Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation

Jochen Kuper^{1,*}, Stefanie C Wolski¹,
Gudrun Michels and Caroline Kisker*

Rudolf Virchow Center for Experimental Biomedicine, Institute for Structural Biology, University of Würzburg, Würzburg, Germany

The XPD protein is a vital subunit of the general transcription factor TFIIF which is not only involved in transcription but is also an essential component of the eukaryotic nucleotide excision DNA repair (NER) pathway. XPD is a superfamily-2 5'-3' helicase containing an iron-sulphur cluster. Its helicase activity is indispensable for NER and it plays a role in the damage verification process. Here, we report the first structure of XPD from *Thermoplasma acidophilum* (taXPD) in complex with a short DNA fragment, thus revealing the polarity of the translocated strand and providing insights into how the enzyme achieves its 5'-3' directionality. Accompanied by a detailed mutational and biochemical analysis of taXPD, we define the path of the translocated DNA strand through the protein and identify amino acids that are critical for protein function.

The EMBO Journal (2012) 31, 494–502. doi:10.1038/emboj.2011.374; Published online 11 November 2011

Subject Categories: genome stability & dynamics

Keywords: DNA repair; helicase; nucleotide excision repair; SF2B; XPD

Introduction

Nucleotide excision repair (NER) is a highly versatile DNA repair pathway (Van Houten, 1990; Sancar, 1996; Goosen and Moolenaar, 2001; Friedberg *et al.*, 2006; Gillet and Scharer, 2006), which is exemplified by its ability to remove a broad range of DNA lesions, including benzo[a]pyrene-guanine adducts caused by smoking, guanine-cisplatin adducts resulting from chemotherapy, and photoproducts generated by ultraviolet light (Sancar, 1994). For the latter lesion, NER is the only pathway in humans that is able to resolve these damages. The importance of NER is reflected in the severe phenotypes of three different human diseases: xeroderma pigmentosum, Cockayne syndrome, and trichothiodystrophy (van Gool *et al.*, 1997; Vermeulen *et al.*, 1997; de Boer and

Hoeijmakers, 2000; Bergmann and Egly, 2001; Friedberg *et al.*, 2006).

It is known that a total of around 30 proteins are involved in eukaryotic NER; however, the individual contribution of each protein is still a matter of debate. A general scheme of action has emerged and the first complex to interact with the damaged site consists of the XPC and Rad23B proteins. XPC–Rad23B most likely recognizes only distortions in Watson–Crick base pairing, thus further damage verification is required to identify the lesion (Tapias *et al.*, 2004; Min and Pavletich, 2007; Sugawara and Hanaoka, 2007). After formation of the XPC–Rad23B–DNA complex additional factors are recruited to the putative site of the damage, the first being the general transcription factor TFIIF. TFIIF not only plays a role as a general transcription factor of RNA polymerase I and II but is also an essential part of the NER machinery (Zurita and Merino, 2003). XPD and XPB are its most important subunits as they open the DNA around the lesion in an ATP-dependent fashion (Zurita and Merino, 2003). Both XPD and XPB are superfamily 2 (SF2) helicases, however, they operate with opposite polarity. The two subunits play very distinct roles since the ATPase and not the helicase activity of XPB is essential to associate TFIIF with the lesion (Coin *et al.*, 2007) while the actual opening of the DNA around the lesion is performed by the helicase activity of XPD (Coin *et al.*, 2007). During this process XPD is in a prime position to verify the damage, ensuring that the lesion is not only a backbone distortion resulting from an unusual DNA sequence. It has been shown that the yeast homologue of XPD is stalled at sites of damage (Naegeli *et al.*, 1993); however, it is not clear whether the action of XPD alone, or a combined action with other NER factors, such as XPA lead to the final damage verification process. Concomitantly, further factors are recruited including XPA, RPA, and the endonucleases XPG and XPF-ERCC1 (Evans *et al.*, 1997a,b; Gillet and Scharer, 2006).

Helicases and translocases can be grouped into six superfamilies (SF1–6) of which only SF1 and SF2 show a tandem repeat of the core RecA-like motor domains (Singleton *et al.*, 2007). SF1 and 2 helicases can each be divided into A and B subfamilies, which display either 3'-5' or 5'-3' polarity, respectively (Singleton *et al.*, 2007). An intriguing question is how translocation polarity is achieved in this enzyme scaffold. The very similar motor domains translate the chemical energy into motion and, in principle, translocation is achieved by altering the position between the motor domains upon ATP hydrolysis (Velankar *et al.*, 1999). In SF1A family enzymes translocation is accomplished using an alternating affinity cycle of the motor domains resulting in motor domain 2 (HD2) as the leading domain binding more tightly to the DNA. ATP hydrolysis causes a reversal, and as a consequence, motor domain 1 (HD1) binds tighter and HD2 is released resulting in translocation (Velankar *et al.*, 1999). To generate the opposite polarity, two possible mechanisms

*Corresponding authors. J Kuper or C Kisker, Rudolf Virchow Center for Experimental Biomedicine, University of Würzburg, Josef-Schneider-Strasse 2, Würzburg 97080, Germany.
Tel.: +49 931 31 80391; Fax: +49 931 31 87320;
E-mail: jochen.kuper@virchow.uni-wuerzburg.de
or Tel.: +49 931 31 80381; Fax: +49 931 31 87320;
E-mail: caroline.kisker@virchow.uni-wuerzburg.de

¹These two authors contributed equally to this work

Received: 1 July 2011; accepted: 20 September 2011; published online: 11 November 2011

can be envisioned. The helicase scaffold could bind the DNA with a different orientation using a similar translocation mechanism, or the DNA is bound in the same orientation in both subfamilies with the translocation mechanism being altered. The latter has been shown recently for SF1A and SF1B enzymes (Saikrishnan *et al*, 2009) with a reversed order of tight binding for HD1 and HD2 resulting in opposite translocation polarity. For SF2 enzymes, it has so far not been possible to derive similar conclusions.

XPD belongs to the SF2B subfamily of helicases and contains a 4Fe–4S cluster (FeS) in an auxiliary domain, that is vital for its helicase activity and might be involved in the damage recognition process (Rudolf *et al*, 2006; Fan *et al*, 2008; Liu *et al*, 2008; Wolski *et al*, 2008). Although apo structures of SF2A and SF2B members are available, no DNA-bound structure of an SF2B enzyme has been reported so far which would allow to pin point the underlying translocation mechanism in this family. We have chosen *Thermoplasma acidophilum* XPD (taXPD) as a representative member of SF2B helicases and a homologue of human XPD as a model enzyme. After determining the structure of apo-taXPD (Wolski *et al*, 2008) we have now solved the first structure of an XPD–DNA complex. To identify functionally important regions of the protein and determine the DNA-binding regions of taXPD, we performed a detailed mutational and biochemical analysis to address mechanistically important questions on translocation polarity, helicase function and damage recognition.

Results

The taXPD DNA structure

In order to obtain insights into the molecular events leading to DNA translocation of taXPD and the detailed nature of the DNA-binding mechanism we have solved the crystal structure of taXPD in the presence of a 22-mer single-stranded DNA (ssDNA) oligonucleotide. The structure was refined at a resolution of 2.2 Å to an *R* factor of 19.5% (*R*_{free} = 25.1%) with good stereochemistry as indicated by the Ramachandran statistics (Table I). After molecular replacement using the apo structure of taXPD (pdb entry 2VSF) as search model we obtained clear electron density for the additional N-terminal helix (residues 5–23) that has been introduced in the expression construct used in this study (Figure 1A and C; Supplementary Figure S1) thus representing the full-length *T. acidophilum* XPD. This additional helix contains the Q-motif that is also present in FancJ and several other helicases. The overall fold of the helicase scaffold does not change significantly when compared with the structure of apo-taXPD (Wolski *et al*, 2008) and can be superimposed with a root mean square deviation of 0.86 Å using 572 Cα atoms. The only major difference is an eight-residue loop (residues 421–429) that is disordered in the taXPD–DNA complex structure. In addition to the electron density for the N-terminal α-helix four other significant electron density features were observed (Supplementary Figure S2). Three of them could be identified as sulphate molecules originating from the crystallization solution. The first sulphate molecule is located in the ATP-binding pocket of the Walker A motif, representing a possible phosphate-binding site, most likely the α-phosphate (Figure 2A). This position is in good agreement with the position of the phosphates of the nucleotide-bound

Table I Data collection and refinement statistics

Space group	P6 ₅
Wavelength (Å)	0.972
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	79.1, 79.1, 175.7
α, β, γ (deg)	90, 90, 120
Resolution (Å)	44.5–2.2 (2.32–2.2)
Completeness (%)	100 (100)
<i>R</i> _{sym}	0.085 (0.583)
<1/σI>	9.4 (2.1)
Redundancy	3.8 (3.8)
Refinement	
Resolution (Å)	52–2.2
No. of reflections	29 606
<i>R</i> _{cryst} / <i>R</i> _{free}	0.195/0.251
No. of atoms	
Protein/ions/DNA	4933/16/96
Water	138
B factors (Å ²)	
XPD	45.6
DNA	75.9
Bond lengths (Å)	0.012
Bond angles (deg)	1.736
Ramachandran statistics (%)	95.4/4.1/0.5

Values in parentheses refer to the highest resolution shell.

$$R_{\text{sym}} = \sum_{\text{hkl}} \{ \sum_i |I_i - \langle I \rangle| / \sum_i \langle I \rangle \}$$

$R_{\text{cryst}} = \sum_i ||F_o| - |F_c|| / \sum_i |F_o|$ where *F*_o and *F*_c are the observed and calculated structure factors, respectively. *I*/σ*I* indicates the average of the intensity divided by its average standard deviation. *R*_{free} is the same as *R*_{cryst} for 5% of the data randomly omitted from refinement. B-factor analysis was performed on the full B factors after TLS refinement. Ramachandran statistics indicate the fraction of residues in the favoured, allowed, and outlier regions of the Ramachandran diagram.

Figure 1 The XPD–DNA complex. (A) Overall structure of XPD with the two RecA-like domains in yellow and red, the FeS cluster domain in cyan, and the arch domain in green. The DNA identified in the electron density is shown in orange. (B) Enlarged view showing the tetranucleotide visualized in the structure with its carbon atoms in light blue. Residues interacting with the DNA are shown with their carbon atoms in grey and hydrogen bonds are indicated by dashed green lines. (C) Side view of the taXPD–DNA complex. Colour scheme is similar to (A). The cleft where the DNA is bound is indicated with arrows. The additional N-terminal helix harbouring the Q-motif is shown in grey. (D) Combination of experimentally verified DNA (orange) with modelled DNA (grey). The colour scheme for XPD is as described above.

structures of NS3 helicase in the presence of the nucleotide analogue AMPPNP and UvrB in the presence of ATP (Supplementary Figure S3). The second sulphate is located directly in the basic groove at the exit of the pore and forms hydrogen bonds with the hydroxyl group of Y166, the N ζ atom of K170, the N η 1 atom of R88, and the N ϵ 2 atom of H198 at equal distances of ~ 2.8 Å (Figure 2B). The last sulphate site is located in the arch domain at α -helix 13 and is additionally involved in crystal contacts (data not shown). It is, therefore, unlikely that it represents a physiologically relevant site and will not be discussed further.

The fourth and largest feature was identified as a four base fragment originating from the 22-mer ssDNA oligonucleotide that has been utilized for co-crystallization. The sequence could be mapped from the electron density as TACG and is positioned in the 5' region of the DNA substrate (Figure 1A and B). The fragment is located at a cleft in HD2 (Figure 1A and B) that has been formerly predicted as a DNA-binding site (Wolski *et al*, 2008). The protein–DNA interface buries an accessible surface area of 370 Å² and is comprised in near equal parts of non-polar and polar residues, with the latter group consisting of equal ratios of charged and uncharged polar residues as analysed by PROTORG (Reynolds *et al*, 2009). The most prominent protein–DNA interactions

are mediated through Trp549 and Arg584 (Figure 1B). The indole ring of Trp549 forms π -stacking interactions with the adenine base at a distance of 3.7 Å and is responsible for most of the non-polar surface interactions. The N η 1 and N η 2 atoms of Arg584 form hydrogen bonds with the phosphate backbone of the adenine and cytosine bases. Asp582 is not directly located at the interface but is stabilizing Arg584 via hydrogen bonds and thereby keeps it in position for DNA binding. Although no direct interaction of Phe538 with the cytosine base can be observed (Figure 1B), one can speculate that this residue is involved in stabilizing larger bases such as adenine or guanine.

Most importantly, however, the electron density enabled us to identify the polarity of the crystallized DNA fragment. This is of paramount interest since there has been no structural information for an SF2B helicase bound to DNA so far, and consequently the detailed mechanism of translocation and unwinding could not be deduced for these enzymes. The 5'-end is pointing away from the HD2 domain into the solvent, whereas the 3' end is pointing towards HD1 (Figure 1A and B). Comparing the DNA fragment with our previously proposed DNA model, one can observe that the DNA strand in the protein–DNA complex is directly extended in our model (Figure 1D). Taking the directionality of the fragment into account it clearly demonstrates how the translocated strand is situated within the helicase scaffold.

Figure 2 Putative phosphate positions in the binary XPD–DNA complex. (A) The first sulphate molecule is located in the ATP-binding pocket of the Walker A motif in HD1 and is shown in all bonds representation. Hydrogen bonds are indicated by dashed green lines. (B) The second sulphate molecule is located in close proximity to the FeS cluster, in the basic groove at the exit of the pore.

Biochemical properties of taXPD variants

Based on the crystal structure of the taXPD–DNA complex, our previously proposed DNA-binding model (Wolski *et al*, 2008), and sequence alignments (Supplementary Figure S4) we have rationally designed several taXPD point mutations. These variants provide a basis to assess how the path of the translocated DNA strand extends from the experimentally verified stretch of ssDNA and identify functionally important regions of the protein that also apply to eukaryotic XPDs. We generated the variants listed in Table II and characterized the ssDNA-binding properties via biolayer interferometry, their ability to hydrolyse ATP and their helicase activity. To rule out artefacts due to misfolding or instability every variant was subjected to thermal unfolding studies and only proteins that exhibited an unfolding transition similar to the wild-type protein were analysed further (Supplementary Figure S5). Wild-type taXPD displays a dissociation constant (K_D) of

Table II Binding constants, ATPase activity, and helicase activity

Variant	K_D (nM)	ATPase activity (mol ATP*s ⁻¹ *mol ⁻¹ XPD)	Helicase activity (Rel. fluorescence change*s ⁻¹)	Class
Wild type	61 ± 3.5	2.0 ± 0.6	333.1 ± 20.7	
R59A	384 ± 83.7	0.1 ± 0.03	218.3 ± 15.6	I +
R88H	298 ± 57.9	3.1 ± 0.5	357.7 ± 30.4	II +
E107A	403 ± 79.7	2.9 ± 0.4	273.6 ± 16.6	II +
F133A	353 ± 43.1	1.8 ± 0.80	901.8 ± 21.5	II (+)
Y166A	145 ± 67.7	4.7 ± 1.8	36.3 ± 4.5	II +
K170A	2122 ± 403	3.8 ± 1.2	1054.2 ± 59.9	II +
F326A	594 ± 48.9	3.7 ± 1.1	17.5 ± 2.3	II (+)
Y425E	3323 ± 850	0.17 ± 0.09	35.6 ± 5.4	I +
W549S	28 256 ± 22 839	ND	106.3 ± 22.8	I (+)
D582N	697 ± 78.5	0.12 ± 0.06	ND	I +
R584E	2260 ± 366.7	0.07 ± 0.02	ND	I +

Activities for the relevant parameters are given as indicated. ND indicates that no activity could be detected, + in the class row refers to conserved residue and (+) indicates residues where a homologue can be found in close vicinity as indicated in Supplementary Figure S4.

Figure 3 Model of the XPD–DNA complex and location of the XPD variants. (A) Model of the XPD–DNA complex using the apo-XPD structure with the protein shown in a transparent surface representation. The different variants are shown as CPK models in green, orange, and red, indicating the strength of reduction for DNA binding compared with the wild-type protein (red ~450-fold reduction; orange ~35-fold reduction; green ~5-fold reduction (Y166 is an exception and displays only an ~2.5-fold reduction) relative to the wild-type protein). The surface of the subdomains of XPD is coloured as follows: HD1 (yellow), HD2 (red), arch (green), and FeS (cyan). The modelled ssDNA is shown in a cartoon representation. Experimental ssDNA is shown with an orange backbone and the ssDNA model in black. (B) Close-up view into the area of HD2 where the cleft is located. (C) Residues located at the pore, or in close proximity to the pore, after rotation of the molecule by 180° as indicated by the arrow.

61 nM for the ssDNA substrate used in our study, indicating a high-affinity binding to this target (Table II). This is comparable to the ssDNA-binding affinity of *Sulfolobus acidocaldarius* XPD (Fan *et al*, 2008).

In our previously published model (Wolski *et al*, 2008), we proposed a path for the translocated strand along the two motor domains (HD1 and HD2) which then leads through a pore created by the arch domain, HD1, and the FeS domain (Figures 1D and 3). The protein–DNA model is shown in a surface representation in Figure 3 and the variants under investigation are depicted on the surface in different colours according to their DNA-binding phenotype (red, ~450-fold reduction; orange at least 35-fold reduction; green at least 5-fold reduction compared with the wild-type protein (the only exception is Y166A which displays only an ~2.5-fold reduction)). The proposed DNA-binding region is in very good agreement with those variants showing a significantly decreased affinity for ssDNA. The strongest effects on ssDNA binding were observed in the second motor domain (HD2) where Tyr425, Trp549, Asp582, and Arg584 are located at the side of a small cleft on its surface (Figure 3A and B). This cleft

is formed by α -helices 22 and 24 (for a reference to the secondary structure elements, see Supplementary Figure S1A and B) on one side and a loop located between β -strand 14 and the 3₁₀-helix 3 on the opposite side. The W549S variant displays the strongest phenotype with a K_D of only 28 μ M (Table II), an affinity which is almost a thousand times reduced compared with the wild-type protein. This result is in very good agreement with our crystallographic data, since the observed ssDNA fragment is bound in this cleft.

In the proposed model, the ssDNA further extends between the two motor domains as indicated by the reduced affinity of the R59A variant. Arg59 is strictly conserved and located in HD1 at the interface between HD1 and HD2 and points towards the channel formed by these two domains (Table II; Figure 3A and B). Our model suggests that the DNA is protruding through the pore formed by HD1, the arch domain, and the iron–sulphur cluster domain. We designed four variants within or in close proximity to the pore to cover the space presumably occupied by DNA, namely E107A, R88H, Y166A, and F326A (Figure 3A and C). All four variants display a decreased binding affinity (2.5-fold for Y166A and between 5-fold and 8-fold for the other three mutants) indicating that these residues are involved in protein–DNA interactions and thus further support the model that the translocating strand is passing through the pore (Table II; Figure 3A and C). After passage through the pore the translocated DNA strand most likely interacts with residues located in a pronounced basic depression that is formed by two α -helices (α 5 and α 8) from the FeS domain and α -helix 10 from HD1. In this depression, the exchange of Lys170 to alanine reduces the DNA-binding affinity at least ~30-fold (Table II; Figure 3A and C), thus strongly supporting the notion that this region also interacts with the DNA. Phe133 is located at the very end of this depression and demarks the outer end of the protein at this side. Interestingly, the F133A variant shows significantly decreased affinity for ssDNA indicating its involvement in DNA binding (Table II; Figure 3A and C). Taken together, the location and binding affinities of the different variants suggest that the path of ssDNA stretches nearly over the complete length of the protein covering a distance of ~67 Å, which could thus accommodate around 20 bases of ssDNA.

In addition to the ssDNA-binding analysis, we performed ATPase activity measurements in the presence and absence of ssDNA. In the absence of ssDNA, the ATPase activity of taXPD is hardly detectable (data not shown). Upon addition of 0.5 μ M ssDNA, the specific activity of wild-type taXPD is 2.0 mol ATP $\cdot$ s^{−1} $\cdot$ mol^{−1} XPD (Table II) and is thus approximately four times higher than that of XPD from *S. acidocaldarius* (Fan *et al*, 2008). Interestingly, the variants can clearly be divided into two classes with respect to their location and phenotype; class I is defined by variants located in the motor domains of the protein and, using the orientation displayed in Figure 1A, in ‘front’ of the pore while class II is defined by variants in the arch domain, the FeS domain, and HD1 which are located in or ‘behind’ the pore. As shown above, all class I variants (Table II) display a significantly decreased affinity for ssDNA which is accompanied by a significant loss of ssDNA-induced ATPase activity. The most strongly affected variant is W549S that displays no measurable ATPase activity and also exhibits the highest K_D value. Following the proposed path of the ssDNA up to the pore and ending with R59A, all variants

Figure 4 Helicase activity of taXPD and its variants. The activity of wild-type XPD and the variants is shown in a bar diagram. The inset shows a typical kinetic measurement; ND indicates that the activity was not detectable.

display at least a ten-fold reduction in ATPase activity. However, mutants located within or behind the pore display a different phenotype, and the correlation between weakened DNA binding and reduced ATPase activities no longer exists. In these variants, the ATPase activity remains high despite impaired ssDNA binding. The first class II variant, Y166A, is located directly in the pore, and displays an elevated activity of $4.7 \text{ mol ATP} \cdot \text{s}^{-1} \cdot \text{mol}^{-1}$ compared with $2.0 \text{ mol ATP} \cdot \text{s}^{-1} \cdot \text{mol}^{-1}$ for the wild type. Likewise the K170A, F326A, R88H, and E107A variants exhibit a significantly elevated ATPase activity that is at least 50% increased compared with the wild-type protein (Table II). The activity of F133A is, however, comparable to wild-type taXPD.

We also investigated the dsDNA unwinding capability of wild-type XPD and its variants (Table II; Figure 4) utilizing a fluorescence-based assay followed by native PAGE analysis to substantiate our results. A typical kinetic measurement of taXPD is shown in the inset of Figure 4. Using the nomenclature of class I and class II variants, it becomes evident that all class I variants (R59A, Y425E, W549S, D582N, and R584E), in addition to their defects in DNA binding and ATPase activity, are also impaired in their ability to unwind dsDNA (Figure 4). Notably, the two variants D582N and R584E that mimic mutations in human XPD that lead to XP display no measurable activity. The phenotype of the class II variants, however, displays a more complex pattern. For the R88H and E107A variants, helicase activity seems to be unaltered compared with wild-type XPD, despite the defects in DNA binding and altered ATPase activity. Interestingly, while F133A and K170A also show a decreased binding affinity for ssDNA accompanied by wild-type (F133A) or elevated (K170A) ATPase activity, both mutants exhibit a significantly increased helicase activity. K170A displays the highest helicase activity (~ 3 -fold elevated compared with wild type) and combined with the elevated ATPase activity it can be concluded that in this case the higher turnover rate for ATP also results in a higher helicase activity. In contrast, the Y166A and F326A variants also display an elevated ATPase activity but show a highly reduced (F326A, ~ 20 -fold and Y166A, ~ 9 -fold) helicase activity.

Discussion

Recently, three structures of apo-XPD from archaeal organisms have been reported which revealed the overall architecture of this important helicase. All groups proposed a

Figure 5 Different functionalities within the XPD protein. The surface representation of XPD and the four views highlight different possible functions of specific residues within the protein. The red area is located in HD2 and may play a role in ratcheting, whereas the yellow area depicts residues abolishing helicase activity. The green areas show residues generating a hyperactive helicase. The arrows indicate the location of important features discussed in the text. The two arrows for the basic groove denote the start and end point of the groove.

similar path for ssDNA within the protein scaffold (Fan *et al*, 2008; Liu *et al*, 2008; Wolski *et al*, 2008). Other studies have implicated that the 3' end of the DNA is located close to the FeS domain (Pugh *et al*, 2008; Honda *et al*, 2009). However, the polarity and exact location of the ssDNA and thus vital information about strand separation and ratcheting elements in the helicase remained unclear. The structure of our taXPD–DNA complex combined with our mutational data defines the path and polarity of the translocated ssDNA with respect to the helicase scaffold (Figure 1D). Furthermore, we have identified variants incapable of unwinding dsDNA thus defining important functional parts of the helicase. This has major implications for the mechanism of translocation along DNA, how directionality is achieved and the elements involved in unwinding of dsDNA (Figure 5).

For SF1 family helicases (exemplified by PcrA and RecD2), it has been shown recently that both subgroups, that is, SF1A and SF1B helicases bind DNA in a similar orientation but translocate in opposite directions along the DNA through inverted actions of the RecA-like motor domains (Saikrishnan *et al*, 2009). XPD displays a comparable DNA-binding orientation as observed in the 3'-5' SF2A helicases NS3 and Hel308 for the translocated strand (Kim *et al*, 1998; Buttner *et al*, 2007). Our structure, therefore, suggests that also in SF2 superfamily helicases the orientation of the DNA is unchanged, whereas directionality is achieved via different means. Since the overall layout of the motor domains is similar between SF1A/B and SF2A enzymes, the mechanism of translocation in the 5'-3' direction could occur by analogous means as in SF1B family helicases, namely the reversed order of gripping between the motor domains.

We previously identified a wedge-like feature in HD2 as a potential DNA-binding element (Wolski *et al*, 2008). Our present data unambiguously show that this feature is critical for DNA binding, since the corresponding amino acids Trp549 and Phe538, which directly interact with the bound DNA, are located in this wedge-like element. However, from the polarity of the bound DNA it can be excluded that strand separation takes place at this position. The 5'-end of the DNA in the XPD–DNA complex at this position is pointing towards the solvent, whereas the 3'-end points towards HD1 (Figure 1A and D). This result, in combination with the binding data of the accompanying mutational analysis, clearly defines the directionality of the DNA and its path within the protein. As a consequence, the 5'-end of the DNA is located at the outer part of HD2 and its 3'-end is situated at the other end of the protein, 'behind' the pore at the aforementioned basic depression (Figure 5). Since XPD is known to translocate in the 5'-3' direction, strand separation most likely occurs at the 3'-end of the translocated strand, namely at the FeS domain which is in line with data implicating that the FeS domain binds close to the 3'-end of DNA (Pugh *et al*, 2008; Honda *et al*, 2009).

The observed DNA-binding site in HD2 could provide the necessary grip on the DNA for ratcheting. Notably, neither Trp549 nor Phe538 is positioned in a strictly conserved region of the sequence alignment (Supplementary Figure S4). However, in very close vicinity to Trp549 other hydrophobic amino acids such as a highly conserved phenylalanine positioned three residues C-terminally could potentially fulfil this role in the eukaryotic enzyme. Additionally, the importance of this feature in XPD is also underlined by the XP-causing mutations D681N and R683W/E that correspond to Asp582 and Arg584 in taXPD, the latter directly interacting with the phosphate backbone of the bound DNA. Both variants show severe effects with respect to DNA binding, ATPase, and helicase activity. Tyr425 also belongs to this feature and displays a strongly decreased binding and ATPase phenotype as well as reduced helicase activity when mutated to a glutamate. Unfortunately, Tyr425 is part of a disordered loop in the protein–DNA complex, and therefore no further conclusions can be derived regarding the role of this residue with respect to the formation of a protein–DNA complex. Taken together, this region comprises amino-acid residues that display the most significant phenotypes when analysed for ssDNA binding. This observation could at least partially explain why only at this position ssDNA from the 22-mer was observed in the crystal structure and not also further down in our predicted path of DNA. The remainder of the ssDNA might just be too flexible to be observed or could have been attacked by trace amounts of DNase which was used during XPD purification. The flexibility could also be enforced not only by the lower impact on binding of residues down the predicted path of ssDNA but also due to the conformation the helicase adopts in this complex. TaXPD seems to be in a preprocessive state when compared with NS3 and Hel308 (Wolski *et al*, 2008), which might also lower the affinity for the ssDNA further down the path.

A highly interesting position in the protein is marked by the putative phosphate-binding site bridging the FeS domain and HD1 (Figure 2B). Three of the four residues (Supplementary Figure S4; Figure 2B) interacting with the sulphate are strictly conserved and have been mutated resulting in three different phenotypes. R88H, Y166A and K170A all

display decreased ssDNA binding, at least to a certain extent, and a significantly elevated ATPase activity; however, the influence of the exchanges on the helicase activity of XPD varies greatly. While the R88H variant seems to be unaffected, the K170A variant displays an increased helicase activity and the Y166A variant displays a significantly reduced activity. At the same time, the latter mutant displays the smallest difference compared with the wild-type protein with respect to ssDNA binding of all three variants. It seems that this area marks a 'hot spot' for helicase regulation in both a positive and negative fashion, indicating a high functional importance of this site. The actual mechanism of regulation remains elusive in the absence of a complete XPD DNA structure that would also cover this region of the protein. However, it is tempting to speculate that Tyr166 could be involved in stacking interactions thereby promoting strand translation. Together with another highly conserved aromatic side chain, Tyr185, it forms a hydrophobic pocket indicating a possible binding site for DNA bases. A double mutant of Tyr166 and Tyr185 to alanine shows a decreased ssDNA-binding phenotype and a severely impaired ATPase activity (data not shown). However, the stability of the variant was affected as well indicating structural effects resulting from these amino-acid exchanges, which renders the exact interpretation of the phenotype difficult. Lys170 on the other hand seems to negatively regulate helicase activity, which was also observed for the fourth residue involved in the putative phosphate-binding site, namely His198. This residue was mutated by Pugh *et al* and shows a behaviour analogous to Lys170 underlining the functional relevance of this motif (Pugh *et al*, 2012). Notably, K170A shows the highest impairment in ssDNA binding among the class II mutants. The reduced interaction with the DNA substrate seems to directly increase the helicase and ATPase activity. Lys170 could therefore act as an anchor negatively regulating activity. Another variant showing a similar phenotype is F133A. Here, we observed a highly elevated helicase activity but without the increase of ATPase activity indicating a different mechanism of regulation.

Arg88, the third residue involved in the 'hot spot' does not show an alteration in helicase activity when mutated to histidine despite its higher ATPase activity. This finding is in contrast to previous studies with the enzyme from *S. acidocaldarius* (Rudolf *et al*, 2006; Liu *et al*, 2008). The difference can be explained by a decay of the FeS cluster that seems to be less stable in the *Solfulobus* enzyme. This would also likely explain the missing helicase activity of the human enzyme (Dubaele *et al*, 2003). Arg88 bridges the 'hot spot' sulphate position directly with the FeS cluster through hydrogen bonds formed between the guanidinium group of Arg88 with the sulphate and the FeS cluster coordinating Cys113, respectively, thus underlining the importance of Arg88. In taXPD, the FeS cluster seems to be more stable so that it is not dramatically affected by the R88H exchange. When mutated to alanine, however, the helicase activity is abolished, which is likely due to an alteration of the FeS cluster (Pugh *et al*, 2012).

The F326A substitution displays a similar phenotype as observed for the Y166A mutant with a significant reduction in helicase activity, an increase in ATPase activity and a moderate ssDNA-binding phenotype. Interestingly, Phe326 is located between the arch and the FeS domain where it is

partly embedded in their connecting interface. The phenyl ring of Phe326 is pointing towards the pore and might be involved in stacking interactions with bases of the DNA. Additionally, this mutation may affect the required flexibility between the arch and the FeS domain, thereby reducing helicase activity. This is the first time that the arch domain is implicated in helicase activity demonstrating that not only the FeS domain is of functional importance with respect to dsDNA unwinding but also the relative orientation of the two domains or the ‘ease’ to separate the two domains may be important as well. It was shown in *in-vitro* studies that XPD unwinds dsDNA substrates containing a bubble (Rudolf *et al*, 2010). Furthermore, *in vivo* the protein has to be capable of binding to DNA without the presence of loose ends in close proximity. Therefore, the pore has to be opened upon formation of the protein–DNA complex, which has to occur at the interface between the arch and the FeS cluster domain.

The exact position of initial dsDNA unwinding often referred to as a ‘wedge’ like feature could not be localized so far. Although it is likely that it is located in the FeS domain, no residue in an exposed position to perform this task could be identified. Phe326 may not be sufficiently exposed to assume this role when compared with similar elements in other structures. However, rearrangements of the arch domain during DNA binding and unwinding are most likely to occur that could lead to differences in the exposure of certain residues, which could then fulfil the task of a wedge.

Class I and II variants are clearly separated by their location within the protein (see above). Interestingly class I variants display mainly impairment in ssDNA binding combined with decreased ATPase and helicase activity, whereas most class II variants display decreased ssDNA and elevated ATPase activity and a variety of helicase phenotypes. Based on the overall architecture of taXPD, it is tempting to speculate that class I variants negatively affect the translocation and ATPase activity of taXPD via impairing the interplay between the two motor domains. Class II variants, despite their impairment in binding or interaction with ssDNA, show mostly elevated ATPase activity, which could be due to a decoupling effect. Since the translocation of ssDNA between the two motor domains is most likely not affected by these variants the reduced interaction with ssDNA upstream of the motor domains could result in a reduction of pulling force on the substrate that results in elevated ssDNA-induced ATPase activity.

During eukaryotic NER, the helicase activity of XPD is exclusively responsible for unwinding the DNA around the damage (Coin *et al*, 2007). *In-vivo* and *in-vitro* data further suggest that XPD is involved in damage verification. It was shown that translocation of the yeast XPD homologue Rad3 is stalled at sites of damage *in vivo* (Naegeli *et al*, 1993). More recently, XPD from *Ferroplasma acidarmanus* was shown to be stalled at CPD lesions on the translocated strand *in vitro* (Mathieu *et al*, 2010). This stalling was accompanied by an elevated ATPase activity. A similar behaviour can be observed for UvrB, the prokaryotic NER protein responsible for damage verification (Wang *et al*, 2006).

In restriction protection assays, Mathieu *et al* (2010) demonstrated that the DNA was susceptible to cleavage 16 bases prior to the damage (5′–3′ direction) and 15 bases after the damage, whereas it was protected at a site close to the damage. In eukaryotic NER, the length of the excised

fragment on average is around 27 bp while displaying an asymmetrically excised damage that is always located 2–9 bases from the 3′-end (Wood, 1999, 2010). An extrapolation of these observations to our DNA-binding model suggests that the site of damage verification should be located in or close to the pore (Figure 5). The pore is placed asymmetrically with respect to the two motor domains and, more importantly, with respect to the translocated DNA strand and puts it closer to the 3′-end of the translocated DNA. This positioning would be in line with the observed restriction protection data of the CPD damage and, therefore, the pore represents a structural feature likely involved in damage verification. The pore size and diameter is not fixed but probably varies in response to ATP hydrolysis (Wolski *et al*, 2010). The above-described ‘hot spot’ and Phe326 are located in the pore in direct proximity to the DNA model and the FeS cluster and show helicase regulating phenotypes. This activity would also be necessary during the damage induced stalling of the helicase that could be regulated using the sites identified in this study.

In conclusion, our data provide the first detailed insights into the translocation mechanism of an SF2B helicase and reveal how polarity is achieved. Combined with the biochemical analysis of XPD important tasks of this helicase during NER can also be rationalized. We have generated loss and gain of function variants that pinpoint towards important functional regions of the protein and will provide the framework for further analyses.

Materials and methods

Mutagenesis, expression, and purification of taXPD

The plasmid used for expression and mutagenesis was described earlier (Wolski *et al*, 2008) using the alternative start codon resulting in a 19 amino-acid N-terminal extension. TaXPD mutants were generated using the Quick-Change site directed mutagenesis kit (Stratagene). The reactions were carried out as suggested by the manufacturer’s instructions. All mutants were verified by double-stranded sequencing. TaXPD WT and variants were expressed as N-terminally His-tagged proteins in *Escherichia coli* BL21-CodonPlus (DE3)-RIL cells (Stratagene) by induction with 0.1 mM isopropyl- β -thiogalactoside at 14°C for 18 h. The proteins were purified to homogeneity by metal affinity chromatography (Ni-NTA, Invitrogen) followed by size-exclusion chromatography (HiLoad 26/60 Superdex 200 prep grade; GE Healthcare) in 20 mM Tris (pH 8), 200 mM NaCl and 10 mM MgCl₂. The proteins were concentrated to 5 mg ml^{−1} based on a theoretical molar absorption coefficient of 65140 M^{−1} cm^{−1}. Expressed proteins were checked for sypro orange-based thermal unfolding (Ericsson *et al*, 2006) using a MX3005P real-time PCR cycler (Stratagene). Protein concentration was between 0.1 and 0.5 mg ml^{−1} and a final concentration 5 × sypro orange in 30 μ l 20 mM Tris (pH 8), 200 mM NaCl and 10 mM MgCl₂ was used.

Crystallization and structure solution

XPD was co-crystallized with a 22-mer oligonucleotide (5′-cccagtagcagccagtgccg-3′). The protein was incubated with the DNA in a 1:1.2 molar ratio at room temperature for 30 min. Crystals were grown by vapour diffusion in hanging drops containing equal volumes of a protein–DNA complex in 20 mM Tris (pH 8), 200 mM NaCl and 10 mM MgCl₂ and a reservoir solution consisting of 50 mM MES (pH 6.5), 10 mM MgSO₄ and 15% 2-Methyl-2, 4-pentanediol equilibrated against the reservoir solution at 20°C. Prior to data collection, the crystals were cryoprotected by sequential transfer into mother liquor containing increasing amounts of MPD at a final concentration of 30%. The crystals were flash cooled in liquid nitrogen and data collection was performed at 100 K. The data set was collected at beamline BM14 (ESRF) at a wavelength of 0.972 Å to a resolution of 2.2 Å. All data were indexed and processed using

Mosflm and Scala (Leslie, 1992, 2006; Evans, 2006). The crystals belong to space group $P6_5$ with unit cell dimensions of $a = b = 79.1$ Å and $c = 175.7$ Å. Initial phases were obtained by molecular replacement with 2VSF as a search model using the program PHASER (McCoy *et al*, 2007). The model was improved by alternating rounds of manual model building using Coot (Emsley and Cowtan, 2004) and maximum likelihood refinement using Refmac5 (Bailey, 1994). All non-water ligands were added at the end of refinement to minimize model bias and use the best electron density for interpretation. The Ramachandran analysis was performed using the program Rampage (Bailey, 1994).

ssDNA taXPD biolayer interferometry binding assay

Real-time binding assays between ssDNA (5'-gactagtactgttaccggctc catctctaccgcaatcaggccagctatgc-3') and purified taXPD wild type and variants were performed using biolayer interferometry on an Octet RED system (Fortebio). This system monitors interference of light reflected from the surface of a fibre optic sensor to measure the thickness of molecules bound to the sensor surface. 3'-Biotinylated DNA was obtained from Biomers and coupled to kinetic grade streptavidin biosensors (Fortebio) at a concentration of 100 nM. Sensors coated with ssDNA were allowed to bind taXPD in reaction buffer (20 mM Tris pH 8, 150 mM NaCl, 10 mM MgCl₂, 1 mM DTT and 1 mg ml⁻¹ BSA) at different taXPD concentrations ranging from 0.5 to 5 µM. Measurements were carried out in triplicates and with different protein batches. Binding kinetics were calculated using the Octet Data Analysis Software 6.3, with a 1:1 binding model to calculate the association rate constants. Binding affinities were calculated as the ratio of dissociation and association rate constants.

In-vitro ATPase assay

taXPD ATPase activity was measured with an *in-vitro* ATPase assay, in which ATP consumption is coupled to the oxidation of NADH via pyruvate kinase and lactate dehydrogenase activities. Activities were measured at 37°C in 50 ml solution, containing 1.5 U pyruvate kinase, 1.9 U lactate dehydrogenase, 2 mM phosphoenolpyruvate and 0.15 mM NADH, 10 mM KCl, 20 mM MgCl₂, 0.5 mM DTT, 0.5 mM CaCl₂ and 60 mM HEPES (pH 7.2). ssDNA (5'-gtcgtgagctctaga ctgcagttgagagcttgctaggacggatccctcagg-3') was added at a final concentration of 0.5 µM. The assay was carried out under saturating concentrations of ATP (2 mM), using XPD wild type and variants at a concentration range of 200–2000 nM. For catalytic measurements, the mix of all reagents, with the exception of ATP, was preincubated at 37°C until a stable base line was achieved. Enzyme catalysis was initiated by the addition of ATP. The activity profiles were measured at 340 nm using an Agilent 8453 diode array spectrophotometer. The initial velocities were recorded and ATP consumption was determined using the molar extinction coefficient of NADH. The measurements were carried out in triplicates and with different protein batches.

In-vitro helicase assay

Helicase activity was analysed utilizing a fluorescence-based helicase assay (Martinez-Senac and Webb, 2005). We used an open fork substrate with a cy3 label at the 3'-end of the translocated strand oligonucleotide (5'-agctaccatgctgcacgaattaagcaattcgtaatcat ggctcatagct-3'-cy3) and a dabcy1 modification on the 5'-end of the opposite strand (Dabcy1-5'-agctatgacctgattacgaattgcttgaatcctgac gaactgtag-3'). This resulted in a quenching of the cy3 fluorescence that is removed upon unwinding of the substrate. Assays were carried out in 20 mM Tris pH 8.5, 10 mM NaCl, 5 mM MgCl₂, 1 mM EDTA and 2 mM DTT. 800 nM of the protein was mixed with 40 nM open fork substrate and 800 nM capture oligonucleotide (5'-caactcgtaatcatggctc-3'). After a stable baseline was reached, the reaction was subsequently started with the addition of 500 µM ATP. In the absence of XPD, no significant increase in fluorescence after the addition of ATP could be observed. Kinetics were recorded with a Fluoromax4 fluorescence spectrometer (Horiba Jobin Yvon) and monitored for at least 10 min or until the reaction was completed where possible. Fluorescence was detected at an excitation wavelength of 550 nm (slit width 2 nm) and an emission wavelength of 570 nm (slit width 2 nm). Initial velocities were fitted with Origin8 and represent the averages of at least two different reactions and two independent protein batches. In order to verify the unwinding of the substrate helicase reactions were monitored on native polyacrylamide gels using a PHAROS FX scanner (Bio-Rad).

Accession codes

Coordinates and structure factors for the XPD–DNA complex have been deposited in the Protein Data Bank (4a15).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

Acknowledgements

We thank the staff of the European synchrotron radiation facility at beamline BM14 for technical support. This research was supported by grants through the Deutsche Forschungsgemeinschaft (KI-562/2 and Forschungszentrum FZ-82) to CK.

Author contributions: SCW, JK and CK conceived and designed the experiments. SCW, JK and GM performed the experiments and analysed the data. CK and JK wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Bailey S (1994) The CCP4 suite—programs for protein crystallography. *Acta Cryst D* **50**: 760–763
- Bergmann E, Egly JM (2001) Trichothiodystrophy, a transcription syndrome. *Trends Genet* **17**: 279–286
- Buttner K, Nehring S, Hopfner KP (2007) Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat Struct Mol Biol* **14**: 647–652
- Coin F, Oksenych V, Egly JM (2007) Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair. *Mol Cell* **26**: 245–256
- de Boer J, Hoeijmakers JH (2000) Nucleotide excision repair and human syndromes. *Carcinogenesis* **21**: 453–460
- Dubaele S, Proietti De Santis L, Bienstock RJ, Keriell A, Stefanini M, Van Houten B, Egly JM (2003) Basal transcription defect discriminates between xeroderma pigmentosum and trichothiodystrophy in XPD patients. *Mol Cell* **11**: 1635–1646
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**: 2126–2132
- Ericsson UB, Hallberg BM, Detitta GT, Dekker N, Nordlund P (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**: 289–298
- Evans E, Fellows J, Coffey A, Wood RD (1997a) Open complex formation around a lesion during nucleotide excision repair provides a structure for cleavage by human XPG protein. *EMBO J* **16**: 625–638
- Evans E, Moggs JG, Hwang JR, Egly J-M, Wood RD (1997b) Mechanism of open complex and dual incision formation by human nucleotide excision repair factors. *EMBO J* **16**: 6559–6573
- Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**: 72–82
- Fan L, Fuss JO, Cheng QJ, Arvai AS, Hammel M, Roberts VA, Cooper PK, Tainer JA (2008) XPD helicase structures and activities: insights into the cancer and aging phenotypes from XPD mutations. *Cell* **133**: 789–800
- Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T (2006) *DNA Repair and Mutagenesis*, 2nd edn, Washington, D.C.: ASM Press
- Gillet LC, Scharer OD (2006) Molecular mechanisms of mammalian global genome nucleotide excision repair. *Chem Rev* **106**: 253–276
- Goosen N, Moolenaar GF (2001) Role of ATP hydrolysis by UvrA and UvrB during nucleotide excision repair. *Res Microbiol* **152**: 401–409

- Honda M, Park J, Pugh RA, Ha T, Spies M (2009) Single-molecule analysis reveals differential effect of ssDNA-binding proteins on DNA translocation by XPD helicase. *Mol Cell* **35**: 694–703
- Kim JL, Morgenstern KA, Griffith JP, Dwyer MD, Thomson JA, Murcko MA, Lin C, Caron PR (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. *Structure* **6**: 89–100
- Leslie AG (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr* **62**: 48–57
- Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **35**: 27–33
- Liu H, Rudolf J, Johnson KA, McMahon SA, Oke M, Carter L, McRobbie AM, Brown SE, Naismith JH, White MF (2008) Structure of the DNA repair helicase XPD. *Cell* **133**: 801–812
- Martinez-Senac MM, Webb MR (2005) Mechanism of translocation and kinetics of DNA unwinding by the helicase RecG. *Biochemistry* **44**: 16967–16976
- Mathieu N, Kaczmarek N, Naegeli H (2010) Strand- and site-specific DNA lesion demarcation by the xeroderma pigmentosum group D helicase. *Proc Natl Acad Sci USA* **107**: 17545–17550
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**: 658–674
- Min JH, Pavletich NP (2007) Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* **449**: 570–575
- Naegeli H, Bardwell L, Friedberg EC (1993) Inhibition of Rad3 DNA helicase activity by DNA adducts and abasic sites: implications for the role of a DNA helicase in damage-specific incision of DNA. *Biochemistry* **32**: 613–621
- Pugh RA, Honda M, Leesley H, Thomas A, Lin Y, Nilges MJ, Cann IK, Spies M (2008) The iron-containing domain is essential in Rad3 helicases for coupling of ATP hydrolysis to DNA translocation and for targeting the helicase to the single-stranded DNA-double-stranded DNA junction. *J Biol Chem* **283**: 1732–1743
- Pugh RA, Wu CG, Spies M (2012) Regulation of translocation polarity by helicase domain 1 in SF2B helicases. *EMBO J* **31**: 503–514
- Reynolds C, Damerell D, Jones S (2009) ProtorP: a protein-protein interaction analysis server. *Bioinformatics* **25**: 413–414
- Rudolf J, Makrantonis V, Ingledew WJ, Stark MJ, White MF (2006) The DNA repair helicases XPD and FancJ have essential iron-sulfur domains. *Mol Cell* **23**: 801–808
- Rudolf J, Rouillon C, Schwarz-Linek U, White MF (2010) The helicase XPD unwinds bubble structures and is not stalled by DNA lesions removed by the nucleotide excision repair pathway. *Nucleic Acids Res* **38**: 931–941
- Saikrishnan K, Powell B, Cook NJ, Webb MR, Wigley DB (2009) Mechanistic basis of 5′-3′ translocation in SF1B helicases. *Cell* **137**: 849–859
- Sancar A (1994) Mechanisms of DNA excision repair. *Science* **266**: 1954–1956
- Sancar A (1996) DNA excision repair. *Annu Rev Biochem* **65**: 43–81
- Singleton MR, Dillingham MS, Wigley DB (2007) Structures and mechanism of helicases and nucleic acid. *Annu Rev Biochem* **76**: 23–50
- Sugasawa K, Hanaoka F (2007) Sensing of DNA damage by XPC/Rad4: one protein for many lesions. *Nat Struct Mol Biol* **14**: 887–888
- Tapias A, Auriol J, Forget D, Enzlin JH, Scharer OD, Coin F, Coulombe B, Egly JM (2004) Ordered conformational changes in damaged DNA induced by nucleotide excision repair factors. *J Biol Chem* **279**: 19074–19083
- van Gool AJ, van der Horst GT, Citterio E, Hoeijmakers JH (1997) Cockayne syndrome: defective repair of transcription? *EMBO J* **16**: 4155–4162
- Van Houten B (1990) Nucleotide excision repair in Escherichia coli. *Microbiol Rev* **54**: 18–51
- Velankar SS, Soutanas P, Dillingham MS, Subramanya HS, Wigley DB (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**: 75–84
- Vermeulen W, de Boer J, Citterio E, van Gool AJ, van der Horst GT, Jaspers NG, de Laat WL, Sijbers AM, van der Spek PJ, Sugawara K, Weeda G, Winkler GS, Bootsma D, Egly JM, Hoeijmakers JH (1997) Mammalian nucleotide excision repair and syndromes. *Biochem Soc Trans* **25**: 309–315
- Wang H, DellaVecchia MJ, Skorvaga M, Croteau DL, Erie DA, Van Houten B (2006) UvrB domain 4, an autoinhibitory gate for regulation of DNA binding and ATPase activity. *J Biol Chem* **281**: 15227–15237
- Wolski SC, Kuper J, Hanzelmann P, Truglio JJ, Croteau DL, Van Houten B, Kisker C (2008) Crystal structure of the FeS cluster-containing nucleotide excision repair helicase XPD. *PLoS Biol* **6**: e149
- Wolski SC, Kuper J, Kisker C (2010) The XPD helicase: XPanDing archaeal XPD structures to get a grip on human DNA repair. *Biol Chem* **391**: 761–765
- Wood RD (1999) DNA damage recognition during nucleotide excision repair in mammalian cells. *Biochimie* **81**: 39–44
- Wood RD (2010) Mammalian nucleotide excision repair proteins and interstrand crosslink repair. *Environ Mol Mutagen* **51**: 520–526
- Zurita M, Merino C (2003) The transcriptional complexity of the TFIIH complex. *Trends Genet* **19**: 578–584