Supporting materials

Methods

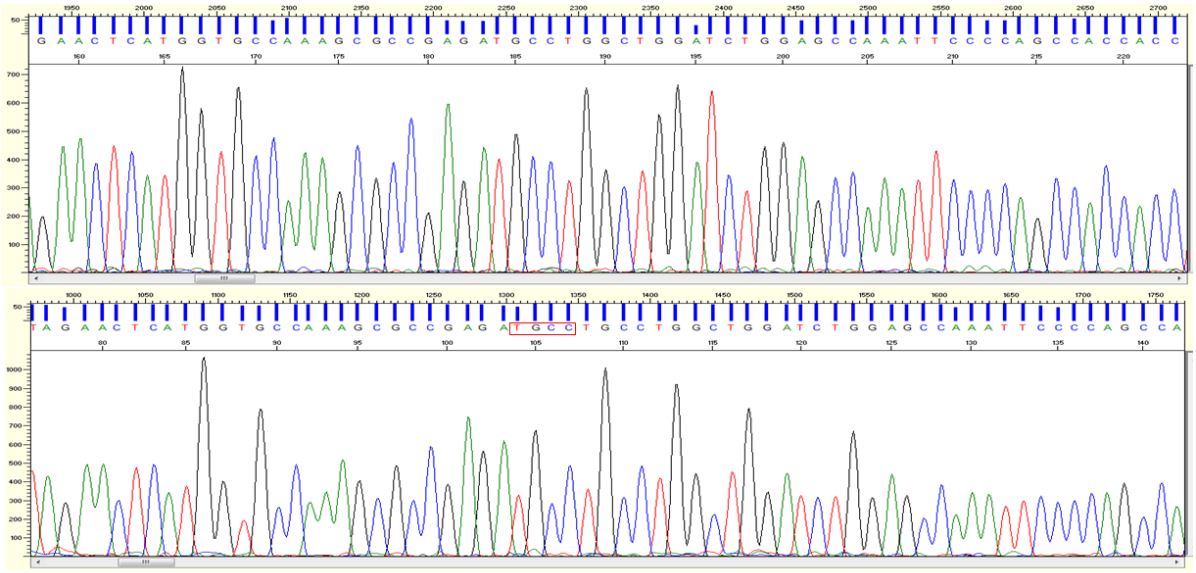
S1. Cloning, Expression, and Purification of XPB-C Recombinant Protein.

The cDNA sequence encoding the amino acid residues 494 through 782 of the human XPB was amplified by polymerase chain reaction (PCR) using Phusion® High-Fidelity DNA Polymerase (NEB BioLabs, USA) using the plasmid pOTB7-HsXPB (The CCSB Human ORFEOME collection) as the template and a pair of oligonucleotide primers HSXPB-F4 (5′-GGATCCGAGCTGCAGAATAATG-3′) and HSXPB-R3 (5′-GTCGACTAATTTCCTAAAGCGCTTG-3′). The underlined sequences in both primer sequences correspond, respectively, to the *Bam*HI and *Sal*I restriction sites. The mutagenesis to prepare the GST-XPBm(494-781) was constructed in a sequential two-step PCR reaction. The first PCR is to amplify two fragments labeled as A and B, respectively. The A fragment was amplified with primers HSXPB-F4 and XPBmutR (5'-GCGCCGAGATGCCTGCCTGGCTGGATCT-3') and the B fragment with primers XPBmutF (5'-AGATCCAGCCAGGCAGGCATCTCGGCGC-3') and HSXPB-R3. The underlined sequences in XPBmutF/R primer correspond to the mutation insertion. Amplified fragments were then gel purified. 100 ng of each fragment DNA were mixed together and incubated at 96°C for 2 min and at 25°C for 5 min. The recombinant strands were extended with Phusion® High-Fidelity DNA Polymerase at 72°C for 5 minutes. After that, primers HSXPB-F4 and HSXPB-R3 were added to amplify the entire mutant sequence (A+B) in the second PCR. The final PCR products corresponding to DNA encoding XPB-C and mutant XPBm(494-781) were cloned separately into the pMOSblue plasmid using the commercial cloning kit (GE Healthcare, USA). After DNA sequencing analysis, both fragments were cloned into the *Bam*HI and *Sal*I restriction sites of the *Escherichia coli* expression vector pGEX-6P1 (GE HealthCare, USA). Protein expression is under the control of the isopropyl-beta-D-thiogalactopyranoside (IPTG) inducible T7 promoter. The final plasmid constructions pGEX6P1-XPB-C and pGEX6P1-XPBm(494-781) were transformed into competent cells of *E. coli* strain Rosetta (DE3) pLys-S (Invitrogen, USA). A single colony was amplified in 10 mL of LB media containing 100 ug mL^{-1} ampicillin and 35 ug mL^{-1} chloramphenicol, in incubator at 37°C overnight shaking (250 rev min^{-1}). Cell aliquots were then used to inoculate larger volume cultures. Protein expression was induced at OD_{600}=0.6 after adding IPTG to a final concentration of 0.1 mM at 25°C. GST-XPB-C was expressed at 25°C, 225 rev

min^{-1}, and overnight expression. Cells from 6x 1L culture were harvested at 5,000 x g for 20 min at 4°C and pellet was re-suspended in 300 mL of buffer A (PBS 2x, pH 7.5, containing 5% glycerol, 0.5% nonidet NP-40, 2 mM EDTA, 2 mM EGTA, 10 mM benzamidine, 1 mM PMSF, 1 mM beta-glycerophosphate, 1 mM sodium pyrophosphate, 1 mM sodium vanadate), and immediately quick-frozen in liquid nitrogen. Cells were thawed in an ice water bath and disrupted by sonication at 100% output power level, 50% duty cycle, for 10 minutes (Branson Sonifier D450 with ½" disruptor horn). Supernatant fraction was clarified at 50,000 x g for 20 min at 4°C and loaded at a flow rate 2 mL min^{-1} onto a 5 mL Glutathione Sepharose column (GE Healthcare) using the Äkta Purifier UPC10 (GE Healthcare) at 8°C. Recombinant GST-tagged protein was eluted in buffer B (50 mM Tris-Cl, pH 8.0, 10 mM reduced glutathione, 100 mM NaCl, 5% glycerol, 1 mM EDTA, 10 mM benzamidine, 1 mM PMSF, 1 mM beta-glycerophosphate, 1 mM sodium pyrophosphate, 1 mM sodium vanadate). Aliquots of the eluate fractions were analyzed by SDS-PAGE 12% (w/v) and fractions containing the GST-XPB-C protein were pooled and concentrated in Amicon Ultra 30K filters (Millipore, USA) at 4°C and 3,000 x g. The concentrated sample was diluted in PreScission Protease buffer (GE Healthcare, USA) and digestion was carried out as described by the manufacturer recommendations. The digestion reaction mixture was passed through a GSTrap FF 5mL column and the flow-through fraction containing the XPB-C protein was collected and concentrated in Amicon Ultra 10K filters (Millipore, USA) at 4°C. Concentrated sample was loaded at the flow rate of 0.1 mL min^{-1} onto a HiPrep 16/60 Sephacryl S-100 High Resolution size exclusion column (GE Healthcare) previously equilibrated in buffer C (10 mM Tris-Cl, pH 8.0, containing 100 mM NaCl, and 5% glycerol). Samples corresponding to the peak fractions were analyzed by SDS-PAGE 15% (w/v). Fractions containing the pure HsXPB-C were pooled, concentrated in Amicon Ultra 10K filters, and stored in aliquots containing protein at 10 mg mL^{-1} at -80°C for crystal preparation.

Supplementary Figures

(a)



(b)

```
XPB     5′-GAACTCATGGTGCCAAAGCGCCGAGA----TGCCTGGCTGGATCTGGAGCCAAATTCCCCAGCCA-3′
XPBmut  5′-GAACTCATGGTGCCAAAGCGCCGAGATGCCTGCCTGGCTGGATCTGGAGCCAAATTCCCCAGCCA-3′
```

(c)

```
wt 736    R   S   S   Q   A       S   R   R   F   G   T   M   S   S   M   S
         AGA TCC AGC CAG GCA     TCT CGG CGC TTT GGC ACC ATG AGT TCT ATG TCT
mutant   AGA TCC AGC CAG GCA GGC ATC TCG GCG CTT TGG CAC CAT GAG TTC TAT GTC
   736   R   S   S   Q   A   G   I   S   A   L   W   H   H   E   F   Y   V

wt 752    G   A   D   D   T   V   Y   M   E   Y   H   S   S   R   S   K   A
         GGG GCC GAC GAC ACT GTG TAC ATG GAG TAC CAC TCA TCG CGG AGC AAG GCG
mutant   TGG GGC CGA CGA CAC TGT GTA CAT GGA GTA CCA CTC ATC GCG GAG CAA GGC
   753   W   G   R   R   H   C   V   H   G   V   P   L   I   A   E   Q   G

wt 769    P   S   K   H   V   H   P   L   F   K   R   F   R   K  782
         CCC AGC AAA CAT GTA CAC CCG CTC TTC AAG CGC TTT AGG AAA TGA
mutant   GCC CAG CAA ACA TGT ACA CCC GCT CTT CAA GCG CTT TAG GAA ATG A
   770   A   Q   Q   T   C   T   P   A   L   Q   A   L  781
```

Fig. S1. Sequence of XPBm(494-781) and XPB-C cloning. (**a**) Chromatograms of DNA sequencing results for XPB(497-782) (top panel) and XPBm(494-781) (bottom panel) at the 4-bp insertion region for the non-coding strand. The 4-bp insertion is framed in red in the bottom panel. (**b**) The DNA sequence of the non-coding strand for XPB-C (XPB) and XPBm(494-781) (XPBmut) at the 4-bp insertion (red letters) region. (**c**) Both DNA and amino acid sequences for XPB-C (wt) and XPBm(494-781) (mutant) for the region starting from residue 736 until the ending residue. The residues after the frameshift in XPBm(494-781) are colored in red. The phosphorylation residue Ser751 in the XPB-C sequence is highlighted in purple with a yellow letter.

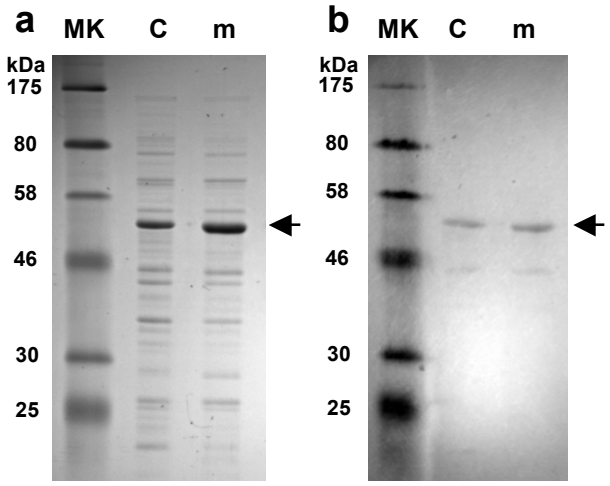


Fig. S2. Western blot analysis of GST-XPB-C and GST-XPBm(494-781) using anti-XPB-C serum. Normalized crude extracts (1μg/lane) containing the recombinant GST-XPB-C (C) or GST-XPBm(494-781) (m) were separated by 12% SDS-PAGE. (**a**) Coomassie blue stained gel. (**b**) Western blot results with rabbit serum against XPB-C and goat anti-rabbit IgG antibody conjugated with alkaline phosphatase (Pierce, USA). Pre-stained protein markers (MK) were used for (a) and (b). Arrows indicate the positions of targeted proteins.