XPD Helicase Structures and Activities: Insights into the Cancer and Aging Phenotypes from XPD Mutations

Li Fan[1,5], Jill O. Fuss[2,5], Quen J. Cheng[2], Andrew S. Arvai[1], Michal Hammel[3], Victoria A. Roberts[4], Priscilla K. Cooper[2] and John A. Tainer[1,2,*]

[1]Department of Molecular Biology, Skaggs Institute of Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037, USA
[2]Life Sciences Division, Department of Genomic Stability
[3]Physical Biosciences Division
Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[4]San Diego Supercomputer Center, University of California at San Diego, La Jolla, CA 92037
[5]These authors contributed equally to this work.
*Correspondence: jat@scripps.edu
DOI 10.1016/j.cell.2008.04.030

Cell 133, 789-800, May 30, 2008

SUMMARY

Mutations in XPD helicase, required for nucleotide excision repair (NER) as part of the transcription/repair complex TFIIH, cause three distinct phenotypes: cancer-prone xeroderma pigmentosum (XP), or aging disorders Cockayne syndrome (CS), and trichothiodystrophy (TTD). To clarify molecular differences underlying these diseases, we determined crystal structures of the XPD catalytic core from *Sulfolobus acidocaldarius* and measured mutant enzyme activities. Substrate-binding grooves separate adjacent Rad51/RecA-like helicase domains (HD1, HD2) and an arch formed by 4FeS and Arch domains. XP mutations map along the HD1 ATP-binding edge and HD2 DNA-binding channel and impair helicase activity essential for NER. XP/CS mutations both impair helicase activity and likely affect HD2 functional movement. TTD mutants lose or retain helicase activity but map to sites in all four domains expected to cause framework defects impacting TFIIH integrity. These results provide a foundation for understanding disease consequences of mutations in XPD and related 4Fe-4S helicases including FancJ.

INTRODUCTION

Inherited genetic defects in XPD helicase provide an opportunity to understand how molecular level DNA repair defects impact cancer and aging phenotypes of the whole organism (de Boer et al., 2002; Hoeijmakers, 2001; Lehmann, 2001). XPD is a 5'-3' SF2 family helicase (Singleton et al., 2007) that opens damaged DNA for bulky lesion repair in NER. XPD is a core component of the transcription and repair factor TFIIH (Laine et al., 2006; Schaeffer et al., 1993; Sung et al., 1993; Tirode et al., 1999). Mutations in the human XPD helicase gene (*ERCC2*) are mainly single residue changes and sometimes at adjacent residues; yet, they cause three strikingly different genetic disorders: XP, CS combined with XP (XP/CS), and TTD (Lehmann, 2001; Ludovic et al., 2006). Although all three diseases share a photosensitivity phenotype, they differ greatly in their predispositions to cancer or accelerated aging. XP patients show several 1000-fold increase in skin cancer, whereas neither CS nor TTD patients show an increase in the cancer incidence despite sun sensitivity. Furthermore, both CS and TTD are premature aging diseases plus developmental disorders, with CS patients being more severely affected and exhibiting severe mental retardation from birth. Despite extensive biochemical and cell biological analysis, key questions remain concerning how point mutations in adjacent residues in a single enzyme can give rise to such different disease phenotypes (Lehmann 2001).

XPD helicase activity is essential for NER but dispensable for transcription (Coin et al., 2007; Laine et al., 2006). XPD protein-protein interactions are critical for both helicase activity and stability of the TFIIH complex (Dubaele et al. 2003). Mutations in the XPD C terminus that cause TTD weaken binding to TFIIH subunit p44 and reduce DNA repair activity (Coin et al. 2007). XPD also interacts

with XPG, and loss of XPG destabilizes TFIIH and its association with XPD (Ito et al., 2007). Nuclear receptor transactivations are inhibited by XPD mutations that reduce p44 interactions (Dubaele et al., 2003) and by XPG loss (Ito et al., 2007) probably due to decreased TFIIH stability. TFIIH from TTD, but not from XP patients, has basal transcription defects in vitro as well as reduced in vivo TFIIH concentrations (Dubaele et al. 2003), suggesting XPD's role in TFIIH stability is impacted by TTD-causing mutations. Cellular and biochemical analyses provide detailed information on XPD activities, patient mutations, and TFIIH stability (Bootsma and Hoeijmakers, 1993; Dubaele et al., 2003; Winkler et al., 2000). However, an understanding of the molecular basis for these effects has proven elusive without combined structural and biochemical analyses of the XPD helicase.

Recent biochemical characterization of the *Sulfolobus acidocaldarius* XPD homolog (SaXPD) and yeast genetic analyses uncovered a unique Fe-S cluster domain conserved among related SF2 helicases important for genomic stability including Chl1, Rtel1, and FancJ (also known as BACH1 and BRIP1), which is defective in Fanconi anemia (Rudolf et al. 2006). These studies showed that these XPD-like helicases require a novel Fe-S cluster region inserted between the Walker A and Walker B motifs, suggesting that the Fe-S region conformation may be controlled by ATP binding and hydrolysis, as an analogously placed insertion is coupled to the ATP binding state in the Rad50 ABC ATPase (Hopfner et al., 2000). Furthermore, recent studies on the *Ferroplasma acidarmanus* XPD protein revealed that the intact cluster acts in the correct orientation of the XPD protein at the ssDNA-dsDNA junction (Pugh et al., 2008). This Fe-S region is biologically critical as a mutation in the XPD Fe-S region causes TTD (Schumacher et al., 2008), and a FancJ mutation in this region causes severe clinical symptoms of Fanconi anemia and a predisposition to early-onset breast cancer (Cantor et al., 2004; Levran et al., 2005). Although unusual in nuclear proteins, Fe-S clusters were discovered to act in DNA binding for DNA repair glycosylases, as originally shown for endonuclease III (Thayer et al., 1995). Fe-S clusters may also act as electron- and oxygen-responsive molecular switches on DNA (Boal et al., 2005; Outten, 2007).

To provide a molecular foundation to address current paradoxes regarding XPD activities and the role of XPD mutations in causing distinct human diseases, we determined structures of SaXPD with and without the Fe-S cluster and analyzed the activities of mutations at conserved sites that cause XP, XP/CS, and TTD diseases. The XPD four-domain fold and architecture, which is substantially different than expected (Bienstock et al., 2003), reveal functional roles for the 4Fe-4S cluster and XPD mutation sites relevant to disease-causing defects in XPD as well as the related 4Fe-4S helicase FancJ. More generally, the relationships of XPD structures and activities characterized here support a unified understanding of XPD activities and interactions in cell biology.

RESULTS

Crystal Structure Determination

To understand the XPD structure, we expressed, purified, and analyzed SaXPD. Sequence alignments show SaXPD represents the XPD catalytic core (XPDcc) with a 4Fe-4S cluster and all the helicase motifs conserved with the human XPD (Figures 1A and S1). The human XPD C-terminal extension, missing in SaXPD, is predicted to be disordered by PONDR (Romero et al., 2001), and may act in TFIIH interactions (Figure 1A). To determine the XPDcc structure and 4Fe-4S cluster role unique to XPD and related helicases such as FancJ (Rudolf et al., 2006), we therefore crystallized SaXPD and solved crystal structures with and without the bound 4Fe-4S cluster.

SaXPD crystallized in space group $P2_12_12_1$ with one molecule per asymmetric unit (Table 1). We solved the SaXPD crystal structure by multiwavelength anomalous diffraction (MAD) with Se-Met substituted protein expressed in bacteria, and refined the structure to 2 Å resolution (R = 22.3%, Rfree = 26.0%). The high-quality composite omit electron density maps allowed us to fit and refine all amino acid residues (1–551). The structure extends results on SaXPD sequence and mutagenesis (Rudolf et al., 2006) by characterizing the XPDcc with all conserved helicase motifs and the 4Fe-4S cluster.

XPDcc Domain Structure and Architecture

The SaXPD structure shows that the XPD catalytic core is comprised of four domains: two Rad51/RecA-like domains (HD1 and HD2) with two additional domains (the 4FeS and Arch domains) inserted into HD1 (Figures 1, S1, and S2). These four XPDcc domains contain 22 out of the 26 known disease causing point mutation sites; only 4 of the XPD sites are positioned in the C-terminal extension from HD2 (Figure 1A). HD1 (175 residues: 1–81, 149–192, and 304–353) and HD2 (198 residues: 354–551) share the α-β fold with a central seven stranded parallel β sheet flanked by α helices (Figure S2) that resemble the ATPase domain in Rad51 and RecA (Shin et al., 2003). The 4FeS (67 residues: 82–148) and Arch (111 residues: 193–303) domains are inserted between adjacent β strands of the central β sheet of HD1, making them closely connected to HD1, but relatively independent of HD2 (Figure S2). In contrast, known helicase structures typically have domains inserted into HD2 (Singleton et al., 2007).

Alignment of the human XPD sequence with the SaXPD sequence, secondary structure, and helicase motifs validates and informs the conserved nature of the XPD fold and domain structure (Figure S1). The helicase motifs are conserved from SaXPD to human XPD. HD1 helicase motif I (residues 31–36) and 1a (50–60) occur before the 4FeS domain insertion followed by HD1 helicase motif II (177–186). The four Cys residues (88, 102, 105, and 137), which act as ligands to the 4Fe-4S cluster, are invariant with human XPD (Figure S1). The Arch domain insertion into HD1 occurs between helicase motifs II and III (317– 327) (Figures S1 and S2). The placement of the four HD1 helicase motifs

results in intimate connections between the ATP-binding and hydrolysis state of HD1 and the conformations of the 4FeS and Arch domains. The C-terminal HD2 contains the remaining three conserved helicase motifs IV (394–408), V (439–455), and VI (501–517). The composite ATP-binding site (motifs I, II, V, and VI) comes together at the HD1-HD2 interface cleft, as expected for the inchworm helicase mechanism whereby ATP binding and hydrolysis drives ssDNA translocation responsible for helicase activity (Singleton et al., 2007). The SaXPD C terminus ends at the outside edge of HD2, suggesting the C-terminal region is an extension protruding from the XPDcc (Figures 1B and 1C), consistent with its role in protein interactions in other SF2 family helicases (Singleton et al., 2007).

The HD1, 4FeS, and Arch domains lie in the same plane and thereby form a shallow pentagonal box shape (60 x 60 x 25 Å). The ellipsoidal HD2 is covalently connected to HD1 (Figures 1B, 1C, and S2) and packs primarily against HD1 (over a 30 x 25 Å interface) and against the Arch domain (15 x 15 Å interface) on one side of this box to create long deep grooves between HD2 and the remaining three domains.

HD2 protrudes about 30 Å from the box to form prominent channels 25 Å long for ATP binding between HD2 and HD1, and 50 Å long for ssDNA binding between HD2 and the Arch and 4FeS domains (Figure 1B). In contrast, the side of the box facing away from HD2 is relatively flat except for a 20 Å diameter depression at the junction of the Arch and 4FeS domains suitable for binding one end of a dsDNA bubble.

The 50-Å-long channel extending along the helicase motifs in HD2 is gated at both ends by the arch and HD2 gateways. The arch gateway is located under the arch formed between the Arch and 4FeS domains. The HD2 gateway lies between the Arch and HD2, and is formed partly by anHD2helix-loop-helix insertion that extends outward to pack around the Arch domain. Each gateway has dimensions of about 10 x 10 x 10 Å, providing a possible means to sense bulky DNA damage.

The SaXPD four-domain fold, domain insertions, relative domain orientations, and overall architecture are different from known helicase structures. The most similar existing helicase structure is UvrD, a helicase acting in the bacterial NER pathway (Lee and Yang, 2006). The previously reported XPD structural model, which was built with rigorous comparative molecular modeling and site-directed mutagenesis of the bacterial repair protein UvrB (Bienstock et al., 2003), is substantially different from the experimentally defined XPDcc (Figure S3).

Arch Domain Structure

The Arch domain, named by its arch-shaped conformation, is a three-stranded antiparallel β sheet with two α helix pairs, one of which has an extended loop interacting with a loop from the 4FeS domain. The Arch domain inserts into the HD1 sequence immediately after helicase motif II and

rejoins HD1 in the α helix preceding helicase motif III (Figures S1 and S2). The β sheet bridges between the HD1 fold and the Arch domain α helices, which give the domain its arched shape. The Arch domain is strategically positioned via its covalent connections to HD1 to join the ATP-binding helicase domain to the far edge of the HD2 motor helicase domain and also to form a small interface (about 15 x 15 Å) with the 4FeS domain to make an enclosed tunnel. This tunnel juxtaposes functionally conserved, charged residues from the Arch domain (R194, R259, and R278) with functionally conserved, charged, and aromatic side chains positioned by the 4FeS domain (K84, K103, Y139, and Y140), consistent with a ssDNA-binding role for the tunnel and Arch-4FeS domain interface. At the opposite face, the junction of the Arch domain with the 4FeS domain forms half of the 20 Å diameter depression on the otherwise relatively flat back face of the box. One consequence of the narrow depth and flat back of the arch is that only about six ssDNA bases would be buried from access to the approaching NER nuclease XPG, assuming XPG interactions with DNA resemble those for Fen-1 (Chapados et al., 2004). In such a situation, the DNA damage could still be accessible for XPA binding. Thus, this architecture could be relevant to damage access during NER.

4FeS Cluster Domain Structure

To characterize the native XPD 4Fe-4S cluster without oxidation, we grew crystals anaerobically and cryocooled them in liquid nitrogen for X-ray diffraction data collection. The experimental electron density for the SaXPD crystals grown anaerobically shows that the 67-residue 4FeS domain contains an 4Fe-4S cluster coordinated by four cysteine ligands (Cys88, Cys102, Cys105, and Cys137) (Figures 1D and S4). All four Fe ions are present based upon their five sigma peaks in unbiased omit maps (Figure S4), so we name this domain the 4FeS domain and the cluster the 4Fe-4S cluster. This 4Fe-4S cluster is sensitive to oxidation, and this redox sensitivity is increased by DNA substrates (data not shown), explaining the previous characterization of SaXPD with a 3Fe cluster (Rudolf et al., 2006) and supporting a potential functional role for cluster oxidation in XPD functions. The existence of an oxygen-sensitive 4Fe-4S cluster also implies that previous biochemical XPD characterizations may reflect a combination of direct mutation effects and indirect effects complicated by the instability of the 4Fe-4S cluster and its associated domain.

The 4FeS domain is composed of four helices connected by loops and stabilized by the interactions of four Cys ligands to the Fe ions. The first cysteine ligand (SaXPD Cys88) is located at the C terminus of a one-turn helix connected to HD1. The 13 residues between Fe ion ligands Cys88 and Cys102 form a loop (here named the Fe cluster loop [FCL]) with a one-turn helix near the middle. Fe ion ligands Cys102 and Cys105 reside at the N and C termini of another one-turn helix, respectively. Cys105 is connected by an extended loop and a helix to Cys137, which is located at the N terminus of a long helix connected back to HD1. The 4FeS domain structure characterized here appears characteristic of a helicase damage response family including FancJ. In our structure, two disease-

causing mutation sites both cause similar defects in the 4Fe-4S cluster reinforcing its functional significance. The A349P mutation in FancJ, which can cause severe Fanconi anemia clinical symptoms (Levran et al., 2005), would disrupt the hydrogen bond between the main-chain nitrogen and Fe ion ligand Cys137. Similarly, the XPD TTD mutation R112H (SaXPD K84) disrupts the charged side-chain hydrogen bond to Fe ion ligand Cys102 (Figure 1D).

To test the structural importance of the 4Fe-4S cluster, we removed the cluster by soaking crystals in a cyrosolution containing ferricyanide under aerobic conditions. Although apo-XPD crystals diffracted to lower resolutions, we were able to solve and refine the apo-SaXPD structure to 3.0 Å resolution (Table 1).

Loss of the Fe-S cluster induced four significant structural changes (Figure 1C). First, the average overall B factor increased from 41 to 107 $Å^2$ (Table 1), suggesting the 4Fe-4S cluster has a role in maintaining the overall stability of the enzyme. Second, the 4FeS domain is disordered except for parts directly connected to HD1. Third, the Arch domain loop (residues 265–270) that forms an interface with the FCL is disordered showing the importance of the 4FeS domain in maintaining the arch and arch gateway. Fourth, the first eight residues at the N terminus also become disordered, revealing an intimate connection of the 4FeS domain conformation with HD1.

In the 4Fe-4S bound SaXPD structure, the interface of the Arch and 4FeS domains involves the interaction of the Arch domain loop with the FCL. This interface primarily involves polar hydrogen bonding and salt-bridge interactions from main-chain and charged side-chain atoms, suggesting it may have functionally important flexibility. Loss of the 4Fe-4S cluster did not change the overall relative orientations among HD1, Arch, and HD2, but resulted in the rotational opening of the distal helical hairpin in the Arch domain. The 4Fe-4S cluster therefore is critical to form the closed interface with the Arch domain and the FCL. The 4Fe-4S cluster appears critical to SaXPD helicase activity, consistent with our results on mutations disrupting the cluster (see below). These structural results suggest that the channel under the arch formed by HD1 and the Arch and 4FeS domains plays an important role in forming a passageway for ssDNA translocation during XPD helicase unwinding (see below). The location, redox sensitivity, and biological roles of XPD in NER are consistent with key roles for Fe-S clusters proposed in DNA damage sensing (Yavin et al., 2006). These experimental results on the XPD 4FeS domain have implications for a possible role of electron transfer along DNA in NER as well as for the function of related helicases including FancJ.

XPDcc Molecular Surface, Helicase Motifs, and DNA Binding

To analyze functional implications of the XPDcc structure, we examined the location of the exposed conserved molecular surface, clusters of positively charged residues along the channel, and the HD2 helicase motifs for ssDNA interactions, and compared these to existing helicase-DNA cocrystal

structures. The low sequence conservation of SaXPD with the human enzyme (22% identity and 39% similarity) provides an advantage for identifying functionally relevant elements, as only regions important to the XPDcc structure and function are likely to be conserved (Figure S1). Thus, the structure-based alignment and position of the functional motifs show that conserved residues cluster in the ssDNA- and ATP-binding grooves between HD1 and HD2, surface areas near the 4Fe-4S cluster, and near the interface between the Arch and HD2 (Figure 2A).

The electrostatic potential mapped onto the molecular surface shows clusters of positive charge near both ends of the gated channel (Figures 2B and 2D), but negative charges from Asp and Glu side chains at the channel rim overlay the channel where they would restrict a negatively charged ssDNA phosphate backbone from moving upward out of the channel. The surprising channel lid formed from negative charges is reminiscent of the use of repulsive charge interactions in trains, where electromagnetic force is used to suspend, guide, and propel the train with possible analogies to DNA. The structure also implies that the overall channel appears likely to contribute to ssDNA binding with no single region contributing most of the binding energy, consistent with the functional need for ssDNA to move.

To objectively and comprehensively evaluate all possible XPD-DNA complexes computationally, we used a complete six-dimensional systematic computational search. In all of the analyzed energetically favorable solutions, the dsDNA is located near both the HD2 and arch gateways. Most energetically favorable dockings place dsDNA at the HD2-Arch domain junction, where the gate formed by the helix-loop-helix protrusion on HD2 would be an appropriate wedge to hold open dsDNA. Some favorable solutions place the dsDNA in the depression at the arch gateway on the flat side of the HD1, Arch, and 4FeS domains. Based upon these objective computational searches, we therefore positioned the two dsDNA ends of the in vivo bubble substrates for XPD at these two sites (Figures 2B–2D). All of these top docking solutions of a dsDNA are located at the ends of the deep active channel groove, identified from the helicase motifs. Interestingly, they overlap well with crystallographically bound glycerol, isopropanol, and citrate ions that appear to mimic DNA backbone phosphate and sugar moieties (see below). Our placement of the dsDNA near the C-terminal extension is also consistent with previous studies showing that human XPD helicase activity is regulated by interactions of this region with other TFIIH proteins such as p44 (Dubaele et al., 2003).

These dsDNA interactions identified by computational docking could be reasonably connected via an ssDNA segment whose position is suggested by the SF2 helicase Hel308 structure (Buttner et al., 2007) and by matching the dsDNA relative to the HD2 helicase motifs. In this model, the placement of helicase motif IV and motif Ia suggests the ssDNA would be pushed away from the restrictive tunnel between the Arch and 4FeS domains by ATP binding (Figure 2C). The resulting distance between the two dsDNA ends suggests that over 10 nucleotides of ssDNA are needed to span the channel

between the two gates. Knowing that substantial domain movements can occur in helicases upon DNA binding (Fan et al., 2006; Lee and Yang, 2006), we avoided computationally optimizing a specific DNA bound conformation. However, the overall observed XPDcc features support a specific working model for DNA interactions for comparisons to the observed effects of disease-causing mutations.

Structural and Functional Placement of XPDcc Mutation Sites

Based upon the XPDcc structure and modeled DNA interactions, we examined the apparent functional roles of representative XP, XP/CS, and TTD mutation sites that appeared to be functionally conserved in the SaXPD structure (Figures 1A and S1). The SaXPD structure contains 22 of the 26 known XPD point mutation sites associated with human disease. At the HD2 gate to the active site channel, residues R531, R373, and K369 protrude into the computationally predicted DNA and furthermore interact with bound citrate, isopropanol, and glycerol from the crystallization buffer in apparent mimicry with DNA components (Figure 3A). R373 (R511, the corresponding HsXPD mutation site, as noted in parentheses from now on) and R531 (R683) are XP mutation sites, yet the adjacent side chain K369 is not a known mutation site, so this channel site tests biochemical impacts (see below). Interestingly, XP site D529N (D681N), which would not seem to impact an obvious DNA binding residue, would remove the charged side-chain interaction with R531 that positions the Arg at the proposed DNA-binding site. Thus, these residues appear to represent DNA-binding site changes that will impact helicase activity by altering XPD binding to DNA or ATP.

Within the active site channel extending toward the tunnel under the Arch and 4FeS domains, XP mutation sites T56 (T76), S402 (S541), Y403 (Y542), and K446 (R601) line one rim of the active site channel, where they are positioned to interact with ssDNA. In contrast, the XP/CS mutation site G447 (G602D), which is adjacent to XP site K446 (R601), will not only impact DNA binding by placing a negative charge into the channel, but also disrupt the main-chain turn structure by replacing Gly with Asp (Figure 3B). Thus, the Asp at this strategic site, which is a joint between two tight turns at the channel rim, restricts functional flexibility compared to Gly, which has great conformational freedom due to its absence of any side chain. This Gly site XP/CS change is distinguished from the XP mutation sites as it impacts the flexibility of HD2 as well as the DNA-binding channel.

Other XP and XP/CS mutation sites are associated with the ATP-binding channel formed between HD1 and HD2. XP/CS site R514 (R666) forms a charged side-chain hydrogen bond to backbone carbonyl oxygen that allows functionally important conformational switching at the HD2 interface with HD1 (Figure 3C). The XP/CS Arg-to-Trp mutation reduces the flexibility at this site because the type of switching observed for Arg, as for example at ATP sites in the ATPase GspE (Yamagata and Tainer, 2007), is restricted for Trp by both its fewer single bond rotations and its larger ring compared to Arg. The XP/CS mutation G34R (G47R) replaces a flexible Gly in the Walker A motif (helicase motif I) with a bulky Arg. This mutation replaces the open ATP binding site with an Arg that permanently fills

this site and can form hydrogen bonds to the adjacent negatively charged side chains. The attached Arg thus replaces the ATP, which would be hydrolyzed and exchanged during helicase cycles, as modeled after structures of UvrD helicase (Lee and Yang, 2006). Therefore, G34R (G47R) will rigidify the HD2-HD1 interface at the ATP site. In contrast, in this same region the XP mutation D180N (D234N) breaks a salt bridge from the Asp to conserved helicase motif I K35 (K48) to impact ATP binding and helicase activity (Figure 3C), without restricting functionally important conformational changes.

At the bottom of HD2, the XP/CS C523R mutation (G675R) replaces a Cys (or Gly) with a more rigid and bulky Arg to greatly reduce the conformational flexibility of HD2 (Figure 3D). In sharp contrast, TTD mutation K438P (R592P) removes two charged side-chain hydrogen bonds from the Lys (or Arg) to main-chain carbonyls of residues 389 and 392, impacting HD2 structural integrity by increasing framework flexibility while also reducing framework stability. TTD mutation D521G (D673G) similarly removes a salt bridge to Arg456 plus charged side-chain hydrogen bonds to main chain amino nitrogens of residues 517 and 518. These lost interactions almost certainly cause substantial increases in framework flexibility that decreases structural integrity in this TTD mutation. The TTD site at K84 (R112) in the 4FeS domain should reduce framework stability as it removes the charged hydrogen bond to the Fe ion ligand Cys102 (Figure 1D). Although we cannot predict their precise molecular defect, additional TTD mutations that are present in the human C-terminal extension, which is missing in SaXPD, are predicted to be accessible for protein interactions within TFIIH.

Taken together, the structural analyses of the XP, XP/CS, and TTD mutations mapped to identical and functionally equivalent residues in SaXPD prompt specific proposals for the defects associated with each type of mutation at the XPDcc level. In general, these mutations appear structurally positioned to fall into three classes by their predicted impacts on binding, conformation, and framework stability: (1) XP mutations should decrease DNA or ATP binding to impact helicase activity; (2) XP/CS mutations should, besides loss of helicase activity, reduce functional flexibility to impact the conformational state of the HD1-HD2 interface; and (3) TTD mutations should damage framework interactions to reduce the structural integrity of XPD and thus its protein interactions. These analyses therefore suggest there should be differential patterns of impacts on ATPase, helicase, and DNA binding measurements for the three classes of XPD mutations.

Mutations in SaXPD Impair ATPase and Helicase Activity

To biochemically test structurally implied roles of XPD mutation sites along with control mutations, we identified and mutated 15 sites in SaXPD corresponding to the best-conserved human disease mutations (Taylor et al., 1997), plus two 4Fe-4S cluster-forming cysteines, and K369Q, a test of the proposed DNA-binding channel (Figure 3A). Our results probed mutations representing five XP, four XP/CS, and three TTD sites plus the channel and 4Fe-4S control mutations.

We found that ssDNA stimulated ATPase activity more than duplex, overhang, or bubble substrates (data not shown), consistent with activities of other helicases (Singleton et al., 2007). Wild-type SaXPD had an ATP-hydrolysis rate of 0.55 mol ATP per second per mol XPD with ssDNA. The majority of our mutations impacted ATP hydrolysis (Figure 4), especially G34R (G47R) (motif I) and R514W (R666W), which totally lacked ATPase activity. In addition, D180N (D234N), G447D (G602D), R531W (R683W), and C102S retained less than 20% of wildtype level ATP hydrolysis.

Helicase assays were performed on a 50-overhang substrate and yielded a wild-type rate of 2.22 base pairs per min per XPD molecule. Most mutations in SaXPD impacted helicase activity more severely than ATPase (Figure 4). In contrast to all of the tested XP and CS mutant enzymes, TTD mutant D521G (D673G) or K438P (R592P) retained over 20% helicase activity, supporting the model that TTD mutations result in TFIIH destabilization rather than a catalytic defect. Consistent with this model, the almost complete loss of helicase activity in the K84H (R112H) and 4Fe-4S cysteine mutations is likely caused by a gross destabilization of the 4FeS domain, as seen in our ferricyanide-oxidized apo structure (Figure 1C).

To test structurally implied DNA-binding sites, we examined the ssDNA-binding activity of the mutant enzymes by fluorescence anisotropy. As expected from the structural analyses suggesting a long binding channel, single-site mutations did not cause a dramatic loss of ssDNA binding in any of the mutant enzymes tested. The most striking decreases in ssDNA binding occurred for TTD mutant K84H (R112H), supporting an important role of the 4FeS cluster domain in binding ssDNA as proposed (Figure 2). Consistent with these ssDNA-binding results, the chemical oxidation of the cluster resulted in a rapid loss of the helicase activity and a more minor reduction in the ATPase activity (Figure S6), consistent with the apo-XPD structure suggesting that complete loss of the cluster can impact the integrity of HD1.

At the base of the channel below the arch gateway, XP mutant T56A (T76A) retained 81% ssDNA-binding activity, suggesting it is involved but not critical for binding, as expected. At the other end of this channel at the HD2 gateway, XP mutant K446L (R601L/W) and our channel-testing mutant K369Q also retained 76% of wild-type DNA binding. Moreover, the XP/CS mutant G447D (G602D), predicted to place a negative charge in the channel, also showed a significant binding drop to 68% of the wild-type levels (Figure 4; Table 2). All of the observed ssDNA-binding changes are consistent with the channel-exposed residues acting in ssDNA binding. On the other hand, not all XP/CS mutants inhibit ssDNA binding, as evidenced by the marked increase in binding of C523R (G675R) and G34R (G47R). As G34R (G47R) is at the ATP-binding site and not associated directly with DNA binding, the increased DNA binding seen in two of the four XP/CS mutants supports our proposal that XP/CS mutants create conformationally restricted XPDcc. Such conformational restriction is predicted

to allow tighter DNA binding as less interaction energy is channeled into opening dsDNA and moving the ssDNA along the channel.

Elegant biochemical characterizations of tested human XPD mutations (Dubaele et al., 2003) are in striking agreement with our SaXPD results. Mutations in human XPD corresponding to G34R (G47R), T56A (T76A), K84H (R112H), D180N (D234N), G447D (G602D), and R531W (R683W) all show greatly reduced helicase activity,matching the results observed with SaXPD mutations (Table 2; Figure 4). Moderate helicase activity (14%–23%) remained for mutations at sites in human XPD corresponding to Y403C (Y542C), K446L (R601L), D521G (D673G), and C523R (G675R), as seen for SaXPD (Table 2). High helicase activity in both human and SaXPD was seen for the SaXPD site K438P (R592P).

These results appear to corroborate the structural and functional models of the three distinct XPD-related diseases. XP mutations primarily impair helicase activity, as none of these mutations display helicase activity above 20% of wild-type (Table 2). Because ATPase activity is retained in some of these mutants, we posit that XP mutations affect the ability of XPD helicase to translocate along the DNA. Likewise, XP/CS mutations also severely impact helicase activity and, based on their crucial location at the HD1-HD2 interface, will prevent functionally important conformational changes. This reduction of conformational flexibility should reasonably impact both helicase activity and critical protein-protein interactions of XPD as part of the TFIIH complex, resulting in its more severe phenotype. In contrast, TTD mutations do not necessarily lose helicase activity. Instead, these mutations are predicted to disrupt framework stability, resulting in reduced stability of XPD interactions with its protein partners. The 4Fe-4S cluster mutations, C88S and C102S, which also destabilize tertiary structure, completely abolish helicase activity, supporting the hypothesis that the cluster is essential for XPD catalytic activity (Rudolf et al., 2006). The patterns of residue roles suggested by our combined structural and mutational analyses are clarified by mapping the three classes of XPD mutation sites onto the Cα positions in the SaXPD structure (Figure 5).

DISCUSSION

Without structural information, XPD patient mutations have been difficult to understand, since patient phenotypes cannot be predicted from the position of an individual mutation along the linear gene sequence and since adjacent mutations can cause different diseases. Our SaXPD crystal structures and biochemical assays, done under conditions to limit oxidation of the 4Fe-4S cluster, provide an informative framework to reduce this complexity. As a working hypothesis that can now be tested in a variety of experimental systems, we have herein divided XPD mutations into three classes by their predicted impacts: (1) ATP and DNA binding for XP, (2) conformation for XP/CS, and (3) framework for TTD (Figure 5).

XP mutations tend to be in the predicted DNA- or ATP-binding channels, but all those tested reduce helicase activity and hence must reduce bubble opening in NER; this observation provides a structural explanation for biochemically based predictions from human XPD (Botta et al., 2002; Dubaele et al., 2003; Theron et al., 2005). XP mutations furthermore tend to reduce rather than increase DNA binding, so XPD mutant complexes are unlikely to stay bound on DNA substrates and act to block other processes.

XP/CS mutations cause the loss of both helicase activity and the functional flexibility of HD1-HD2, but possess substantial DNA binding affinity in two of four mutants analyzed (Figure 4). The aberrant conformations thus produce a loss of NER function without necessarily reducing DNA binding at target sites. Furthermore, the conformational switching driven by ATP- and DNA-binding states is likely to affect dynamic protein-protein interactions within the TFIIH complex and with other critical protein partners including XPG. A conformationally restricted state of XPD may affect protein interactions that normally differ between TFIIH functions in transcription initiation versus NER or transcription-coupled repair (TCR), resulting in context-inappropriate interactions and activities (Sarker et al., 2005). If so, then this might explain the otherwise perplexing biological observation that XP/CS mutations in XPD, but not XP or TTD mutations, result in NER-dependent inappropriate incisions at transcription sites distant from DNA damage (Theron et al., 2005). If XP/CS mutations cause HD1-HD2 to become locked in an abnormal conformation, this could promote changes in protein-protein interactions that specifically affect TFIIH functions in TCR, defects in which are the molecular hallmark of CS. Our structural results would predict that at least some of these interactions should be with HD2, so this can now be tested by experiments to map the interaction domains of XPD with RNA Pol II, CSB, and XPG. The high-resolution SaXPD structure fits into yeast and human TFIIH electron microscopy reconstructions, suggesting that the conserved XPD catalytic core informs the overall TFIIH architectural arrangement with the HsXPD Arch and C-terminal extensions likely involved in interactions (Figure S7). The computational placement of the XPD and XPB crystal structures within the TFIIH ring suggests XPD and XPB could cooperate in opening the DNA for NER, consistent with known XPB and XPD activities (Fan et al., 2006; Coin et al., 2007).

TTD mutations do not necessarily reduce helicase activity but are predicted to cause framework defects expected to impact levels of XPD as well as the stability of TFIIH, matching prior observations on human XPD biochemistry (Vermeulen et al., 2001; Dubaele et al., 2003). XPD C-terminal modifications weaken the interaction with p44 and thereby disturb the conformation of TFIIH (Dubaele et al., 2003), consistent with the XPD C-terminal extension being available for protein interactions. Our results also suggest that many TTD mutations impact XPD-protein interactions indirectly by primarily acting as framework defects. This structure-based interpretation agrees with the observed

cell biology, as there is reduced TFIIH in cells homozygous for the R112H mutation, which does not affect the XPD interaction with its p44 partner in TFIIH (Dubaele et al., 2003).

The mutation R112H (K84H in SaXPD), which involves a loss of the hydrogen bond to a Cys ligand of the 4Fe-4S cluster, also highlights the importance of the 4FeS domain. The gated channel and position of the 4Fe-4S cluster in XPD appear ideal for efficient damage sensing. The 4Fe-4S placement makes sense if the cluster is to be a detector of different types of bulky damage in DNA, consistent with experiments showing that 4Fe-4S cluster proteins are held at damaged sites where the clusters become oxidized (Yavin et al., 2006). The controlled oxidation of the 4Fe-4S cluster provides an elegant way for the cluster to potentially substantially augment DNA detection by the gateways at both ends of the XPD DNA-binding channel.

These XPDcc mutant analyses characterize primary defects. However, biological and clinical features of XPD mutations will depend upon their local severity combined with their impacts on interactions and functions at the next level. Yet our results suggest that most TTD and XP/CS mutations impact macromolecular interactions indirectly and in opposing ways, both of which may reduce TFIIH integrity, as shown experimentally (Vermeulen et al., 2001). Whereas TTD mutations should increase framework flexibility, XP/CS mutations appear to decrease HD1-HD2 functional flexibility. These structural results provide a basis to evaluate the likely impacts of such changes and to understand differences observed between cellular and clinical phenotypes. Consistent with activity analyses (Clarkson and Wood, 2005), these structural results would not, for example, support a functional repair role for XPD polymorphism D312N, which is in a surface-exposed position pointing away from the DNA-binding channel. These new results thus broaden our understanding of how XPD structural changes might impact cancer risks or result in developmental/aging phenotypes.

EXPERIMENTAL PROCEDURES

Cloning and Recombinant Protein Production

The XPD gene was amplified from *Sulfolobus acidocaldarius* genomic DNA and cloned into the pET15b vector for expression of untagged recombinant protein in E. coli. Protein expression and purification procedures were based on those published (Rudolf et al. 2006) with minor modifications. Mutants were generated using the Quikchange II XL Kit (Stratagene). SaXPD wildtype and mutant protein expression was carried out in BL21 Rosetta2 cells (Invitrogen) with details as described in Supplemental Data.

Crystallization, Data Collection, Structure Determination, Refinement, and Analysis

Purified SaXPD protein was concentrated to 10–20 mg/mL for crystallization experiments by vapor diffusion in an anaerobic glovebox for data collection using synchrotron radiation. The initial phases for SaXPD structure were calculated from the MAD data (Table 1), and the structures determined and

refined as described in Supplemental Data. Docking analyses were done with DOT as described in Supplemental Data.

ATPase, Helicase, and DNA-Binding Assays

ATPase activity was measured by incubating SaXPD with γ-^{32}P-ATP at 45^{o}C and separating free phosphate by thin-layer chromatography. Helicase activity was measured by incubating SaXPD with 50-overhang DNA substrates at 55^{o}C and resolving unwound labeled product by native PAGE. To minimize exposure of the protein to oxygen, all pipetting steps except for setting up the final reaction mixture were carried out in a nitrogen glove bag. SaXPD-DNA interactions were measured by fluorescence anisotropy. Details for all activity assays are described in the Supplemental Data.

ACCESSION NUMBERS

Atomic coordinates and structure factors have been deposited in the Protein Data Bank with the ID code 3CRV for SaXPD and 3CRW for the apo SaXPD.

SUPPLEMENTAL DATA

Supplemental Data include seven figures, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at http://www.cell.com/cgi/content/full/133/5/789/DC1/.

ACKNOWLEDGMENTS

We thank Steve Yannone for *S. acidocaldarius* DNA, Brian Chapados for aiding data processing, Karl-Peter Hopfner and Malcolm White for discussions, and Michael Pique and Arthur Olson for creating three-dimensional physical models for analyses. For the electron microscopy (EM) reconstructions, we thank Wei-Hau Chang and Roger Kornberg for providing the yeast TFIIH EM map and Arnaud Poterszman and Jean Marc Egly for providing the human TFIIH EM map. This work was supported by NIH grants R01 CA112093 and P01 CA92584 (both to J.A.T. and P.K.C.) and GM070996 (V.A.R.), plus NRSA Fellowship 1F32CA108393 to J.O.F. We thank the SIBYLS beamline 12.3.1 staff at the Berkeley Lab Advanced Light Source (ALS) and Stanford Synchroton Radiation Laboratory (SSRL) beamline 11.1 staff for aiding diffraction data collection and U.S. Department of Energy for support of ALS efforts under Contract Number DE-AC02-05CH11231.

Received: February 13, 2008
Revised: April 8, 2008
Accepted: April 21, 2008
Published: May 29, 2008

REFERENCES

Bienstock, R.J., Skorvaga, M., Mandavilli, B.S., and Van Houten, B. (2003). Structural and functional characterization of the human DNA repair helicase XPD by comparative molecular modeling and site-directed mutagenesis of the bacterial repair protein UvrB. J. Biol. Chem. 278, 5309–5316.

Boal, A.K., Yavin, E., Lukianova, 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. Biochemistry 44, 8397–8407.

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

Botta, E., Nardo, T., Lehmann, A.R., Egly, J.M., Pedrini, A.M., and Stefanini, M. (2002). Reduced level of the repair/transcription factor TFIIH in trichothiodystrophy. Hum. Mol. Genet. 11, 2919–2928.

Buttner, K., Nehring, S., and Hopfner, K.P. (2007). Structural basis for DNA duplex separation by a superfamily-2 helicase. Nat. Struct. Mol. Biol. 14, 647–652.

Cantor, S., 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. USA 101, 2357–2362.

Chapados, B.R., Hosfield, D.J., Han, S., Qiu, J., Yelent, B., Shen, B., and Tainer, J.A. (2004). Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. Cell 116, 39–50.

Clarkson, S.G., and Wood, R.D. (2005). Polymorphisms in the human XPD (ERCC2) gene, DNA repair capacity and cancer susceptibility: An appraisal. DNA Repair (Amst.) 4, 1068–1074.

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

de Boer, J., Andressoo, J.O., de Wit, J., Huijmans, J., Beems, R.B., van Steeg, H., Weeda, G., van der Horst, G.T.J., van Leeuwen, W., Themmen, A.P.N., et al. (2002). Premature aging in mice deficient in DNA repair and transcription. Science 296, 1276–1279.

Dubaele, S., Proietti, D., 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, 1635–1646.

Fan, L., Arvai, A.S., Cooper, P.K., Iwai, S., Hanaoka, F., and Tainer, J.A. (2006). Conserved XPB core structure and motifs for DNA unwinding: implications for pathway selection of transcription or excision repair. Mol. Cell 22, 27–37.

Hoeijmakers, J.H.J. (2001). Genome maintenance mechanisms for preventing cancer. Nature 411, 366–374.

Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P., and Tainer, J.A. (2000). Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. Cell 101, 789–800.

Ito, S., Kuraoka, I., Chymkowitch, P., Compe, E., Takedachi, A., Ishigami, C., Coin, F., Egly, J.M., and Tanaka, K. (2007). XPG stabilizes TFIIH, allowing transactivation of nuclear receptors: Implications for Cockayne syndrome in XP-G/CS patients. Mol. Cell 26, 231–243.

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

Lee, J.Y., and Yang, W. (2006). UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. Cell 127, 1349–1360.

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

Levran, O., Attwooll, C., Henry, R.T., Milton, K.L., Neveling, K., Rio, P., Batish, S.D., Kalb, R., Velleuer, E., Barral, S., et al. (2005). The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. Nat. Genet. 37, 931– 933.

Ludovic, C., Gillet, J., and Scharer, O.D. (2006). Molecular Mechanisms of Mammalian Global Genome Nucleotide Excision Repair. Chem. Rev. 106, 253–276.

Outten, F.W. (2007). Iron-sulfur clusters as oxygen-responsive molecular switches. Nat. Chem. Biol. 3, 206–207.

Pugh, R.A., Honda, M., Leesley, H., Thomas, A., Lin, Y., Nilges, M.J., Cann, I.K., and 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.

Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence complexity of disordered protein. Proteins 42, 38–48.

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. Mol. Cell 23, 801–808.

Sarker, A.H., Tsutakawa, S.E., Kostek, S., Ng, C., Shin, D.S., Peris, M., Campeau, E., Tainer, J.A., Nogales, E., and Cooper, P.K. (2005). Recognition of RNA polymerase II and transcription bubbles by XPG, CSB, and TFIIH: Insights for transcription-coupled repair and Cockayne Syndrome. Mol. Cell 20, 187– 198.

Schaeffer, L., Roy, R., Humbert, S., Moncollin, V., Vermeulen, W., Hoeijmakers, J.H.J., Chambon, P., and Egly, J.M. (1993). Dna-repair helicase - a component of btf2 (tfiih) basic transcription factor. Science 260, 58–63.

Schumacher, B., Garinis, G.A., and Hoeijmakers, J.H. (2008). Age to survive: DNA damage and aging. Trends Genet. 24, 77–85.

Shin, D.S., Pellegrini, L., Daniels, D.S., Yelent, B., Craig, L., Bates, D., Yu, D.S., Shivji, M.K., Hitomi, C., Arvai, A.S., et al. (2003). Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. EMBO J. 22, 4566–4576.

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

Sung, P., Bailly, V., Weber, C., Thompson, L.H., Prakash, L., and Prakash, S. (1993). Human xeroderma-pigmentosum group-d gene encodes a dna helicase source. Nature 365, 852–855.

Taylor, E.M., Broughton, B.C., Botta, E., Stefanini, M., Sarasin, A., Jaspers, N.G.J., Fawcett, H., Harcourt, S.A., Arlett, C.F., and Lehmann, A.R. (1997). Xeroderma pigmentosum and trichothiodystrophy are associated with different mutations in the XPD (ERCC2) repair/transcription gene. Proc. Natl. Acad. Sci. USA 94, 8658–8663.

Thayer, M., Ahern, H., Xing, D., Cunningham, R.P., and Tainer, J.A. (1995). Novel DNA binding motifs in the DNA-repair enzyme endonuclease III structure. EMBO J. 14, 4108–4120.

Theron, T., Fousteri, M.I., Volker, M., Harries, L.W., Botta, E., Stefanini, M., Fujimoto, M., Andressoo, J.O., Mitchell, J., Jaspers, N.G., et al. (2005). Transcription- associated breaks in xeroderma pigmentosum group D cells from patients with combined features of xeroderma pigmentosum and Cockayne syndrome. Mol. Cell. Biol. 25, 8368–8378.

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, 87–95.

Vermeulen, W., Rademakers, S., Jaspers, N.G., Appeldoorn, E., Raams, A., Klein, B., Kleijer, W.J., Hansen, L.K., and Hoeijmakers, J.H. (2001). A temperature- sensitive disorder in basal transcription and DNA repair in humans. Nat. Genet. 27, 299–303.

Winkler, G.S., Araujo, S.J., Fiedler, U., Vermeulen, W., Coin, F., Egly, J.M., Hoeijmakers, J.H., 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, 4258–4266.

Yamagata, A., and Tainer, J.A. (2007). Hexameric structures of the archaeal secretion ATPase GspE and implications for a universal secretion mechanism. EMBO J. 26, 878–890.

Yavin, E., Stemp, E.D., O'Shea, V.L., David, S.S., and Barton, J.K. (2006). Electron trap for DNA-bound repair enzymes: a strategy for DNA-mediated signaling. Proc. Natl. Acad. Sci. USA 103, 3610–3614.

Table 1. Data Collection and Refinement Statistics

Data collection	wt	SeMet MAD			apo
Wavelength (Å)	0.9800	0.97931	0.97899	0.91837	1.0000
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions (Å)	a=53.54 b=70.22 c=144.3	a=53.68 b=70.06 c=145.3			a=53.39 b=69.81 c=145.2
Resolution (Å)	50-2.00 (2.07-2.00)	50-2.45 (2.54-2.45)	50-2.45 (2.54-2.45)	50-2.45 (2.54-2.45)	45-3.0 (3.11-3.0)
Rsym (%)	2.7 (41.0)	7.5 (65.5)	8.0 (50.7)	7.1 (54.8)	6.6 (32.3)
Completeness (%)	99.0 (97.1)	97.1 (96.0)	97.3 (98.7)	97.1 (98.6)	81.5 (37.7)
I/sigma(I)	35.3 (2.3)	9.6 (4.8)	16.0 (3.9)	15.3 (3.8)	26.0 (2.2)
Unique reflections	37145	20313	24303	20274	9404
Total reflections	181139	190108	224711	200497	57268

Refinement	wt	apo
Resolution (Å):	50-2.00	50-3.00
R-factor/ R-free (%):	22.3/26.0	24.5/31.0
Average B-factor ($Å^2$)	41.2	107.3
Protein atoms Cofactor atoms Ion/ solvent/water	4508 8 0/9/279	3967 0 1/0/44
Rms Deviations: Bonds (Å) Angles (^{o})	 0.008 1.28	 0.008 4.44

Data in parentheses are those of the highest resolution shell

Table 2. Catalytic Activity and DNA binding of SaXPD Mutants

Human mutation	SacXPD mutation	Disease	Motif	ATPase mol/sec (%wt)	Helicase bp/min (%wt)	ssDNA Kd (nM) (%wt)
Wild-type	--	--		0.55 (100%)	2.22 (100%)	46.3 (100%)
T76A	T56A	XP	Ia	0.21 (38%)	0.15 (7%)	57.1 (81%)
D234N	D180N	XP	II	0.05 (9%)	0.05 (2%)	56.3 (82%)
Y542C	Y403C	XP	IV	0.42 (76%)	0.42 (19%)	46.0 (101%)
R601L/W	K446L	XP	V	0.31 (56%)	0.32 (14%)	59.0 (78%)
R683W/Q	R531W	XP		0.01 (2%)	0.04 (2%)	Ind.
K507	K369Q	--	Channel	0.16 (29%)	0.29 (13%)	60.9 (76%)
G47R	G34R	XP/CS	I	0.00 (0%)	0.04 (2%)	29.9 (155%)
G602D	G447D	XP/CS		0.07 (13%)	0.03 (1%)	68.1 (68%)
R666W	R514W	XP/CS	VI	0.00 (0%)	0.03 (1%)	58.8 (79%)
G675R	C523R	XP/CS		0.23 (42%)	0.31 (14%)	28.2 (164%)
R112H	K84H	TTD		0.48 (87%)	0.10 (5%)	79.8 (58%)
R592P	K438P	TTD	V	0.31 (56%)	2.28 (103%)	52.1 (89%)
D673G	D521G	TTD		0.19 (35%)	0.52 (23%)	49.2 (94%)
C116	C88S	--	4Fe-4S	0.18 (33%)	0.06 (3%)	Ind.
C134	C102S	--	4Fe-4S	0.10 (18%)	0.05 (2%)	N.A.

Numbers in parentheses indicate percentage of wild-type activity level. “Ind.” – indeterminate value, “N.A.” – not assayed.

Figure 1. XPDcc Conservation, Functional Motifs, Mutation Sites, Domains, and Structure (A) XPDcc domains. Sequence comparison indicates that SaXPD contains the XPD conserved catalytic core, so the four XPDcc domains are shown schematically in boxes for HD1 (cyan), HD2 (green), 4FeS (orange), and Arch (purple) domains with conserved helicase motifs (red bars with white labels). Disease XP (red), XP/CS (yellow), and TTD (purple) mutation sites are labeled for human (Hs) and corresponding SaXPD (in parentheses) sites. Residue F136 corresponding to a FancJ mutation is highlighted by a blue-green flag. For the detailed sequence alignment with mutation sites, secondary structure, and domain fold, see Supplemental Data (Figure S1). The human enzyme has a partly ordered (gray) C-terminal extension (CTE) (as predicted by PONDR) that is a probable TFIIH p44 interface. (B) XPDcc fold and domains (ribbons). Helicase domains HD1 (cyan) and HD2 (green) form the ATP-binding interface. Front view (left) shows the arch formed by the 4FeS (orange) and Arch (purple) domains, which are inserted into HD1. Side view (right) shows HD2 protruding from the flat box formed by HD1, 4FeS, and Arch as well as the HD2 helix-loop-helix insertion (green). Domain boundaries are indicated by residue numbers. (C) Apo XPDcc structure. Disordered regions (dashed lines with the boundaries indicated by residue numbers) show the 4Fe-4S cluster acts in ordering the 4FeS domain, the Arch interface, and parts of HD1. (D) Electron density for the 4Fe-4S cluster and key residues forming hydrogen bonds (green dashed lines) to the Cys ligands. Mutations at these sites cause TTD in XPD and Fanconi anemia in FancJ. Composite omit maps calculated from the model are displayed at 1 sigma level.

Figure 2. Structure-Based Model for DNA Interactions (A) Conserved molecular surface (front view). Conserved residues are shown on the surface from deep blue (identical) to light blue (highly similar) to greenish light blue (similar) based on sequence alignment between HsXPD and SaXPD (Figure S1). (B) Electrostatic molecular surface of the arch region (back view). Electrostatic potential (blue positive to red negative) suggests a dsDNA binding site (arrow). (C) XPD-DNA binding model (DNA phosphate backbone as purple tube). ssDNA binding was located by superimposing known helicase-DNA complex structures (2P6R.pdb) (Buttner et al., 2007) to the SaXPD structure over conserved HD2 helicase motifs IV and VI. The SaXPD dsDNA-binding sites were located by a complete six-dimensional search. Helicase motifs are in red. (D) Electrostatic molecular surface for the DNA-binding HD2 gateway. Electrostatic potential is calculated as in (B) and DNA (yellow tube) is modeled as in (C).

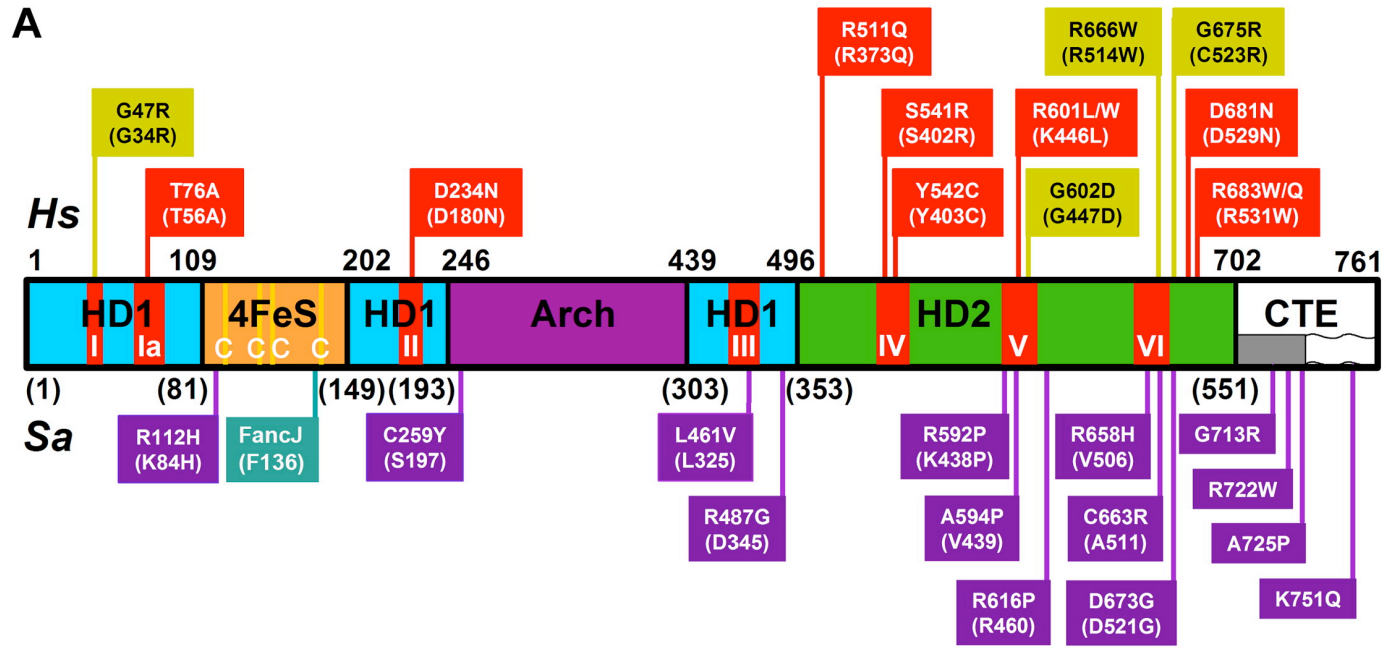
Figure 3. Structural Roles of Amino Acid Residues Associated with Disease-Causing XPD Mutations Disease-causing mutations (Ca colored spheres: XP are red, XP/CS are gold, and TTD are purple) mapped in the SaXPD structure with hydrogen bonds (red dashed lines). (A) Mutation sites at the HD2 gateway DNA-binding channel shown in Figures 2C and 2D. Solvent molecules including citrate (CIT), isopropanol (IPA), and glycerol (GOL) that mimic DNA backbone (purple tube, see Figure 2C),

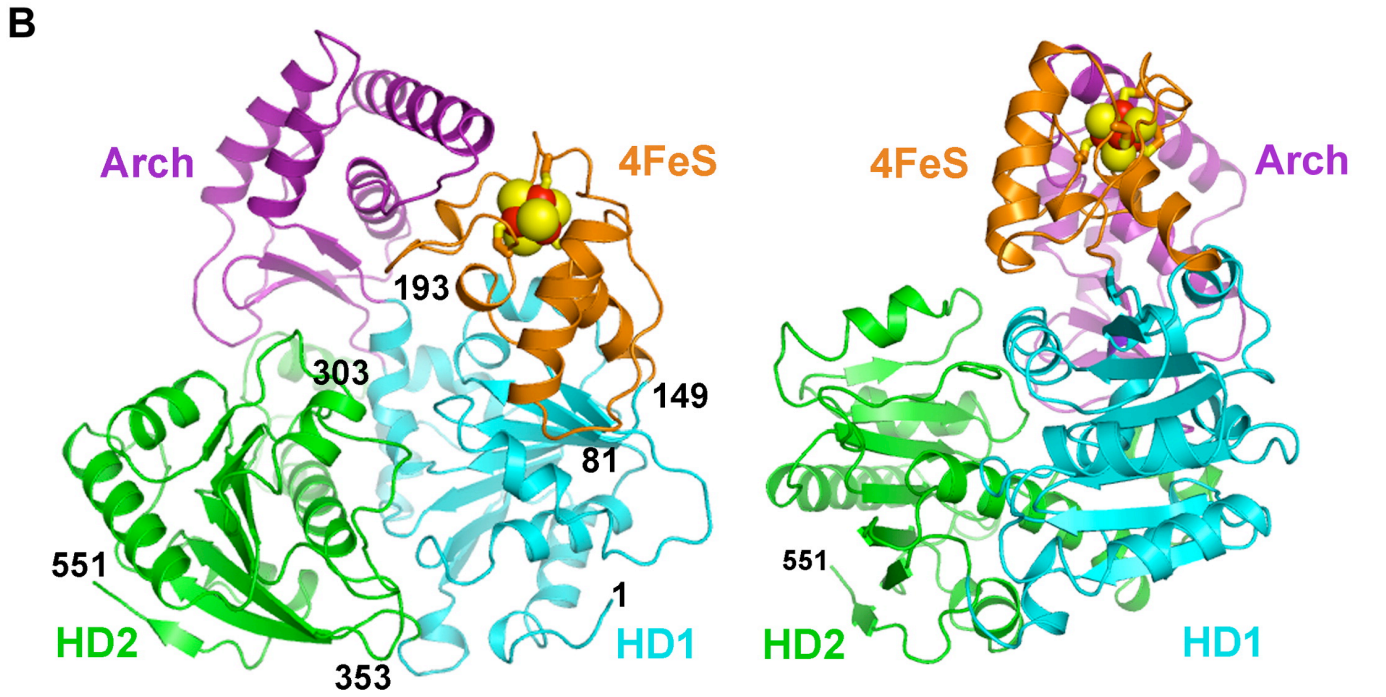
interact with charged residues K369, R373, and D529. (B) Mutation sites at the groove along HD2, Arch, and HD1 domains. (C) Mutation sites at the ATP-binding groove between HD1 and HD2 domains. (D) Mutation sites at the edge of HD2 domain.

Figure 4. Effects of Disease-Causing XPD Mutations Catalytic activity and DNA-binding of SaXPD mutants, as percentage of wild-type activity. DNA-binding affinities were not determined for mutants R531W, C88S, and C102S. See Table 2 for rate and Kd measurements and Figure S5 for representative data.

Figure 5. Structural Placement of XPDcc Disease-Causing Mutations Mapping the three classes of mutations onto the SaXPD structure reveals patterns associated with each disease defect. (A) Stereo pair mapping the distribution of disease causing mutations on a XPDcc Ca trace. Disease causing mutation sites (Ca colored sphere): red (XP), gold (XP/CS), and purple (TTD). Residue F136 (a FancJ mutation) is also shown (cyan). (B) XPDcc fold and domain architecture (ribbons) with labeled disease-causing mutation sites as spheres colored as in (A). (C) XP mutations impact DNA and ATP-binding regions. (D) XP/CS mutations impact HD1-HD2 conformational changes. (E) TTD mutations impact overall framework stability.

Fig. 1


A
Hs
Sa
G47R (G34R)
T76A (T56A)
D234N (D180N)
R511Q (R373Q)
S541R (S402R)
Y542C (Y403C)
R601L/W (K446L)
G602D (G447D)
R666W (R514W)
G675R (C523R)
D681N (D529N)
R683W/Q (R531W)
1
109
202
246
439
496
702
761
HD1
I
Ia
4FeS
C C C C
HD1
II
Arch
HD1
III
IV
HD2
V
VI
CTE
(1)
(81)
(149)
(193)
(303)
(353)
(551)
R112H (K84H)
FancJ (F136)
C259Y (S197)
L461V (L325)
R487G (D345)
R592P (K438P)
A594P (V439)
R616P (R460)
R658H (V506)
C663R (A511)
D673G (D521G)
G713R
R722W
A725P
K751Q


B
Arch
4FeS
193
303
149
81
551
1
HD2
HD1
353
4FeS
Arch
551
HD2
HD1

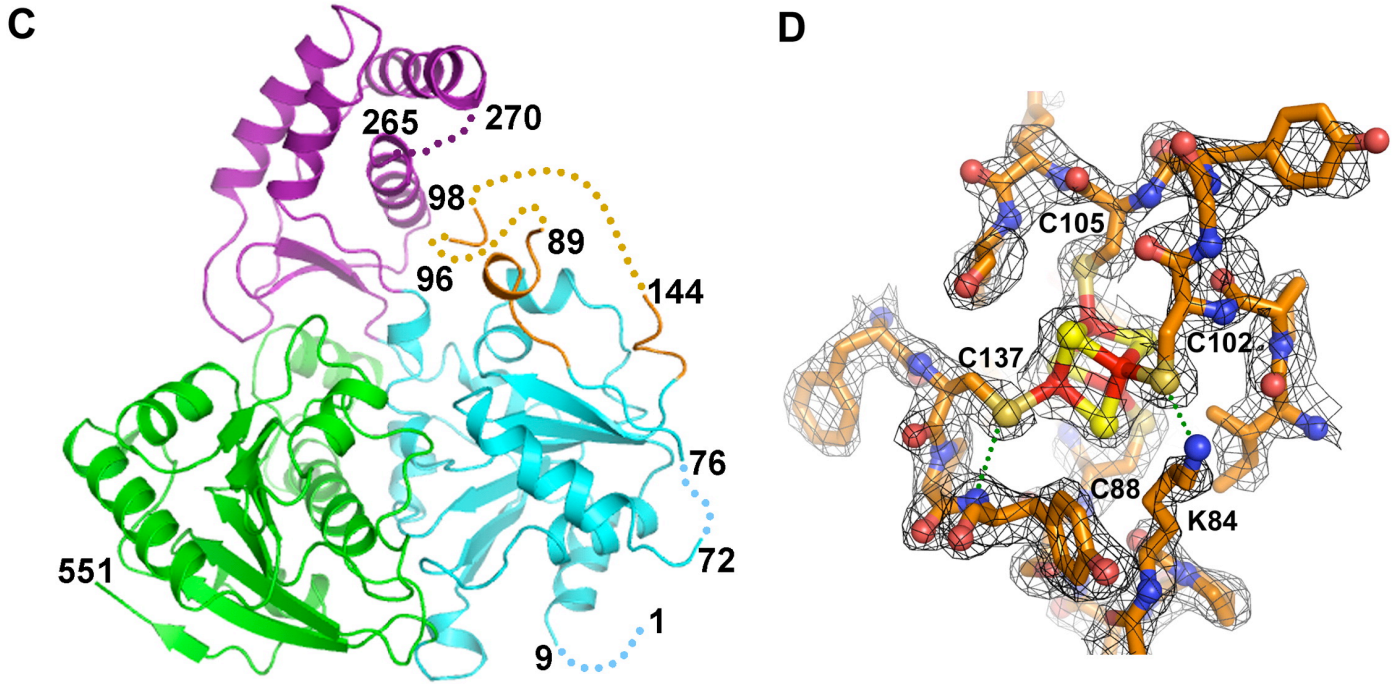

C
265
270
98
89
96
144
76
551
72
1
9
D
C105
C137
C102
C88
K84

Fig. 2

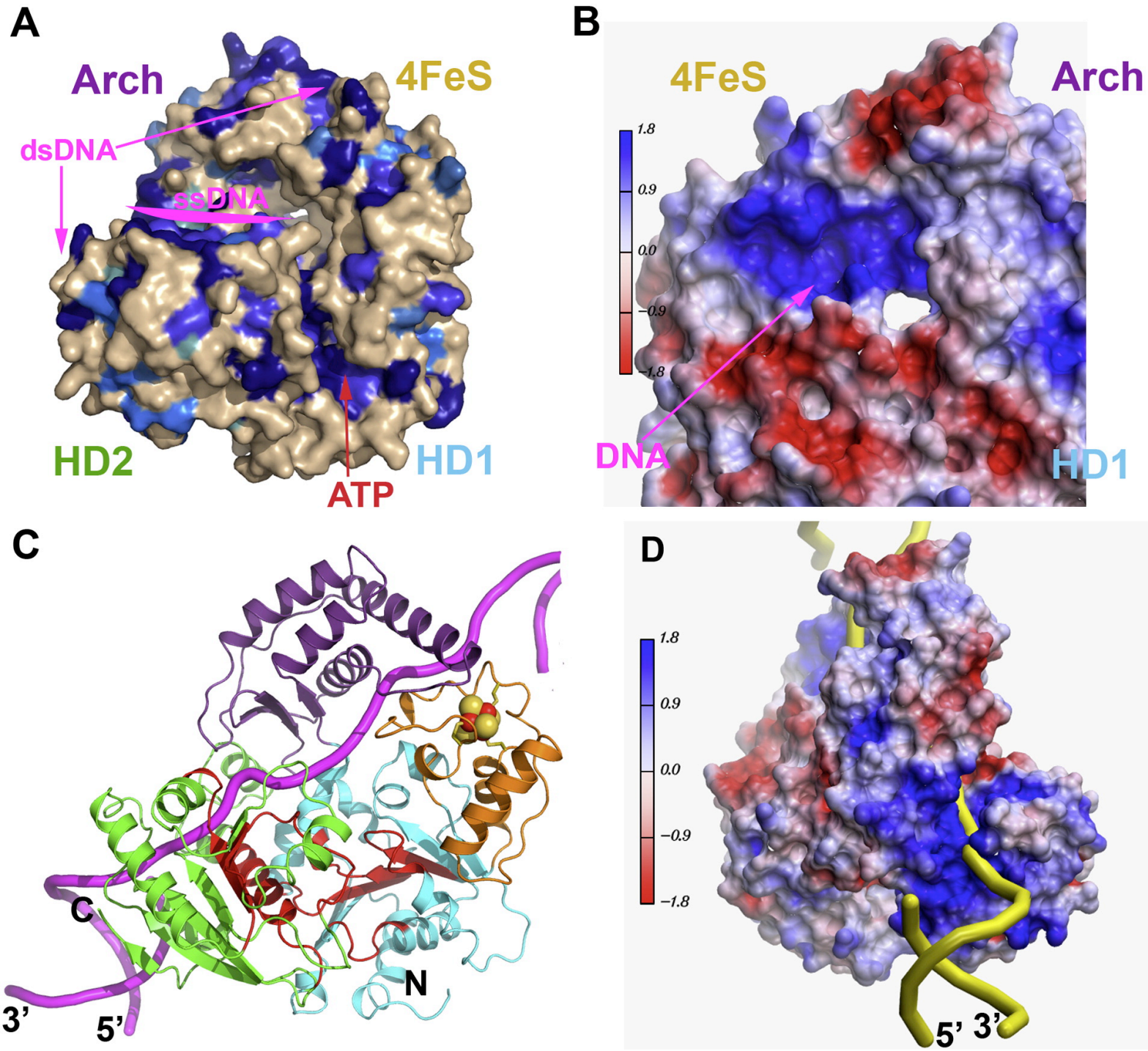

A
Arch
4FeS
dsDNA
ssDNA
HD2
HD1
ATP
B
4FeS
Arch
1.8
0.9
0.0
−0.9
−1.8
DNA
HD1
C
C
N
3’
5’
D
1.8
0.9
0.0
−0.9
−1.8
5’ 3’

Fig. 3

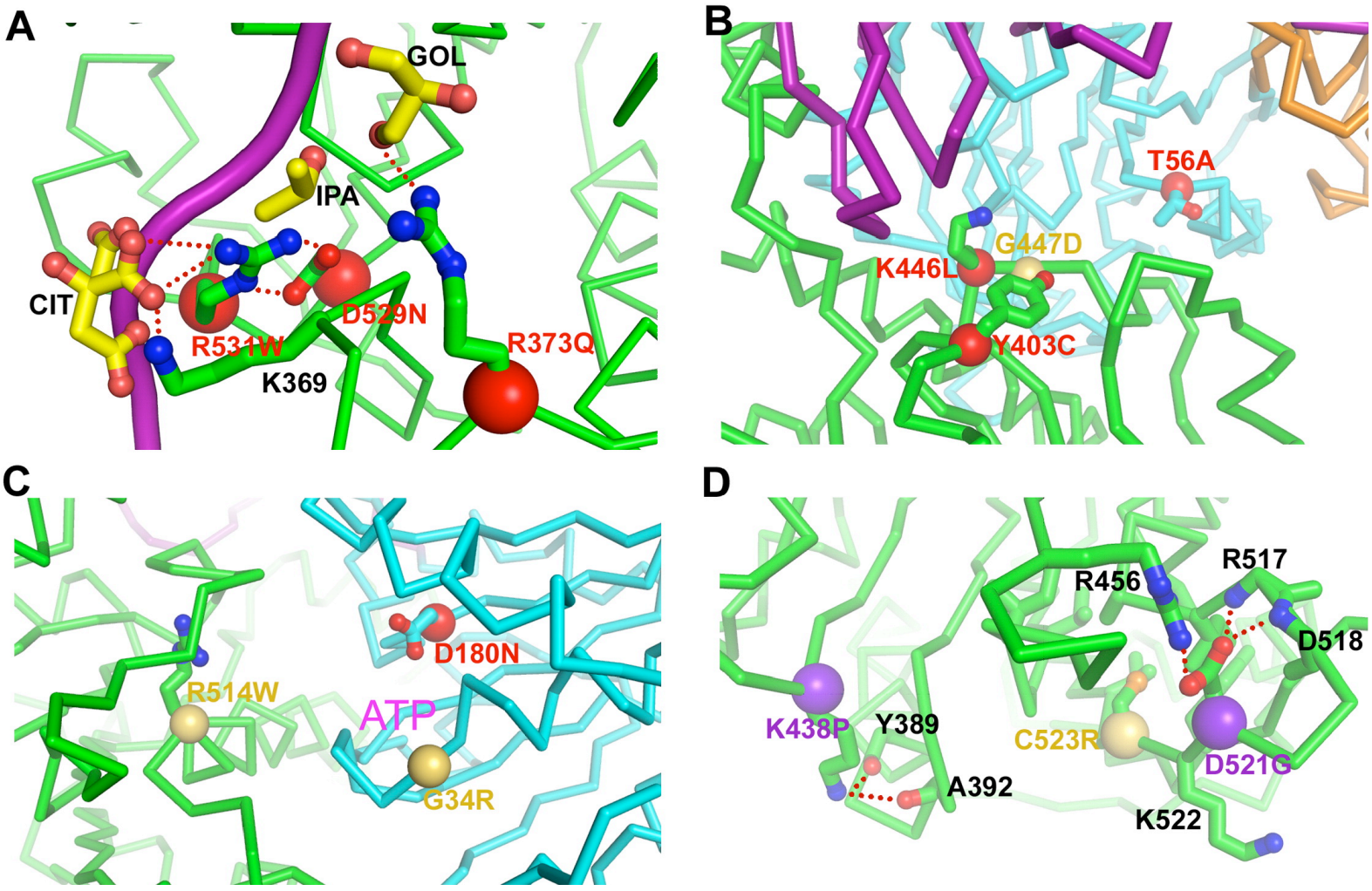

A
GOL
IPA
CIT
R531W
D529N
R373Q
K369
B
T56A
G447D
K446L
Y403C
C
D180N
R514W
ATP
G34R
D
R517
R456
D518
K438P
Y389
C523R
D521G
A392
K522

Fig. 4

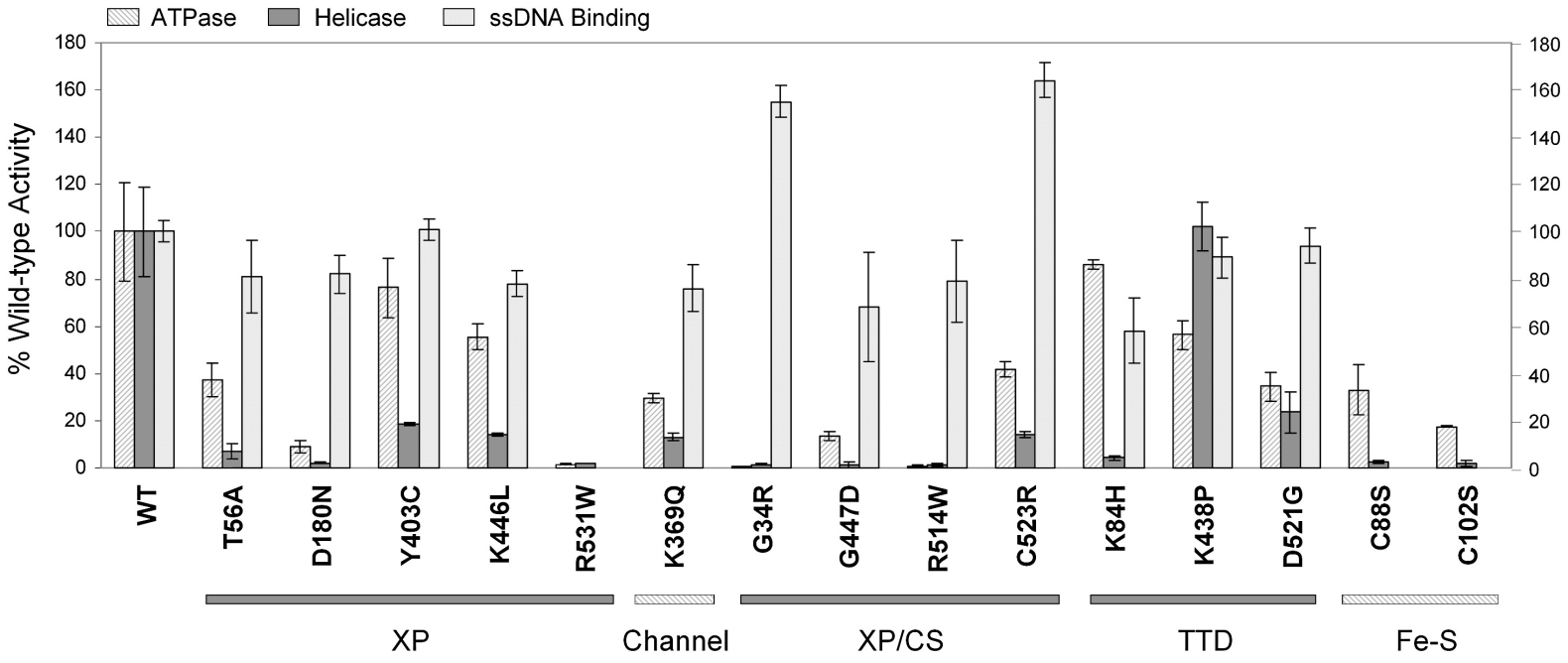
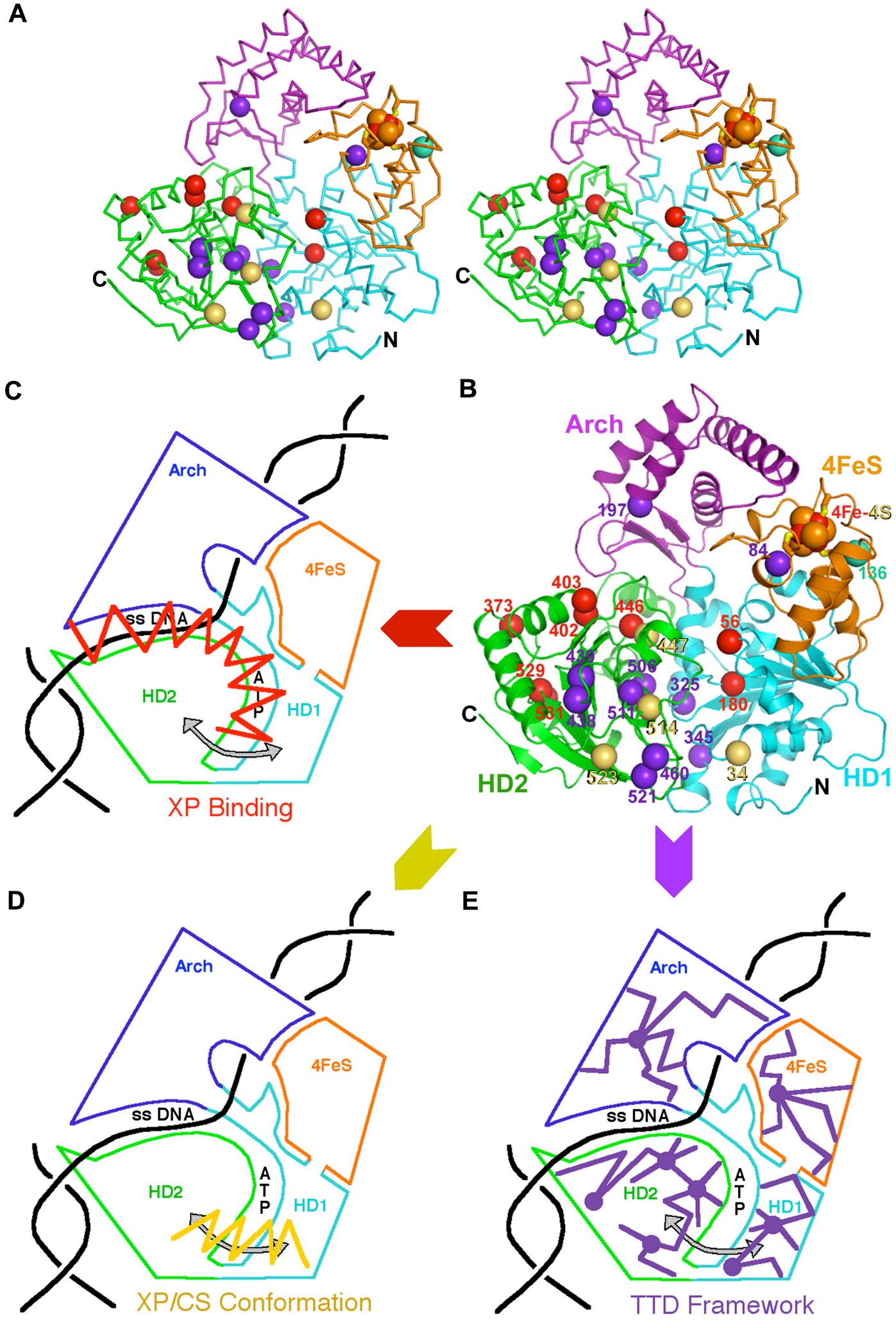

ATPase
Helicase
ssDNA Binding
% Wild-type Activity
0
20
40
60
80
100
120
140
160
180
WT
T56A
D180N
Y403C
K446L
R531W
K369Q
G34R
G447D
R514W
C523R
K84H
K438P
D521G
C88S
C102S
XP
Channel
XP/CS
TTD
Fe-S

Fig. 5


A
C
N
C
N
C
Arch
4FeS
ss DNA
HD2
ATP
HD1
XP Binding
B
Arch
4FeS
4Fe-4S
197
84
136
403
373
446
402
56
447
439
506
529
325
180
531
438
511
514
345
C
HD2
523
460
34
521
N
HD1
D
Arch
4FeS
ss DNA
HD2
ATP
HD1
XP/CS Conformation
E
Arch
4FeS
ss DNA
HD2
ATP
HD1
TTD Framework