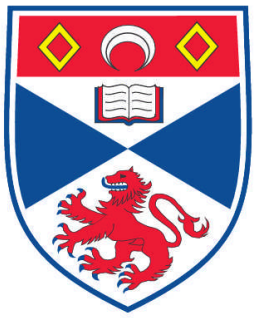
CHARACTERISATION OF XPD FROM SULFOLOBUS ACIDOCALDARIUS : AN IRON-SULPHUR CLUSTER CONTAINING DNA REPAIR HELICASE

Jana Rudolf

A Thesis Submitted for the Degree of PhD
at the
University of St. Andrews



2007

CHARACTERISATION OF XPD FROM *SULFOLOBUS ACIDOCALDARIUS*: AN IRON-SULPHUR CLUSTER CONTAINING DNA REPAIR HELICASE

JANA RUDOLF

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

OCTOBER 2006

University
of
St Andrews

CONTENT

FIGURES AND TABLES

ABBREVIATIONS

ADP	adenosine 5'-diphosphate
AEBSF	4-(2-aminoethyl)-benzenesulfonylfluoride-HCl
Ape	*Aeropyrum pernix*
ATP	adenosin 5'-triphosphate
ATP-γS	adenosin 5'-(O)-(3-thio)triphosphate
BA	bathophenanthroline
BER	base excision repair
bp	base pair
BSA	bovine serum albumine
CAK	Cdk activating kinase
cdk	cyclin dependent kinase
CS	Cockayne syndrome
DinG	damage inducible helicase G
dsDNA	double stranded desoxyribonucleic acid
DT	dithionite
DTT	dithiothreitol
EDTA	ethylene diamine teraacetate
EMSA	electrophoretic mobility shift assay
EPR	electron paramagnetic resonance
FA	Fanconi anemia
FancJ	Fanconi anemia complementation group J
FCL	iron-sulphur cluster loop
FeS	iron-sulphur
GTP	guanosine 5'-triphosphate
HhH	helix-hairpin-helix
ICP-OES	inductively coupled plasma – optical emission spectroscopy
IPTG	isopropyl β-D-thiogalactopyranoside
IRE	iron responsible element

IRP	iron regulatory protein
ITC	isothermal titration calorimetry
MALDI-ToF	Matrix-Associated Laser Desorption Ionization – Time of Flight
NA	nucleic acid
NER	nucleotide excision repair
nt	nucleotide
NTP	nucleoside triphosphate
PCNA	proliferating cell nuclear antigene
RNPII	RNA polymerase II
ROS	reactive oxygen species
RT	room temperature
Sac	*Sulfolobus acidocaldarius*
SDS	sodium dodecyl sulfate
SDSL	side-directed spin-label(ling)
SDS-PAGE	SDS-polyacrylamide gelelectrophoresis
ssDNA	single stranded desoxyribonucleic acid
Sso	*Sulfolobus solfataricus*
TBE	tris-borate-EDTA buffer
TCA-DOC	trichloroacetic acid – deoxycholate
TCC	tricarbonic acid cycle
TFIIH	transcription factor II H
TTD	trichothiodystrophy
UV	ultra violet
XP	*xeroderma pigmentosum*
XPA/B/C/D/E/F/G	*xeroderma pigmentosum* complementation group A/B/C/D/E/F/G

DECLARATION

I, Jana Rudolf, hereby certify that this thesis, which is approximately 40000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date _____ Signature of candidate _____

I was admitted as a research student in October 2003 and as a candidate for the degree of PhD in October 2004; the higher study for which this is a record was carried out in the University of St Andrews between 2003 and 2006.

Date _____ Signature of candidate _____

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date _____ Signature of supervisor _____

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date _____ Signature of candidate _____

ABSTRACT

DNA is constantly damaged by a variety of exogenous and endogenous sources. To maintain the integrity of the genome, different DNA repair mechanisms have evolved, which deal with different kinds of DNA damage. One of the DNA repair pathways, Nucleotide Excision Repair (NER), is highly conserved throughout the three kingdoms of life and deals mainly with lesions arising in the DNA duplex after exposure to UV-light. The NER pathway in archaea is more closely related to that of eukarya, although the overall process is not yet well understood. This thesis describes the isolation and characterisation of one of the repair factors, XPD, from the crenarchaeon *Sulfolobus acidocaldarius* (SacXPD).

SacXPD was first identified due to its homology with the eukaryal XPD protein. In eukarya XPD is the 5'→ 3' helicase involved in opening the DNA duplex around a damaged site. In eukarya, XPD is part of a 10-subunit complex, where it fulfils important structural roles and takes part in NER, transcription initiation from RNA polymerase II promoters and cell cycle regulation. The archaeal protein on the contrary is a monomer and a 5'→ 3' SF2 DNA helicase as its eukaryal counterpart. Its cellular functions, however, are unclear.

Upon purification of SacXPD, it was discovered that the protein binds an iron-sulphur cluster (FeS), which is essential for its helicase activity, but not for any other enzymatic functions, such as the ATP hydrolysing activity. The FeS cluster domain was not only identified in archaeal XPD, but also in eukaryal XPD and other related eukaryal helicases, such as FancJ. The presence of the FeS cluster was confirmed in the eukaryotic XPD homologue Rad3 from *Saccharomyces cerevisiae*. Mutagenesis studies were used to investigate a possible function of the FeS cluster, which may be used to engage ssDNA during the duplex unwinding process. This would actively distort the ss/ ds DNA junction. In addition, the resulting bending of the clamped single DNA strand could be used to avoid reannealing. The consequences of some human mutations introduced into the SacXPD gene were investigated and could contribute to our understanding of the development of human diseases.

ACKNOWLEDGEMENTS

The success of this PhD compromises the involvement and support of a range of different people.

First of all I have to thank my supervisor, Prof Malcolm F. White, who is an excellent teacher, for his support, encouragement and guidance. I would like to thank present and past members of the Mass spec team for analyzing all these protein samples. Thanks to Dr W. John Ingledew for EPR analysis and useful discussions on FeS proteins. I am grateful to Prof Michael J. Stark and Dr Vasso Makrantoni for excellent work and discussion on yeast Rad3. Thanks also to Prof Stephen K. Chapman (University of Edinburgh) for providing access to the ICP-OES, to Dr Rona Ramsey (University of St Andrews) for help with the FeS cluster reconstitution, to Dr Sharon Kelly for initial CD-spectroscopy, Prof Alan Cooper and Mrs Margarete Nutley for the DSC experiments (University of Glasgow), Dr David Keeble (University of Dundee) and Dr Hassane ElMkami (University of St Andrews) for access to the EPR spectrometer, Dr Uli Schwarz-Linek for help with the ITC, Mr Paul Talbot for N-terminal sequencing, Mrs Biljana Petrovic-Stojanovska for some of the XPD mutants (University of St Andrews) and the AIRC for funding. Thanks to Liza, Vasso, Sonia, Robert and Richard for reading this thesis correction!!!

Thanks to everyone in the past and present MFW and the PJC group for their support and for putting up with me for the past three years. Especially to Liza Cubeddu, who is a great teacher and friend and contributed a lot to this work. Thanks to Vasso Makrantoni for her friendship, help and expertise in allen Lebenslagen. Coffee? Thanks to Dorothee Goetz (including family) for her help, friendship und dass sie immer ein offenes Ohr fuer mich hatte. Thanks to Robert Dorazi who always had to sustain my complains, for his friendship, tolerance and support.

Thanks to the Viking for being a great friend, his support and for being more Scottish than any Scot I have ever met. Thanks to Federica Catti, who was an excellent flat mate and friend. To Torsten Seyfarth, he is a great friend and always supported me and without him I would not have come to St Andrews (so for everyone: it's mainly HIS fault). Thanks to Yves Cossette, for his friendship and his English lessons. Thanks to all

my friends for being there in up's and down's, especially to Anja Poehlein, Cristina Lucas, Till Blum, Martin Schueler, Anja Frisch, Cosimo Cadicamo and Martial Hille. Thanks to the French community and my basketball team for loads of great matches and injuries. And last but not least, a big thanks to my family, who were always supporting me and were always there for me.

1 INTRODUCTION

1.1 ARCHAEA-THE THIRD DOMAIN OF LIFE

Life as we know it is divided into the prokaryotic (life forms without nucleus) and the eukaryotic (life forms with nucleus) world. When Fox and Woese reformed the taxonomy of the bacteria on the basis of similarities of their 16S-rRNA, rather than physical markers, it led to the discovery of the third domain of life, which they named "Archaebacteria" (Woese & Fox (1977)). Archaea share many similarities with bacteria. They are single celled organisms, possess no real nucleus and usually have a single circular chromosome. On the other hand archaeal DNA processing pathways are more related to the eukarya. This includes the DNA repair machineries, gene transcription, DNA replication and recombination and the mRNA translation process. These pathways in archaea tend to be stripped down versions of the eukaryal equivalents. Understanding fundamental features in archaea is helpful to unravel the complexity of the eukaryal pathways. Archaea are also easier to handle and provide useful insights into gene and protein evolution.

The domain of the archaea is divided into two kingdoms: the crenarchaea and the euryarchaea. The latter consists mainly of methanogenes, extreme halophiles, (hyper)thermophiles and marine mesophiles. They are thought to be the closer relatives of the eukarya (White (2003)). The crenarchaea include many (hyper)thermophiles as well as marine and terrestric mesophiles. Spontaneous mutations that could drive evolution are less tolerable under extreme environmental conditions. Hence, the evolution of extremophiles is slower than evolution of organisms living under mesophile conditions. Therefore, the archaea are thought to be the closest living relatives to the last universal common ancestor (LUCA) of life on earth.

It is worth mentioning, that there may be another kingdom: The korarchaea. The existence of the korarchaea was only suggested on the basis of 16S-rRNA analysis of samples from a hyperthermophile biocoenosis (Barns et al. (1996)). However, their existence was never proved nor disproved, but it is likely that the identified 16S-rRNA was an artifact. The Nanoarchaea were first thought to represent a separate kingdom as well, on the basis of the isolation of *Nanoarchaeum equitans*, a parasite of another archaeon, which also possesses the smallest genome of any known organism (Huber et al. (2002); Waters et al. (2003)). Later phylogenetic studies, however, suggested that *N.*

equitans may be a member of a fast-evolving euryarchaeal lineage (Brochier et al. (2005)).

In this project, proteins from the two crenarchaeal hyperthermophiles *Sulfolobus acidocaldarius* DSM639 (XPD) and *S. solfataricus* P2 (XPF) are described. Both organisms grow aerobically in acidic (pH 2-4), sulphur-rich environments at around 80 ˚C. Genomes of *S. solfataricus* and *S. acidocaldarius* have been fully sequenced (She et al. (2001), Chen et al. (2005)). Recombinant proteins were expressed and purified using *Escherichia coli* as a heterologous expression host. *Sulfolobus* proteins are generally thermostable and easily amenable to biochemical and structural studies. XPD (Saci_0192) was first identified on the basis of its sequence similarity to its eukaryotic counterpart; XPF (Sso0729) was already extensively studied by J. Roberts, a former PhD student.

1.2 DNA DAMAGE AND REPAIR

DNA is constantly damaged by exogenous (e.g. UV-light, ionizing radiation) and endogenous (e.g. reactive oxygen species produced in metabolic processes) agents. In humans, the number of damages occurring per cell per day was estimated to be around 10^4-10^6 (Lindahl (1993)). If modifications persist, they can give rise to mutations, cancer and cell death.

The simplest reaction leading to DNA damage is hydrolysis. Hydrolysis of the N-glycosylic bond between a purine and its deoxyribose sugar leads to depurination. The generated abasic site can be further fragmented to single strand breaks. Hydrolysis can also cause deamination of bases and can alter the genome sequence. For example, the deamination of cytosine leads to uracil, which base pairs with adenine. If not repaired, replication of this mispair will cause a C-G to T-A transversion, permanently altering a genetic sequence.

Oxidation is a major threat to DNA. Reactive oxygen species (ROS) are frequent by-products of the aerobic metabolism. Oxidised bases (e.g. 8-oxoguanine, thymine glycol) block transcription and replication and are highly mutagenic. Other metabolic

hazards can be found in form of enzyme cofactors such as S-adenosylmethionine, which methylates bases accidentally.

Environmental factors, like UV radiation, ionizing radiation, cigarette smoke or anticancer drugs are potentially mutagenic. They cause bulky helix-distorting lesions or damage to the DNA sugar phosphate backbone. This can provoke single strand breaks, and eventually double strand breaks during replication. Double strand breaks also occur at replication fork collapses, but also during meiosis and recombination. Rare incidences of interstrand crosslinks are also highly cytotoxic. (reviewed in Scharer (2003))

To cope with such a variety of DNA damages, cells have developed several DNA repair systems, most of which are highly conserved throughout evolution. An overview is given in Figure 1.1.

XPD and XPF are part of the Nucleotide Excision Repair (NER) machinery, a highly versatile and conserved repair pathway. It deals with a variety of bulky helix-distorting lesions mainly arising from exposure to UV light. The most common adducts are cyclobutane pyrimidine dimers (CPD) and (6-4) photoproducts (6-4-PP).

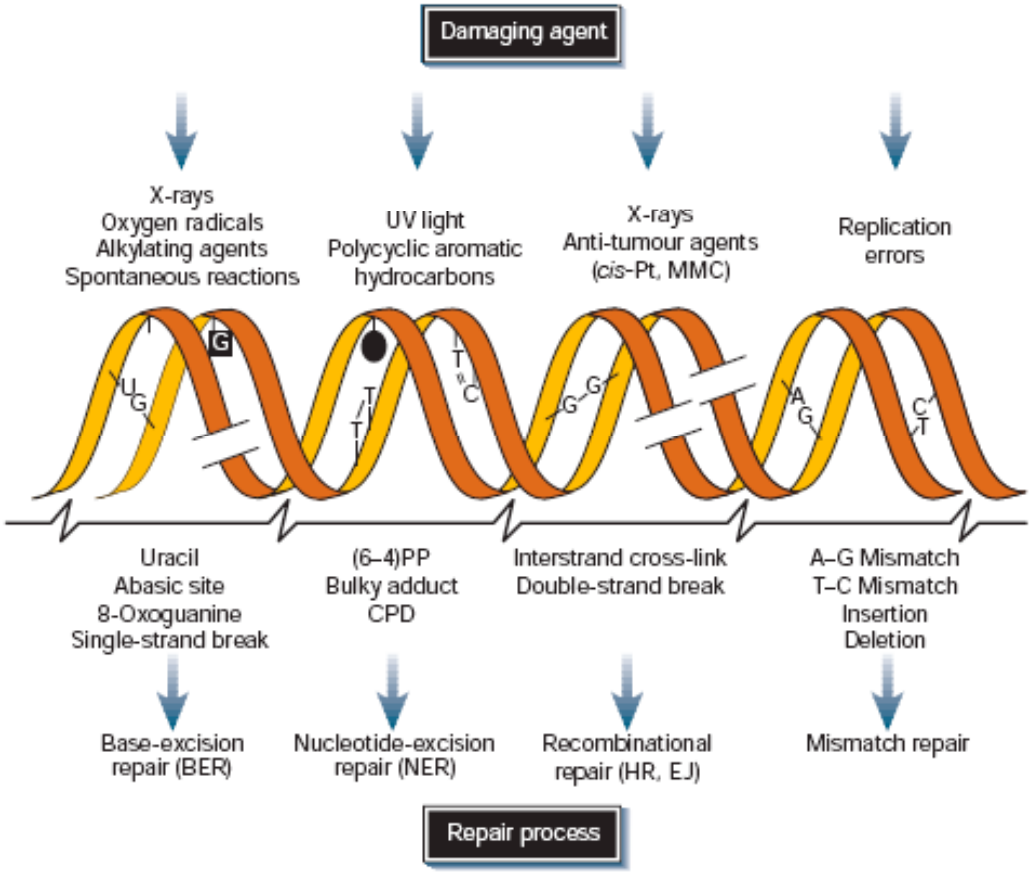


Figure 1.1: Overview of DNA damaging factors, lesions and repair mechanisms

DNA is constantly damaged by exogenous and endogenous agents. Different repair pathways have evolved, which deal with different types of lesions. Figure adapted from Hoeijmakers (2001).

1.2.1 Nucleotide excision repair in the three domains of life

The overall mechanism of NER is highly conserved among all organisms, although the participating factors are not. In bacteria the process basically involves three proteins (UvrA, UvrB and UvrC), while in eukarya circa 15 proteins constitute the core repair process. It is thought that the NER system in archaea could resemble a simplified version of the eukaryal NER pathway, because homologues of some of the eukaryal NER factors were identified in archaea. Other archaea also possess the UvrABC system or use a combination of both NER systems (White (2003)). It was suggested, that some archaea

acquired the bacterial NER factors through lateral gene transfer (Aravind et al (1999)). Two distinct NER pathways have evolved: Global Genome Repair (GGR) and Transcription Coupled Repair (TCR). As the names indicate, the GGR system repairs lesions occurring throughout the genome, while TCR repairs damage arising in the actively transcribed DNA strand (van Hoffen et al. (2003)).

Most organisms, apart from placental mammals, encode a photolyase, an enzyme that directly reverses CPD (the most common lesion arising following UV-exposure) in a light dependent reaction. In human cells NER is the only way to deal with DNA photoproducts, since no photolyase homologue has been identified in the genome (Yasui et al. (1994); Sancar (2003)).

1.2.1.1 GGR

In bacteria, UvrA binds specifically to damaged DNA (Seeberg & Steinum (1982)), although the damage signalling factor is UvrB (Theis et al. (2000)). UvrA dimerises upon binding to DNA and binds one or two molecules of UvrB (Orren & Sancar (1989); Verhoeven et al. (2002b)). UvrA delivers UvrB to the damaged site (Orren et al. (1992)), where UvrB unwinds the DNA around the lesion, locks itself using a β-hairpin structure and actively kinks the DNA around 130 ˚ (Theis et al. (2000); Shi et al. (1992)). UvrA leaves the complex in an ATP-dependent manner and the endonuclease UvrC, which has high affinity for the UvrB-DNA complex, binds UvrB and performs a dual incision. The 3' incision is first made at the 4^{th} to 6^{th} phosphodiester bond before the lesion and the following 5' incision occurs at the 8^{th} phosphodiester bond (Sancar & Rupp (1983)). The 12-13 nt long excised oligonucleotide together with the bound UvrC are removed by the helicase UvrD. UvrB, which is bound to the strand opposite the damage, remains at the damaged site until DNA polymerase I fills in the gap and removes it at the same time (Orren et al. (1992)). Finally, the gaps between old and new strands are sealed by a DNA ligase.

E. coli harbours another 3' endonuclease, called Cho (Uvr*C ho*mologue), which is able to act in combination with UvrA and UvrB to make the 3' incision at the 9^{th} phosphodiester bond. Cho is not essential for bacterial NER, but it is thought to make

some contribution to DNA repair within this pathway (Moolenaar et al. (2002); van Houten (2002)).

In eukarya, NER is carried out by six repair factors called RPA, XPA, XPC, TFIIH (Transcription Factor II H), XPG and XPF-ERCC1. Three factors appear first at the damaged site (RPA, XPA, XPC), although it is unclear in which order. Another dilemma is that none of the three are specific for CPDs, the major substrate for human NER. XPC prefers to bind undamaged DNA over CPDs, while RPA and XPA have a non-physiologically relevant preference for CPDs (Reardon & Sancar (2003)). Another DNA damage binding protein, DDB, was proposed to initiate excision of CPDs (Wagasugi et al. (2001)), but it was found that DDB-knockout Chinese hamster cell extracts still efficiently repaired CPD lesions (Reardon et al. (1997)). Therefore, it was suggested that any of these repair initiation factors, which have affinities for each other, could appear first at the damaged site and assemble in a random order (Figure 1.2). Eventually, protein-protein interactions and cooperativity would lead to the assembly of the NER machinery. The interaction of XPC with TFIIH leads to formation of a rather unstable closed complex (Uchida et al. (2002)). The two helicases of TFIIH, XPB and XPD, unwind the DNA around the lesion in 3'→5' and 5'→3' direction, respectively. This leads to conformational changes within all bound proteins and the resulting stable open complex is called preincision complex 1 (PIC1). XPC recruits XPG to the 3' incision site and leaves the complex (PIC2), similar to how UvrA swaps with UvrC in bacteria. XPF-ERCC1 is recruited and strongly bound by XPA (PIC3) and XPG facilitates the 3' incision at the $6^{th} \pm 3$ phosphodiester bond from the damage, followed by the 5' incision by XPF-ERCC1 at the $20^{th} \pm 5$ phophodiester bond. The excised oligonucleotide has a typical length of 24-32 nt. Subsequently, all repair factors except RPA leave and the remaining gap becomes subject to repair synthesis carried out by repair factor C (RFC)-PCNA and polymerase δ/ε which fills the gap. DNA ligase 1 finally seals the repair patch (Huang et al. (1992); Mu et al. (1997); Evans et al. (1997); Reardon & Sancar (2005)).

Little is known about the overall process of NER in archaea. Although homologues of RPA, XPB, XPD, XPF and XPG have been identified, the damage recognition factors XPA and XPC have not been detected in archaeal genomes (White (2002)). Interestingly,

lower eukaryotes and plants do not possess homologues of XPA either, and it is thought that this gene was added later to the repair pathway. The single stranded DNA binding (SSB) protein (RPA in eukarya) is a potential candidate for damage recognition and NER initiation (Cubeddu & White (2005)).

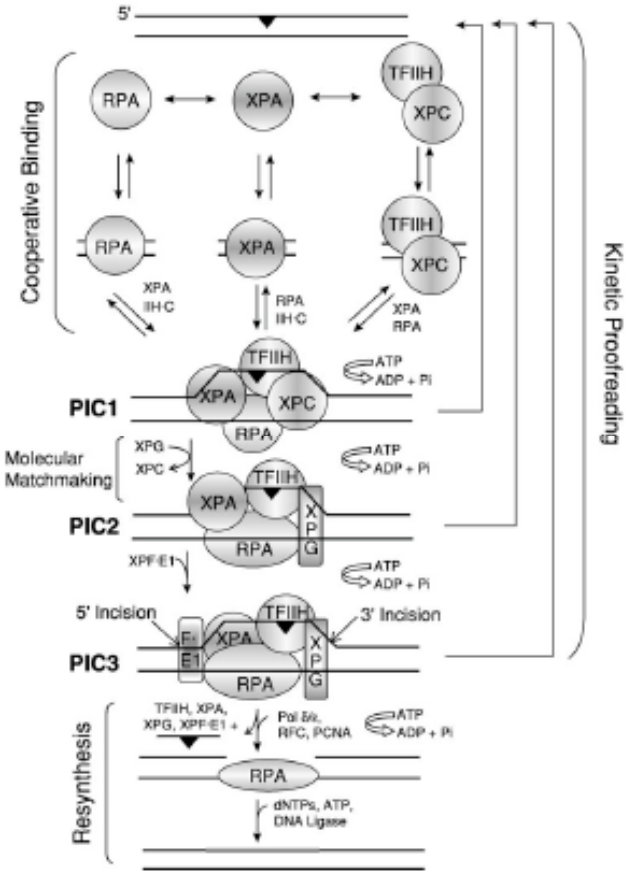


Figure 1.2: Model of nucleotide excision repair in human cells

The damage recognition factors XPA, XPC and RPA assemble in a random but cooperative manner at the damaged site (black triangle). XPC recruits TFIIH whose two helicases (XPB and XPD) unwind the DNA around the damage in opposite directions leading to formation of presincision complex 1 (PIC1). XPC recruits XPG to PIC1 in a process called molecular matchmaking and leaves PIC1. The resulting complex (PIC2) is recognised by XPF-ERCC1 leading to formation of PIC3. Dual incision takes place and a damage containing oligonucleotide of a typical length of 24-32 nt is released. Steps leading to the formation of the three PIC complexes may be aborted at any time (kinetic proofreading). The gap is filled in by polymerase δ/ε and the repair patch sealed by ligase (figure adapted from Reardon & Sancar (2005)).

1.2.1.2 TCR

Damages in the transcribed strand during transcription stall the RNA polymerase (RNAP) in *E. coli*. Stalled RNAP complexes are recognised by a transcription repair coupling factor (TRCF), also known as Mfd. TRCF functions as a translocase and release factor for RNAP at blocking lesions and is also thought to recruit the NER factors, which remove the damage from the transcribed strand (Selby & Sancar (1993)).

The mechanism in eukarya is not very clear. NER of transcription blockages does not require XPC, while the transcription coupling factors involved are CSA and CSB. The latter seems to be the orthologue of bacterial TRCF and is involved in recruitment of the NER machinery, but does not lead to release of RNAPII once it is stalled at a lesion. The release of RNAPII may be facilitated by transcription termination factor 2 (TTF2). (reviewed in Reardon & Sancar (2005)).

There are no published data on TCR in archaea and a *mfd* homologue has not yet been identified in archaeal genomes.

1.3 HELICASES

Helicases are powerful motor proteins which transfer the free energy derived from NTP hydrolysis onto translocation along nucleic acid (NA). This results in unwinding of duplex NA and provides single stranded DNA or RNA. Helicases are found in all DNA processing pathways and their general function is to provide access to genomic information.

On the basis of conserved residues, helicases are categorised into three superfamilies (SF1-3) and one family (F4) (Gorbalenya & Koonin (1993)). These (super)families are characterised by seven (SF1 and 2), three (SF3) and five (F4) conserved motifs, respectively. Two motifs are common to all helicases (motifs I and II/ A and B) and they are called Walker A and B motifs (Walker et al. (1982)). Generally, lysine and glycine residues of the Walker A box bind the NTP and coordinate the β-γ phosphoanhydride bond, while acidic residues in the Walker B box direct a divalent cation (mostly a Mg^{2+} or Mn^{2+}) to activate water for hydrolysis of the terminal phosphoanhydride bond (Velankar et al. (1999)). The other motifs are specific for each

family and were thought to facilitate specific functions in respective helicases, such as substrate binding and unwinding. However, crystal structures and biochemical studies revealed that all motifs are somehow involved in NTP binding and hydrolysis. They come together in the three dimensional structure of the protein to form the NTP binding and hydrolysing pocket within the two RecA-like motordomains (Korolev et al. (1997); Velankar et al. (1999)). It is worth mentioning that for the majority of putative helicases, helicase activity has not been shown. It is generally accepted that the so-called helicase motifs are actually characteristic for NTPases, rather than helicases. Hence, specific properties of a helicase are encoded outside the conserved motifs (reviewed in Hall & Matson (1999); Singleton& Wigley (2002)).

Superfamily 1 and 2 are the largest families and are thought to derive from a common ancestor. They harbour most of the eukaryotic, prokaryotic and viral DNA or RNA helicases and NTPases. With a few exceptions, all members of these superfamilies are ATPases. In addition to the seven conserved motifs, another motif was recently discovered in DEAD-box RNA helicases, called the Q-motif. It is located prior to motif I and is involved in the coordination of the β-γ phosphoanhydride bond of the NTP (Cordin et al. (2004)).

XPD, both in archaea and eukarya, is a SF2 DNA-helicase. It contains the seven conserved motifs and the conserved sequence Asp-Glu-Ala-His (DEAH) in the Walker B box, typical for SF2 helicases (Gorbalenya & Koonin (1993); see also Chapter 3, section 1.3.5).

1.3.1 Models of nucleic acid unwinding

Helicases unwind duplex nucleic acid to produce single stranded DNA/ RNA. Duplex unwinding mostly occurs with a specific polarity, either in 3’ to 5’ or 5’ to 3’ direction. How directionality is facilitated is still unclear, but how the enzymes manage to unwind nucleic acid became much clearer in recent years. There are three different models of nucleic acid unwinding, which depend on the oligomeric state of the helicase. They are called the Inchworm Model, the Active Rolling Model and the Model for hexameric Helicases.

1.3.1.1 The Inchworm Model

The inchworm model was first proposed for the Rep helicase (Yarranton & Genfter (1979)) before it was discovered that the active form of Rep was a dimer. Since then the model has been thoroughly refined, especially with studies on the PcrA helicase. The inchworm model (Figure 1.3A) describes an unwinding mechanism for monomeric helicases assuming two sites of DNA binding per monomer. The first site contacts single stranded DNA, which induces conformational changes, leading to contact of the second motor domain with dsDNA. Binding of ATP increases the affinity for dsDNA and decreases that for ssDNA. When ATP is hydrolysed another large domain movement is induced, which can be imagined as a contraction. Since the enzyme is still tightly bound to the dsDNA region, this contraction leads to a movement of the ssDNA binding domain one step size in the direction of the ss/ ds fork. Bound ADP increases the affinity for ssDNA and decreases that for dsDNA. When ADP and P_i is released, the enzyme "snaps" open, which leads to a movement of one step size of the dsDNA binding domain away from the ss/ ds fork. Overall, this would look like the enzyme inches along the DNA strand (Velankar et al. (1999)).

Step sizes

The step size of a helicase per hydrolysed ATP can be defined as the step size of translocation or the step size of duplex unwinding. Duplex unwinding requires more energy and therefore step sizes are smaller. For monomeric helicases the number of base pairs unwound per hydrolysed ATP varies from one for the PcrA helicase (Dillingham et al. (2000)) to up to six for the NPH-II helicase (Jankowsky et al. (2000)).

Modifications of the general Inchworm Model

A few modifications of the general inchworm model are described: the Mexican wave model (Figure 1.3B) and the cooperative inchworm model. For the SF1 helicase PcrA it was found that the enzyme flips out one base for every hydrolysed ATP near the ss/ ds fork and releases another base on the opposite end. This would simulate a Mexican wave, while the enzyme translocates along the single DNA strand (Velankar et al. (1999)). Other helicases, such as the SF2 RNA-helicase NPH-II, do not interact with the

bases, they only bind the ribose phosphate backbone of one strand and "strip off" the opposite strand during translocation (Kawaoka et al. (2004)). The cooperative inchworm model was proposed for the Dda-helicase, because it was found that the rate of duplex unwinding increases with increasing concentrations of protein (Byrd & Raney (2004)).

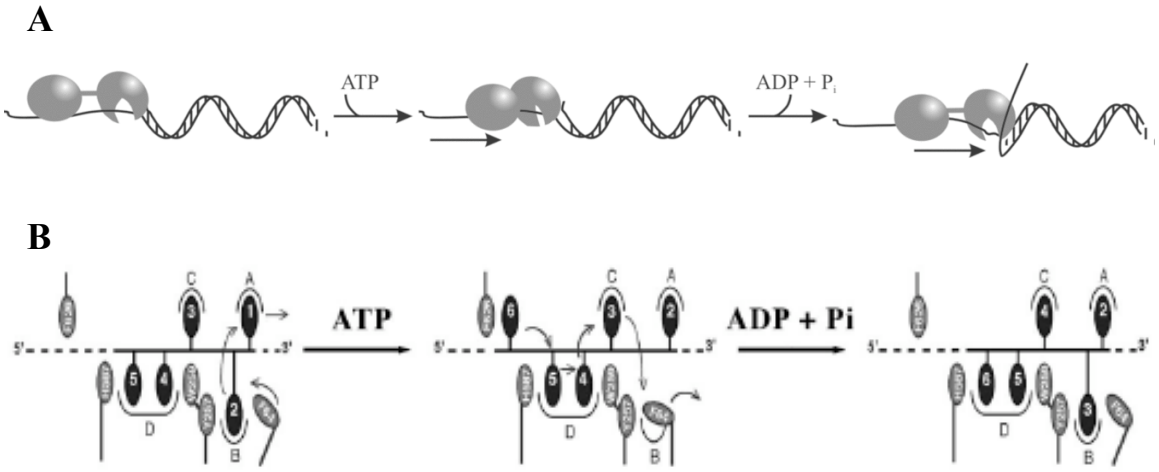


Figure 1.3: DNA unwinding mechanism for monomeric helicases

(A) The inchworm model proposes two DNA binding sites per monomer, a single stranded and a double stranded DNA binding site. Cyclic changes of ATP binding, hydrolysis and ADP + P_i release induce conformational changes in the enzyme and alternating changes in affinities for ssDNA and dsDNA. The movement resembles that of a moving inchworm and leads to duplex unwinding (adapted from Eoff & Raney (2005)). (B) The Mexican Wave model was proposed for the SF1 helicase PcrA and presumes a direct interaction of the ssDNA binding domain with the bases of the DNA strand. For each hydrolysed ATP a base near the ss/ ds fork is flipped out and released on the other end of the bound DNA strand (adapted from Soultanas & Wigley (2000)).

1.3.1.2 The Active Rolling Model

This model, although debatable, was proposed for dimeric helicases with two sites of nucleic acid binding, one per monomer (Figure 1.4). The mechanism suggested for the *E. coli* Rep helicase is as follows (Lohman (1993)): one or both monomers are bound to ssDNA with a specific polarity. Binding of ATP induces a negative cooperativity and increases the affinity of one monomer to ssDNA, while it decreases the affinity of the other one. The latter has now an increased affinity for dsDNA and binds the DNA duplex ahead of the ss/ ds fork. ATP hydrolysis induces duplex melting, leading to strand

displacement with the result that both subunits are bound to ssDNA. Exchange of ATP for ADP completes the cycle.

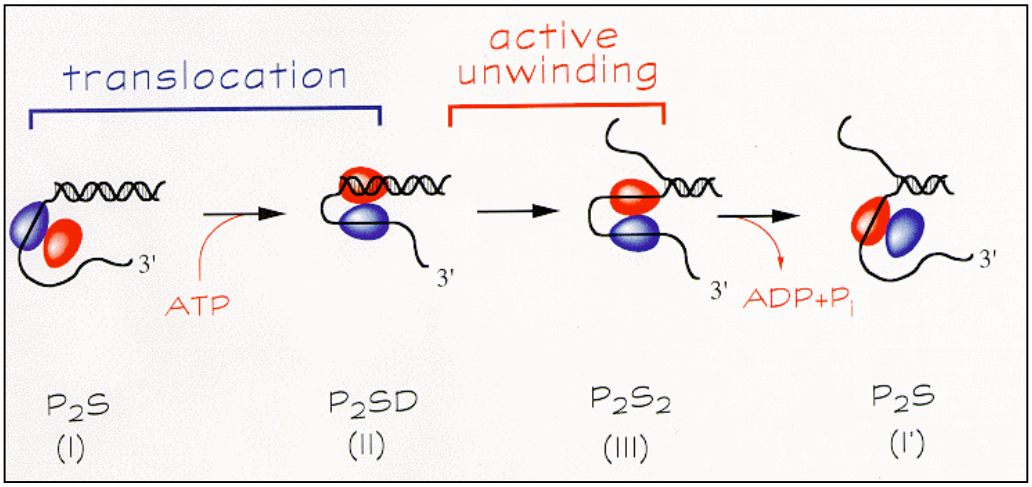


Figure 1.4: DNA unwinding mechanism for dimeric helicases

The active rolling model for dimeric helicases assumes one DNA binding site per monomer. Conformational changes and changes in affinities to ss/ dsDNA are induced by ATP hydrolysis. The emerging protein–DNA complexes are shown in the lower panel: P, protein monomer; S, ssDNA; D, dsDNA (adapted from Korolev et al. (1997)).

1.3.1.3 Hexameric Helicases

Hexameric helicases are complicated and can either consist of identical monomers, trimers of dimers or dimers of trimers. They form a ring-like structure through whose central hole single stranded nucleic acid can pass, while the helicase translocates along its nucleic acid substrate. NTP hydrolysis is carried out in a sequential manner. Allosteric effects induced by NTP hydrolysis cycles change the conformations of the enzyme subunits and the affinity to nucleic acid (reviewed in Soultanas & Wigley (2000); Eoff & Raney (2005)).

The crystal structure of the E1 protein from papillomavirus suggested a general mechanism for translocation of hexameric helicases (Figure 1.5; Enemark & Joshua-Tor (2006)). The subunits bind the ssDNA backbone in a helical right-handed staircase-like manner. Alternating states of ATP binding, hydrolysis and ADP release change the conformation of the central hairpins interacting with the DNA. Each subunit tracks one nucleotide per hydrolysed ATP. The subunit on the bottom of the staircase releases ADP and upon binding of ATP moves to the top of the staircase. Subunits in intermediate positions obtain varying states of ATP-hydrolysis and product formation ($ADP+P_i$). The ring structure and interdomain interactions avoid early dissociation from the nucleic acid and ensures completion of events like replication.

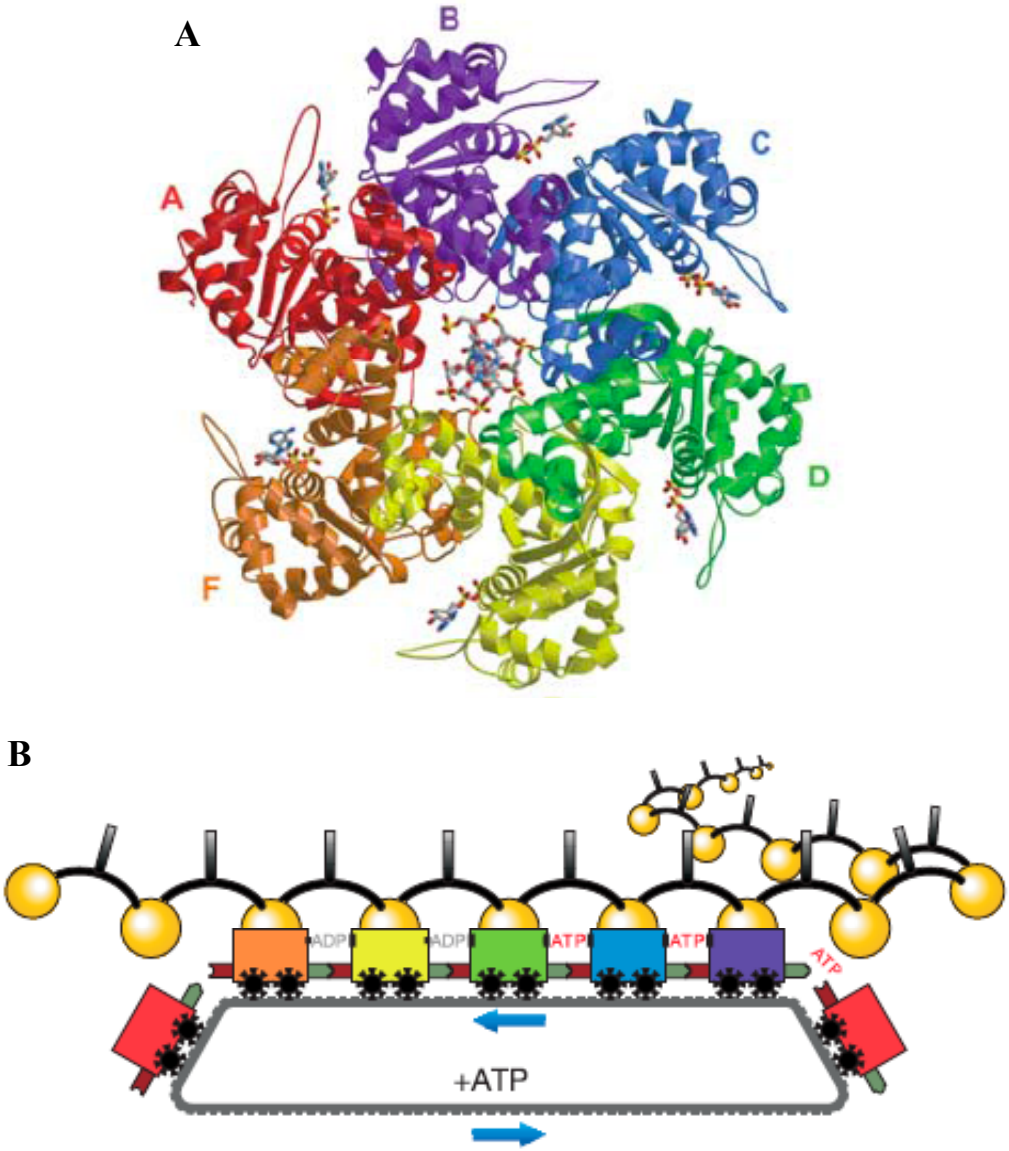


Figure 1.5: Model of nucleic acid unwinding by hexameric helicases

(A) The crystal structure of hexamer 1 of the E1 protein from papillomavirus is shown. The six subunits are colour-coded. The DNA in the central hole and the ADP molecules at the subunit interfaces are depicted as stick representations. (B) The mechanism of staircasing for translocation along ssDNA is illustrated. The six individual subunits are represented as wagons with the same colour code as in (A). Each wagon transports one DNA nucleotide from the right to the left. The wagon couplers depict staircasing interactions. The left wagon (red) releases its DNA nucleotide. Upon ATP binding it moves to the right, where it picks up the next DNA nucleotide, couples to the purple wagon and moves with the DNA nucleotide to the left again (Figure adapted from Enemark & Joshua-Tor (2006)).

1.3.2 Human XPD

In humans, XPD (ERCC2) is the 5'→3' helicase (Weber et al. (1988)) and as a component of TFIIH it takes part in NER, transcription initiation of genes regulated by RNA polymerase II (RNPII) promoters and cell cycle regulation (Orphanides et al. (1996); Lehmann (2001)). The orthologues and aliases are known as Rad3 (*Saccharomyces cerevisiae*) and rad15 (*Schizosaccharomyces pombe*). XPD is an SF2 helicase and it contains the seven conserved helicase motifs and a highly conserved D-E-A-H signature sequence in motif II (Gorbalenya & Koonin (1993)). It bridges the two subcomplexes of TFIIH by binding to the p44 subunit of the TFIIH core with its C-terminus, and with its N-terminus it contacts the MAT1 subunit of the CAK subcomplex (Coin et al. (1998); Tirode et al. (1999)).

1.3.2.1 Functions of human XPD and its role in TFIIH

The core TFIIH complex comprises the five subunits XPB, p34, p44, p52 and p62 (Figure 1.6). As part of this complex, hXPD is the major helicase in NER (Coin et al. (1998)).

The C-terminus of hXPD tethers the p44 subunit, which stimulates its helicase activity up to 10-fold during NER (Coin et al. (1998)). The amino terminal region interacts with MAT1 (Busso et al. (2000); Feaver et al. (2000)), a subunit of the Cdk Activating Kinase (CAK) complex, which consists of three subunits (MAT1, cyclin-H, cdk-7; Figure 1.6). The remaining factor is called TTD-A (p8), a 8 kDa protein, that is not as strongly associated with the complex as the other five proteins. It targets XPD and plays an essential role in stabilising TFIIH (Cleaver (2005)). CAK and the TFIIH core in concert with XPD and TTD-A represent the TFIIH holocomplex.

XPD is catalytically active during NER and plays an important structural role during transcription initiation, by bridging the TFIIH core and the CAK subcomplex (Tirode et al. (1999)). The cdk-7 subunit of the CAK complex phosphorylates the C-terminus of the large subunit of RNPII, thereby facilitating promoter escape and the recruitment of mRNA processing enzymes (Winkler et al. (2000)). Furthermore, as part of the CAK complex, XPD regulates cdk-7, which in turn regulates cell cycle through

phosphorylation of other downstream CDKs. Excess XPD was found to inhibit mitosis, while low levels of XPD lead to cell over-proliferation (Chen et al. (2003)). XPD probably releases the CAK complex during interphase and is degraded to allow cellular redistribution of cdk-7 during the cell cycle. High levels of XPD also have dramatic effects on viability in *Drosophila melanogaster* embryos. Hence, expression levels of XPD need to be tightly regulated during cell cycle. Accordingly, downregulation of expression levels of XPD during prophase coincides with repression of basal transcription of genes regulated by RNPII promoters (Chen et al. (2003)). Thus, the importance of XPD in eukaryotic cells is obvious. It acts not only as a helicase in NER, but has also very important structural and regulatory functions during transcription and cell cycle.

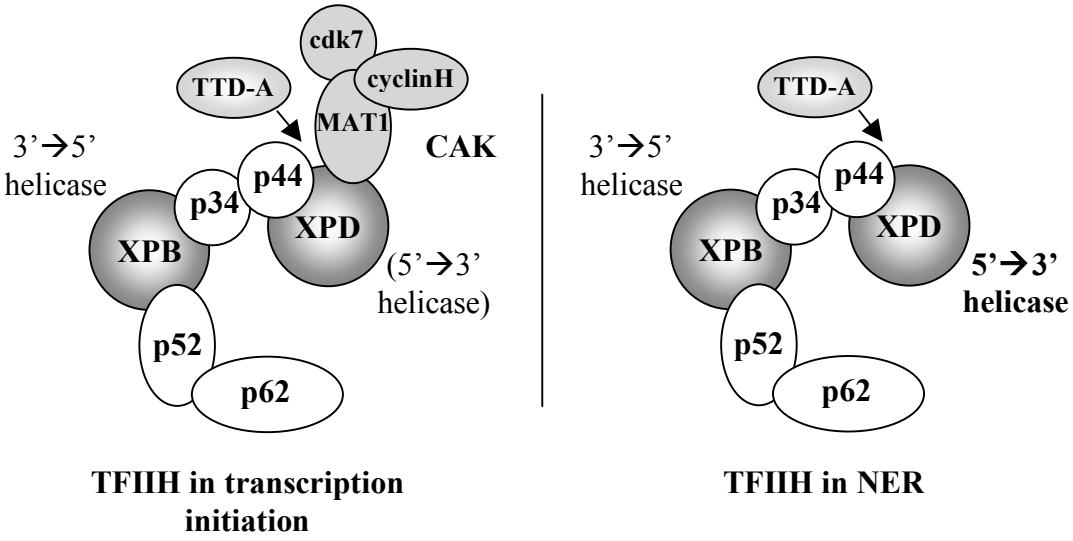


Figure 1.6: Subunit arrangement of TFIIH

Electron microscopic studies revealed a circular form of the human TFIIH holocomplex, which can accommodate dsDNA in its central hole. The CAK complex protrudes from the top of the complex. The two helicases are located on opposite sites. The exact positions of p34, p52 and p62 are not yet clear and are assumed on the basis of biochemical data (Schultz et al (2000)). TTD-A stabilizes the TFIIH complex by binding to p44 and XPD. The TFIIH core is not connected to the CAK subcomplex during NER and XPD is enzymatically active.

1.3.2.2 Mutations of human XPD and diseases

The enzymatic activity of XPD is only required for DNA repair. It is dispensable for transcription, but has a stimulatory effect probably because it brings both the TFIIH core and the CAK complex in close proximity (Reardon et al. (1996)). Therefore, the XPD gene is rather tolerant to mutations.

Most of the human mutations are located in the C-terminal region (Lehmann (2001)). The reasons for this could be that either mutations on the N-terminus are more devastating, and lead to early abortion of the human embryo or they are not clinically significant, so that respective patients are not registered. The C-terminus of XPD interacts with the p44 subunit of TFIIH and stimulates the helicase activity of XPD 10-fold (Coin et al. (1998)). Without stimulation by p44, XPD still exhibits a reduced helicase activity in analogy to the stimulation of UvrB by UvrA in *E. coli* (Seeley & Grossman (1990)). Furthermore, three of the conserved motifs are found in the C-terminus and mutations most likely lead to reduced or abolished ATPase activities. Mutations in the *xpd* gene can cause trichothiodystrophy (TTD) or *xeroderma pigmentosum* (XP) (Lehmann (2001)).

xeroderma pigmentosum (XP)

XP is generally characterised by increased photosensitivity with pigmentation anomalies and predisposition to skin cancer, as well as varying grades of neurological abnormalities induced by primary neuronal degeneration. Certain mutations in XPD induce one of the most severe XP phenotypes found in the XP complementation group, with the exception of XP-A. XP-D has an early onset from the age of 12 to the early twenties. Cells also show surprisingly high unscheduled repair synthesis with 20-45 % increase over normal cells (reviewed in Lehmann (2001)).

Trichothiodystrophy (TTD)

TTD is characterised by sulphur-deficient brittle hair, arising from a reduced number of cysteine-rich matrix proteins in the hair shaft. Varying grades of mental retardation, skeletal abnormalities and ichthyotic skin are described (Itin & Pittelkow (1990)). Some patients also show increased UV-sensitivity, but no pigmentation changes or skin cancers have been reported. Repair defects are very heterogeneous varying from

severe (10 % of normal repair) to no repair defect at all. The majority of patients have low levels of TFIIH (reviewed in Lehmann (2001)). It was first thought that this was due to a decreased stability of the multisubunit complex caused by an abolished or reduced interaction of XPD with p44. However, mutagenesis studies revealed that although most TTD cell lines showed reduced levels of TFIIH, one mutation that abolishes p44 interaction (R683W) led to nearly normal TFIIH levels. In addition, the mutation R112H does not affect p44 interaction but still showed reduced levels of the transcription factor (Dubaele et al. (2003)).

Another interesting discovery was made with mice harbouring a TTD mutation in the XPD gene. They exhibited a phenotype similar to humans with the characteristic brittle hair, but also showed premature aging with all the typical symptoms, such as reduced life span, osteoporosis, cutaneous aging and early greying (de Boer et al. (2002)). This was thought to be an effect of accumulation of DNA damage in TTD cells.

Cockayne syndrome (CS)

CS, another disease associated with XP genes, has not yet been described for XPD patients. There are only some rare cases of combined XP/CS phenotypes. CS is a disease, which affects transcription but, as was stated before, XPD is dispensable for transcription. It is likely that there may be CS mutations in XPD, which lead to clinically not relevant phenotypes and hence are not recognised. CS is a multisystem disorder with severe dwarfism, skeletal abnormalities, microcephaly and progressive neuronal dysfunctions arising from demyelination (Nance & Berry (1992)). Unlike XP, CS does not lead to pigmentation changes in the skin or skin cancer. Interestingly, although cells are sensitive to UV-killing, they are proficient in removing photoproducts (van Hoffen et al. (1993)).

Apart from inherited disorders like XP or TTD, XPD gene polymorphisms seem to be involved in the development of almost every kind of cancer. Specific polymorphisms have frequently been found for example in patients with lung (Benhamou & Sarasin (2005)), bladder (Schabath et al. (2005)) and breast (Debniak et al. (2006)) cancers.

1.3.3 Other related helicases and pathways

1.3.3.1 Human FancJ and the Fanconi anemia DNA repair pathway

Fanconi anemia (FA) is named after the Swiss paediatrician Guido Fanconi, who described the disease for the first time in 1927. It is a rare (1-5/ million) autosomal (apart from FancB: X-chromosomal) recessive inherited disorder, leading to bone marrow failure, increased cancer susceptibility and cellular hypersensitivity to DNA cross-linking agents with accumulation of chromosomal breakages (Taniguchi & D'Andrea (2006)). The disease shows an early onset between the ages of six and eight with an average survival age of 30 years. FA shows clinical heterogeneity and 25 - 40 % of the patients appear phenotypically normal (Auerbach et al. (2001)).

The FA pathway is a DNA repair mechanism involved in the maintenance of chromosomal stability and is induced after treatment with DNA damaging agents such as ionizing radiation, UV-light, DNA cross-linking agents or hydroxyurea (Auerbach et al. (2001)). Twelve complementation groups (Figure 1.7) have been described, eight of which form the FA-core complex (Fanc: A, B, C, E, F, G, L and M). This multisubunit complex is required for mono-ubiquitination of FancD2 (FancD2-Ub), which is the critical step for induction of the FA DNA repair pathway. FancD2-Ub moves to the chromatin fraction and accumulates at damaged sites. FancD2 has no known conserved motifs and its exact function is unknown. The FA DNA repair pathway becomes more important, as proteins involved in other inherited disorders were found to colocalise with Fanc-proteins. This includes ATM (*Ataxia telangiectasia*), BLM (Bloom syndrome), NBS1 (Nijmegen breakage syndrome) or ATR (Seckel syndrome) (reviewed in: Taniguchi & D'Andrea (2006)).

The Fanconi anemia pathway itself is not yet well understood, and scientists are busy adding missing pieces to the puzzle. One missing piece was discovered only recently, the helicase FancJ. FancJ was found to be identical to the gene product of the helicase called BRIP1 or BACH1 (*BR*CA1 *i*nteracting *p*rotein 1/ *B*RCA1 *a*ssociated *C*-terminal *h*elicase) (Litman et al. (2005); Levitus et al. (2005)). FancJ binds the C-terminal BRCT-domains of BRCA1. Phosphorylation of FancJ at serine 990 is critical for a stable BRCA1-FancJ interaction, which leads to binding of FancJ to the phosphopeptide

binding motif located at the interface between BRCTs (Cantor et al. (2001); Yu et al. (2003)). Germline mutations in BRCA1 and BRCA2 (FancD1; Howlett et al. (2002)) are associated with early onset of breast and ovary cancer (Yoshida & Miki (2004)). BRCA1 takes part in homologous recombination (HR) and the repair of DNA double stranded breaks. Subsequently, FancJ deficient cell lines are defective in HR, hypersensitive to mitomycin C (MMC) induced chromosomal instability and patients show early onset of breast cancer (similar to BRCA1). Furthermore, FancJ is not required for FancD2 mono-ubiquitination and it was suggested that it functions further downstream of the pathway (Litman et al. (2005)).

Like XPD, FancJ is a 5'→3' SF2 helicase with a conserved DEAH signature motif in the Walker B box (Cantor et al. (2004)). It is 25 % identical and shows 43 % similarity to hXPD. The mutation of K52 to arginine in the Walker A box led to a loss of ATPase and helicase activity and mutant cell lines exhibited a FA deficiency phenotype (Cantor et al. (2001); Bridge et al. (2005)). FancJ prefers to unwind forked duplexes and can also resolve D-Loop structures, which regularly occur during HR. No helicase activity was observed on holiday junctions (Gupta et al. (2005)).

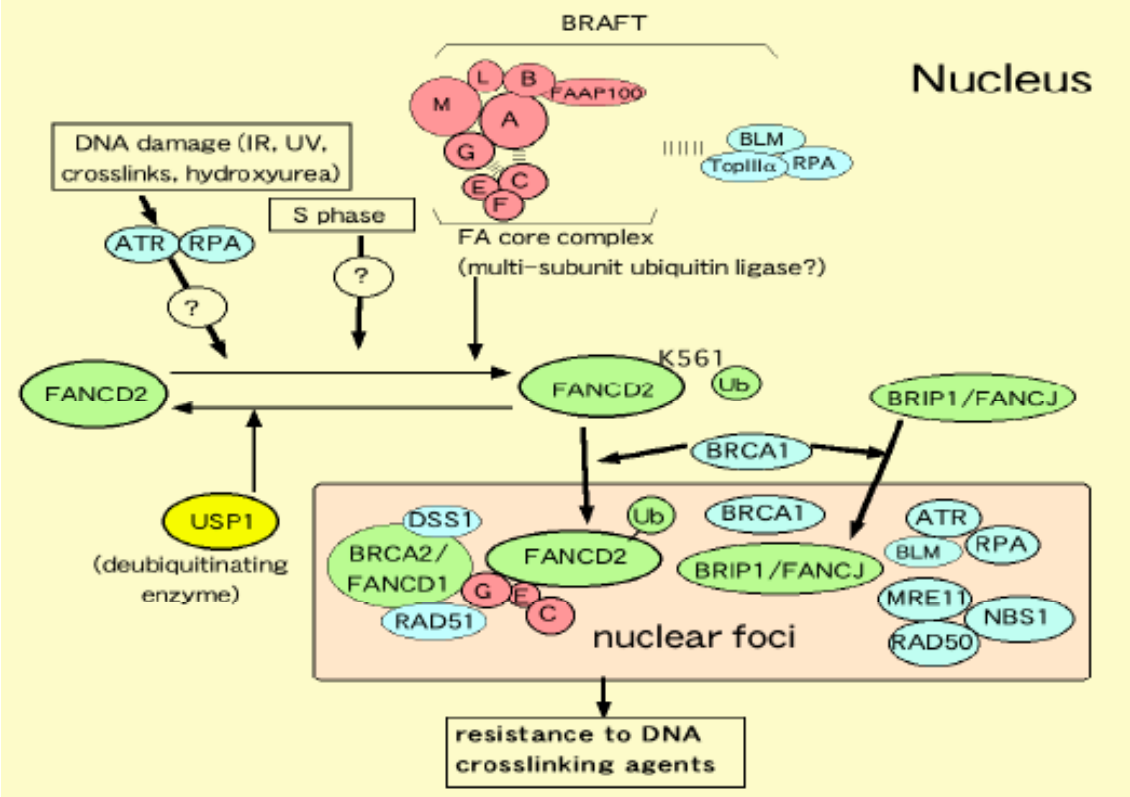


Figure 1.7: Current model of the Fanconi anemia DNA repair pathway

Upon DNA damage or during normal S-phase progression FancD2 is mono-ubiquitinated at K561 (FancD2-Ub). Several proteins are necessary for efficient mono-ubiquitination including ATR, RPA and the FA-core complex. FancD2 then localises in the nuclear foci where it interacts with FancD1 (BRCA2). The overall process of FA damage repair is not yet fully understood. USP1 is a negative regulator and deubiquitinates FancD2 (adapted from Taniguchi & D'Andrea (2006)).

1.3.3.2 Bacterial DinG

The *d*amage-*in*ducible helicase DinG is the closest bacterial homologue of eukaryotic and archaeal XPD helicases (Figure 1.8), although not participating in the same DNA repair pathway (Koonin (1993)). *DinG* is repressed by LexA, which suggested its involvement in the bacterial SOS response mechanism (Lewis et al. (1992)). Upregulation of *dinG* was observed after treatment with mitomycin C and nalidixic acid (Lewis et al. (1992), van Dyk et al. (2001)). Experiments using the cold-sensitive mutant *DinD68* suggested DinG may be a negative regulator of the bacterial SOS-response system (Yasuda et al. (1998)).

Similar to XPD, DinG is a SF2 helicase with 5'→3' directionality. The ATPase activity is stimulated by single stranded DNA and a k_{cat} of 24 sec^{-1} makes it one of the most efficient ATPases (Voloshin et al. (1993)).

Figure 1.8: Sequence alignment of archaeal XPD and bacterial DinG helicases

The figure shows a sequence alignment of the related helicases XPD from archaea and DinG from bacteria. Eco, *E. coli* DinG; Sty, *Salmonella typhimurium* DinG; Sac, *S. acidocaldarius* XPD; Pto, *Pyrococcus tokodaii* XPD. Highly conserved residues are white on black and similar residues are white on grey. The motifs are indicated on the top of the respective sequence. The cysteines which are involved in FeS cluster coordination are highlighted in yellow (for a better alignment of the FeS cluster domain in DinG helicases see supplementary figure 6 in Rudolf et al. (2006)). The C-terminal 21 (*E. coli*) and 19 residues (*S. typhimurium*), respectively, are omitted.

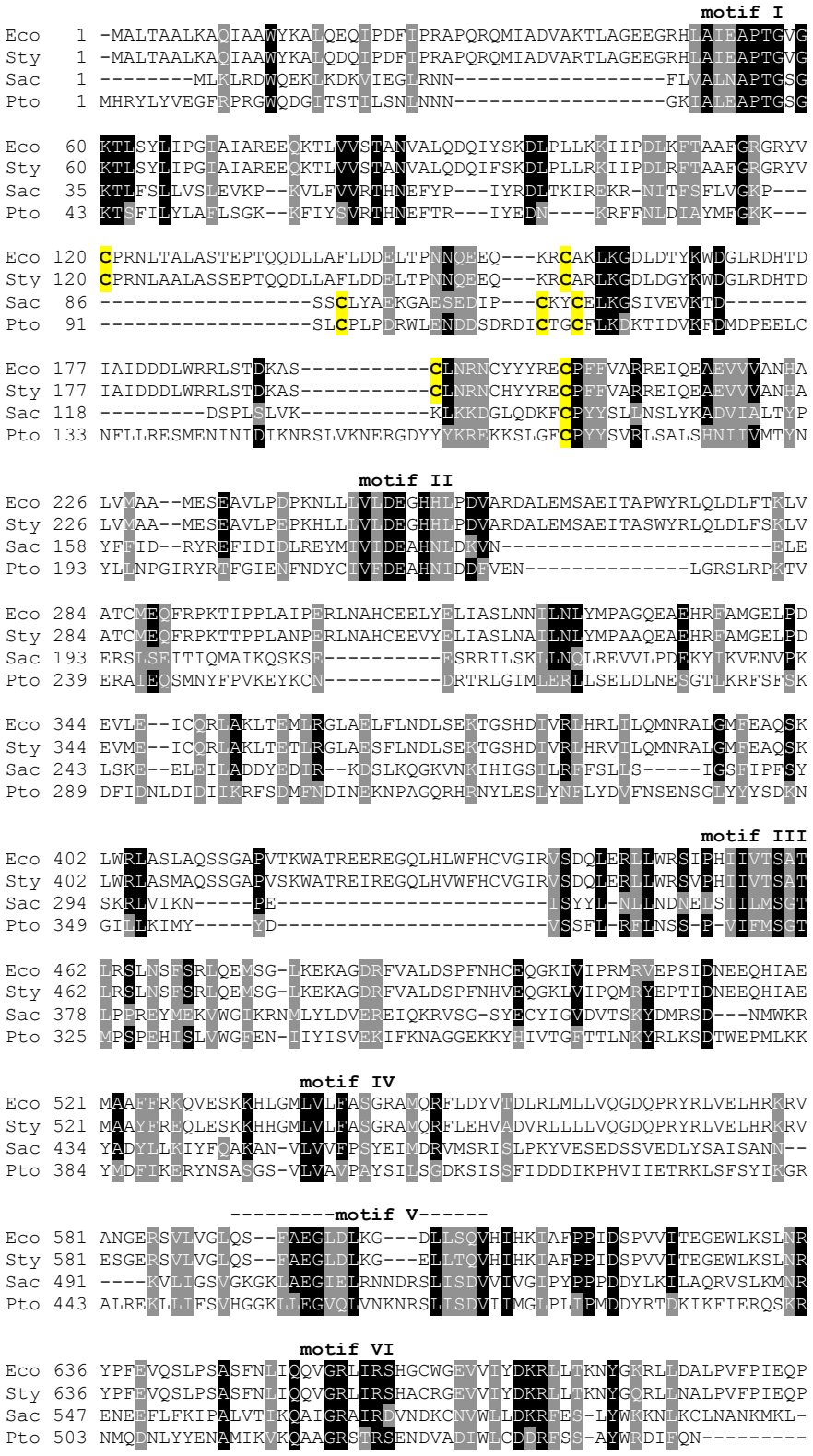


Figure 1.8

1.4 IRON-SULPHUR CLUSTER PROTEINS

Iron-sulphur (FeS) clusters are survivors of an anaerobic world, when reduced iron and sulphide were available en masse and FeS clusters could assemble spontaneously. With the creation of photosynthesis and the presence of oxygen, evolving aerobic organisms faced a new problem. The major problem with oxygen is that it readily oxidises transition metals, because they are excellent univalent electron donors. Ferric iron precipitates and it is no longer available for biological uptake. Furthermore, reactive oxygen species ($NO\bullet$, O_2^-, $OH\bullet$, H_2O_2) inside the cell are dangerous. In this context, free iron represents a specific problem for cells by reacting with H_2O_2 thus creating hydroxyl radicals ($OH\bullet$), which in turn damage proteins and DNA (Fenton Reaction). FeS cluster proteins are particularly sensitive to oxygen, as in most of the cases, oxidation leads to cluster destruction and subsequent enzyme inactivation.

The accumulation of oxygen in the atmosphere is assumed to have happened slowly, giving organisms the time to adapt. The general metabolism needed little *de novo* synthesis, the problem was the large number of FeS proteins, which took part in almost every metabolic pathway. The adjustments needed included acquisition of siderophores and other systems to capture free iron, iron transport and storage systems, regulatory systems that adjust the metabolism in response to the availability of iron and the presence of reactive oxygen species, repair systems and FeS cluster assembly machineries. FeS clusters present a particular problem in an aerobic world and cells have gone to a lot of effort to maintain them. Although, it seems that evolution is driving especially aerobic organisms to replace their FeS cluster proteins by more resistant or clusterfree isozymes; examples include the isozymes fumarase C or aconitase A. [reviewed in Imlay (2006)]

1.4.1 Structure of FeS clusters

The simplest form of a FeS cluster consists of one iron coordinated by four protein ligands. The most common ligand is cysteine, due to the fact that iron has a high affinity for thiolate groups, followed by histidines and aspartates. The most frequently found clusters in nature are [2Fe-2S], [3Fe-4S] and [4Fe-4S] (Figure 1.9). More complex

clusters derive from metal substitutions and fusions between the simpler clusters (Rees (2002)).

FeS clusters can exist in different oxidation states, which are calculated from the overall oxidation state of all irons (Fe^{2+} or Fe^{3+}) and all sulphurs (S^{2-}) involved, without the coordinating ligands. They are referred to as the oxidised or the reduced form of the FeS cluster. Apart from a few exceptions, the all-ferrous and all-ferric state of an FeS cluster does not occur in nature. In general, only two oxidation states have been observed for [2Fe-2S] (2+/ 1+), [4Fe-4S] low potential (2+/ 1+) and [4Fe-4S] high potential (3+/ 2+) FeS clusters. For some of the $[3Fe\text{-}4S]^{1+/\,0}$ clusters a third all-ferrous state is naturally occurring (2-) (Duff et al. (1996); Beinert et al. (1997)).

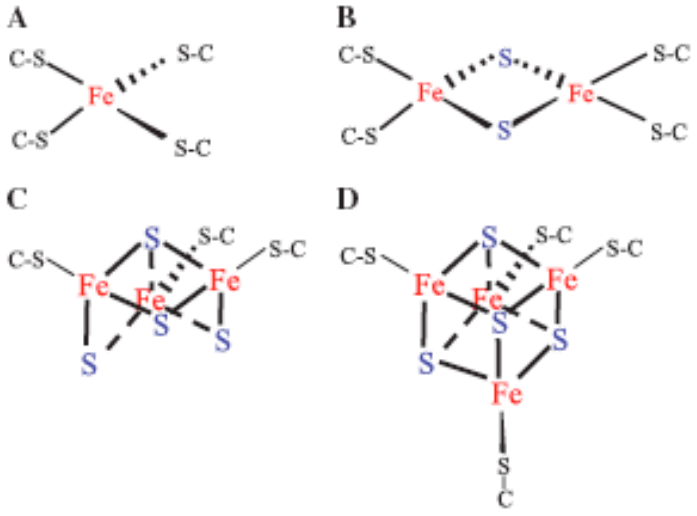


Figure 1.9: Structure of the most common FeS clusters

The most common FeS clusters found in nature are: (A) $[1Fe\text{-}0S]^{3+/2+}$, (B) $[2Fe\text{-}2S]^{2+/1+}$, (C) $[3Fe\text{-}4S]^{1+/0}$ and (D) $[4Fe\text{-}4S]^{2+/1+}$ or $[4Fe\text{-}4S]^{3+/2+}$. The irons (red) are coordinated by protein ligands, mostly cysteines (black) and bridged by inorganic sulphide (blue). Figure adapted from Imlay (2006).

1.4.2 Functions of FeS cluster proteins

FeS clusters cover a wide range of redox potentials from around -0.6 to +0.45 V (Capozzi et al (1998)). This makes them excellent electron donors or acceptors. Therefore, FeS clusters are used as a bridge between redox couples which are physically separated. Some enzymes even contain several FeS clusters, which serve as an "electric wire" through the enzyme. One example is fumarate reductase, which contains three FeS clusters with a typical spacing of 10-14 Å. The electrons are transferred from the membrane bound menaquinone to cytosolic fumarate by electron hopping between the clusters (Iverson et al. (1999); Page et al. (2003)).

FeS clusters of nitrogenase – An example for electron transfer

Another example is nitrogenase, a large protein complex used by diazotrophic microorganisms to transform atmospheric dinitrogen into ammonia. Microorganisms fix up to 2 x 10^{11} kg N_2/ year and they are highly important for the maintenance of the natural nitrogen cycle. The nitrogenase contains 19 iron and 20 sulphurs and under nitrogen limiting conditions it makes up to 10 % of the total soluble protein of *Azotobacter vinelandii* (Zheng et al. (1998)). To satisfy this high demand of iron and sulphur, the nitrogenase has a specialised FeS cluster assembly machinery (see section 1.4.3). To split the dinitrogen triple bond (960 kJ/ mol) microorganisms require 16 ATP molecules. The homodimeric Fe-protein of the nitrogenase binds a [4Fe-4S] cluster between its interfaces. Upon binding of 2 ATPs, it forms a complex with the MoFe-protein of nitrogenase and delivers 2 electrons at once to the P-cluster of the MoFe-protein. This double electron transfer is unique and changes the oxidation state of the [4Fe-4S] cluster of the Fe-protein from +2 to 0. Subsequently ATP is hydrolysed, which refills the electron "lack" of the Fe-protein. The P-cluster (8 Fe:7 S) transfers the 2 electrons to the M-cluster. The M-cluster (7 Fe:9 S:1 Mo) transfers the electrons in the final step to split the dinitrogen bond (Muhlenhoff & Lill (2000); Seefeld et al. (2004))

There are a huge number of FeS proteins with electron transfer functions. Most famous are the FeS proteins, which participate in the electron transfer during photosynthesis or respiration. But FeS cluster proteins also fulfil other roles, which do not involve an electron transfer step.

FeS clusters in sensing and regulation of gene expression

One of these enzymes is aconitase. There are three forms of aconitase ([4Fe-4S], [3Fe-4S] and the apoprotein) in two different locations (cytosol: c-aconitase; mitochondria: m-aconitase). C- and m-aconitase have 30 % identity in their amino acid sequence. Depending on how many irons are bound to the enzyme, the function changes. The m-aconitase in the [4Fe-4S] form takes part in the tricarbolic acid cycle (TCC) and one of the 4 Fe is probably directly coordinated by the substrate citrate, which is converted into isocitrate. The enzyme is only active in the 4 Fe state, the 3 Fe protein represents an inactive enzyme. Citrate is a strong iron chelator, so it is possible that the fourth iron is delivered to the enzyme by the substrate itself. C-aconitase in its 4 Fe form also takes part in the TCC, the 3 Fe form is inactive and the apoprotein is generated under iron-limiting conditions. As apoprotein this enzyme is known as iron regulatory protein 1 (IRP1) and it binds to iron responsible elements (IREs) in the untranslated region of several m-RNAs. IREs are m-RNA secondary loop structures, which are highly conserved. Depending on whether IRP1 binds to an IRE in the 3' or 5' untranslated region, it stabilises or destabilises transcripts and is involved in cellular iron homeostasis, which leads for example to an increased iron uptake. This is one example where the presence or absence of a FeS cluster is used for sensing intracellular iron levels and for the regulation of protein expression (reviewed in: Beinert et al. (1996); Beinert & Kiley (1999)).

In other proteins, such as SoxR, the oxidation state of the cluster is used to sense reactive oxygen species (ROS) and induces the transcription of respective genes (only in the oxidised form of the FeS cluster), whose products are necessary for the defence against ROS (Gaudu & Weiss (1996)).

Structural roles of FeS clusters

In DNA repair enzymes like MutY and EndoIII - both contain a [4Fe-4S] cluster - the cluster stabilises a loop structure known as the iron-sulphur-cluster loop (FCL), which interacts with DNA. In the absence of the FeS cluster, these enzymes are inactive due to an inability to bind DNA (Thayer et al. (1995); Chepanoske et al. (2000)). In addition to its structural function, the FeS cluster in these glycosylases was also suggested to have a

catalytical role, by facilitating electron charge transport on the DNA, leading to local accumulation of these repair enzymes near to damaged sites (Boal et al. (2005)).

1.4.3 The FeS cluster assembly machinery

Under anaerobic conditions and in the presence of reduced iron and sulphide, FeS clusters assemble spontaneously. In a biological system spontaneous assembly does not occur and three distinct biogenesis systems have been described for bacteria. The Nif (*ni*trogen-*f*ixation) system is responsible for formation and assembly of the FeS clusters in the enzyme nitrogenase (section 1.4.2). All nitrogen-fixing bacteria possess the Nif system (Rees & Howard (2000)). The Isc (*i*ron-*s*ulphur *c*luster assembly) system is more a general machinery for FeS cluster assembly and is responsible for biosynthesis of the majority of FeS clusters of cellular FeS proteins (Zheng et al. (1998)). The third system Suf (*sulf*ur-mobilisation) is active under iron-limiting and oxidative stress conditions. The genes of the respective systems are coexpressed and arranged in clusters. All systems are essential and deletions are lethal or lead to severe growth defects. In general these systems are interchangeable and highly conserved throughout the three domains of life. [reviewed in: Muhlenhoff & Lill (2000)]

<u>The general mechanism of action of FeS cluster assembly</u>

The general mechanism of action for all three systems is as follows: A pyridoxal-phosphate dependent enzyme activates sulphur from L-cysteine. The resulting enzyme bound cysteine-persulphide is the ultimate sulphur source for FeS cluster formation. The sulphur is delivered to a protein, which serves as a molecular scaffold for the formation of the FeS cluster precursor. The delivery of sulphur is followed by iron incorporation. The correct mechanism is not yet fully understood. For most of the factors found to be essential for FeS cluster formation and assembly, an exact function has not yet been described. It is also unclear how the FeS cluster precursor is transferred to the respective apoproteins.

Peculiarities of FeS cluster assembly in eukaryotes

In eukaryotes, the location of the FeS cluster assembly machinery is in the mitochondria, yet FeS cluster proteins are also abundant in the cytosol. It is still unclear, how cytosolic FeS cluster proteins are assembled in eukarya. It was suggested that a FeS cluster precursor is somehow transferred from the mitochondrial matrix into the cytosol. The yeast ATM1 (human: hABC7) ABC transporter seems to be involved in this process, however, the modality of the transfer and how the FeS cluster is incorporated into the apoprotein in the cytosol is still unknown (Kispal et al. (1999)).

1.4.4 Iron metabolism and human diseases

Iron is an essential element and the majority of the iron in humans is used for incorporation in erythrocytes and FeS cluster containing proteins. Due to the toxicity of iron for the aerobic metabolism, tight regulation of iron homeostasis is necessary. Most common diseases in human are caused by mutations in factors that regulate iron uptake (e.g. defective hepcidin or transferring receptors). This leads to iron accumulation in specific tissues (Fe-overload) and causes hemachromatosis. Atransferrinemia, a rare inherited disease, is caused by the absence of the Fe-transport protein transferrin in the patients blood. This leads to anemia and iron deposition in the heart and liver, causing heart and liver failures. Aceruloplasminemia is caused by defective ceruloplasmin, a ferroxidase, and leads to accumulation of iron especially in neural tissue. Following lipid peroxidation, patients suffer from progressive neurodegeneration (reviewed in: Hentze et al. (2004)). The most common inherited iron-related disorder is Friedreich's ataxia, caused by a triplet expansion in intron 1 of the *frataxin* gene. The severity of the phenotype caused by Friedreich's ataxia correlates with the length of this triplet expansion. Frataxin is located in the mitochondria and the related disorder is characterised by accumulation of iron within the mitochondria and FeS cluster protein deficiency. The disease causes slow progressive degeneration of the nervous system leading to an inability of voluntary muscle movement (ataxia) and heart conditions (Rouault & Tong (2005)). The exact function of frataxin, however, is not yet clear.

This thesis will describe the characterisation of the archaeal homologue of human XPD. The protein was isolated from the crenarchaeon *S. acidocaldarius* and is shown to be a monomeric SF2 ATP dependent 5'→ 3' DNA helicase. *S. acidocaldarius* XPD was found to coordinate an essential FeS cluster. Bioinformatic analysis showed that the FeS cluster domain was also present in human XPD and other related human helicases such as FancJ and the bacterial homologue DinG.

2 MATERIALS AND METHODS

2.1 ISOLATION OF GENOMIC DNA FROM *SULFOLOBUS*

The isolation of genomic DNA from *Sulfolobus solfataricus* (Sso) and *Sulfolobus acidocaldarius* (Sac) cells (frozen cell pellet, stored at -80 ˚C) was carried out using the QIAGEN DNeasy ® Tissue Kit (50) following the protocol for Animal Tissues. For polymerase chain reaction (PCR), a 1:10 dilution of genomic DNA was used.

2.2 CLONING AND PROTEIN EXPRESSION

2.2.1 Cloning procedures and vectors

The gene sequence of the *S. acidocaldarius xpd* (SacXPD) was kindly provided by Roger Garret (Copenhagen), since at that time the genome sequence was not yet available. The gene was amplified from genomic DNA using proof-reading KOD HiFi polymerase (Novagen) and the following primers:

5'-CGTCGGATCCCCATGGAACTTAGAGAGTGGCAAAGC-3'

5'-CCGGGGATCCGTCGACTTACAGTTTAATTTTCCTTG-3'

The PCR product was cloned into vector pET28c (Novagen) using *Bam*H1/*Nco*I recognition sites for expression of untagged recombinant protein. The insertion was fully sequenced.

The truncated his-tagged form of SacXPD (E72 → D162) was obtained as described above using the following primers:

5'-CGTCGGATCCCCATGGCAGAGAAGAGAAATATCACTTTTTCTTTTCTTGTAG G-3'

5'-CCGGGGATCCGTCGACTTAATCTATAAAGAAATATGGATAAGTTAGGGC-3'

The PCR product was cloned into vector pET28c (Novagen) using *Sal*1/*Nco*I recognition sites for expression of his-tagged recombinant protein.

The *Saccharomyces cerevisiae* (strain BY4741a) RAD3 gene was amplified from genomic DNA, kindly provided by Dr P. Dennison (CBMS, St Andrews), using Long PCR Enzyme Mix (Fermentas) and the following primers:

5'-GAATCGGCCGGCATGCAAAGTGGGATTTTTGGTACCGTCTCTGGG-3'

5'-GCCGGCATGCAGATTATTACAAAAAAAGTGGG-3'

The PCR product was gel purified and cloned into the TOPO 2.1 vector using the TOPO TA Cloning system (Invitrogen). The gene was excised from the TOPO vector using *Sph*I and *Xho*I, the Rad3 fragment gel purified and subcloned into the *Sph*I/*Sal*I cloning sites of the yeast shuttle vectors YCplac33 and YCplac111 (donated by Prof Michael J. Stark, Dundee). The insertion was fully sequenced.

The three PCNA subunits from *S. solfataricus* were amplified from genomic DNA using proof-reading KOD HiFi polymerase and the following primers:

PCNA1: 5'-GGCCCATGGTTAAGATTGTTTACCCTAATGC-3'
5'-GGCGTCGACTTATAACCTTGGCGCTATCCAAAAGATC-3'
PCNA2: 5'-GGCCCATATGATGAAAGCTAAGGTAATTGACG-3'
5'-GGCCTCGAGTTAGTCTGCCCTTGGTGCAATGTAAAAGTCC-3'
PCNA3: 5'-GGCCCCATGGCATATCTTAAATCTTTTGAAAGG-3'
5'-GGCCGTCGACTTAAACTTTTGGAGCTAATAAATAAG-3'

The PCR products were cloned into vector pETDuetI (Novagen) using *Sal*1/*Nco*I (PCNA1 and PCNA3) and *Nde*I/*Xho*I (PCNA2) recognition sites, respectively, for expression of untagged recombinant protein. The insertions were fully sequenced.

The sequence of the C-terminal Helix-hairpin-Helix (HhH) DNA interaction motif of SsoXPF was amplified from genomic DNA using proof-reading KOD HiFi polymerase and the following primers:

5'-CGTCGGATCCCCCATGGCTAAATTGGAAAGTGTTTCTG-3'
5'-CCGGGGATCCGTCGACCTAAAGGAAATCAAATAAAG-3'

The PCR product was cloned into vector pET28c (Novagen) using *Sal*1/*Nco*I recognition sites for expression of his-tagged recombinant protein. The insertion was fully sequenced.

The vectors containing the his-tagged constructs of the SsoPCNA subunits (pET33b-PCNA1, pET30a-PCNA2, pET30a-PCNA3) and the untagged form of SsoXPF (in pET19b) as well as the untagged form of ApeXPF (in pET28c) were provided by J. A. Roberts.

2.2.2 Site-directed mutagenesis

Mutations the the SsoXPF, SacXPD and yeast RAD3 (in YCplac111) genes were generated using the Stratagene XL QuikChange Mutagenesis Kit and the primers indicated in Appendix 1 (Table 1). The mutations were confirmed by DNA sequencing. Amino acid substitutions were confirmed by mass spectrometry on pure protein preparations.

2.2.3 Protein expression

All proteins were expressed in *Escherichia coli* BL21 Rosetta cells (Novagen) and cultures grown in Luria Bertani (LB) medium containing the appropriate antibiotics (ampicillin: 100 µg/ mL or kanamycin: 35 µg/ mL). For expression of full-length forms of SacXPD, cultures were grown at 37 ˚C to optical densities (ODs) of 1.0 to 1.2. After induction of protein expression with IPTG (of 0.2 mM), cultures were incubated overnight (~14 hours) at 28 ˚C with shaking at 180 rpm. Cultures for expression of the truncated form of SsoXPD, the HhH motif of SsoXPF, full-length forms of SsoXPF and the his-tagged and untagged SsoPCNA subunits were grown at 37 ˚C until the OD_{600} was between 0.7 - 0.9. Protein expression was induced with IPTG (0.2 mM) and cultures were grown at 37 ˚C with shaking for further three hours.

2.3 PROTEIN PURIFICATION

2.3.1 Purification of untagged recombinant SacXPD

The bacterial cell pellet from 2 L of culture was resuspended in 30 mL buffer A (20 mM MES, pH 6.0, 200 mM NaCl, 1 mM EDTA, 1 mM DTT) containing benzamidine (1 mM) and sonicated 3 x 4 min (at an amplitude of 15 microns) with cooling on ice. The lysate was centrifuged (48000 x g, 20 min, 4 ˚C), heat-treated (65 ˚C, 20 min) and centrifuged for a further 20 min. The extract was filtered (Aerodisc® 0.45 µm syringe filter), loaded onto a heparin column (HiTrap™ 5 mL Heparin HP, Amersham Biosciences), and eluted with a linear gradient of sodium chloride (1 M). Fractions containing the SacXPD protein were identified by SDS-PAGE (Invitrogen), pooled and purified to homogeneity using a HiLoad® 26/60 Superdex® 200 size exclusion column (Amersham Biosciences) equilibrated with Gel Filtration buffer (20 mM Tris, pH 7.8, 200 mM NaCl, 1

mM EDTA, 1 mM DTT). Identities of mutant and wildtype proteins were confirmed by mass spectrometry.

2.3.2 Purification of His-tagged E72-D162 truncated SacXPD

Bacterial cell lysis was achieved as described in section 2.3.1 using binding buffer with 20 mM sodium phosphate, pH 7.4, 500 mM NaCl and 1 mM benzamidine. Cell lysate was not heat-treated and loaded onto a Nickel chelating column (HiTrap™ 5 mL Chelating HP, Amersham Biosciences) after filtration through an Aerodisc® 0.45 µm syringe filter, and eluted with a linear gradient of imidazole (0.5 M). Fractions containing the SsoXPF-HhH protein were identified by SDS-PAGE (Invitrogen), pooled and purified to homogeneity using a HiLoad® 26/60 Superdex® 200 size exclusion column (Amersham Biosciences) equilibrated with Gel Filtration buffer. The protein was confirmed by mass spectrometry.

2.3.3 Purification of his-tagged SsoXPF-HhH motif

Bacterial cell lysis was achieved as described in section 2.3.1 using binding buffer with 20 mM sodium phosphate, pH 7.4, 500 mM NaCl and 1 mM benzamidine. Soluble cell extract was loaded onto a Nickel chelating column (HiTrap™ 5 mL Chelating HP, Amersham Biosciences) and SsoXPF-HhH protein was purified to homogeneity as described in section 2.3.2.

2.3.4 Purification of untagged recombinant SsoPCNA subunits

The bacterial cell pellet of 2 L culture of each expressed subunit of PCNA was resuspended in 30 mL buffer A (20 mM Tris, pH 7.8, 1 mM EDTA, 1 mM DTT) and 1 mM benzamidine and sonicated 3 x 4 min with cooling on ice. The lysate was centrifuged (48000 x g, 20 min, 4 ˚C). Heat-treatment (65 ˚C, 20 min) took place only for subunit 2, PCNA1 and PCNA 3 were not heat stable. Lysate containing PCNA2 was centrifuged for a further 20 min following heat-treatment. All extracts were filtered (Aerodisc® 0.45 µm syringe filter) before being loaded onto a sepharose anion exchange column (HiTrap™ Q HP 5 mL, Amersham Biosciences). Proteins were eluted with a linear gradient of sodium chloride (1 M) and fractions containing PCNA protein were identified by SDS-PAGE

(Invitrogen), pooled and purified to homogeneity using a HiLoad® 26/60 Superdex® 200 size exclusion column (Amersham Biosciences) equilibrated with buffer (30 mM Tris, pH 7.8, 200 mM NaCl, 1 mM EDTA, 1 mM DTT). Proteins were confirmed by mass spectrometry. Equivalent amounts of each subunit were mixed together and incubated at 4 ˚C on a rotary mixer for approximately one hour to form the heterotrimer. The heterotrimer was separated from dimers and monomers using size exclusion chromatography as described above.

2.3.5 Purification of untagged SsoXPF, ApeXPF and his-tagged SsoPCNA subunits

These proteins, including the SsoXPF mutant proteins, were purified as described in the PhD thesis of J. A. Roberts (2004).

2.3.6 Determination of protein concentrations

Protein concentrations were either estimated from measuring the absorbance at 280 nm (A_{280}) or using Bradford Reagent (Bio-Rad). Theoretical molar extinction coefficients were obtained from ProtParam analysis (http://ca.expasy.org/tools/protparam.html). Bradford reagent was used as described in the manufacturer's manual and protein concentrations were calculated using a BSA standard curve.

2.4 TRYPTIC DIGESTION OF SACXPD

A total reaction volume of 20 µL contained 6.5 µg recombinant SacXPD, 2.5 mM KCl and different concentrations of trypsin (200, 100, 50, 25, 12.5, 6.25 ng). Reactions were incubated for 90 min at 30 ˚C and stopped by addition of AEBSF protease inhibitor to a final concentration of 10mM. Fragments were separated on SDS-PAGE and analysed by MALDI-ToF mass spectrometry.

For N-terminal protein sequencing, reactions were set up as described above with 25 µg SacXPD and 200 ng trypsin. Peptide bands were transferred onto a PVDF membrane in transfer buffer (10 mM CAPS, pH 11.0, 10 % methanol) and visualized by using amido black protein stain. Major bands were cut from the membrane and protein fragments N-terminal sequenced using the Edman-Reaction.

2.5 GLUTARALDEHYDE CROSS-LINKING OF PROTEINS

Reactions containing 25 µg protein in 50 mM sodium phosphate buffer, pH 7.6, were incubated at room temperature (RT) for 10 min. Glutaraldehyde was added to a final concentration of 1 % and reactions were quenched after two minutes by addition of $NaBH_4$ (100 mM). Proteins were TCA-DOC precipitated and analysed by SDS-PAGE.

2.6 WESTERN BLOT

Sheep antiserum containing polyclonal antibodies against SacXPD was purchased from Alba Bioscience. Protein samples were separated on SDS-PAGE and transferred onto Hybond Nitrocellulose membrane (Amersham Biosciences) soaked in transfer buffer (25 mM Tris, pH 8.3, 250 mM glycine, 15 % methanol). Protein was transferred under semi-dry conditions for 25 min at 25 V/ 400 mA. The membrane was blocked one hour (5 % milk, 0.1 % Tween 20 in PBS buffer) at room temperature and incubated with the primary antibody (1:20000) overnight at 4 ˚C. The membrane was washed with blocking buffer (3x 10 min, RT) and the secondary antibody added (1:200000; Anti-Goat, Pierce ImmunoPure® Antibody) and incubated for one hour (RT). The membrane was washed in blocking buffer (3x 10 min) and ultrapure water (3x 5 min). Chemiluminescence detection was initiated by using SuperSignal® West Pico Chemiluminescent Substrate or Femto Maximum Sensitivity Substrate Kits (Pierce). The signal was visualized using the Fuji Luminescent Image Analyzer LAS-1000 and Image Reader software.

2.7 BATHOPHENANTHROLINE ASSAY

The number of irons bound to SacXPD were quantified using a 100 µL aliquot of 50 µM protein in gel filtration buffer, mixed with 30 µL concentrated HCl and heat treated for 15 min at 100 ˚C. The control reaction contained buffer only. After 2 min centrifugation at 13000 x g, 100 µL of the reaction was transferred into a clean sterile 2 mL tube and 1.3 mL 500 mM Tris-HCl, pH 8.5 was added. Freshly prepared 5 % ascorbic acid (100 µL) and 0.1 % bathophenanthroline (400 µL) were then added to the reaction. Samples were thoroughly mixed after each addition and reactions were incubated for 1 h at room temperature. The

absorbance of the reaction was measured at 535 nm against the control reaction. Iron bound to XPD was calculated using the molar extinction coefficient of 22369 mol^{-1} cm^{-1} (Pieroni et al. (2001)).

2.7.1 Iron-sulfur cluster reconstitution

To reconstitute the FeS cluster, all buffers and solutions were purged with argon for at least 30 min. Tubes and other containers were as well purged with argon before usage. The protein was dialyzed against 20 mM Tris-HCl, pH 7.8 and 200 mM NaCl to remove the EDTA present in the storage buffer. The protein solution was divided into two 200 µL samples and freshly prepared DTT was added to a final concentration of 1 mM. The protein concentration was determined (127 µM) by measuring the absorbance at 280 nm (ND-1000 Spectrophotometer, NanoDrop Technologies). Protein was transferred to anaerobic cuvettes and purged with argon. 10 mM $Fe(II)(NH_4)_2(SO_4)_2$ was made up in anaerobic buffer (20 mM Tris-HCl, pH 7.8 and 200 mM NaCl) and further incubated in argon atmosphere.

Three reactions were set up: (1) Control containing buffer only, (2) control reaction containing 200 µL protein solution with addition of 40 µL buffer, (3) sample reaction containing 200 µL protein solution with 40 µL 10 mM $Fe(II)(NH_4)_2(SO_4)_2$. Every step was followed by absorbance readings between 600 and 300 nm. 4 µL of 5 mM dithionite (DT) was added 3 times subsequently to reaction (3) with incubations of 10 min before absorbance readings. This was done to ensure that all irons were present in the reduced form. After the third reading, the solution was incubated another 30 min at room temperature and a final spectra was taken. Microspin G-50 columns were prepared, placed in 15 mL Greiner tubes and incubated in argon atmosphere for 30 min. 100 µL of the protein solution was applied to each column and centrifuged for 2 min (800x g) to remove excess iron. Sodium chloride was added to a final concentration of 200 mM to the recovered samples to stabilize the protein. Absorbance measurements were taken on the ND-1000 spectrophotometer and the protein concentration determined. Numbers of irons bound to the protein were determined using the bathophenanthroline method (2.7).

2.8 SPECTROSCOPIC METHODS

2.8.1 Circular dichroism (CD)

Protein was dialysed into CD buffer (10 mM Tris, pH 7.8, 200 mM NaF, 0.5 mM EDTA) overnight at 4 ˚C and protein concentrations adjusted. Spectra were taken on a JASCO J-810 Spectralpolarimeter using Spectra Manager Software. For near UV readings, the protein solutions were adjusted to a concentration of 1 mg/mL. The spectra were analysed between 450 nm and 260 nm at a scan rate of 20 nm/min using a cuvette with a pathlength of 0.5 cm. For far UV measurements, the protein solutions were adjusted to concentrations of 0.5 mg/mL. The spectra were taken between 260 nm and 180 nm at a scan rate of 50 nm/min using a cuvette with a pathlength of 0.02 cm. Spectra for each protein were taken six times, averaged and the blank (CD buffer) subtracted. Spectra were analysed using the DichroWeb (Lobley et al. (2002)) server (http://www.cryst.bbk.ac.uk/cdweb/html/home.html).

2.8.2 Electron Paramagnetic Resonance (EPR)

For EPR spectroscopy, XPD (95 μM) was treated with either 1 mM potassium ferricyanide or 1 mM dithionite. The samples were incubated on ice for 5 min in 3 mm quartz EPR tubes before being quick frozen and stored in liquid nitrogen until usage. Spectra were obtained on a Bruker X-Band (cw) EPR Spectroscope at 30 K with the following settings: modulation frequency, 100 KHz; modulation amplitude, 0.32 mT; microwave frequency 9.49 GHz; microwave power 0.2 mW. A liquid helium transfer system was employed to regulate the temperature of the samples.

2.8.3 Inductively Coupled Plasma – Optical Emission Spectroscopy (ICP-OES)

The elementary analysis of the Fe-S cluster bound to SacXPD was done on a Perkin-Elmer Optima 5300 ICP-OES at the Chemistry Department of the University of Edinburgh. Protein was purified as described above (2.3.1) with the following changes: Gel Filtration buffer was free of DTT and the NaCl concentration was lowered to 100 mM. Standards used for standard curves contained 0.5, 1, 2, 3, 5 and 10 ppm of iron and sulfur. The Fe:S ratio of 10 μM SacXPD in Gel Filtration buffer was measured by comparing emission intensities with the standard curves. The blank contained Gel Filtration buffer alone.

2.9 CATALYTIC ASSAYS

2.9.1 DNA substrates

The oligonucleotides (Operon) used for constructing DNA substrates were 5'-[^{32}P]-end-labelled and assembled into various structures (see Appendix 1, Table 2) by slow cooling from 85 °C to room temperature for a minimum of 3 hours to overnight. Substrates were purified on a native 8 % acrylamide:TBE gel, eluted overnight at 37 ˚C into buffer TE (10 mM Tris, pH 8.0, 1 mM EDTA) and ethanol precipitated. Constructs were resuspended in ultrapure water and stored at -20 ˚C.

2.9.2 SsoXPF endonuclease assay

Endonuclease reactions were carried out in buffer containing 30 mM HEPES, pH 7.6, 5 % glycerol, 40 mM KCl, 1 mM EDTA, 0.1 mg/mL BSA, 0.1 mg/mL calf thymus DNA and 80 nM DNA substrate. Samples were incubated at room temperature for 5 min with 1 μM PCNA heterotrimer, before 1 μM XPF protein was added. The samples were incubated further for 2 min at 55 ˚C and each reaction was initiated by adding $MgCl_2$ to a final concentration of 10 mM. Aliquotes of 5 μL were taken at selected time points and added to 90 μL chilled stop solution (10 mM Tris, pH 8.0, 5 mM EDTA) to terminate the reaction. The DNA was ethanol precipitated and resuspended in formamide loading dye (95 % formamide, 0.25 % bromphenol blue, 0.1 % xylene cyanol). Samples were heated at 95 ˚C for 5 min and loaded onto a denaturating (7 M Urea) 12 % polyacrylamide gel and electrophoresed at 50 ˚C, 90 W for 30-45 min. Gels were exposed to phosphoscreen, imaged and quantified using Fuji Image Gauge software.

2.9.3 SacXPD helicase assay

Helicase assays were carried out in 20 mM MES, pH 6.5, 1 mM DTT, 0.1 mg/mL BSA, 10 nM ^{32}P labelled DNA-substrate and 200 nM protein. The reaction mix was incubated at 45 ˚C for 1 min and the reaction started with the addition of a 1:1 mix of $MgCl_2$ and ATP to a final concentration of 1 mM (for reactions without ATP only 1 mM $MgCl_2$ was added). At specific time points 10 μL samples were taken and immediately added to 20 μL of chilled stop solution (10 mM Tris-HCl, pH 8, 5 mM EDTA, 10 μM competitor DNA (unlabelled DNA oligo corresponding to DNA oligo that was labelled),

0.5 % SDS, 1 mg/mL proteinase K). The samples were incubated 15 min at room temperature to allow proteinase K digestion. The samples were mixed with 6 µL loading buffer (30 % ficoll, 0.25 % bromphenol blue, 0.1 % xylene cyanol) and 10 µL sample loaded onto a native 12 % acryl amide:TBE gel. Samples were electrophoresed for 2 hours at 130 V. Gels were exposed to phosphoscreen, imaged and quantified using Fuji Image Gauge software.

Control reaction contained the reaction mix without protein and 1 mM $MgCl_2$:ATP. A 10 µL sample was immediately added to 20 µL stop solution. The dsDNA control was left at room temperature during the entire procedure. The ssDNA control was heat treated at 95 ˚C for 2 min and cooled on ice. The third control, an indicator for the stability of the DNA substrate, was incubated at 45 ˚C for 5 min.

2.9.4 Quantification and data processing of helicase assays

Single stranded (ss; unwound) and double stranded DNA (ds) fractions were quantified using Image Gauge software (Fuji). The DNA unwound (%) was calculated: (ss/ total) × 100 and plotted against the time (s) using KaleidaGraph software.

2.10 STREPTAVIDIN DISPLACEMENT ASSAY

Assays were carried out either in buffer A (20 mM MES, pH 6.5, 1 mM DTT, 50 mM KCl) or buffer B (20 mM MES, pH 6.5, 1 mM DTT). If not otherwise stated, all concentrations are final concentrations. The reaction mix (buffer A or B, 0.1 mg/ mL BSA, 1 µM Streptavidin, 10 nM 5' ^{32}P-labeled DNA substrate containing a biotin molecule on the 3' end) was incubated at 55 ˚C for 3 min. 600 nM XPD protein was added followed by a further incubation for 3 min. Reactions were initiated by addition of a mixture of 1 mM ATP or ADP, 1 mM $MgCl_2$, 30 µM competitor DNA, and 10 µM biotin. Samples (5 µL) were taken at specific time points (1, 3, 5, 10 min) and transferred into 10 µL chilled EDTA (350 mM). Samples were mixed with 3 µL ficoll loading dye and 10 µL loaded onto a 12 % native acrylamide gel and run for 4 hours. [Byrd & Raney (2004)]

2.11 SacXPD ATPase Hydrolysis Reactions

2.11.1 Standard reactions

The Malachite Green Phosphate Assay Kit (BioAssay Systems) was used for detection of free phosphates after ATP hydrolysis. The standard curve was made with increasing concentrations of phosphate as described in the manufacture's manual. Standard reactions contained 20 mM MES, pH 6.5, 1 mM DTT, 0.1 mg/mL BSA, 5 nM DNA (oligonucleotide) and 200 nM protein. When PhiX174 DNA was used, the concentration was lowered to 0.5 nM due to an increased number of available binding sites. Samples were preincubated at 45 ˚C for one minute. Reactions were started by adding a 1:1 mixture of $MgCl_2$:ATP to a final concentration of 1 mM. At specific time points 40 µL samples were taken and immediately added to 40 µL chilled 0.3 M perchloric acid in a 96 well plate. Malachite green solution (20 µL) was added to each sample and incubated for 12 min at room temperature. Absorbance at 650 nm was measured using a plate reader (BIO-TEK®, PowerWave XS).

2.11.2 Different lengths of oligonucleotides

The reactions were carried out as described above. Oligonucleotide concentrations were either adjusted to identical concentration of oligonucleotide (5 nM) or identical concentrations of nucleotides (2 µM). PhiX174 virion DNA (NEB) was used as a indefinite length DNA substrate (5 nM). The absorbances were read on a SpectraMAX 250 Microplate Reader (Molecular Devices) after 12 min incubation at room temperature. The sequence of the oligonucleotides can be found in the appendix.

2.12 DNA Interactions

2.12.1 Electrophoretic Mobility Shift Assay for XPF (EMSA)

For DNA-binding studies, different concentrations of SsoXPF-HhH or SsoXPF-D52A mutant protein were incubated in binding buffer (50 mM Tris, pH 7.5, 50 mM NaCl, 1 mM EDTA, 0.1 mg/mL BSA, 10 nM DNA substrate and 20 µM PCNA if required) at 55 ˚C for 20 min. Reactions were cooled on ice, mixed with loading dye (6x : 60 mM Tris, pH 7.6, 6 mM EDTA, 30 % glycerol, 0.03 % bromphenol blue, 0.03 % xylene cyanol), loaded

onto a 8 % native acrylamide:TBE gel and electrophoresed at 130 V for 2.5 hours. Control reactions were set up without protein. Gels were exposed to phosphoscreen, imaged and quantified using Fuji Image Gauge software.

2.12.2 EMSA studies of XPD

For DNA binding studies of XPD, different concentrations of protein were incubated 30 minutes at room temperature in binding buffer (20 mM HEPES, pH 7.6, 2 mM DTT, 50 mM NaCl, 0.002 % Triton-X, 0.1 mg/ mL BSA) with 10 nM ^{32}P-5'-radiolabelled DNA substrate (10 µL final volume). Reactions were mixed with 3 µL ficoll loading dye (30 % ficoll, 0.25 % bromphenol blue, 0.1 % xylene cyanol) and 10 µL sample immediately applied to a 6 % acrylamide:TBE gel and run for 2 hours at 130 V. Gels were exposed to phosphoscreen, imaged and quantified using Fuji Image Gauge software.

2.12.3 Quantification and data processing of EMSA

Unbound DNA was quantified using Image Gauge software (Fuji). The fractions of bound DNA were calculated 1 – (unbound/ total) and plotted against the logarithm of the protein concentration (M) using KaleidaGraph software. The data points were fitted to the following equation: $y = 1/ (1 + (m1/ m0)^{m2})$, where m1 is the apparent dissociation constant, m0 is the protein concentration (M) and m2 is a correction factor for deviations in the slope, which could arise from cooperativity.

2.12.4 Analytical gel filtration (Superose 6)

Analytical gel chromatography was used to investigate if XPD binds DNA in form of a monomer or oligomer. For this a Superose 6 column (Amersham) was equilibrated in column buffer (50 mM Tris, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 5 mM DTT). Samples were incubated in column buffer and 200 µL injected onto the column and run at 0.5 mL/ min for 60 min. The absorbance at 280 nm was recorded.

First a standard curve was generated using the following proteins of known molecular weight: blue dextran (2000 kDa), beta-amylase (200 kDa), alcohol dehydrogenase (150 kDa), albumin (66 kDa), carbonic anhydrase (29 kDa) and cytochrome c (12.4 kDa). The void volume (V0) of the column was determined by using blue dextran. The retention times

of the protein markers Beta-amylase/ cytochrome c, alcohol dehydrogenase/ carbonic anhydrase and albumin were recorded and elution volumes (Ve) calculated. The standard curve was obtained by plotting Ve/ V0 against the logarithm of the molecular weight of the proteins. The equation of the obtained standard curve was as follows: $y = 6.3839 – 2.0157x$ ($R = 0.98344$).

All samples were prepared in column buffer and 50 µM ATPγS and incubated at room temperature for 30 minutes before they were applied to the Superose 6 column. The protein concentration was 25 µM and DNA concentrations were at 10 and 30 µM. A 5 nt and a 14 nt dT oligonucleotide, respectively, was used. The following samples were set up and analysed by gel filtration: (1) XPD, (2) 10 µM DNA (dT5 or dT14), (3) XPD + 10 µM DNA (dT5 or dT14) and (4) XPD + 30 µM DNA (dT5 or dT14). Retention times were recorded, Ve and the resulting molecular masses were calculated from the standard equation.

2.12.5 Isothermal Titration Calorimetry (ITC)

For ITC experiments a MicroCal™ VP-ITC micro calorimeter was used (for further information see: http://www.microcal.com). XPD protein and ssDNA oligonucleotides (6 nt, 14 nt, 20 nt dT) were dialysed overnight at 4 ˚C into ITC-buffer (20 mM sodium phosphate buffer, pH 7.4, 100 mM NaCl, 10 % glycerol). Protein and oligonucleotide concentrations were determined by measuring the absorbance at 280 and 260 nm, respectively. The reaction temperature was set to 25 ˚C (no binding was observed at higher temperatures tested, e.g. 45 ˚C and 50 ˚C) and all solutions were degassed for 15 minutes prior to use (ThermoVac, MicroCal™). The reference power was set to 10 and the spacing between injections to 180 s. The syringe contained the ligand (ssDNA: dT6: 97.6 µM; dT14: 99.1 µM; dT20: 100.9 µM), which was injected into the protein solution (10.6 µM). The initial injection of 2 µL (within 4 s) was followed by 28 injections of 10 µL (duration: 20 s). Control reactions were performed (titration of ligand into buffer) and subtracted from the integrated binding data.

In preliminary experiments the heat of dilution of the control reactions were too high and caused the binding curve to change from an exothermal to an endothermal reaction. Since it was not certain if this effect is real, a baseline was estimated of -1600 cal $mole^{-1}$

(20 nt dT) and -1500 cal $mole^{-1}$ (14 nt dT), respectively and subtracted from the original integrated binding data. The data were fitted to a one site of binding model using MicroCal™ ORIGIN software, because it was expected that XPD binds to DNA in a 1:1 ratio.

3 PURIFICATION AND CHARACTERISATION OF XPD FROM *SULFOLOBUS ACIDOCALDARIUS*

3.1 INTRODUCTION

Human XPD (hXPD) is a 5'→3' helicase (Weber et al. (1988)) and it is part of the basal Transcription Factor II H (TFIIH) complex, which consists of 10 subunits (Schaeffer et al. (1993); Ranish et al. (2004); Giglia-Mari et al. (2004)). As part of the TFIIH core, hXPD is the major helicase in Nucleotide Excision Repair. As part of the TFIIH holocomplex it forms a structural bridge between the TFIIH core and the CAK subcomplex and participates in transcription initiation from RNA polymerase II (RNPII) promoters (Reardon et al. (1996); Coin et al. (1998)).

The functional homologue of XPD in bacteria is UvrB; extensive research has been carried out on this enzyme. UvrB binds to UvrA and the complex moves along the DNA, scanning for lesions by base pair flipping. Once damage is identified, UvrA is exchanged for UvrC, which performs a dual incision. Hence, in bacteria UvrB is thought to be the damage recognition factor (reviewed in van Houten et al. (2005)).

Because of its many complex interactions and regulations, which are still not fully understood, characterisation of XPD in eukarya remains difficult. At the initiation of this PhD project, nothing was known about XPD in archaea. Whole genome sequencing and sequence homologies made it possible to identify XPD in the hyperthermophilic crenarchaea *Sulfolobus solfataricus* and *S. acidocaldarius*.

This chapter deals with the initial characterisation of wildtype XPD from *S. acidocaldarius* (SacXPD). General structural and enzymatic features are described. The characterisation of the iron-sulphur (FeS) cluster is supported by some general and technical explanations. If not otherwise stated, XPD refers to SacXPD.

3.2 EXPRESSION AND PURIFICATION OF WILDTYPE *SULFOLOBUS ACIDOCALDARIUS* XPD

The gene encoding *S. acidocaldarius* XPD (Saci_0192) was PCR amplified from genomic DNA and cloned into the pET28c vector to produce untagged recombinant protein. The gene was fully sequenced to confirm it was free of base pair alterations and accurately cloned. Protein was expressed in *Escherichia coli* and purified using a heat step method to remove the majority of *E. coli* proteins, followed by a two-step purification scheme, consisting of affinity (heparin sepharose) and size exclusion chromatography. The whole protein mass was determined and confirmed by mass spectrometry.

Protein was recovered at 5-10 mg/ L culture after heparin sepharose and size exclusion chromatography (Figure 3.1A and B). As seen from the SDS-PAGE in Figure 3.1C, only a small fraction of the total expressed protein was purified. This was a general observation with all XPD proteins and one possible explanation is that most of the protein denatures during sonication, thus not binding the heparin column. XPD eluted with a calculated mass of 49 kDa from the gel filtration column, which was smaller than the expected mass of around 64 kDa (Figure 3.1D). It indicated that SacXPD is a monomer in solution and tight folding of the protein may cause the observed mass differences. Mass spectrometry analysis gave a molecular weight of 63,842 Da which is in good agreement with the expected mass of 63,837 Da (Figure 3.2).

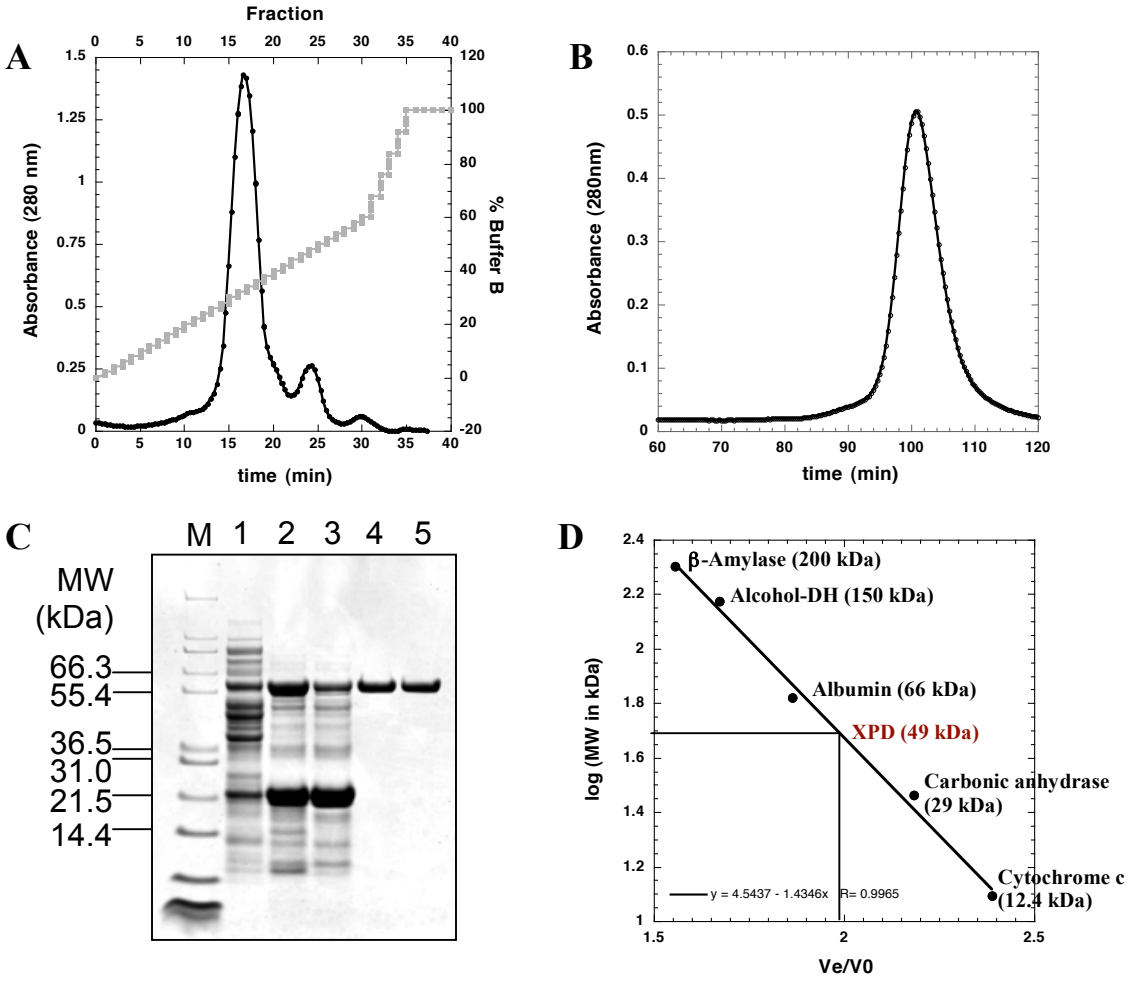


Figure 3.1: Purification of wildtype SacXPD

Elution chromatograms from (A) heparin and (B) gel filtration columns. (C) SDS-PAGE showing the purification steps. From left to right: *E. coli* cell lysate before (1) and after heat treatment at 65 ˚C (2), flow through after heparin column (3), pooled fractions containing SacXPD after heparin column (4) and gel filtration column (5). Marker (M): Mark12 Standard Marker (Invitrogen). (D) Calculation of the molecular mass of XPD from the elution volume (201.5 mL) after gel filtration.

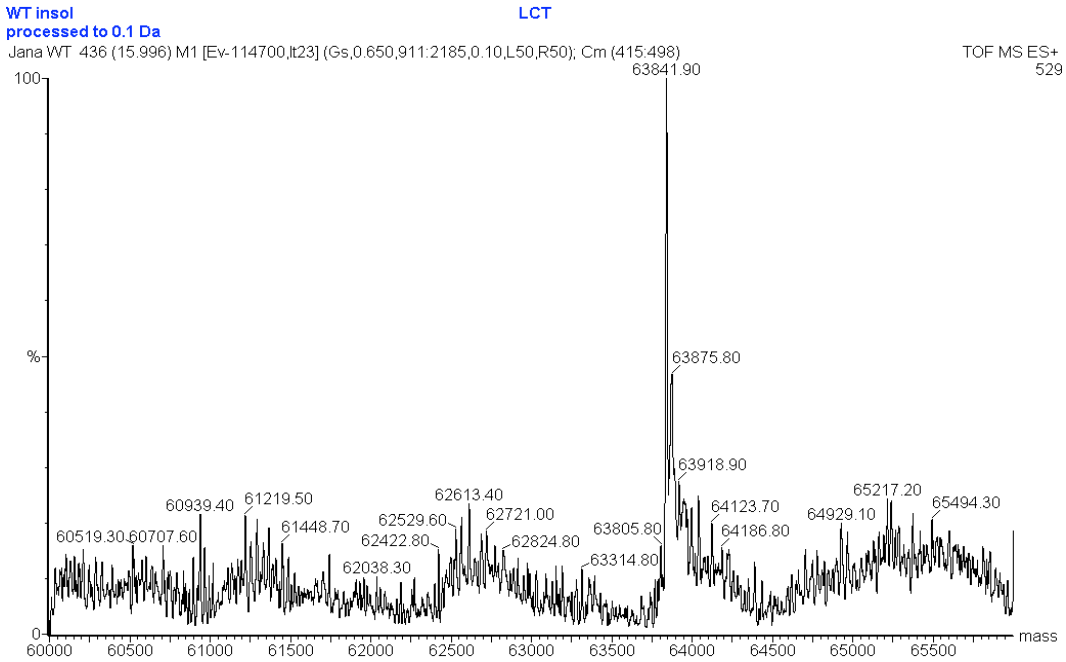


Figure 3.2: Whole mass determination of wildtype XPD by LCT mass spectrometry

The molecular mass of wildtype XPD was determined to confirm an intact full-length protein and the absence of amino acid modifications. The major peak was calculated to a molecular mass of 63842 Da, which was in good agreement with the theoretical mass of 63837 Da.

3.3 IRON-SULPHUR CLUSTER CHARACTERISATION

The protein solution of purified XPD exhibited a brown/ yellowish colour, suggesting the presence of a bound cofactor (Figure 3.3A). UV/ visible spectrophotometry revealed a broad shoulder between 500 and 300 nm (Figure 3.3B), which was found to be characteristic for Iron-Sulphur (FeS) cluster proteins (e.g. Eady et al. (1988); Coldren et al. (1997); Wollenberg et al. (2003)). The absorbance in this region increased after addition of ferricyanide (an oxidant). On the other hand, no change was observed upon addition of the reductant dithionite (DT), implying that the cluster naturally exists in the reduced form.

The nature of the FeS cluster bound to XPD was investigated using three different methods: Electron Paramagnetic Resonance (EPR) spectroscopy, Inductively Coupled

Plasma – Optical Emission Spectroscopy (ICP-OES) and the bathophenanthroline (BA) method.

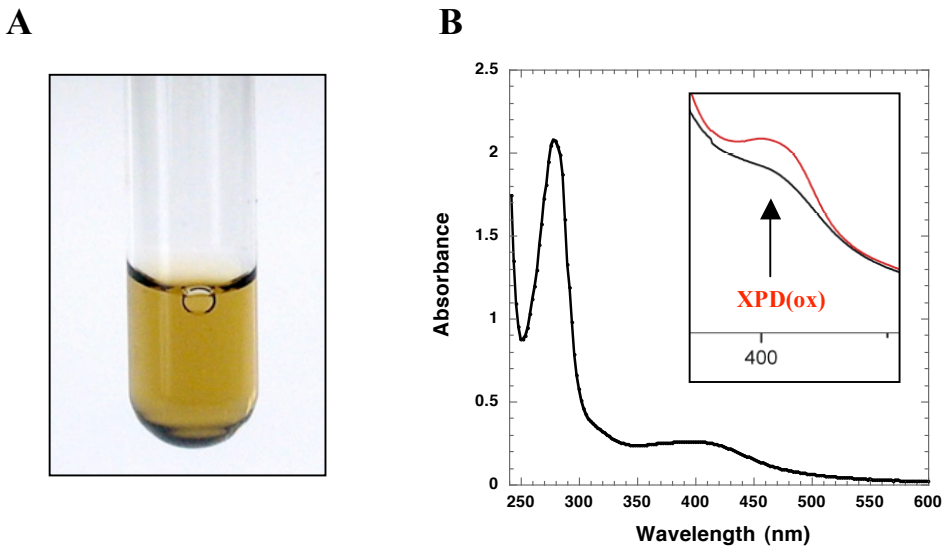


Figure 3.3: Colour and UV/ visible spectra of wildtype XPD

(A) Preparation of purified XPD (~150 µM). (B) UV/ vis spectra. The red line in the inset represents the slight absorbance increase, which was observed after oxidant (ferricyanide) treatment on the shoulder at around 400 nm.

3.3.1 Introduction to EPR

Apart from mass and charge, a fundamental property of an electron is its angular momentum (spin). Electrons are spin 1/2 particles and the resulting spin angular momentum (S) may be clockwise (+1/2) or anticlockwise (-1/2). The angular momentum together with the charge of the electron induces a magnetic dipole (µ) which makes it act like a small bar magnet (Figure 3.4A). If a strong external magnetic field (B_0) is applied, µ forces the electrons to align either parallel or antiparallel to the B_0 (Figure 3.4B). This splits the energy of the ground level in a state of lowest (parallel) and highest (antiparallel) energy depending on the electron spin quantum number m_S (as described above) with $m_S = -1/2$ (lowest energy level) and $m_S = +1/2$ (highest energy level). EPR measures the energy differences (ΔE) between these two energy states (Figure 3.4C).

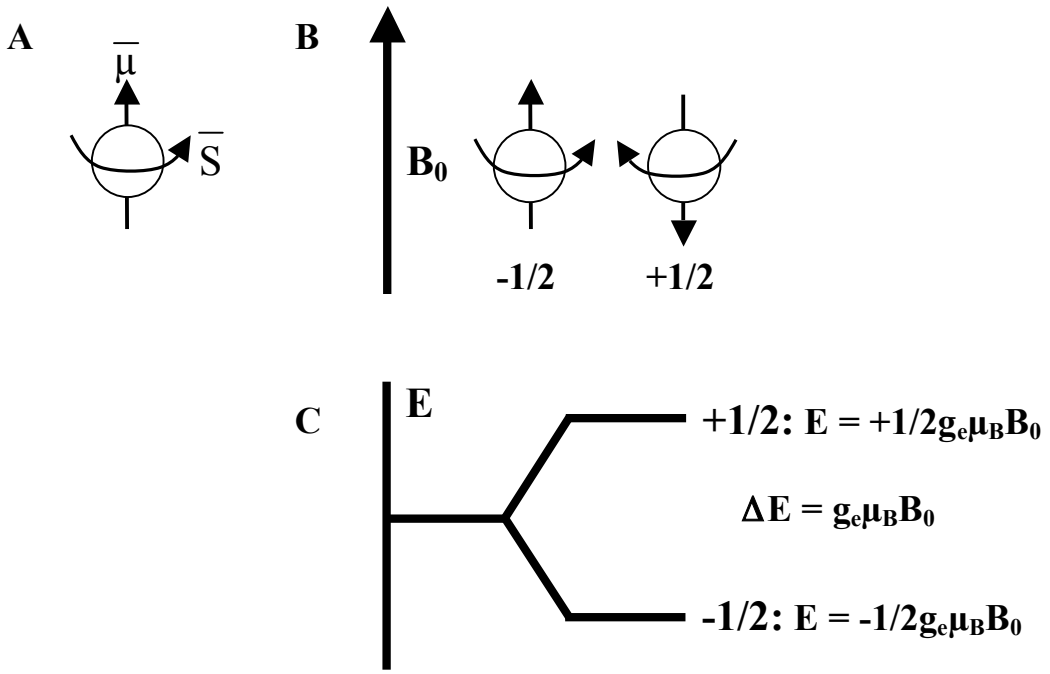


Figure 3.4: The basics behind an EPR experiment: electrons, spins and energy levels

(A) A free, unpaired electron in space shows a specific spin angular momentum (S) which induces, due to the charge of the electron, a magnetic moment μ. (B) If a strong external magnetic field is applied (B_0) the unpaired electrons are forced to align either parallel or antiparallel to it, inducing a state of lowest ($m_s = -1/2$) or highest ($m_s = +1/2$) energy for the electrons. (C) EPR measures the energy differences (ΔE) between both energy levels.

More complex interactions can occur, because electrons do not exist free in space but form part of atoms and molecules. One of these interactions is called the orbital angular momentum L and describes the movement of the electron around the nucleus along three principle axes x, y and z. The probability distribution of an electron is called its orbital. An orbital is occupied by a maximum of two electrons with opposite spins. When electron spins are paired, the resulting net spin is zero and energy absorption does not occur during an EPR experiment, which is called "EPR silent". Hence, for a successful EPR experiment it is necessary that a sample contains unpaired electrons.

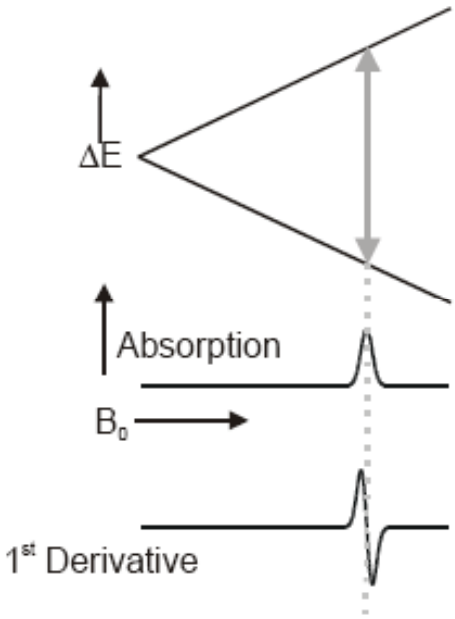


Figure 3.5: Example of how an EPR signal is obtained

The energy levels of the ground state of the unpaired electrons are split by the amount of ΔE, which is a function of the external magnetic field. As the external magnetic field strength is increased, the separation of the energy levels increases. Absorption occurs when the energy separation of the excited states equals that of the impinging microwave photon. The amplitude of the signal is transformed into its first derivative.

The energy difference between the excited states of an electron can be calculated from: $\Delta E = h\nu = g\mu_B B_0$ (h, Planck constant; g, spectroscopic factor; ν, frequency of radiation (Ghz); μ_B, Bohr magneton; B_0, magnetic field (Gauss)) An EPR spectrum is obtained by sweeping the magnetic field at constant microwave frequency (it is technically more difficult to vary the latter in a broad range). Irradiation of the EPR sample with microwave frequencies generates a microwave photon. When ΔE equals the energy of the impinging microwave photon, energy absorption occurs and an EPR signal is generated (Figure 3.5).

The spectroscopic factor g of a free electron equals ~ 2.00, while deviations from this number usually arise due to spin-orbit coupling between the ground state and the excited states of an electron and the specific orientation of the orbital (anisotropy).

Therefore, the spectroscopic factor measures along the three principle axes of a molecule, resulting in g-factors labelled as g_x, g_y and g_z (Figure 3.6).

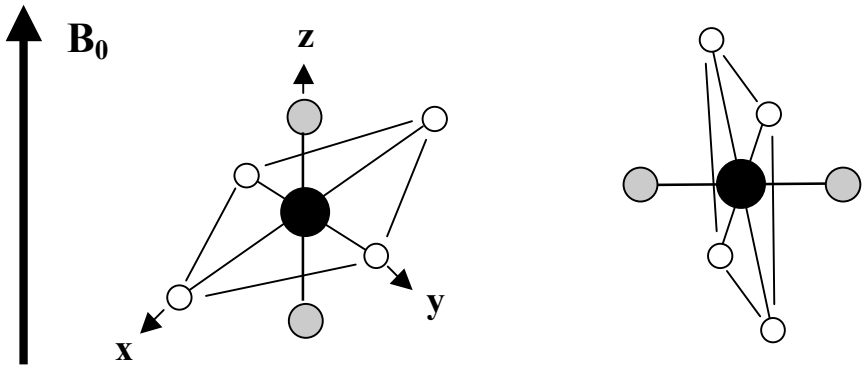


Figure 3.6: Example of a transition metal coordinated by ligands in the three principle axes relative to B_0

The transition metal is coordinated by two equal ligands in z-direction and by two different but equal ligands in either x and y direction ($g_z > g_y = g_x$). The resulting g-factor differs depending on whether B_0 is parallel to the z-axis or the x and y-axes.

The temperature at which an EPR experiment is carried out is an additional variable. It can vary from high temperature gas phase to cryogenic temperatures. At which temperature EPR measurements are made, depends on the relaxation time of the species under observation. Transition metals for example have very high relaxation rates, hence the EPR experiment has to be carried out at very low temperatures to slow relaxation times. This generates a spectrum that is called a powder spectrum, because in a frozen solution the orientation of the unpaired electron along x, y and z relative to B_0 is randomised (anisotrop). Examples of anisotropy are given in Figure 3.7.

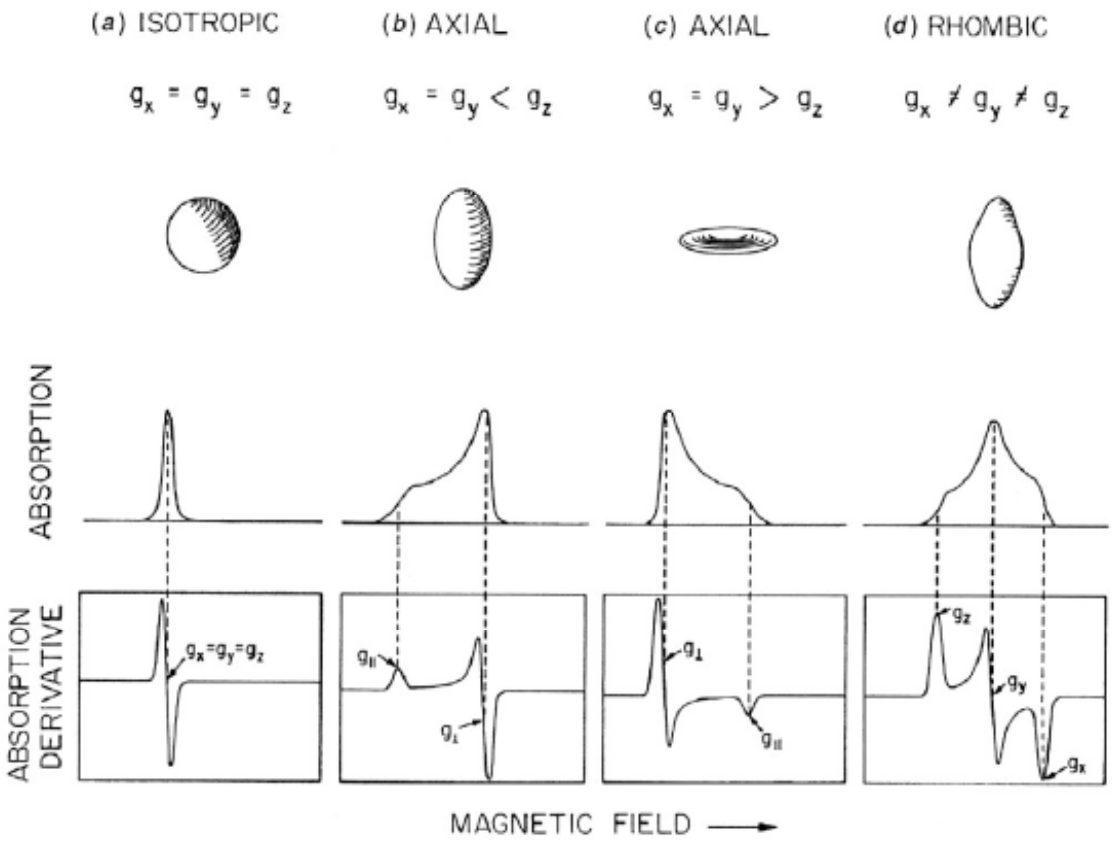


Figure 3.7: Absorption and first derivative of three different classes of anisotropy

For the first class of anisotropy $g_x=g_y=g_z$ all principle g-factors are the same (isotropy). The second class is called axial, meaning that one axis differs from the other two. In the third class (rhombic) all three g-factors differ from one another (image adapted from: http://www.auburn.edu/~duinedu/5%20samples.pdf.).

If two unpaired electrons from different ions or molecules are close enough together, they can exchange-couple, resulting in a net spin S = 0. This is the case for the iron ions in FeS clusters. A FeS cluster is a mixture of ferrous (Fe^{2+}, reduced form, six electrons in the outer shell, S = 2) and ferric (Fe^{3+}, oxidised form, five electrons in the outer shell, S = 5/2) iron. Depending on the number of irons present in a cluster, their oxidation states, electron coupling and coordinating ligands, specific signals are observed in an EPR experiment (Figure 3.8). In general, FeS clusters exist in two different oxidation states, whereas one of the oxidation states is always EPR silent due to spin-coupling. For characterisation of the cluster, the g-values are important as they have

specific values for a specific type of FeS cluster. These values are calculated by the following equation (as shown above): $h\nu = \Delta E = g\mu_B B_0$, where B_0 and ν can be measured (h and μ_B are constants) resulting in $g = 714.484\ \nu / B_0$.

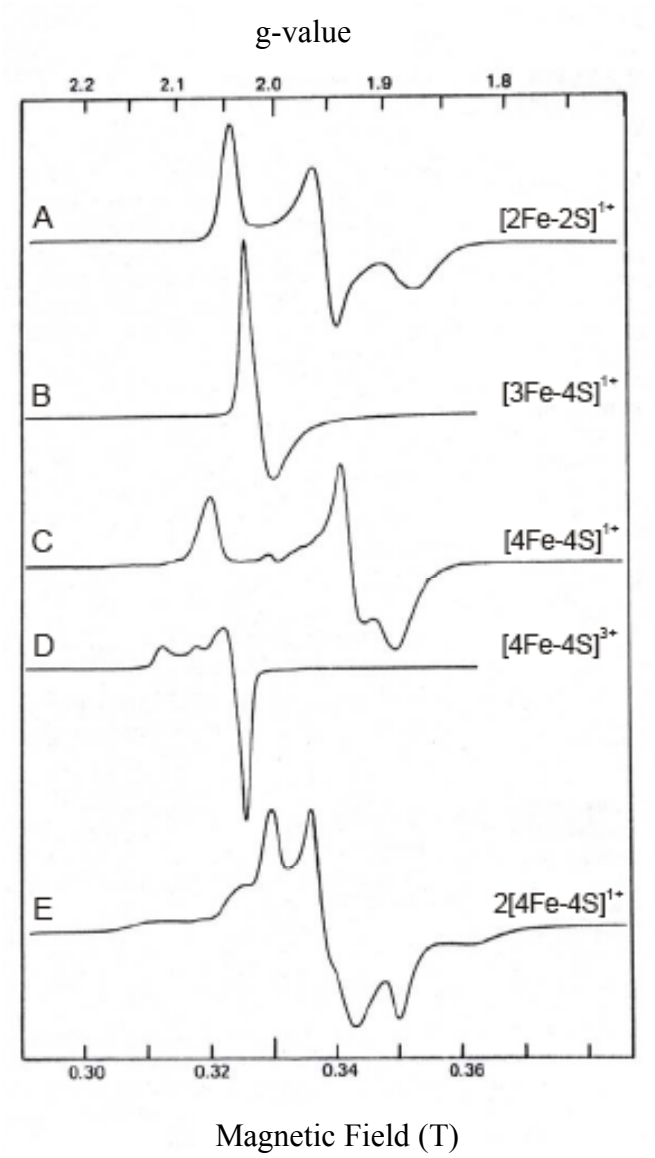


Figure 3.8: Examples of EPR signals of different types of FeS clusters

Typical EPR signals obtained for different FeS clusters of ferredoxins from *Mastigocladus laminosus* (A), *Desulfovibrio gigans* (B), *Bacillus stearothermophilus* (C), *Chromatium vinosum* high potential FeS cluster (D) and *Clostridium pasteurianum* 8 Fe ferredoxin (E).
(image adapted from: http://www.auburn.edu/~duinedu/5%20samples.pdf.)

3.3.2 EPR of XPD

EPR spectroscopy experiments on XPD only gave a signal for the oxidised form of the FeS cluster (Figure 3.9). There are two possible forms of FeS clusters which give EPR signals in the oxidised form: $[3Fe\text{-}4S]^{1+}$ and high potential iron protein type (HiPIP) $[4Fe\text{-}4S]^{3+}$ clusters (reviewed in: Beinert et al. (1997)). The shape of the signal and g-values of both oxidised forms differ (Figure 3.8). For XPD, the signal and g-values (2.02 and 1.99) indicated the presence of a [3Fe-4S] cluster bound to XPD. Obtaining an EPR spectrum turned out to be rather difficult and the EPR signal was very weak. The cluster is likely to be deeply buried within the enzyme and oxidation becomes difficult, as oxidising agents only have limited access to the FeS cluster. This was also found to be the case for other FeS cluster proteins and represents a natural protection mechanism. FeS clusters are sensitive to oxidising conditions and oxidation most likely leads to destruction. As these metallo-clusters fulfil important structural and catalytic roles as part of an enzyme, damage has to be avoided (reviewed in Imlay (2006)).

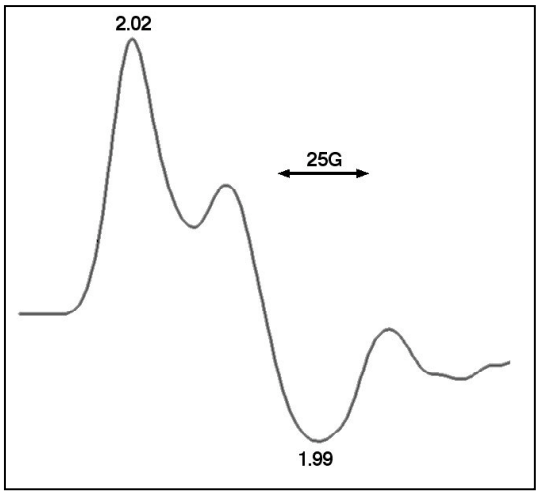


Figure 3.9: EPR Spectrum of oxidant treated XPD

The shape of the EPR signal is typical of a $[3Fe\text{-}4S]^{1+}$ cluster. The g-values were calculated as 2.02 and 1.99. The second peak which is not labelled with a g-value may be generated by a free radical.

3.3.3 ICP-OES

ICP-OES uses ionized argon gas. Temperatures of the created argon plasma reach approximately 10000 Kelvin. The sample is injected as an aerosol directly into the centre of the plasma where it completely atomises. The resulting emissions are detected and compared to standards of known concentrations (Figure 3.10). To calculate the iron to sulphur ratio of a protein, all bound iron (usually only the ones which are part of the FeS cluster) and all bound sulphurs (inorganic sulphurs as part of the cluster, as well as sulphurs incorporated in cysteines and methionines) have to be included. Expected ratios are calculated and compared to ratios obtained experimentally. [From: PerkinElmer: Guide to Inorganic Analysis]

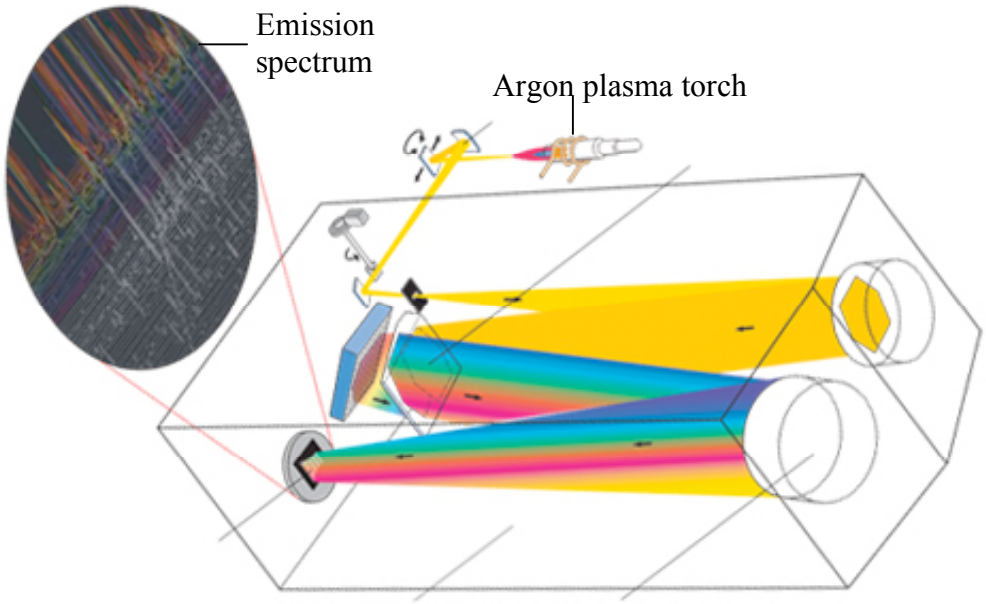


Figure 3.10: Schematic diagram of the mode of operation of an ICP-OES instrument

The sample is atomised in the argon plasma and the emitted light is directed and split through a mirror systems until it arrives at the detector where the emission spectrum is recorded (image adapted from: http://marine.rutgers.edu/LAICPMSintro/ICP-OES%20mainframe.htm.htm)

XPD contains 18 sulphurs bound to methionines or cysteines and four inorganic sulphurs as part of the FeS cluster. The expected Fe:S ratio for an 4Fe-4S cluster was calculated to be 0.182 and for a 3Fe-4S cluster 0.136. ICP-OES yielded an iron to sulphur ratio for XPD of 0.133, consistent with the value of 0.136 expected for a 3Fe-4S cluster.

3.3.4 The bathophenanthroline-method

BA (4,7-diphenyl-1,10-phenanthroline) is a Fe^{2+} chelator and forms a red complex with ferrous iron in solution. Iron becomes easily oxidised and a strong reductant (e.g. ascorbic acid) is used to bring all iron ions in the reduced form before BA is added. BA is one the most sensitive agents for the detection of ferrous iron and it forms a $[Fe^{II}(BA)_3]^{2+}$ complex similar to the ruthenium-BA complex shown in Figure 3.11.

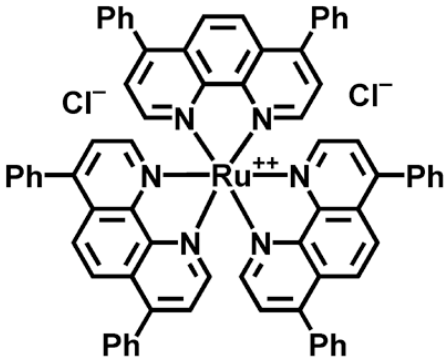


Figure 3.11: Bathophenanthroline -ruthenium complex
Fe^{2+} and Ru^{2+} are bound in the same way by BA. The ion is coordinated by three molecules of BA through their nitrogen atoms.

The number of irons bound to one molecule of XPD were determined by measuring the absorbance of the sample at 535 nm and calculated using the molar extinction coefficient of 22369 M^{-1} cm^{-1} (Pieroni et al. (2001)) as described in Materials and Methods.

Table 3.1: Calculation of the number of irons bound to XPD from absorbance readings using the bathophenanthroline method

Protein	Abs_{535}_1	Abs_{535}_2	Abs_{535}_3	Average	Total Fe (pmol)	Fe/ XPD
WT-XPD	0.1412	0.1380	0.1655	0.1482	16364	3.27

The volume taken for analysis after protein denaturation was 100 µL out of the total volume of 130 µL (100 µL protein solution + 30 µL concentrated HCl). The volume of the final reaction was 1.9 mL. Therefore, the total number of irons (pmol) in the sample was determined with: (Average absorbance/ 0.022369 μM^{-1}) x 1900 µL x 1.3 (Table 3.1). Initially, 5000 pmol XPD protein (100 µL of a 50 µM sample) was used. Hence, the number of irons per molecule XPD equals 16364 pmol Fe/ 5000 pmol XPD, which leads to a ratio of approximately 3 Fe per XPD (Table 3.1).

3.3.5 Location of the FeS cluster domain.

Iron-sulphur clusters are mainly coordinated by cysteine residues due to the high affinity of iron for thiolate groups. The iron atoms are bridged by sulphide ions and thus a cluster is built. The main types of clusters found in nature are $[2Fe\text{-}2S]^{2+/1+}$, $[3Fe\text{-}4S]^{1+/0}$, $[4Fe\text{-}4S]^{2+/1+}$ or $[4Fe\text{-}4S]^{3+/2+}$ (reviewed in Imlay (2006)).

Sequence alignments of XPD with other archaeal XPD sequences as well as related human helicases, revealed four conserved cysteine residues near the N-terminus of the protein between Walker A and B boxes (Figure 3.12), which could be involved in the coordination of the FeS cluster (Rudolf et al. (2006)). These cysteines were also conserved in hXPD raising the possibility that hXPD contains an unrecognised FeS cluster.

Related human helicases other than XPD include FancJ, RTel1 and Chl1. The FancJ helicase takes part in the Fanconi anemia DNA repair pathway, which was introduced in Chapter 1. The RTel1 helicase is one of the helicases involved in regulation of telomere maintenance. Lack of RTel1 led to genomic instability and embryonic lethality in mice (Ding et al. (2004)). It is thought that RTel1 resolves secondary structures in telomeric repeats and hence prevents illegitimate recombination between telomeric DNA and between other intrachromosomal sequence repeats during

recombination induced by attempted DNA repair. The Chl1 helicase shows high sequence similarity to FancJ and XPD/ Rad3p (Amann et al. (1997)). In yeast it is involved in sister-chromatid cohesion and segregation. Although not essential for viability, yeast Chl1 null mutants show increased rates of chromatid missegregation with chromosome loss and non-disjuctions (Gerring et al. (1990)). Therefore, Chl1 may be involved in avoiding illegitimate recombination events (like RTel1) during sister-chromatid cohesion (Skibbens (2004)).

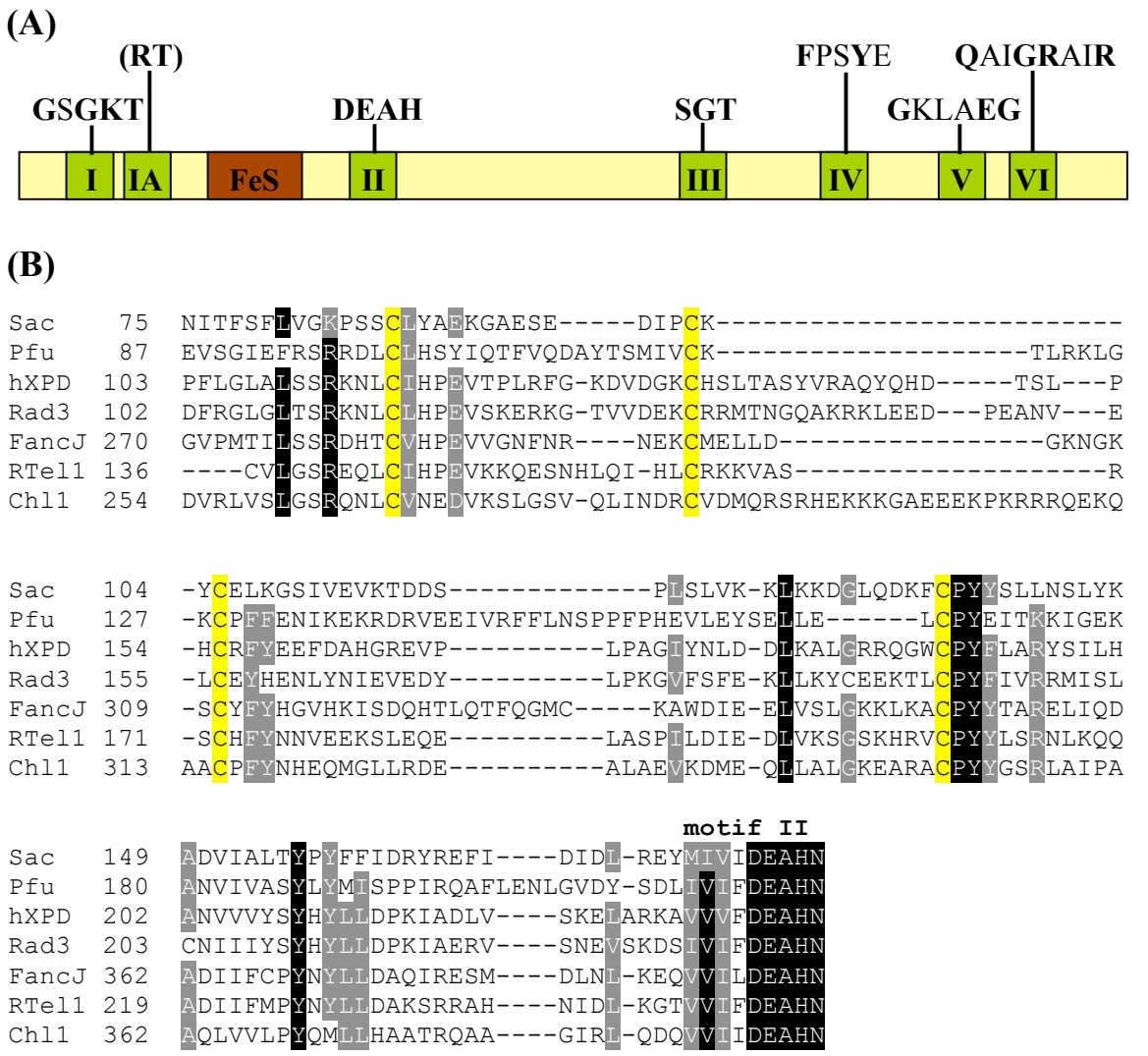


Figure 3.12: Location of the FeS cluster domain in XPD and other related human helicases

(A) The location of the FeS cluster domain relative to the conserved helicase motifs of SacXPD is illustrated. The seven conserved motifs are shown in green and the location of the FeS cluster domain is presented in brown. The core amino acid sequence of the helicase motifs of *S. acidocaldarius* XPD is shown across the top of the figure, strictly conserved residues within archaeal and eukaryal XPD proteins are shown in bold. (B) A sequence alignment of the FeS cluster domain of archaeal and eukaryal XPD and other related human helicases is shown. The FeS cluster domain is conserved within XPD proteins from different archaea and eukarya, including human XPD. The conserved cysteine residues, which could be involved in FeS cluster coordination are highlighted in yellow; similar amino acid residues are in grey and absolutely conserved residues are highlighted in black. Sac, *S. acidocaldarius* XPD; Pfu, *P. furiosus* XPD; hXPD, human XPD; Rad3, *S. cerevisiae* XPD homologue Rad3; FancJ, human FancJ; RTel1, human RTel1; Chl1, human Chl1.

3.4 IRON-SULPHUR CLUSTER RECONSTITUTION

As described in section 3.3, XPD contains a $[3Fe\text{-}4S]^0$ cluster. It was also mentioned that FeS clusters are sensitive to oxidising agents. Reconstitution of the FeS cluster was carried out because the enzyme was purified under aerobic conditions. This could lead to the loss of one iron during the purification process (Tilley et al. (2001)). However, it does not necessarily imply a loss of activity. Several enzymes were shown to be active in the 3Fe and the 4Fe form, including *E. coli* MutY and ribonucleotide reductase and the F420-reducing hydrogenase from *Methanococcus voltae* (Mulliez et al. (1999); Bingemann & Klein (2000); Lu & Wright (2003))

The reconstitution was carried out under semi-anaerobic conditions and the method is described in Chapter 2. Simply, reduced iron in the form of $(Fe(II)(NH_4)_2(SO_4)_2)$ was added to the protein solution and the reconstitution progress was followed by measuring the absorbance between 600 and 300 nm. The iron quantification was carried out by using the bathophenanthroline method. A control was made containing buffer and 1.7 mM $Fe(II)(NH_4)_2(SO_4)_2$. This solution was applied to a G50-microspin column and the flow through analyzed for remaining iron. The absorbance measurement of the control was subtracted from the absorbance of the sample containing the protein with the reconstituted cluster. The final results were compared to a sample containing protein with a non-reconstituted FeS cluster. Even though this method is not very accurate and it has a high error rate, the results indicated, that XPD has a 3Fe-4S cluster bound, because both the unreconstituted and the reconstituted cluster revealed 3 moles of iron per one mole of protein (Table 3.2).

Table 3.2: Determination of iron bound to XPD before and after FeS cluster reconstitution

	Control+Fe	unreconstituted	reconstituted
A (535 nm)	0.0513	0.1290	0.1916
	0.0857	0.1225	0.1722
	0.0461	0.1246	0.1835
Average absorbance	0.061	0.1254	0.1824
Average minus (control + Fe)	-	-	0.1214
Total Fe (pmol)	-	14075.25	13630
Fe/ XPD	-	2.82	2.73

Table 3.2 shows data of triplicate absorbance readings ($A_{(535nm)}$). The iron concentration was determined using the bathophenanthroline method. The control sample contained buffer and 1.7 mM $Fe(II)(NH_4)_2(SO_4)_2$ (Control+Fe). The unreconstituted sample contained XPD protein only while the reconstituted sample contained XPD and 1.7 mM $Fe(II)(NH_4)_2(SO_4)_2$. Before the iron content was determined, all samples were purified using a G50 microspin column followed by NaCl addition to a final concentration of 200 mM. The total amount of irons in the sample and the number of irons bound to the XPD protein was calculated as described in section 3.3.4.

3.5 CHARACTERISATION OF WILDTYPE HELICASE ACTIVITY

3.5.1 Polarity and ATP-dependence

Most helicases unwind nucleic acids unidirectional, either 5'→3' or 3'→5'. Two different DNA substrates were made to investigate the directionality of XPD: a 5'-overhang and a 3'-overhang DNA substrate (Figure 3.13A). XPD can load either on the 5' or 3' single stranded region to unwind the double stranded part in the opposite direction. If XPD is a 5'→3' helicase, unwinding should be observed with the 5'-overhang DNA substrate, while a 3'→5' helicase would displace the strand of the 3'-overhang. Helicase activity for XPD from *S. acidocaldarius* was only detected by using a 5'-overhang DNA substrate (Figure 3.13B), confirming a 5'→3' directionality consistent with findings for hXPD (Weber et al. (1988)).

The energy for the motion of a helicase is provided by hydrolysis of NTP and the majority of helicases utilize ATP (Levin & Patel (2003)). The ATP dependence of XPD

for DNA unwinding was tested by performing reactions with and without ATP and detecting appearance of the ^{32}P-radiolabelled single DNA strand. For this, and all following experiments, a 5'-overhang DNA substrate was used. As seen in Figure 3.13B strand displacement was only observed in the presence of ATP.

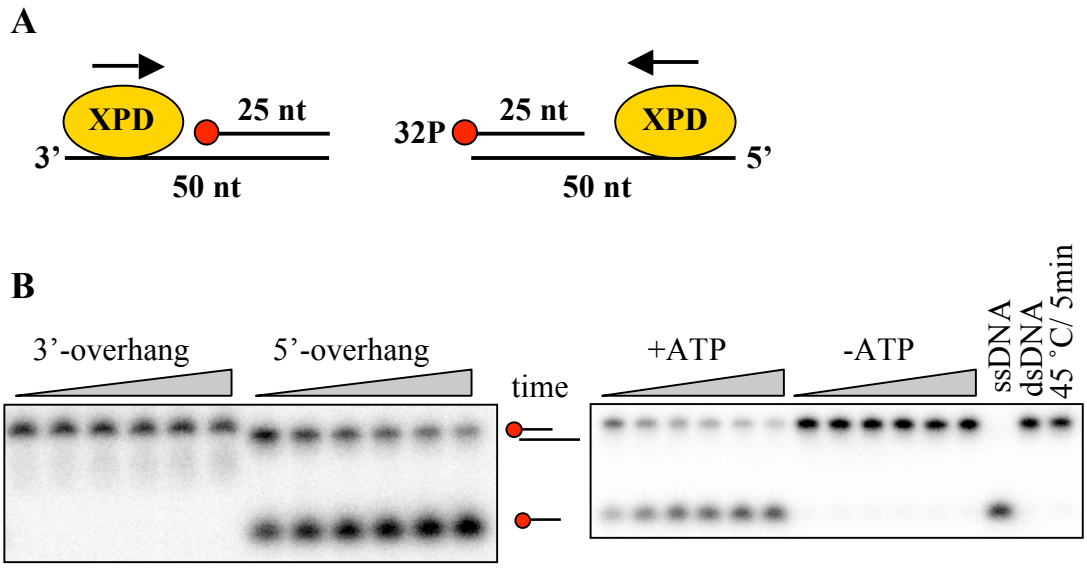


Figure 3.13: Helicase activity of wildtype XPD: Polarity and ATP dependence

(A) 3'-overhang (left) and 5'-overhang (right) DNA substrates. The protein binds to the single stranded region and displacement of the 25 nt oligonucleotide, which is ^{32}P-5'-labelled, is expected for only one of the two substrates, depending on the polarity of XPD. (B) 12 % native acrylamide gels showing a time course (30, 60, 90, 120, 180 and 300 s) of XPD unwinding activity. The amount of dsDNA (upper lane) is expected to decrease from left to right, while the amount of ssDNA (lower lane) should increase, if helicase activity takes place.

3.5.2 Influence of protein concentration on XPD helicase activity

The influence of the protein concentration (200, 100, 50, 25, 10 and 5 nM) at constant DNA concentrations (10 nM) on the helicase activity of XPD was investigated.

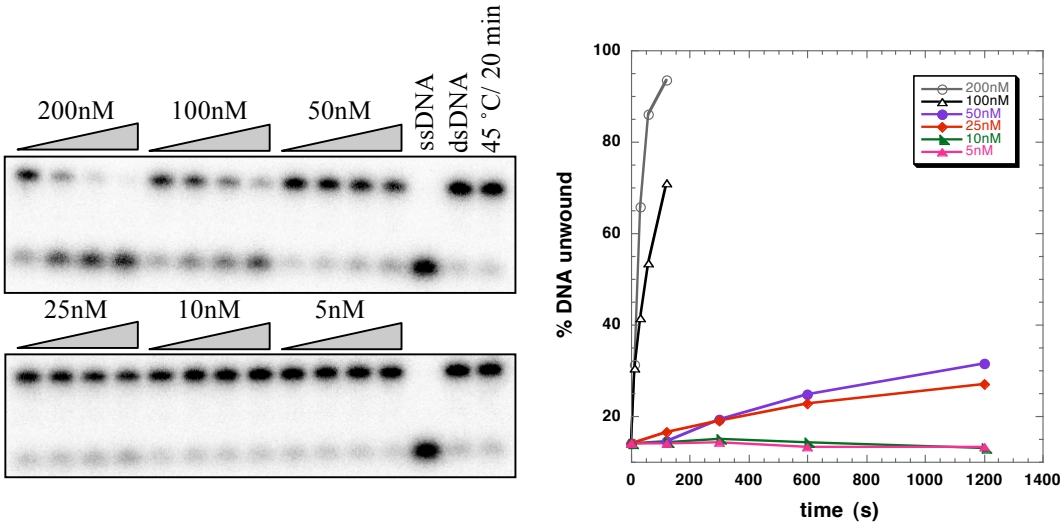


Figure 3.14: Helicase activity of wildtype XPD: Influence of varying protein concentrations

Representative 12 % native acrylamide gels showing time courses of DNA unwinding experiments with different concentrations of XPD (left). Time points were: 10, 30, 60, 120 s (200 and 100 nM protein) and 2, 5, 10, 20 min (50, 25, 10, 5 nM protein). The quantification results from a single experiment are shown in graphical format on the right. Data points were connected by lines.

Sufficient helicase activity under the applied conditions was only observed with at least 10-fold excess of protein over DNA substrate (Figure 3.14). Limited activity was measured at protein concentrations of 50 and 25 nM, but no activity was seen at 10 and 5 nM XPD. For all further experiments 200 nM protein was used. These observations could suggest a limited affinity of XPD for the DNA, with dissociation constants of around 100 nM. On the other hand, it could also imply a positive cooperativity for the mechanism of action for XPD, as was described for the Dda helicase from bacteriophage T4 (Byrd & Raney (2004)).

3.5.3 Influence of the Temperature

Enzymatic reactions, as all reactions occurring in nature, are temperature dependent. XPD is an enzyme from a hyperthermophilic organism and one would expect the rate of DNA unwinding to increase with increasing temperature. Experiments with enzymes from *Sulfolobus* were usually carried out in our laboratory at 55 ˚C. However, the DNA unwinding activity of XPD was too fast at 55 ˚C to be able to compare different conditions in an appropriate way. Therefore, a slightly lower temperature, where XPD shows adequate helicase activity was sought.

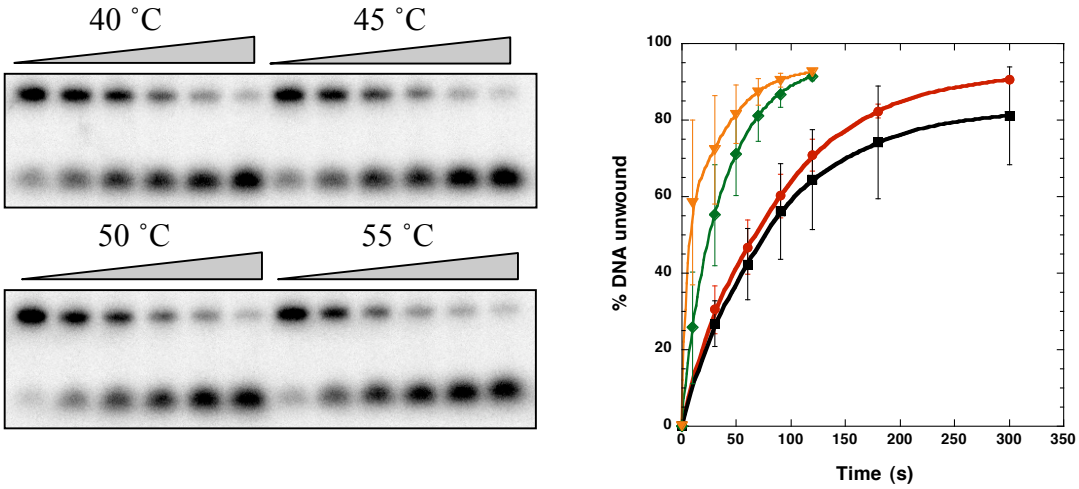


Figure 3.15: Helicase activity of wildtype XPD: temperature dependence

Representative 12 % native acrylamide gels are shown (left). Time points for the upper gel are: 30, 60, 90, 120, 180, 300 s; lower gel: 10, 30, 50, 70, 90, 120 s. The graph (right) shows quantification results of three experiments with indicated standard errors. Data points were fitted to a smooth curve. Black, 40 ˚C; red, 45 ˚C; green, 50 ˚C; orange, 55 ˚C.

As expected, XPD showed temperature dependence. There was not much difference in helicase activities between 40 ˚C and 45 ˚C, as well as between 50 ˚C and 55 ˚C (Figure 3.15). All subsequent experiments were carried out at 45 ˚C, because sufficient and reliable enzyme activity was obtained at this temperature.

3.5.4 Influence of NTP/ Me^{2+} on the helicase activity of XPD

As mentioned before, helicases are fueled by the energy derived from NTP hydrolysis. The NTP forms a 1:1 complex with a divalent metal ion. The helicase uses this metal ion to activate water to split the β-γ-phosphoanhydride bond of the NTP.

The effect of different concentrations of NTPs and divalent metal ions on the helicase activity of XPD was tested. Reactions were initiated by adding the following concentrations of NTPs or divalent cations: ATP or GTP and $MgCl_2$ in equivalent concentrations of 5, 2, 1, 0.5 and 0.1 mM and $MgCl_2$, $MnCl_2$, $CaCl_2$ at 1 and 2 mM with 1 mM ATP. Control reactions contained 1 mM $MgCl_2$ and 1 mM ATP only, respectively. Reactions were incubated 3 min at 45 ˚C.

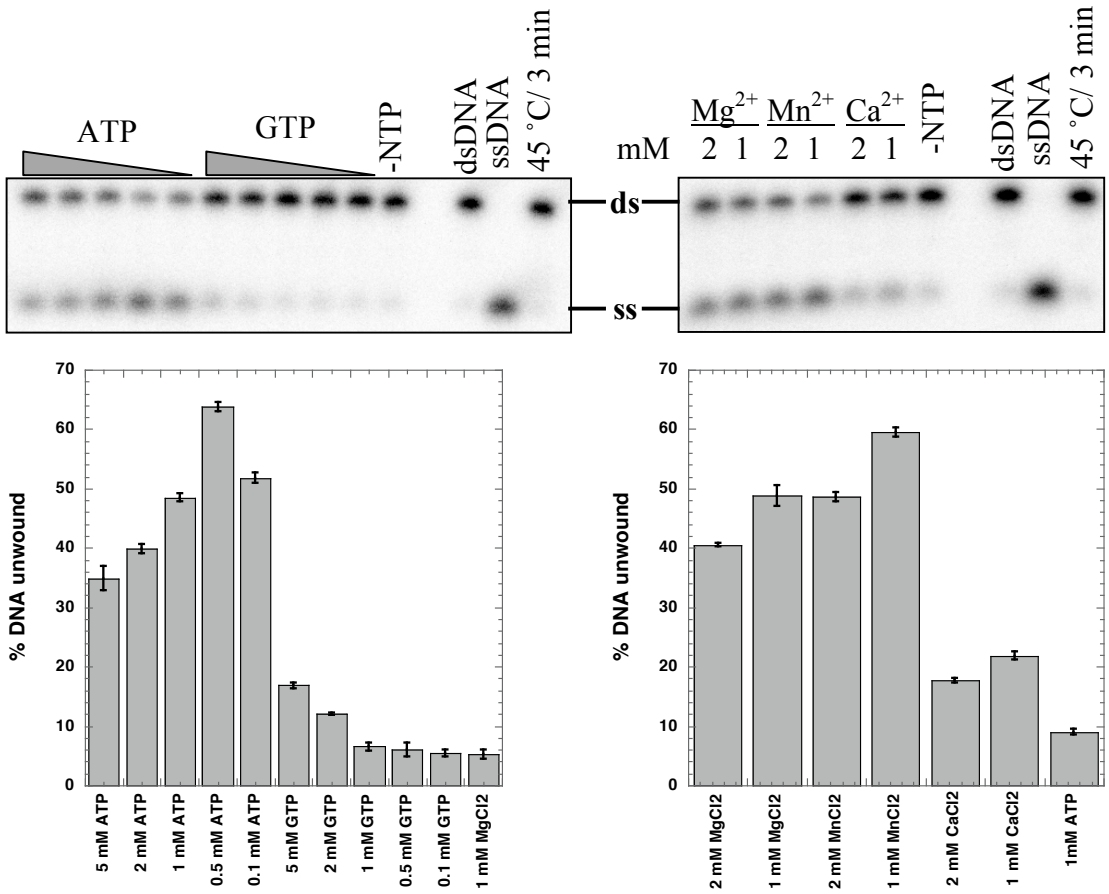


Figure 3.16: Helicase activity of XPD with different concentrations of NTPs and Me^{2+}

The top figures are examples of native 12 % acrylamide gels showing XPD helicase activity under various conditions as shown across the top. ds, double stranded DNA substrate; ss, 5'-^{32}P radiolabelled single stranded DNA appearing after DNA substrate unwinding. The bottom figures are the quantifications of duplicate experiments (corresponding to the gels above) showing percentage DNA unwound; standard errors are indicated. ATP or GTP/ $MgCl_2$ concentrations are: 5, 2, 1, 0.5, 0.1 mM.

XPD utilizes mainly ATP for its enzymatic activity (Figure 3.16). The highest activity was observed with 0.5 mM ATP, which matches results obtained for the ATPase activity (see Chapter 4). It seems that XPD can also utilize GTP at high concentrations, since the helicase activity increases with increasing concentrations of GTP.

Generally, most helicases and NTPases are able to use both Mg^{2+} and Mn^{2+} ions. From the quantification data (Figure 3.16) it seems that manganese is preferred over magnesium ions, a property mainly found for viral (RNA-dependent) NTPases/ helicases

(Warrener & Collett (1995), Borowski et al (1999)). The difference, however, was not significant. Ca^{2+} was used as well, but led to a much lower enzymatic activity.

3.5.5 Influence of buffer conditions on the helicase activity of XPD

The influence of different salt concentrations on the DNA unwinding activity of XPD was tested by using reaction buffer (20 mM MES, 1 mM DTT) and different concentrations of KCl or NaCl (150, 100, 75, 50 and 25 mM final concentrations). Furthermore, the effect of different pH-values was investigated by using 20 mM of the following buffer solutions (including 1 mM DTT): MES, pH 6.0 and 6.5; Tris, pH 7.0, 7.5 and 8.0. Reactions were incubated for 5 min at 45 ˚C.

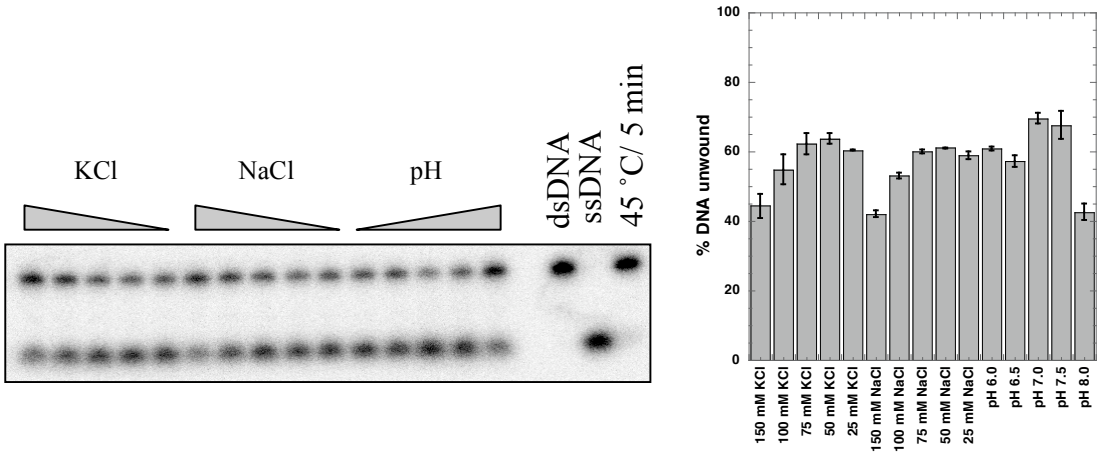


Figure 3.17: Helicase activity of WT-XPD under different buffer conditions

The figure on the left shows a representative 12 % native acrylamide gel of helicase activity under different salt and pH conditions (from left to right). The respective quantification data are shown on the right. Experiments were carried out in duplicate and standard errors are indicated.

As illustrated in Figure 3.17 the preferred buffer conditions were: no salt and pH 7-7.5. In general, the enzyme activity did not vary much in different buffers. Salt concentrations above 100 mM and pH > 7.5 lead to a decrease in activity, probably through weakening the DNA interaction (ionic interactions).

3.5.6 Influence of oxidising and reducing conditions

XPD is a FeS cluster protein (3.3). These metallo clusters are sensitive to oxidizing and reducing agents. The influence of the oxidant potassium ferricyanide ($K_3Fe(CN)_6$) and the reductant dithiothreitol (DTT) was tested (Figure 3.18). Stock solutions were stored at -20 ˚C (DTT) or 4 ˚C ($K_3Fe(CN)_6$) and freshly added to the reactions.

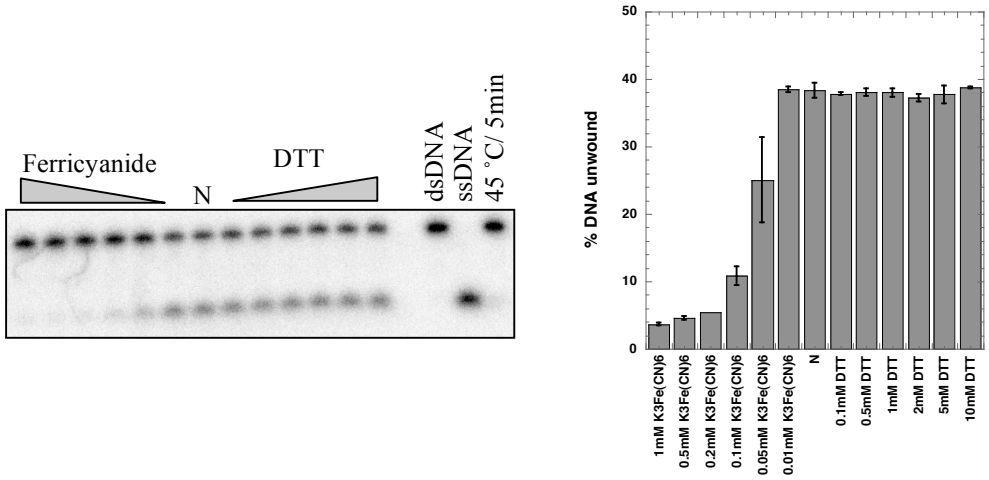


Figure 3.18: Helicase activity of wildtype XPD: Influence of reducing and oxidising agents

A representative 12 % acrylamide gel of the helicase experiments is shown on the left. Concentrations were: 1, 0.5, 0.2, 0.1, 0.05, 0.01 mM ferricyanide and 10, 5, 2, 1, 0.5, 0.1 mM DTT. The control reaction (N) contained neither of them. Quantification results are shown on the right. Experiments were carried out in duplicate and standard errors are indicated.

Reducing conditions did not have an influence on the enzyme activity, while oxidant-treatment showed a dramatic effect (Figure 3.18). Most likely protein precipitation upon oxidant treatment can be excluded as an explanation for enzyme inactivation, because EPR experiments were set up using 1 mM ferricyanide and no cluster destruction or protein precipitation was observed. Although, it is possible that the enzyme has a more compact structure at room temperature (setup for EPR) than at 45 ˚C (setup for helicase assays), which could limit the access of oxidant at lower temperatures. It could also account for the difficulties we had in obtaining EPR spectra for the oxidised form of the FeS cluster. Another possible reason for the observed enzyme inactivation upon oxidant treatment is that cyclic conformational changes in XPD induced by binding

to DNA and/ or ATP, lead to solvent exposure of the FeS cluster. Usually, as it was hypothesized for the glycosylases of the MutY/ EndoIII family (Boal et al. (2005)), FeS clusters are deeply buried within the enzymes. Access of oxidant to the FeS domain may eventually lead to a loss of irons. It is possible that one of the three irons is more exposed when bound to DNA and/ or ATP. Oxidation would change the overall oxidation state of the cluster and subsequently lead to instability. This can cause the loss of the oxidised iron, followed by loss of enzymatic activity and perhaps cluster destruction (Imaly (2006)). It remains to be shown whether the effect observed upon oxidant treatment is reversible, for example by addition of a reductant and checking for recovery of the helicase activity.

3.6 HELICASE ACTIVITY ON DIFFERENT LENGTHS OF ssDNA OVERHANGS

This experiment was conducted to examine the minimum binding site necessary to obtain efficient substrate unwinding. Double stranded DNA substrates (25 nt; see Appendix 1, Table 2) with different lengths of 5'-dT overhangs were made and helicase reactions carried out under standard conditions. Earlier time points were fitted to a linear regression and the slope represented the first order rate constant (k_{cat}). The percentage of duplex unwinding per second was calculated into pM of unwound duplex DNA per second (Figure 3.19).

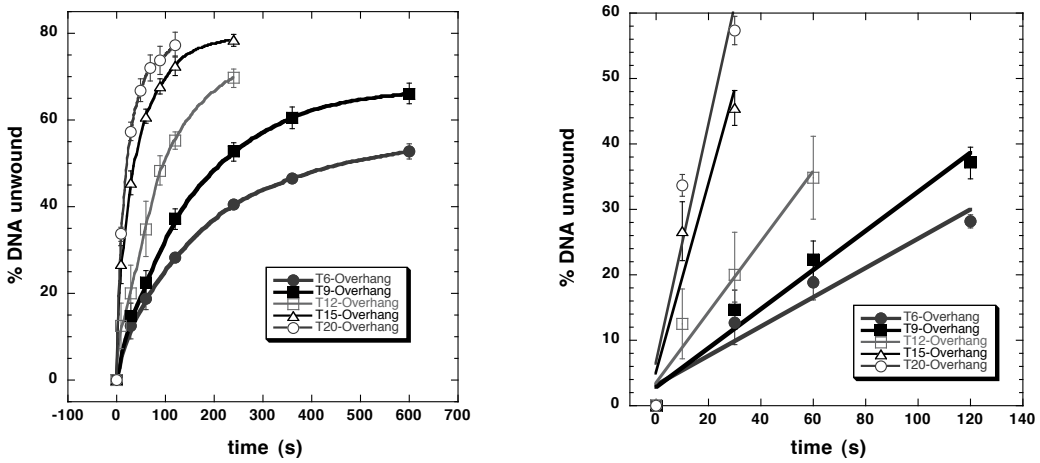


Figure 3.19: XPD helicase activity on different lengths of dT 5'-overhangs

The graph shows quantification results of at least duplicate experiments where XPD was tested on different lengths of dT 5'-overhangs (see legend) Time points for the figure on the left were 10, 30, 50, 70, 90, 120 s (T20); 10, 30, 60, 90, 120, 240 s (T15, T12); 30, 60, 120, 240, 360, 600 s (T9, T6). Data were fitted to a smooth fit and standard errors are indicated. In the figure on the right the first three (T15, T20) or the first four (T6, T9 and T12) data points were fitted to a linear fit to obtain a comparable rate. The slope represented the k_{cat} (% DNA unwound s^{-1}) and was calculated in pM DNA unwound s^{-1}: T6, 20; T9, 30; T12, 50; T15, 140; T20, 180.

As seen from Figure 3.19, the XPD helicase was found to be active with all DNA constructs used. Although, the rates of duplex unwinding decreased with decreasing lengths of the 5'-overhang. Six nucleotides were thought to be too short to load the enzyme properly onto the single strand overhang, however, helicase activity was observed. This could be most likely explained by the fact that monomeric helicases

possess a single and a double stranded DNA binding domain. Engagement of the duplex region may somewhat stabilize the enzyme on the DNA substrate, the single stranded binding domain would still be able to unwind DNA with lower rates in the beginning, increasing to faster rates once the extending single strand is long enough to engage it appropriately. Therefore, higher rates with longer single strand overhangs could be explained with a higher affinity and more stable binding of the enzyme to the DNA substrate. But lower initial rates with shorter 5'-overhangs should lead to a sigmoidal curve, which was not the case (Figure 3.19). Another explanation for the increased rates of duplex unwinding seen when the 5'-overhang is longer could also be due to a cooperativity of XPD (3.5.2). This would mean that, the more XPD molecules accumulate on a DNA strand, the faster the unwinding reaction becomes. Thermal fraying could also be a reason for the observed helicase activity on shorter overhangs, leading to a longer 5' ss-overhang for enzyme loading.

3.7 UNWINDING OF SUBSTRATES CONTAINING A LESION

The SF2 RNA helicase NPH-II was shown to have reduced unwinding activity if there was an alteration or a discontinuity in the ribose-phosphate backbone of the loading strand. Modifications in the displaced strand did not have an effect on enzyme activity (Kawaoka et al. (2004)).

Similar experiments were carried out to investigate if a lesion (fluorescein) in the loading or displaced strand has a similar effect on the helicase activity of XPD. A 5' overhang DNA substrate was constructed from a 45mer and a 60mer oligonucleotide (see Appendix 1, Table 2 for oligonucleotide sequences). A fluorescein label was attached to either of the single strands and annealed with an intact complementary strand.

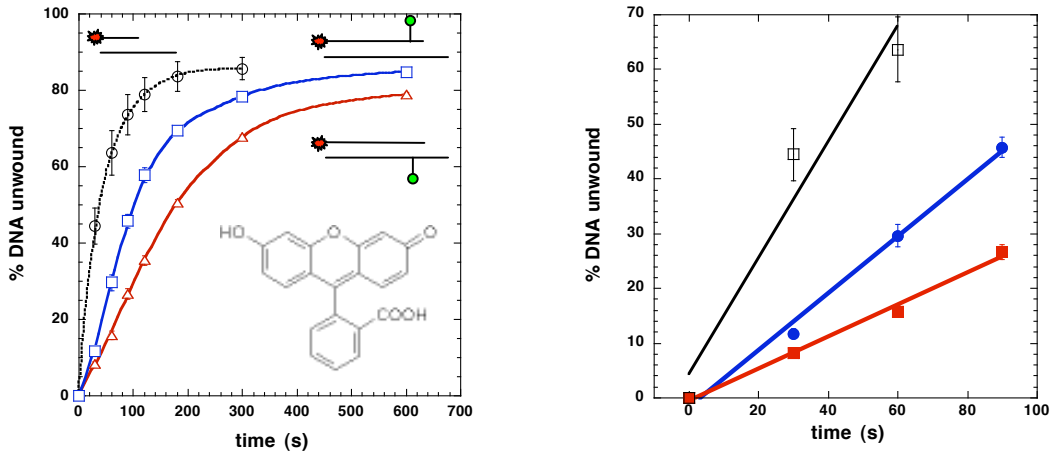


Figure 3.20: Examining the unwinding activity of XPD on fluorescein-labelled DNA strands

Helicase activity of XPD on constructs with a fluorescein label on the displaced (45mer, blue) or loading strand (60mer, red) are shown. The 45mer single strand was ^{32}P-5'-radiolabelled. As a comparison, displacement of 25 nt oligonucleotide without fluorescein is shown (control, black). The figure on the left shows data fitted to smooth curves and the chemical structure of fluorescein is presented on the bottom right of the figure. In the right figure the first four (45mer (blue) or 60mer (red)) or the first three (control, black) data points were fitted to a linear fit. The slope represented the k_{cat} (% DNA unwound s^{-1}) and was calculated in pM DNA unwound s^{-1}: 45mer, 50; 60mer, 30; control, 110. Data points represent an average from triplicate experiments and standard errors are indicated.

The results shown from Figure 3.20 indicate that XPD does not stall at a fluorescein lesion attached to either strand of the DNA substrate. It was not possible to synthesize the respective control DNA substrate without fluorescein lesion (the 45mer annealed 3' to the 60mer oligonucleotide without fluorescein attached to either of the two DNA strands). Instead, the displacement of a shorter 25mer annealed 3' to a 50 nt oligonucleotide (oligonucleotides B1-25 and B50, see Appendix 1, Table 2) is shown in the graph as a comparison. A small difference was observed between the two constructs with a fluorescein lesion. Unwinding seems to be slower if the lesion is on the loading strand (30 pM s^{-1} compared to 50 pM s^{-1} in the displaced strand, Figure 3.20), indicating a similar, but attenuated effect as shown for NPH-II.

These data are still preliminary and it may be useful to repeat them using an appropriate control substrate as a comparison. Furthermore, other DNA substrates have to

be tested to investigate if XPD interacts directly with the bases of the DNA or the DNA backbone (e.g. using DNA substrates containing abasic sites or base modifications in either the leading or displaced strand). It also remains possible that XPD simply bypasses damage.

3.8 TRANSLOCATION – DISPLACEMENT OF STREPTAVIDIN

The ability of XPD to translocate along DNA was tested by using a streptavidin displacement assay. Streptavidin and biotin form one of the strongest protein-protein interactions ($K_D = 10^{-15}$ M) known in nature. The DNA substrate was ^{32}P-5'-radiolabelled and a biotin was attached to its 3' end (Figure 3.21A). Streptavidin was prebound to biotin and displacement monitored by gel shifting as described in Byrd & Raney (2004).

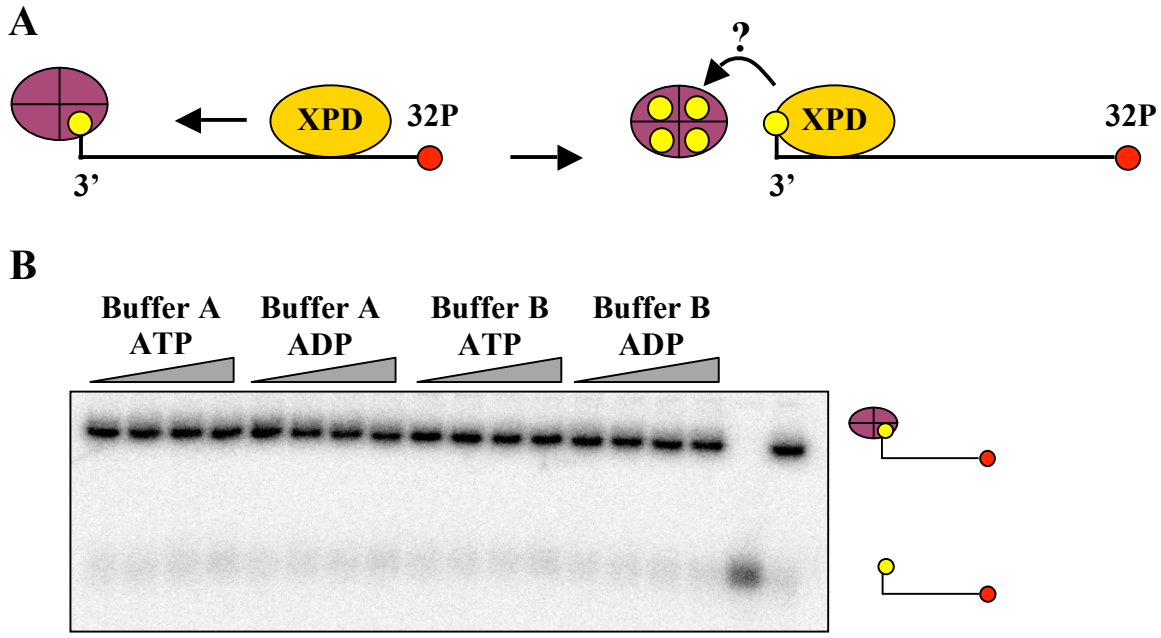


Figure 3.21: Streptavidin displacement by XPD

(A) Scheme of streptavidin removal: Streptavidin was prebound to the 3' end of a biotin labelled oligonucleotide (^{32}P-5'-radiolabelled). If XPD displaces streptavidin, the displaced molecule will immediately bind free biotin, which is in excess in the buffer solution, to avoid rebinding to the biotin attached to the oligonucleotide. (B) Representative 12 % acrylamide gel of the displacement assay. The protein concentration was 600 nM and assay temperatures 55 ˚C. Buffer A contained 50 mM KCl, while buffer B did not contain any salt. Time points were 10, 30, 60 and 120 s.

Initially, this assay was intended to visualize translocation of XPD along ssDNA, but no streptavidin displacement was observed. This does not rule out that XPD does translocate along ssDNA. As discovered in later DNA binding experiments, the affinity of XPD to DNA is very weak. Therefore, this experiment was not optimal to show translocation because XPD would not be associated strongly enough with the DNA to even displace streptavidin. However, the experiment showed that XPD does not remove proteins that are tightly associated with DNA, which might include proteins such as SSB or Alba in *Sulfolobus* (Cubeddu & White (2005); Jelinska et al. (2005)), for example. The removal of DNA-bound proteins during NER is probably completed before the repair machinery assembles. Hence, there would be no need for XPD to have this function. There is still the possibility that XPD is able to remove molecules, which bind to DNA with lower affinities, but this requires further investigation.

3.9 INFLUENCE OF TEMPERATURE ON THE ATPASE ACTIVITY

All helicases are also NTPases and they transfer the free energy derived from NTP hydrolysis directly into translocation along nucleic acid. The next chapter will address the characterisation of ATPase activity of wildtype XPD in more detail; it will also describe that XPD shows single stranded DNA stimulated ATPase activity.

Temperature dependence assays were carried out using 50 nM WT protein and a 70mer single stranded oligonucleotide (see Appendix 1, Table 3) was included in the reaction mix to stimulate ATPase activity. Samples and the ATP/ $MgCl_2$ stock solution were incubated 2 min at the indicated temperatures prior to the initiation of the reactions. After addition of ATP/ $MgCl_2$ only one time point was taken after 2 min incubation. Control reactions without protein were run in parallel and absorbance readings subtracted as background from the respective samples with protein.

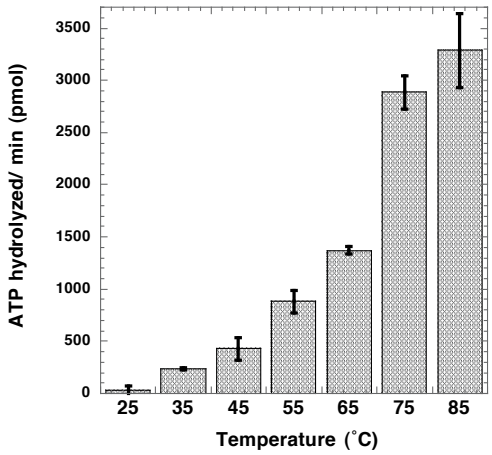


Figure 3.22: Temperature dependence of wildtype XPD ATPase activity

The graph shows pmol of ATP hydrolysed per minute by XPD at different temperatures. Data from three independent experiments were averaged and standard errors are indicated.

As shown in Figure 3.22 the ATPase activity of XPD increased with increasing temperature, as expected for an enzyme from a hyperthermophilic organism. The activity approximately doubles every 10 ˚C up to 75 ˚C, which is the optimum growth temperature of *S. acidocaldarius*. Rates at 75 ˚C were determined as approximately 2900 pmol ATP min^{-1} ([19 moles ATP/ mol XPD] sec^{-1}). The enzyme activity observed at 85 ˚C probably represents an equilibrium between active XPD and heat-inactivated enzyme. For the duration of the experiments, the enzyme appeared stable and no aggregation was observed. Enzyme activities observed at temperatures as high as 85 ˚C also confirmed that the ATPase activities measured can be solely attributed to XPD and did not originate from a contaminating highly active *E. coli* helicase.

3.10 SUMMARY AND CONCLUSION

This chapter described the characterisation of the wildtype XPD protein from *S. acidocaldarius*, a hyperthermophile crenarchaeon. In contrast to hXPD, SacXPD is a monomer in solution and does not form stable complexes with other proteins. Like its eukaryotic counterpart SacXPD is an ATP-dependent DNA helicase with a 5'→3' polarity. As a monomeric helicase, XPD would follow the inchworm model as the proposed mechanism for nucleic acid unwinding (reviewed in e.g.: Soultanas & Wigley (2000)).

This chapter showed that the *in vitro* helicase activity of XPD was concentration dependent. Duplex unwinding was only observed with XPD in at least 10-fold excess over DNA. One possible explanation is that XPD has a low affinity for DNA. Another explanation could be positive cooperativity, as described for the Dda helicase T4 (Byrd & Raney (2004)) or a combination of both.

XPD contains the seven conserved helicase motifs and a highly conserved DEAH motif in the Walker B box, classifying it as a SF2 helicase (Gorbalenya & Koonin (1993); Rudolf et al. (2006)). This study presents XPD as the first helicase described containing an iron-sulphur (FeS) cluster. Elemental analysis and EPR identified the cluster to be of the 3Fe-4S type and the FeS cluster domain is located near the N-terminus of the protein between Walker A and B motif (Figure 3.12). Analysis of sequence alignments revealed that the cluster may be present in the related *E. coli* helicase DinG (see Chapter Introduction), hXPD and other related human helicases (Figure 3.12; Rudolf et al. (2006)). Preliminary experimental data suggested that DinG has indeed a FeS cluster bound (R. D. Camerini-Otero, personal communication).

Iron-sulphur clusters usually occur in a wide range of redox active proteins, involved in rapid electron tunnelling, for example in photosystem I and II or the nitrogenase complex. Other enzymes use the high affinity of iron for oxygen as sensor to modulate oxygen levels within the cell. They induce downstream regulatory processes or they are directly involved in regulation of gene transcription. Other iron proteins take part in iron metabolism (transport, storage, regulation and FeS cluster assembly) or iron can be used to bind substrate to an enzyme (e.g. aconitase). Furthermore, some FeS clusters fulfil exclusively structural roles (for reviews see: Rouault & Klausner (1996); Sicht &

Roesch (1998); Beinert & Kiley (1999); Rees (2002); Kiley & Beinert (2003)). It was shown in this chapter that the FeS cluster of XPD is redox sensitive, as oxidant treatment (ferricyanide) abolishes helicase activity. This is the first evidence for a possible function of the FeS cluster, which will be further investigated in the following chapter.

Streptavidin displacement assays showed that XPD does not remove proteins that are tightly associated with the DNA. XPD was able to bypass lesions (fluorescein) present in either of the two strands of a DNA duplex, although with lower rates if the lesion was located on the loading strand. It was suggested that SF1 and SF2 helicases differ in the way they bind nucleic acids. While SF1 (PcrA) contact the bases directly, it was thought that SF2 (NPH-II, NS3) helicases mainly interact with the sugar-phosphate backbone (reviewed in: Singleton & Wigley (2002)). But only limited biochemical evidence is available yet. The helicase experiments using DNA substrates with a fluorescein molecule either on the loading or displaced strand showed, that an intact desoxyribose-phosphate backbone on the loading strand is important, but not crucial. How DNA is bound by XPD during the duplex unwinding process, however, requires further investigation.

4 PURIFICATION AND CHARACTERISATION OF XPD MUTANTS

4.1 INTRODUCTION

The previous chapter described the characterisation of wildtype XPD from *S. acidocaldarius*, the first helicase described with an iron-sulphur cluster. This cluster may also be present in hXPD and other related human helicases.

Mutations in the *XPD* gene are found in patients with three distinct disorders: *xeroderma pigmentosum* (XP), Cockayne Syndrome (CS) and Trichothiodystrophy (TTD) (Lehmann (2001)). XP-patients show increased UV-sensitivity and are prone to develop skin cancers. CS and TTD patients do not develop skin cancer but show severe photosensitivity. CS is a multisystem disorder and causes mental retardation and skeletal abnormalities, whereas TTD leads to characteristic sulphur-deficient brittle hair, mental retardations and dwarfism. These multiple phenotypes caused by mutations in a single gene are explained by the fact that XPD is involved in NER, transcription and cell cycle regulation.

Mutations in the hXPD protein are mainly found at the C-terminus, which is the domain interacting with the p44 subunit of the transcription factor TFIIH (Figure 4.1; Coin et al. (1998)). Mutations located outside the conserved motifs are likely to influence the helicase activity, which is stimulated by p44, or the general protein stability. Some mutations are also found at the N-terminus, mainly in the conserved motifs. Only one mutation outside the conserved motifs was frequently found in TTD patients (reviewed in Lehmann (2001)). Dubaele et al. (2003) showed that this mutation (R112H) results in a loss of unwinding activity. The reason for the loss of activity was unclear.

This chapter will describe a set of different mutant proteins, which were used to assess the importance of several conserved and non-conserved residues for the enzyme activity of SacXPD. It also sought to find an answer to the question: what is the function of the FeS cluster in a helicase?

The FeS cluster was not previously described for eukaryotic XPD. This was due to the fact that eukaryotic XPD can only be purified in extremely low concentrations. This makes it impossible to detect the FeS cluster visually, because this was how the FeS cluster in SacXPD was initially discovered and a colour is only visible if the protein is at high concentrations. Furthermore, apart from the conserved cluster coordinating cysteines

there are not many conserved residues within the FeS cluster domain. The spacing between the cysteines also differs within the different helicases identified with a FeS cluster. Therefore, it is difficult to detect the FeS cluster domain using bioinformatics.

Specific mutations were made in the yeast Rad3p orthologue of hXPD. These mutations were made on the basis of mutations introduced in SacXPD and specific expected phenotypes proved the existence of the FeS cluster in eukaryotic XPD.

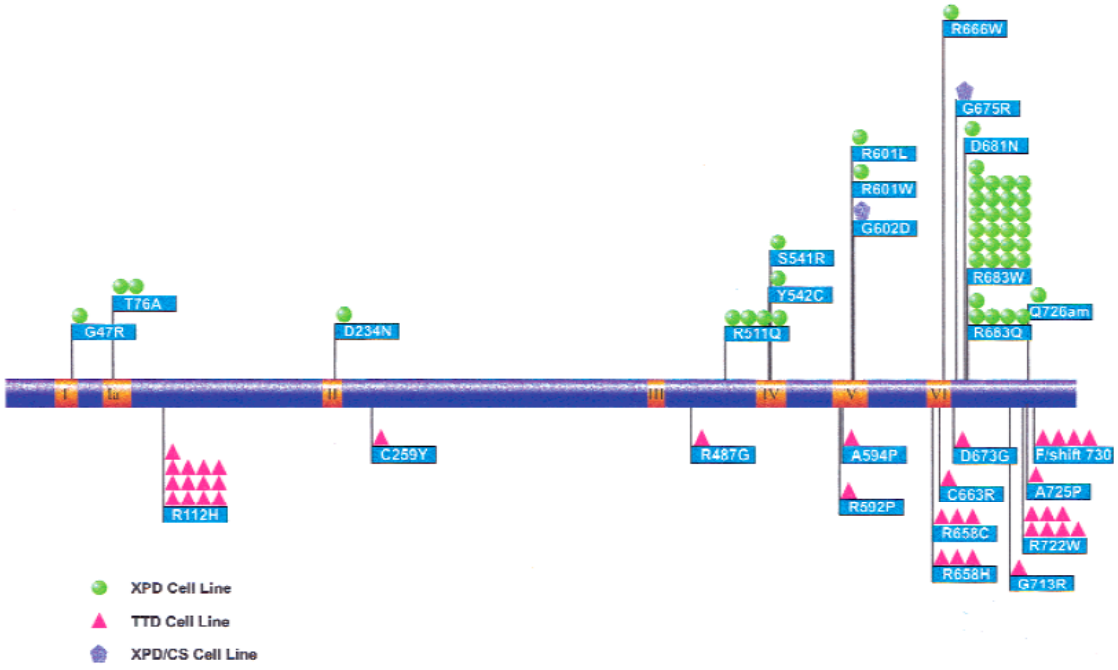


Figure 4.1: Common mutations in the human XPD gene

The schematic shows the domain organisation of hXPD (blue) with the seven conserved helicase motifs in orange. The flags indicate amino acid changes found in different patients and the number of symbols equals the number of times a particular allele has been found mutated in patients (adapted from Lehmann (2003)).

4.2 SELECTION OF MUTANT PROTEINS

For characterisation of the enzyme activity and the function of the FeS cluster, a set of mutant proteins was generated by site directed mutagenesis. Residues were selected with regard to a possible influence on the FeS cluster formation. Other mutations were introduced in SacXPD, because the analogous amino acid modification in human helicases, such as XPD and FancJ, were found to cause certain diseases (Figure 4.2).

Lysine 35 is located in the Walker A box and mutation to an alanine should abolish the ATPase activity (Sung et al. (1988); Velankar et al. (1998)). This mutant was used as a negative control. The four conserved cysteine residues (C88, C102, C105, C137), which were thought to be involved in FeS cluster coordination, were mutated to serines. The importance of conserved hydrophobic residues around the FeS cluster was investigated by generating the mutants Y139F and Y140F in addition to the two double mutants Y139A/Y140A and Y156A/Y158A. Hydrophobic residues were found to surround FeS clusters and to have important roles in protecting the FeS cluster from the access of solvents and oxidising agents and to stabilise the cluster through hydrogen bonding of the aromatic ring with the cluster (Iwagami et al. (1995)). These residues are also involved in the maintenance of an anaerobic environment around the cluster and may have other structurally important features as well. It was described that the prosthetic centre of FeS clusters is sensitive to side chain substitutions. This means, that changing tyrosine to a polar residue would lead to instability, while a change to an unpolar residue, especially with an aromatic side chain, would have fewer consequences and is likely to maintain the stability of the cluster (Agarwal et al. (1995)).

The remaining mutants that were engineered represented human mutations, which were found in XP-D (T56A, K84H, G447D, C523R, R531W) or Fanconi anemia patients (FancJ; SacXPD: F136P) (Table 4.1). A truncated version of the protein was generated, which only contained the FeS cluster domain from residues Glu72 to Asp162.

Other mutant proteins were made by B. Petrovic-Stojanovska (CBMS, University of St Andrews) for spin-labelling experiments. All non-essential cysteines were changed to serines generating mutant XPD-3CS (XPD-C360S/C523S/C543S). This mutant was used to introduce the following amino acid substitutions: F38C, N58C, D186C and

A511C. The enzymatic activity was tested, but the phenotypes were not further characterised.

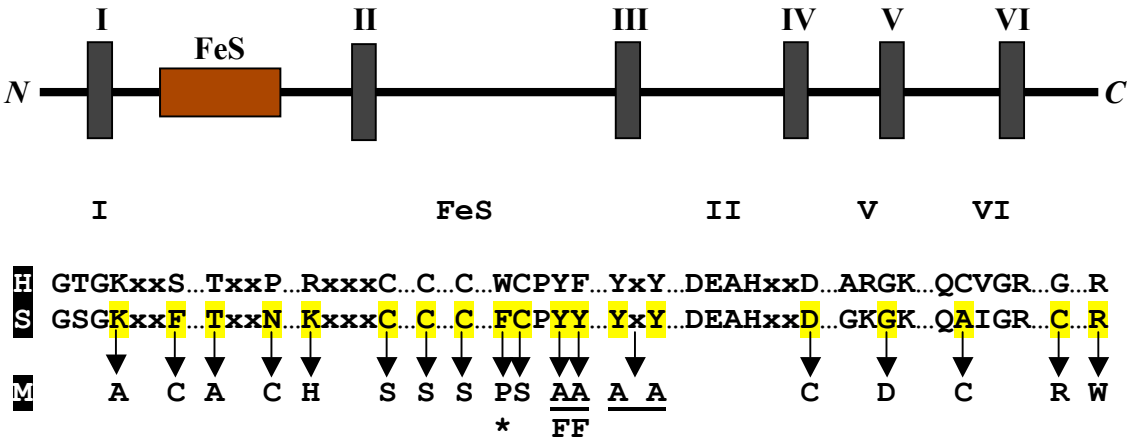


Figure 4.2: Overview of mutations introduced into *S. acidocaldarius* XPD

In the top panel a schematic of the XPD gene is shown, including the six main helicase motifs (labelled I-VI) and the FeS cluster domain. The alignment in the bottom panel shows the abbreviated XPD sequences of the human (H) compared to the *S. acidocaldarius* (S) protein. Residues that were selected for mutation in SacXPD are highlighted in yellow and the residues they were mutated to are shown in the bottom line (M). Double mutants are underlined and (*) indicates the position of the FancJ mutation A349P (the equivalent residue in human FancJ is an alanine).

4.3 PURIFICATION OF MUTANT PROTEINS

Mutants were purified as described for the wildtype protein. The mutations and the whole protein mass were confirmed by mass spectrometry. Three different species were observed and are summarized in Table 4.1. Class I mutants were devoid of the FeS cluster and completely colourless on purification. Class II mutants expressed with FeS cluster, but lost it during purification or upon storage. A fading of the initial wildtype colour could be followed by eye, until the protein solution was completely colourless. These mutants are referred to as mutants containing an unstable FeS cluster. Class III mutants bound a stable cluster and showed wildtype colour. T56A was not characterised, because the recombinant protein could not be expressed.

In general, class II and III mutants purified as wildtype and eluted from the heparin column at ~380 mM NaCl. Only the R531W mutant showed very weak binding and eluted at around 260 mM NaCl. This could suggest that this mutation influences DNA binding. Class I mutants showed lower expression, but bound tighter to the affinity column and eluted at ~420 mM NaCl. The only exception from that scheme was the C523R mutant which, although possessing a stable FeS cluster, eluted at ~420 mM NaCl from the heparin column. The histidine tagged truncated mutant showed extremely low expression and from the absence of colour it can be concluded that no cluster was bound (data not shown). Therefore, this mutant was not characterised further. All proteins purified as monomers by gel filtration (GF) chromatography.

Table 4.1: Characteristics of XPD mutant proteins

Mutation in *S. acidocaldarius*	Mutation in Human	Yield compared to WT (%)	Cluster?
K35A	-	100	stable
T56A	T76A (XP)	0	-
K84H	R112H (TTD)	100	unstable
C88S	-	30	non
C102S	-	100	stable
C105S	-	30	non
F136P	A349P (FA)	100	unstable
C137S	-	30	non
Y139F	-	100	stable
Y140F	-	100	stable
Y139A/ Y140A	-	100	unstable
Y156A/ Y158A	-	100	stable
G447D	G602D (XP/ CS)	100	stable
C523R	G675R (XP/ CS)	100	stable
R531W	R683W (XP)	30	stable
XPD-trunc	-	10	non

Table 4.1 shows mutated residues of *S. acidocaldarius* XPD and the equivalent mutations in human, when the mutation was selected because of clinical relevance in humans. The clinical symptoms in humans are labelled: XP, *xeroderma pigmentosum*; CS, Cockayne Syndrome; TTD, trichothiodystrophy; FA, Fanconi anemia. Purification results, presence and stability of the FeS clusters are also indicated.

4.4 PURITY AND FOLDING

The purity of XPD mutants was analysed by SDS-PAGE. The folding of selected proteins was studied by circular dichroism (CD) spectroscopy. This included one example of a class I mutant (C105S), two class II mutants (K84H, F136P), which represent human mutations and a class III mutant (C102S). The latter was of particular interest, because it was the only one out of four cysteine mutants that possessed a stable

FeS cluster. Differences in the FeS cluster due to change of ligands could induce changes in the overall folding of the protein. The wildtype protein in Figure 4.3A (right gel) shows a band at ~120 kDa, which may be a non-specific protein carried over during protein purification.

The proteins appeared pure after gel filtration chromatography. Breakdown products seen for some of the mutants (Figure 4.3A) appeared typically upon thawing after long-term storage at -20 ˚C. These bands were also seen after tryptic digestion and varied between mutants with and without cluster. The formation of these different proteolytically sensitive fragments indicates different domain organisation between cluster and "clusterless" proteins, which will be discussed in Chapter 5.

Far UV-CD spectroscopy confirmed that all mutants maintained the overall secondary structure (Figure 4.3B). Minor changes in the spectrum of the C105S mutant enzyme reflects local perturbations in the Fe-S domain, but suggested that the mutant enzyme adopts a stable folded structure. A general observation was, that the K84H mutant showed an increased instability after the complete loss of the FeS cluster leading to precipitation. In contrast, the F136P mutant appeared stable. These differences in stability led to small differences in the far UV-CD spectra of both mutants as seen in Figure 4.3B.

It seemed that in general the FeS cluster had no crucial function in protein folding or structural stability, because all proteins could be purified and were folded. Problems with expression, purification or stability were most likely brought about by the mutation itself and not through a loss or instability of the cluster. As seen from the C105S mutant, the cluster may have a very localized structural influence.

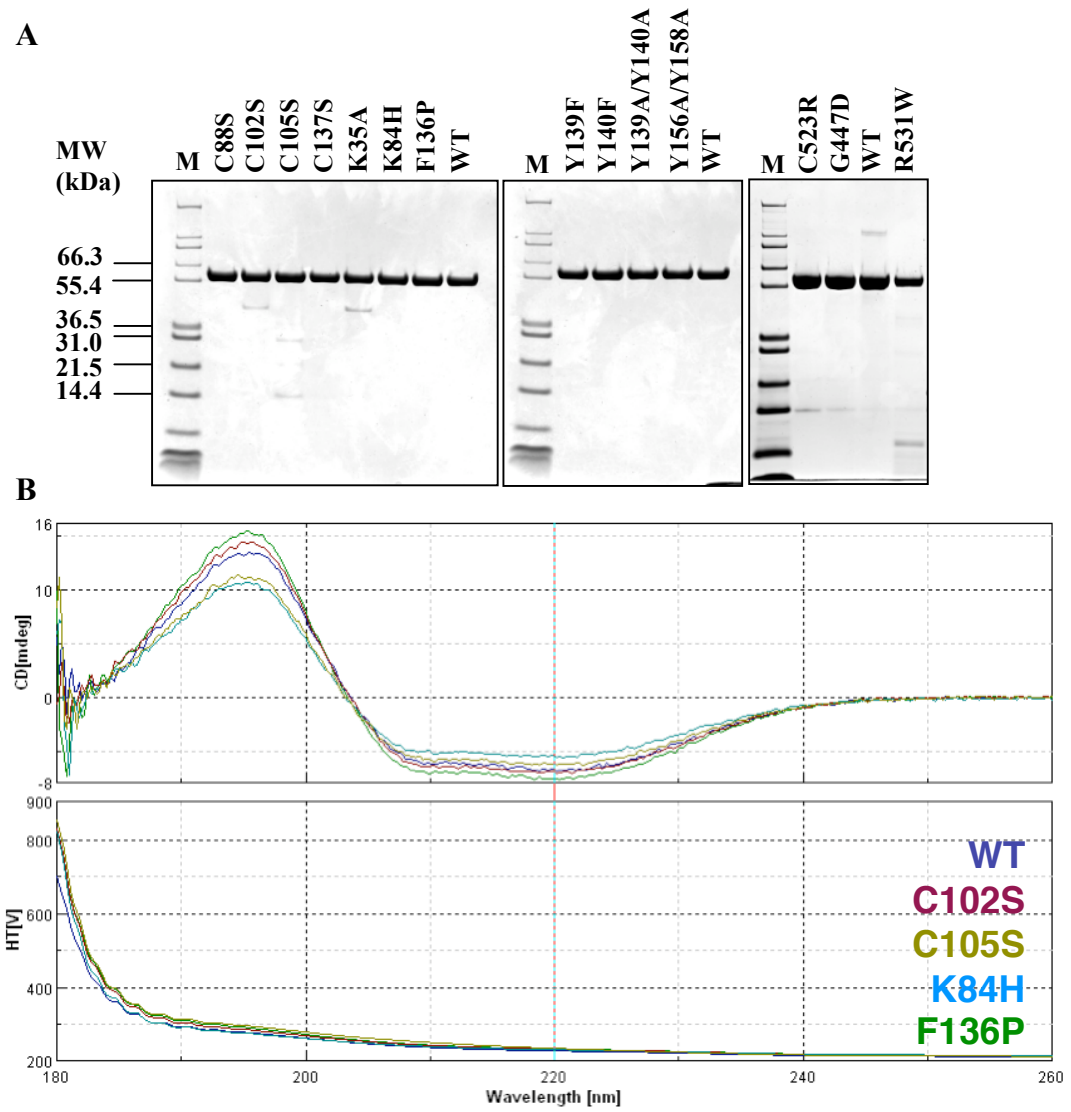


Figure 4.3: Mutant protein purification and analysis

SDS-gels of purified proteins after gel filtration chromatography are shown in (A). The Mark12™ Standard Marker (Invitrogen) is indicated with M. CD spectroscopy analysis are shown in panel (B). Colour scheme for WT and mutant proteins are indicated in the figure.

4.5 IRON QUANTIFICATION

From the purification of XPD mutant proteins, it became evident, that only three of the four conserved cysteines may be involved in FeS cluster coordination. Although WT XPD was described in the previous chapter to bind a 3Fe-4S cluster, there was still the possibility that in position 102 a serine could substitute for a cysteine, as was found in other cases (Kowal et al. (1995)). Indeed, in the amino acid sequence of *Thermoplasma volcanium* XPD, a serine instead of a cysteine is present at the equivalent position of *S. acidocaldarius* C102 (Figure 4.4).

To confirm the presence of three iron ions, the iron content was determined for wildtype (control), C102S and C105S mutants by using the bathophenanthroline method. Experiments were carried out in triplicate and the average absorbance used to calculate the numbers of irons bound to 1 molecule of XPD. As expected, no iron was detected for the C105S mutant (A_{535}: 0.0005), which was colourless on purification. The C102S (A_{535}: 0.129 = 2.85 Fe/XPD) and the wildtype (A_{535}: 0.1482 = 3.27 Fe/ XPD) protein both lead to a ratio of 3 iron bound to 1 molecule XPD.

This experiment proved that colourless mutant proteins did not have iron bound, while coloured mutant proteins possessed a 3Fe-4S cluster. Serine does not substitute for a cysteine in position 102 and no iron ion is coordinated by C102. Only three cysteines are essential for coordination of the cluster, consistent with spectroscopic and elemental analysis carried out on wildtype protein. The 3Fe-4S cluster was confirmed for the wildtype XPD by EPR and ICP-OES. Also, FeS cluster reconstitution strengthened the hypothesis that XPD has only a 3 Fe not a 4 Fe cluster bound (Chapter 3).

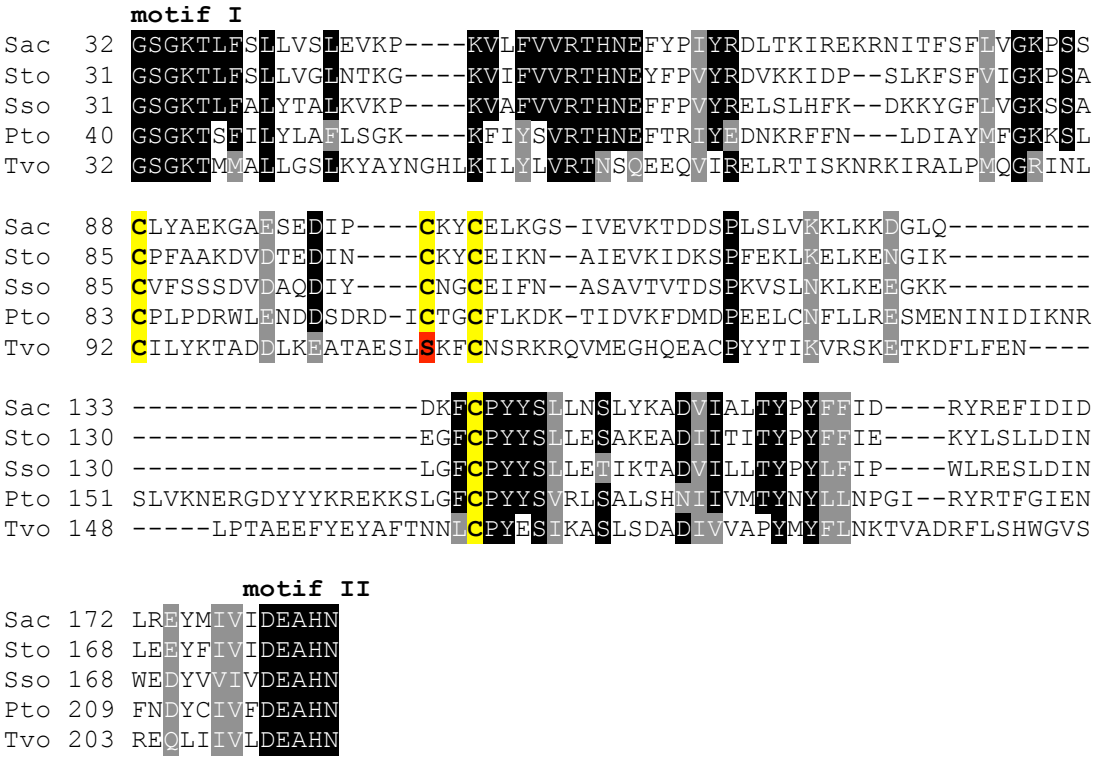


Figure 4.4: Sequence alignment of XPD homologues from different archaea

Amino acid sequence alignment of XPD proteins (from helicase motif I to II) from different archaea. Highly conserved residues are highlighted black and conserved similar residues are highlighted dark grey. The four conserved cysteine residues are highlighted in yellow, and the serine substitution at the equivalent position of SacXPD-C102 is highlighted in red. Sac, *S. acidocaldarius*; Sto, *S. tokodaii*; Sso; *S. solfataricus*; Pto, *Picrophilus torridus*; Tvo, *Thermoplasma volcanicum*.

4.6 FUNCTION OF THE FeS CLUSTER DOMAIN

4.6.1 Helicase activity

The influence of the FeS cluster on the helicase activity was investigated using the four cysteine mutants (C88S, C102S, C105S and C137S) and the four tyrosine mutants (Y139F, Y140F, Y139A/ Y140A and Y157A/ Y159A) (4.2). Duplex unwinding activities were compared to the WT enzyme. The K35A mutant served as negative control, because it should not show helicase activity.

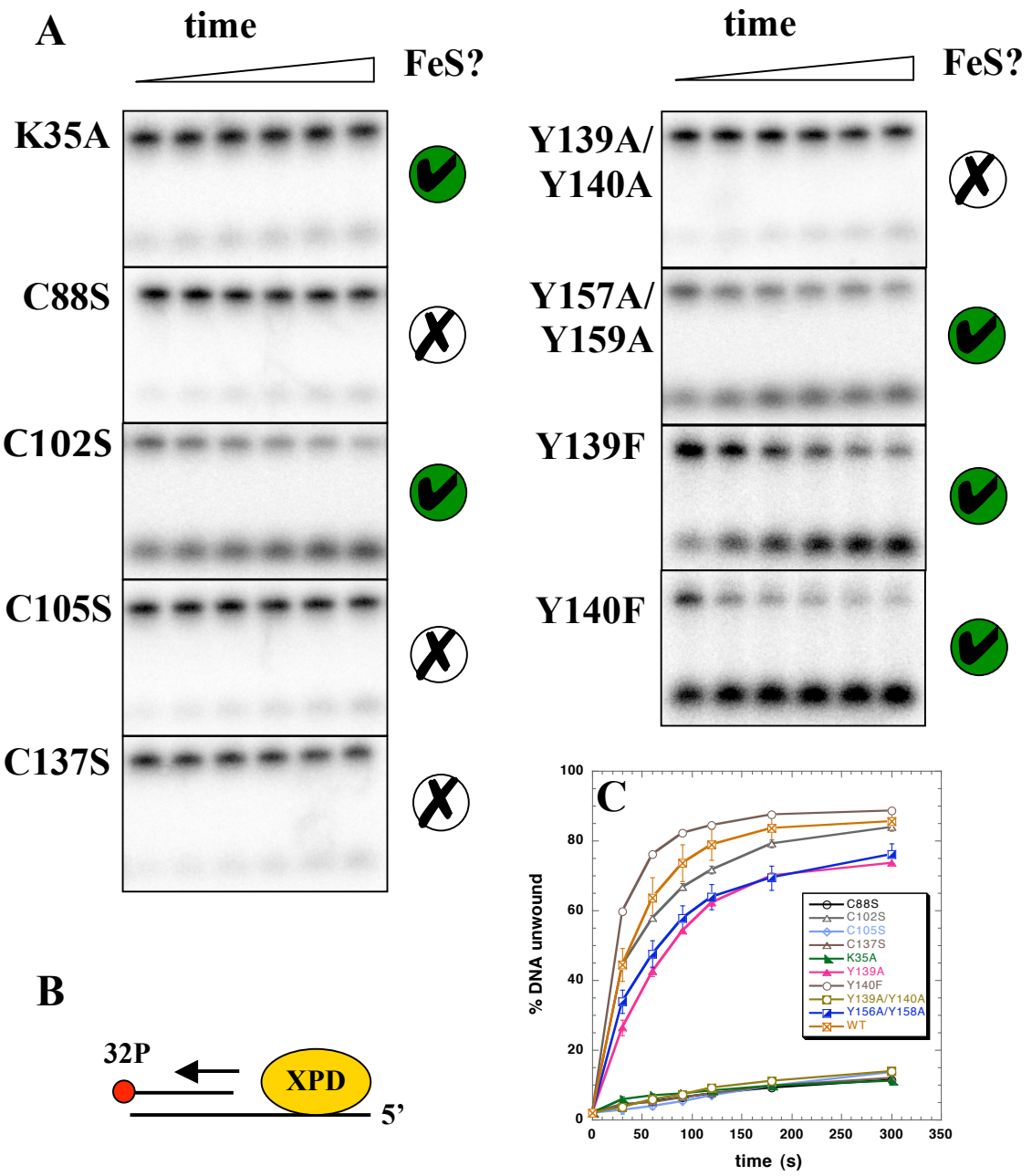


Figure 4.5: Helicase activity of XPD FeS cluster mutants

(A) Representative 12 % native acrylamide gels showing helicase activity of mutant XPD proteins. Upper bands represent dsDNA and lower bands are the ^{32}P-labeled ssDNA. The latter arises if unwinding takes place. Time points were: 30, 60, 90, 120, 180, 300 s. Green circles with a tick indicate the presence of a stable FeS cluster, while white circles and a cross indicate the absence of the FeS cluster. (B) The schematic shows the setup of the helicase assays. (C) Quantification results of triplicate experiments are shown and standard errors are indicated. The zero time point was given a value of DNA unwinding = zero. WT and K35A mutants are shown as positive and negative controls, respectively.

The results of the helicase assays using N-terminal mutant enzymes are shown in Figure 4.5. The Walker A box mutant K35A had a stable cluster, but was inactive due to an inability to hydrolyse ATP. This mutant was used as a negative control. The side chain of this invariant lysine residue at this position is located at the same position as the magnesium ion after binding of the ATP-Mg^{2+} complex. It was found that this side chain binds the β-phosphoryl group of ATP and is thought to stabilize the transition state during hydrolysis (Velankar et al. (1999)). Mutation of this residue leads to enzyme inactivation.

The three cysteine mutants (C88S, C105S, C137S), which did not coordinate a FeS cluster, did not show helicase activity. In contrast, the C102S, which retained a stable cluster showed wildtype helicase activity.

The FeS cluster was destabilised in the double mutant Y139A/Y140A. Single mutations of these two residues to phenylalanine did not show an equivalent effect, although, a slight decrease in general protein stability was observed with the Y139F compared to the Y140F mutant. Both residues are highly conserved, even if Y140 is conserved either as a tyrosine or as a phenylalanine. Structurally, there is not much difference between tyrosine and phenylalanine and to maintain the hydrophobic environment around the cluster, Phe could fulfil this function just as well. To gain a more profound effect, both residues have to be mutated to a different amino acid, such as alanine. The Y140F mutant, which is a stable protein containing a stable FeS cluster, showed wildtype activity (Figure 4.5). The mutation in position 139 to Phe decreased the stability of the protein, but did not directly influence that of the FeS cluster. The unwinding experiments mirrored this effect, leading to a reduced activity relative to the wildtype. The double mutant Y139A/Y140A lost the FeS cluster and did not exhibit unwinding activity. Similar to the Y140F mutation, the double mutant Y156A/Y158A, however, did not influence the structure of the FeS cluster, but decreased the stability of the protein. This is reflected in the rate of unwinding analogous to Y139F.

These experiments indicate an essential role of the FeS cluster for the helicase activity of XPD.

4.7 CHARACTERISATION OF HUMAN MUTATIONS

Most of the known human mutations in XPD are located at the C-terminus. Human XPD is a functional bridge between the TFIIH core complex and the CAK subunit. The TFIIH holocomplex is only formed when taking part in transcription, while the TFIIH core, including hXPD, acts in NER. The C-terminus of hXPD interacts with the p44 subunit of the TFIIH core, while the N-terminus binds to the MAT1 subunit of the CAK subcomplex. No mutations have been reported on the amino-terminal end, which would abolish the interaction with MAT1. However, mutations located on the C-terminus of the hXPD influence the binding to p44. Apart from that, three of the seven conserved helicase motifs are located in this area and mutations are likely to affect the ATPase activity of XPD. Lastly, this region may be involved in DNA binding and substrate specificity (Lehmann (2001); Sandrock & Egly (2001), Chen & Suter (2003)). Some of the C-terminal mutants (G447D, C523R, R531W) were purified and characterised by M. J. Stacey, an Honours student.

Fanconi anemia is a recessively inherited disorder leading to chromosomal instability and subsequently to developmental abnormalities, bone marrow failure and increased risk of cancer. The Fanconi anemia pathway is a fairly recently discovered DNA repair mechanism and not very much is known about the factors involved and the mechanism of action. The gene corresponding to FancJ was only discovered recently (Levitus et al. (2005), Levran et al. (2005)) and turned out to be an already known helicase called BACH1/ BRIP1, which interacts with BRCA1 (Cantor et al. (2001); see Chapter 1).

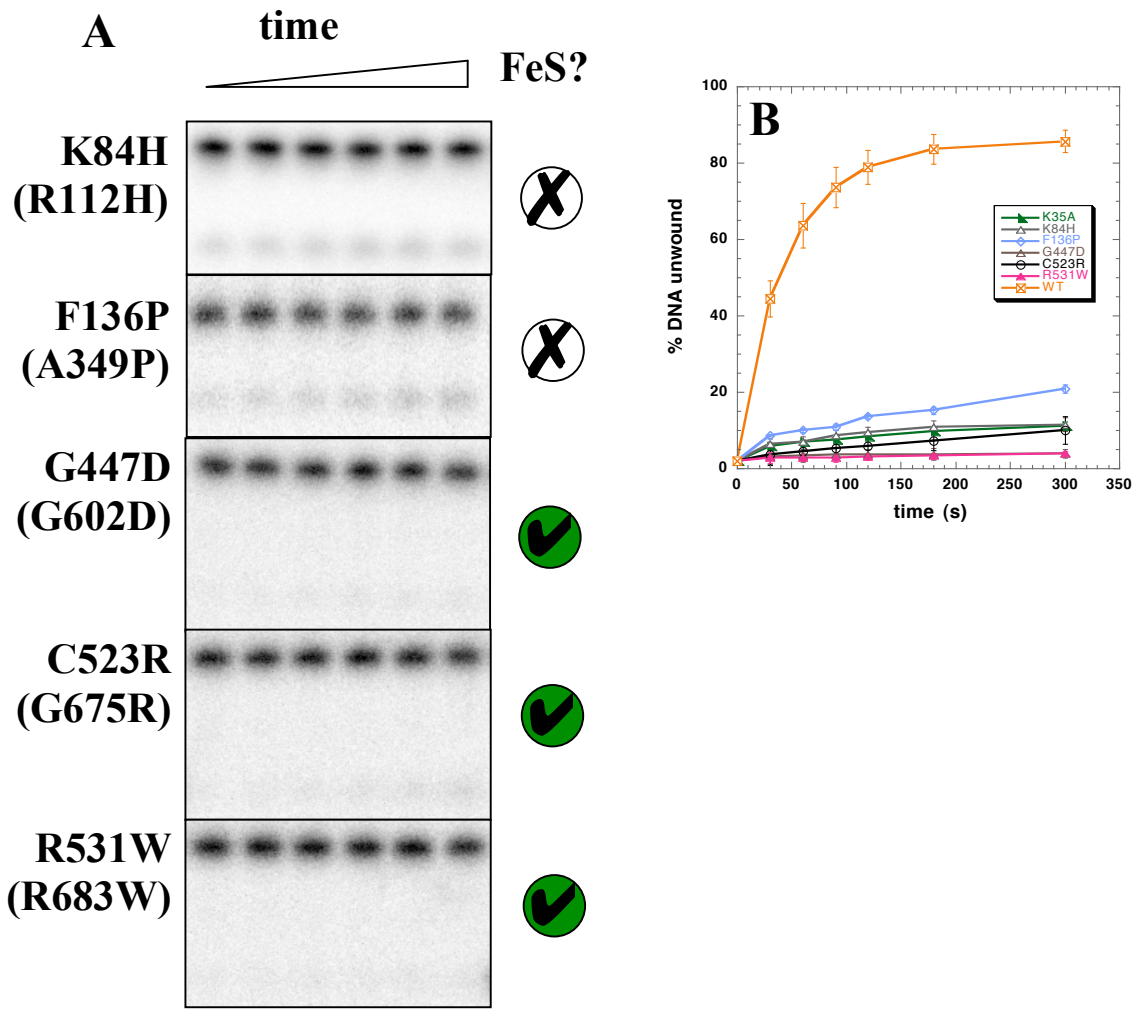


Figure 4.6: Helicase characterisation of mutants resembling human mutations

(A) Representative 12 % acrylamide gels showing helicase activity of mutant proteins. Upper bands represent dsDNA and lower bands are ssDNA. The latter arises if unwinding takes place. Time points were: 30, 60, 90, 120, 180, 300 s. Green circles indicate the presence of a stable FeS cluster, while white circles indicate the absence of the FeS cluster. (B) Quantification results of triplicate experiments are shown and standard errors are indicated. Data points were connected by lines. The zero time point was given a value of DNA unwinding = zero. WT and K35A mutants are shown as positive and negative controls, respectively.

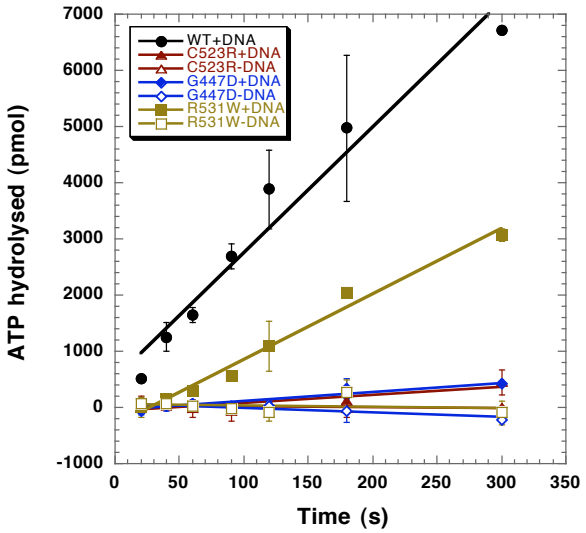
The two human mutations located at the amino-terminus of XPD were frequently found in TTD (R112H) and FancJ (A349P) patients, respectively. Although, they are not located in one of the conserved motifs, these two mutants represent inactive helicases (Dubaele et al. (2003); Levran et al. (2005)). The mutants K84H (R112H) and F136P (A349P) of SacXPD led to disruption of the FeS cluster and subsequent enzyme inactivation (Figure 4.6). Dubaele et al. (2003) found that the R112H mutant protein shows no helicase activity and was therefore deficient in *in vitro* reconstituted NER. In addition hXPD-R112H showed partial inhibition of transcription, but exhibited wildtype interaction with the TFIIH holocomplex. The mechanism of helicase inactivation, however, remained unclear. Furthermore, Berneburg et al. (2000a) showed that human cell lines harbouring the R112H XPD mutation resemble phenotypes which are closer related to phenotypes observed with XP-cell lines, rather than other TTD cell lines. They found decreased accumulation of incision products during NER and lower rates of CPD removal, similar to other XP-cell lines tested. In addition, in vitro experiments using a UV-sensitive *S. pombe* rad15.P strain, rad15-R111H (human: R112H) managed to restore UV-sensitivity only partially, in contrast to TTD-cell lines where almost full restoration was observed. An explanation for these differences could not be provided. The discovery that SacXPD contains an essential FeS cluster, which is also present in the human homologue, provided a possible explanation for the effect this mutation has on the enzyme activity. The same applied for the FancJ mutation A349P (Levran et al. (2005)). This residue is not conserved at all, however it is located prior to the last cysteine of the FeS cluster domain. It was found in patients with severe symptoms of Fanconi anemia. An amino acid change to a rigid proline can induce structural alterations and destabilise the cluster. Indeed, the F136P mutant of XPD possessed an instable FeS cluster. The loss of the FeS cluster was associated with a loss of the helicase activity of F136P (Figure 4.6).

The remaining human mutations are situated at the C-terminus of XPD, where most of the human mutations are located. All three mutants (G447D, C523R, R531W) showed abolished helicase activity (Figure 4.6). The residue G447 (human: G602) is located in motif V. The affected patient was identified to be hemizygote for the G602D mutation and showed symptoms of XP and CS (Broughton et al. (1995)). The human mutant

protein was shown to interact with p44, but represented a deficient helicase leading to a deficiency in NER (Berneburg et al. (2000b); Dubaele et al. (2003)). In accordance with these findings, the mutation G447D in *Sulfolobus* abolished the ATPase activity (Figure 4.7), leading to an inactive helicase (Figure 4.6).

Residue C523 (human: G675) is located C-terminal from motif VI. The respective patient showed a combined XP/ CS phenotype with a hemizygote genetic background for the hXPD-G675R mutation (Broughton et al. (1995)). In contrast to the above mutation, the hXPD-G675R mutation does not affect the helicase activity, but disrupts the interaction with p44. Therefore, the helicase activity is similar to that observed with the unstimulated wildtype protein and the mutant was shown to support low but significant NER *in vitro* while having no influence on transcription initiation (Dubaele et al. (2003)). *In vivo* experiments using a UV-sensitive *S. pombe* strain showed that the analogous mutation in the rad15 protein could not restore UV-sensitivity. Furthermore, hXPD-G675R cells from cell lines of the respective patient were shown to be deficient in 3' incision and hence unable to complete NER (Berneburg et al. (2000a & b)). The analogous mutation in SacXPD, C532R, led to an almost complete loss of ATPase and helicase activity (Figure 4.6 and Figure 4.7), indicating differences in the mechanism of action between SacXPD and human XPD.

XPD patients harbouring the R683W mutation (Sac: R531W) exhibit an XP phenotype. The mutation is located outside the conserved motifs at the C-terminus of the protein. It abolishes the interaction with the p44 subunit of TFIIH, which leads to low levels of NER (Dubaele et al. (2003)). Helicase activity of hXPD-R683W was similar to the basal wildtype activity, but much lower than the p44 stimulated wildtype activity. *In vivo* (*S. pombe*) this mutant partially restored UV-sensitivity in a rad15.P strain (Berneburg et al. (2000a)). The mutation in *S. acidocaldarius* led to an inactive helicase (Figure 4.6). This may be caused by weak DNA interaction (Chapter 5), which caused a half-maximal ATPase activity (Figure 4.7).



	no DNA	ssDNA
G447D	0	1.6
C523R	0	1.4
R531W	0	11.6
WT	-	22.3

Figure 4.7: ATPase activity of the XPD mutants G447D, C523R and R531W

The graph shows results from duplicate experiments for wildtype and mutant enzymes (200 nM) with ssDNA (0.5 nM PhiX175 virion) and without DNA. The ssDNA stimulated ATPase activity of the wildtype enzyme is shown as a control. Time points were: 20, 40, 60, 90, 120, 180, 300 s. Data were fitted to a linear regression and standard errors are indicated. The table summarises the results of the ATPase assays for the wildtype and the three mutant proteins. Numbers indicate ATP hydrolysed/ second in pmol.

4.8 YEAST RAD3 & THE PRESENCE OF THE FES CLUSTER IN EUKARYOTES

Bioinformatic analysis showed a strong conservation of the FeS cluster domain not only within the domain of the archaea, but also in hXPD, other related human helicases and yeast Rad3p (Chapter 3, Figure 1.13). In order to prove the presence of the FeS cluster in an eukaryotic model organism, *Saccharomyces cerevisiae* strains were constructed and characterised by Dr. V. Makrantoni in the laboratory of Prof M. J. Stark (Wellcome Trust Center, University of Dundee).

On the basis of these sequence alignments three amino acid residues were selected in the yeast Rad3p orthologue of hXPD and mutated in analogy to the already characterised mutations in SacXPD. These amino acid substitutions were expected to induce specific degrees of UV-sensitive phenotypes.

The first mutant, Rad3-K48A, represents the control mutant, which is completely inactive due to an inability to hydrolyse ATP (Sung et al. (1988); Velankar et al (1999)).

This mutation is analogous to XPD-K35A, located in motif I and impairs the ATPase activity. Rad3-C115S is the analogue of XPD-C88S. In *S. acidocaldarius* this mutation led to an inability in coordinating the FeS cluster and subsequently to enzyme inactivation. The third mutation was one of the human mutations (R112H), which was found in TTD patients. The analogue in yeast is Rad3-R111 and K84 in *S. acidocaldarius*. Mutation to a histidine was shown to lead to cluster destabilisation and enzyme inactivation of SacXPD (section 4.3 and 4.6.1).

In yeast RAD3 is an essential gene. A RAD3 knockout mutant and two shuttle vectors (YCplac33 and YCplac111) were used to create the Rad3 strains. One copy of the *S. cerevisiae* RAD3 gene in the wildtype diploid strain was deleted, generating a strain containing a *HISMX* marker cassette. This strain was then transformed with an *URA3*-marked plasmid (YCplac33) generating a wildtype RAD3 (YCplac33-RAD3) strain with a selective marker. After sporulation and tetrad dissection, strains containing the *URA3/HIS3* markers were selected. The positive segregants were subsequently transformed with derivatives of a *LEU2*-marker plasmid YCplac111 carrying either wildtype RAD3 or one of the RAD3 mutant genes (K48A, R111H and C115S). Following two rounds of YCplac33-RAD3 counter selection using 5-fluoroorotic acid, six independently-derived isolates of each RAD3 construct dependent solely on the YCplac111-RAD3 plasmids for Rad3p function were identified. All RAD3 genes were fully sequenced. Evaluation of UV-sensitivity was performed by irradiating cells with different doses of UV-light. Following UV-treatment, cells were grown for two days in the dark to avoid photolyase activation (Rudolf et al. (2006)). The results are shown in Figure 4.8.

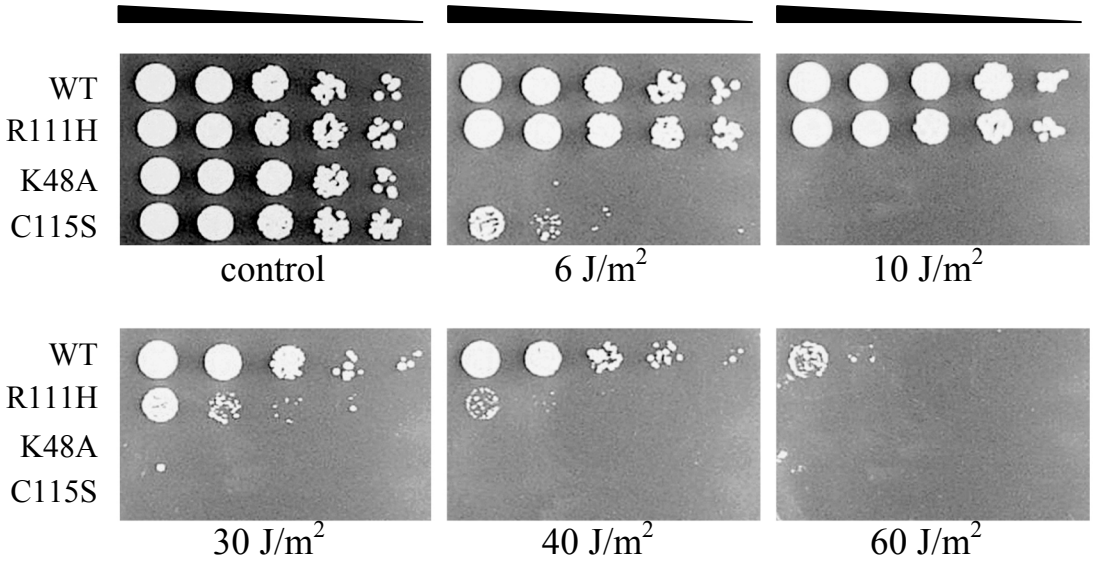


Figure 4.8: Mutations in yeast Rad3p show influence on the FeS cluster and leads to an UV-sensitive phenotype

Each image shows YPD-agar plates with serial dilutions (1:10) of a yeast culture. Plates were irradiated with different doses of UV-light and grown in the dark for 2 days. As the UV dose increases K48A and C115S mutants show highly sensitive UV phenotypes, while R111H is only mildly affected.

The results obtained from these experiments (Figure 4.8) proved the presence of the FeS cluster in yeast *in vivo*. The K48A negative control showed no resistance to UV-light and all cells died after exposure to only 6 J/ m^2 UV-irradiation.

The residue C88 of SacXPD is the first cysteine in the amino acid sequence, which is involved in coordination of the FeS cluster. Mutation to a serine led to cluster loss and subsequently abolished the helicase activity. The analogous mutation in yeast, C115S, caused a severe NER deficiency phenotype with extreme UV-sensitivity. The majority of yeast cells were not able to survive the lowest level of UV exposure (6 J/ m^2) and showed a similar UV-sensitive phenotype as the control mutant K48A, which has an impaired ATPase activity.

The last mutation, Rad3-R111H, is the equivalent of the human R112H mutation, which causes TTD, and K84H in *S. acidocaldarius* XPD. The results obtained in the yeast system showed moderate UV-sensitivity, in accordance with data obtained for the *S. pombe* rad15-R111H mutant, which partially restored UV-sensitivity in a rad15 null

background (see section 4.7; Berneburg et al. (2000a)). In human cell lines from different patients harbouring the R112H mutation, cells showed a range of very mild to extremely severe effects of UV-light sensitivity. Yeast cells exposed to low doses of UV-light can cope relatively well with DNA damage that arises. However, if the number of damages increases with increasing UV-irradiation, the mutated helicase is not proficient enough to support NER and cells die due to accumulation of DNA damage. In *S. acidocaldarius* the cluster is initially there, which means that initially there is an efficient helicase available for DNA damage repair. The cluster also appeared to be stable *in vivo*, because cluster destruction was only observed after purification. It is possible, that the labile bound cluster is lost during the DNA unwinding process, caused by alternating conformational changes of the enzyme. This could lead to delayed damage repair, which only becomes evident after acute DNA damage.

4.9 SPIN-LABELLED MUTANTS

Spin-labels are paramagnetic due to an unpaired electron and can easily be monitored by EPR (see introduction to EPR in previous chapter). The reporter is mostly a nitroxyl radical (the unpaired electron is mainly localized to the N-O bond), which is incorporated in a heterocyclic structure (the label; Figure 4.9). The functional group of the spin-label reacts at neutral pH with thiol-groups in cysteines of the protein generating a covalent bond. Therefore, it is possible to attach spin-labels at every position in the protein by introducing cysteines by site-directed mutagenesis (site-directed spin-label, SDSL). If two or more spin labels are present in a protein structure, the distances between them can be measured by EPR. Since proteins are flexible macromolecules, they change their conformations in response to changes in the environment or in the presence of an interacting partner molecule. These conformational changes can be monitored by measuring the changing distances between SDSL. Therefore, SDSL is a powerful tool to study structure and dynamics of macromolecules (Hubbell et al. (1996)).

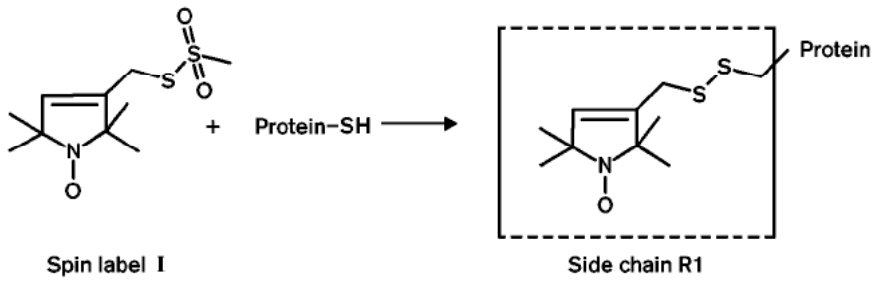


Figure 4.9: Side-directed spin-labeling (SDSL) of proteins

The most common procedure to attach a spin label (nitroxide side chain) to a protein is as follows: Cysteines are specifically introduced into the protein primary structure by site-directed mutagenesis. The nitroxide side chain reacts specifically with the sulfhydryl group of the cysteine and generates a site-directed spin-label attached at a specific position in the protein sequence. The most commonly used reagent is methanethiosulfonate which generates side chain R1 (Figure adapted from Hubbell et al. (1996)).

SDSL was used to observe conformational changes of XPD upon binding of ligands, e.g. ATP or DNA. Distance measurements were carried out relative to the cysteines of the FeS cluster domain. In that respect, all cysteines outside the FeS cluster domain were changed to serines, creating the triple mutant C360S/C523S/C543S (XPD-3CS). This triple mutant was used to introduce cysteines for SDSL at the following positions: F38, N58, D186 and A511. Respective XPD mutants were recombinantly expressed, purified and spin-labelled by B. Petrovic-Stojanovska (CBMS, St Andrews). These mutants were tested for helicase activity and the influence of the SDSL was also investigated (Figure 4.10 and Figure 4.11).

The XPD-3CS mutant was tested for helicase and ATPase activity and showed rates identical to the wildtype (data not shown). Subsequently introduced mutations were tested for their influence on the helicase activity before and after spin-labelling (Figure 4.10 and Figure 4.11). Mutants N58C and D186C showed wildtype activity, while F38C and A511C exhibited abolished and lower activity, respectively. The spin-label seemed not to alter the general properties of the mutants, since there were no significant differences observed between labelled and unlabelled mutant proteins.

F38 is not a conserved residue and is located in motif I. This mutant was also tested for ATPase activity and no free phosphates could be detected (data not shown). Hence, the mutation to a cysteine abolishes the ATPase activity due to its proximity to the Walker A motif.

N58 is a non-conserved residue within motif IA. Substitution by a cysteine did not alter the enzyme activity relative to the wildtype. The same was observed for the D186C mutant. The residue at this position is conserved, either as an aspartic acid or a proline and it is located close to the Walker B box. One would assume that amino acid replacements at this position may influence ATP hydrolysis, but it seemed that this residue is not critical for catalysis.

A511 is also not conserved and is situated in motif VI. The mutant protein exhibited wildtype helicase activity during the first minute, then strand displacement activity appeared to come to a halt. The assays were carried out with an excess of protein therefore one would assume that all DNA substrates are occupied by proteins before the reaction is initiated. Slower rebinding after substrate release can thus be excluded. One possibility is that the protein sample used represented a mixture of active and inactive enzyme and only the DNA substrates occupied with active XPD proteins contributed to the observed helicase activity.

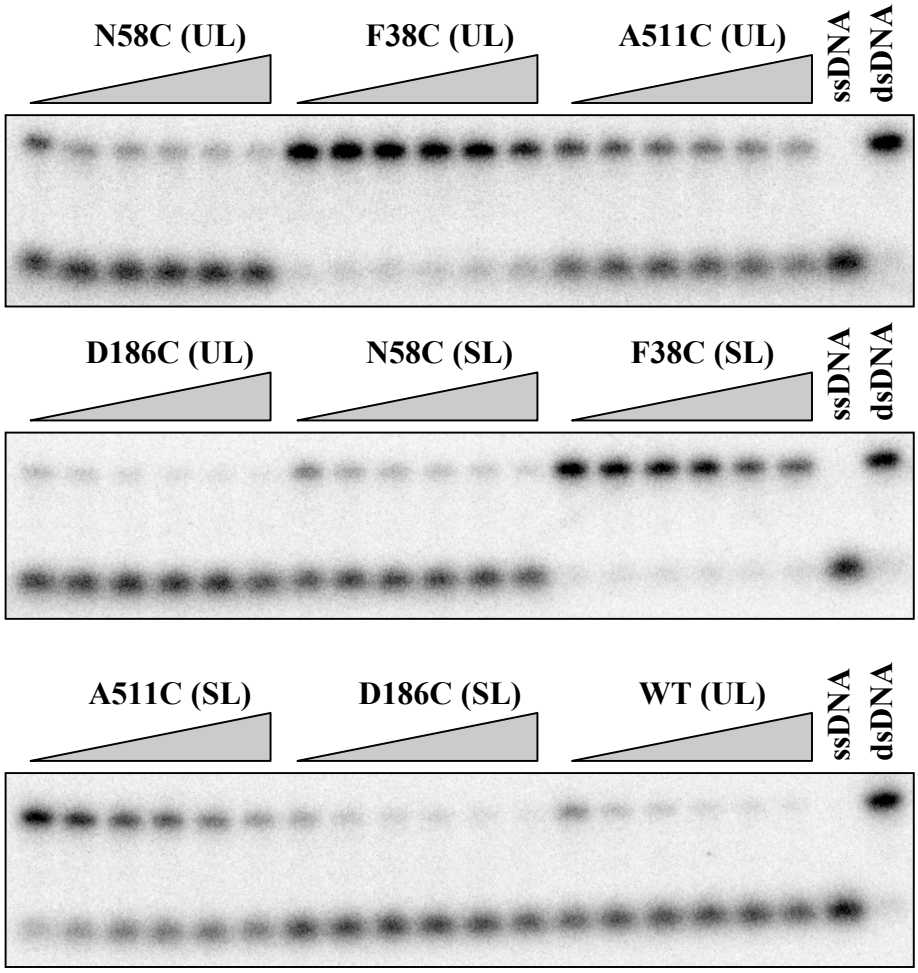


Figure 4.10: Native acrylamide gels showing helicase activities of mutants with and without spin-label

Representative 12 % acrylamide gels of helicase assays. The upper bands show dsDNA, lower bands represent the radiolabelled ssDNA arising if unwinding takes place. UL, unlabelled (no SDSLs attached), SP, spin-labelled (SDSLs attached). Time points were: 30, 60, 90, 120, 180, 300 s.

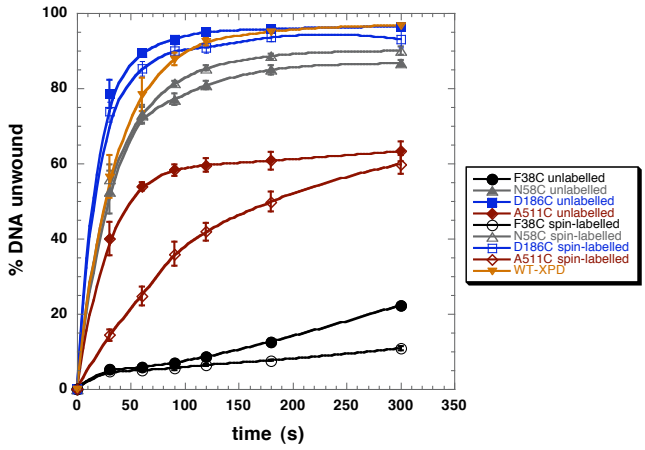


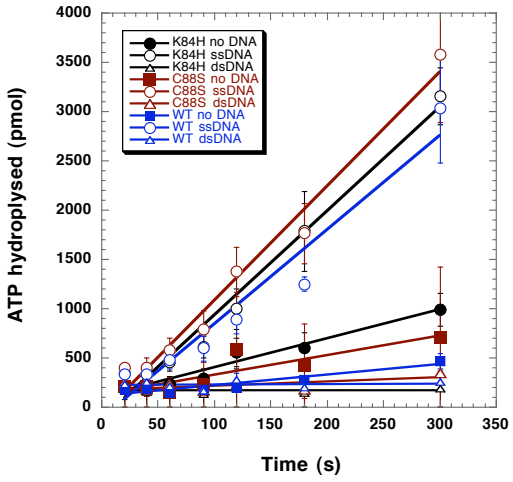
Figure 4.11: Quantification results showing helicase activities of mutants with and without spin-label

Quantification results of triplicate helicase experiments comparing XPD mutants with and without spin-label are shown and standard errors are indicated. The zero time point was given a value of DNA unwinding = zero. WT is shown as positive control. Data points were fitted to a smooth curve.

4.10 MONITORING ATPASE ACTIVITY UNDER DIFFERENT CONDITIONS

4.10.1 Single and double stranded DNA stimulation of the ATPase activity

Most NTPases show nucleic acid stimulation of their NTP hydrolysing activity. The majority of helicases are stimulated by single stranded nucleic acid although a few show double stranded nucleic acid stimulation (Singleton & Wigley (2002); Tuteja & Tuteja (2004)). Many helicases also show a very low basic level of NTP hydrolysis activity (Levin & Patel (2003)). The stimulation of the ATPase activity of XPD was investigated by using PhiX174 virion (ssDNA) and RFII DNA (dsDNA). The protein concentration used was 50 nM. This ensured accurate absorbance readings at later time points and that all proteins were bound to DNA. The C88S and K84H mutants were chosen as representative examples for enzymes without FeS cluster (Class I) and those with an instable cluster (Class II). Results were compared to the wildtype protein and are shown in Figure 4.12.



	no DNA	ssDNA	dsDNA
K84H	3	10.6	0.03
C88S	2	11.6	0.4
WT	1	9.6	0.06

Figure 4.12: ATPase activity of XPD wildtype and mutants in the absence and presence of DNA

The graph shows results from triplicate experiments for wildtype and mutant enzymes with ssDNA, dsDNA and without DNA. Time points were: 20, 40, 60, 90, 120, 180, 300 s. Data were fitted to a linear regression and standard errors are indicated. The table summarises the results of the ATPase assays for the wildtype and the two mutant proteins. Numbers indicate ATP hydrolysed/ second in pmol.

Wildtype and mutant XPD showed similar rates of ATP hydrolysis with ssDNA and dsDNA, as well as without DNA (Figure 4.12). This indicates that the ATPase activity is not dependent on the presence of the FeS cluster. This means, that the loss of the FeS cluster uncouples the ATPase and helicase activity. Mutant and wildtype XPD showed ssDNA stimulated ATPase activity with a rate of approximately 10 moles ATP hydrolysed per second (~ 0.7 mol ATP sec^{-1} per mol XPD). There was a general trend where mutants with uncoupled helicase and ATPase activity showed a slightly higher ATP hydrolysing activity than the wildtype (data not shown). A very low background was observed without DNA (1-3 ATP sec^{-1}) and binding to dsDNA completely inhibited ATP hydrolysis (0.03-0.4 ATP sec^{-1}). An initial lag phase, typical for pre-steady state kinetics, was observed (Eoff et al. (2005); Figure 4.12). It arises from the fact that the enzymes pre-bound to DNA have to undergo several conformational changes after addition of ATP/ $MgCl_2$ before they can hydrolyse ATP and release the product (ADP and P_i).

XPD is a monomer in solution. Monomeric helicases have two sites of DNA (ds/ ss) binding per monomer. Binding to ssDNA with the ssDNA binding site induces the correct domain movements and enables the enzyme to efficiently bind and hydrolyse ATP. Binding to dsDNA with the dsDNA binding site induces domain movement that leads to an incorrectly formed ATP binding pocket. This leads to an inability of the enzyme to bind and/ or hydrolyse ATP and inhibition of the ATPase activity is observed. In solution, a basal ATP hydrolysing activity is seen, due to the fact that the flexible enzyme may occasionally adopt the correct conformation of the ATP binding pocket, and hydrolyses ATP (Korolev et al. (1997), Velankar et al. (1999), Soultanas & Wigley (2000)).

All other mutants were then tested and all showed ssDNA dependent ATPase activities similar or slightly higher than the wildtype, with a few exceptions (data not shown). No free phosphates were detected for XPD-K35A. This mutation is located in the Walker A box and leads to an inability to hydrolyse ATP (Sung et al. (1988)). The C523R and G447D mutants showed extremely low activities, with 1.5 ATP hydrolysed sec^{-1} (Figure 4.7). The R531W mutant showed an ATP hydrolysing rate approximately half of that of the wildtype enzyme (see also Table 4.2.).

4.10.2 ATPase activity on different lengths of oligonucleotides

Reactions were carried out to investigate if the FeS cluster of XPD is necessary for translocation of the enzyme along ssDNA. For this experiment two representative mutants were chosen, which contained an instable (K84H) and no FeS cluster (C88S), respectively, and were compared to the wildtype XPD. It was found that the rate of ATP hydrolysis is directly proportional to the length of the oligonucleotide (Young et al. (1994)). Hydrolysis of one ATP molecule is directly coupled to movement of the helicase of one step-size (a specific number of bases) along the DNA. Thus, the longer the oligonucleotide, the more ATP is hydrolysed before the enzyme releases the DNA strand. The two different approaches to test this are visualised in Figure 4.13.

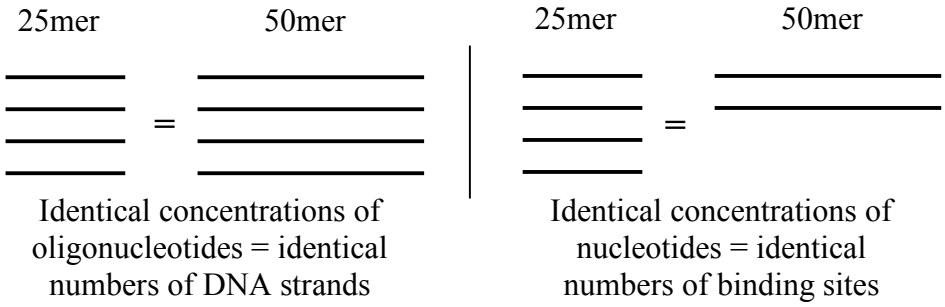


Figure 4.13: Schematic of the two different approaches to investigate translocation by measuring ATPase rates

The figure depicts the different setups described in 4.10.2.1 and 4.10.2.2. The black lines symbolise the numbers of available single DNA strands of an 25mer compared to an 50mer oligonucleotide.

4.10.2.1 ATPase activity of XPD using identical concentrations of oligonucleotide

Different lengths of ssDNA using identical concentrations of oligonucleotides (5 nM; Figure 4.13) led to an increase of ATPase activity, which was directly proportional to the length of the oligonucleotide (Figure 4.14). The maximum rate appeared to be reached with 100 nt for mutants K84H and C88S, because no further increase was observed by using circular PhiX174 virion DNA. The wildtype exhibited slightly lower ATP hydrolysing rates, as was observed before (4.10.1). All tested XPD proteins reached a maximum rate of between 30 and 35 mol ATP hydrolysed per second (~ 0.5 - 0.7 mol ATP sec^{-1} per mol XPD).

The translocation rate for the PcrA helicase has been calculated to be 50 bases/ second, with a step size of 1 ATP/ base. In these experiments different lengths of ssDNA with identical concentrations of oligonucleotide were used and the oligonucleotide was in excess (8 μM) to PcrA (0.1 μM) (Dillingham et al. (2000)).

In the experiments described in this section, the concentration of oligonucleotide was low (5 nM), due to limited availability compared to an excess (200 nM) of protein. This could lead to the effect that the longer the oligonucleotide, the more proteins can bind and the observed increase in ATPase activity is solely due to the fact that more ATPases are stimulated. Hence, translocation cannot be proven. Considering that XPD

was in 40-fold excess over DNA substrate, even the 100 nt oligonucleotide would not provide enough binding sites for all proteins to bind. Nonetheless, maximum ATP hydrolysing rates were reached for the mutant proteins with an oligonucleotide length of 100 nt and no further increase with longer ssDNA substrates was observed (Figure 4.14). There are several possibilities to explain these results. It could mean that when the 100 nt oligonucleotides were used, all proteins were already bound to nucleic acid and ATPase activity is stimulated, but no translocation occurs. On the other hand it could mean that all protein molecules were bound to the 100 nt oligonucleotide and maximum translocation rates are already reached and therefore no further increase in ATPase activity is observed using the PhiX174 ssDNA. But in contrast to this, the wildtype enzyme (which is assumed to be able to translocate along ssDNA) showed only approximately 60 % of its maximum ATPase activity when the 100 nt oligonucleotide was used, while the endpoint of maximum ATP hydrolysing rate was slightly higher than the ones observed for the mutant proteins. This could indicate, that mutants and wildtype have different step sizes (ATP hydrolysed/ base) when they are tracking along ssDNA. The loss of the FeS cluster in mutant proteins might not only lead to an uncoupling of ATPase and helicase activity, but also to the effect that mutants hydrolyse more ATP to track one base further along ssDNA than the wildtype enzyme, while maximum ATPase rates are approximately the same. This would lead to the observed effect that mutants reach maximum rates faster under the experimental conditions than the wildtype. It would also indicate that mutants without FeS cluster are able to translocate, although in a less efficient manner compared to the wildtype enzyme. Further experiments have to be carried out to fully prove translocation ability of mutant XPD proteins.

Oligonucleotide length	WT	K84H	C88S
	ATP hydrolysed sec^{-1}		
0	1	3	2
6	1.3	2.7	2.6
14	3.8	6.1	3.6
25	4	8.9	7.1
50	8.4	16.6	14.8
70	12.7	21.5	19.7
100	19.1	31	31.6
5800	36.3	32.1	34.3

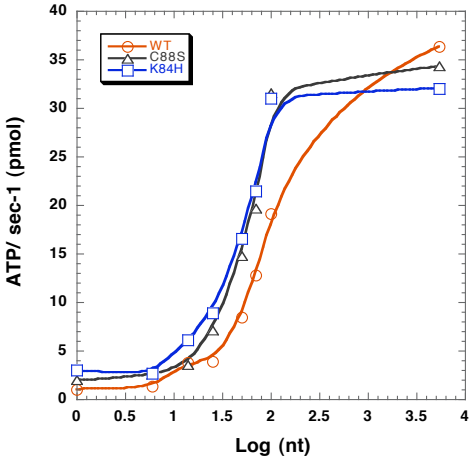


Figure 4.14: Rates of ATP hydrolysis with equal numbers of single DNA strands

The table shows obtained rates with different lengths of ssDNA. Concentrations were: DNA, 5 nM (oligonucleotide); protein, 200 nM. The assays were carried out at 45 ˚C. Time points (20, 40, 60, 90, 120, 180, 300 s) were plotted against the free phosphates, which were released after ATP was hydrolysed. The rate of ATPs hydrolysed sec^{-1} was obtained by linear regression. The rates were plotted against the length of the DNA strands and visualized on a half-logarithmic scale. Experiments were carried out in duplicate.

4.10.2.2 ATPase activity of XPD using identical concentrations of nucleotides

Compared to the setup described in the previous section, this experiment keeps available binding sites constant, rather than the number of single DNA strands (Figure 4.13). The concentration of nucleotides used was 2 µM, the protein concentration was reduced to 20 nM and the experiments carried out at 55 ˚C.

By using identical concentrations of nucleotides the ATPase activity only increased until a length of 25 nt (Figure 4.15). One possibility is that there were too few oligonucleotides available for binding as the length of the DNA strands increased, which led to the observed decrease in ATP hydrolysis with longer oligonucleotides.

The more likely explanation is that the concentration of the oligonucleotides used was inaccurate, as they were not gel purified prior to use. Therefore it is possible, that the oligonucleotide samples presented a mixture of different oligonucleotide lengths that

would lead to an incorrect calculation of the concentration of nucleotides after UV absorbance readings. The nucleotide concentration of the PhiX174X ssDNA was correct, as supplied by the manufacturer.

To make a valid comparison of the data shown in Figure 4.15 with experiments described in section 4.10.2.1 the protein concentration used has to be considered 10 times higher in the latter experiment. Hence, maximum ATPase rates observed with PhiX174 ssDNA in the experiments described in this section are approximately double compared to the ones in the previous section. This is an effect of the increased temperature used in the setup described in this section (55 ˚C instead of 45 ˚C). The only obvious difference when the rates of ATP hydrolysis are compared between the two different experimental setups is that the mutant proteins showed higher maximum ATPase rates (Figure 4.15) (1.4-1.5 ATP sec^{-1} per molecule XPD for PhiX174 DNA) than the wildtype enzyme (1.1 ATP sec^{-1} per molecule XPD for PhiX174 DNA). This would match the smaller step sizes of mutants without FeS cluster compared to the wildtype, as described in section 4.10.2.2, only that maximum ATPase rates also increase.

As stated for the previous section, these experiments do not prove translocation of mutants without FeS cluster and experiments to elucidate this are ongoing.

Oligonucleotide length	WT	K84H	C88S
	ATP hydrolysed sec^{-1}		
0	1	3	2
6	8.4	7	4.5
14	13.5	11.7	11.7
25	14.7	13	11.2
50	9.4	7.1	5
70	10	9.6	7.6
100	6.6	6.6	7.1
5800	6.7	8.5	8.8

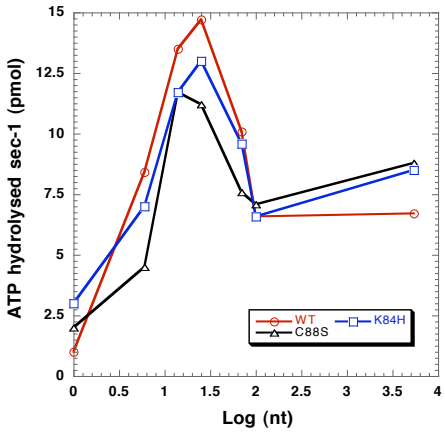
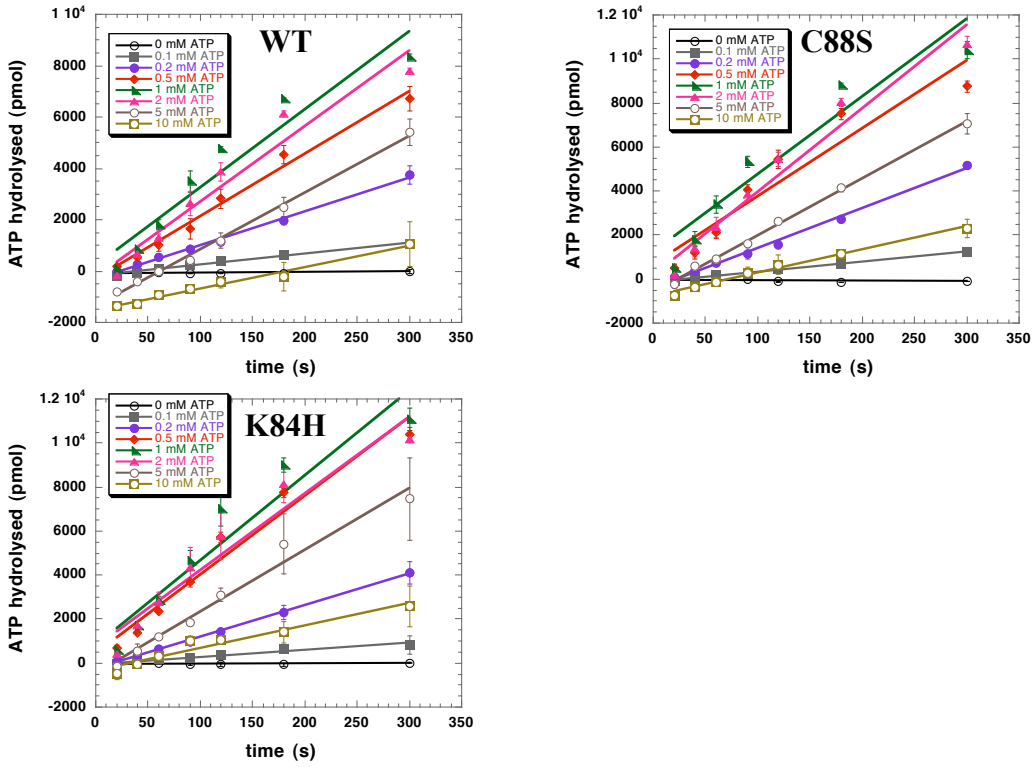


Figure 4.15: Rates of ATP hydrolysis with equal numbers of nucleotides

The table shows ATP hydrolysis rates obtained with different lengths of ssDNA. Concentrations were: DNA, 2 μM (nucleotides); protein, 20 nM. The assays were carried out at 55 ˚C. Time points (20, 40, 60, 90, 120, 180, 300 s) were plotted against the free phosphates, which were released after ATP was hydrolysed. The rate of ATPs hydrolysed sec^{-1} was obtained by linear regression. The rates were plotted against the length of the DNA strands and visualized on a half-logarithmic scale. This experiment was only carried out once.

4.10.3 ATPase activity using different concentrations of ATP

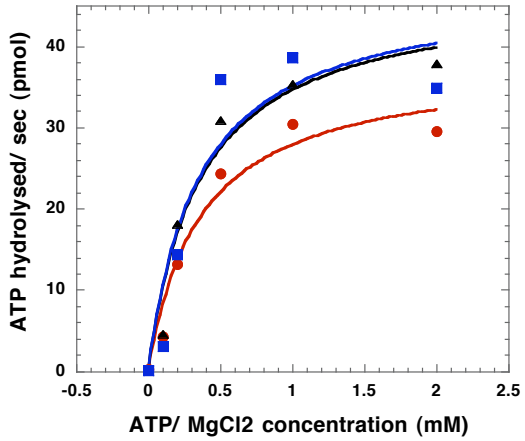
Reactions were carried out as described previously using 200 nM protein and 0.5 nM PhiX174 virion DNA with different concentrations of ATP:$MgCl_2$ (0, 0.1, 0.2, 0.5, 1, 2, 5, 10 mM). Reactions were stopped at specific time points (20, 40, 60, 90, 120, 180, 300 s) and results fitted to a linear fit (Figure 4.16).



ATP concentration (mM)	ATP hydrolysed s^{-1} WT	C88S	K84H
0	0.3	-0.1	0.1
0.1	4.3	4.6	3.1
0.2	13.2	18.1	14.3
0.5	24.4	30.9	35.9
1.0	30.4	35.3	38.6
2.0	29.5	37.9	34.9
5.0	22.2	25.8	28.1
10.0	8.3	10.6	10.2

Figure 4.16: Influence of the ATP concentration on the ATPase activity of XPD

The plots on the top of the figure show data of ssDNA stimulated ATPase activity of XPD wildtype and two representative mutants from duplicate experiments. Reactions were initiated using different concentrations of ATP/ $MgCl_2$. Time points were 20, 40, 60, 90, 120, 180 and 300 s. Data were fitted to a linear fit and standard errors are indicated. The rates obtained from the linear regression are summarized in the table on the bottom of the figure.



	K_m ± standard error (mM)	V_{max} ± standard error (pmol s^{-1})
WT	0.36 ± 0.13	38 ± 4.5
C88S	0.35 ± 0.12	46.9 ± 5.4
K84H	0.35 ± 0.22	47.5 ± 9.8

Figure 4.17: Rates of ATP hydrolysis depending on the ATP concentration

The rates (per second) shown in the table in Figure 4.16 were plotted against the ATP concentration and data points up to 2 mM ATP fitted to Michaelis-Menten kinetics. K_m (mM ATP) and V_{max} values (pmol ATP hydrolysed per second) are shown in the table. The graph shows the WT with red circles, C88S in black triangles and K84H in blue squares. All reactions were carried out in duplicate.

As seen from the standard errors in Figure 4.17, the Michaelis-Menten fit was not very good. This was due to an inhibition of the ATPase activity above 1 mM ATP/ $MgCl_2$. The K_m values are in good agreement with rates measured for wildtype XPD using different concentrations of ATP in helicase assays (Chapter 3). The maximum ATPase activity was obtained with approximately 0.35 mM ATP/ $MgCl_2$ for wildtype and mutant enzymes. As previously observed, the mutants showed slightly higher ATP hydrolysing activity compared to the wildtype protein (4.10.1).

The inhibition at higher concentration of ATP/ $MgCl_2$ may be due to a general inhibitory effect of high concentrations of magnesium ions on the enzymatic activity. It was reported that magnesium concentrations above 4 mM inhibit the enzymatic activity of hXPD, while increasing concentrations of ATP increased the enzyme activity (Dubaele et al. (2003)). Therefore, it has yet to be tested if the inhibitory effect is due to an inhibition by ATP or by Mg^{2+}-ions.

4.11 SUMMARY AND CONCLUSIONS

Archaea are excellent models for studying eukaryotic DNA repair pathways. This chapter showed how work on an archaeal protein can lead to the discovery of completely new features of the human homologue and can help to understand the pathology of human disorders such as TTD and Fanconi anemia.

This chapter demonstrated the importance of the FeS cluster for the helicase activity of XPD. It showed that three of the four conserved cysteines are involved in coordination of the FeS cluster consistent with findings that wildtype XPD binds a 3Fe-4S cluster as described in the previous chapter. The cluster was shown to be essential for DNA unwinding but not for ATPase activity. Indeed, mutants which do not coordinate a (stable) FeS cluster were still able to hydrolyse ATP and most likely translocate along DNA, while they were no longer able to unwind DNA (a summary is given in Table 4.2). Hence, destruction of the FeS cluster uncouples helicase and ATPase activity.

Strong evidence was provided for the existence of the FeS cluster in hXPD and other related human helicases such as FancJ, which takes part in the Fanconi anemia pathway. Indeed, the human mutation R112H was shown to abolish the helicase activity of hXPD, the reason for which was unclear (Dubaele et al. (2003)). Here it was shown, that this mutation and the FancJ-A349P mutation disrupt the FeS cluster and inactivate the helicase activity. The presence of the cluster was also confirmed *in vivo*, using *S. cerevisiae* as an eukaryotic model. The UV sensitivity of corresponding mutations, which were known to disrupt the FeS cluster in *S. acidocaldarius* were evaluated in yeast. This study was also strengthened by an earlier study on Chinese hamster ovary cell lines, where one cell line (UV-5) was found to possess the XPD mutation C116Y (equals Rad3-C115; Table 4.2), which caused a highly UV-sensitive phenotype (Weber et al. (1994)).

Other human mutations on the C-terminus of hXPD were introduced in *S. acidocaldarius* XPD and characterised. All mutations abolished the helicase activity in accordance with findings in hXPD, but no uncoupling of helicase and ATPase activity was observed. The two mutations G447D and C523R abolished the ATPase activity of SacXPD, although the latter is not located in one of the conserved motifs. In human, the R683W mutation abolishes the interaction with the p44 subunit of TFIIH, which stimulates the helicase activity of XPD (Dubaele et al. (2003)). On the contrary, archaeal

XPD is functional as a monomer and its helicase activity is not likely to be further stimulated by binding to a partner protein. The equivalent mutation in *S. acidocaldarius* (R531W) leads to a low ATPase activity, which could be explained by a weak interaction with ssDNA (Chapter 5). This weak interaction could cause the mutant protein to easily fall off the DNA substrate, which would in turn disable the helicase from unwinding dsDNA. Indeed, SacXPD-R531W did not exhibit helicase activity.

The FeS cluster most likely plays a structural role by stabilising a fold similar to the UvrB beta-hairpin (Figure 4.18; reviewed in Theis et al. (2000)). This fold extrudes from the protein structure and could be inserted into the DNA substrate near the ss/ dsDNA fork. When XPD translocates along DNA, this hairpin structure could either be used to physically disrupt the hydrogen bonds between base pairs or to lock one single DNA strand to the enzyme and actively distort the ss/ dsDNA fork, which would subsequently lead to duplex destabilisation. If the FeS cluster is not present, the hairpin cannot be formed and the ability to unwind DNA will be lost.

This chapter has shown, how mutagenesis studies can be used effectively to investigate the function of the FeS cluster of XPD. The results presented have also shown the FeS cluster to be essential for the strand displacement activity of the helicase.

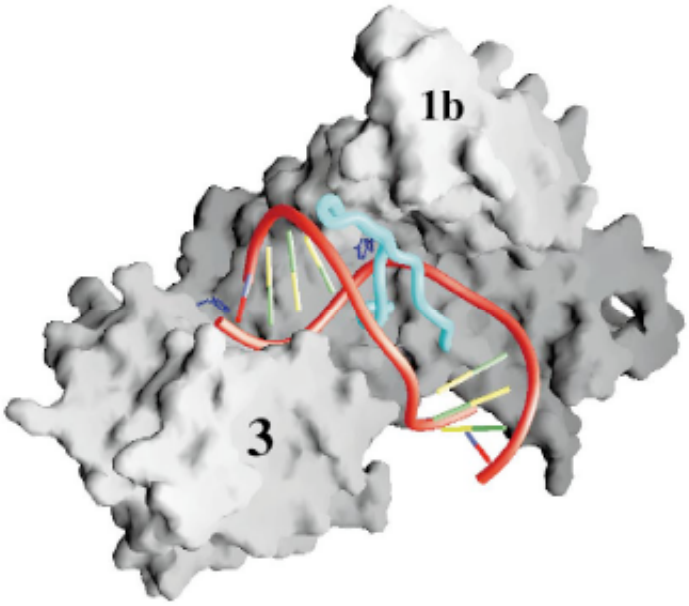


Figure 4.18: Hypothetical model of UvrB-DNA pre-incision complex

The UvrB protein is shown as a space-filled model in grey, the DNA phosphate backbone is shown in red and the beta-hairpin extruding from the protein in cyan. [adapted from Theis et al. (2000)]. The beta-hairpin is inserted into the DNA duplex and locks UvrB near the site of damage, which leads to the release of UvrA and the recruitment of UvrC. In XPD a similar structure could be used to actively disrupt base pairing during translocation.

Table 4.2: Summary of XPD mutations

Mutation			Phenotype			
SacXPD	hXPD	Rad3p	(human/ yeast)	Helicase?	ATPase?	FeS?
K35A	ND/ (G47R)	K48A	(XP)/ high UV sensitivity	N	N	stable
F38C	ND	ND	ND	N	N	stable
N58C	ND	ND	ND	Y	Y	stable
K84H	R112H	R111H	TTD/ mild UV-sensitivity	N	Y	unstable
C88S	ND	C115S	ND / High UV sensitivity	N	Y	non
C102S	ND	ND	ND	Y	Y	stable
C105S	ND	ND	ND	N	Y	non
F136P	A349P	ND	FA/ ND	N	Y	unstable
C137S	ND	ND	ND	N	Y	non
Y139F	ND	ND	ND	Y*	Y	stable
Y140F	ND	ND	ND	Y	Y	stable
Y139A/ Y140A	ND	ND	ND	N	Y	unstable
Y156A/ Y158A	ND	ND	ND	Y*	Y	stable
D186C	ND	ND	ND	Y	Y	stable
G447D	G602D	ND	CS + XP/ ND	N	N	stable
A511C	ND	ND	ND	Y**	Y	stable
C523R	G675R	ND	CS + XP/ ND	N	N	stable
R531W	R683W	ND	XP/ ND	N	Y**	stable
X-3CS	ND	ND	ND	Y	Y	stable

Table 4.2 shows a summary of all XPD/ Rad3 mutants described in this chapter. SacXPD, *S. acidocaldarius* XPD; hXPD, human XPD; Rad3p, orthologue of XPD in *S. cerrevisiae*; FancJ, human helicase FancJ (Fanconi anemia); FeS?, FeS cluster present?; X-3CS, XPD-C360S/ C523S/ C543S; ND, not defined; Y, mutant shows WT activity; Y*, activity slightly reduced relative to the WT; Y**, activity ~ 50 % reduced compared to the WT; N, no activity detected; XP, *xeroderma pigmentosum*; TTD, trichothiodystrophy; CS, Cockayne Syndrome; FA, Fanconi anemia.

5 DOMAIN ARRANGEMENT OF XPD AND INTERACTIONS WITH DNA AND OTHER PROTEINS

5.1 INTRODUCTION

In solution helicases are highly flexible enzymes. Large domain movements are observed upon binding to nucleic acid and during NTP binding and hydrolysis cycles. Some helicases also oligomerise following binding to nucleic acid, which changes their mode of action as was described in Chapter 1. This will be investigated for the case of XPD from *S. acidocaldarius*.

This chapter will also investigate changes in domain organisation when XPD is in the open or closed conformation. One of the best studied helicases is the SF1 helicase PcrA. Crystal structures revealed four domains, which were entitled 1A, 1B, 2A and 2B (Subramanya et al. (1998); Figure 5.1). It was hypothesised that domains 1A and 2A are common to all SF1 and SF2 helicases (Bird et al. (1998)). They are generally called the RecA-like motor domains, due to their similarity and first identification in the RecA enzyme (Story et al. (1992)). Domains 1A and 2A contain the seven conserved motifs and form a cleft for nucleotide binding. This cleft opens and closes while ATP is bound and hydrolysed (Velankar et al. (1999)). Domain 1A and 2A mediate ssDNA binding, while domains 1B and 2B, at least in PcrA and Rep helicases, are involved in dsDNA binding (Korolev et al. (1997); Velankar et al. (1999)). The two latter domains show high variability between helicases and are sometimes completely absent (Figure 5.2; reviewed in Singleton & Wigley (2002)).

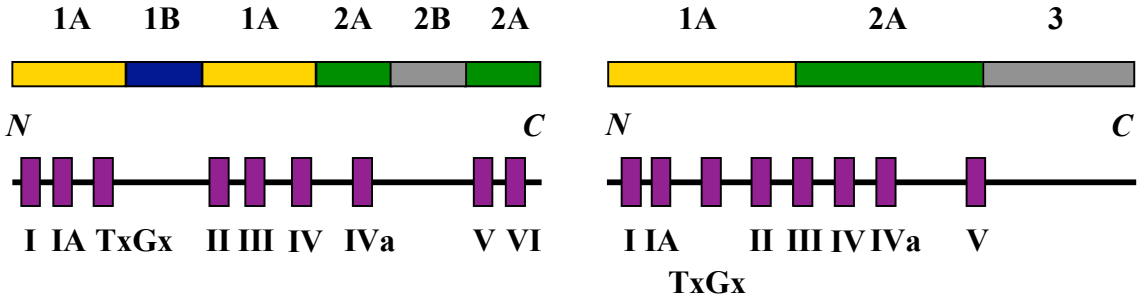


Figure 5.1: Contribution of the helicase motifs and their affiliations to the RecA-like motor domains

On the left the domain and helicase motif distribution of the SF1 helicase Rep is shown and compared to an SF2 helicase (NS3 from hepatitis C virus) on the right. The TxGx motif was originally identified by Pause & Sonenberg (1992) for members of the DEX(D/H) helicase family to be involved in ssDNA binding (Figure adapted from Korolev et al. (1998))

As described in Chapter 3, the FeS cluster domain of XPD is located between helicase Walker A and B boxes. Hence, it could either represent domain 1A or 1B. This chapter will investigate DNA binding of XPD wildtype and mutants to different DNA constructs, as well as the influence of ATP.

In eukarya, XPD is part of a multi-subunit complex which regulates and stimulates its enzymatic activity. In archaea it is a monomer in solution and the question, whether XPD interacts with other proteins will be addressed in this study.

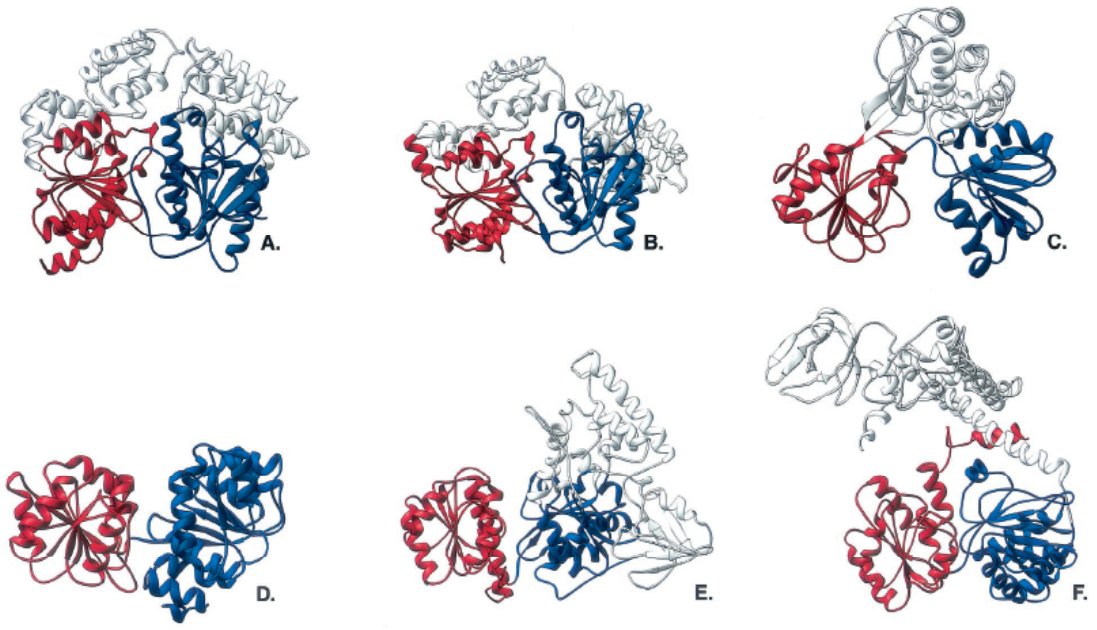


Figure 5.2: Structural conservation of the RecA-like motor domains

Crystal structures of different SF1 and SF2 helicases are shown. Domain 1A (blue) and 2A (red) are indicated. A, PcrA from *Bacillus stearothermophilus* (SF1); B, Rep from *E. coli* (SF1); C, NS3 helicase domain from hepatitis C virus (SF2); D, eIF4A-type RNA helicase from *Methanococcus jannaschii* (SF2); E, UvrB from *Thermus thermophilus* (SF2); F, RecG from *Thermotoga maritima* (SF2) (Figure adapted from Singleton & Wigley (2002)).

5.2 TRYPTIC DIGESTION

5.2.1 Domain mapping of wildtype protein and a "clusterfree" cysteine mutant

Tryptic digestion was used to identify possible domains of the XPD wildtype protein and examine differences in a mutant without the FeS cluster (C88S, for characterisation see Chapter 4). The wildtype protein showed seven major fragments after a 90 min incubation, indicating a highly flexible domain organization. There were obvious differences in the digestion pattern between wildtype and mutant. The mutant showed only five major protein bands arising after 90 min, some of which are completely different to the wildtype (Figure 5.4). Major protein fragments were analyzed by N-terminal sequencing and/ or mass spectrometry (Table 5.1).

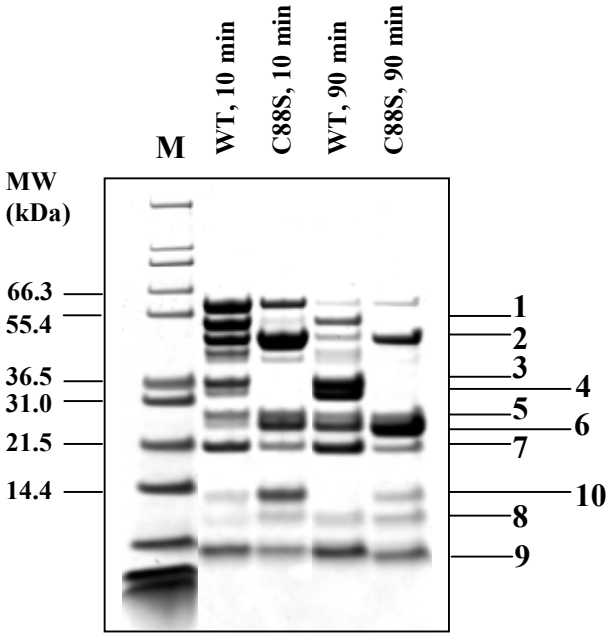


Figure 5.3: Analysis of XPD domain structure by tryptic digestion

The protein gel shows the migration of tryptic fragments of XPD. Numbers on the right indicate protein bands, which were analysed by N-terminal sequencing or mass spectrometry. Samples of wildtype (with cluster) and C88S (no cluster) were incubated 10 and 90 min, respectively, at 30 ˚C, before reactions were stopped.

Table 5.1: Results of analysed peptides after tryptic digestion

Fragment	N-Terminal Sequence	Mass spec, possible fragment	MW (kDa) estimated from gel	MW (kDa) from protparam analysis [1]	Origin of the band
1	(K_{73})RNITF	$N_{(75)} \rightarrow K_{(538)}$	55.5	54.2	wildtype
2	-	$G_{(94)} \rightarrow K_{(538)}$	50.0	52.1	C88S
3	(K_{73})RNITF	$N_{(75)} \rightarrow R_{(373)}$	37.8	35.1	wildtype
4	(K_{73})RNITF	$N_{(75)} \rightarrow K_{(369)}$	35.8	34.2	wildtype
5	-	$G_{(94)} \rightarrow K_{(352)}$	30.3	29.2	C88S
6	(K_{128})KDGLQ	$K_{(129)} \rightarrow K_{(352)}$	27.4	26.6	wildtype
7	(M_{372})RSDNM	$S_{(374)} \rightarrow K_{(538)}$	22.1	19.4	wildtype
8	(M_{372})RSDNM	$S_{(374)} \rightarrow R_{(460)}$	10.4	10.1	wildtype
9	(M_1)AKLRD	$L_{(4)} \rightarrow K_{(49)}$	5.5	5.5	wildtype

[1] http://www.expasy.ch/tools/protparam.html

Table 5.1 summarizes the results obtained by mass spectrometry and N-terminal sequencing for selected peptide bands from both wildtype XPD and C88S mutant. The numbers of the fragments in the first column on the left refer to the indicated bands in Figure 5.3. The accurate molecular masses of the bands from the SDS-gel were obtained by linear regression by plotting the molecular weight of the marker bands against the logarithm of the distance the marker proteins migrated in the gel.

Figure 5.4: Possible tryptic digestion pattern of XPD wildtype and C88S mutant

Schematic domain structure of the XPD protein (wildtype and mutant) is shown. The seven conserved helicase motifs are indicated in green and the FeS domain is shown in brown (wildtype) and gold (C88S), respectively. The arrows indicated trypsin sensitive sites relative to the helicase motifs and the FeS cluster domain. Sites confirmed by N-terminal sequencing are indicated with black arrows and assumed sites (mass spectrometry data only) with grey arrows. The major cleavage site on the N-terminus of the wildtype and mutant protein is indicated with a "X".

Without N-terminal sequence information for the tryptic fragments generated from the C88S mutant it would be difficult to determine with any certainty where the trypsin sensitive sites are positioned if only using the mass spectrometry data. As seen from Figure 5.3 the first incision site in C88S changes the position in the "clusterfree" mutant when compared to the first incision site in the wildtype. This leads to a 5 kDa smaller C-terminal fragment and an intense protein band of about 13 kDa, which resembles the N-terminal peptide (number 10, Figure 5.3). However, the peptide could not be identified by mass spectrometry. This experiment confirmed local structural changes upon loss of the FeS cluster as seen in the spectrum obtained by CD spectroscopy (Chapter 4, Figure 1.3). Data for other proteolytic fragments of the mutant protein obtained by mass spectrometry were not very conclusive. It would be useful to obtain N-terminal sequence information on these fragments.

The domain organisation of XPD can only be assumed (Figure 5.5). Most likely the FeS cluster domain is distinct from the domains harbouring motifs I, IA and II, because incisions are made by trypsin on both sides of the cluster. Since all helicase motifs must be located in domains 1A and 2A, the following domain organisation could be possible (Figure 5.5): Motifs I, IA and II are part of domain 1A; the FeS cluster represents an

insertion and resembles domain 1B; the C-terminal motifs are located in domain 2A. This is, despite the fact that there is a major tryptic sensitive site between motif III and IV (Figure 5.4), which could indicate that motif III does not belong to domain 2A. The insertion between motif II and III could represent domain 2B.

The reason for the observed trypsin sensitive site between motifs III and IV in XPD could be a loop structure, which was found to couple ATP hydrolysis to DNA binding (and subsequently translocation) in SF1 helicases. In the helicases PcrA (motif III loop) and Rep (loop2) this loop structure was found to be part of motif III. In the SF2 helicase RecQ, the so-called aromatic-rich loop was located between motifs II and III. It is possible that this loop structure is also present in XPD, located between motifs III and IV, leading to an additional tryptic sensitive site within domain 2A (Korolev et al. (1997); Velankar et al. (1999); Dillingham et al. (1999); Zittel & Keck (2005)).

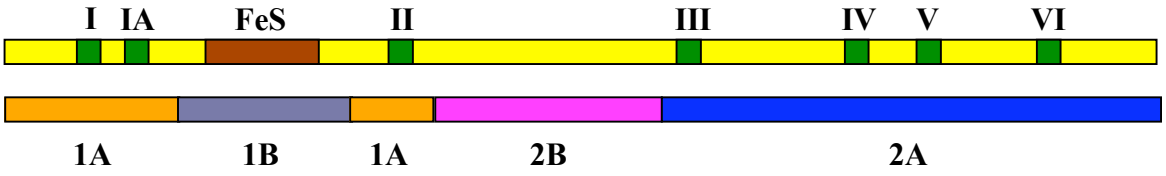


Figure 5.5: Possible domain organisation of *S. acidocaldarius* XPD

The upper figure represents a schematic representation of the secondary structure elements in XPD. The conserved helicase motifs are in green, the FeS cluster domain is in brown. The lower figure demonstrates a possible distribution of the two RecA-like motor domains (1A and 2A) in orange and blue, respectively. The additional domains 1B and 2B are presented in grey and magenta. The FeS cluster domain could represent an insertion (domain 1B) into domain 1A.

5.2.2 Influence of ADP on the tryptic digestion pattern

To investigate the domain structure of wildtype XPD and the C88S mutant protein upon binding of ADP, samples were incubated 90 min with and without 1 mM ADP/ $MgCl_2$.

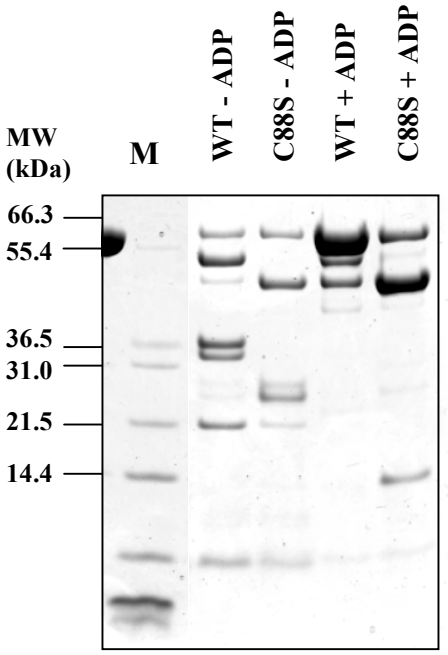


Figure 5.6: Influence of ADP on the domain structure of XPD

SDS-PAGE showing tryptic digest fragments of XPD in the absence and presence of MgADP. Samples were incubated 90 min at 30 ˚C before inactivation of trypsin.

The binding of ADP led to a striking structural change in the enzyme (Figure 5.6) and protection of almost all trypsin sensitive sites was observed. Most of the wildtype protein persisted as full length after 90 min incubation with trypsin. This is consistent with the hypothesis that helicases undergo large domain movements following binding of ATP/ ADP and transform from an open conformation to a more compact closed conformation. This event brings the helicase motifs in close three-dimensional proximity, ready to detach the terminal phosphate from ATP. It was also suggested that the gamma phosphate group functions as a molecular switch, which induces domain swivelling and

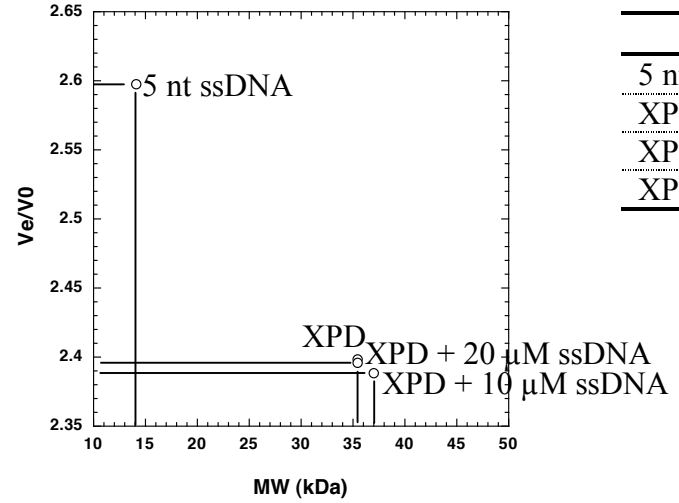
eventually results in translocation along nucleic acids (Korolev et al. (1997); Lorsch & Herschlag (1998)).

Interestingly, the first tryptic sensitive site from the N-terminus of the C88S mutant remained unprotected and available for tryptic digestion, while all other protein fragments disappeared (Figure 5.6). This site is located within the FeS cluster domain in clusterfree mutants (Figure 5.4). The ATP binding and hydrolysing pocket is still correctly formed, as this mutant showed ssDNA stimulated ATPase activity (Chapter 4). DNA binding is also not affected (section 5.4.3). Therefore, the hypothesis of stabilizing a region or fold by the FeS cluster is strengthened by the observation that ADP binding protects this area completely when the cluster is present. In mutants without cluster the amino acid chain may extrude from the protein and present a target for trypsin.

5.3 OLIGOMERIC STATE OF XPD UPON BINDING TO DNA

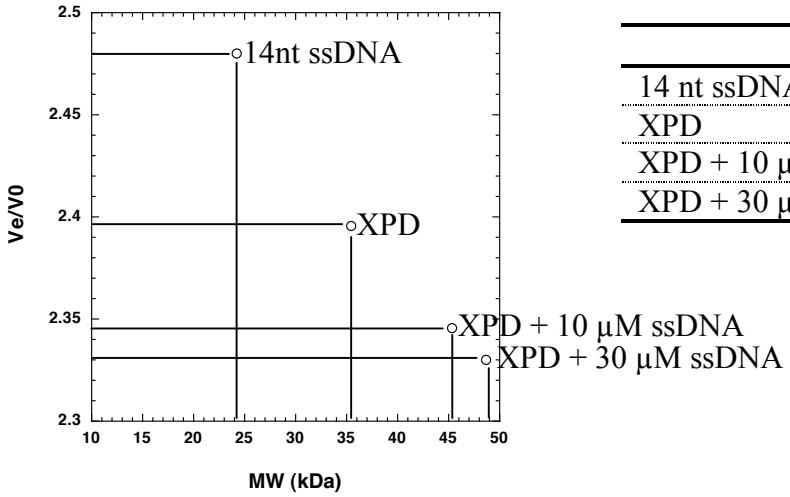
Some helicases, such as Rep or UvrB are monomers in solution, but the active form of the enzymes is a dimer, which assembles upon binding to nucleic acid (Chao & Lohman (1991); Verhoeven et al. (2002b)). The following experiments were carried out to investigate if the active form of *S. acidocaldarius* XPD is a monomer or an oligomer. This was performed by analytical gel filtration chromatography. XPD was incubated for 30 min at room temperature with ATP-γS and a 5 nt or 14 nt oligonucleotide before the mixture was applied to a Superose 6 analytical column. The control samples (XPD without DNA and DNA without XPD) were set up in the same way and ATP-γS was added to the reaction buffer.

A



	MW (kDa)
5 nt ssDNA	14.1
XPD	35.4
XPD + 10 µM ssDNA	37.1
XPD + 20 µM ssDNA	35.5

B



	MW (kDa)
14 nt ssDNA	24.3
XPD	35.4
XPD + 10 µM ssDNA	45.3
XPD + 30 µM ssDNA	48.6

Figure 5.7: Oligomeric state of XPD upon DNA binding

Superose 6 column elution volumes and calculated masses derived from a standard curve using proteins of known molecular masses. Controls were XPD without DNA and DNA without XPD and all samples were setup using the same conditions (including ATP-γS). Ve, elution volume; V0, void volume; MW, molecular mass. (A) Binding of XPD to a 5 dT single stranded oligonucleotide. (B) Binding of XPD to a 14 dT single stranded oligonucleotide. Calculated masses are shown in the tables on the right.

The molecular mass of XPD estimated from analytical gel filtration is rather low (35 kDa) compared to its expected mass of 64 kDa. Nevertheless, the intact protein mass was identified as full-length XPD by mass spectrometry. One likely reason for this size discrepancy could be, that after binding to the non-hydrolysable ATP-analogue ATP-γS, XPD changes into its closed conformation thus leading to a smaller diameter of the protein globular shape and retarding its passing through the superose column matrix. A lower mass for XPD was also observed in gel filtration chromatography using the Superdex 200 matrix, in the last step of protein purification. Here the buffers were virtually ATP-free and the calculated molecular mass of 48 kDa represents the open conformation of XPD, which is also likely to be tightly folded.

XPD bound ssDNA as a monomer. XPD-DNA binding was only observed in the presence of the 14 nt oligonucleotide, suggesting that the minimum DNA binding site of XPD exceeds 5 nt and stable protein-DNA interactions do not occur with oligonucleotides of 5 nt length (Figure 5.7). As seen from Figure 5.7B, XPD binds to the 14 nt ssDNA oligonucleotide, inducing a molecular mass shift of around 10 kDa relative to the sample that contained XPD only. This mass shift was not observed upon addition of a 5 nt oligonucleotide (Figure 5.7A). Small molecular mass differences between samples containing 10 and 30 μM dT14 DNA can arise due to differences in sample loading (Figure 5.7B). In general, it does not imply that XPD forms higher complexes upon binding to ssDNA indicating that the active form is also a monomer. From these experiments, there is a high possibility that the active form of XPD is a monomer. If dimer formation was to occur upon binding DNA, the expected molecular mass of the complex would be at least 70-80 kDa based on the elution volume.

Preliminary data obtained using Isothermal Titration Calorimetry (ITC) showed similar results. No XPD-DNA binding was obtained using a 6 nt long oligonucleotide, while 14 nt and 20 nt were bound efficiently. An increase in binding affinity was observed when using the 20 nt instead of the 14 nt oligonucleotide, consistent with a higher affinity to DNA with increasing length. (For ITC data see Appendix 2. Results are preliminary and experiments have to be repeated.)

5.4 XPD-DNA BINDING – ELECTROPHORETIC MOBILITY SHIFT ASSAYS (EMSA)

5.4.1 Defining buffer conditions for EMSA

Conditions were tested using high (20 mM HEPES, pH 7.6, 2 mM DTT, 200 mM NaCl, 0.002 % Triton-X) and low (20 mM HEPES, pH 7.6, 2 mM DTT, 50 mM NaCl, 0.002 % Triton-X) salt buffer, 10 nM duplex or 5' overhang DNA substrate (B50/ B50comp and B1-25/ X50, respectively; see Appendix 1 Table 2) with and without 0.1 mg/ mL calf thymus DNA. XPD protein concentrations used were 0 and 10 μM and BSA at 0.1 mg/ mL.

As shown in Figure 5.8, duplex and 5' overhang DNA substrates were bound with similar affinities to XPD. As expected, salt concentrations had the main influence on binding. For further experiments 50 mM NaCl was used. Calf thymus DNA also inhibited binding of XPD to DNA substrates, probably due to an increase of available DNA binding sites. Therefore, calf thymus DNA was excluded from further experiments.

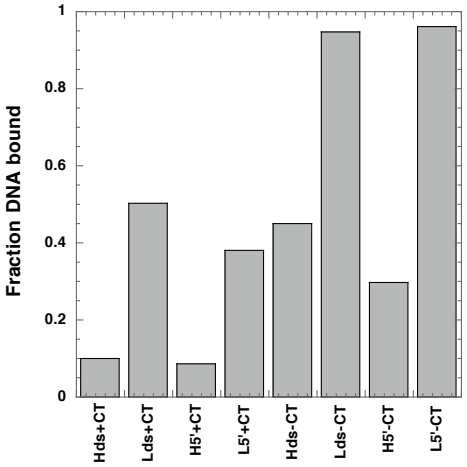


Figure 5.8: Wildtype-XPD DNA binding under different buffer conditions

Fractions of DNA bound are plotted against the reaction conditions. H, high salt (200 mM NaCl); L, low salt (50 mM NaCl); ds, double strand DNA (50 nt length); 5', 5' overhang DNA (50 nt length with 25 nt 5' ssDNA overhang); +CT, with calf thymus DNA; -CT, without calf thymus DNA.

5.4.2 The FeS cluster in XPD is not essential for DNA binding

The glycosylases MutY/ EndoIII take part in the Base Excision Repair (BER) pathway and contain a [4Fe-4S] cluster. This cluster stabilises a loop structure (FeS cluster loop, FCL), which coordinates positively charged residues for interaction with the DNA backbone. Indeed, it was found that MutY mutants without cluster lose their ability to bind DNA and to remove 8-oxo-guanine lesions from the DNA (Thayer et al. (1995); Chepanoske et al. (2000)). Experiments in this section were carried out to determine if the FeS cluster of XPD could have a similar structural function. The importance of the FeS cluster for binding of DNA was tested using mutants C88S (lacking the cluster), K84H and F136P (unstable cluster) and C102S (stable cluster). DNA binding of the mutants was tested using a duplex and a 5' overhang DNA substrate and compared to the wildtype enzyme. A representative gel is shown in Figure 5.9.

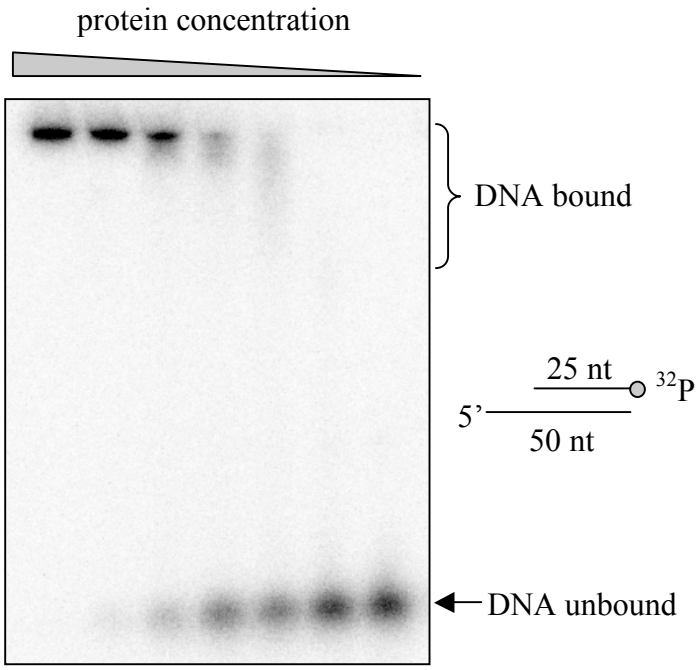
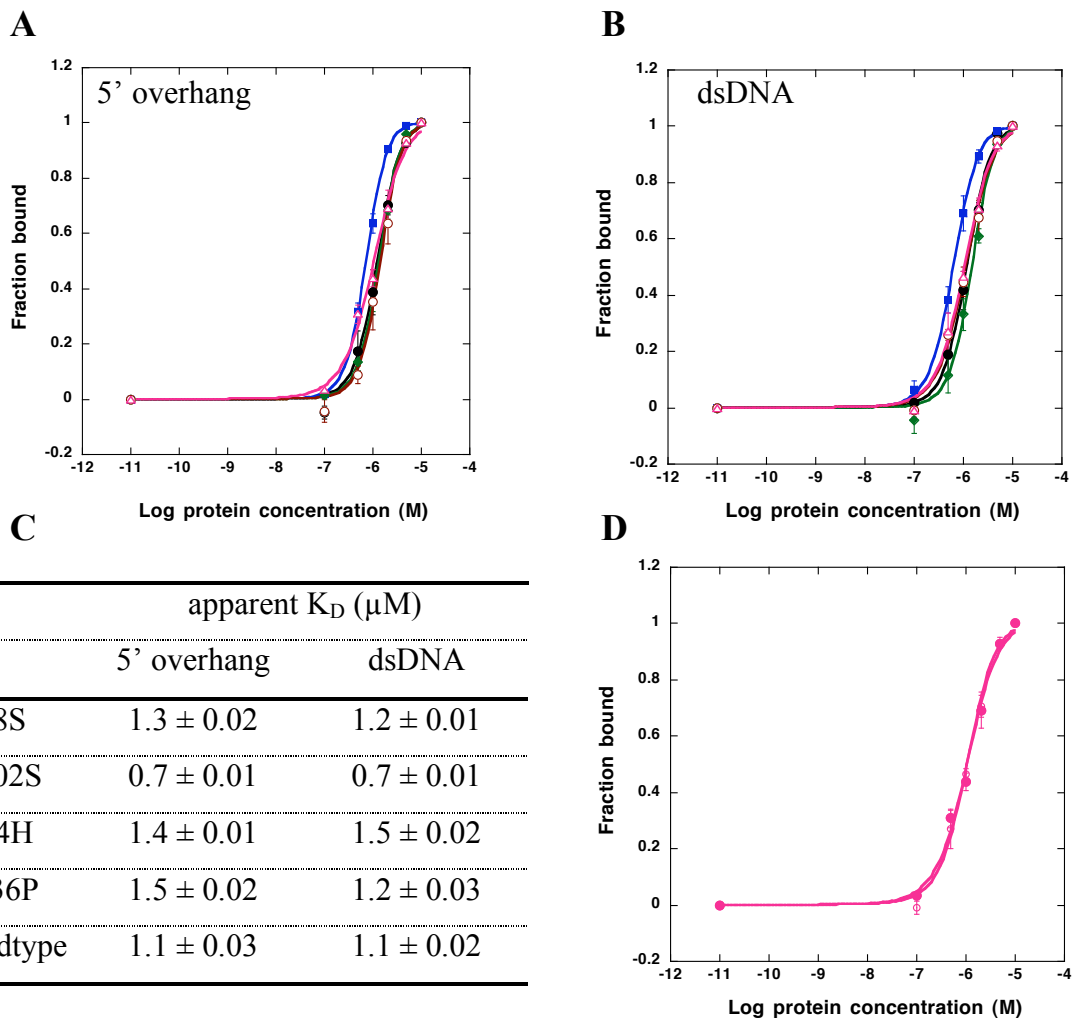


Figure 5.9: EMSA of wildtype XPD binding to a 5' overhang DNA substrate

The figure shows a representative 6 % native acrylamide gel of wildtype XPD binding to a 5' overhang DNA substrate. Protein concentrations were: 10, 5, 2, 1, 0.5, 0.1 and 0 µM XPD. Unbound DNA was quantified and Fractions bound calculated as described in Materials and Methods.



	apparent K_D (μM)	
	5' overhang	dsDNA
C88S	1.3 ± 0.02	1.2 ± 0.01
C102S	0.7 ± 0.01	0.7 ± 0.01
K84H	1.4 ± 0.01	1.5 ± 0.02
F136P	1.5 ± 0.02	1.2 ± 0.03
wildtype	1.1 ± 0.03	1.1 ± 0.02

Figure 5.10: DNA binding of XPD mutants and wildtype to dsDNA and 5' overhang substrate

Binding curves show fraction DNA bound versus the logarithm of XPD protein concentration. Binding of wildtype XPD and mutants to (A) 5' overhang and (B) dsDNA; colours are as follows: C88S, black; C102S, blue; K84H, green; F136P, brown; wildtype, magenta. Data were fitted to the following equation (for explanation see Materials and Methods): Fraction bound = 1/ 1+ $(K_D/ [XPD])^m$). Apparent K_Ds (in μM) from these binding curves are tabulated in (C). In (D) the wildtype XPD binding affinity to dsDNA (open circles) and 5' overhang (closed circles) DNA were compared directly. Experiments were carried out in triplicate and standard errors are indicated.

As seen from Figure 5.10, the FeS cluster is not essential for XPD-DNA binding and a similar role to the FCL of MutY/ EndoIII can thus be excluded. XPD wildtype and mutant proteins bound DNA with similar affinities. A 5' overhang DNA substrate is the preferred substrate for the 5'→3' helicase XPD, as the enzyme can load onto the 5' single stranded DNA overhang and unwind duplex DNA in 3' direction. Duplex DNA on the other hand was shown not to be a substrate for XPD, as dsDNA did not stimulate the ATPase activity (Chapter 4). Nonetheless, 5' overhang and dsDNA were both bound with the same affinities by wildtype and mutant XPD. Dissociation constants were calculated to be around 1 μM, indicating that XPD binds DNA weakly.

There may be several reasons for the observed weak binding affinity of XPD for DNA. Bandshift assays can give an idea about protein-DNA interactions; they are straightforward; different conditions and substrates can be tested quickly and easily. However, moving a large complex through a matrix of networked polymers for hours could potentially lead to complex disassembly, especially if binding is not very tight. Therefore, the observed K_D can appear much higher than the real one and is therefore called an apparent K_D. Another possibility is that XPD is missing a ligand or a partner protein to bind DNA more strongly. Lastly, there is the possibility that XPD has high on/ off rates or that tight DNA binding by XPD is not required as this would hinder ssDNA tracking of the helicase.

During the course of the aforementioned experiments, another observation made was that XPD-DNA complexes tended to remain in the wells of the polyacrylamide gels. In addition, two different slower migrating complexes were observed within the gel, which appeared and disappeared with increasing protein concentrations. This was seen for all DNA constructs of 50 nt length, no matter if the substrate was dsDNA, 5' overhang or different bubble substrates (Figure 5.14). The formation of several complexes in EMSA was also observed for the P143 helicase from the virus *Autographa californica* (McDougal & Guarino (2001)). The authors suggested, that multiple molecules bound to the DNA substrate in a non-cooperative manner, thus accounting for the various protein-DNA complexes seen.

The R531W mutant was tested separately. It was thought that this mutation could interfere with DNA binding (see Chapter 4), as this mutant bound only weakly to the heparin affinity column and its ATPase activity was reduced. These observations could arise from a decreased ability of the R531W mutant to bind to DNA resulting in a loss of helicase activity. The mutant was tested on a 5' overhang DNA substrate and preliminary data confirmed a weaker ability to bind DNA (Figure 5.11). The apparent K_D was calculated with 12.2 ± 0.03 µM, which is about 10-fold higher than the wildtype (1.1 µM). Further investigations are necessary to determine if the XPD-R531W mutation affects dsDNA or ssDNA binding.

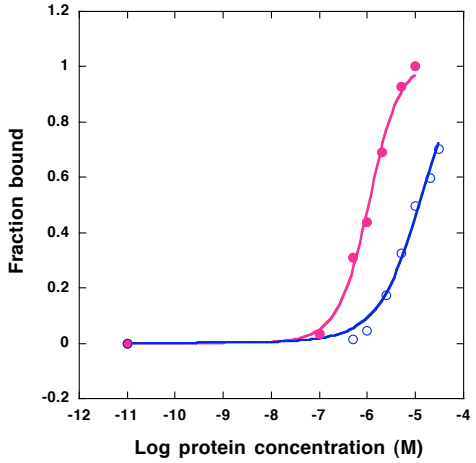


Figure 5.11: Comparison of DNA binding affinity of XPD-R531W and wildtype to a 5' overhang substrate.

DNA binding curves showing fractions of bound DNA versus the concentration of XPD-R531W and wildtype on a semilogarithmic scale. The wildtype is shown in magenta and the R531W mutant in blue. Data were fitted to the following equation (for explanation see Materials and Methods): Fraction bound = $1/ (1+ (K_D/ [XPD])^m)$. Experiments were carried out once.

5.4.3 Influence of ATP on binding to a 5'-overhang DNA substrate

It was shown in section 5.4.2 that the FeS cluster in XPD is not essential for its interaction with DNA, but these experiments were carried out in the absence of ATP. The following experiments tested whether ATP has an influence on the observed DNA binding affinity of XPD. Accordingly, 1 mM $MgCl_2$/ ATP-γS was included in the

reaction. The DNA binding affinity of a representative mutant protein lacking the FeS cluster (C105S) was compared to the wildtype using a 5' overhang DNA substrate (Figure 5.12).

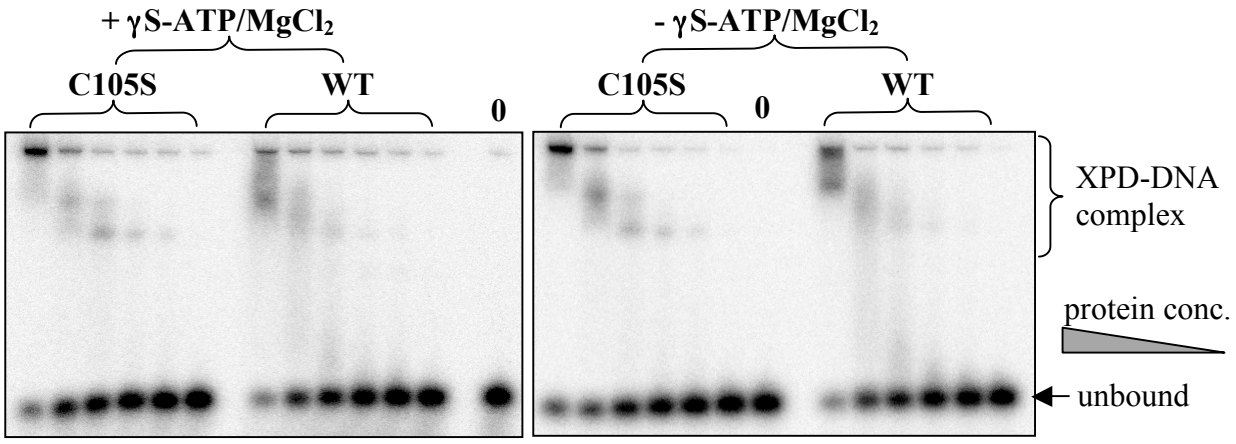


Figure 5.12: Effect of ATP on the binding affinity to a 5' overhang DNA substrate

Native EMSA showing XPD-DNA binding of wildtype and C105S mutant to a 5' overhang DNA substrate in the presence (left) and absence (right) of 1 mM MgATP-γS. Protein concentrations for each experiment were 10, 5, 2, 1, 0.5, 0.1 µM. "0" indicates the control without protein. Experiments were carried out once.

Unbound DNA was quantified and calculated to fractions of DNA bound as described in Materials and Methods. No significant differences in DNA binding by wildtype XPD and C105S mutant were observed in the presence or absence of ATP-γS, indicating that ATP does not increase the affinity to a 5' overhang DNA substrate (Figure 5.12) and is therefore not necessary for the XPD-DNA interaction.

5.4.4 XPD binding to single stranded DNA

The 14mer dT oligonucleotide, which had already been used for gel filtration XPD-DNA binding analysis (section 5.3), and a 25mer (B25comp, see Appendix 1, Table 2 for oligonucleotide sequence) of a mixed DNA sequence were used to investigate the ssDNA binding affinity of wildtype XPD. It was assumed that XPD would bind ssDNA in a sequence independent manner, as it has been described for other helicases (Ayora et al. (2002); Levin & Patel (2002); Tuteja & Tuteja (2004)).

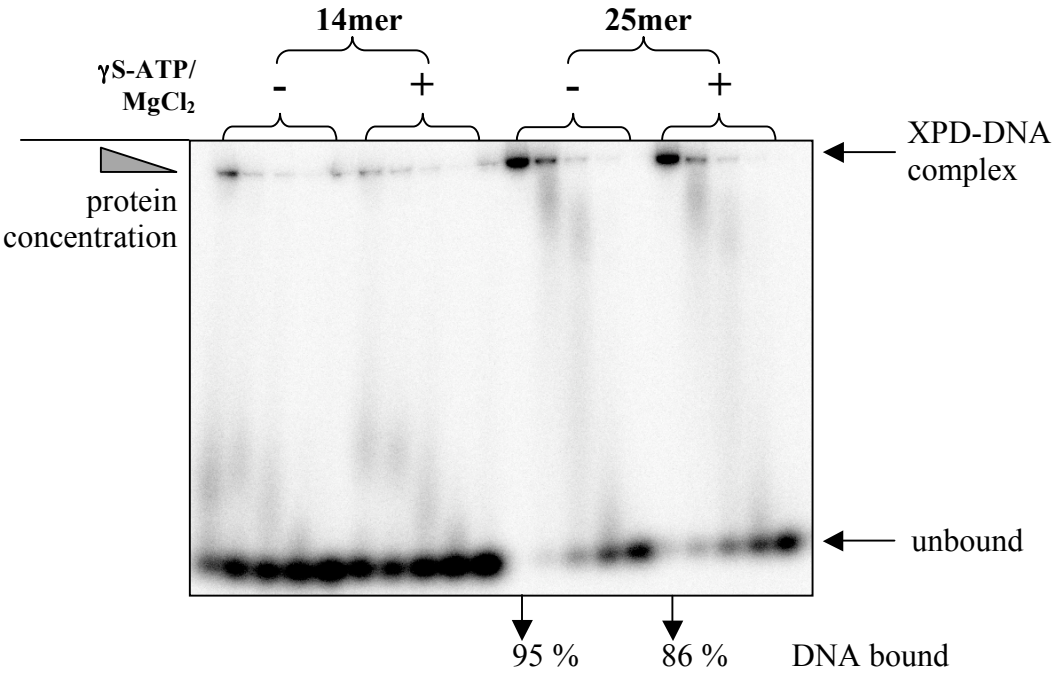


Figure 5.13: Single stranded DNA binding of wildtype XPD

Native EMSA showing binding of wildtype XPD to a 14mer (left) and a 25mer (right) ssDNA with and without MgATP-γS. Protein concentrations were: 50, 25, 12.5, 5, 0 μM. $MgCl_2$ and ATP-γS were at a concentration of 1 mM. Experiments were carried out in duplicate.

XPD binding affinities for ssDNA were considerably weaker than those observed for double stranded or partial duplex DNA (section 5.4.2). The strength of single stranded DNA binding was also dependent upon the length of the oligonucleotide (Figure 5.13). Generally, most helicases show tighter interactions with ssDNA than with dsDNA (Tuteja & Tuteja (2004)). In the case of XPD, however, the above data show that for stable binding to occur only the length of the oligonucleotide is important. Thus, it could be possible that XPD binds a 50mer ssDNA with the same affinity as a 50mer dsDNA (section 5.4.2).

Binding assays were also carried out in the presence and absence of MgATP-γS. From gelshift experiments it appears that the presence of ATP slightly decreases the affinity of XPD for ssDNA (Figure 5.13). However, these differences were not significant

(e.g. apparent K_Ds for the 25 nt oligonucleotide from duplicate experiments: 10.5 μM (-ATP-γS) vs. 12.1 μM (+ATP-γS)).

It will be necessary to carry out the same experiment on a fully complementary DNA duplex and a ssDNA oligonucleotide of the same length in the presence and absence of ADP and ATP-γS, to investigate if changes occur in the DNA binding affinities during ATP hydrolysis cycles of XPD. It would also be more useful to obtain data in solution (e.g. using fluorescence polarisation experiments), rather than using bandshift assays, because the protein-DNA complex is not likely to be stable enough to observe small differences in protein-DNA binding affinities using gel shifts.

5.4.5 DNA binding affinities of XPD to bubble vs. 5' overhang substrates

In humans, XPD takes part in DNA repair. Single stranded regions in a DNA duplex can arise after DNA damage has occurred or during the repair process by formation of the so-called repair bubble (Evans et al. (1997)). In the following experiments a possible preference of XPD for bubble DNA substrates, compared to a 5' overhang, was tested as bubble DNA structures naturally occur during DNA repair. Two different DNA constructs were made containing a central uncomplementary region of 4 and 10 nt, respectively. These constructs are referred to as "bubble" substrates.

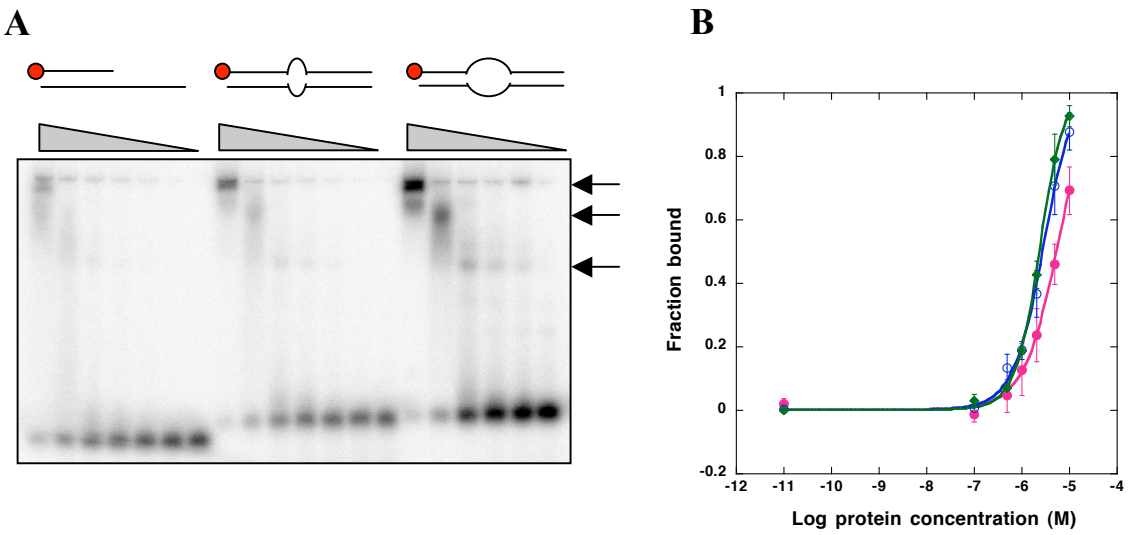


Figure 5.14: Comparison of XPD binding to bubble and 5' overhang DNA substrates

Native gelshift experiments (A) show binding of wildtype XPD to a ^{32}P-radiolabeled 5' overhang (left panel), 4 nt bubble (middle panel) and 10 nt bubble (right panel) DNA substrate. Schematics of substrates are shown across the top. Protein concentrations in descending order are 10, 5, 2, 1, 0.5, 0.1, (0) µM. The arrows indicate the three different complexes formed within the gel. (B) Binding curves of triplicate experiments showing total fraction XPD bound to each substrate on a semilogarithmic scale. Green, 10 nt bubble; blue 4 nt bubble; magenta, 5' overhang. Data were fitted to the following equation (for explanation see Materials and Methods): Fraction bound = 1/ (1+ (K_D/ [XPD])^m). Standard errors are indicated.

Bandshift experiments showed that XPD preferentially bound the bubble substrates (Figure 5.14). Dissociation constants of triplicate experiments were determined as 5.5 ± 0.02 µM (5' overhang), 2.8 ± 0.02 µM (Bubble 4 nt) and 2.4 ± 0.01 µM (Bubble 10 nt). It appeared that the wildtype enzyme used in this series of experiments was not fully active. The same sample was tested for helicase and ATPase activity and showed much lower activity than normally observed, although mass spectrometry confirmed the correct mass for wildtype XPD. This is reflected in higher K_D values for DNA binding in this set of experiments, compared to previous experiments using 5' overhang DNA substrates (section 5.4.2). A direct comparison between the three substrates was still meaningful, as the reactions were treated in the same way, using the same reaction mix. Therefore, XPD binding affinities obtained for these three DNA substrates were comparable to each other.

It seems that XPD prefers to bind to unpaired regions within a DNA duplex, rather than long stretches of ssDNA or dsDNA. As shown in section 5.4.2 dsDNA and 5' overhang substrates were bound with similar affinities by XPD. If a perturbation in the DNA duplex is introduced in form of a bubble, the affinity of XPD to DNA doubles. This is consistent with the observation that human XPD takes part in NER and the active conformation of the repair complex is a bubble structure, which varies in size during the different stages of DNA repair (Evans et al. (1997)). These data will nonetheless require confirmation using different experimental approaches, as described in section 5.4.4.

Damage recognition factors have not yet been identified in archaea. The single stranded DNA binding (SSB) protein has been proposed to be involved in the early stages of damage recognition, as it is able to melt duplex DNA and preferentially binds damaged DNA while stabilising appearing single stranded regions (Cubeddu & White (2005)). *In vivo*, such local unwinding of duplex DNA results in the formation of a bubble. This could be a possible reason why XPD prefers to bind bubble substrates to 5' overhangs. After SSB has melted part of the DNA duplex, the presence of a bubble and SSB could lead to positive cooperativity and recruitment of the other NER factors. SSB, when bound to immobilized ssDNA was also shown to interact specifically with certain NER factors, such as XPB, but not XPD. XPD is the major helicase in NER (Coin et al. (1998)). Therefore, strong interactions of SSB with XPD would probably reduce the processivity of the helicase. In contrast, XPB has almost no detectable helicase activity and interaction with SSB could contribute to positioning XPB within the DNA repair complex.

5.5 XPD PROTEIN-PROTEIN INTERACTIONS

In humans, XPD is part of the transcription factor IIH complex and interacts with the p44 and the MAT1 subunit of the TFIIH holocomplex (Coin et al. (1998); Busso et al. (2000); Feaver et al. (2000)). These interactions regulate the enzymatic activity of hXPD. In archaea, XPD appears to be a monomer in solution and has no obvious necessity for a binding partner to stimulate its helicase activity (Rudolf et al. (2006)). There is no current evidence supporting the idea that the homologue of hXPD in archaea has the same function in DNA repair. One avenue to investigate its role in the archaea, would be to

identify interacting proteins. As part of the NER machinery, some interactions would be expected in analogy to hXPD, for example with the 3' endonuclease XPG or the 3'→5' helicase XPB (Iyer et al. (1996)).

XPD protein-protein interactions were investigated using Affi-Gel 10 affinity matrix to immobilize untagged recombinant XPD on gel agarose beads.

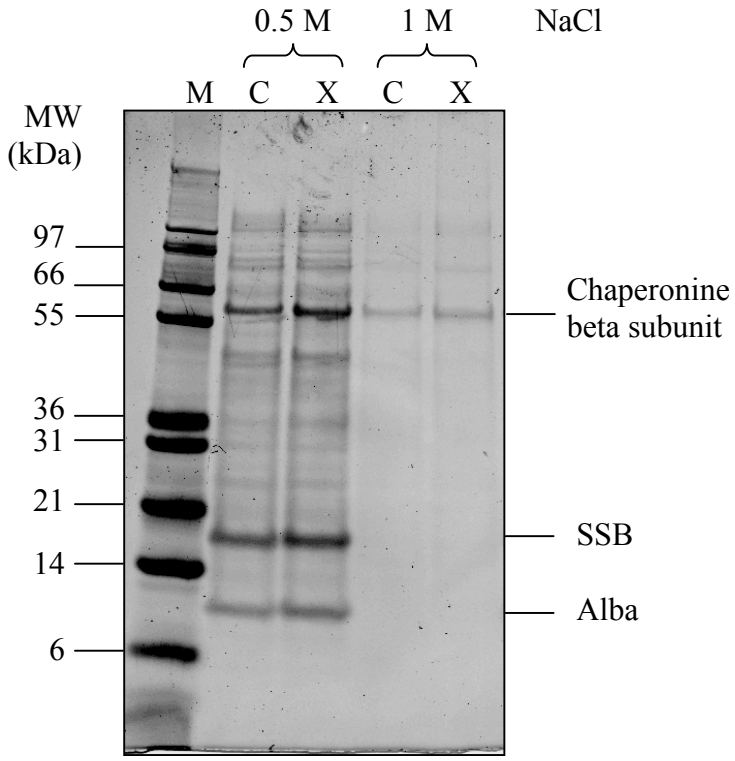


Figure 5.15: Protein-protein interactions of XPD

A Sypro-stained SDS-gel of Affi-Gel column elutions is shown. *S. acidocaldarius* cell lysate was applied to a column containing Affi-Gel with prebound recombinant XPD (X) or a control column without XPD (C). Lanes show wash steps with 0.5 and 1 M NaCl. M, Mark12 Standard Marker (Invitrogen). Major bands and corresponding proteins are indicated.

After XPD was bound to the Affi-Gel matrix, remaining free active ester groups were blocked with ethanolamine (pH 8.0). After the column was extensively washed with binding buffer (20 mM HEPES, pH 7.8, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 10 % glycerol), agarose beads were analysed by SDS-PAGE to confirm binding of XPD to the column matrix. The control was treated in the same way, only that no protein was

prebound to the column matrix and all active ester groups were blocked with ethanolamine. *S. acidocaldarius* cell lysate was applied to each column and after extensive washing, proteins bound to the column were eluted with increasing concentrations of NaCl (Figure 5.15).

No significant protein-protein interactions were observed for archaeal XPD (Figure 5.15). There are several reasons why no binding partner was detected: (1) the experimental conditions need to be further optimised, (2) the interaction is weak, transient and unstable or (3) the interaction with other proteins only occurs if XPD is bound to DNA. Another possibility that a binding partner could not be detected because it is not (highly) expressed under normal growth conditions, was investigated using UV-irradiated *S. acidocaldarius* cells. This experiment led to the same result and no XPD protein-protein interactions could be detected (data not shown).

5.6 SUMMARY AND CONCLUSIONS

This chapter described possible domain structures of XPD, as well as interactions with DNA and other proteins.

The tryptic digestion pattern of wildtype XPD and a mutant without FeS cluster (C88S) revealed subtle differences, based on an altered trypsin sensitive site near the N-terminus of the proteins. This clearly indicated local changes in secondary or tertiary structure in this region depending on the presence of the FeS cluster. Putative domains in XPD were mapped according to the configuration of domains found in other SF1 and 2 helicases. Domain mapping suggested that the FeS cluster domain represents domain 1B as an insertion within RecA motor domain 1A and could be involved in the interaction with the ss/ ds junction. In the previous chapters it was proposed that the FeS cluster domain of XPD forms a loop structure similar to the beta-hairpin of UvrB. Accordingly, XPD could use this loop structure as a “clamp” to engage ssDNA and actively distort the ss/ dsDNA region. Alternatively, it could clamp the single DNA strand to the two motor domains in such a way that it comes in close proximity to the aromatic-rich loop, which can then couple the free energy derived from ATP hydrolysis onto translocation.

Binding of ADP protected all trypsin sensitive sites in wildtype XPD, suggesting domain rearrangement and a change from the open to the closed conformation of the enzyme (Korolev et al. (1997)). In contrast, in the clusterfree mutant C88S the first tryptic sensitive site was unprotected and located within the FeS cluster domain. Mutants without FeS cluster were able to efficiently hydrolyse ATP (see previous chapter) indicating the correct formation of the ATP binding pocket. Therefore, the cluster is not involved in promotion of the correct domain movement of the Walker A box relative to the other motifs.

The [4Fe-4S] clusters of the glycosylases MutY/ EndoIII are involved in the formation of the FeS cluster loop (FCL), which facilitates DNA interaction (Thayer et al. (1995); Chepanoske et al. (2000)). The FeS cluster domain of XPD was shown not to promote DNA binding to 5' overhang and duplex DNA substrates as no changes in affinities were observed between wildtype and mutant proteins with and without FeS cluster. Therefore, it is unlikely that the FeS cluster domain in XPD has an analogous structural function to the FCL in MutY/ EndoIII.

The mutation R531W, which is located at the C-terminus of the protein led to around 10-fold reduction in DNA binding compared to the wildtype. In humans, this mutation (hXPD: R683W) abolishes the interaction with p44, which stimulates the helicase activity of hXPD 10-fold (Coin et al. (1998)). This leads to a significantly reduced helicase activity (Dubaele et al. (2003)). In archaea, XPD does not interact with a TFIIH analogous complex, however the higher K_D seen for DNA binding leads to reduced ATPase activity and abolishes the helicase activity.

The binding of XPD to DNA was very weak in general, subsequently leading to a weak helicase activity. At least a 5-fold excess of XPD was required to detect any helicase activity (Chapter 3). Single stranded DNA was shown not to be a good substrate for XPD-DNA binding, as observed affinities were very low. This effect could arise from the fact that the ssDNA oligonucleotides used in these experiments were much shorter than dsDNA constructs. In fact it was found that ssDNA binding affinities increased with increasing length of the oligonucleotide. ATP did not dramatically alter binding affinities to ds or ssDNA. Possible reasons could be non-optimal experimental conditions to

produce the desired effects or the effects were too subtle to be observed by gel shift assays.

Some helicases exist as monomers in solution (inactive form) and form oligomers upon binding to DNA (active form). In this chapter, it was confirmed that the active form of XPD in archaea is a monomer, strengthening the hypothesis that the mechanism of action of this helicase follows the Inchworm Model. Preliminary data also suggest that XPD binds tighter to DNA constructs containing a bubble, compared to duplex DNA or duplex DNA containing a 5' overhang. This does not necessarily imply that XPD is a damage recognition factor, yet it could imply indirect damage recognition by defined substrate preference. The SSB protein was proposed to be a possible candidate for early stage damage recognition (Cubeddu & White (2005)) and the emerging bubble around the damage after duplex melting by SSB could lead to repair factor recruitment. However, no protein-protein interactions could be shown for XPD under the experimental conditions used, suggesting that the XPD interaction with other proteins is rather weak or only transient. Different approaches will be undertaken to further address the question of identifying putative interacting partners of XPD.

6 THE ARCHAEAL ENDONUCLEASE XPF

6.1 INTRODUCTION

XPF is a structure specific endonuclease, which generates the 5' incision during NER (Sijbers et al. (1996); Komori et al. (2002); Roberts et al (2003)) and takes part in recombination (Davis et al. (1995)) and interstrand cross-link repair (Kuraoka et al. (2000)). XPF belongs to a large nuclease superfamily which includes type II restriction enzymes, Holliday junction resolving enzymes and endonuclease I (Nishino et al. (2003)). These enzymes are dependent on divalent metal ions for catalysis and possess the highly conserved signature sequence V/I-E-R-K-X_3-D, which represents the catalytic center (Aravind et al. (1999)).

The eukaryal XPF (*Saccharomyces cerevisiae*: Rad1) forms a heterodimer with ERCC1 (*S. cerevisiae*: Rad10), a feature that is essential for stability of XPF, protein-protein interactions with XPA and DNA binding. XPF contains an inactive N-terminal SF2-like helicase domain, a central nuclease and a C-terminal double Helix-hairpin-Helix (HhH_2) domain (Figure 6.1). ERCC1 consists only of the nuclease domain, which is inactive, and the HhH_2 motifs (Bardwell et al. (1992); Park et al. (1995); Sgouros et al. (1999)). Helix-hairpin-helix motifs are found in a variety of DNA-interacting proteins and are thought to mediate non-sequence specific DNA interactions (Shao & Grishin (2000)).

In Euryarchaea, XPF has essentially the same domain structure as its eukaryal counterpart, but displays an active helicase fused to the N-terminus (Figure 6.1). The enzyme forms a homodimer (Komori et al. (2002)). XPF is also present as a homodimer in crenarchaea, but consists only of the nuclease and HhH_2 domains. In addition, the nuclease activity of crenarchaeal XPF is stimulated and has an absolute requirement for the proliferating cell nuclear antigen protein (PCNA). PCNA is bound via a C-terminal PCNA interaction motif (PIP; Figure 6.1) that is found in a variety of proteins. Deletion of the PIP motif abolishes the nuclease activity of *S. solfataricus* XPF (Roberts et al. (2003)).

PCNA forms a heterotrimer in archaea (Dionne et al. (2003)), while it is a homotrimer in eukarya (Krishna et al. (1994)). It forms a ring-like structure encircling DNA and functions generally as a sliding clamp, mediating interactions between several

proteins and DNA. Interacting proteins include FEN1/ XPG, DNA polymerase δ and DNA-ligase (Warbrick (2000)).

Archaeal XPF shows substrate specificity for stalled replication forks and nicked Holliday junctions suggesting a role in recombination and replication. In eukarya these substrates are processed by Mus81, which dimerises with Mms4 in *S. cerevisiae* and Eme1 in *Schizosaccharomyces pombe*. Crenarchaeal XPF also cleaves XPF-like substrates with lower preference. Eukaryal XPF favours cleavage of DNA 3' to pyrimidines, whereas Mus81 and probably also crenarchaeal XPF shows no sequence preference. XPF from crenarchaea is thought to be the ancestral version of XPF-like endonucleases (deLaat et al. (1998a); Kaliraman et al. (2001); Boddy et al. (2001); Bastin-Shanower et al. (2003); Roberts et al. (2003); Roberts & White (2005a)).

A third eukaryal XPF-like homologue has been described recently and is called FancM. It consists of an N-terminal active helicase and and C-terminal inactive nuclease domain. FancM is an essential component of the Fanconi anemia DNA repair pathway (see Chapter 1). It most likely forms a heterodimer, but the binding partner has not yet been identified. A function for FancM has also not yet been assigned (Meetei et al. (2005)).

In this chapter the cloning and purification of the HhH_2 motif of XPF from *S. solfataricus* is described complemented by a short general characterization. DNA binding of full-length XPF is also addressed as well as the influence of several amino acid residues on the catalytic activity.

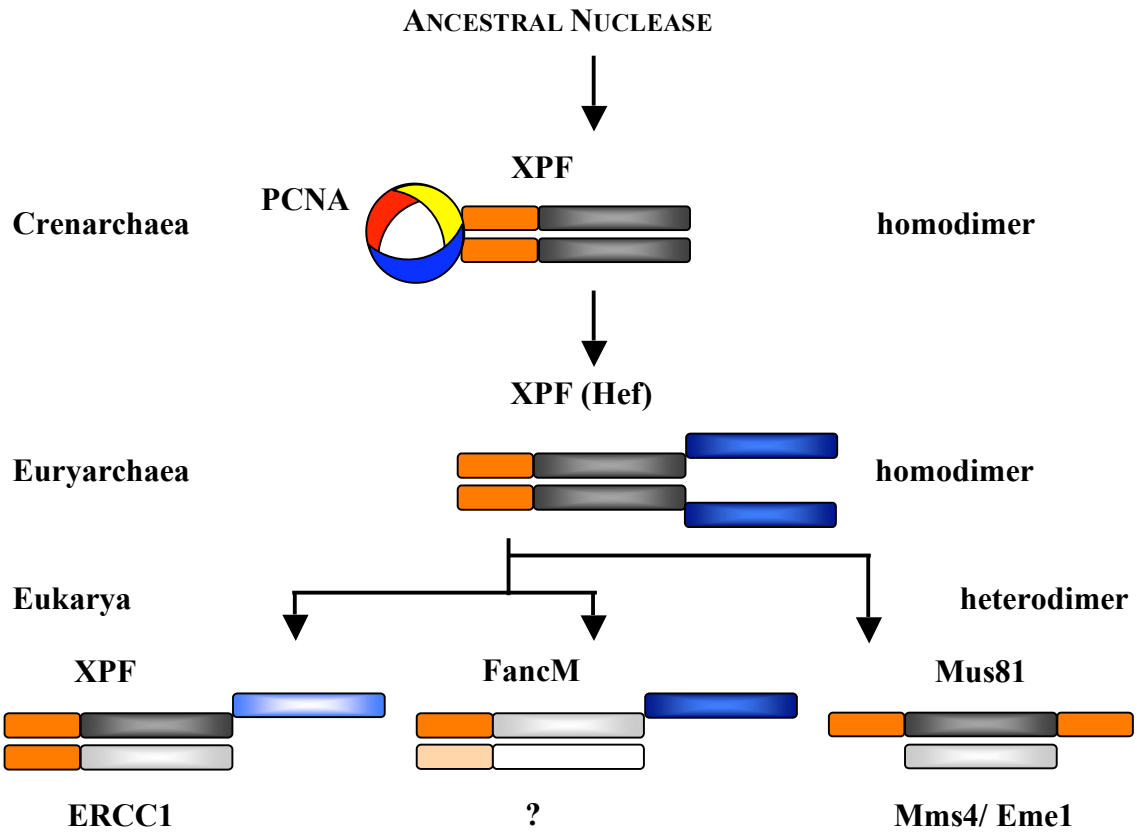


Figure 6.1: Domain organisation of XPF and Mus81 in Archaea and Eukarya

The schematic shows the domain organization of XPF within the domains of the archaea and the eukarya. Crenarchaeal XPF probably represents an ancestral form of XPF from which all other XPF types evolved. The domains are indicated as follows: Orange, the Helix-hairpin-Helix domain; dark gray, intact nuclease domain; light gray, inactive nuclease domain; dark blue, active helicase; light blue, inactive helicase. A binding partner has not yet been described for FancM. The heterotrimeric PCNA is shown as a circle in yellow-red-blue.

6.2 PURIFICATION OF THE SSO-XPF-HHH MOTIF

The Helix-hairpin-Helix (HhH) motif of *Sulfolobus solfataricus* XPF (SsoXPF) was PCR amplified and cloned into the *Bam*H1/ *Sal*1 cloning sites of the pET28c vector for expression of histidine tagged protein. The gene fragment was fully sequenced to confirm the absence of PCR mediated mutations. Protein was expressed in *Escherichia coli* and purified using a heat step method to remove the majority of *E. coli* proteins followed by a two-step purification scheme, consisting of nickel-chelating and size exclusion chromatography (Figure 6.2C). The calculated molecular weight (MW) of the his-tagged monomer was 13.1 kDa. The elution volume of the peak fraction in gel filtration chromatography was 218 mL (Figure 6.2A), corresponding to an approximate MW of 29 kDa (Figure 6.2B). Although the molecular masses calculated from size exclusion chromatography are only approximate, the mass of the HhH_2 protein in solution was clearly that of a dimer.

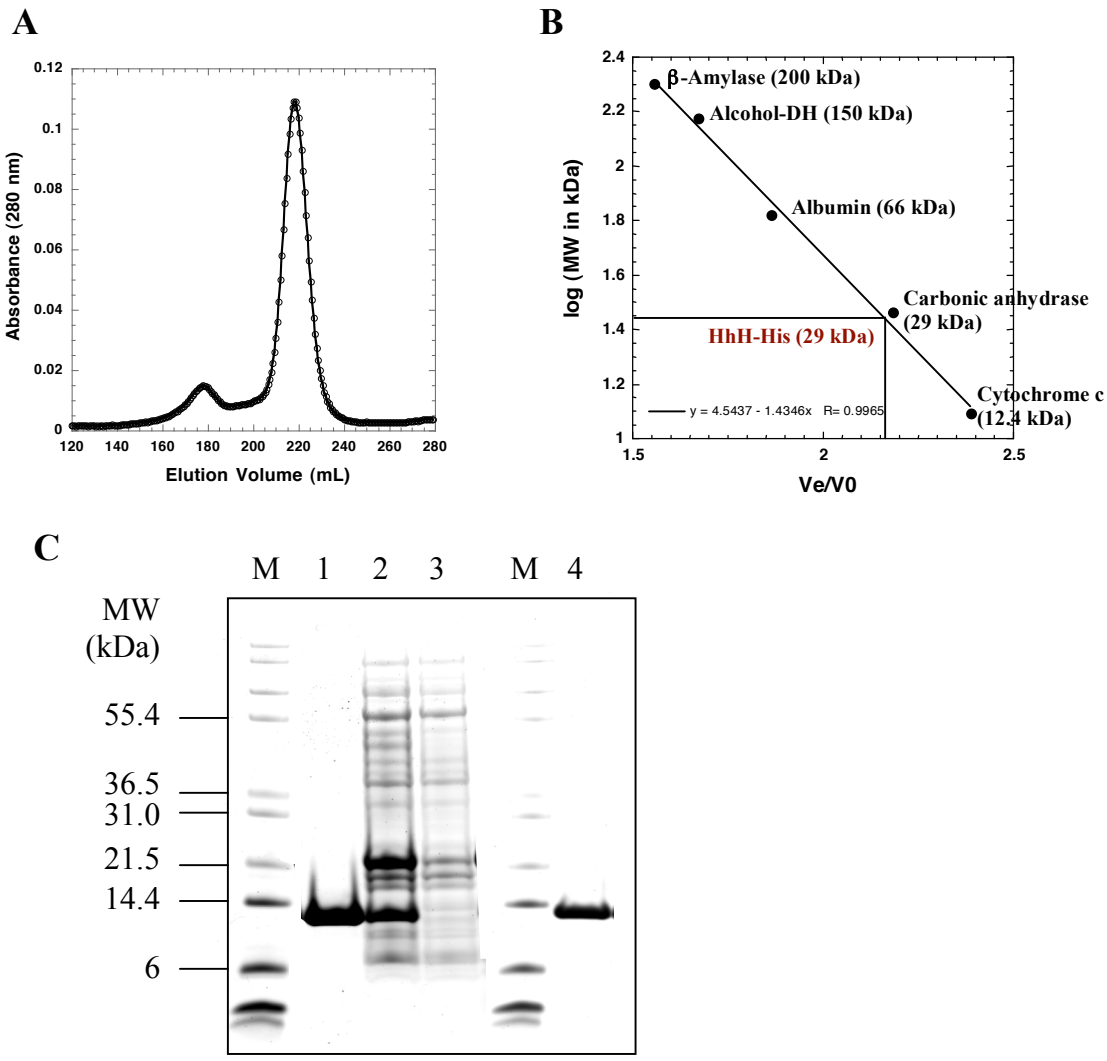


Figure 6.2: Purification of the histidine tagged HhH_2 domain of SsoXPF

The elution chromatogram (A) and the calculation of the molecular mass of SsoXPF-HhH_2 from the elution volume (218 mL) after gel filtration (B). SDS-PAGE analysis showing HhH purification steps (C). From left to right: pooled fractions after elution from nickel chelating column (1), *E. coli* lysate after heat treatment at 65 ˚C (2), flow through after Ni-column (3) and pooled fractions containing SsoXPF-HhH_2 after gel filtration chromatography (4). Marker (M): Mark12 Standard Marker (Invitrogen).

6.2.1 Oligomeric state of the HhH_2-domain

The isolated nuclease domains and full-length SsoXPF form homodimers (Roberts (2004)). Dimerization is a common theme in the XPF protein family. The C-terminal part of the euryarchaeal Hef protein from *Pyrococcus furiosus*, including the nuclease and HhH_2 domains, forms a homodimer as well (Nishino et al. (2003)). In contrast the eukaryal XPF/ Mus81 proteins exist as heterodimers with only one active nuclease (Park et al. (1995)). Glutaraldehyde cross-linking was used to confirm the hypothesis that both the HhH_2 and the nuclease domains dimerise in crenarchaeal XPF.

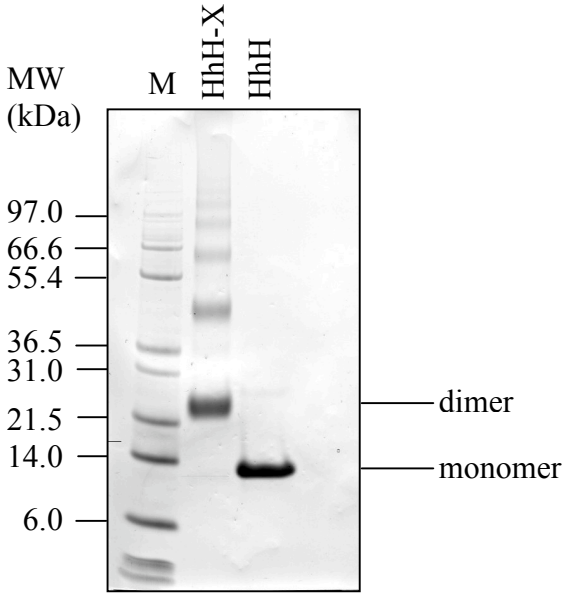


Figure 6.3: Glutaraldehyde cross-linking of SsoXPF-HhH_2

Cross-linking experiments show dimer-formation of the HhH_2 domains upon treatment with glutaraldehyde (HhH-X), compared to the control (HhH). M, Mark12 Standard Marker (Invitrogen).

Glutaraldehyde cross-linking confirmed what was already clear from gel filtration (Figure 6.2) that the HhH-domains of XPF dimerise (Figure 6.3). Therefore, the HhH_2 motif not only mediates DNA interaction, but also contributes to the interaction between both XPF monomers, to form a stable dimeric molecule.

6.3 STRUCTURE AND FUNCTION OF THE HHH-MOTIFS

HhH-motifs are found in many different DNA binding proteins, including polymerases, ligases and glycosylases. HhH motifs show more structural than sequence conservation and are generally arranged in pairs consisting of five helices (Thayer et al. (1995); Doherty et al. (1996)). Two antiparallel α-helices bridged by a hairpin form the Helix-hairpin-Helix structure (Figure 6.4B). The hairpin promotes contact with the DNA double helix. The two HhH-pairs are connected by a fifth helix. A HhH-structure allows an asymmetric protein to obtain a pseudo-2-fold symmetry to be able to interact with the highly symmetric DNA molecule. The DNA is bound via the phosphodiester backbone, mainly by hydrogen bonds between peptide backbone nitrogens and oxygens of the DNA phosphate groups. The presence of two consecutive HhH motifs also allows the protein to contact both DNA strands at the same time (Figure 6.4A). Most proteins containing this motif exist as dimers, hence the HhH-motifs also mediate protein-protein-interactions. This enables the protein to bind two different sites of the DNA simultaneously, which leads to bending of the DNA molecule (Shao et al. (2000).

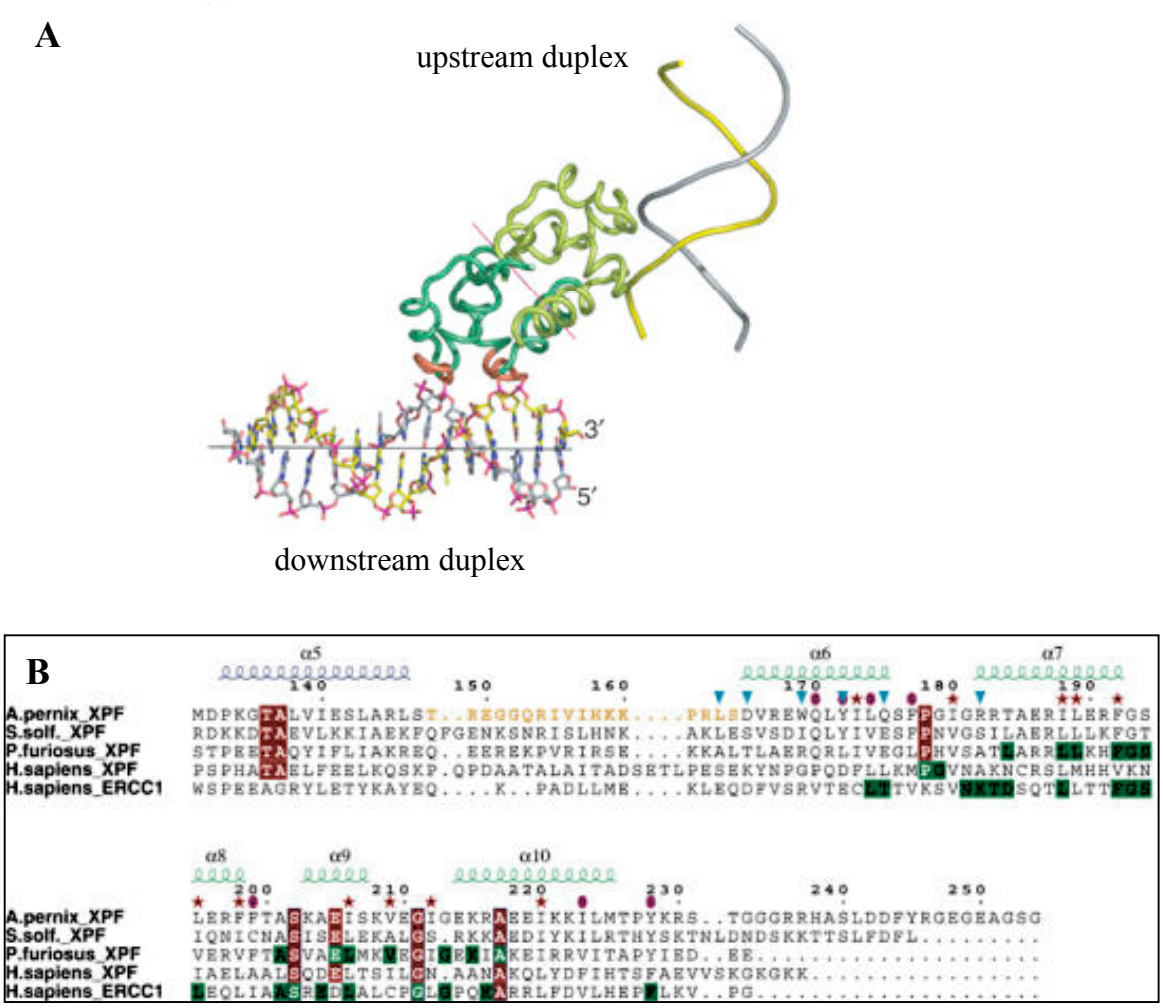


Figure 6.4: The HhH_2 domains of crenarchaeal XPF show structural rather than sequence conservation

(A) Crystal structure of *A. pernix* XPF showing the two HhH_2 motifs in dark and light green bound to a DNA duplex. The two loops of each HhH_2 motif contact both strands of the DNA duplex in the minor groove. (B) Sequence alignment of XPF from different archaea and human. The positions of the five helical structures are referred to as α6 to 10 (adapted from Newman et al. (2005)).

Indeed, for the *Aeropyrum pernix* (Ape) XPF structure it was shown that the HhH_2 domains of each monomer contact the downstream and upstream duplex simultaneously (Figure 6.4A). Initial binding of the downstream duplex by protomer A triggers the linker (between HhH_2 and nuclease domain) to flex and kink the DNA duplex around 90 °. Each hairpin contacts one phosphate of different strands in the minor groove of the DNA backbone, which leads to binding of XPF with a specific polarity. The HhH_2 domains

were not only found to be responsible for protein-DNA interactions, but also for positioning the nuclease domain through the action of the flexible linker region (Newman et al. (2005)).

6.4 DNA BINDING OF SSOXPF AND THE SSOXPF-HHH$_2$ DOMAINS

6.4.1 HhH$_2$-DNA interactions

The HhH$_2$ domains are thought to promote non-sequence specific DNA binding (Shao & Grishin (2000)). The DNA binding affinity of the isolated HhH$_2$ domain of SsoXPF was tested by gel shift experiments using a 3'-flap (Mus81-like) and a splayed duplex (XPF-like) DNA substrate. Roberts & White (2005a) showed, that SsoXPF has a preference for Mus81-like DNA substrates and cleaves XPF-like structures with lower rates. This preference was expected to be reflected in the DNA binding affinity of the HhH$_2$ domains of SsoXPF.

The observed binding affinity of the HhH$_2$ domains to DNA was very weak (Figure 6.5). Apparent K_Ds were around 15 μM for splayed duplex DNA substrates and 19 μM (-PCNA and 26 μM (+PCNA), respectively, for 3'-flap DNA substrates. PCNA did not increase the binding affinity to DNA. It is widely accepted that the HhH motifs mediate DNA binding (Shao et al. (2000)). The crystal structure of ApeXPF revealed how these domains bind to DNA (Newman et al. (2005)). It is possible that the HhH$_2$ domains without the nuclease domain cannot engage the DNA substrate in a proper and efficient way. It was shown that ApeXPF also makes interdomain contacts upon binding of DNA, which probably stabilizes the enzyme on the DNA (Newman et al. (2005)). It is possible that PCNA has no influence on DNA binding at all and only stabilizes the complex once XPF is already bound. It may induce some conformational changes enabling the nuclease to engage the 3' single DNA strand properly, which will subsequently be subject to incision activity.

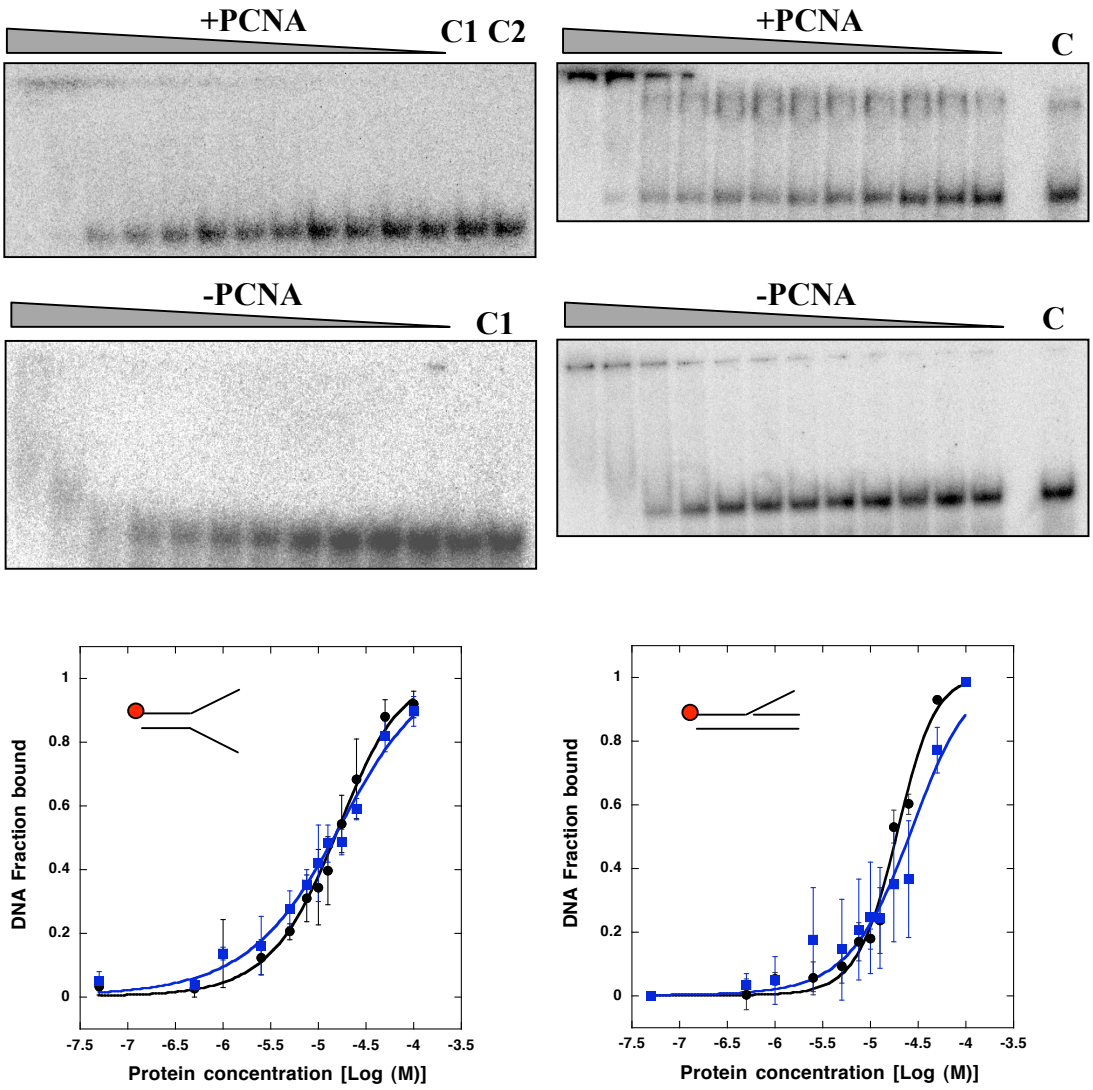


Figure 6.5: DNA binding of the HhH_2 domains

Representative 8 % acryl amide gels are shown on the top, for splayed duplex DNA substrates (left) and 3'-flap DNA substrates (right). Protein (HhH_2) concentration were: 100, 50, 25, 17.5, 12.5, 10, 7.5, 5, 2.5, 1, 0.5, 0.05 and 0 (C) µM. C1, no HhH_2/ no PCNA; C2, no HhH_2/ 20 µM PCNA. The respective quantification results of triplicate experiments are shown beneath the gels. The DNA fractions bound to the HhH_2 protein are plotted against the logarithm of the protein concentration (M). The data points were fitted to the following equation: Fraction bound = 1/ (1+ (K_D/ [XPD])^m) (for explanation see Chapter 2) Standard errors are indicated. Apparent K_Ds (µM) were: 14.8 ± 1.1 (-PCNA, black) and 15.1 ± 1.1 (+PCNA, blue) for splayed duplex substrates and 18.9 ± 1.0 (-PCNA, black) and 25.8 ± 1.1 (+PCNA, blue) for 3'-flap substrates.

6.4.2 DNA binding of full-length SsoXPF

To investigate whether HhH_2-mediated DNA binding depends on the presence of the nuclease domains of the XPF protein, untagged full-length wildtype SsoXPF was used in the presence and absence of PCNA. No DNA binding could be observed on a splayed duplex substrate (Figure 6.6). It is known that individual repair factors have very low affinities for DNA, but cooperativity, molecular matchmaking and kinetic proof-reading increases the affinity of an individual protein to DNA and other proteins and thus facilitates assembly of the repair machinery (reviewed in Reardon & Sancar (2005)). In solution experiments (e.g. using ITC or fluorescence polarization experiments) could give better data for XPF-DNA binding. It is possible that the protein-DNA complex was not stable for the duration of the gelshift experiments.

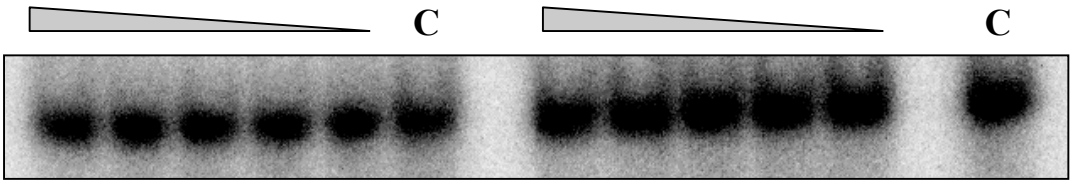


Figure 6.6: DNA binding of untagged WT-SsoXPF

Gel shift experiments showing amounts of unbound DNA in WT-XPF DNA binding experiments. WT-XPF was tested on a splayed duplex substrate with (right) and without (left) PCNA. Protein concentrations were in descending order (from left to right): 100, 50, 25, 12.5 and 6.25 μM WT-XPF. C, control without XPF protein.

An inactive mutant of SsoXPF was also tested for its ability to bind DNA. XPF-D52 is a highly conserved residue and the mutant (D52A) was shown to be completely inactive (Roberts et al. (2003)). The equivalent residue in ApeXPF, D68, is part of the hydrophobic cleft, which was proposed to accommodate the free 3' single DNA strand which becomes subject to incision by the nuclease domain (Newman et al. (2005)).

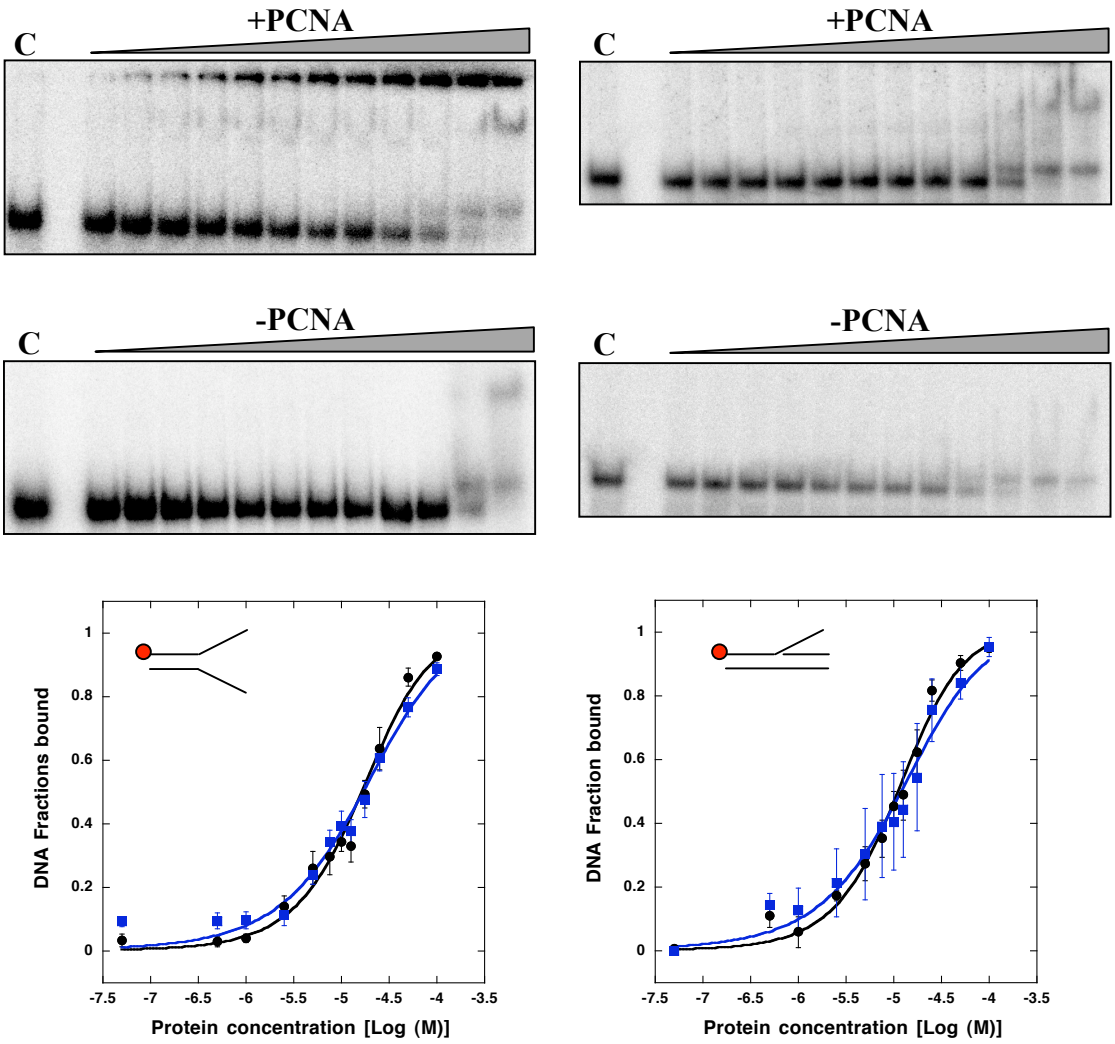


Figure 6.7: DNA binding of the SsoXPF-D52A mutant

Representative 8 % acryl amide gels are shown on the top, for splayed duplex DNA substrates (left) and 3'-flap DNA substrates (right). Protein (XPF-D52A) concentration were: 100, 50, 25, 17.5, 12.5, 10, 7.5, 5, 2.5, 1, 0.5, 0.05 and 0 (C) µM. The respective quantification results of triplicate experiments are shown beneath the gels. The DNA fractions bound to the XPF-D52A protein are plotted against the logarithm of the protein concentration (M). The data points were fitted to the following equation: Fraction bound = 1/ (1+ (K_D/ [XPD])^m) (for explanation see Chapter 2) Standard errors are indicated. Apparent K_Ds (µM) were: 16.6 ± 1.1 (-PCNA, black) and 17.1 ± 1.1 (+PCNA, blue) for splayed duplex substrates and 11.3 ± 1.1 (-PCNA, black) and 12.2 ± 1.1 (+PCNA, blue) for 3'-flap substrates.

In contrast to the wildtype protein, the inactive SsoXPF mutant D52A was able to bind DNA, although the reason for this is unclear (Figure 6.7). Apparent K_Ds for splayed duplex DNA binding were similar to those observed for the HhH_2 domains alone and K_Ds for binding a 3'-flap were approximately half as those obtained for the HhH_2 (Figure 6.5 and Figure 6.7). However, these differences are not significant and further experiments have to be carried out. As shown for the HhH_2 motif, PCNA did not show an influence on DNA binding affinities.

The D52A mutant was shown to lack nuclease activity (Roberts et al. (2003)). The same result was obtained for the analogous human XPF mutant D720A (Enzylin & Scharer (2002)). Residue D720 is not involved in metal ion coordination and showed no influence on DNA binding activity. Thus, the function of this acidic residue is not clear. It was suggested that it could be involved in the activation of water for the nucleophilic attack on the phosphodiester bond. Therefore, DNA binding effects can almost certainly be attributed solely to the HhH_2 domains. There is still no explanation why DNA binding could not be observed with the wildtype XPF protein. It is possible that the D52A mutant obtains a slightly different structure which enables it to bind DNA tighter compared to the wildtype protein.

DNA binding was also tested towards dsDNA and ssDNA. Both for the D52A and the HhH_2 dimer binding was observed, although only very weak. Apparent K_Ds were around ~60 μM for the HhH_2 domains and ~50 μM for the D52A mutant.

6.5 SSOXPF MUTAGENESIS

6.5.1 Selection of XPF mutants

The ApeXPF crystal structure was solved in complex with duplex DNA (Figure 6.8A). Although duplex DNA is not a substrate for XPF, one of the HhH_2 domains of the ApeXPF homodimer bound to the DNA (position referred to as downstream duplex). This XPF monomer is referred to as protomer A. Binding of one of the HhH_2 domains to DNA induced conformational changes in the XPF homodimer mediated by a flexible linker region that made it possible to model the missing upstream duplex bound to the HhH_2 domain of protomer B. In addition, the position of the nuclease domain of protomer A allowed for the proposition of a model of DNA substrate binding and incision. Thus, it was suggested that the free 3' single DNA strand, which becomes subject to incision, enters a hydrophobic cleft in the nuclease domain of protomer A lined by the active site residues E-R-K (Figure 6.8B; highlighted in yellow in Figure 6.9). The conserved residues of the hydrophobic cleft were thought to make contacts to the phosphodiester backbone of the DNA stabilizing the 3' single strand. Here the importance of some of the conserved residues in the hydrophobic cleft was investigated by mutating analogous residues in SsoXPF to alanine.

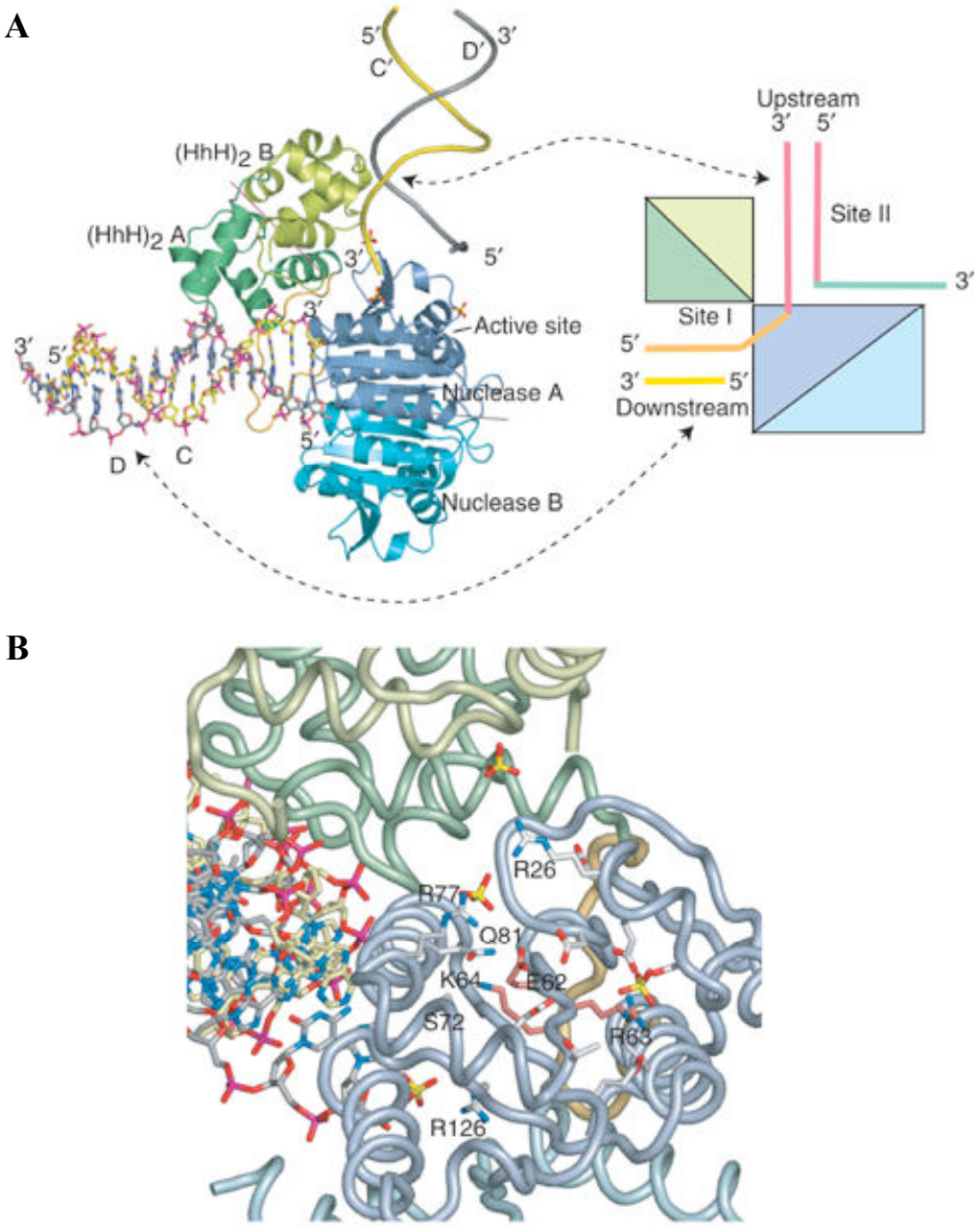


Figure 6.8: Crystal structure of ApeXPF, modes of DNA binding and the hydrophobic cleft

(A) The HhH_2 domain of protomer A (dark green) contacts the downstream duplex of the DNA substrate, which induces a rotation of about 95 ˚ relative to the nuclease domain. The DNA substrate is kinked so that the free 3' ssDNA can enter the hydrophobic cleft of the nuclease domain of protomer A (dark blue). (B) The residues S72, R77, F79 (not shown) and Q81 are believed to be involved in stabilizing the single strand and coordinate it relative to the active site E(62)-R(63)-K(64) which is on the bottom of the hydrophobic cleft. The nuclease domain of protomer B is inactive (cyan (A)), but is involved in stabilisation of the whole protein complex on the DNA substrate by facilitating interdomain contacts (adapted from Newman et al. (2005))

XPF from *S. solfataricus* was preferentially used for mutagenesis and kinetic assays, because it showed a more robust enzymatic activity and the *Sulfolobus* PCNA heterotrimer was readily available. Mutations in SsoXPF were selected according to conserved residues in the ApeXPF crystal structure (highlighted in green, Figure 6.9).

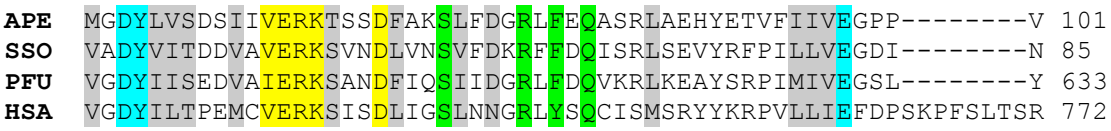


Figure 6.9: Sequence alignment of archaeal and human XPF

Sequence alignment showing the active site residues of XPF proteins from different archaea and human. The active site motif V/I-E-R-K-X_3-D is highlighted in yellow, absolutely conserved residues are in blue, similar residues are in grey and residues that were subject to mutagenesis are highlighted in green. APE, *A. pernix* XPF; SSO, *S. solfataricus* XPF; PFU, *Pyrococcus furiosus* hef; HSA, *Homo sapiens* XPF.

As described in Newman et al. (2005), ApeXPF also has a strict requirement for the PCNA sliding clamp (Figure 6.10). Endonuclease activity of ApeXPF was shown to be stimulated by the SsoPCNA heterotrimer, while no activity was observed without PCNA.

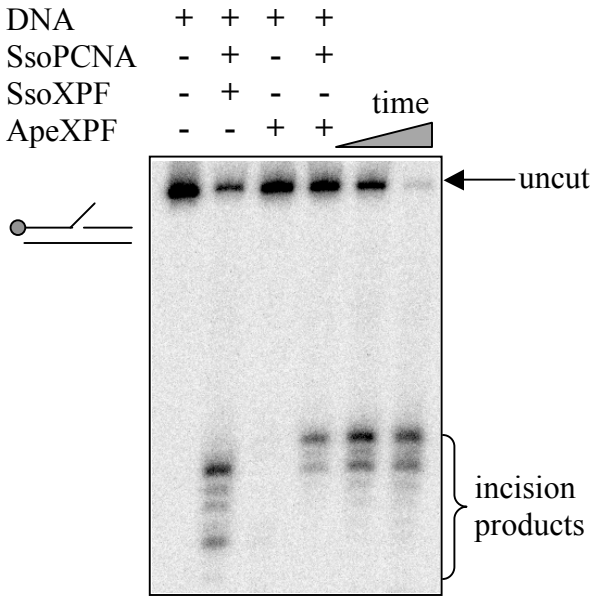


Figure 6.10: Stimulation of ApeXPF nuclease activity by SsoPCNA

Products of endonuclease reactions were analysed on a denaturing 12 % polyacrylamide gel. Radiolabeled DNA substrates and incision products were detected by phosphoimaging. Nuclease activity of SsoXPF and ApeXPF was monitored using a 3'-flap DNA substrate. SsoPCNA heterotrimer was used for stimulation of ApeXPF nuclease activity. Time points were: 10 sec (SsoXPF + SsoPCNA); 10 min (ApeXPF); 1, 3, 10 min (ApeXPF + SsoPCNA).

XPF-mutants were generated using the QuikChange Mutagenesis Kit (Stratagene) and respective mutant proteins (Figure 6.11) were purified as described for the wildtype XPF protein (see Material and Methods). The mutations and the whole protein mass were confirmed by mass spectrometry.

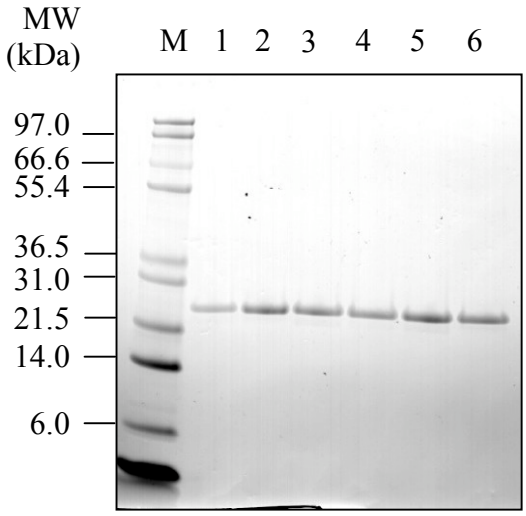


Figure 6.11: SDS-PAGE of purified XPF mutants

The gel shows mutant XPF proteins after purification. M, Mark12 Standard Marker (Invitrogen); XPF mutants: 1, S56A; 2, R61A; 3, F63A; 4, Q65A; 5, Y108A; 6, wildtype.

6.6 CHARACTERISATION OF THE CATALYTIC CORE OF SSOXPF/ APEXPF USING MUTAGENESIS

The XPF mutant and the wildtype proteins were characterised by their ability to process different DNA substrates. DNA constructs containing a 3' single stranded flap or nick are preferrentially cleaved by SsoXPF, while a splayed duplex DNA substrate is cleaved with lower rates (Roberts & White (2005a)). Initially, a 3'-flap and a splayed duplex substrate were used for endonuclease assays, as they are the preferred substrates for human Mus81/ Eme1 and XPF/ ERCC1, respectively. The nicked duplex was used to confirm data obtained for the 3'-flap substrate. Firstly, all assays were carried out using 1 μM protein. To confirm the data under more realistic single turnover conditions, another set of experiments was performed using 7 μM SsoXPF protein. The Y108A mutant was

not further characterised as it showed wildtype activity in initial nuclease assays. Experiments were carried out as described in Material and Methods.

Single substrate reactions of enzymatic catalysis can be easily described with Equation 6.1, where $k_{1/-1}$ and $k_{2/-2}$ are equilibrium rate constants and k_{cat}, the catalytic constant that describes the catalytic step of product formation. Under conditions where the enzyme (E) is in excess over the substrate (S) the enzyme-substrate (ES) complex is favored and the catalytic constant k_{cat} can be obtained directly. Cleavage products of SsoXPF were separated on a denaturing polyacrylamide gel, visualized by phosphoimaging and quantified (Image Gauge Software). The mean ratio of total DNA counts (cut + uncut) to uncut counts was used to determine k_{cat}. k_{cat} was directly obtained by linear regression (ln (total DNA/ uncut) versus time) using the equation y = mx, where m corresponds to k_{cat}.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} EP \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P \qquad \text{(Equation 6.1)}$$

Representative gels for XPF mutants and wildtype nuclease activity with different DNA substrates are shown in Figure 6.13 and the quantification results are presented in Table 6.1 and Table 6.2. Previous results of a preference of SsoXPF for 3'-flaps and nicked duplexes were confirmed, indicating different modes of DNA binding (Roberts & White (2005a)). As explained in earlier sections, for stable protein-DNA binding it seems important that both the regions downstream and upstream of the incision site are double stranded (see also Figure 6.8A). Single stranded DNA in the downstream region, as present in a splayed duplex substrate, reduces the stability of the complex and therefore the enzymatic activity.

A reduction in activity was observed for all mutants, ranging from modest (S56A, F63A) to very severe effects (R61A, Q65A). In the crystal structure of ApeXPF the equivalent residues of SsoXPF-R61 (ApeXPF-R77) and Q65 (ApeXPF-Q81) were found to be located just outside the catalytic cleft (Figure 6.8) and a substrate recognition role was attributed to them. This would explain the severe effects a mutation to alanine had on catalysis. There was also a difference in the reduction of the incision rates between 3'-

flap and splayed duplex. While the wildtype showed approximately 10-fold reduction, the R61A and Q65A mutants showed 280-fold and 80-fold reduction, respectively (Table 6.1), confirming a role in DNA substrate recognition for these residues. The hydrophobic cleft is centered around F63 (ApeXPF-F79; Figure 6.12), however, mutation of this residue to alanine has only mild effects on the nuclease activity, suggesting that loss of only one hydrophobic residue can be tolerated. The reduction of the catalytic rates between 3'-flap versus splayed duplex is similar to the wildtype. The residue SsoXPF-S56 (ApeXPF-S72) is also located at the entrance to the hydrophobic cleft, but mutation to an alanine showed relatively mild effects on catalysis of a 3'-flap DNA substrate, in contrast to R61A or Q65A. However, a 100-fold reduction in endonuclease activity was observed when using a splayed duplex DNA construct instead of a 3'-flap, suggesting an involvement of S56 in DNA substrate recognition.

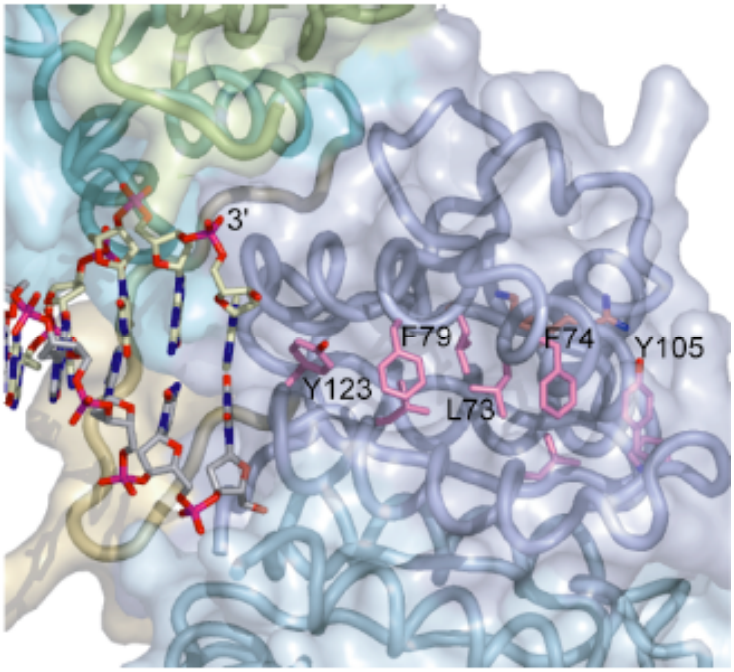


Figure 6.12: The hydrophobic cleft of the nuclease domain of ApeXPF

The hydrophobic cleft, centered around F79, could accommodate ssDNA. The hydrophobic residues may intercalate between bases, while positively charged residues could interact with the DNA backbone (adapted from Newman et al. (2005)).

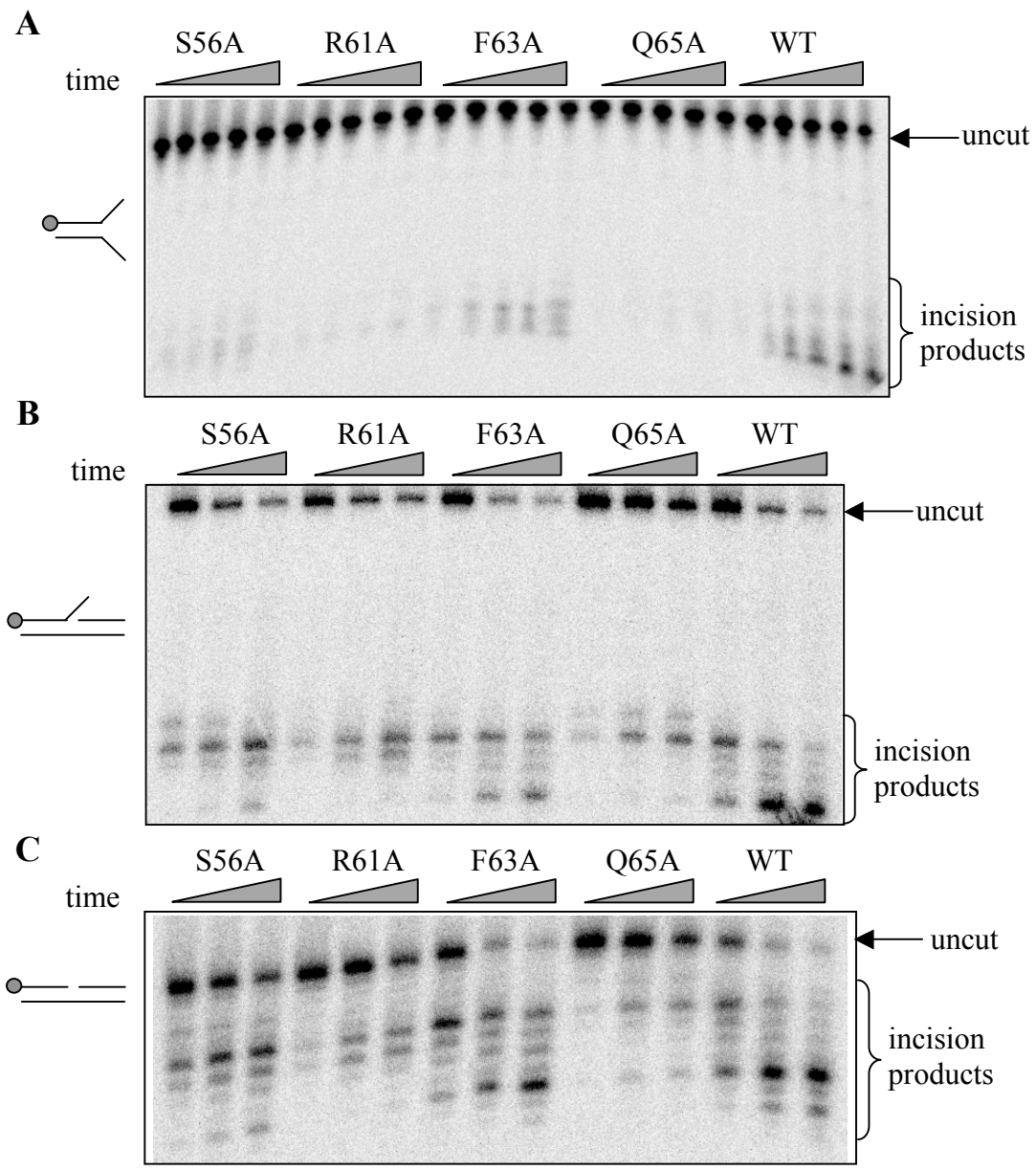


Figure 6.13: Nuclease activity of XPF mutants and wildtype on different DNA substrates

Selected examples of denaturing polyacryamide gels showing incision products of XPF mutants and wildtype on different DNA substrates. (A) splayed duplex, (B) 3'-flap and (C) nicked duplex. A schematic of the DNA constructs is shown left of the gel images. The different XPF proteins are indicated across the top of each gel. Time points were 5, 10, 15, 20, 30 min (S56A, R61A, Q65A); 1, 2, 3, 4, 5 min (F63A); 0.5, 1, 1.5, 2, 3 min (wildtype) for splayed duplex substrates. For 3'-flaps and nicked duplexes the time points were as follows 20, 40, 60 sec (S56A, F63A); 60, 120, 180 sec (R61A, Q65A); 10, 20, 30 sec (wildtype).

Table 6.1: Reaction rates of SsoXPF mutants and wildtype (7 µM) on splayed duplex, nicked duplex and 3'-flap DNA substrates

Mutation in SsoXPF	Equivalent residue in ApeXPF	Splayed duplex k_{cat} ± s.e. (min^{-1} x 10^3)*	Nicked duplex k_{cat} ± s.e. (min^{-1} x 10^3)*	3'-flap k_{cat} ± s.e. (min^{-1} x 10^3)*
WT	-	720 ± 49	9760 ± 1570	7200 ± 880
S56A	S72	33 ± 1	1730 ± 120	3120 ± 210
R61A	R77	3 ± 0.1	510 ± 13	840 ± 23
F63A	F79	277 ± 27	3190 ± 260	4730 ± 88
Q65A	Q81	3.3 ± 0.4	230 ± 10	260 ± 11

Table 6.2: Reaction rates of SsoXPF mutants and wildtype (1 µM) on splayed duplex, nicked duplex and 3'-flap DNA substrates

Mutation in SsoXPF	Equivalent residue in ApeXPF	Splayed duplex k_{cat} ± s.e. (min^{-1} x 10^3)**	Nicked duplex k_{cat} ± s.e. (min^{-1} x 10^3)*	3'-flap k_{cat} ± s.e. (min^{-1} x 10^3)**
WT	-	260 ± 47	5900 ± 400	2380 ± 230
S56A	S72	2.1 ± 0.5	1155 ± 35	810 ± 50
R61A	R77	0.3 ± 0.1	230 ±1	380 ± 7
F63A	F79	35 ± 11	2350 ± 106	1520 ± 220
Q65A	Q81	0.7 ± 0.2	170 ± 15	160 ± 11

Table 6.1 and Table 6.2 show k_{cat} of SsoXPF mutants and wildtype using different DNA substrates and protein concentrations. Sso, *S. solfataricus*; Ape*, A. pernix*; s.e., standard error; * data from a single experiment; ** data from triplicate experiments.

6.7 SUMMARY AND CONCLUSIONS

This chapter described the purification of the C-terminal HhH_2 domain of *S. solfataricus* XPF. When this project was initiated it was known that the isolated nuclease domain as well as the whole XPF protein forms a homodimer (Roberts (2004)). Although not likely to be the case, it still needed to be proven that the HhH_2 domains of XPF also dimerise. Glutaraldehyde cross-linking experiments confirmed this.

The HhH_2 motif is found in a variety of DNA binding proteins and facilitates non-sequence specific DNA binding (Doherty et al. (1996)). Different DNA constructs were tested, including a 3'-flap, the preferred substrate for SsoXPF endonuclease activity. The HhH_2 domains showed very weak DNA binding, with an apparent K_D of around 20-25 μM for a 3'-flap. Splayed duplexes were bound with K_Ds of 15 μM. Double and ssDNA was bound as well, with apparent K_Ds which were much higher. PCNA was shown not to have an influence on DNA binding. The observed weak DNA binding affinity could arise from improper folding of the isolated HhH_2 domains, from missing interdomain interactions with the nuclease domain or from the absence of other binding partners.

Subsequently, wildtype XPF was tested, but no DNA binding could be observed. On the contrary, when the mutant SsoXPF-D52A, which is deficient in nuclease activity was used, DNA binding was observed in gelshift experiments. Apparent K_Ds were similar to those of the isolated HhH_2 domains. The reason why DNA binding could be observed only for the mutant and not for the WT is not clear. One possibility is that the mutant obtains a slightly different structure enabling a stronger interaction with DNA. PCNA did not have any influence on D52A-DNA binding, matching the results obtained for the HhH_2 domains. This suggests that PCNA does not contribute in loading the enzyme onto the DNA, but it helps to keep it attached to the site of DNA damage repair (Dionne et al. (2003)). Binding to PCNA via its C-terminal interaction motif stimulates the nuclease activity of XPF, which seems not to be necessary for the initial step of DNA binding and substrate coordination.

With the solved crystal structure of ApeXPF, Newman et al. (2005) proposed that the two HhH_2 domains of each monomer engage the DNA substrate downstream and upstream of a lesion (Figure 6.4A). This results in large domain movement and the coordination of one of the nuclease domains relative to the 3' single DNA strand, which

will become subject to incision. It showed that the enzyme kinks the DNA around 90 °, which is thermodynamically easier if there is a lesion present in one of the DNA strands. The ssDNA enters a hydrophobic cleft within the nuclease domain of one of the XPF monomers (Figure 6.12). This hydrophobic cleft contains several conserved residues, which are thought to interact with the DNA and help to coordinate this strand to the active site which is located on the bottom of the hydrophobic cleft. Mutagenesis studies carried out on the XPF homologue from *S. solfataricus* showed that these residues are indeed important for the endonuclease activity. How they are involved in ssDNA binding has yet to be investigated by crystal structures of XPF for example in complex with longer and 3'-flapped DNA constructs, as the published structure was in complex with dsDNA (Newman et al. (2005)).

7 CONCLUSIONS AND FUTURE WORK

This PhD thesis described the characterisation of XPD from *S. acidocaldarius*. It showed that SacXPD, like its human homologue, functions as a 5'→3' monomeric SF2 DNA helicase. In contrast to eukarya, SacXPD does not appear to be associated with a higher multimeric protein complex. Helicase activity was detected in the presence of ATP and Mg^{2+} (Mn^{2+}) without the need for an additional binding partner. Gel filtration experiments have also shown that the active form of SacXPD is a monomer, as no higher complexes were observed upon binding to ssDNA. Therefore, the mechanism of action for XPD should follow the Inchworm model (Velankar et al. (1998)). Respective ssDNA and dsDNA binding domains of XPD still have to be located, using mutagenesis for example. A possible domain model was proposed in Chapter 5. It would be useful to obtain crystal structures of XPD in complex with ATP and ADP bound to a 5' overhang DNA substrate. More work will also be carried out on the DNA substrate specificity of XPD. Although SacXPD is a homologue of human XPD, it does not necessarily imply that it has an identical function in NER. In addition, as shown for XPF for example, XP proteins tend to participate in different DNA processing pathways.

SacXPD is the first described example of a helicase containing an iron-sulphur (FeS) cluster. A number of different methods were applied to investigate the nature of the FeS cluster, revealing that in solution XPD has a $[3Fe\text{-}4S]^0$ cluster bound. The EPR spectrum obtained showed a weak but significant signal for the oxidised form (1+) of the cluster. It was suggested that the FeS cluster is buried within the enzyme and is not easily accessible for oxidants. In contrast, helicase experiments showed a high sensitivity of XPD to even very low concentrations of the oxidant ferricyanide. This may be caused by conformational changes the enzyme undergoes while unwinding DNA. These changes could expose the FeS cluster to solvents, resulting in oxidation and subsequent enzyme inactivation. It will be necessary to optimise EPR experiments to obtain stronger signals for wildtype XPD and to clarify existing data. It would also be useful to obtain an EPR signal for the C102S mutant, which coordinated a stable FeS cluster. Despite several attempts, the FeS cluster of the C102S mutant could not be oxidised in solution and no EPR spectrum could be recorded. A successful EPR experiment would fully prove that the wildtype and C102S mutant bound exactly the same type of FeS cluster and that cysteine 102 is not involved in FeS cluster coordination in the wildtype protein.

The FeS cluster is located between the Walker A and B boxes and it has been shown to be essential for helicase activity. It was also shown that the FeS cluster is not necessary for ATP hydrolysis, DNA binding or general structural stability of the protein. Tryptic digestion experiments showed different digestion patterns for wildtype XPD and a mutant without the FeS cluster, indicating a local structural role. Sequence alignments revealed the presence of the FeS cluster domain in other helicases, including *E. coli* DinG and hXPD. Preliminary experimental data on DinG suggested the presence of a FeS cluster (R. D. Camerini-Otero, personal communication). As for the human helicases, the discovery of the FeS cluster opens completely new insights into the human pathology. Two human mutations found in TTD (hXPD-R112H) and Fanconi anemia (FancJ-A349P) patients were successfully introduced into SacXPD. These mutations destabilise the FeS cluster and subsequently abolish the helicase activity. The presence of the cluster was also confirmed *in vivo* in *S. cerevisiae* Rad3p, the orthologue of hXPD. The function of the FeS cluster has still to be investigated. A likely function may be to help stabilise a loop or hairpin structure, which supports the engagement of one of the single DNA strands during DNA duplex unwinding. This could actively distort the ss/ ds fork, facilitating base pair disruption during translocation. A possible model is shown in Fig 7.1.

The ATPase activity of SacXPD was stimulated by ssDNA, however, no stimulation was observed with dsDNA. All of the mutants, apart from those with mutations in one of the conserved helicase motifs (K35A, F38C, G447D, C523R), showed ssDNA dependent ATPase activity similar or slightly higher than the wildtype. Mutations that affected the coordination or stability of the FeS cluster showed an uncoupling of helicase and ATPase activity. The K_M for ATP hydrolysis was measured to be 0.35 to 0.36 mM ATP for the wildtype enzyme and representative mutants, while maximum rates of ATP hydrolysis were slightly higher for mutants compared to the wildtype.

Translocation could not be fully proven, but it is assumed that the FeS cluster is not necessary for ssDNA tracking. One possible way to test translocation would be to use two oligonucleotides of identical sequence and different lengths and compare their K_M by measuring the ATPase activity using different concentrations of oligonucleotides. The

same oligonucleotides containing a central blocking lesion would then be used in subsequent experiments. The presence of the blocking lesion should reduce the ATPase activity and increase the K_M (Raney & Benkovic (1995)).

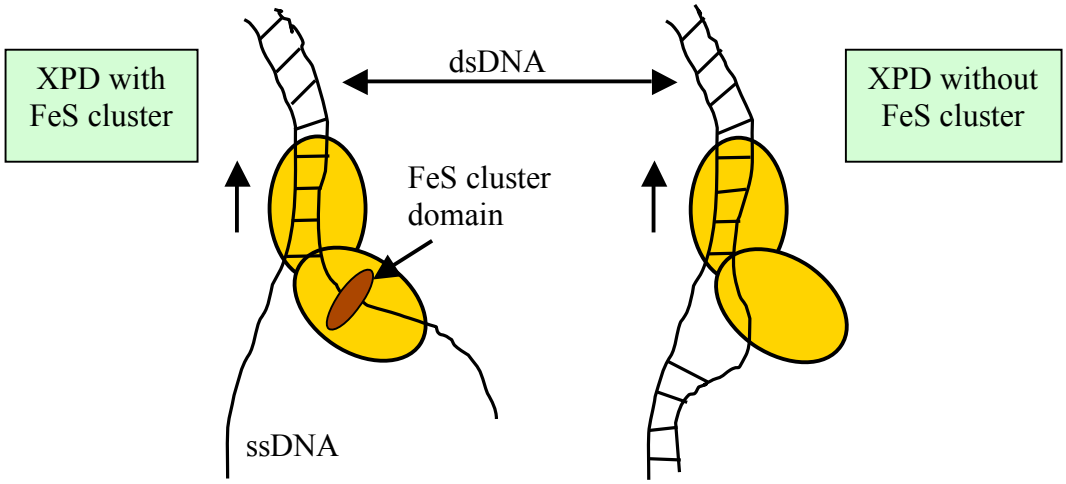


Fig 7.1: Hypothetical model of XPD duplex DNA unwinding and the function of the FeS cluster

The figure shows a schematic of XPD duplex DNA unwinding. Single and double stranded DNA binding domains are shown in yellow. The FeS cluster supports a loop or hairpin structure (in brown, left), which clamps the single DNA strand of the loading strand and kinks it around 90 ˚. In this way, the ss/ ds fork ahead is actively distorted, facilitating duplex unwinding. Furthermore, this prevents reannealing of the two single DNA strands. In mutants that do not coordinate a FeS cluster (right), the DNA is still unwound with very low rates, probably similar to XPB. Since the loading strand is not bend away from the complementary strand, this leads to reannealing and formation of the duplex DNA.

XPD was shown not to stall on a fluorescein lesion in either the loading or displaced strand, although slightly lower DNA unwinding rates were detected when the lesion was on the loading strand. Further experiments using different DNA substrates will need to be carried out to investigate if XPD interacts directly with the bases of the loading strand or rather with the sugar-phosphate backbone, as proposed for SF2 helicases (Singleton & Wigley (2002)). These DNA substrates can also be used to probe if XPD stalls at a lesion in one of the DNA strands. If XPD recognises a lesion while

unwinding DNA and stalls, this could imply a mechanism for the regulation of XPD and positioning other proteins (e.g. XPG/ FEN1) during DNA repair.

Interactions of XPD with other proteins during DNA repair are likely to occur, however, no interactions could be detected for XPD. This could be caused by weak binding affinities of XPD to other proteins (e.g. in the case of transient interactions) or that interactions were not detected under the conditions or methods used. If XPD takes part in NER, interactions with XPB and XPG/ FEN1 would be anticipated. Further experiments investigating protein-protein interactions of XPD are likely to involve different experimental conditions and/ or XPD constructs (e.g. GST-, His-tag). In another approach, fractionating *S. acidocaldarius* cell lysate by gel filtration chromatography and detection of fractions containing XPD via antibodies could lead to the identification of interacting proteins in a possible complex (Richard et al. (2004)). It is also possible that binding to DNA induces conformational changes in XPD that lead to exposure of a new protein-binding surface. Pull-down experiments in the presence of an appropriate DNA substrate could also reveal XPD binding partners.

XPD binding to DNA appeared weak, which is not unusual for DNA repair enzymes (Reardon & Sancar (2005)). It is possible that XPD does not interact strongly with DNA as this could hinder translocation. It seemed, that bubble DNA substrates were bound preferentially compared to a 5'-overhang DNA substrate. This could imply that SacXPD specifically recognises structures similar to an emerging repair bubble indicating its involvement in DNA repair. But these data are only preliminary and experiments have to be continued, preferentially using fluorescence based in-solution assays. The same applies for DNA binding studies using single and double stranded DNA in the presence of ATP-γS and ADP, respectively. According to the binding model for monomeric helicases, binding of ATP should decrease the affinity for ssDNA and increase that for dsDNA, while ADP is supposed to increase ssDNA and decrease dsDNA binding affinities (Velankar et al. (1999)).

This work showed how an archaeal enzyme can help to unravel completely new features of eukaryotic proteins. The discovery of the FeS cluster in *S. acidocaldarius* XPD and the presence of the FeS cluster domain in several human helicases contribute to our understanding of the development of different human diseases, including cancer.

REFERENCES

Agarwal A., Li D. and Cowan J.A. (1995) Role of aromatic residues in stabilization of the [Fe4S4] cluster in high-potential iron proteins (HiPIPs): physical characterization and stability studies of Tyr-19 mutants of Chromatium vinosum HiPIP. *Proc Natl Acad Sci U S A* **92 (**21): 9440-4.

Amann J., Kidd V.J. and Lahti J.M. (1997) Characterization of putative human homologues of the yeast chromosome transmission fidelity gene, CHL1. *J Biol Chem* **272 (**6): 3823-32.

Aravind L., Walker D.R. and Koonin E.V. (1999) Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res* **27 (**5): 1223-42.

Auerbach A.D., Buchwald M. and Joenje H. (2001) Fanconi Anemia. Scriver C.R., Beaudet A.L., Valle D., Sly W.S., Childs B., Kinzler K. and Vogelstein B. The metabolic and molecular bases of inherited diseases. *McGrew-Hill* **1 (**8): 753-68

Ayora S., Weise F., Mesa P., Stasiak A. and Alonso J.C. (2002) Bacillus subtilis bacteriophage SPP1 hexameric DNA helicase, G40P, interacts with forked DNA. *Nucleic Acids Res* **30 (**11): 2280-9.

Bardwell L., Cooper A.J. and Friedberg E.C. (1992) Stable and specific association between the yeast recombination and DNA repair proteins RAD1 and RAD10 in vitro. *Mol Cell Biol* **12 (**7): 3041-9.

Barns S.M., Delwiche C.F., Palmer J.D. and Pace N.R. (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci U S A* **93 (**17): 9188-93.

Bastin-Shanower S.A., Fricke W.M., Mullen J.R. and Brill S.J. (2003) The mechanism of Mus81-Mms4 cleavage site selection distinguishes it from the homologous endonuclease Rad1-Rad10. *Mol Cell Biol* **23 (**10): 2487-96.

Beinert H., Holm R.H. and Muenck E. (1997) Iron-sulfur clusters: nature's modular, multipurpose structures. *Science* **277 (**5326): 653-9.

Beinert H., Kennedy M.C. and Stout C.D. (1996) Aconitase as Iron-sulfur Protein, Enzyme, and Iron-Regulatory Protein. *Chem Rev* **96 (**7): 2335-74.

Beinert H. and Kiley P.J. (1999) Fe-S proteins in sensing and regulatory functions. *Curr Opin Chem Biol* **3 (**2): 152-7.

Benhamou S. and Sarasin A. (2005) ERCC2 /XPD gene polymorphisms and lung cancer: a HuGE review. *Am J Epidemiol* **161 (**1): 1-14.

Berneburg M., Clingen P.H., Harcourt S.A., Lowe J.E., Taylor E.M., Green M.H., Krutmann J., Arlett C.F. and Lehmann A.R. (2000a) The cancer-free phenotype in trichothiodystrophy is unrelated to its repair defect. *Cancer Res* **60 (**2): 431-8.

Berneburg M., Lowe J.E., Nardo T., Araujo S.J., Fousteri M.I., Green M.H., Krutmann J., Wood R.D., Stefanini M. and Lehmann A.R. (2000b) UV damage causes uncontrolled DNA breakage in cells from patients with combined features of XP-D and Cockayne syndrome. *19* **5 (**1157-66):

Bianco P.R., Brewer L.R., Corzett M., Balhorn R., Yeh Y., Kowalczykowski S.C. and Baskin R.J. (2001) Processive translocation and DNA unwinding by individual RecBCD enzyme molecules. *Nature* **409 (**6818): 374-8.

Bingemann R. and Klein A. (2000) Conversion of the central [4Fe–4S] cluster into a [3Fe–4S] cluster leads to reduced hydrogen-uptake activity of the F420-reducing hydrogenase of Methanococcus voltae. *Eur J Biochem* **267** 6612-8.

Bird L.E., Subramanya H.S. and Wigley D.B. (1998) Helicases: a unifying structural theme? *Curr Op Struct Biol* **8 (**1): 14-8.

Boal A.K., Yavin E., Lukianove O.A., O'shea V.L., David S.S. and Barton J.K. (2005) DNA-bound redox activity of DNA repair glycosylases containing [4Fe-4S] clusters. *Biochem* **44 (**23): 8397-407.

Boddy M.N., Gaillard P.H., Mcdonald W.H., Shanahan P., Yates J.R.R. and Russell P. (2001) Mus81-Eme1 are essential components of a Holliday junction resolvase. *Cell* **107 (**4): 537-48.

Borowski P., Kuehl R., Mueller O., Hwang L.H., Schulze Zur Wiesch J. and Schmitz H. (1999) Biochemical properties of a minimal functional domain with ATP-binding activity of the NTPase/helicase of hepatitis C virus. *Eur J Biochem* **266 (**3): 715-23.

Botta E., Nardo T., Broughton B.C., Marinoni S., Lehmann A.R. and Stefanini M. (1998) Analysis of mutations in the XPD gene in Italian patients with trichothiodystrophy: site of mutation correlates with repair deficiency, but gene dosage appears to determine clinical severity. *Am J Hum Genet* **63 (**4): 1036-48.

Brochier C., Gribaldo S., Zivanovic Y., Confalonieri F. and Forterre P. (2005) Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? *Genome Biol.* **6 (**5): R42.

Broughton B.C., Thompson A.F., Harcourt S.A., Vermeulen W., Hoeijmakers J.H., Botta E., Stefanini M., King M.D., Weber C.A., Cole J., Arlett C.F. and Lehmann A.R. (1995) Molecular and cellular analysis of the DNA repair defect in a patient in xeroderma pigmentosum complementation group D who has the clinical features of xeroderma pigmentosum and Cockayne syndrome. *Am J Hum Genet* **56 (**1): 167-74.

Busso D., Keriel A., Sandrock B., Poterszman A., Gileadi O. and Egly J.M. (2000) Distinct regions of MAT1 regulate cdk7 kinase and TFIIH transcription activities. *J Biol Chem* **275 (**30): 22815-23.

Byrd A.K. and Raney K.D. (2004) Protein displacement by an assembly of helicase molecules aligned along single-stranded DNA. *Nat Struct Mol Biol* **1 (**6): 531-8.

Cantor S.B., Bell D.W., Ganesan S., Kass E.M., Drapkin R., Grossman S., Wahrer D.C., Sgroi D.C., Lane W.S., Haber D.A. and Livingston D.M. (2001) BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* **105 (**1): 149-60.

Cantor S.B., Drapkin R., Zhang F., Lin Y., Han J., Pamidi S. and Livingston D.M. (2004) The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. *Proc Natl Acad Sci U S A* **101 (**8): 2357-62.

Capozzi F., Ciurli S. and Luchinat C. (1998) Coordination sphere versus protein environment as determinants of electronic and functional properties of iron-sulfur proteins. *Struct Bond* **90** 127-60.

Chao K.L. and Lohman T.M. (1991) DNA-induced dimerization of the Escherichia coli Rep helicase. *J Mol Biol* **221 (**4): 1165-81.

Chen J., Larochelle S., Li X. and Suter B. (2003) Xpd/Ercc2 regulates CAK activity and mitotic progression. *Nature* **424 (**6945): 228-32.

Chen J. and Suter B. (2003) Xpd, a structural bridge and a functional link. *Cell Cycle* **2** (6): 503-6.

Chen L., Brugger K., Skovgaard M., Redder P., She Q., Torarinsson E., Greve B., M. A., Zibat A., Klenk H.P. and Garrett R.A. (2005) The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota. *J Bacteriol* **187** (14): 4992-9.

Chepanoske C.L., Golinelli M.P., Williams S.D. and David S.S. (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): 11-9.

Cleaver J.E. (2005) Splitting hairs - discovery of a new DNA repair and transcription factor for the human disease trichothiodystrophy. *DNA Repair* **4** (2): 285-7.

Coin F., Marinoni J.C., Rudolfo C., Fribourg S., Pedrini A.M. and Egly J.M. (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): 184-8.

Coldren C.D., Hellinga H.W. and Caradonna J.P. (1997) The rational design and construction of a cuboidal iron-sulfur protein. *Proc Natl Acad Sci U S A* **94** (13): 6635-40.

Cordin O., Tanner N.K., Doere M., Linder P. and Banroques J. (2004) The newly discovered Q motif of DEAD-box RNA helicases regulates RNA-binding and helicase activity. *EMBO J* **23** (13): 2478-87.

Cubeddu L. and White M.F. (2005) DNA damage detection by an archaeal single-stranded DNA-binding protein. *J Mol Biol* **353** (3): 507-16.

Davies A.A., Friedberg E.C., Tomkinson A.E., Wood R.D. and West S.C. (1995) Role of the Rad1 and Rad10 proteins in nucleotide excision repair and recombination. *J Biol Chem* **270 (**42): 24638-41.

De Boer J., Andressoo J.O., De Wit J., Huijmans J., Beems R.B., Van Steeg H., Weeda G., Van Der Horst G.T., Van Leeuwen W., Themmen A.P., Meradji M. and Hoeijmakers J.H. (2002) Premature aging in mice deficient in DNA repair and transcription. *Science* **296 (**5571): 1250-1.

De Laat W.L., Appeldoorn E., Jaspers N.G. and Hoeijmakers J.H. (1998a) DNA structural elements required for ERCC1-XPF endonuclease activity. *J Biol Chem* **273 (**14): 7835-42.

Debniak T., Scott R.J., Huzarski T., Byrski T., Masojc B., Van De Wetering T. and Al. E. (2006) XPD Common Variants and their Association with Melanoma and Breast Cancer Risk. *Breast Cancer Res Treat* **98 (**2): 209-15.

Dillingham M.S., Soultanas P. and Wigley D.B. (1999) Site-directed mutagenesis of motif III in PcrA helicase reveals a role in coupling ATP hydrolysis to strand separation. *Nucleic Acids Res* **27 (**16): 3310-7.

Dillingham M.S., Soultanas P. and Wigley D.B. (2000) Demonstration of unidirectional single-stranded DNA translocation by PcrA helicase: measurement of step size and translocation speed. *Biochem* **39 (**1): 205-12.

Ding H., Schertzer M., Wu X., Gertsenstein M., Selig S., Kammori M., Pourvali R., Poon S., Vulto I., Chavez E., Tam P.P., Nagy A. and Lansdorp P.M. (2004) Regulation of murine telomere length by Rtel: an essential gene encoding a helicase-like protein. *Cell* **117 (**7): 873-86.

Dionne I., Nookala R.K., Jackson S.P., Doherty A.J. and Bell S.D. (2003) A heterotrimeric PCNA in the hyperthermophilic archaeon Sulfolobus solfataricus. *Mol Cell* **11 (**1): 275-82.

Dip R., Camenisch U. and Naegeli H. (2004) Mechanisms of DNA damage recognition and strand discrimination in human nucleotide excision repair. *DNA Repair* **3 (**11): 1409-23.

Doherty A.J., Serpell L.C. and Ponting C.P. (1996) The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res* **24 (**13): 2488-97.

Dubaele S., Proietti De Santis L., Bienstock R.J., Keriel A., Stefanini M., Van Houten B. and Egly J.M. (2003) Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Mol Cell* **11 (**6): 1635-46.

Duff J.L.C., Breton J.L.J., Butt J.N., Armstrong F.A. and Thomson A.J. (1996) Novel Redox Chemistry of [3Fe-4S] Clusters: Electrochemical Characterization of the All-Fe(II) Form of the [3Fe-4S] Cluster Generated Reversibly in Various Proteins and Its Spectroscopic Investigation in Sulfolobus acidocaldarius Ferredoxin. *J Am Chem Soc* **118 (**36): 8593-603.

Eady R.R., Richardson T.H., Miller R.W., Hawkins M. and Lowe D.J. (1988) The vanadium nitrogenase of Azotobacter chroococcum. Purification and properties of the Fe protein. *Biochem J* **256** 189-96.

Enemark E.J. and Joshua-Tor L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442 (**7100): 270-5.

Enzlin J.H. and Scharer O.D. (2002) The active site of the DNA repair endonuclease XPF-ERCC1 forms a highly conserved nuclease motif. *EMBO J* **21 (**8): 2045-53.

Eoff R.L., Spurling T.L. and Raney K.D. (2005) Chemically modified DNA substrates implicate the importance of electrostatic interactions for DNA unwinding by Dda helicase. *Biochem* **44 (**2): 666-74.

Evans E., Moggs J.G., Hwang J.R., Egly J.M. and Wood R.D. (1997) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J* **16 (**21): 6559-73.

Feaver W.J., Huang W., Gileadi O., Myers L., Gustafsson C.M., Kornberg R.D. and Friedberg E.C. (2000) Subunit interactions in yeast transcription/repair factor TFIIH. Requirement for Tfb3 subunit in nucleotide excision repair. *J Biol Chem* **275 (**8): 5941-6.

Fischer C.J., Maluf N.K. and Lohman T.M. (2004) Mechanism of ATP-dependent translocation of E.coli UvrD monomers along single-stranded DNA. *J Mol Biol* **344 (**5): 1287-309.

Gaudu P. and Weiss B. (1996) SoxR, a [2Fe-2S] transcription factor, is active only in its oxidized form. *Proc Natl Acad Sci U S A* **93 (**19): 10094-8.

Gerring S.L., Spencer F. and Hieter P. (1990) The CHL 1 (CTF 1) gene product of Saccharomyces cerevisiae is important for chromosome transmission and normal cell cycle progression in G2/M. *EMBO J* **9 (**13): 4347-58.

Giglia-Mari G., Coin F., Ranish J.A., Hoogstraten D., Theil A., Wijgers N., Jaspers N.G., Raams A., Argentini M., Van Der Spek P.J., Botta E., Stefanini M., Egly J.M., Aebersold R., Hoeijmakers J.H. and Vermeulen W. (2004) A new, tenth subunit of TFIIH is responsible for the DNA repair syndrome trichothiodystrophy group A. *Nat Genet* **36 (**7): 714-9.

Gorbalenya A.E. and Koonin E.V. (1993) Helicases: amino acid sequence comparison and structure-function relationships. *Curr Op Struct Biol* **3** 419-29.

Gupta R., Sharma S., Sommers J.A., Jin Z., Cantor S.B. and Brosh R.M.J. (2005) Analysis of the DNA substrate specificity of the human BACH1 helicase associated with breast cancer. *J Biol Chem* **280 (**27): 25450-60.

Hall M.C. and Matson S.W. (1999) Helicase motifs: The engine that powers DNA unwinding. *Mol Microbiol* **34 (**5): 867-77.

Hentze M.W., Muckenthaler M.U. and Andrews N.C. (2004) Balancing acts: molecular control of mammalian iron metabolism. *Cell* **117 (**3): 285-97.

Hoeijmakers J.H. (2001) Genome maintenance mechanisms and preventing cancer. *Nature* **411 (**6835): 366-74.

Howlett N.G., Taniguchi T., Olson S., Cox B., Waisfisz Q., De Die-Smulders C., Persky N., Grompe M., Joenje H., Pals G., Ikeda H., Fox E.A. and D'andrea A.D. (2002) Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* **297 (**5581): 606-9.

Huang J.C., Svoboda D.L., Reardon J.T. and Sancar A. (1992) Human nucleotide excision nuclease removes thymine dimers from DNA by incising the 22nd phosphodiester bond 5' and the 6th phosphodiester bond 3' to the photodimer. *Proc Natl Acad Sci U S A* **89 (**8): 3664-8.

Hubbell W.L., Mchaourab H.S., Altenbach C. and Lietzow M.A. (1996) Watching proteins move using site-directed spin labeling. *Structure* **4 (**7): 779-83.

Huber H., Hohn M.J., Rachel R., Fuchs T., Wimmer V.C. and Stetter K.O. (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417 (**6884): 63-7.

Ilyina T.V., Gorbalenya A.E. and Koonin E.V. (1992) Organization and evolution of bacterial and bacteriophage primase-helicase systems. *J Mol Evo* **34 (**4): 351-7.

Imlay J.A. (2006) Iron-sulphur clusters and the problem with oxygen. *Mol Microbiol* **59 (**4): 1073-82.

Iverson T.M., Luna-Chavez C., Cecchini G. and Rees D.C. (1999) Structure of the Escherichia coli fumarate reductase respiratory complex. *Science* **284 (**5422): 1961-6.

Iwagami S.G., Creagh A.L., Haynes C.A., Borsari M., Felli I.C., Piccioli M. and Eltis L.D. (1995) The role of a conserved tyrosine residue in high-potential iron sulfur proteins. *Protein Sci* **4 (**12): 2562-72.

Iyer N., Reagan M.S., Wu K.J., Canagarajah B. and Friedberg E.C. (1996) Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. *Biochem* **35 (**7): 2157-67.

Jankowsky E., Gross C.H., Shuman S. and Pyle A.M. (2000) The DExH protein NPH-II is a processive and directional motor for unwinding RNA. *Nature* **403 (**6768): 447-51.

Jelinska C., Conroy M.J., Craven C.J., Hounslow A.M., Bullough P.A., Waltho J.P., Taylor G.L. and White M.F. (2005) Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. *Structure* **13 (**7): 963-71.

Kaliraman V., Mullen J.R., Fricke W.M., Bastin-Shanower S.A. and Brill S.J. (2001) Functional overlap between Sgs1-Top3 and the Mms4-Mus81 endonuclease. *Genes Dev* **15 (**20): 2730-40.

Kawaoka J., Jankowsky E. and Pyle A.M. (2004) Backbone tracking by the SF2 helicase NPH-II. *Nat Struct Mol Biol* **11 (**6): 526-30.

Kelly S.M., Jess T.J. and Price N.C. (2005) How to Study Proteins by Circular Dichroism. *Biochim Biophys Acta* **1751** 199-139.

Kiley P.J. and Beinert H. (2003) The role of Fe-S proteins in sensing and regulation in bacteria. *Curr Opin Microbiol* **6 (**2): 181-5.

Kispal G., Csere P., Prohl C. and Lill R. (1999) The mitochondrial proteins Atm1p and Nfs1p are essential for biogenesis of cytosolic Fe/S proteins. *EMBO J* **18 (**14): 3981-9.

Komori K., Fujikane R., Shinagawa H. and Ishino Y. (2002) Novel endonuclease in Archaea cleaving DNA with various branched structure. *Genes Genet Syst* **77 (**4): 227-41.

Koonin E.V. (1993) Escherichia coli dinG gene encodes a putative DNA helicase related to a group of eukaryotic helicases including Rad3 protein. *Nucleic Acids Res* **21 (**6): 1497.

Korolev S., Hsieh J., Gauss G.H., Lohman T.M. and Waksman G. (1997) Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell* **90 (**4): 635-47.

Kowal A.T., Werth M.T., Manodori A., Cecchini G., Schroder I., Gunsalus R.P. and Johnson M.K. (1995) Effect of cysteine to serine mutations on the properties of the [4Fe-4S] center in Escherichia coli fumarate reductase. *Biochem* **34 (**38): 12284-93.

Krishna T.S., Kong X.P., Gary S., Burgers P.M. and Kuriyan J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79 (**7): 1233-43.

Kuraoka I., Kobertz W.R., Ariza R.R., Biggerstaff M., Essigmann J.M. and Wood R.D. (2000) Repair of an interstrand DNA cross-link initiated by ERCC1-XPF repair/recombination nuclease. *J Biol Chem* **275 (**34): 26632-6.

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

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

Levin M.K. and Patel S.S. (2002) Helicase from hepatitis C virus, energetics of DNA binding. *J Biol Chem* **277 (**33): 29377-85.

Levin M.K. and Patel S.S. (2003) Helicases as molecular motors. Schliwa M. Molecular Motors. *Wiley-VCH Verlag-GmbH*

Levitus M., Waisfisz Q., Godthelp B.C., De Vries Y., Hussain S., Wiegant W.W., Elghalbzouri-Maghrani E., Steltenpool J., Rooimans M.A., Pals G., Arwert F., Mathew C.G., Zdzienicka M.Z., Hiom K., De Winter J.P. and Joenje H. (2005) The DNA helicase BRIP1 is defective in Fanconi anemia complementation group J. *Nat Genet* **37 (**9): 934-5.

Levran O., Attwooll C., Henry R.T., Milton K.L., Neveling K., Rio P., Batish S.D., Kalb R., Velleuer E., Barral S., Ott J., Petrini J., Schindler D., Hanenberg H. and Auerbach A.D. (2005) The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. *Nat Genet* **37 (**9): 931-3.

Lewis L.K., Jenkins M.E. and Mount D.W. (1992) Isolation of DNA damage-inducible promoters in Escherichia coli: regulation of polB (dinA), dinG, and dinH by LexA repressor. *J Bacteriol* **174 (**10): 3377-85.

Lindahl T. (1993) Instability and decay of the primary structure of DNA. *Nature* **362 (**6422): 709-15.

Litman R., Peng M., Jin Z., Zhang F., Zhang J., Powell S., Andreassen P.R. and Cantor S.B. (2005) BACH1 is critical for homologous recombination and appears to be the Fanconi anemia gene product FANCJ. *Cancer Cell* **8 (**3): 255-65.

Lobley A., Whitmore L. and Wallace B.A. (2002) DICHROWEB: An Interactive Website for the Analysis of Protein Secondary Structure from Circular Dichroism Data. *Bioinformatics* **18** 211-212.

Lohman T.M. (1993) Helicase-catalyzed DNA unwinding. *J Biol Chem* **268 (**4): 2269-72.

Lorsch J.R. and Herschlag D. (1998) The DEAD box protein eIF4A. 2. A cycle of nucleotide and RNA-dependent conformational changes. *Biochem* **37 (**8): 2137-206.

Lu A.L. and Wright P.M. (2003) Characterization of an Escherichia coli mutant MutY with a cysteine to alanine mutation at the iron-sulfur cluster domain. *Biochem* **42 (**13): 3742-50.

Mcdougal V.V. and Guarino L.A. (2001) DNA and ATP binding activities of the baculovirus DNA helicase P143. *J Virol* **15** 7206-9.

Meetei A.R., Medhurst A.L., Ling C., Xue Y., Singh T.R., Bier P., Steltenpool J., Stone S., Dokal I., Mathew C.G., Hoatlin M., Joenje H., De Winter J.P. and Wang W. (2005) A human ortholog of archaeal DNA repair protein Hef is defective in Fanconi anemia complementation group M. *Nat Genet* **37 (**9): 958-63.

Moolenaar G.F., Van Rossum-Fikkert S., Van Kersten M. and Goosen N. (2002) Cho, a second endonuclease involved in Escherichia coli nucleotide excision repair. *Proc Natl Acad Sci U S A* **99 (**3): 1467-72.

Mu D., Wakasugi M., Hsu D.S. and Sancar A. (1997) Characterization of reaction intermediates of human excision repair nuclease. *J Biol Chem* **272 (**46): 28971-9.

Muhlenhoff U. and Lill R. (2000) Biogenesis of iron-sulfur proteins in eukaryotes: a novel task of mitochondria that is inherited from bacteria. *Biochim Biophys Acta* **1459 (**2-3): 370-82.

Mulliez E., Ollagnier-De Choudens S., Meier C., Cremonini M., Luchinat C., Trautwein A.X. and Fontecave M. (1999) Iron-sulfur interconversions in the anaerobic ribonucleotide reductase from Escherichia coli. *J Biol Inorg Chem* **4 (**5): 614-20.

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

Newman M., Murray-Rust J., Lally J., Rudolf J., Fadden A., Knowles P.P., White M.F. and Mcdonald N.Q. (2005) Structure of an XPF endonuclease with and without DNA suggests a model for substrate recognition. *EMBO J* **24 (**5): 895-905.

Nishino T., Komori K., Ishino Y. and Morikawa K. (2003) X-ray and biochemical anatomy of an archaeal XPF/Rad1/Mus81 family nuclease: similarity between its endonuclease domain and restriction enzymes. *Structure* **11 (**4): 445-57.

Orphanides G., Lagrange T. and Reinberg D. (1996) The general transcription factors of RNA polymerase II. *Genes Dev* **10 (**21): 2657-83.

Orren D.K., Selby C.P., Hearst J.E. and Sancar A. (1992) Post-incision steps of nucleotide excision repair in Escherichia coli. Disassembly of the UvrBC-DNA complex by helicase II and DNA polymerase I. *J Biol Chem* **267 (**2): 780-8.

Page C.C., Moser C.C. and Dutton P.L. (2003) Mechanism for electron transfer within and between proteins. *Curr Op Chem Biol* **7 (**5): 551-6.

Park C.H., Bessho T., Matsunaga T. and Sancar A. (1995) Purification and characterization of the XPF-ERCC1 complex of human DNA repair excision nuclease. *J Biol Chem* **270 (**39): 22657-60.

Pause A. and Sonenberg N. (1992) Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *EMBO J* **11 (**7): 2643-54.

Pieroni L., Khalil L., Charlotte F., Poynard T., Piton A., Hainque B. and Imbert-Bismut F. (2001) Comparison of bathophenanthroline sulfonate and ferene as chromogens in colorimetric measurement of low hepatic iron concentration. *Clin Chem* **47 (**11): 2059-61.

Porter D.J., Short S.A., Hanlon M.H., Preugschat F., Wilson J.E., Willard D.H.J. and Consler T.G. (1998) Product release is the major contributor to kcat for the hepatitis C virus helicase-catalyzed strand separation of short duplex DNA. *J Biol Chem* **273 (**30): 18906-14.

Raney K.D. and Benkovic S.J. (1995) Bacteriophage T4 Dda helicase translocates in a unidirectional fashion on single-stranded DNA. *J Biol Chem* **270 (**38): 22236-42.

Ranish J.A., Hahn S., Lu Y., Yi E.C., Li X.J., Eng J. and Aebersold R. (2004) Identification of TFB5, a new component of general transcription and DNA repair factor IIH. *Nat Genet* **36 (**7): 707-13.

Reardon J.C., Ge H., Gibbs E., Sancar A., Hurwitz J. and Pan Z.Q. (1996) Isolation and characterization of two human transcription factor IIH (TFIIH)-related complexes: ERCC2/CAK and TFIIH. *Proc Natl Acad Sci U S A* **93 (**13): 6482-7.

Reardon J.T. and Sancar A. (2003) Recognition and repair of the cyclobutane thymine dimer, a major cause of skin cancers, by the human excision nuclease. *Genes Dev* **17 (**20): 2539-51.

Reardon J.T. and Sancar A. (2005) Nucleotide Excision Repair. *Prog Nucleic Acid Res Mol Biol* **79** 183-235.

Reardon J.T., Thompson L.H. and Sancar A. (1997) Rodent UV-sensitive mutant cell lines in complementation groups 6-10 have normal general excision repair activity. *Nucleic Acids Res* **25 (**5): 1015-21.

Rees D.C. (2002) Great metalloclusters in enzymology. *Annu Rev Biochem* **71** 221-46.

Rees D.C. and Howard J.B. (2000) Nitrogenase: standing at the crossroads. *Curr Op Chem Biol* **4 (**5): 559-66.

Richard D.J., Bell S.D. and White M.F. (2004) Physical and functional interaction of the archaeal single-stranded DNA-binding protein SSB with RNA polymerase. *Nucleic Acids Res* **32 (**3): 1065-74.

Roberts J.A. (2004) Characterisation of Sulfolobus solfataricus XPF. An archeal DNA repair endonuclease. CBMS (University of St Andrews) St Andrews.

Roberts J.A., Bell S.D. and White M.F. (2003) An archaeal XPF repair endonuclease dependent on a heterotrimeric PCNA. *Mol. Microbiol.* **48 (**2): 361-371.

Roberts J.A. and White M.F. (2005a) An archaeal endonuclease displays key properties of both eukaryal XPF-ERCC1 and Mus81. *J Biol Chem* **280 (**7): 5924-8.

Rouault T.A. and Klausner R.D. (1996) Iron-sulfur clusters as biosensors of oxidants and iron. *Trends Biochem Sci.* **21 (**5): 174-7.

Rouault T.A. and Tong W.H. (2005) Iron-sulphur cluster biogenesis and mitochondrial iron homeostasis. *Nat Rev Mol Cell Biol* **6 (**4): 345-51.

Rudolf J., Makrantoni V., Ingledew W.J., Stark M.J. and White M.F. (2006) The DNA Repair Helicases XPD and FancJ Have Essential Iron-Sulfur Domains. *EMBO J* **23 (**6): 801-8.

Sancar A. (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem Rev* **103 (**6): 2203-37.

Sancar A. and Rupp W.D. (1983) A novel repair enzyme: UVRABC excision nuclease of Escherichia coli cuts a DNA strand on both sides of the damaged region. *Cell* **33 (**1): 249-60.

Sandrock B. and Egly J.M. (2001) A yeast four-hybrid system identifies Cdk-activating kinase as a regulator of the XPD helicase, a subunit of transcription factor IIH. *J Biol Chem* **276 (**38): 35328-33.

Schabath M.B., Delclos G.L., Grossman H.B., Wang Y., Lerner S.P., Chamberlain R.M., Spitz M.R. and Wu X. (2005) Polymorphisms in XPD exons 10 and 23 and bladder cancer risk. *Cancer Epidemiol Biomarkers Prev* **14 (**4): 878-84.

Schaeffer L., Moncollin V., Roy R., Staub A., Mezzina M., Sarasin A., Weeda G., Hoeijmakers J.H. and Egly J.M. (1994) The ERCC2/DNA repair protein is associated with the class II BTF2/TFIIH transcription factor. *EMBO J* **13 (**10): 2388-92.

Schaeffer L., Roy R., Humbert S., Moncollin V., Vermeulen W., Hoeijmakers J.H., Chambon P. and Egly J.M. (1993) DNA repair helicase: a component of BTF2 (TFIIH) basic transcription factor. *Science* **260 (**5104): 58-63.

Scharer O.D. (2003) Chemistry and biology of DNA repair. *Angew Chem Int Ed Engl* **42 (**26): 2946-74.

Seeberg E. and Steinum A.L. (1982) Purification and properties of the uvrA protein from Escherichia coli. *Proc Natl Acad Sci U S A* **79 (**4): 988-92.

Seefeld L.C., Dance I.G. and Dean D.R. (2004) Substrate interactions with nitrogenase: Fe versus Mo. *Biochem* **43 (**6): 1401-9.

Seeley T.W. and Grossman L. (1990) The role of Escherichia coli UvrB in nucleotide excision repair. *J Biol Chem* **265 (**13): 7158-65.

Selby C.P. and Sancar A. (1993) Molecular mechanism of transcription-repair coupling. *Science* **260 (**5104): 53-8.

Sgouros J., Gaillard P.H. and Wood R.D. (1999) A relationship betweena DNA-repair/recombination nuclease family and archaeal helicases. *Trends Biochem Sci.* **24 (**3): 95-7.

Shao X. and Grishin N.V. (2000) Common fold in helix-hairpin-helix proteins. *Nucleic Acids Res* **28 (**14): 2643-50.

She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G., Awayez M.J., Chan-Weiher C.C., Clausen I.G., Curtis B.A., De Moors A., Erauso G., Fletcher C., Gordon P.M., Heikamp-De Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X., Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N., Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T., Garrett R.A., Ragan M.A., Sensen C.W. and Van Der Oost J. (2001) The complete genome of the crenarchaeon Sulfolobus solfataricus P2. *Proc Natl Acad Sci U S A* **98 (**14): 7835-40.

Shi Q., Thresher R., Sancar A. and Griffith J. (1992) Electron microscopic study of (A)BC excinuclease. DNA is sharply bent in the UvrB-DNA complex. *226* **2 (**425-32):

Sijbers A.M., De Laat W.L., Ariza R.R., Biggerstaff M., Wei Y.F., Moggs J.G., Carter K.C., Shell B.K., De Jong M.C., Rademakers S., De Rooij J., Jaspers N.G., Hoeijmakers J.H. and Wood R.D. (1996) Xeroderma pigmentosum group F caused by a defect in a structure-specific DNA repair endonuclease. *Cell* **86 (**5): 811-22.

Singleton M.R. and Wigley D.B. (2002) Modularity and specialization in superfamily 1 and 2 helicases. *J Bacteriol* **184 (**7): 1819-26.

Skibbens R.V. (2004) Chl1p, a DNA helicase-like protein in budding yeast, functions in sister-chromatid cohesion. *Genetics* **166 (**1): 33-42.

Soultanas P. and Wigley D.B. (2000) DNA helicases: 'inching forward'. *Curr Opin Struct Biol* **10 (**1): 124-8.

Soultanas P. and Wigley D.B. (2001) Unwinding the 'Gordian knot' of helicase action. *Trends Biochem Sci.* **26 (**1): 47-54.

Sticht H. and Roesch P. (1998) The structure of iron-sulfur proteins. *Prog Biophys Mol Biol* **70 (**2): 95-136.

Story R.M., Weber I.T. and Steitz T.A. (1992) The structure of the E. coli recA protein monomer and polymer. *Nature* **355 (**6360): 318-25.

Subramanya H.S., Bird L.E., Brannigan J.A. and Wigley D.B. (1996) Crystal structure of a DExx box DNA helicase. *Nature* **384 (**6607): 379-83.

Sung P., Higgins D., Prakash L. and Prakash S. (1988) Mutation of lysine-48 to arginine in the yeast RAD3 protein abolishes its ATPase and DNA helicase activities but not the ability to bind ATP. *EMBO J* **7 (**10): 3263-9.

Taniguchi T. and D'andrea A.D. (2006) Molecular pathogenesis of Fanconi anemia: recent progress. *Blood* **107 (**11): 4223-33.

Thayer M.M., Ahern H., Xing D., Cunningham R.P. and Tainer J.A. (1995) Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J* **14 (**16): 4108-20.

Theis K., Skorvaga M., Machius M., Nakagawa N., Van Houten B. and Kisker C. (2000) The nucleotide excision repair protein UvrB, a helicase-like enzyme with a catch. *Mutat Res* **460 (**3-4): 277-300.

Tilley G.J., Camba R., Burgess B.K. and Armstrong F.A. (2001) Influence of electrochemical properties in determining the sensitivity of [4Fe-4S] clusters in proteins to oxidative damage. *Biochem* **360 (**3): 717-26.

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

Tuteja N. and Tuteja R. (2004) Prokaryotic and eukaryotic DNA helicases. *Eur J Biochem* **271** 1835-48.

Uchida A., Sugasawa K., Masutani C., Dohmae N., Araki M., Yokoi M., Ohkuma Y. and Hanaoka F. (2001) The carboxy-terminal domain of the XPC protein plays a crucial role in nucleotide excision repair through interactions with transcription factor IIH. *DNA Repair* **1 (**6): 449-61.

Van Dyk T.K., Derose E.J. and Gonye G.E. (2001) LuxArray, a high-density, genomewide transcription analysis of Escherichia coli using bioluminescent reporter strains. *J Bacteriol* **183 (**19): 5496-505.

Van Hoffen A., Balajee A.S., Van Zeeland A.A. and Mullenders L.H. (2003) Nucleotide excision repair and its interplay with transcription. *Toxicology* **193 (**1-2): 79-90.

Van Hoffen A., Natarajan A.T., Mayne L.V., Van Zeeland A.A., Mullenders L.H. and Venema J. (1993) Deficient repair of the transcribed strand of active genes in Cockayne's syndrome cells. *Nucleic Acids Res* **21 (**25): 5890-5.

Van Houten B., Croteau D.L., Della Vecchia M.D., Wang H. and Kisker C. (2005) 'Close-fitting sleeves': DNA damage recognition by the UvrABC nuclease system. *Mutat Res* **577 (**1-2): 92-117.

Van Houten B., Eisen J.A. and Hanawalt P.C. (2002) A cut above: discovery of an alternative excision repair pathway in bacteria. *Proc Natl Acad Sci U S A* **99 (**5): 2581-3.

Velankar S.S., Soultanas P., Dillingham M.S., Subramanya H.S. and Wigley D.B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97 (**1): 75-84.

Verhoeven E.E., Wyman C., Moolenaar G.F. and Goosen N. (2002b) The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands. *EMBO J* **21 (**15): 4196-205.

Voloshin O.N., Vanevski F., Khil P.P. and Camerini-Otero R.D. (2003) Characterization of the DNA damage-inducible helicase DinG from Escherichia coli. *J Biol Chem* **278 (**30): 28284-93.

Wakasugi M., Shimizu M., Morioka H., Linn S., Nikaido O. and Matsunaga T. (2001) Damaged DNA-binding protein DDB stimulates the excision of cyclobutane pyrimidine dimers in vitro in concert with XPA and replication protein A. *J Biol Chem* **276 (**18): 15434-40.

Walker J.E., Saraste M., Runswick M.J. and Gay N.J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1 (**8): 945-51.

Warbrick E. (2000) The puzzle of PCNA's many partners. *Bioessays* **22 (**11): 997-1006.

Warrener P. and Collett M. (1995) Pestivirus NS3 (p80) protein possesses RNA helicase activity. *J Virol* **69 (**3): 1720-6.

Waters E., Hohn M.J., Ahel I., Graham D.E., Adams M.D., Barnstead M., Beeson K.Y., Bibbs L., Bolanos R., Keller M., Kretz K., Lin X., Mathur E., Ni J., Podar M., Richardson T., Sutton G.G., Simon M., Soll D., Stetter K.O., Short J.M. and Noordewier M. (2003) The genome of Nanoarchaeum equitans: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci U S A* **100 (**22): 12984-8.

Weber C.A., Kirchner J.M., Salazar E.P. and Takayama K. (1994) Molecular analysis of CXPD mutations in the repair-deficient hamster mutants UV5 and UVL-13. *Mutat Res* **324 (**4): 147-52.

Weber C.A., Salazar E.P., Stewart S.A. and Thompson L.H. (1988) Molecular cloning and biological characterization of a human gene, ERCC2, that corrects the nucleotide excision repair defect in CHO UV5 cells. *Mol Cell Biol* **8 (**3): 1137-46.

White M.F. (2003) Archaeal DNA repair: paradigms and puzzles. *Biochem Soc Trans* **31** 690-3.

Winkler G.S., Araujo S.J., Fiedler U., Vermeulen W., Coin F., Hoeijmakers J.H., Egly J.M., Wood R.D., Timmers H.T. and Weeda G. (2000) TFIIH with inactive XPD helicase functions in transcription initiation but is defective in DNA repair. *J Biol Chem* **275 (**6): 4258-66.

Woese C.R. and Fox G.E. (1977) Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc Natl Acad Sci U S A* **(**74):

Wollenberg M., Berndt C., Bill E., Schwenn J.D. and Seidler A. (2003) A dimer of the FeS cluster biosynthesis protein IscA from cyanobacteria binds a [2Fe2S] cluster between two protomers and transfers it to [2Fe2S] and [4Fe4S] apo proteins. *Eur J Biochem* **270 (**8): 1662-71.

Yarranton G.T. and Genfter M.L. (1979) Enzyme-catalyzed DNA unwinding: studies on Escherichia coli rep protein. *Proc Natl Acad Sci U S A* **76 (**4): 1658-62.

Yasuda T., Morimatsu K., Nagata T. and Ohmori H. (1998) Inhibition of Escherichia coli RecA coprotease activities by DinI. *EMBO J* **17 (**11): 3207-16.

Yasui A., Eker A.P., Yasuhira S., Yajima H., Kobayashi T., Takao M. and Oikawa A. (1994) A new class of DNA photolyases present in various organisms including aplacental mammals. *EMBO J* **13 (**24): 6143-51.

Yoshida K. and Miki Y. (2004) Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer Sci* **95 (**11): 866-71.

Young M.C., Kuhl S.B. and Von Hippel P.H. (1994) Kinetic theory of ATP-driven translocases on one-dimensional polymer lattices. *J Mol Biol* **235 (**5): 1436-46.

Yu X., Chini C.C., He M., Mer G. and Chen J. (2003) The BRCT domain is a phospho-protein binding domain. *Science* **302 (**5645): 639-42.

Zheng L., Cash V.L., Flint D.H. and Dean D.R. (1998) Assembly of iron-sulfur clusters. Identification of an iscSUA-hscBA-fdx gene cluster from Azotobacter vinelandii. *J Biol Chem* **273 (**21): 13264-72.

Zittel M.C. and Keck J.L. (2005) Coupling DNA-binding and ATP hydrolysis in Escherichia coli RecQ: role of a highly conserved aromatic-rich sequence. *Nucleic Acids Res* **33 (**22): 6982-91.

APPENDIX 1: OLIGONUCLEOTIDE SEQUENCES

Table 1: Mutagenesis oligonucleotide primer

Name	Sequence 5' to 3'
XPF-S56A	GATCTAGTAAATGCAGTTTTTGACAAAAGG
XPF-S56A-r	CCTTTTGCTAAAAACTGCATTTACTAGATC
XPF-R61A	GTTTTTGACAAAGCGTTTTTTTGACCAA
XPF-R61A-r	TTGGTCAAAAAAACGCTTTGTCAAAAAC
XPF-F63A	GACAAAAGGTTTGCTGACCAAATTAGC
XPF-F63A-r	GCTAATTTGGTCAGCAAACCTTTTGTC
XPF-Q65A	AGGTTTTTTGACGCAATTAGCAGGCTT
XPF-Q65A-r	AAGCCTGCAATTGCGTCAAAAAACCT
XPF-Y108A	GCTACAATTGATGCTGATGTTAAAGTG
XPF-Y108A-r	CACTTTAACATCAGCATCAATTGTAGC
XPD-K35A	CACTGGTAGTGGGGCGACTCTATTTTC
XPD-K35A-r	GAAAATAGAGTCGCCCCACTACCAGTG
XPD-T56A	CCGAAAGTATTGTTCGTAGTCAGAGCACATAATGAG
XPD-T56A-r	CTCATTATGTGCTCTGACTACGAACAATACTTTCGG
XPD-K84H	CTTTTCTTGTAGGTCACCCCTCATCATGTTT
XPD-K84H-r	AAACATGATGAGGGGTGACCTACAAGAAAAG
XPD-C88S	CCCTCATCATCTTTATACGCTGAGAAGGG
XPD-C88S-r	CCCTTCTCAGCGTATAAAGATGATGAGGG
XPD-C102S	GAGCGAAGATATTCCATCCAAATATTGTGAGC
XPD-C102S-r	GCTCACAATATTTGGATGGAATATCTTCGCTC
XPD-C105S	CCATGCAAATATTCCGAGCTTAAAGGCTCT
XPD-C105S-r	AGAGCCTTTAAGCTCGGAATATTTGCATGG
XPD-F136P	GACGGACTTCAGGATAAACCCTGCCCCTACTACTCA
XPD-F136P-r	TGAGTAGTAGGGGCAGGGTTTATCCTGAAGTCCGTC
XPD-C137S	CTTCAGGATAAATTCTCCCCCTACTACTC
XPD-C137S-r	GAGTAGTAGGGGGAGAATTTATCCTGAAG

XPD-Y139F	AAATTCTGCCCCTTCTACTCATTACTC
XPD-Y139F-r	GAGTAATGAGTAGAAGGGGCAGAATTT
XPD-Y140F	AAATTCTGCCCCTACTTCTCATTACTC
XPD-Y140F-r	GAGTAATGAGAAGTAGGGGCAAGAATTT
XPD-Y139A/Y140A	AAATTCTGCCCCGCCGCCTCATTACTC
XPD-Y139A/Y140A -r	GAGTAATGAGGCGGCGGGGCAGAATTT
XPD-Y157A/Y159A	CGTTATAGCCCTAACTGCTCCAGCTTTCTTTATAGATCG
XPD-Y157A/Y159A -r	CGATCTATAAAGAAAGCTGGAGCAGTTAGGGCTATAACG
Rad3 K48A	CCTTCAGGAACAGGTGCAACGGTCTCATTAC
Rad3 K48A-r	GTAATGAGACCGTTGCACCTGTTCCTGAAGG
Rad3 R111H	CTTGGCTTGACATCACACAAAAATTTGTGTTTGC
Rad3 R111H-r	GCAAACACAAATTTTTGTGTGATGTCAAGCCAAG
Rad3 C115S	CAAGAAAAAATTTGAGTTTGCATCCCGAAGTG
Rad3 C115S-r	CACTTCGGGATGCAAACTCAAATTTTTTCTTG
Rad3 Y208A/Y210A	CTCTTTGTAACATTATTATTGCTTCTGCCCATTATCTATTA GATCCTA
Rad3 Y208A/Y210A-r	TAGGATCTAATAGATAATGGGCAGAAGCAATAATAATGT TACAAAGAG

Table 2: Oligonucleotides for DNA substrates

Name	Sequence 5' to 3'
B1-25	CCTCGAGGGATCCGTCCTAGCAAGC
B25comp	GCTTGCTAGGACGGATCCCTCGAGG
B50	CCTCGAGGGATCCGTCCTAGCAAGCCGCTGCTACCGGAAGCTTCTGGACC
B50 comp	GGTCCAGAAGCTTCCGGTAGCAGCGGCTTGCTAGGACGGATCCCTCGAGG
R26-50	TCTCAACTGCAGTCTAGACTCGAGC
T 6	TTTTTTGCTTGCTAGGACGGATCCCTCGAGG
T 9	TTTTTTTTTGCTTGCTAGGACGGATCCCTCGAGG
T12	TTTTTTTTTTTTGCTTGCTAGGACGGATCCCTCGAGG
T15	TTTTTTTTTTTTTTTGCTTGCTAGGACGGATCCCTCGAGG
T20	TTTTTTTTTTTTTTTTTTTTGCTTGCTAGGACGGATCCCTCGAGG
X26-50	GCTTGCTAGGACGGATCCCTCGAGG
X50	GCTCGAGTCTAGACTGCAGTTGAGAGCTTGCTAGGACGGATCCCTCGAGG
60mer Fluor	CAGCGGCTTGCTAGTACGGA[Fl~dT]CCCTCGAGGCCGCAGGTACTCGTGAATTCCGCTATTCGG
60mer	CAGCGGCTTGCTAGTACGGATCCCTCGAGGCCGCAGGTACTCGTGAATTCCGCTATTCGG
45mer Fluor	CCGAATAGCGGAATTCACGAGTACCTGCGGCCTCGAGGGA[Fl~dT]CCGT
45mer	CCGAATAGCGGAATTCACGAGTACCTGCGGCCTCGAGGGATCCGT
Bubble4	GGTCCAGAAGCTTCCGGTAGCAGACCGTTGCTAGGACGGATCCCTCGAGG
Bubble10	GGTCCAGAAGCTTCCGGTAGGCTACCGCACCTAGGACGGATCCCTCGAGG

Table 3: Oligonucleotides for length dependent ATPase assays

Name	Sequence 5’ to 3’
6nt	GGCGCG
14nt	GTACCAGCATGAAC
25nt	GCTTGCTAGGACGGATCCCTCGAGG
50nt	GCTCGAGTCTAGACTGCAGTTGAGAGCTTGCTAGGACGGATCCCTCGAGG
70nt	GGTCCAGAAGCTTCCGGTAGCAGCGGCTTGCTAGGACGGATCCCTCGAGGCCGCAGGTACTCGTGAATTC
100nt	GCTCGAGTCTAGACTGCAGTTGAGAGGTCCAGAAGCTTCCGGTAGCAGCGGCTTGCTAGGACGGATCCCTCGAGGCGTCGAGGAATTCAACCACCGCTCC

APPENDIX 2: ITC-DATA

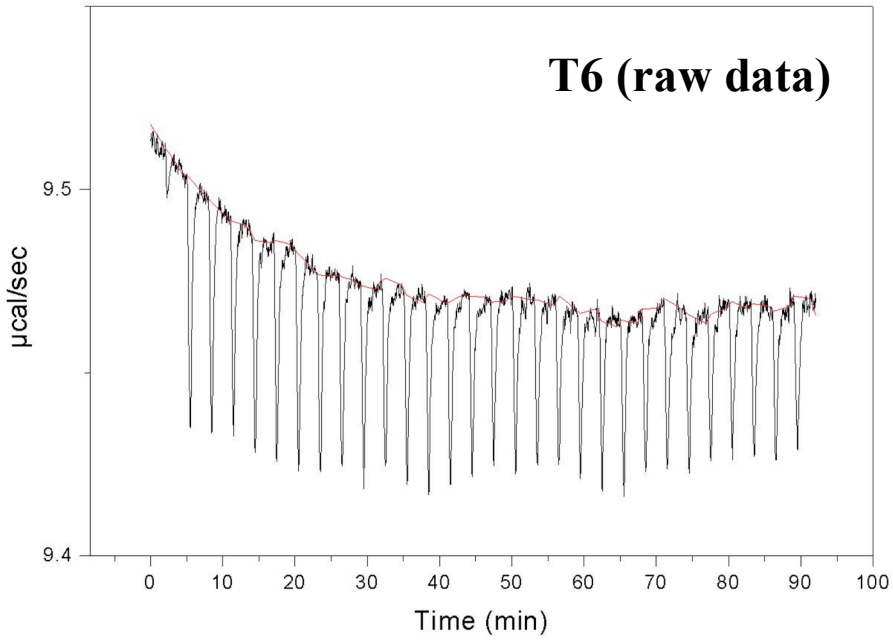


Figure 1: ITC data of XPD binding to a 6mer oligonucleotide

The figure only shows the raw ITC data of XPD binding to a 6 nt dT oligonucleotide. No DNA binding was observed at 25 ˚C, hence data not further processed.

Table 4: Summary of XPD-DNA binding data obtained by ITC

	T20	**T14**
Chi^2	34870.6	15881.6
Binding sites (N)	0.3915±0.01314	0.4810±0.0419
Binding constant (K)	4.441E5±2.918E4 l mol^{-1}	1.936E5±1.959E4 l mol^{-1}
K_D	2.5±0.3 µM	5.2±0.5 µM
Change in Enthalpy (ΔH)	-2.339E4±993.3 cal mol^{-1}	-1.596E4±1722 cal mol^{-1}
Change in Entropy (ΔS)	-52.6 cal mol^{-1} K^{-1}	-29.34 cal mol^{-1} K^{-1}

Table 4 shows protein-DNA binding data obtained by ITC. Experiments were carried out at 25 ˚C and binding of XPD to a 14 nt or 20 nt dT oligonucleotide was investigated.

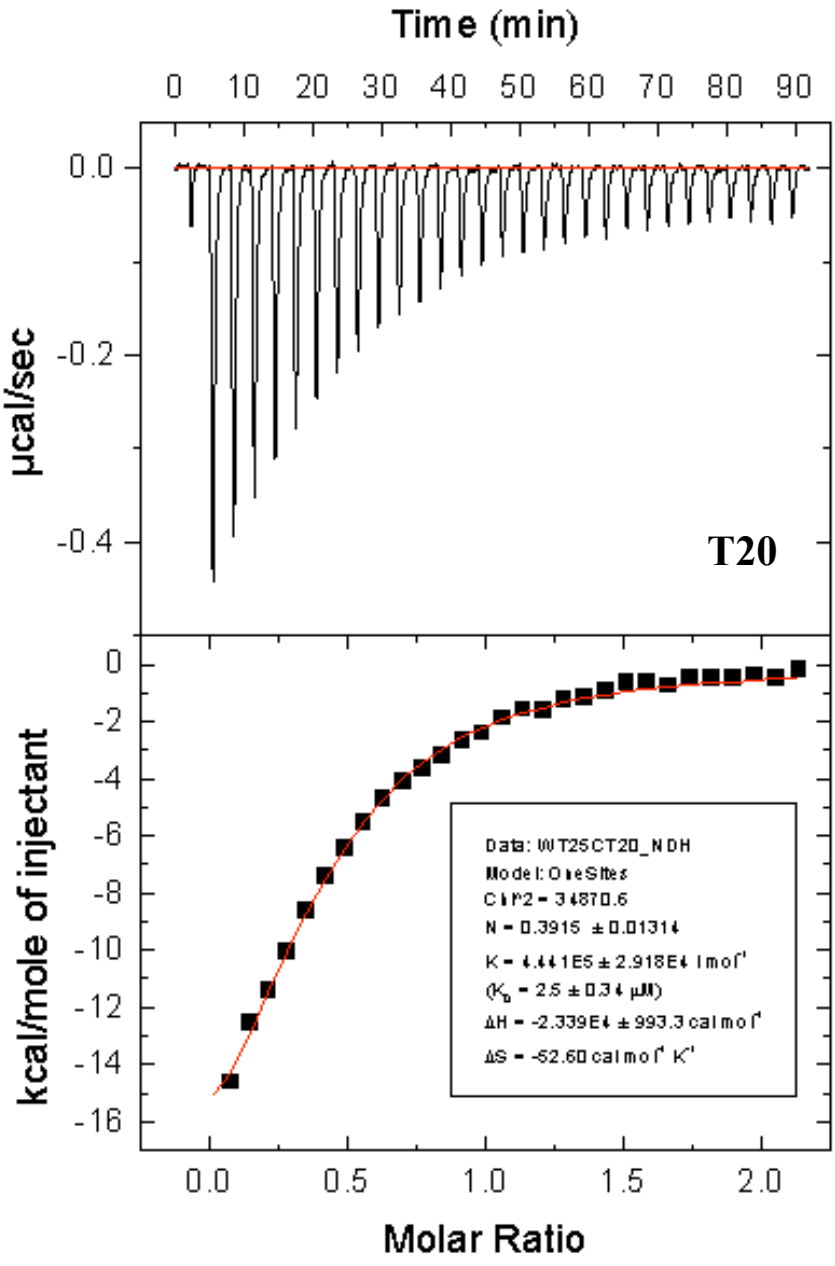


Figure 2: ITC data of XPD binding to a 20mer oligonucleotide

The figure shows processed ITC data after baseline correction (-1600 cal mole^{-1}) of XPD (10.6 µM) binding to a 20 nt dT oligonucleotide (100.9 µM). Experiments were carried out at 25 ˚C. No DNA binding was observed at 45 ˚C. Data were fitted to a one site of binding model using MicroCal^{TM} ORIGIN software, assuming XPD binds DNA in a 1:1 ratio. The results are summarised in Table 4.

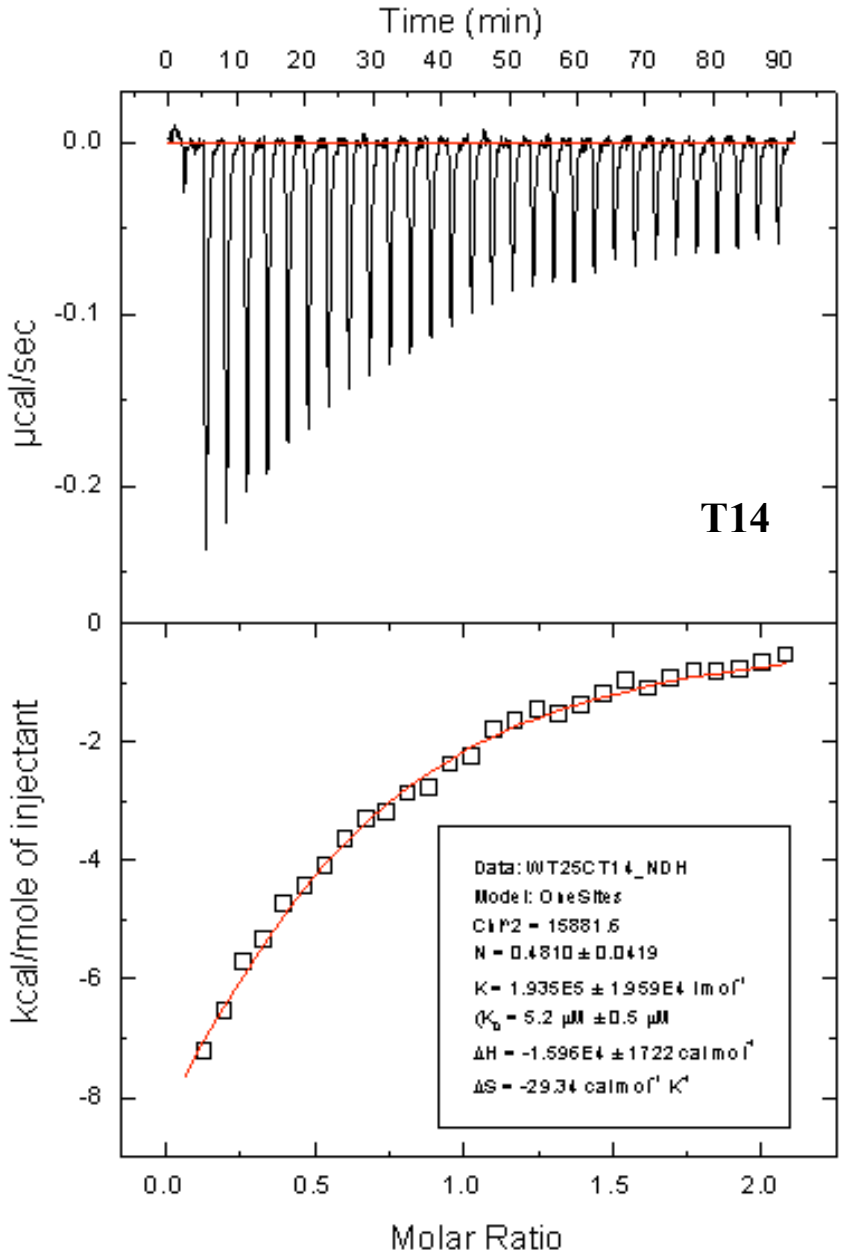


Figure 3: ITC data of XPD binding to a 14mer oligonucleotide

The figure shows processed ITC data after baseline correction (-1500 cal mole^{-1}) of XPD (10.6 µM) binding to a 14 nt dT oligonucleotide (99.1 µM). Experiments were carried out at 25 ˚C. No DNA binding was observed at 50 ˚C. Data were fitted to a one site of binding model using MicroCal^{TM} ORIGIN software, assuming XPD binds DNA in a 1:1 ratio. The results are summarised in Table 4.