt ttc atc gga tgc tat ctt gaa g

XPD_K134A_fwd; cgt act tca act tcg cga taa gat ccg atg

XPD_K134A_rev; cat cgg atc tta tcg cga agt tga agt acg

XPD_M115R_fwd; gca aaa ttc tgc aac cgg aag aag cgt gag

XPD_M115R_rev; ctc acg ctt ctt ccg gtt gca gaa ttt tgc

XPD_V324R_fwd; ggt ggg caa gcg acc ttt cag tta ttg c

XPD_V324R_rev; gca ata act gaa agg tcg ctt gcc cac c

XPD_E319R_fwd; cgt gga aaa cga gaa gcg aaa ggt ggg caa g

XPD_E319R_rev; ctt gcc cac ctt tcg ctt ctc gtt ttc cac g

XPD_R88M_fwd; cca atg cag ggc atg gtc aat atg

XPD_R88M_rev; cat att gac cat gcc ctg cat tgg

XPD_P530E_fwd; gcc ttt ccc gag gga gga cgc cat c

XPD_P530E_rev; gat ggc gtc ctc cct cgg gaa agg c

XPD_Y46A_fwd; tgc cct gca ggc ttc aag tga aag g

XPD_Y46A_rev; cct tta ctt gaa gcc tgc agg gcc ag

XPD_R88A_fwd; cca atg cag ggc gcg gtc aat atg tgc

XPD_R88A_rev; cac ata ttg acc gcg ccc tgc att gg

XPD_K170E_fwd; ctt atg aaa gca tgg agg cag cac tgc cag

XPD_K170E_rev; ctg gca gtg ctg cct cca tgc ttt cat aag g

XPD_C148S_neu_fwd; gga aca acg tgt ccc ctt atg aaa gc

XPD_C148S_neu_rev; gct ttc ata agg gga cac gtt gtt cc

XPD_E107A_fwd; cga gat aaa tgc agc atc tct tgc

XPD_E107A_rev; gca aga gat gct gca ttt atc tcg

XPD_E315A_fwd; ggc gaa tac gtg gca aac gag aag g

XPD_E315A_rev; cct tct cgt ttg cca cgt att cgc c

XPD_Y540S/Y545S_rev; ccc ttg ccg gat ttc ctc tca tag gtg tc

XPD_Y540S/Y545S_fwd; gac tcc tat gag agg aaa tcc ggc aag gg

XPD_F538S_fwd; tca aca gat cgc tgt ctg act

XPD_F538S_rev; agt cag aca gcg atc tgt tg

XPD_W549S_fwd; ggc aag ggc tcg gaa tac agc g

XPD_W549S_rev; cgc tgt att ccg agc cct tg

C73S fwd; cag ggt caa tat gtc cat ac

C73S rev; gta tgg aca tat tga ccc tg

C94S fwd; ctt gca aaa ttc tcc aac atg

C94S rev; cat gtt gga gaa ttt tgc aag

C108S fwd; tga ggc tgc atc ccc gta ctt caa c

C108S rev; gtt gaa gta cgg gga tgc agc ctc a

neu_Y545S/W549S_fwd; tat gag agg aaa tcc ggc aag ggc tcg gaa tac

neu_Y545S/W549S_rev; gta ttc cga gcc ctt gcc gga ttt cct ctc ata

neu_K170A_fwd; gcc ctt atg aaa gca tgg cgg cag cac tg

neu_K170A_rev; cag tgc tgc cgc cat gct ttc ata agg gc

neu_Y185A_fwd; gtt ata gca cct tat ggc agc ttt tct caa c

neu_Y185A_rev; gtt gag aaa agc tgc cat aag gtg cta taa c

neu_F226A_fwd; ggc aga tcc ata gga tcc gcc agg aat atc tg

neu_F226A_rev; cag ata ttc ctg gcg gat cct atg gat ctg cc

XPD_ R59A_fwd; ctt gtc gcc acc aat tcg

XPD_ R59A_rev; gaa ttg gtg gcg aca agg

XPD_R59E_rev; gaa ttg gtc tcg aca agg

XPD_R88A_fwd; gca ggg cgc ggt caa tat g

XPD_R88A_rev; cat att gac cgc gcc ctg c

XPD_R88H_fwd; gca ggg cca cgt caa tat g

XPD_R88H_rev; cat att gac gtg gcc ctg c

XPD_F326A_fwd; caa ggt acc tgc cag tta ttg

XPD_F326A_rev; caa taa ctg gca ggt acc ttg

XPD_Y425E_fwd; gtc atc aaa aga gga cac gc

XPD_Y425E_rev; gcg tgt cct ctt ttg atg ac

XPD_F538S/Y540S_fwd; gat cgc tgt ctg act cct atg a

XPD_F538S/Y540S_rev; tca tag gag tca gac agc gat c

XPD_D582N_fwd; gtg atc ctg aac aag agg

XPD_D582N_rev; ccc tct tgt tca gga tca c

XPD_R584E_fwd; gga caa gca ggc cgg cca g

XPD_R584E_rev; ctg gcc ggc ctg ctt gtc c

Acknowledgements

First of all, I would like to thank my supervisor Prof. Caroline Kisker for her guidance, advice, encouragement, and giving me the opportunity to work on this interesting project. Special thanks to the members of my advisory committee Prof. Thomas Müller, Prof. Yves Muller and Prof. Bennett van Houten for many scientific discussions and suggestions in my committee meetings. Furthermore I would like to thank Prof. Thomas Müller for accepting to be co-referent of my thesis. Special thanks to Prof. Sheila David for the great time I had during my research stay in her lab.

I am grateful to many people in the lab for their contributions to this present work: Florian Rohleder cloned the XPD pBadM11 construct. Gudrun Michels performed ATPase and helicase assays with the XPD mutants. Dr. Heide Marie Roth taught me how to perform electro mobility shift assays and helicase assays. Dr. Jochen Kuper for his supervision and scientific discussions. He solved the apo XPD structure and helped me with data collection and structure solution of the XPD-DNA complex. Furthermore he optimized the helicase assay. Prof. Hermann Schindelin I would like to thank for helping with building the apo XPD structure. I am grateful to Dr. Petra Hänzelmann for teaching me how to work in the anaerobic tent. Moreover I am thankful to Lisa Engstrom for having a great time working together on the electron paramagnetic resonance experiments. Korbinian Heil and Prof. Thomas Carell I would like to thank for kindly providing me CPD containing oligonucleotides. I would like to thank all members of the Kisker and Schindelin groups for scientific discussions and creating a good working environment.

Special thanks to Mike Friedrich for his constant support and his help with formatting and correcting the draft of this thesis.

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

Affidavit
(eidesstattliche Erklärung)

I hereby declare that my thesis entitled

“Structural and functional characterization of nucleotide excision repair proteins”

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

List of publications

Kuper J*, **Wolski SC***, Michels G, Kisker C: Structural and functional characterization of the XPD helicase defines a mechanism of DNA-translocation with implications for nucleotide excision repair. *EMBO accepted*

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

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

*both authors contributed equally