Molecular Cell, Volume *29*

Supplemental Data

Crystal Structure of *Bacillus stearothermophilus* UvrA Provides Insight into ATP-Modulated Dimerization, UvrB Interaction, and DNA Binding

Danaya Pakotiprapha, Yoshihiko Inuzuka, Brian R. Bowman, Geri F. Moolenaar, Nora Goosen, David Jeruzalmi, and Gregory L. Verdine

Supplemental Experimental Procedures

Expression of *Bacillus stearothermophilus* UvrA, UvrB, and UvrC

The genes for *Bacillus stearothermophilus* UvrA, UvrB, and UvrC were identified by searching with the *Escherichia coli* sequences against the genome of *B. stearothermophilus* strain 10 (The University of Oklahoma Advance Center for Genome Technology). PCR was used to amplify the genes from genomic DNA, and the resulting PCR products were cloned into pET-28a (+) (Novagen) (Table S1). The expression constructs contained full-length Uvr proteins, an N-terminal His_6 tag and a thrombin cleavage site and were confirmed by sequencing. In the present work, protein residues are numbered starting from the initiator methionine in the wild-type sequence.

UvrA, UvrB, and UvrC proteins were expressed in *E. coli* BL21(DE3) pLysS. The cells were grown in LB broth at 37°C until OD_{600} reached 0.5-0.6, at which point expression was induced by addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The cells were allowed to grow at 30°C for 4 hours and harvested by centrifugation. The cell pellet was resuspended in lysis buffer (50 mM $NaPO_4$ pH 8.0, 500 mM NaCl, 10 mM imidazole, 5 mM β-ME), flash frozen in liquid nitrogen, and stored at -80°C.

Selenomethionine-substituted UvrA was expressed in *E. coli* BL21(DE3) pLysS using a protocol adapted from Van Duyne *et al* (Van Duyne et al., 1993). Cells were

grown in M9 minimal medium at 37°C until OD_{600} reached 0.5-0.6, at which point amino acid mix and selenomethionine were added. After additional 30 minutes at 30°C, expression was induced with 1 mM IPTG. The cells were allowed to grow at 30°C for 14 hours and processed as the native protein except that lysis buffer contained 15 mM β-ME.

Protein Purification

Thawed cells were supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and lysed by sonication. The clarified cell lysate was incubated for 45 minutes with nickel-nitrilotriacetic acid (Ni-NTA) agarose beads (Qiagen). Unbound proteins were washed away with 20 mM imidazole in 50 mM $NaPO_4$ pH 8.0, 500 mM NaCl, 5 mM β-ME. Partially purified UvrA was eluted with 250 mM imidazole in the same buffer.

UvrA was further purified using Heparin HiTrap column (GE Healthcare) with a linear gradient of 0.3-1.2 M NaCl in Buffer A (25 mM Tris·HCl pH 7.4, 5 mM β-ME), and size exclusion chromatography (Superdex200, GE Healthcare) using 400 mM NaCl in Buffer A. Ni-NTA purified UvrB was buffer-exchanged into 100 mM NaCl in buffer A before further purified over Mono-Q (GE Healthcare) using a linear gradient of 100-600 mM NaCl in buffer A and completed using size exclusion chromatography (Superdex200, GE Healthcare) using 400 mM NaCl in Buffer A. UvrC was purified in the same way as UvrA except that 0.5-1.5 M NaCl linear gradient was used for Heparin HiTrap column and 0.5 M NaCl in Buffer A was used for the size exclusion step. Isolated domain constructs of UvrA and UvrB (UvrA131-245 and UvrB149-250) were purified by Ni-NTA and size exclusion chromatography (Superdex75, GE Healthcare, 200 mM NaCl in Buffer A).

For biochemical experiments, the proteins were dialyzed into 25 mM Tris-HCl pH 7.4, 250 mM NaCl, 20% (v/v) glycerol, 5 mM β-ME, flash frozen in liquid nitrogen and stored at -80˚C.

ATPase Assay

The rate of ATP hydrolysis was measured using a coupled enzyme assay system consisting of pyruvate kinase and lactate dehydrogenase, in which ATP hydrolysis is coupled to the oxidation of NADH (Kiianitsa et al., 2003). The reaction mixture (200 µl) contained 50 mM K-HEPES pH 7.5, 150 mM potassium acetate, 8 mM magnesium acetate, 5 mM β-ME, 250 µg/ml bovine serum albumin (BSA), 0.3 mM NADH, 2 mM PEP, 1.5 U pyruvate kinase, 1.8 U lactate dehydrogenase, 50-200 mM *Bst*UvrA and 1-2 mM ATP. The mixture was incubated at 55°C for 5 minutes before addition of ATP. The reaction was followed by monitoring the decrease in A_{340} over a 30-minute period using SpectraMax M5 microplate reader (Molecular Devices). The rate of ATP hydrolysis was calculated from the linear change in A_{340}, with correction for NADH oxidation in the absence of protein. The data are reported as mean turnover number (k_{cat}) (mol ATP/min/mol UvrA) ± standard error of the mean (n=8). The effect of DNA on ATP hydrolysis was studied by adding 0.2 µM 50-mer dsDNA containing fluorescein at the central position to the assay mixture.

Incision and Electrophoretic Mobility Shift Assays (EMSA)

The incision assay employed a 50-bp DNA fragment containing an N3-menthol lesion (Verhoeven et al., 2002) that was labeled at the 5′ end of the damaged strand using [γ-^{32}P]-ATP and T4 polynucleotide kinase. Damaged DNA fragment (0.2 nM) was incubated with 0.675 nM UvrA, 50 nM UvrB, and 12.5 nM UvrC in 20 µl Uvr-endo buffer (50mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM KCl, 0.1 µg/µl BSA, 1 mM ATP) for 15 minutes at 57°C. The reactions were terminated by using 3 µl EDTA/SDS (0.33 M EDTA, 3.3% SDS) and 2.4 µl glycogen (4 µg/µl) followed by ethanol precipitation. The incision products were visualized on a 15% denaturing polyacrylamide gel.

Loading of UvrB onto damaged DNA was measured using EMSA. Damaged DNA (0.2 nM) was incubated with 2 nM UvrA and 100 nM UvrB in 10 µl Uvr-endo buffer for 10 minutes at 57°C. The reaction mixtures were analyzed on a cooled 3.5% native polyacrylamide gel containing 1 mM ATP and 10 mM $MgCl_2$ in 1X Tris-borate/EDTA as described (Visse et al., 1992).

Analysis of DNA binding by UvrA mutants was carried out using EMSA. Fluorescein-containing 50-mer DNA (0.2 nM) was titrated with UvrA in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM ATP, 5% glycerol, 0.1 mg/mg BSA at 55°C for 30 minutes and then chilled on ice. The reaction mixtures were analyzed on a 6% native polyacrylamide gel containing 1 mM ATP and 10 mM $MgCl_2$ in 1X Tris-borate/EDTA. For the reactions with DNA containing N3-menthol lesion (Verhoeven et al., 2002), UvrA was preincubated at 57°C in the Uvr-endo buffer for 5 minutes prior to addition of DNA. The reactions were incubated at 57°C for 10 minutes, chilled on ice, and analyzed on a cooled 3.5% native polyacrylamide gel containing 1 mM ATP and 10 mM $MgCl_2$ in 1X Tris-borate/EDTA.

Supplemental References

Kiianitsa, K., Solinger, J.A., and Heyer, W.D. (2003). NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. Anal Biochem *321*, 266-271.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680.

Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol *229*, 105-124.

Verhoeven, E.E., van Kesteren, M., Turner, J.J., van der Marel, G.A., van Boom, J.H., Moolenaar, G.F., and Goosen, N. (2002). The C-terminal region of Escherichia coli UvrC contributes to the flexibility of the UvrABC nucleotide excision repair system. Nucleic Acids Res *30*, 2492-2500.

Visse, R., de Ruijter, M., Moolenaar, G.F., and van de Putte, P. (1992). Analysis of UvrABC endonuclease reaction intermediates on cisplatin-damaged DNA using mobility shift gel electrophoresis. J Biol Chem *267*, 6736-6742.

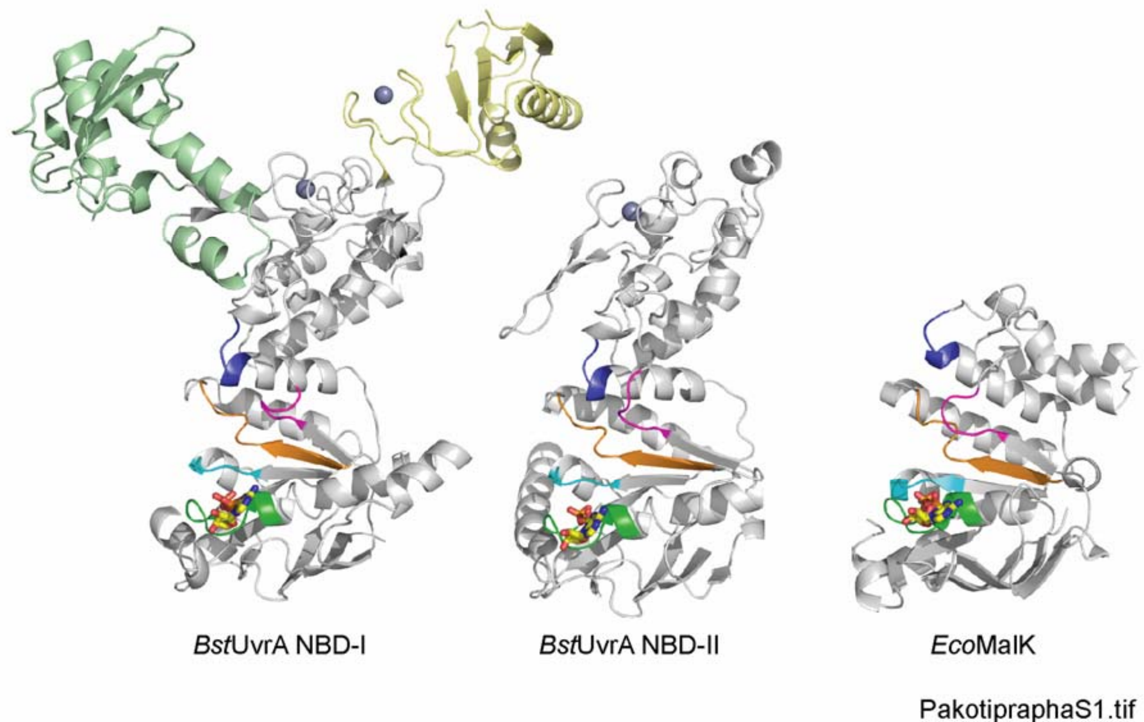
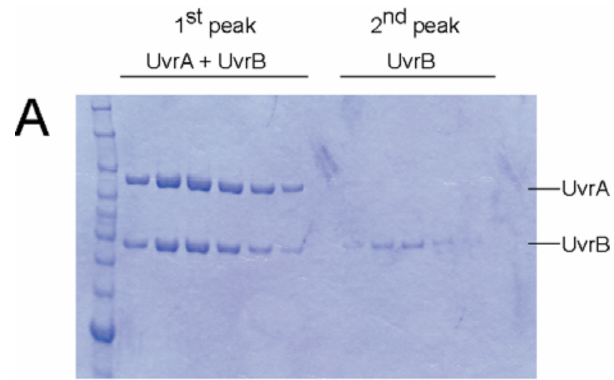


Figure S1. Structural comparison of the nucleotide-binding domains (NBDs) of *Bst*UvrA and *E.coli* MalK. NBD-I and NBD-II of UvrA were superimposed with the NBD of the maltose transporter MalK (ADP-bound, PDB code 2AWO) The conserved ATPase motifs are colored as follows: Walker A/P-loop, green; Walker B and D-loop, orange; ABC signature motif, blue; Q-loop, magenta; and H-loop/switch, cyan. For clarity, the structures are shown side-by-side. The bound ADP molecules are shown as sticks and the Zn atoms in the UvrA structure are shown as spheres. The UvrB-binding domain and the insertion domain, which are inserted in the NBD-I of UvrA, are shown in pale yellow and pale green, respectively.



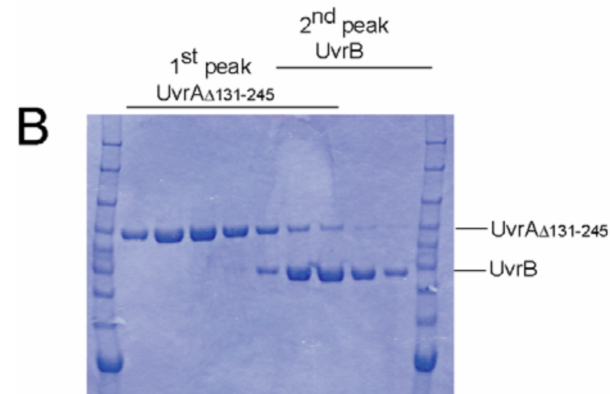


Figure S2. SDS gel electrophoretic analysis of size exclusion chromatography of the following: wild-type UvrA·UvrB or $UvrA_{\Delta 131\text{-}245}$·UvrB. We observe that only wild-type UvrA complexed with UvrB, whereas $UvrA_{\Delta 131\text{-}245}$ failed to interact with UvrB. The UV absorption profiles are shown in Figure 5C.

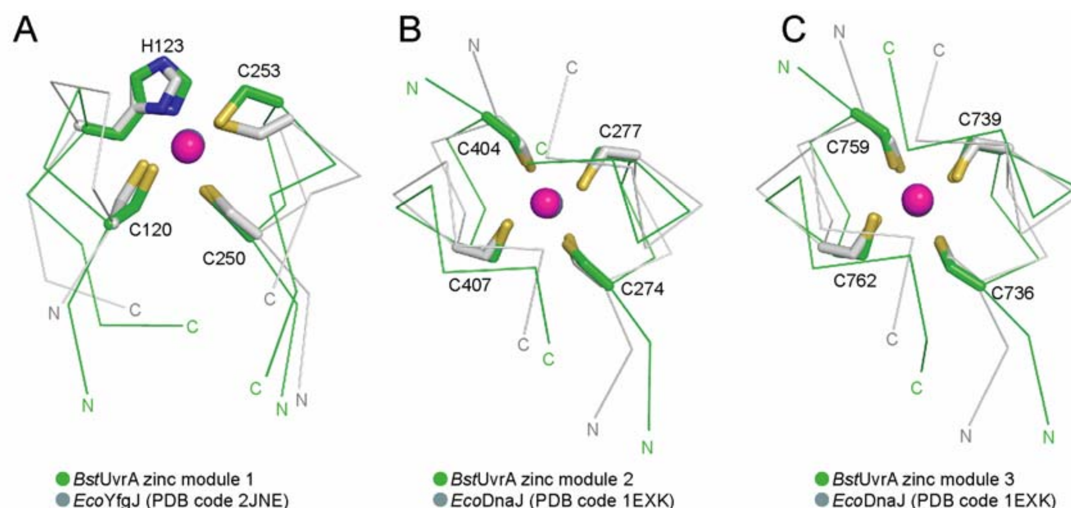


Figure S3. Zinc structural modules in UvrA. Each Zn structural module of UvrA (green) is shown in an overlay with its closest structural neighbor (gray). The zinc atoms in UvrA are colored magenta. (A) UvrA Zn module 1 and the Zn module from YfgJ (PDB code 2JNE), RMSD 0.798 Å. (B) and (C) UvrA Zn modules 2 and 3 and the Zn ribbon 2 from DnaJ (PDB code 1EXK), RMSD 0.655 and 0.574 Å, respectively. Residue numbers are indicated for UvrA. The amino and carboxy termini are labeled. The RMSD was calculated using Zn; S_γ, C_β, and C_α (for cysteine residues); N_δ, C_γ, C_β, and C_α (for histidine residues); as well as C_α of the intervening residues.

PakotipraphaS4.tif

Figure S4. Multiple sequence alignment of UvrA orthologs. Sequences of 135 UvrA orthologs were aligned using ClustalW v.1.83 (Thompson et al., 1994). Only sequences

from *Bacillus stearothermophilus* and four other bacteria in which NER is most extensively studied are shown: *Bst, Bacillus stearothermophilus; Bca, Bacillus caldotenax; Eco, Escherichia coli; Tth, Thermus thermophilus; Tma, Thermotoga maritima.* The secondary structure according to the *Bst*UvrA model is shown above the sequences, colored by domains as in Figure 1. Disordered regions are depicted as dashed lines. Locations of the conserved ABC ATPase motifs, glycine-rich loops, and zinc-coordinating residues are shown on the amino acid sequences.

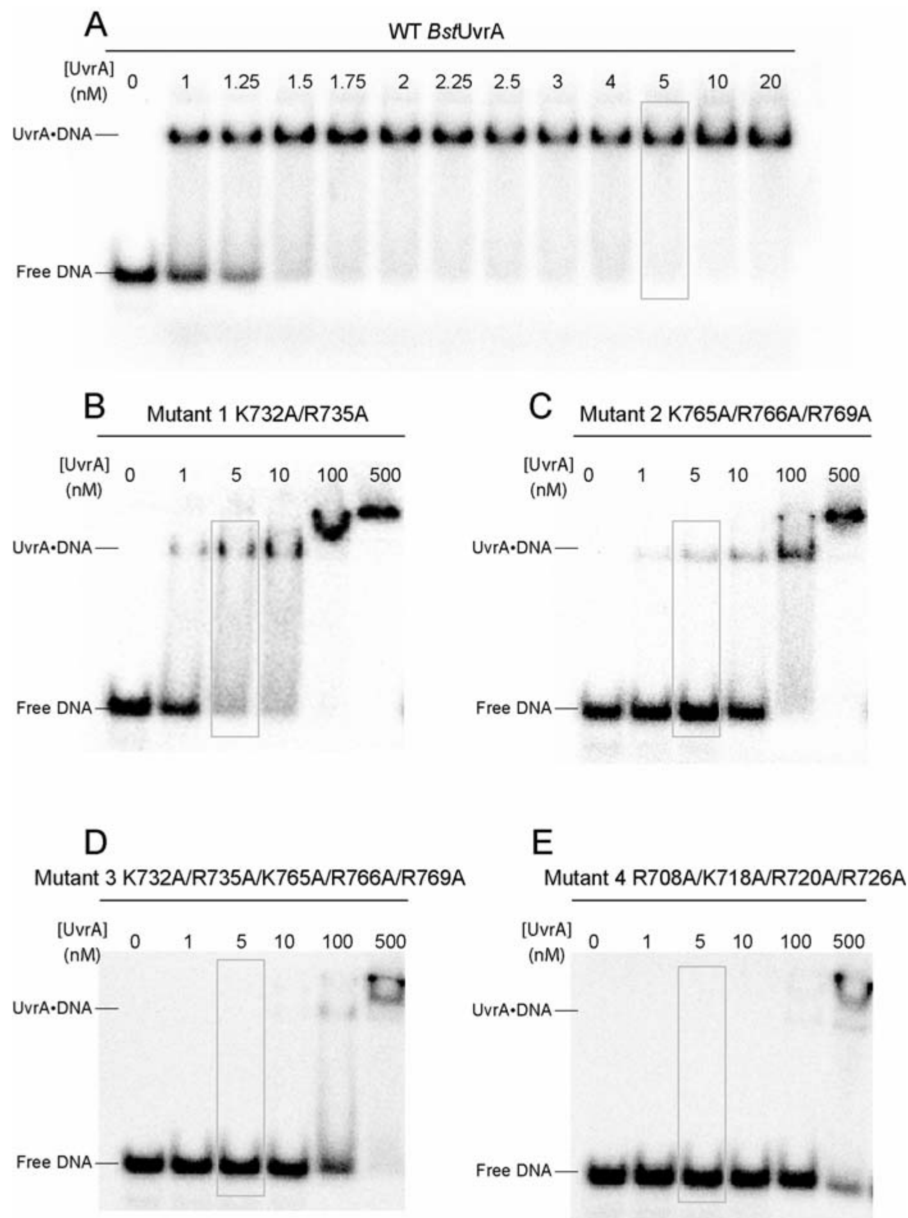


Figure S5 Binding of wild-type and mutant *Bst*UvrA to the 50-bp duplex containing the fluorescein lesion. The positions of the free DNA and the UvrA·DNA complex are indicated. Mutants 1, 2, 3, and 4 contain mutations in regions a, b, a+b, and c, respectively (see Figure 6C). The 5 nM experimental point (light gray boxes) prominently illustrates the difference in affinity for damaged DNA between the wild-type and mutant proteins.

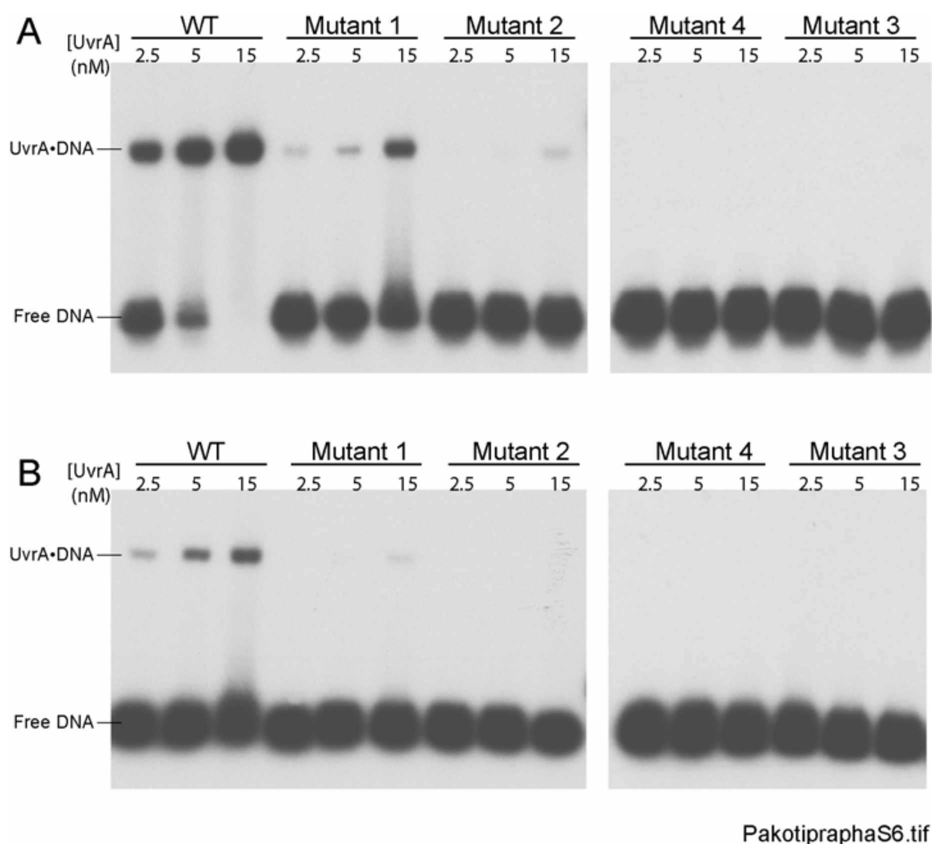


Figure S6 Binding of wild-type and mutant *Bst*UvrA to the 50-bp duplex containing the N3-menthol lesion (Verhoeven et al., 2002) (A) and the undamaged 50-bp duplex (B). The positions of the free DNA and the UvrA·DNA complex are indicated. Mutants 1, 2, 3, and 4 contain mutations in regions a, b, a+b, and c, respectively (see Figure 6C).

Table S1. Sequences of primers used in the amplification of *Bacillus stearothermophilus uvr* genes and the construction of UvrA mutants.

Primers	Sequences (5' → 3') [a, b]
uvrA fwd	GGTAATCCATATGGATAAAATTATCGTCAAAGGGGCGCGCGCCCACAACTTG
uvrA rev	GGTAATCAAGCTTTCACGCCTTCGCCGCTTCATACCGCGCCTGCATGCGCG
uvrB fwd	GGTAATCCATATGGGTCCGAAGAAAGTGGAGGGCCGTTTTCAATTAGTGTCGC
uvrB rev	GGTAATCAAGCTTTCACCCTTCCGCTTTCAATTCGAAAATGATATCGCGC
uvrC fwd	GGTAATCGCTAGCAAAAAGAACGAGCGGCTGAAAGAAAAGCTGG
uvrC rev	GGTAATCAAGCTTTCATTCATGCAGTTTTTCATAGATTTTCTCCGCCACC
UvrA$_{\Delta 131\text{-}245}$ fwd	CACGGCATTGAAATCCAATCGCAG**GGTACC**GAAAAGCACGCTTGTCCGTACTG
UvrA$_{\Delta I131\text{-}245}$ rev	CAGTACGGACAAGCGTGCTTTTC**GGTACC**CTGCGATTGGATTTCAATGCCGTG
UvrA$_{\Delta 285\text{-}400}$ fwd	CTGCGACGGGCTCGGGGCGAAGCTC**GGTACC**GAACAACCATGTCCGACATGCCAAG
UvrA$_{\Delta 285\text{-}400}$ rev	CTTGGCATGTCGGACATGGTTGTTC**GGTACC**GAGCTTCGCCCCGAGCCCGTCGCAG
UvrA 131-245 fwd	CGCGGCAGCCATATGCCCATTTGCCCGACGCAC
UvrA 131-245 rev	GGCCGCAAGCTTTTACGAAAAGCCGCAGTACGGAC
UvrB 149-250 fwd	CGCGGCAGCCATATGGGGTCGCCGGAAGAATATCGG
UvrB 149-250 rev	GGCCGCAAGCTTTTACACGAAGTGCGACGCCGG
UvrA-R708A fwd	ACCGGGGTGTTTGACGACATC**GCG**GATGTGTTTGCCTCGACGAAC
UvrA-R708A rev	GTTCGTCGAGGCAAACACATC**CGC**GATGTCGTCAAACACCCCGGT
UvrA-K718A/R720A/R726A fwd	GCCTCGACGAACGAAGCG**GCG**GTG**GCG**GGCTACAAAAAAGGG**GCG**TTCAGCTTCAATGTCAAAGG
UvrA-K718A/R720A/R726A rev	CCTTTGACATTGAAGCTGAA**CGC**CCCTTTTTTGTAGCC**CGC**CAC**CGC**CGCTTCGTTCGTCGAGGC
UvrA-K732A/R735A fwd	GGGCGGTTCAGCTTCAATGTC**GCG**GGCGGG**GCG**TGCGAGGCCTGCCATGGCGAT
UvrA-K732A/R735A rev	ATCGCCATGGCAGGCCTCGCA**CGC**CCCGCC**CGC**GACATTGAAGCTGAACCGCCC
UvrA-K765A/R766A/R769A fwd	CCGTGCGAAGTGTGCCACGGC**GCGGCG**TACAAC**GCG**GAGACGCTCGAGGTGACGTAT
UvrA-K765A/R766A/R769A rev	ATACGTCACCTCGAGCGTCTC**CGC**GTTGTA**CGCCGC**GCCGTGGCACACTTCGCACGG

[a] Recognition sequences of the restriction enzymes used for cloning of the PCR products are underlined: *uvrA* and *uvrB*, *Nde*I and *Hin*dIII; *uvrC*, *Nhe*I and *Hin*dIII

[b] The sequence encoding the GT dipeptide inserted in place of the deleted domains, and the positions of K→A, R→A mutations are in bold.