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*

Rachelle J. Bienstock‡, Milan Skorvaga¶, Bhaskar S. Mandavilli§, and Bennett Van Houten∥

From the ‡Scientific Computing Laboratory and §Laboratory of Molecular Genetics, NIEHS, National Institutes of Health, Research Triangle Park, North Carolina 27709 and ∥Department of Molecular Genetics, Cancer Research Institute, Slovak Academy of Sciences, Vlarska 7, 833 91 Bratislava, Slovakia

A molecular model for the human nucleotide excision repair protein, XPD, was developed based on the structural and functional relationship of the protein with a bacterial nucleotide excision repair (NER) protein, UvrB. Whereas XPD does not share significant sequence identity with UvrB, the proteins share seven highly conserved helicase motifs that define a common protein structural template. They also have similar functional roles in their ATPase activity and the ability to unwind DNA and verify damaged strands in the process of NER. The validity of using the crystal structure of UvrB as a template for the development of an XPD model was tested by mimicking human disease-causing mutations (XPD: R112H, D234N, R601L) in UvrB (E110R, D338N, R506A) and by mutating two highly conserved residues (XPD, His-237 and Asp-609; UvrB, H341A and D510A). The XPD structural model can be employed in understanding the molecular mechanism of XPD human disease causing mutations. The value of this XPD model demonstrates the generalized approach for the prediction of the structure of a mammalian protein based on the crystal structure of a structurally and functionally related bacterial protein sharing extremely low sequence identity (<15%).

Nucleotide excision repair (NER) is a process by which damaged nucleotides, from UV light and chemical carcinogens, are removed from DNA. NER is one of the most highly conserved biochemical pathways, and enzymes that mediate this process appear in prokaryotes, archaea, and eukaryotes (1, 2). NER can be viewed in the following five continuous steps: 1) damage recognition and verification, 2) incision, 3) excision, 4) repair synthesis, and 5) ligation. During the damage recognition step, a protein complex first identifies a structural perturbation in the helical DNA, which is then verified by additional damage processing and strand opening proteins. Once the lesion has been verified, endonucleases are recruited to the damaged strand performing the initial incision followed by a 5’ incision facilitating the excision of an oligonucleotide containing the damage. Repair synthesis fills in the resulting gap. Finally DNA ligase seals the newly completed repair patch. The evolutionary conservation of proteins involved in NER is exemplified by human XPD, which maintains 56 and 52% (protein) sequence identity, respectively, with the equivalent Schizosaccharomyces pombe, Rad15, and Saccharomyces cerevisiae, RAD3, proteins over the length of the entire protein (BLAST alignments). Human and mouse XPD retain 98% (protein) sequence identity (3).

XPD is one of the constituent proteins forming the TFIIH complex that is responsible for damage processing and strand verification (4, 5). XPD and XPB have intrinsic helicase activities that are absolutely required for NER (4). TFIH is also required for initiation of RNA polymerase II transcription. Mutations in the XPD gene can lead to one of the following three disease syndromes in humans: xeroderma pigmentosum (XP), trichothiodystrophy (TTD), and Cockayne syndrome (CS) (6). Analysis of specific point mutations associated with each disease has revealed that a majority of patients with XPD have mutations in one of seven helicase motifs (6).

Helicases from bacteria to humans are believed to have evolved from a common ancestor (7). Helicase motifs may be so highly conserved because they represent a fundamental mechanism for DNA processing necessary for transcription, replication, repair, and recombination (8). Mutations in several helicase proteins have been linked to cancer-related and accelerated aging disorders (Werner’s syndrome, Bloom’s syndrome, xeroderma pigmentosum, Cockayne’s syndrome, and xeroderma pigmentosum) (9–11).

Helicases have been classified into three large superfamilies (designated SF1, SF2, and SF3) and two smaller families (F4 and F5) (12). Proteins belonging to SF1 and SF2 are characterized as such based on possessing a set of seven conserved motifs ranging in length from about 20 to 40 amino acids. These helicase motifs demonstrate both a sequence and three-dimensional structure conservation. SF1 and SF2 helicase proteins were predicted to have a conserved β-pleated sheet structure based on other solved NTPase structures sharing these motifs (12). Structural studies have confirmed this prediction. Helicase systems contain the Walker box A and B motifs found in many NTP-binding proteins (in helicase motif I and II). Individual specificity of the protein function is conferred by the insertion of other domains not shared between the proteins. Crystal structures of three known monomeric/dimeric helicases and one repair protein that shares a common helicase fold have been determined as follows: 1) Escherichia coli Rep with ssDNA alone, and a ternary complex with ADP (13, 14); 2) Bacillus stearothermophilus PcrA with ADP-PNP, ssDNA, and...
Mg$^{2+}$ (15–17); 3) hepatitis C virus NS3 with single-stranded oligonucleotide (14, 18–20); and 4) UvrB, from Thermus thermophilus (21, 22), and Bacillus caldotenax with ATP and Mg$^{2+}$ (23, 24).

In general, structural homology between related proteins is much more highly conserved than sequence homology (25). Thus the helicase motifs possess similar folds and relative three-dimensional positional relationships within the proteins despite their low overall sequence identity. Over the entire length of the protein, XPD shares 12.8% sequence identity and 61% sequence similarity with PcrA, 15% identity and 59% similarity with Rep, 15% identity and 62% similarity with UvrB (Fig. 1), and 12.5% identity and 56% similarity with NS3. Both UvrB and XPD belong to SF2 and exhibit greater similarity in sequence as compared with the other three helicases whose structures have been solved. In addition, UvrB serves a homologous function as TFIIH in bacterial NER (24, 26, 27) as well as TFIIH in bacterial NER (24, 26, 27) in that it assists in verifying a damaged site and processing the lesion.

Three-dimensional relational positions within the proteins have been solved. In addition, UvrB serves a homologous function as TFIIH in bacterial NER (24, 26, 27) as well as TFIIH in bacterial NER (24, 26, 27) in that it assists in verifying a damaged site and processing the lesion. Two of these helicase families have been shown to be related based on their common helicase structural motifs and biological function, to open double-stranded DNA during NER (28). Accuracy of the XPD model was confirmed experimentally by expressing B. caldotenax UvrB proteins containing mutations corresponding to known disease-causing mutations in XPD. In addition, two residues highly evolutionarily conserved throughout all bacterial UvrB species and also RAD3 and XPD (mouse and human) were also mutated. The behavior of these mutant UvrB proteins was examined through incision, DNA binding, and oligonucleotide destabilizing assays in a reconstituted system containing purified UvrA and UvrC from B. caldotenax. The XPD model serves as a guide for structural-functional studies of the protein and assists in gaining a molecular understanding of the NER process in humans, XPD disease-causing mutations, and XPD protein–protein and protein–DNA interactions.

**MATERIALS AND METHODS**

**Molecular Modeling of XPD**

**Multiple Sequence Alignment**—A structure-based multiple sequence alignment was developed for the DNA helicase proteins, using the location of domains and helicase motifs, as these sequences do not possess high sequence identity. The alignment included sequences of all the solved monomeric helicase structures (Rep, PcrA, NS3, and UvrB), five bacterial UvrB sequences, human and mouse XPD, and RAD3. The yeast XPD homologue sequences. Part of this multiple alignment, the alignment of UvrB and XPD sequences, is included in Fig. 1. The alignment method used was based on the work of Gorbalenya et al. (29) and Gorbalenya and Koonin (30) using the programs DIAGON (31) and OPTAL (32) and the Sankoff algorithm (33) for multiple sequence alignment. The XPD model was aligned by hand to the multiple sequence alignment described for SP2 helicases by Gorbalenya et al. (29) and Gorbalenya and Koonin (30) using the RAD3 sequence as a guide.

**XPD Model Construction**—The XPD protein model was constructed using the bacterial B. caldotenax UvrB crystal structure coordinates (Fig. 2A (Protein Data Bank code 1D9Z) (23)) as the template. The sequence alignment indicated that XPD shared closer sequence similarity within the helicase motifs with UvrB, and the alignment resulted in the fewest deletions and insertions throughout the protein sequence as compared with Rep, PcrA, and NS3. UvrB and XPD also share an analogous function as nucleotidyl transferase repair proteins. The XPD homology model was constructed using the Accelrys Homology software package (MSI Homology Users Guide, San Diego). There were 15 loop insertions constructed using the Molecular Simulations homology modeling loop generation software. Because the structure for the loops is less well determined than the rest of the model, as they are not built from experimentally determined structure, they are indicated in black in the model structure in Fig. 2B. UvrB residues 181–301 were deleted from the XPD model structure, as there are no corresponding residues in the XPD helicase. XPD contains an additional domain, XPD residues 283–448, not present in UvrB. This domain is not similar to any other helicase protein domains and does not possess sequence homology to any other solved protein or protein domain structures. A suggestion of its function is made from the use of folding algorithms in combination with secondary structure prediction methods (as described in the Supplemental Material). The structural model of this domain is shown in Fig. 3.

**Modeling the Inserted (Non-homologous) XPD Domain (Residues 283–448)**—Three protein-threading algorithms were applied independently to identify a protein fold to model the XPD domain represented by residues 283–447 (not present in UvrB) as follows: the GenThreader algorithm (34, 35), the 3D-PSSM algorithm (36, 37), and the Bioinbug algorithm (42). The top protein folds for this XPD domain given by all three algorithms independently were all α-helical proteins, and all shared a fold with nucleic acid-binding proteins. The consensus from these three methods is that this domain has a protein fold similar to that of the AraC family transcriptional activator (Protein Data Bank entry 1BL0) or DNA-binding domain of the centromere binding protein B (CENP-B) (Protein Data Bank entry 1BW6). Both of these proteins are members of the same SCOP database protein fold class characteristic of all α-helical DNA/RNA binding, three α-helical bundle proteins. The structural model of this domain based on this fold is shown in Fig. 3.

**Modeling the XPD C Terminus (Residues 713–761)**—The NMR and crystal structure of the UvrB C terminus were solved independently of the rest of the protein (38, 39). This part of the protein was found to have a helix-turn-helix fold and to interact with a homologous domain in the other subunit. There was no homology or sequence similarity between the XPD and UvrB C termini.

The C-terminal XPD sequence does not exhibit high sequence identity to any previously solved structure, threading was used to search for a “structural” similarly folded homologue. Consensus results from seven protein threading algorithms (L0OPP (40), The Sausage Machine (41), 3D-PSSM (36, 37), Bioinbug (42), GenThreader (34, 35), COBLATH (43), and Fugue (44)) were used in concert with secondary structure prediction algorithms (PSIPRED (45), PHD (46), SSSP PRO (47), PROF (48), and PREDATOR (49)) to develop a threaded model for the XPD C terminus (49 residues, 713–761). The top structural hits from all threading methods were all α-helical proteins (all members of the CATH 1.10.10.60 class, and secondary structure prediction methods also predicted the 49 C-terminal residues would form an α-helical bundle. Similar protein structures that were selected by more than one of the seven different threading methods are as follows: 1OCTC, the oct-1 POU domain (a DNA binding protein); 1BH9A (TBP associated factor); and 4HB1, a designed four-helical bundle.

A structural model was developed for the XPD C terminus based on the coordinates from two of the helices were used to model the XPD C terminus. The fold of the XPD C terminus is shown in Fig. 3. The threaded model structure developed for the XPD C terminus is shown and compared with the UvrB experimentally determined C-terminal structure (38, 39) in Fig. 3. A significant number of frequently occurring XPD disease-causing mutations of XP or TTD patients have mutations occurring in the C terminus of the protein. A structural model was aligned by hand to the XPD model, and experimentally to bind to the N-terminal domain of another component TFIH protein, p44 (50). p44 residues 68–76, 683, 713, and 722 have been shown experimentally to interact with XPD (51).

**Experimental Assays**

**Construction and Expression of UvrB Mutants—**UvrA, UvrB, and UvrC proteins from B. caldotenax were purified by standard procedures (New England Biolabs IMPACT™ T7 system manual) with some modifications for each protein, which will be published elsewhere. T4 polynucleotide kinase was purchased from Invitrogen. Pfu DNA polymy
The sequence alignment of UvrB (B. caldotenax and T. thermophilus) and XPD (human) that served as the basis for the development of the XPD model. The helicase motifs are colored corresponding to their color in the UvrB and XPD structures shown.
erase was purchased from Stratagene. All UvrB mutants were con-
structed by PCR using QuickChange site-directed mutagensis method
(Stratagene) with Pfu Turbo DNA polymerase, pUC18TurboB<sup>Bsu</sup> as
a template DNA and two oligonucleotide primers (Genosys), each com-
plementary to opposite strands of the template DNA, containing the
desired mutation. Resulting mutant plasmids were verified by sequenc-
ing for the absence of additional PCR-induced mutations in the uvrB
gene sequence. Mutated uvrB gene was finally subcloned into pTYB1
expression vector.

**DNA Substrates**—Fluorescein-containing DNA substrates were syn-
thesized by Sigma Genosys. The DNA sequence of a 50-bp dsDNA
substrate containing a single internal fluorescein adduct (F<sub>exo</sub>-50
dsDNA) is shown in Fig. 2A. For 5' labeling 10 pmol of 50-mer fluores-
cein-containing top strand was incubated with 25 units of T4 polynu-
cleotide kinase in 70 mM Tris/Cl (pH 7.6), 10 mM MgCl<sub>2</sub>, 100 mM
KCl, 1 mM 2-mercaptoethanol, and 15 pmol of [γ<sup>32</sup>P]ATP (3000 Ci/mmol).
After incubation at 37 °C for 1 h, the reaction was terminated by
incubation at 80 °C for 10 min in the presence of 20 mM EDTA. Anneal-
ing of the top and the bottom strand was performed in the presence of
50 mM NaCl, followed by purification through Bio-Spin P-30 polyacryl-
amide gel column (Bio-Rad) for removal of unincorporated nucleotides.
The double-stranded character and homogeneity of the 50 bp substrate
were examined by a restriction assay (38) and analyzed on a 10%
polyacrylamide sequencing gel under denaturing conditions. The cho-
lesterol containing substrate was a generous gift from N. Goosen
(Leiden, The Netherlands) (52).

The DNA sequence of the helicase substrate (HSIF-M13mp19) was
described previously (26). Five pmol of a 26-mer containing an internal
fluorescein adduct (HSIF) were labeled at its 5'-terminus under the
same conditions as the F<sub>exo</sub>-50 top strand. The helicase substrate was
constructed by hybridizing 0.4 pmol of 5'-labeled HSIF oligonucleotide
with equimolar amounts of M13mp19 (++) strand (Invitrogen) and pu-
rified as described above.

**Gel Mobility Shift Assay**—Binding reactions were performed with 2
nM DNA substrate (5'-<sup>32</sup>P-labeled fluorescein-50 dsDNA) with 20 nM
*B. caldotenax* UvrA and 60 nM *B. caldotenax* UvrB in 20 μl of UvrABC
buffer (50 mM Tris/Cl (pH 7.5), 10 mM MgCl<sub>2</sub>, 50 mM KCl, 1 mM ATP, 5
mM DTT) for 20 min at 55 °C. Glycerol was then added to the reaction
(8% v/v), and the reaction mixture was loaded onto a 4% native poly-
acrylamide gel (80:1). The gel and the running buffer (89 mM Tris, 89
mM boric acid, 2 mM EDTA) contained 1 mM ATP and 10 mM MgCl<sub>2</sub>. The
electrophoresis was performed for 3 h at 100 V at room temperature.
The gel was dried and exposed against Storage Phosphor Screen (Am-
ersham Biosciences) overnight at room temperature.

**UvrABC Incision Assay**—The 5'-terminally labeled cholesterol-50 or
fluorescein-50 dsDNA substrate (2 nM) was incubated with UvrABC (20 nM
UvrA, 60 nM UvrB, and 50 nM UvrC) in 20 μl of UvrABC buffer at 55
°C for 1 h. The reaction was terminated by ethanol precipitation. The
samples were denatured with formamide and heated to 90 °C for 5 min
and then quick-chilled on ice. The incision products were analyzed by
electrophoresis on a 10% polyacrylamide gel under denaturating
conditions at 400 V with TBE buffer. In the case of the helicase
substrate incision, the reaction mixture (15 μl) contained ~8
fmol (in ssDNA circles) of DNA substrate, 50 nM UvrA, 100 nM UvrB,
and 100 nM UvrC in buffer A2 (50 mM Tris/Cl (pH 7.5), 100 mM KCl, 15
mM MgCl<sub>2</sub>, 1 mM EDTA, 5 mM ATP, 2 mM DTT) and was incubated at
42 °C for 1 h. The reaction was quenched with 5 μl of stop solution (25%
(v/v) Ficoll, 1% SDS, 0.1 mM EDTA, 0.25% orange G) and heated for 2
min at 85 °C, and the entire sample was then loaded onto a 15%
denaturing polyacrylamide gel equilibrated with TBE running buffer.
Electrophoresis was carried out at 400–500 V for 1–2 h. The gels were
processed as described above.

**Oligonucleotide Destabilizing Assay**—The reaction mixture con-
tained 50 nM UvrA, 100 nM UvrB, and ~8 fmol (in ssDNA circles) of
helicase substrate (HSIF-M13mp19) in buffer A1 (50 mM Tris/Cl (pH 7.5),
150 mM NaCl, 10 mM MgCl<sub>2</sub>, 2 mM ATP, 5 mM DTT) and was
incubated at 37 °C for various time intervals. The reaction was stopped
with 5 μl of stop solution (50% (v/v) glycerol, 1% SDS, 0.1 mM EDTA,
0.25% orange G), and the entire sample was loaded on 10% non-den-
aturing acrylamide gel in TBE running buffer. Electrophoresis was run
at 120–150 V for 1–2 h, and the gels were processed as described
previously.

**ATP Hydrolysis Assay**—The conversion of ATP to ADP by the
UvrABC system was determined by a coupled enzyme assay system
consisting of pyruvate kinase and lactate dehydrogenase to link the
hydrolysis of ATP to the oxidation of NADH. The assay mixture con-
sisted of 50 mM Tris/Cl (pH 7.5), 50 mM NaCl, 4 mM MgCl<sub>2</sub>, 1 mM
DTT, 20 units/ml lactate dehydrogenase, 20 units/ml pyruvate kinase, 2 mM
phosphoenolpyruvate, 0.15 mM NADH, and 200 mM Uvr proteins in the
presence or absence of 50 ng of UV-irradiated DNA substrate. The
continuous assay was performed at 55 °C for 60 min; the ATPase
activity was determined from the slope of the turnover of NADH versus
time.

**RESULTS AND DISCUSSION**

**The XPD Model**

The structure of the molecular model developed for the hu-
man XPD protein, based on the x-ray crystal structure of *B.
caldotenax* UvrB (23, 24), is shown in Fig. 2B. This model
includes all the defined SF2 helicase motifs, domains 1a, 1b, 2,
and 3, and structure for XPD residues 10–282 and 448–712
based on the UvrB structure coordinates as the template. The
model of the additional XPD domain, residues 283–448, absent
in UvrB, and location relative to the helicase motifs are indi-
cated in Fig. 3. The XPD C-terminal model structure (residues
713–761) and location relative to the helicase motifs are shown
in Fig. 3. Fig. 4 indicates the location of all the currently known
disease (XP, TTD, and CS)-associated mutations found in XPD
on the structural model. It is interesting to note that the
majority of the disease-associated mutations are found in do-
main 1a and 3 whose structure is highly conserved across all
helicases. Other frequently occurring disease-causing muta-

![FIG. 2. A, the solved x-ray crystal structure of B. caldotenax UvrB. Domains 1a, 1b, 2, and 3 are shown in yellow, green, dark blue, and red, respectively. The β-hairpin bridging the gap between domains 1a and 1b is shown in cyan. The ATP molecule is shown in stick format (colored by atom type) bound at the interface between domains 1 and 3. B, structural model developed for XPD based on the UvrB crystal structure. XPD domains are colored similarly to the UvrB domains. The 15 loops, which were inserted in the XPD protein model, are shown in black.](image-url)
tions are located in helicase motifs III, V, and VI (3, 6, 53, 54). No disease-associated mutations have been observed in the XPD-inserted threaded domain (residues 283–447).

Overall comparison of the modeled XPD structures with other solved monomeric helicase structures illustrates the striking conservation of the central $\beta$-pleated sheet structure of domain 1a with surrounding $\alpha$-helices common to NTPases as first evidenced in the structure of the nucleotide binding domain of eIF4A (55). This domain contains the Walker A and B motifs common to NTPases and is involved in ATP binding and hydrolysis. Domain 3 of the XPD model also conserves the central $\beta$-pleated sheet structure with surrounding $\alpha$-helices of
the other solved SF1 and SF2 family helicases. Significant differences between the XPD modeled structure and UvrB structure are observed in domain 2, with the insertion of a large loop in place of the antiparallel β-sheets in UvrB, and domain 1B where the α-helices of UvrB are absent in XPD. In addition, the β-hairpin believed to play an essential role in DNA binding in UvrB (26) is not as distinct and extended in the XPD model structure as the two proteins exhibited low sequence similarity in this region. The overall root mean square deviation between the backbone UvrB structure and XPD model (superposition of 3392 atoms) is 1.89 Å (excluding the XPD-inserted domain and C terminus).

Validation of XPD Model through Site-directed Mutagenesis of UvrB

Experimental Strategy, Selection of Mutants, and Biochemical Analyses—The validity of our modeling approach was tested through biochemical experiments with mutant UvrB proteins containing homologous changes as some of the disease-causing mutations in the XPD gene. Two categories of experimental mutations were made (Table I). One group of UvrB mutants was selected as they represent the residues analogous to the XPD disease-causing mutations (E110A/E110R, D338N, and E506A). The second group of mutants was selected based on their high conservation in UvrB, RAD3, and XPD. This second group of UvrB mutants (H341A, D510N/D510A) can be thought of as “predictive” in that they are based on their conservation to have functional significance. The structural locations of the UvrB mutants constructed and purified are shown in Fig. 3. The biochemical properties of these UvrB mutants were studied by incision (Fig. 5), gel mobility shift, oligonucleotide destabilization (Fig. 6), and ATP hydrolysis assays (Table II).

Mutations in Helicase Motif II—The D338N and H341A mutants were selected due to their location in the “DEAD” box Walker B highly conserved helicase II motif found in many NTPases. In addition, His-341 (motif II) is believed to interact with residues in motif VI and III (56). Asp-338 corresponds in sequence alignment with XPD to a high prevalence D234N XP disease-causing mutation. Overall, the UvrB mutations in the helicase motif regions identical to XPD disease-causing mutations had reduced DNA repair capacity. Previously, XPD disease-causing mutations (XPDPD675R, R722W) have been isolated and purified and have been shown experimentally to perform poorly in helicase activity assays (50). The RAD3 K48R ATP-binding site mutant, which is part of the GXXGKT Walker A Box, helicase I motif, has been shown to lose ATPase and DNA helicase activities but still binds ATP (57). This same mutation has been made previously in E. coli UvrB and shown to reduce greatly its ATPase activity (58). Helicase SF2 members contain the conserved consensus “DEAD/H” box in helicase motif II. The first (Asp) residue of this motif binds the Mg\(^{2+}\) ion; the second residue (Glu) is proposed to bind water during ATP hydrolysis. Studies in other helicases have shown mutations in motif II to be ATPase-deficient and incapable of DNA unwinding (7). The UvrB D338N mutant 1) is strongly impaired in its ability to support incision of damaged fluorescein or cholesterol-containing DNA (Fig. 5); 2) did not form a complex with DNA (Fig. 6A); 3) is defective in the destabilization of a damaged oligonucleotide (Fig. 6B); and 4) displayed significantly reduced ATP hydrolysis activity (Table II). The D338A mutation has been reported previously (59) in E. coli UvrB, although it was made under the assumption that the protein was functioning as a nuclease and not to test unwind activity. In contrast mutating the nearby His-341 to alanine (conserved in 29/29 bacterial species, Table I) in helicase motif II resulted in a milder defect in the four functional assays.

Mutations in Helicase Motif V—The UvrB Arg-506 mutation corresponds to the high prevalence R601L/R601W XP disease-causing mutation. Mutating this residue to alanine had a mild phenotype in the four biochemical assays. However, mutating the nearby Asp-510 residue to alanine (corresponding to XPD Asp-609), in the highly conserved helicase V motif, resulted in drastic reduction in activity across all the biochemical assays. The Asp-510 residue was mutated because it is part of a highly conserved motif found in helicase motif V, EG(D/L)/D that is

| XPD | Disease | UvrB | Conservation | Helicase motif |
|-----|---------|------|--------------|---------------|
| R12H | TTD     | E110A/R | 29/29 E, 1D  |               |
| D234N | XPD    | D338N | 29/29 D     | II            |
| His-237 |       | His-341 | 29/29 H, II |               |
| R601L/R601W | XPD | R506A | 29/29 R     | V             |
| D609A |         | D510A/D510A | 29/29 | V             |

Fig. 5. Incision activity of UvrB carrying XPD-like mutations. A, 5′ incision of 50-bp duplex containing fluorescein (2 nM) by UvrA (20 nM), UvrB (60 nM), and UvrC (50 nM). B, histogram showing incision data of 50 bp containing fluorescein, mean ± S.D. of three separate experiments. C, histogram of incision data of 50-bp duplex containing cholesterol lesion, mean ± S.D. of three separate experiments. wt, wild type.

TABLE I

XPD mutations and corresponding residues in UvrB

Mutations His-237 and Asp-609 have not been detected in XPD patients; however, both are highly conserved. His-237 is part of a highly conserved ATP-binding motif (Walker B/motif II) in both proteins: XPD – AspD EAH; UvrB – AspD EAH. Asp-609 is part of a highly conserved sequence found in helicase motif V EG(D/L)/D conserved in UvrB through to RAD3 and XPD. Conservation in 29 bacterial species.
conserved from UvrB to RAD3 (*S. cerevisiae*) and XPD (human). The Glu residue of this motif is believed to be involved in long distance interactions across the protein with motif Ia (56). This mutation was reported previously by Lin et al. (59) in *E. coli* UvrB and shown to reduce its activity. Other residues in helicase motifs V and VI have been shown to be important for ATP hydrolysis; UvrB cannot hydrolyze ATP when Arg-540 and Arg-543 are mutated (24). It is interesting to note that the D510N mutation has a mild phenotype in most biochemical assays but has a significantly higher UvrA/UV DNA-stimulated ATPase activity than the wild type UvrB.

**TTD-like Mutation in the β-Hairpin**—The E110A/E110R UvrB mutant was selected for analysis due to its correspondence (in location according to the sequence alignment) to the R112H high prevalence TTD disease-causing mutant and location in the UvrB β-hairpin proposed to participate in DNA-binding interactions (23, 24, 26, 52). The Glu-110 to arginine or alanine UvrB mutations behaved the same as wild type protein in all assays and did not exhibit any deficiency in the ability to bind, cleave, or destabilize oligomeric DNA. Surprisingly, this mutation which lies outside the ATP binding domain leads to a significant increase in the UvrA/UV DNA-stimulated ATPase activity. Although this mutation was selected because it corresponded to the XPD R112H TTD-causing mutation, this region of the protein, however, shows little sequence conservation between the UvrB and XPD proteins. The β-hairpin turn has already been proposed to be one of the structural loop regions of the XPD model predicted with lower reliability than the rest of the structure and where differences in behavior between UvrB and XPD were anticipated. Additionally, the positively charged arginine residue would be expected to exhibit different biochemical behavior than the negatively charged glutamic acid. The Arg-112 XPD mutation is a TTD disease-causing mutation unlike the others selected which are XP disease-causing mutations. It is possible because TTD is viewed as a disease of transcription and not NER that some other mechanism is involved in the TTD disease syndrome for which the bacterial NER enzyme UvrB is an inappropriate model (6).

**CONCLUSIONS**

The goal of these studies was to construct a molecular model of human DNA repair protein XPD based on the experi-

| Table II | ATPase activity in UvrB mutants |
|----------|-------------------------------|
| Corresponding XPD residue | Sample composition | Activity (µmol/min/µg protein) |
| UvrA | | 20.45 ± 2.3 |
| UvrB (wt)* | | 7.8 ± 0.4 |
| UvrA + UV DNA | | 58.5 ± 2.5 |
| UvrA + UvrB (wt) | | 17 ± 0.3 |
| UvrA + UvrB (wt) + UV DNA | | 82.7 ± 0.2 |
| Arg-112 | | |
| E110R (UvrB mutant) | | 7.7 ± 0.1 |
| E110R + UvrA + UV DNA | | 99 ± 6 |
| His-237 | | |
| H341A (UvrB mutant) | | 2.35 ± 1 |
| H341A + UvrA + UV DNA | | 65.5 ± 2 |
| Asp-234 | | |
| D358N (UvrB mutant) | | ND |
| D358N + UvrA + UV DNA | | 31 ± 1.3 |
| Arg-601 | | |
| R506A (UvrB mutant) | | ND |
| R506A + UvrA + UV DNA | | 40 ± 1.7 |
| Asp-609 | | |
| D510A (UvrB mutant) | | ND |
| D510A + UvrA + UV DNA | | 29 ± 1.6 |
| Asp-609 | | |
| D510N (UvrB mutant) | | 10 ± 1 |
| D510N + UvrA + UV DNA | | 104 ± 3.2 |

* wt, wild type.
+ p < 0.001 significantly different compared with UvrA + UvrB (wt).
+ p < 0.05 significantly different compared with UvrA + UvrB (wt).
+ ND, none detected over background ATP hydrolysis (~2.4 µmol/min).
tally determined crystal structure of the bacterial DNA repair protein, UvrB. Model building was validated by experimental verification of XPD mutations in corresponding residues in the XPB protein. The XPD model helps to give structural insights into some of the mutations in the XPD gene which when mutated yield the XP phenotype. For example, one prediction of the XPD model is that Arg-683, which is located on the “backside” of XPD on domain 3, is part of a binding surface for interaction with p44. This residue is in close proximity to another mutation, D681N, which is the most common XPD mutation leading to TTD, XP, or CS phenotypes.

In summary, we found that XP disease-like mutations could be successfully mimicked experimentally in UvrB (D383N, H341A, and R506A) and had a greater impact on the activity of UvrB than a TTD disease-like mutation (E110A/E110R). This model, and the concept that XPD and UvrB both play a vital step in damage-strand verification, is supported by recent DNA cross-linking experiments with XPD (60). Reardon and Sancar (60) show that XPD is making intimate contact with a psoralen cross-linking experiments with XPD (60). Reardon and Sancar (60)

REFERENCES

1. Eisen, J. A., and Hanawalt, P. C. (1999) *Mutat. Res.* 435, 171–213
2. Van Houten, B., Eisen, J. A., and Hanawalt, P. C. (2002) *Proc. Natl. Acad. Sci. U. S. A.* 99, 2581–2583
3. Botta, E., Nardo, T., Broughton, B. C., Marini, S., Lehmann, A. R., and Stefani, M. (1998) *Am. J. Hum. Genet.* 63, 1036–1048
4. Friedberg, E. C. (2001) *Nat. Rev. Cancer* 1, 22–33
5. Hoeijmakers, J. H. (2001) *Nature* 411, 366–374
6. Lehmann, A. R. (2001) *Genes Dev.* 15, 15–23
7. Hall, M. C., and Matson, S. W. (1999) *Mol. Microbiol.* 34, 867–877
8. Cleaver, J. E. (2001) *Mutat. Res.* 465, 25–36
9. Lehmann, A. R. (2001) *Genes Dev.* 15, 15–23
10. Ellis, N. A. (1997) *Curr. Opin. Genet. & Dev.* 7, 354–363
11. van Brabant, A. V., Stan, R., and Ellis, N. A. (2000) *Annu. Rev. Genomics Hum. Genet.* 1, 409–459
12. Gorbaneva, A., and Koynin, E. (1993) *Curr. Opin. Struct. Biol.* 3, 419–429
13. Korslev, S., Heide, J., Siggia, G. H., Lehman, T. M., and Waksman, G. (1997) *Cell* 89, 635–647
14. Korslev, S., Yao, N., Lehman, T. M., Weber, P. C., and Waksman, G. (1998) *Protein Sci.* 7, 605–610
15. Soultanas, P., Dillingham, M. S., Velankar, S. S., and Wigley, D. B. (1999) *J. Mol. Biol.* 290, 137–148
16. Subramanya, H. S., Bird, L. E., Brannigan, J. A., and Wigley, D. B. (1996) *Nature* 384, 379–383
17. Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S., and Wigley, D. B. (1999) *Cell* 97, 75–84