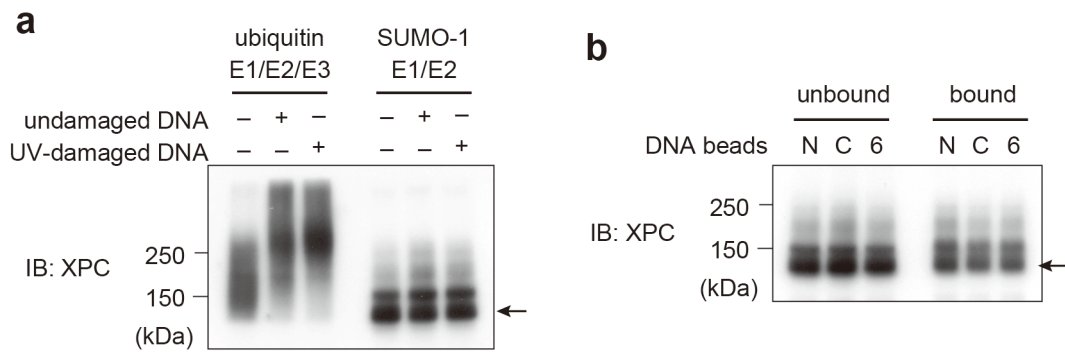


SUPPLEMENTARY INFORMATION

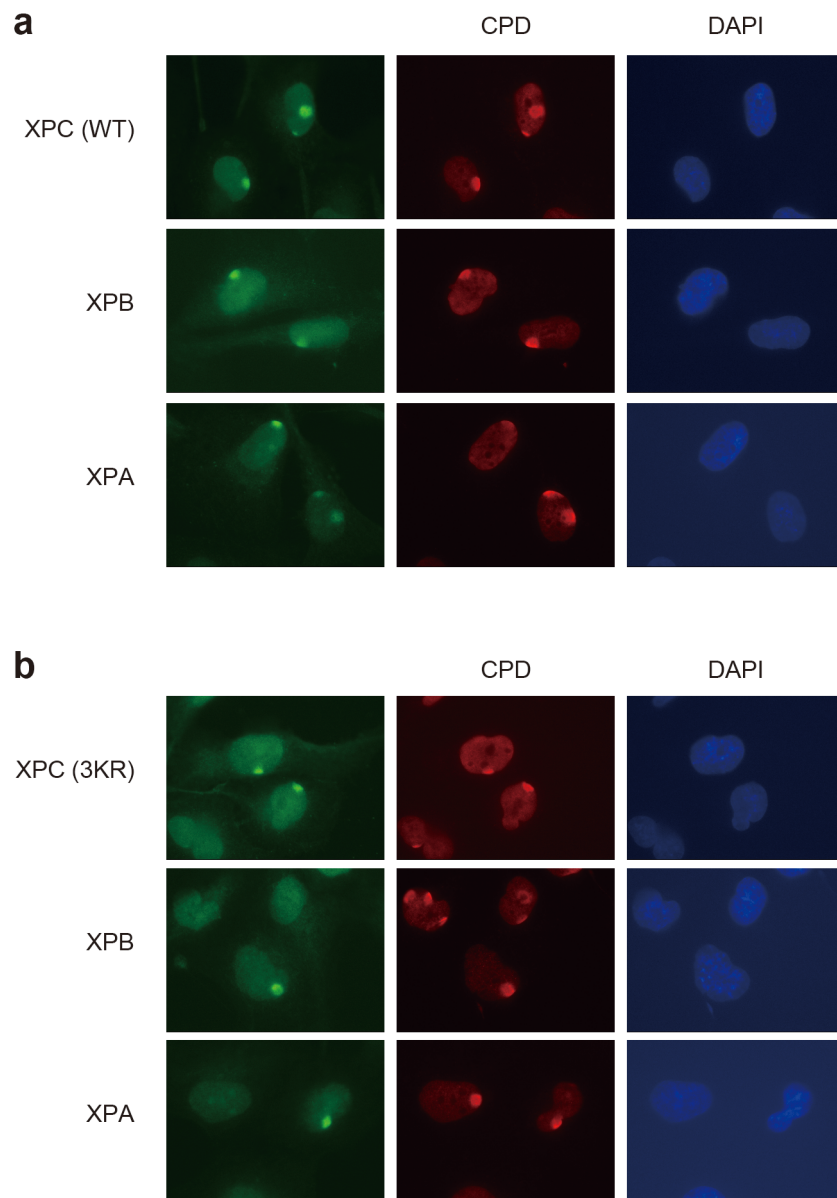
SUMOylation of xeroderma pigmentosum group C protein regulates DNA damage recognition during nucleotide excision repair

Masaki Akita¹, Yon-Soo Tak¹, Tsutomu Shimura¹, Syota Matsumoto^{1,2}, Yuki Okuda-Shimizu³, Yuichiro Shimizu³, Ryotaro Nishi^{1,3}, Hisato Saitoh⁴, Shigenori Iwai⁵, Toshio Mori⁶, Tsuyoshi Ikura⁷, Wataru Sakai^{1,2}, Fumio Hanaoka^{3,8} & Kaoru Sugasawa^{1,2,3}

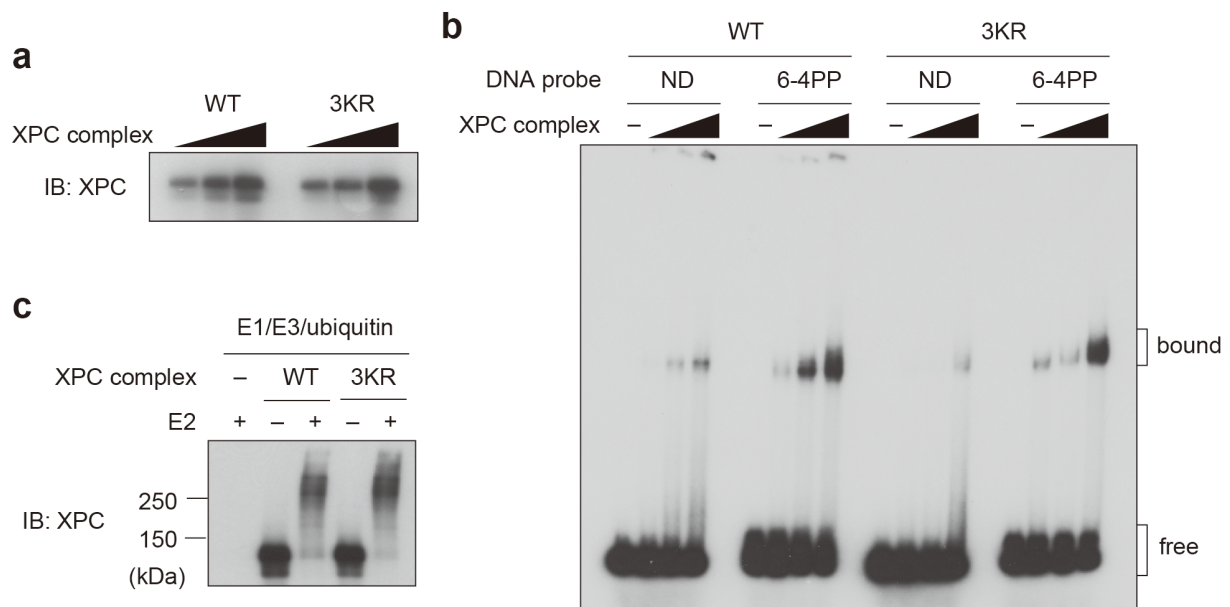
¹Biosignal Research Center, Organization of Advanced Science and Technology, Kobe University, Kobe 657-8501, Japan, ²Graduate School of Science, Kobe University, Kobe 657-8501, Japan, ³Cellular Physiology Laboratory, RIKEN, Wako 351-0198, Japan, ⁴Graduate School of Science and Technology, Kumamoto University, Kumamoto 860-8555, Japan, ⁵Graduate School of Engineering Science, Osaka University, Toyonaka 560-8531, Japan, ⁶Radioisotope Research Center, Nara Medical University, Kashihara 634-8521, Japan, ⁷Radiation Biology Center, Kyoto University, Kyoto 606-8501, Japan, ⁸Faculty of Science, Gakushuin University, Tokyo 171-8588, Japan.



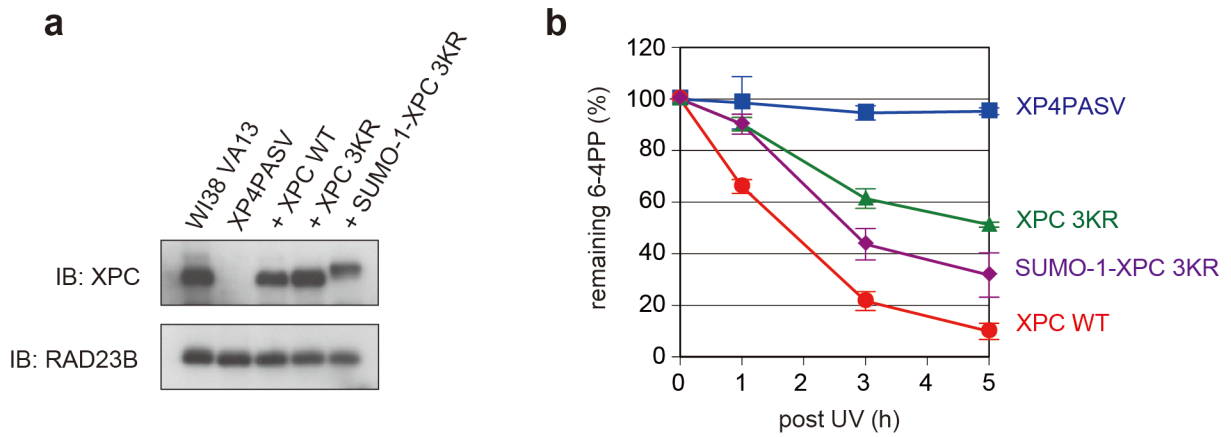
Supplementary Figure S1 | XPC SUMOylation is not stimulated by DNA and does not affect its DNA-binding activity. (a) Immunoblot analyses of XPC from *in vitro* ubiquitination and SUMOylation reactions performed in the presence or absence of undamaged or UV-damaged DNA. (b) Immunoblot analyses of XPC from *in vitro* SUMOylation reactions performed in the presence of paramagnetic beads on which DNA containing no damage (N), cyclobutane pyrimidine dimers (C), or 6-4PP (6) was immobilized. After incubation, the proteins that were bound and unbound to the DNA beads were separated and subjected to immunoblot analyses. The arrows indicate unmodified XPC.



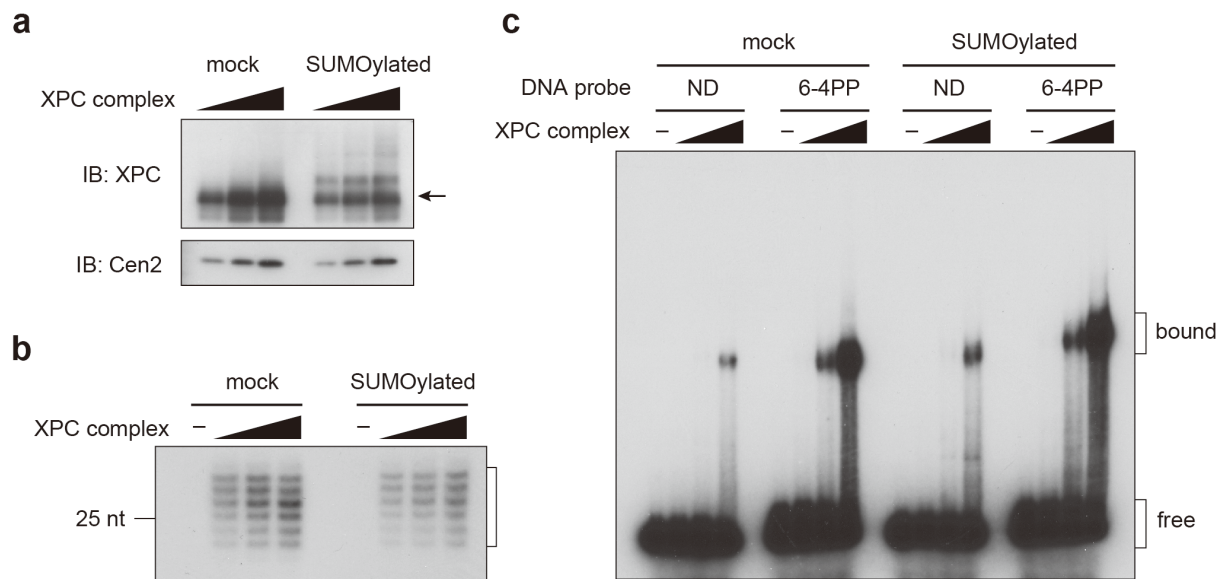
Supplementary Figure S2 | The XPC 3KR mutant is normally recruited to the sites with UV-induced photolesions but compromises recruitment of the downstream NER factors. The transformed XP4PASV cell lines stably expressing the XPC WT (a) or 3KR (b) protein were irradiated with UVC (at 100 J/m^2) through isopore membrane filters. At 30 min post irradiation, the cells were subjected to immunofluorescence staining with antibodies against the indicated NER proteins (green). The presence of UV-induced CPD was also visualized (red), and nuclei were counterstained with DAPI (blue).



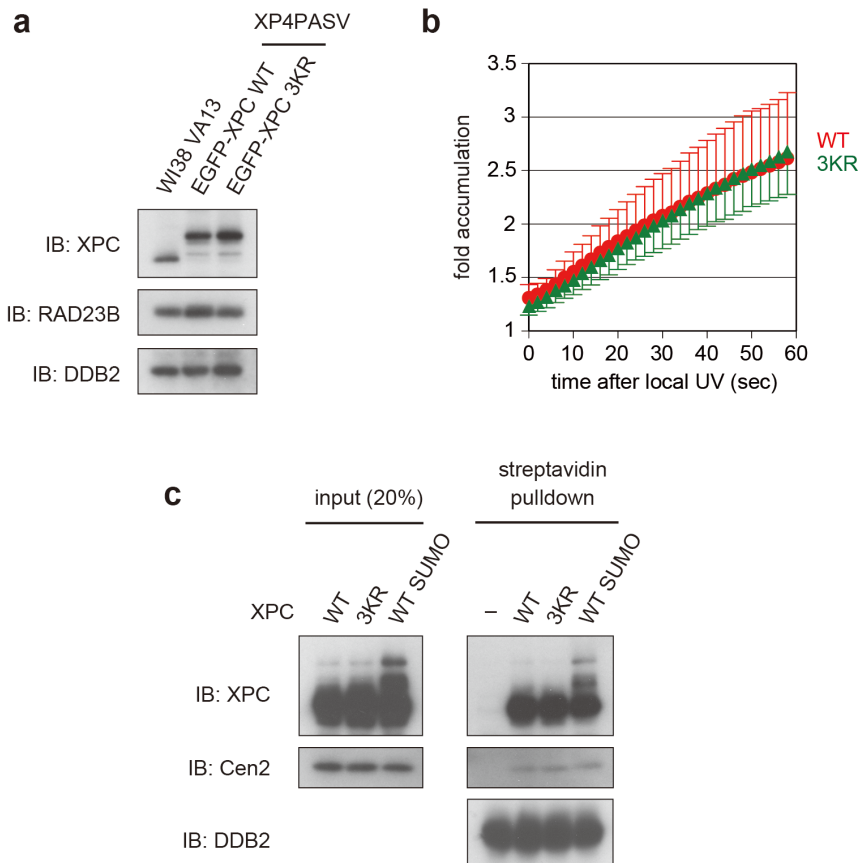
Supplementary Figure S3 |The XPC 3KR mutant protein exhibits normal damage recognition activity and CRL4^{DDB2}-mediated ubiquitination *in vitro*. (a) Immunoblot analyses of the XPC WT and 3KR protein complexes confirming that comparable amounts of proteins were included in the electrophoretic mobility shift assays. (b) Electrophoretic mobility shift assays using the XPC WT or 3KR complex and DNA probes (approximately 180 bp) containing no damage (ND) or a single 6-4PP at a specific site. The positions of the free DNA and XPC-DNA complex (bound) are indicated. (c) Immunoblot analyses of XPC WT and 3KR in cell-free ubiquitination reactions performed in the presence (+) or absence (-) of E2.



Supplementary Figure S4 | SUMO-1 fusion partially alleviates the NER defect caused by the XPC 3KR mutations. (a) Immunoblot analyses comparing the levels of the ectopically expressed XPC proteins. RAD23B was detected on the same blot as a loading control. (b) DNA repair assays of the indicated cell lines following exposure to UVC (10 J/m^2). The percentages of 6-4PP remaining in the genomic DNA were quantified and plotted as a function of time. The mean values and standard errors were calculated from two independent experiments.



Supplementary Figure S5 | The effects of SUMOylation on NER and the DNA damage recognition activity of the XPC protein *in vitro*. (a) Immunoblot analyses of the recombinant FLAG-XPC WT protein bound to RAD23B-His and centrin-2 (Cen2) and then immobilized on anti-FLAG M2 affinity beads and subjected to *in vitro* SUMOylation reactions. For the mock treatment, a similar reaction was carried out in parallel without E1. After an extensive wash, the proteins were recovered by incubating with the FLAG peptide, and varying amounts were subjected to immunoblot analyses, which confirmed that comparable amounts of proteins were included in the following assays. (b) Cell-free NER dual incision assays using increasing amounts of the protein complexes shown in (A). The excised oligonucleotides containing a 6-4PP are indicated. (c) Electrophoretic mobility shift assays using SUMOylated and unmodified XPC WT proteins and DNA probes containing no damage (ND) or a single 6-4PP at a specific site. The positions of the free DNA and XPC-DNA complex (bound) are indicated.



Supplementary Figure S6 | XPC SUMOylation does not affect its recruitment to local UV damage or physical interaction with UV-DDB. (a) Immunoblot analyses of the transformed XP4PASV cells stably expressing EGFP-tagged XPC WT or 3KR. The expression levels of RAD23B and DDB2 in the same samples were measured as controls. (b) Time-lapse images of the cells stably expressing EGFP-tagged XPC WT or 3KR were obtained after a pulse of focused deep UV light was applied to a certain area within a cell nucleus. For each image, the relative fluorescence intensity of the irradiated area relative to that of a non-irradiated area within the same nucleus was plotted as a function of time. The mean values and standard errors were calculated from analyses of 20 cells per condition. (c) The UV-DDB complex containing biotinylated DDB2 was immobilized on streptavidin-coated paramagnetic beads and used to pull down the purified XPC/RAD23B/centrin-2 heterotrimeric complex containing XPC WT, SUMOylated XPC WT, or XPC 3KR protein. The bound proteins and input samples were subjected to immunoblot analyses using the indicated antibodies.

81 89

H.sapiens ----- --MARKRA AGGEPGRREL RSQSKSAKSK ARREEEEDDA FEDEK--PPK KSLLSKVSQG KRKRGCSHPG GSADGPAKKK VAKVTVKSE- ----NLKVI KDFALSDGDD
M.musculus ----- --MAPKRT ADGR--RRK RGQKTEDNKV ARHEESVADD FEDEKQKPRR KSSFPKVSQG KRKRGCSDPG DPTNGAAKKK VAKATAKSK- ----NLKVL KEFALSDGDD
G.gallus ----- --MARKRK ASVPRAPAGK RRPGGAGVRR EREEEREEDG FVEEK--RSV KKSRAQVAHV KEDDGGAGAA KSANTPSKQK GVKRQAKENE ISPENKGTTH KEQTCDSDRK
D.ferio ----- --MA KRKDTQKSD TKPKPKQIANT KSGSKTQKAK ENGMETKKNL KNSKVASRRS RKVKDVLDEV TSKYFQDSE- ----YKTE EPEDLSDHSE
A.thaliana ----- --MKSRS ESKNCRLAQA SRVAVNKVLD KSSARGSRGK KKQDNCDS LKLVKGVNKG KQALDARLID NVLDEDCGG- ----YKTE EPEDLSDHSE
C.elegans METRRSSRL QQQTSDLNA REEPQPSEPP VKRARGAKKN EAPSTAMPMP KTTKSTKPS RKQGSPIEAE LDEMGEFLEN EEEAEIVVQK SKKNRGLKVK IDENLRISAE NGSKSSNFLE

183

H.sapiens LRDFPSDLKK AHH LKATM NEDSNEEEEE SENDWEEVEE L----SEPV ----- -LGDVRESTA FSRSLPVKP VEIEIETPEQ AKTRERSEK KLEFETYLRR AMKRFNKGVH
M.musculus FRDSPADCKK AKKHPKSKVV DQG--TDED SEDDWEEVEE L----TEPV ----- -LDMG-ENSA TSPSDMPVKA VEIEIETPQQ AKERERSEK KMEFETYLRR MMKRFNKEVQ
G.gallus APRKEIKLKR ESRPKKEMDE DNTDDDDDE SEDEWEDVEE L----QEPAL ----- -TDKLEQAV LPVAVLPSNP VEIEIETPEQ LKKRREKRK KAEFETYVRR MIKRFTEKVR
D.ferio ERMIIEDTSL SKQVKEE- ----EEDSE DEDDWEEVEE M----AGPL GPVDSSELAL E-----SKP VEIEIETPDM IRKRQKKEKR KAEFETYLRR MMNRFNKDLL
A.thaliana ----- --NVDDDEM NDSWEDCPI P----SLDS -----T VDDNVDDTR ELTIEFDVDD PDAKKQKNAY RATAEDKVRV
C.elegans ENRMEIDEKP GNLAKKSSKN GSRVKSDEK SENLVQSVPK STTNGS----- -KVAIIEDDP EIRAENGVKS SKSDEKPDFS AQNGSKLAQN APNRISRPRR SVTTAKKVSY

H.sapiens EDTHKVHLLC LLANGFYRNN ICSQPDLHAI GLSIIIPARFT RVL-PRVDVT YYLSNLVKWF IGTFTVNAEL SASEQDNLQT TLERRFAIYS ARDDEELVHI FLLILRALQL LTR-----
M.musculus ENMHKVHLLC LLASGFYRNS ICRQPDLLAI GLSIIPIRFT KVP-LQORDA YYLSNLVKWF IGTFTVNAEL SASEQDNLQT TLERRIAIYS ARDNEELVHI FLLILRALQL LTR-----
G.gallus EDTHKVHLLC LLANGFYRNN ICSQPDLHAI GLSIIPIHFT KVP-AQVDL LYLSNLVKWF VGTFTVNDL STEKGEPLQS TLERRFAIYA ARDDEELVHI FLLILRALQL LCR-----
D.ferio VDTHKVHLLC LMASGLFRNR LLCPELDLAV ALSLLPSHFT TVS-LKRINN FGLEGLKWF QATFTLNPAL PEEKEVDLRT VLEKRMGCLS ARNHEEMTLV FLLYLRLRL FCR-----
A.thaliana ELVHKVHLLC LLARGRIVDS ACNDPLIQAA LLSLLPSYLT KVSNLKVVTV KDIAPLLRVW RENFSVSCSP SSES--FRT SLAFALESRK G-TAEELAAV AVALLRALK TTR-----
C.elegans VPSDDQLLES SSSSELESSS EDEDEIRPK TGSKIACKRE KSKFISESES SSESPPDESE ASEASEDPSI GPGSEPRKRR KIQRKSTLSS GGATTKDLHW PKCSKASTAR KTGNGPSKNA

H.sapiens ----LVLSQL PIPLKSATAK GKKPSKERLT ADPGGSSETS SQVLENHTKP KTSKGTG--- --QETF AKGTCRPSAK GKRNGGRKK RSKPSSSEED EPGDKQKEA TQRPHGRER
M.musculus ----LVLSQL PIPLKSATVK GRKSKSETS EGGGSSSELS SNSPESHNP TTSRR LK--- EETL SEGRGKATAR GKRGTGTAGS RQRKRSCSE ----GEEAEQK VQGRPHARR
G.gallus ----LVLSQL PIPLKETAOK EKSTSKQSL SSTSEGEQSS GTPPKAVAKK CPCKK LK--- KSSGSEEDNE ESCKTKSAQT ERTHSKLTA NGREKQETRN VDSGLREKDV PVRPKNSRWR
D.ferio ----LVLSQL PLP LKATP KSKTTPSKSS SEKAQSEKSS P----- ELKVSFGSKR PSSATAAKS RGGKRRKKT GGGDKAEAG AQPKNSRRR
A.thaliana ----FVSTLD VAS LKADNR NESSGQNRK MHG----- -IFRTSLMV PKQQAISYSP KSKSSHVKNK SPFEKPOLGN PLGSDQVQDN
C.elegans LKVAKSSLR MAQKQPKDQP WKKNLKYDET DRKL LKRR MLEYRHVAHA YVAR LKVI TYDEAYKLE MMKAYFAGR SLLDAVLDP VEFEKSQKNV KKSEKNDKN TAGDSSSESD

H.sapiens RVASRVSYKE ESGSDEAGSG SDFELSSGEE SD-----
M.musculus RVAAKVSYKE ESESDGAGSG SDFEPSSGEG QH-----
G.gallus RVASKCYKYE ESGD-EGSV SDFEIS-GEE SD-----
D.ferio SVASKVSYK- EVGSEEEEEQ SEEEFQPSNE DD-----
A.thaliana AVNSSCEAGM LKSDGTRRK GDVEFERQIA MA-----
C.elegans EWEEMEHPQ PIIDDNI EVS IDHEGGGGDD GEEVVKDWWA IYLRQENIRK IREMMENTH VHLFCMAHL KFVVKIALDE SLVPSLMSQ LPNGYKFIG EPVVPIDIMK NLVKWFADAF

H.sapiens ----- --PSDESEPG PPKQRKAPAP QRTKAGSKSA SRTHRGSHRK DPSSLPAASS SSSSKRGKKM
M.musculus ----- --SSDEDECPG PRKQKRASAP QRTKAGSKSA SKTQRGSQCE PSSFPEASS SSGCKRGGKV
G.gallus ----- --TSDEDFETV SKKRRSSQGA QKSKVMT LK--- KSETSESLR SRNSLGEVPR PHAQRKRNI
D.ferio ----- --SEDSGAVK ICRKSKVSR RSKV KQER SEEEEEEEE EEEKEVKKK RRKKKQKQ-
A.thaliana ----- --LSATADNQQ SSQVNTTKV REITKISNSS SVSDQVSTA FGSKKVSDP-
C.elegans RPLNGVSVV SIEQDSLLEG HEARYPETRR LTALVDAKCF ETDLDRAITL FCLLRGLET TRLVVVRAI PRRWDTQQK ELQNELSKFR ELSRSRSTP GEKSLTEQEK SGEKSQPKAK

H.sapiens CSDGEKAEK- --RSTAGIDQ WLEVFCEQE- --EKWVCVDC VHGTVG--QP LTYCYKATP MTYVVGIDSD GWVRDVTQRY DPVMTVTRK CRVDAEWAE TLRPYQ----
M.musculus SSGAEEMAD- --RKPAGVDQ WLEVYCEPQ- --AKWVCVDC VHGTVG--QP VACYKYATP MTYVVGIDSD GWVRDVTQRY DPVMTATRK CRVDAEWAE TLRPYR----
G.gallus ISSDEDDGQQ MVRKVGTDQ WLEVFLERE- --DRWVCVDC VHGTVG--QP QCQFYATPK LSYLVGFOND GSVKDVQRY DPVMTMTRK KRVDPEWNEE TLQYK----
D.ferio ----- --GADE WLEVYLESS- --GRWVCVDC DQGVG--QP QLCSQATLP ITYVVLGDE GFMKDLGSRY DPTLTSRR RVVSEWNEE TMAPYK----
A.thaliana ----- --LC WLEVYCNEM MDGKWHVHDA WGNDAIEQN TEAAAAAACK VLRVYVAFAA GGAKDVTRRY CTKWHITSSK DTKTKERWLR VSLAPLHLES GATHDEDIAL
C.elegans KGQKSEKAAK KVVVEERYN WVEYWPQRE- --KRWICVDP LHKVSD--EP LSIHEHASP ISYVFAIDNK QGICEVSQRY AMDK VQDFR RRRTPNKWA WTLFLP----

655

H.sapiens ----- --SPF MDREKKEDLE FQAKHMDQPL PTAIGLYKNH PLYALKRHL LKYEAIYPETA AILGY--CRG EAVYSRDCVH TLHSRDTWLK KARVVRGVE PYKMKVGFNS
M.musculus ----- --SLL TEREKKEDQE FQAKHLDDPL PTSISTYKNH PLYALKRHL KFQAIYPETA AVLGY--CRG EAVYSRDCVH TLHSRDTWLK QARVVRGVE PYKMKVGFNS
G.gallus ----- --SPF VDRDKKEETE FQVQLDQDPL PTAIGEYKNH PLYALKRHL KYQAIYPESA AILGY--CRG EAVYSRDCVH TLHSRDTWLK QARVVRGVE PYKMKVGFNS
D.ferio ----- --SPD TERGQKEDQE MQAKLLDKPL PTSVSEYKNH PLYALKRHL LKYEALYPATA AVLGY--CRG EPVYSRDCVH TLHSRDTWLK EARTVLRGEE PYKMKVGFNS
A.thaliana RNFNGLNVPV SRASSSSSF GIRSALEDME LATRALTESL PTNQAYKSH EYIAIEKWLH KNQTLHPK-G PVLGF--CSG HPVYPRCTVQ VLTKEKRWLR DSNLQKANE PSKILKRNSK
C.elegans ----- --PFAAN SERKKWEMMQ MREDL LK LK PTWMEYKNH PLYALEKDL LKFEAIYPPPA TQKPLGQIRG HNVYPRSTVF TLQGENNWLK LARS LK LK PYKIVKARP

H.sapiens RARKARLAEQ QLEREND-- LGLFGYQTE EYQPPVAVDG KVRNFEFNV YLFLPSMMPI GCVQLNPLN HRVARKLID CVQAITGFDF HGGYSHPVTD GYIVCEEFD VLLTAWENEQ
M.musculus RARKARLSEP QLHDHND-- LGLYGHWQTE EYQPIAVDG KVRNFEFNV YLFLPSMMPV GCVQMLPLN NRVARKLID CVQAITGFDF HGGYCHPVTD GYIVCEEFR VLLAAWENEQ
G.gallus QARKARLAEQ ANRKAD-- LALFGRWQTE EYQPIAVDG KVRNFEFNV YLFLPSMPLI GCVQLRPLN NLRARKLID CAQAVTGFDF HGGYSHAVTD GYIVCEEFE VLLAAWENEQ
D.ferio RSRKARMEQ Q-KNVK-- LALFGTWQTE EYQPIAVDG KVRNFEFNV YMFKSCMLPI GCVHVHPLN HRVARKLID CALAVTGFDF HCGFAHAWD GYIVCEEHE ILKAAWENEQ
A.thaliana FKKVKDFEDG DNN LKSSC MELYGKQME PLCLPPAVNG IVPKNERGV DWVSEKCLPP GTVHLRFPRI FAVAKRFGID YAPAMVGFY RSGGATPIFE GIVCTEFDK TLEAYAEQ
C.elegans PRIPVEDRED KFLDVYQ- ----YQWTE KYRRPP LK LK YMFNEMCPL DTVLKLKSLG VQISRLKLGQ CIPAVVGFV DGGFTHPVID GAIVLEKDAI DFINAWEKLE

H.sapiens AVIERKEKEK KEKRALGNWK LLAAGLLIRE RLKRRYGPKS EAAAP----- --HTDAGG LSSDEEETS SQAEARILA ASWPQNREDE EKQK LK LK PK KTKREKAAA SHLFPFEQL-
M.musculus AIIKEKEKEK KEKRALGNWK LLVRLGLIRE RLKRYG LK LK EAAAP----- --HAAGG- LSSDEEETS SQAEARVLA ASWPQNRDP EQK--SEYTK MTRKRAAEA SHLFPFEKL-
G.gallus AEIEKEKEKEK KEKRALGNWK LLTKGLLIRE RLKQRY LK LK EPSAP----- --ETEKGG LSSDEEETS GTAGG--ME IFWPRNRQAE KQK--EKATR KSKQEKKEEA AQLFPFEKL-
D.ferio EIQQKKEQEK REKRAVNTW LLVKGLLKE RLKRRYQGG LASGTG LK LK --KEGGAEL SDEEKEGGA QSAQAPPSL ASWPQNRKE AEK--IVRR VSNRKEEKE KHLFPFEKV-
A.thaliana EKKEEEERR NEAQAASRWQLSSLLTRE RLKNRYANNS NDVEAKSLEV NSETVYKAKN LK LK LK EKQVRA KRGEKSRVRK SRNEDESHEH VFLDEEETF EETSVKTRC KCGSVSEVEQ
C.elegans SGRAEKEEQ RVEKITHENK KLIKGMLRLA YVRKQGHPA AEKPTK----- --RQIG LK LK LK EIEENEGGA GPPSSVDHIT DNTQITPMH GFLSDFINK K-----

H.sapiens -
M.musculus -
G.gallus -
D.ferio -
A.thaliana M
C.elegans -

Supplementary Figure S7 | Evolutionary conservation of the XPC SUMOylation motifs. Amino acid sequences of the XPC orthologs from indicated species are aligned by ClustalW, in which putative SUMOylation motifs predicted by the SUMOplot analyses are highlighted. The motifs with high probability (score 0.9 or higher) are shown in red, whereas those with moderate probability (higher than 0.65 and lower than 0.9) are in blue. The four SUMOylation motifs in human XPC are numbered. Accession numbers of the reference sequences used are as follows: *H. sapiens*, NP_004619.3; *M. musculus*, NP_033557.2; *G. gallus*, XP_414379.3; *D. rerio*; NP_001038675.1; *A. thaliana*, NP_197166.2; *C. elegans*, NP_500156.2.