An atypical BRCT-BRCT interaction with the XRCC1 scaffold protein compacts human DNA ligase IIIα within a flexible DNA repair complex

Michal Hammel
Ishtiaque Rashid
Aleksandr Sverzhinsky
Yasin Pourfarjam
Miaw-Sheue Tsai

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs
An atypical BRCT–BRCT interaction with the XRCC1 scaffold protein compacts human DNA Ligase IIIα within a flexible DNA repair complex

Michal Hammel1,†, Ishtiaque Rashid2,†, Aleksandr Sverzhinsky3,†, Yasin Pourfarjam4, Miaw-Sheue Tsai1, Tom Ellenberger5, John M. Pascal3, In-Kwon Kim4,*, John A. Tainer6,* and Alan E. Tomkinson2,*

1Molecular Biophysics & Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, 2Departments of Internal Medicine, Molecular Genetics & Microbiology and the University of New Mexico Comprehensive Cancer Center, University of New Mexico, Albuquerque, NM 87131, USA, 3Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, Québec, Canada, 4Department of Chemistry, University of Cincinnati, 301 Clifton Ct, Cincinnati, OH 45221, USA, 5Department of Biochemistry, Washington University, St. Louis, MO, USA and 6Departments of Cancer Biology and Molecular & Cellular Oncology, University of Texas, MD Anderson Cancer Center, Houston, TX 77030, USA

Received June 06, 2020; Revised November 21, 2020; Editorial Decision November 23, 2020; Accepted December 13, 2020

ABSTRACT

The XRCC1–DNA ligase IIIα complex (XL) is critical for DNA single-strand break repair, a key target for PARP inhibitors in cancer cells deficient in homologous recombination. Here, we combined biophysical approaches to gain insights into the shape and conformational flexibility of the XL as well as XRCC1 and DNA ligase IIIα (LigIIIα) alone. Structurally-guided mutational analyses based on the crystal structure of the human BRCT–BRCT heterodimer identified the network of salt bridges that together with the N-terminal extension of the XRCC1 C-terminal BRCT domain constitute the XL molecular interface. Coupling size exclusion chromatography with small angle X-ray scattering and multiangle light scattering (SEC-SAXS–MALS), we determined that the XL is more compact than either XRCC1 or LigIIIα, both of which form transient homodimers and are highly disordered. The reduced disorder and flexibility allowed us to build models of XL particles visualized by negative stain electron microscopy that predict close spatial organization between the LigIIIα catalytic core and both BRCT domains of XRCC1. Together our results identify an atypical BRCT–BRCT interaction as the stable nucleating core of the XL that links the flexible nick sensing and catalytic domains of LigIIIα to other protein partners of the flexible XRCC1 scaffold.

GRAPHICAL ABSTRACT

INTRODUCTION

The DNA repair gene XRCCI was first identified in a screen of human cDNAs that complemented the alklylation and X-ray sensitive phenotypes of the mutant Chinese Hamster Ovary cell line EM9 (1). Besides DNA damage sen-
sitivity, a hallmark feature of these cells is a high spontaneous level of sister chromatid exchanges, indicative of hyperrecombination (1). XRCC1 protein lacks enzymatic activity; yet, it plays a key role in multiple DNA repair pathways by acting as a scaffold that binds to both DNA single strand breaks and gaps, poly (ADP-ribose) and to numerous DNA repair enzymes (2). DNA ligase IIIα (LigIIIα) was the first XRCC1-interacting protein to be identified (3), and is critical to bridge two DNA ends to promote specific intermolecular DNA end joining (4). This was followed by DNA glycosylases (5–10), AP endonuclease I (11), and DNA polymerase β (Polβ) (12,13), implicating XRCC1 in coordinating observed product complexes (14) for the short patch subpathway of base excision repair. An interaction with automodified PARP1 serves to recruit XRCC1 and other interacting proteins that also include the end-processing enzymes, polynucleotide kinase phosphorylase (PNKP) (15), aprataxin (APTX) (16) and APLF (17) to DNA single strand breaks (18–20). While these protein-protein interactions and the phenotype of xrc1 cells show XRCC1 is a key factor in single-strand break repair (SSBR), other SSBR and base excision repair (BER) pathways involving DNA ligase I (LigI) and other DNA replication proteins are at least partially functionally redundant (21–24). This redundancy between LigI- and LigIIIα-dependent subpathways also occurs in nucleotide excision repair and in back-up pathways of DNA double-strand break repair, also known as alternative end joining (25–27).

XRCC1-deficient cells have reduced steady-state levels of both Polβ and LigIIIα, indicating that XRCC1 interactions promote the stability of these DNA repair enzymes along with their recruitment to DNA damage sites (28,29). In contrast, the steady-state levels of the other XRCC1-interacting proteins appear unaffected by XRCC1 absence. Interestingly, the binding sites for XRCC1-interacting proteins appear to be arranged linearly along the XRCC1 polypeptide with Polβ interacting with the N-terminal region and LigIIIα with the C-terminal region (2). While most binding regions appear non-overlapping, enabling XRCC1 to assemble multiprotein complexes, PNKP, MRE11, aprataxin and APLF binding sites are mutually exclusive and dependent upon XRCC1 phosphorylation, explaining the detection of distinct XRCC1 complexes containing either PNKP or apratxin (2,15,17,30–34). Interestingly, mutations in the genes encoding XRCC1, PNKP and aprataxin are the causative factor in inherited neurodegenerative diseases, indicating that XRCC1-dependent repair of single strand breaks with non-ligatable ends plays a key role in neuronal cells (35–37).

Three folded domains have been identified within XRCC1 (Figure 1A); an N-terminal domain that interacts with Polβ (38), a central BRCT domain (X1BR1) that interacts with poly(ADP-ribose)ylated (PARylated) PARP1 (18,39) and a C-terminal BRCT domain (X1BR2) that interacts with the C-terminal BRCT domain of LigIIIα (L3BR) (40–42). While inactivation of the central X1BR1 domain results in DNA damage hypersensitivity due to a defect in S phase-dependent DNA repair, disruption of the interaction between the C-terminal BRCTs of XRCC1 and LigIIIα, which results in destabilization of LigIIIα, does not, indicating that XRCC1 has nuclear DNA repair functions that are independent of LigIIIα (43–45). In contrast, LigIIIα is essential for mitochondrial DNA metabolism but XRCC1 is not present in these organelles (21,23,46,47).

For these various XRCC1 biological functions, it is important to understand the XRCC1 architectural structure and its relationship to its scaffolding functions. Since the C-terminal BRCT domains of XRCC1 and LigIIIα can form homodimers (48,49), we reasoned that these proteins may function as homodimers as well as heterodimers. Yet, although the structure of a hybrid complex between human L3BR and mouse X1BR2 has been determined (48), the structure of full-length complexes of dimeric XRCC1 and XRCC1-LigIIIα (XL) are unknown. Here, we solved the crystal structure of human L3BR complexed with human X1BR2 and furthermore define the architectures and structural flexibility of human XRCC1, LigIIIα and the XL heterodimer by combining small-angle X-ray scattering (SAXS) and electron microscopy (EM).

**MATERIALS AND METHODS**

**Protein expression and purification**

Full-length XRCC1 and LigIIIα [in pRSFduet (Novagen)] and XRCC1ΔN [residues 294–633 in pET32a (Novagen)] were expressed in Escherichia coli Rosetta cells, and purified using the same protocol as LigIIIβ described previously (50). L3BR (residues 825–922) and X1BR2 (residues 538–633) domains were expressed from pET28a with N-terminal His-tag in E. coli Rosetta cells. After Ni-NTA (Qiagen) affinity chromatography, proteins were purified by Superdex 200 (GE Healthcare) size exclusion column. For pull-down assay, L3BR, X1BR2 and X1BR1 (residues 294–417) were cloned into pGEX-6p1 (GE Healthcare). GST-tagged L3BR, X1BR2 and X1BR1 were expressed in E. coli Rosetta cells. After glutathione-sepharose chromatography (GE Healthcare), proteins were further purified by Superdex 200 column. X1BR1 was further purified by Sephadex 100 (GE Healthcare) after GST-X1BR1 was cleaved with precission protease. GFP- and MBP-tagged X1BR1 and L3BR were expressed from pET28a with N-terminal His-tag in E. coli Rosetta cells and purified using Ni-NTA affinity chromatography, followed by Superdex 200 size exclusion chromatography.

For crystallization, L3BR (residues 844–922 with C922S mutation) and X1BR2 (residues 538–633) were fused by a 22-amino acid flexible linker containing a His-tag and precission protease cleavage site, and the fusion protein was expressed in E. coli Rosetta cells. The L3BR-X1BR2 fusion protein was cleaved by purified precission protease after Ni-NTA column, and then further purified by Superdex 200 column. Selenomethionine labeled L3BR-X1BR2 fusion protein was expressed in Rosetta cells using amino acids to suppress methionine biosynthesis, as described (51), and was purified by the same protocol as the native protein. Protein crystallization conditions and structure determination methods are described in Supplementary Material.

Full-length XRCC1-LigIIIα and XRCC1 were purified from insect cells after infection with baculoviruses encoding streptactin-tagged LigIIIα and His-tagged XRCC1 (52,53). The XL complex was purified sequentially by HisTrap HP (GE Healthcare), Superdex 200 16/60 (GE Healthcare) and
Figure 1. Human BRCT–BRCT heterodimer structure stably links XRCC1 and LigIIIα via a mutationally verified interface. (A) Domain organization of human XRCC1 and LigIIIα with the C-terminal BRCT domains used for crystallization (red and blue-boxes) (B) GFP-fluorescence-based competition assay (see Supplementary Figure S3) measuring specific binding affinity of L3BR832 and L3BR844 for X1BR2. The data shown represent the mean values and standard deviations from three independent experiments. (C) Structure of human XLBR-BR complex (X1BR2, blue; L3BR, red) is superimposed on to previously reported structures of heterodimers between mouse X1BR2 and human L3BR with two different lengths of N-stretch region. The different lengths of N-stretch regions of X1BR2 and L3BR constructs are highlighted. The XLBR-BR interfaces, which are classified based on the location and main type of interaction [electrostatic (E region, top, green), hydrophobic (H region, middle, purple) and polar (P region, bottom, blue) interactions], are indicated. (D) Panels show close-up of three main binding interfaces between X1BR2 and L3BR. (E) Effects of substituting residues that are located at the BRCT–BRCT interface on X1BR2–L3BR complex formation measured by native gel analysis; upper left panel, substitution of amino acids in the E region (upper left panel) or H and P regions (lower panel), gels were run with negative control (X1BR2 or L3BR alone) and positive control (with wild-type X1BR2, WT) reactions, and then combined in a single panel to compare their effects on L3BR binding. Representative gels from two independent experiments are shown. Since the theoretical pI for L3BR is 9.22 compared with 4.90 for X1BR2, L3BR does not enter the gels, which are run at pH 7.5, unless it is complexed with X1BR2.
dsDNA cellulose (Sigma Aldrich) column chromatography. 6His-XRCC1 was purified sequentially by HisTrap HP and Superdex 200 16/60 column chromatography.

**Protein–protein interaction assays**

In a fluorescence-based competition assay, GST-L3BR (7 μM) and X1BR2-GFP (8 μM) were pre-incubated in a buffer containing 50 mM Tris–HCl pH 7.5, 150 mM NaCl, 10% glycerol and 2 mM DTT for 20 min on ice, and complexes were immobilized to GSH-coated magnetic beads (Promega). The beads were washed three times with the same buffer. The pre-bound complexes were competed off by adding MBP, MBP-L3BR<sup>852</sup> and MBP-L3BR<sup>360</sup> (0–25 μM). After the beads were pelleted using a magnet, the release of the X1BR2-GFP was monitored by the increased GFP fluorescence of the supernatant in 96-well plates using Synergy-2 microplate reader (BioTek).

To confirm the predicted key role of specific amino acids in the heterodimer interface, mutant versions of X1BR2 (residues 538–633) and L3BR (residues 825–922) were made using the Quickchange site-directed mutagenesis kit (Stratagene). Due to the high tendency of X1BR2 to aggregate in the absence of L3BR, we changed Leu626 in the C-terminus of X1BR2, which contributes to crystal packing interactions in both X1BR2 and the X<sub>LBR</sub> complex structure (data not shown) to Glu. Since this substitution did not affect L3BR binding and it also significantly reduced aggregation, it was designated as wild type (WT) in Figure 1E and used as the template for further mutational analyses. Mutant proteins were expressed and purified by the same protocol as the native protein. Purified X1BR2 (8 μM) and L3BR proteins (0–12 μM) were incubated in 20 μl binding buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 2 mM DTT and 10% glycerol) for 30 min at 22°C. Samples were separated by electrophoresis on a native 10% polyacrylamide gel as described by Kang et al. (54).

**Multi-angle light scattering coupled to size-exclusion chromatography**

SEC-MALS was carried out either in parallel with SAXS (see below) or stand-alone. For stand-alone SEC-MALS, XRCC1 and the XL complex were injected at 0.35 ml/min with an AKTAmicro (GE Healthcare) onto a Superdex 200 Increase 10/300 GL column equilibrated in 40 mM HEPES pH 7.5, 200 mM NaCl, 10% glycerol and 0.1 mM TCEP. Molecular masses were determined using the Dawn HELEOS II MALS and OptiLab T-rEX online refractive index detectors (Wyatt Technology). The detectors were calibrated with the BSA monomer and data processed with Astra Version 6.1.6.5 (Wyatt Technology). The size-exclusion column was calibrated using protein standards (Bio-Rad).

**Small angle X-ray scattering**

SAXS was performed at the SIBLYS beamline at the Advanced Light Source (55,56). For SAXS coupled with multi-angle light scattering in line with size-exclusion chromatography (SEC-SAXS-MALS) experiments, 60 μl containing either 4.3 mg/ml 6His-XRCC1 in 200 mM NaCl, 20 mM Tris–HCl pH 7.5, 0.2% glycerol (Figure 2); 5.96 mg/ml Strp-LigIIIα/6His-Flag-XRCC1 in 50 mM Tris–HCl pH 7.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 0.2 mM PMSF, 1 mM benzamidine (Figure 5). SEC-SAXS-MALS were collected at the ALS beamline 12.3.1 LBNL Berkeley, California. X-ray wavelength was set at λ = 1.2 Å and the sample-to-detector distance was 2100 mm resulting in scattering vectors, q, ranging from 0.01 to 0.4 Å<sup>−1</sup>. The scattering vector is defined as q = 4πsinθ/λ, where θ is the scattering angle. All experiments were performed at 20°C and data was processed as described (57). Briefly, a SAXS flow cell was directly coupled with an online Agilent 1260 Infinity HPLC system using a Shodex KW804 or Superdex 200 300/10 column giving similar SAXS results for Strp-LigIIIα/6His-Flag-XRCC1. The column was equilibrated with running buffer as indicated above with a flow rate of 0.5 ml/min. 55 μl of each sample was run through the SEC and three second X-ray exposures were collected continuously during a 30-min elution. The SAXS frames recorded prior to the protein elution peak were used to subtract all other frames. The subtracted frames were investigated by radius of gyration (R<sub>G</sub>) derived by the Guinier approximation I(q) = (0) exp(−q<sup>2</sup>R<sub>G</sub><sup>2</sup>/3) with the limits qR<sub>G</sub> < 1.5 (58). The elution peak was mapped by comparing the integral of ratios to the background and R<sub>G</sub> relative to the recorded frame using the program SCATTER. Non-uniform R<sub>G</sub> values across an elution peak represent a heterogeneous assembly. Final merged SAXS profiles, derived by integrating multiple frames at the peak of the elution peak, where used for further analysis including Guinier plot which determined aggregation free state. The program SCATTER was used to compute the pair distribution function (P(r)). The distance r where P(r) approach zero intensity identifies the maximal dimension of the macromolecule (D<sub>max</sub>). P(r) functions were normalized based on the molecular weight of the assemblies as determined by SCATTER using volume of correlation V<sub>c</sub> (59). Eluent was subsequently split 3 to 1 between the SAXS line and a series of UV at 280 and 260 nm, multi-angle light scattering (MALS), quasi-elastic light scattering (QELS) and refractometer detectors. MALS experiments were performed using an 18-angle DAWN HELEOS II light scattering detector connected in tandem to an Optilab re- refractive index concentration detector (Wyatt Technology). System normalization and calibration was performed with BSA monomer using a 45 μl sample at 10 mg/ml in the same SEC running buffer and a dn/dc value of 0.19. The light scattering experiments were used to perform analytical scale chromatographic separations for M<sub>n</sub> determination of the principle peaks in the SEC analysis. UV, MALS, and differential refractive index data was analyzed using Wyatt Astra 7 software to monitor the homogeneity of the sample across the elution peak complementary to the above-mentioned SEC-SAXS signal validation.

Full-length LigIIIα, XRCC1ΔN, phosphorylated XRCC1ΔN (XRCC1ΔN-p), L3BR, X1BR2 and XRCC1ΔN-p/L3BR complex samples were dialyzed with a buffer containing 25 mM Tris–HCl pH 7.5, 150 mM NaCl, 10% glycerol and 2 mM DTT. SAXS data were collected at three different concentrations on the ALS beamline 12.3.1 LBNL Berkeley, California in high-
throughput (HT) mode (60). Incident X-rays were tuned to a wavelength of \( \lambda = 1.0 \) Å at a sample-to-detector distance of 1.5 m, resulting in scattering vectors \( (q) \) ranging from 0.01 to 0.31 Å\(^{-1}\). All experiments were performed at 10°C and data was processed as described (55). Briefly, the data were acquired at short and long-time exposures (0.5 s, 5 s) then scaled and merged to obtained interference free SAXS profile. Further data processing was identical with process described for SEC-SAXS merged profiles.

Solution structure modeling
An initial atomistic model of XRCC1\(\Delta N\) was built by connecting the NMR structure of X1BR1 (PDB ID: 2DSM) and our crystal structure of X1BR2 by MODELLER (61). A full length atomistic model of XRCC1 was built by connecting NMR structure of XRCC1 N-terminal domain (PDB ID: 1XNA) (62) to the XRCC1\(\Delta N\) model. An initial homology model of full length LigIII\(\alpha\) was built based on the crystal structure of LigIII\(\beta\) (PDB ID: 3L2P) (63), the NMR structure of the N-terminal zinc finger (ZnF; PDB ID: 3OD8) (64) and our crystal structure of the C-terminal L3BR domain by connecting the regions with linkers modeled by MODELLER (61). Two initial atomistic models of dimeric XL complexes were built by adding missing linkers into two EM-derived conformers (see below). Minimal molecular dynamics (MD) simulations were performed on flexible regions in the models by the rigid body modeling strategy BILBOMD in order to optimize the conformational space of domain linkers alone (XL complex) or individual domains (XRCC1\(\Delta N\), XRCC1, LigIII\(\alpha\) and XL complex) (65). The experimental SAXS profiles were then compared to theoretical scattering curves generated from atomistic models using FOXS (66,67) followed by multistate model selection by MultiFOXS (68).

Crosslinking, negative-stain electron microscopy and image analysis
Chemical crosslinking of XRCC1 and the XL complex (dsDNA-cellulose-purified) was carried out in the dark, on ice, for 5 min with 0.05% (v/v) glutaraldehyde in 40 mM HEPES pH 7.5, 200 mM NaCl, 10% glycerol. Reactions were quenched with 1 M Tris pH 7.5 and buffer exchanged to EM buffer (40 mM HEPES pH 7.5, 200 mM NaCl, 0.1 mM TCEP). Crosslinked and non-crosslinked samples were diluted in EM buffer and deposited onto glow-discharged (Agar Scientific) carbon-coated copper grids (Electron Microscopy Sciences) for 1 min before blotting excess liquid with Whatman No. 1 filter paper. Freshly-prepared uranyl formate (1.5%; Electron Microscopy Sciences) was immediately applied for 1 min, followed by blotting excess liquid and air drying. Micrographs were collected using a FEI Tecnai T12 transmission electron microscopy fitted with a LaB6 filament and operated at 120 kV at room temperature. Low-dose mode was implemented in SerialEM (69) to acquire images at 67 000× magnification (corresponding to a pixel size of 1.64 Å) using a FEI Eagle 4k×4k CCD camera with defocus values between 0.5 and 2 μm.

Image processing was carried out with the Scipion framework (70). None of the processing steps corrected for contrast transfer function nor applied symmetry. Automatic particle-picking using a reference-free algorithm (71) resulted in 78,400 coordinates of XRCC1 (non-crosslinked) and 12,700 coordinates of the crosslinked XL complex. Particles were extracted with two-fold downsampling in 264 Å boxes. Bad picks were discarded through iterative 2D classification using CL2D from Xmipp (72). The XL dataset was pruned to remove non-XL particles, resulting in 7648 particles that were used to generate an initial model through a stochastic gradient descent algorithm (73). The initial model was low-pass filtered to 60 Å and used for 3D classification (74) into three classes using a 200 Å circular mask. Classes 1 and 3 (31% and 28%, respectively) were grouped based on their overall similarity and the resulting merged class 1 (low-pass filtered to 60 Å) was refined using the grouped particles without applying a mask (EM conformer 1). Class 2 (41% of particles) was refined separately (EM conformer 2). The resolutions of the XL complex conformers were determined using the gold-standard Fourier shell correlation at the 0.5 criterion (75). Crystal and solution structures were docked into the EM maps using UCSF Chimera (76).

RESULTS
Human XRCC1 and DNA ligase III\(\alpha\) C-terminal BRCT domains determine homodimer and heterodimer formation
XRCC1 and LigIII\(\alpha\) are elongated, multidomain proteins that form a stable complex through interactions between their C-terminal BRCT domains (48,49,63) (Figure 1A). While there is significant evidence that the XL complex is a key factor in BER and SSBR, the ability of these proteins to also form homodimers that may have specific functions (21,23,43,44,48,49), raises questions about the mechanisms that underlie formation of heterodimers and homodimers. In initial size exclusion experiments with the purified BRCT domains of human XRCC1 and LigIII\(\alpha\), L3BR\(832\) (residues 832–922) expressed either as an N-terminal MBP-fusion protein (Supplementary Figure S1A) or fused to C-terminal GFP (Supplementary Figure S1B) forms a dimer, whereas the dimerization of X1BR2 fused to C-terminal GFP appears to be transient under these conditions, consistent with the reported concentration-dependent dimerization of XRCC1 (49). Notably, deletion of 12 N-terminal residues from L3BR\(832\) results in the MBP-fusion protein behaving as a monomer (Supplementary Figure S2A). Despite efficient homodimerization, untagged L3BR\(832\) readily formed a 1:1 complex with MBP or GFP-tagged X1BR2 (Supplementary Figure S2). This finding suggests that the N-terminal linker region of L3BR (resides 832–843) is important for homodimerization but is dispensable for heterodimerization. Notably, in all size-exclusion chromatography experiments, the formation of heterodimers of X1BR2 and L3BR was considerably more favorable than homodimerization (Supplementary Figure S2).

To measure the apparent dimerization affinity of BRCT–BRCT interactions, a GFP-fluorescence-based, competition binding assay was developed (see Materials and Methods and Supplementary Figure S3). For this assay, GST-tagged L3BR\(832\) was immobilized on glutathione-coated
magnetic beads in complex with a GFP-labeled X1BR2 domain. The release of the X1BR2–GFP protein into the supernatant was monitored by the increase in GFP fluorescence as MBP-tagged L3BR832 protein was added to the binding reaction for competition. In this competition binding assay (Figure 1B), L3BR844 shows a slightly higher X1BR2-binding affinity than L3BR832 (EC50 values of 0.67 ± 0.08 and 1.24 ± 0.06 μM for L3BR844 and L3BR832, respectively), whereas MBP alone showed no interaction (Supplementary Figure S3). Together, we conclude that LigIIIα and XRCC1 specifically heterodimerize with subto low micromolar affinity through their C-terminal BRCT domains, and the N-terminal linker of L3BR contributes to the determination of homodimer and heterodimer formation, prompting us to further investigate the interface of human XL and to identify key determinants of the specific XL complex formation.

**BRCT heterodimer atomic resolution structure**

We solved and refined the human X1BR2–L3BR844 complex (XLBR–BR) structure at 2.4 Å resolution (Supplementary Table S1, Figure 1C and Supplementary Figure S4). Overall, the structure of this complex resembles the previously reported hybrid complex between human L3BR and mouse X1BR2 (48) (Figure 1C). In the human X1BR–BR complex, there is a V-shaped ‘head-to-head’ arrangement of interacting L3BR–X1BR2 domains that buries 1164 Å2 of solvent accessible surface. Complementary surfaces of the BRCT domains are mostly contributed by helix α1 and the proximal ‘α1–β2 loop’, together with a short N-terminal extension (N-stretch) from each BRCT domain (Figure 1C). In comparison with the hybrid XLBR–BR structure, there is an increase in the tilt angle between L3BR and X1BR2 domains in the human structure. Notably, the structure of the hybrid X1BR–BR heterodimer with a longer X1BR2 N-stretch (Figure 1C, cyan, PDB ID: 3QVG) is more compact than the human XLBR–BR complex with a shorter N-stretch region (Figure 1C, red/blue), suggesting that an additional electrostatic interaction between the X1BR2 N-stretch residue Asp531 and L3BR residues Arg914 and Lys915 contributes to compaction of the complex (Figure 1C and Supplementary Figure S5). However, there are also differences in the main interaction interface between the similar-length hybrid XLBR–BR (PDB ID: 3PC8) (48) and the human complex, presumably reflecting differences in the amino acids of mouse and human X1BR that contribute to the interface.

In the human complex, three protruding side chains of X1BR (Leu539 from the N-stretch and Arg560 and Arg564 from helix α1) contact the hydrophobic C-terminal half of α1 in L3BR (Figure 1D). Together they make a network of electrostatic interactions with nearby residues Asp849, Asp876 and Asp878 of L3BR. Similar, but not strictly equivalent, reciprocal interactions are made by analogous residues of L3BR. Leu847 within the N-stretch of L3BR and Arg870 from α1 contact the hydrophobic C-terminal half of α1 and its proximal surfaces of X1BR2 (Figure 1D). Overall, the two BRCT domains are interlinked by protruding Leu and Arg residues from each subunit that stabilize the human XLBR–BR heterodimer by a combination of electrostatic (E region), hydrophobic (H region) and polar interactions (P region) (see Figure 1C and D). In the E region, Arg560 and Arg564 from α1 of X1BR2 form a network of salt bridges with Asp849, Asp876 and Asp878 residues from L3BR (Figure 1D). In addition, Arg564 from α1 of X1BR2 makes a hydrogen bond with the main chain carbonyl oxygen of Ala874 of L3BR. The center of the BRCT–BRCT interface (H region) is contributed by hydrophobic side chains from each subunit that form an interdomain hydrophobic core (Figure 1D). X1BR2 Leu539 appears critical for this hydrophobic core by virtue of extensive hydrophobic interactions with Leu847, Tyr871, Ala874 and Phe875 of L3BR. In addition, Leu539 main-chain amide hydrogen bonds with the Tyr871 side-chain of L3BR. Similarly, L3BR Leu847 makes hydrophobic interactions with XRCC1 residues Leu539, Tyr565 and Phe569. Finally, in the P region, Arg870 of L3BR makes hydrogen bonds with the side-chain amide of Asn570 and the main-chain carbonyl oxygens of Leu539 and Ala568 in X1BR2 (Figure 1D, P region).

**X1BR2 N-stretch and α1 are critical for XL heterodimerization**

To test the structurally-implicated critical roles of the E, H and P regions in XL heterodimerization, a series of amino acid substitutions were introduced into all three regions of the XLBR–BR interface (Figure 1E). Substitutions of Leu539 from the N-stretch of X1BR2 with Ala or Glu completely disrupted complex formation with L3BR as detected by native gel electrophoresis (Figure 1E). This reveals the functional importance of this region extending from the BRCT core and also supports and extends the previous finding that the N-terminal segment of X1BR2 significantly expands the buried surface area upon binding to L3BR (48). In contrast, L3BR Leu847, a structural equivalent of Leu539 in the N-stretch of XRCC1, appears to be dispensable for XL complex formation (Figure 1E), despite the similar appearance of both Leu residues and their local interactions with either side of the dimer interface. Although contradictory to a previous report (48), this is consistent with the dispensable role of the N-terminal linker of L3BR for XRCC1 binding and is indicative of functional asymmetry in this heterotypic BRCT–BRCT interaction. Given the dispensable role of L3BR Leu847, the severe defect in the X1BR2 Y565A/F569A double mutant (Figure 1E) is likely to be explained by some other interaction, perhaps a change in packing between XRCC1 helices α1 and α3 due to the replacement of bulky aromatic side chains. While X1BR2 Y565A, F569A and N570A mutants exhibited little or no effect on complex formation with L3BR (Figure 1E), the R560A, R564A and R560A/R564A mutants either failed to or had greatly diminished ability to form complexes with L3BR. Taken together, these mutational analyses defined specific hydrophobic interactions involving the N-stretch (Leu539) of XRCC1 and corresponding surfaces of LigIIIα, along with electrostatic interactions between α1s from XRCC1 (Arg560 and Arg564) and LigIIIα, as key determinants of XRCC1–LigIIIα heterodimer formation.
Atypical asymmetric BRCT–BRCT interaction promotes XL heterodimerization

The α helices of X1BR2 and L3BR domains share a remarkable sequence similarity (50% identity in α1 and surrounding residues, versus 29% overall sequence identity for the whole BRCT domains; Supplementary Figure S6), including similar patterns of charged and hydrophobic patches of surface residues around the dimer interface. Given the pseudo two-fold symmetry of interacting subunits in the XLBR–BR heterodimer (Supplementary Figure S6), this raises the question of why the XL heterodimer is preferred over the L3BR homodimer (Supplementary Figure S6), including similar patterns of charged and hydrophobic residues, versus 29% overall sequence identity.

Given the pseudo two-fold symmetry of interacting subunits in the XLBR–BR heterodimer (Supplementary Figure S6), including similar patterns of charged and hydrophobic residues, versus 29% overall sequence identity. The XLBR–BR complex deviates significantly from the XRCC1 heterodimer (PDB ID: 1CDZ) (48) [2.4 Å resolution, r.m.s.d. of 0.8 Å for Cα atoms]. The XLBR–BR complex deviates significantly from the XRCC1 heterodimer (PDB ID: 1CDZ) (48) [2.4 Å resolution, r.m.s.d. of 0.8 Å for Cα atoms]. The XLBR–BR complex deviates significantly from the XRCC1 heterodimer (PDB ID: 1CDZ) (48) [2.4 Å resolution, r.m.s.d. of 0.8 Å for Cα atoms]. The XLBR–BR complex deviates significantly from the XRCC1 heterodimer (PDB ID: 1CDZ) (48) [2.4 Å resolution, r.m.s.d. of 0.8 Å for Cα atoms].

Collectively, these experimental data are consistent with a transient dimerization of XRCC1 via the X1BR2 domains.

To explore the conformational space that can be occupied by XRCC1, we performed rigid body modeling that was validated by SAXS fitting. Initially we determined the atomistic model of XRCC1 by connecting the NMR structure of X1BR1 (PDB ID: 2D8M) and the crystal structure of X1BR2 (see Materials and Methods). A selected minimal ensemble model (Figure 2D), that includes 40% extended monomer, 35% extended dimer and 25% compact dimer, was the best fit for the experimental SAXS profile and P(r) function (Figure 2C and Supplementary Figure S7A). Next, we modeled full length XRCC1 by connecting the NMR structure of XRCC1 N-terminal domain (PDB ID: 1XNA) (62) to the XRCC1N model. An ensemble model (Figure 2E) that includes 61% extended monomer, 14% extended dimer and 25% compact dimer closely matched the experimental SAXS profile and P(r) function (Figure 2C and Supplementary Figure S7A).

To gain insights into the structure of full-length human LigIIIα, we initially determined a solution state of the L3BR (83). Based on the estimated molecular mass and SAXS fitting of rigid body modeling, the solution structure model for L3BR closely matches the two-fold symmetric homodimer of L3BR (PDB ID: 3PC7) (Figure 3A and Supplementary Figure S7C, Table S2). The results of the Kratky plot (Figure 3B) and the P(r) function (Figure 3A) analyses of full-length LigIIIα indicate that, like LigIIIβ (63), LigIIIα has a disordered multidomain architecture. Unlike the L3BR domain that is exclusively dimeric in solution (Supplementary Figure S7C, Table S2), full-length LigIIIα is a mixture of monomer and dimers in solution, with max-
Figure 2. Human XRCC1 is an elongated, disordered protein that transiently forms homodimers. (A) SEC-SAXS-MALS chromatographs for XRCC1. Solid lines represent the UV 280 nm (light magenta) or SAXS signal (magenta) in arbitrary units, while symbols represent molecular mass (light magenta) and Rg values for each collected SAXS frame (magenta) versus elution time. The SEC-SAXS-MALS results, which show full-length XRCC1 is a mixture of monomer/dimer, are representative of at least two independent preparations of XRCC1. (B) Normalized Kratky plot of XRCC1 (magenta) in comparison with XRCC1ΔN-p (dark green) and XRCC1ΔN (light green). (C) Normalized P(r) functions of XRCC1 (magenta) in comparison with XRCC1ΔN-p (dark green) matching theoretical P(r) functions of atomistic models (black) shown in panels D and E. Weighted ensemble atomistic model shown in molecular surface representation were used to fit experimental SAXS curves for XRCC1ΔN-p (D) and XRCC1 (E) shown in Supplementary Figure S7A and represented as theoretical P(r) functions in panel C.

imal elongation of ~200 Å (Figure 3C and Supplementary Figure S7B, Table S2). A comparison of normalized Kratky plots of LigIIIβ with LigIIIα indicates a similar level of protein disorder (Figure 3B) despite Lig IIIβ behaving as a monomer due to the absence of the L3BR domain (63). As expected (63), a comparison of LigIIIβ (LigIIIβ 1–755) missing the disordered C-terminal 107 residues and LigIIIβ missing the N-terminal ZnF domain (LigIIIβ 1–755 ΔZnF) (Figure 3B) confirmed the significant contribution of the ZnF region to the disordered character of LigIIIβ (63). Together, our rigid body modeling analysis indicates that the N-terminal ZnF extension and the C-terminal BRCT domain of LigIIIα, contribute significantly to the elongated shape of the DNA ligase III polypeptides.

Visualization of XL complex by negative stain electron microscopy reveals a loosely packed domain arrangement

To gain insights into the possible domain organization of the XL complex, we co-purified the full-length proteins and confirmed that they form a ~180 kDa heterodimer by SEC-MALS despite eluting earlier than expected from the gel filtration column (Figure 4A and B). XRCC1 and the XL complex were each stabilized with chemical crosslinking (Figure 4C), buffer exchanged and visualized using negative-stain EM. Micrograph quality was greatly improved by the crosslinking and subsequent buffer exchange steps (Supplementary Figure S9). Yet, 2D classification of the single particles from the XRCC1 sample revealed a conformationally heterogenous population of XRCC1 (Figure 4D) that is, based on our SAXS data (Figure 2A), likely to reflect a mixture of monomers and dimers. Crosslinking did not reduce this heterogeneity (data not shown). Efforts to generate an initial 3D model of XRCC1 failed, probably due to its intrinsic structural and oligomeric heterogeneity. Despite efficient crosslinking, classes of particles representing XRCC1-only particles were detected in the population of particles from the XL sample (Figure 4E). After removal of the XRCC1-only particles, the uniformly-sized XL particles (Figure 4F) were subjected to 3D classification to identify conformationally different forms of the XL complex (Supplementary Figure S10). Particles belonging to two visually similar classes were pooled and used to refine EM conformer 1. Particles from a third class were used to create EM conformer 2 (both conformers shown in Figure 4G).

Some structural heterogeneity for the XL complex was expected given the flexible and dynamic nature of both proteins. The conformers presented here do not repre-
Figure 3. Human LigIIIα is an elongated, disordered protein that transiently forms homodimers. (A) Normalized $P(r)$ functions for experimental SAXS curves for full length LigIIIα (cyan) and L3BR (green) are fitted to the theoretical $P(r)$ functions (black) of atomistic models of full length LigIIIα (panel C) and the two-fold symmetric crystal structure of the L3BR homodimer (PDBID: 3PC7, black dots) and L3BR monomer (red). (B) Comparison of normalized Kratky plots of LigIIIα (cyan), L3BR dimer (green) with the SAXS curves of LigIIIβ (dark violet), LigIIIβ 1–755 (violet), and LigIIIβ 1–755 ΔZnF (dark green) (taken from (63)). Kratky plots show persistent disorder of full length LigIIIα and LigIIIβ but significant less disorder in the LigIIIβ construct with truncated ZnF domain. (C) Weighted ensemble atomistic models of LigIIIα monomer (left) and homodimer (right) are shown in surface representation. Corresponding fits of the SAXS curves for the atomistic models are shown in the Supplementary Figure S7D and as a theoretical $P(r)$ function in panel A.

sent definitive models of domain organization, but rather present a framework for modelling potential domain arrangements. Although other conceivable domain configurations might also be consistent with the EM results, our domain docking EM map interpretation was guided by known crystal structures of LigIIIα domains and our insights into the conformation of XRCC1 alone from 2D classes (Figure 4D). The XRCC1 monomer often adopts a linear arrangement that can be observed in both EM conformers (Figure 4G). The published atomic coordinates of individual XRCC1 domains were docked within this linear protrusion so as to have the N-terminal domain (XRCC-N) distal to the rest of the complex and the L3BR–X1BR2 heterodimer at the interface with the remaining volume. At the other side of the complex, we interpreted the isolated spherical density as the LigIIIα ZnF. Linking the two sides of the complex would be the LigIIIα catalytic fragment composed of the DNA binding, adenylation and OB-fold domains (63). These conserved domains are known to adopt an open conformation, as seen for the Saccharolobus solfataricus DNA Ligase (81) as well as a closed conformation with nicked DNA (63). The EM map for conformer 1 accommodates the catalytic domains in an approximately open conformation whereas a more compact LigIIIα is a better fit for the wider volume seen in conformer 2. Overall, our interpretation of two EM maps suggest a dynamic LigIIIα catalytic fragment that in turn orients XRCC1 at different positions within the complex. The full dynamics of XRCC1 are probably underestimated due to limitations in visualizing unstructured linkers by EM. However, the observed conformers allowed us to build a framework of domains upon which we added quantitative assessment of flexibility by SAXS as described below.

The XL complex is more structured and compact than XRCC1 and LigIIIα

Prior to investigating the solution structure of full-length XL, we determined that the interaction of L3BR532 with XRCC1ΔN-p did not change the extended shape of XRCC1ΔN as judged by the Kratky plot, $P(r)$ function and SAXS parameters of the complex ($D_{\text{max}}$ of $\sim$200 Å and $R_g$ of 54 Å (Figure 5A and B, Supplementary Table S2). We then examined the conformational flexibility of the full-length XL complex by performing SEC-SAXS-MALS experiments (Supplementary Figure S7D). Consistent with Figure 4A, the major peak eluting from the gel filtration column corresponded to the XL heterodimer, as judged by molecular estimates from MALS and SAXS (Supplementary Figure S7D, Table S2). The decrease of $R_g$ val-
Figure 4. Visualization of XL particles by electron microscopy reveals a dynamic domain arrangement. (A) SEC-MALS analysis shows that XRCC1 either alone or in complex with full-length LigIIIα elutes earlier than globular protein standards despite molar masses of ~77 and ~180 kDa, respectively. (B) The XL peak seen in (A) is composed of a heterodimer. (C) The XL dimer was chemically crosslinked prior to negative-stain EM imaging. (D) 2D classification reveals oligomeric and conformational heterogeneity in non-crosslinked XRCC1. (E) Particles belonging to non-XL 2D classes were removed from the cross-linked XL dataset. (F) The final dataset contained uniformly-sized 2D classes of particles. 3D classification and subsequent refinement resulted in two conformers of the XL complex. (G) The 3D EM maps are rotated to each display a linear protrusion from a larger area of density. The linear extension potentially represents XRCC1 domains X1BR2 (PDB ID: 3PC8 chain A), X1BR1 (PDB ID: 2D8M) and the N-terminus (PDB ID: 3K77). The LigIIIα ZnF (PDB ID: 1UW0), three domain catalytic fragment (DBD, NTase and OBD, PDB ID: 3L2P) and L3BR (PDB ID: 3PC8 chain C) domains were docked into the larger area of density of each conformer.
ues determined for individual SAXS frames across the elution peak (Figure 5C) suggests XL conformational heterogeneity through transient interactions between multiple XL domains. However, the Kratky plots revealed significant conformational exchange of the XL heterodimer relative to the XRCC1 dimer (Figure 5A), consistent with the reduced heterogeneity the XL particles compared with XRCC1 particles seen by EM. To determine whether the compaction extends beyond the interacting C-terminal BRCT domains in the XL complex, we examined the protease-sensitivity of labeled, adenylated LigIII. Notably, the labeled catalytic fragment of LigIII was more resistant to proteolysis when complexed with XRCC1 (Supplementary Figure S11), suggesting that the BRCT–BRCT heterodimer interacts with and influences the conformation of the LigIII catalytic region.

To test how the EM conformers 1 and 2 (Figure 4G) match the solution state, we compared models based on the available atomic structures that do not include the linker regions between domains to the SAXS curve. Without linker regions that are missing from the atomic resolution structures, fits of the EM models were poor ($\chi^2_{\text{conformer1}} = 11.6$ and $\chi^2_{\text{conformer2}} = 15.1$, see Supplementary Figure S7E). To determine whether this discrepancy was due to the flexibility of the missing regions, these sections were built using MODELLELLER (61) and conformational sampling was applied using BILBOMD (65) to mimic the flexibility of the added regions (Supplementary Figure S12). While the SAXS fit was improved by including linker flexibility ($\chi^2_{\text{conformer1}} = 2.6$ and $\chi^2_{\text{conformer2}} = 3.0$), the remaining discrepancy between theoretical and experimental SAXS curves in the low resolution range of SAXS ($q \approx 0.05–0.15 \text{ Å}^{-1}$, see Figure 5C) suggested larger conformational rearrangements of the XL domains in solution. Closer inspection of the EM-derived models showed that the N-terminal domain of XRCC1 and the LigIII N-terminal ZnF domain are spatially separated from the XL core (X1BR1, X1BR2, L3BR and LigIII catalytic domains) (Figure 4G). Thus, we performed rigid body modeling that included flexibility of the N-terminal domains of XRCC1 and LigIII using EM conformers 1 and 2, respectively, as initial models. From approximately 10,000 conformers derived by conformational sampling (Supplementary Figure S7F), a multistate model containing two conformers was selected by MultiFOXS (68) (Figure 5D and Supplementary Figure S12). Displacement of the N-terminal domain of XRCC1 and the LigIII N-terminal ZnF domain from their initial locations (Figure 5D and Supplementary Figure S12) significantly improved the SAXS fit ($\chi^2_{\text{conformer1}} = 1.6$ and $\chi^2_{\text{conformer2}} = 1.7$, Figure 5C) in the low-resolution range, in accord with the flexibility of these regions inferred by our EM studies and shown by others (49,50) (Figure 5D and Supplementary Figure S12). Notably, the most dominant conformers in both conformational samplings (Figure 5D) adopted a similar arrangement to the EM-derived conformers. Taken together, the SAXS-derived model of full-length XL supports the domain arrangement predicted by EM in which the BRCT–BRCT heterodimer is the stable core of the XL complex that increases the compaction of the LigIII catalytic region and links the flexible N-terminal regions of XRCC1 that are accessible for additional protein-protein interactions with the flexible N-terminal nick sensing region of LigIIIα.

**DISCUSSION**

Based on previous studies, both XRCC1 and LigIIIα are elongated proteins composed of some well-folded domains connected by flexible, unstructured regions (2,49,63). While there is atomic resolution structural information for the well-folded domains (39,48,62,63,82,83), less is known about the functional architecture of the full length proteins, particularly the XL complex that plays a key role in coordinating base excision and single strand break repair (84,85). Accurate structural analyses of large and flexible complexes can require combined approaches to assess both flexibility (by SAXS), shape (by SAXS and EM) and detail (by X-ray crystallography and NMR) (86,87). Here we integrated results of crystallographic, EM and SAXS experiments for LigIIIα and XRCC1 with existing structural information (39,48,62,63,82,83) to build a complete data-based model of the highly flexible XL heterodimer that is anchored by a unique BRCT–BRCT interaction (Figures 4 and 5).

XRCC1 and LigIIIα BRCT domains are critical for the stability and function of nuclear LigIIIs (29,42,44). Yet, XRCC1 has LigIIIα-independent functions in nuclear DNA repair (42,44,45). Furthermore, the essential function of LigIIIα in mitochondrial DNA metabolism occurs in the absence of XRCC1 and is not dependent upon its C-terminal BRCT domain (21,23,47). In accord with published studies concluding that the C-terminal BRCT domains also homodimerize (48,49), we observed formation of XRCC1 and LigIIIα homodimers albeit to a lesser extent than XL heterodimers. Overall, our atomic resolution structure of the human BRCT–BRCT heterodimer closely resembles that formed by mouse X1BR2 with human L3BR and supports and extends the notion that interactions with the region N-terminal to X1BR2 contribute to the preferred formation of hetero- versus homodimers (48). However, differences between the human and mouse-human hybrid structures including an increase in the tilt angle between the X1BR2 with L3BR domains in the human structure evidently reflect differences in the amino acids contributing to the interaction interface. Residue substitutions confirmed the key role of specific amino acids in the structurally-defined interface. They furthermore highlighted the structural and functional asymmetry, in particular rotation of L3BR α1 helix in the XL heterodimer relative to the corresponding α1 helix of X1BR2 in the homodimer (48), that seems to favor heterodimer formation.

While the XL complex is flexible, it is significantly more compact than either XRCC1 or LigIIIα (Figure 5A). The linker region between the X1BR1 and X1BR2 domains is extensively phosphorylated by CK2, enhancing interactions with PNKP and APTX via their FHA domains (30,33,80,88). However, CK2 phosphorylation had no detectable effect on XRCC1 conformation. Since the catalytic region of LigIII undergoes a large conformational change when it engages nicked DNA (50), we examined the effect of complex formation with XRCC1 on the LigIII catalytic region. Notably, it was much more resistant to proteolysis.
Figure 5. Increased XL compaction from comparison of EM and SAXS results. (A) Normalized Kratky plot of XRCC1 (magenta) in comparison with XRCC1ΔN-p/L3BR (blue) and XL complex (red) revealed compaction of larger XL complexes. The SAXS profiles are representative of two independent experiments. (B) Normalized $P(r)$ functions obtained for experimental SAXS shown in panel A fitted to the theoretical $P(r)$ functions (black) of atomistic models of XRCC1 (Figure 2E) and the multistate model of XL complex derived from the EM conformer 1 shown in panel D. $P(r)$ function of XRCC1 ΔN-p/L3BR complex revealed that the XRCC1ΔN has an elongated, flexible conformation, even when bound to L3BR domain. (C) Experimental (black) and theoretical (colored as indicated) SAXS profiles for the two XL EM conformers with optimized conformations of the linker regions (red and magenta) and optimized locations of the N-terminal domain of XRCC1 and the LigIII N-terminal ZnF domain (XRCC-N, cyan and ZnF, green). SAXS fits are shown together with the fit residuals in the lower panel and $\chi^2$ values indicating goodness of fit. (D) The top weighted model from the multistate SAXS-model of XL complex was modelled using EM conformer 1 (top) and conformer 2 (bottom) as initial model. Both conformers of multistate models are shown in Supplementary Figure S12. Models are superimposed on to the 3D EM maps. Corresponding SAXS fits for the atomistic models are shown in panel C and further shown as a theoretical $P(r)$ functions in panel B. XRCC1 is constitutively phosphorylated by CK2 (30), and the phosphorylation site is indicated by P within a green circle. Interacting regions of XRCC1 with partner proteins are indicated.
within the XL heterodimer. Thus, the BRCT-BRCT heterodimer not only anchors these proteins but also presumably initiates secondary interactions that reduce the flexibility of the LigIII catalytic region, an effect that may contribute to the ability of XRCC1 to stimulate ligation by and maintain the steady state levels of LigIIIα (29,42).

During SSBR, the stable heterodimeric core of XL defined here effectively tethers the flexible XRCC1 scaffold that mediates recruitment of LigIIIα, as well as other XRCC1-interacting proteins, including Polβ, PNPK and APTX (12,13,15,16) to sites of DNA damage via an interaction between X1BR1 and PARylated PARP1 (18,20,39) (Figure 3D). The N-terminal ZnF of LigIIIα is part of a flexible SSB sensing module that may facilitate location of SSBs in the context of local high concentrations of poly(ADP-ribose) (50,89,90). If the SSB is non-ligatable, we envision that the flexibility of XRCC1 and LigIIIα will enable dynamic sensing and processing of the SSB termini by the LigIIIα ZnF and XRCC1-associated proteins within the XL complex, culminating in the circularization of the LigIIIα catalytic fragment around a ligatable nick to complete repair (63). Future structural and biochemical studies are needed to address how the binding of other XRCC1 partner proteins impact the flexibility of the XRCC1 scaffold and the repair of base damage as well as both single- and double-strand breaks. For example, in BRCA2-defective cells, there are increased interactions between XRCC1 and both MRE11 and DNA polymerase theta that may contribute to the observed genome rearrangements in BRCA-deficient cancers (91), so these structural results may inform ongoing investigations of existing and novel chemotherapeutic strategies.

Here, we have used crystal structures combined with negative stain EM and SAXS to obtain complementary insights into the structure and conformational flexibility of an intact highly dynamic protein complex that could not be obtained by either X-ray crystallography or NMR alone. The combined results show that accurate determination of XL structures must include its dynamic and functional flexibility as was recently found for the DNA PK complex (92). Our integrative modeling from EM and SAXS plus atomic resolution domain structures revealed the functional compaction of the XL complex that extended beyond the interacting C-terminal BRCT-BRCT domains to the catalytic region of LigIIIα. The consensus models of the XL complex suggest that the interconversions of the flexible N-terminal extensions of LigIIα and XRCC1 facilitate functional interactions with diverse protein partners during BER and strand break repair.

DATA AVAILABILITY

Coordinates and structure factors have been deposited in the PDB with accession codes 6WH1 for the XLBR-BR complex and 6WH2 for X1BR2. The EM maps have been deposited in the EMDB with accession codes 22130 and 22307 for EM conformer 1 and EM conformer 2, respectively. SAXS data and atomistic models of XRCC1ΔN-p, XRCC1, LigIIIα and XL complex have been deposited in the SASBDB with accession codes SASDJ52, SASDJ62, SASDJ72 and SASDJ82.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Mark Glover kindly provided us with pGEX-CK2α plasmid. Drs Leslie Hicks and Sophie Alvarez provided expert mass spectrometry analysis of proteins at the Proteomics & Mass Spectrometry Facility of the Donald Danforth Plant Science Center (Saint Louis, MO).

Author contributions: M.H., T.E., J.M.P., I-K.K., J.A.T. and A.E.T. designed the overall research. M-ST, I-K.K., A.S. and I.R. were responsible for protein expression and purification; I-K.K., I.R. and A.S. performed the biochemistry and mutagenesis studies, I-K.K., Y.P. and T.E., the X-ray crystallography and analysis, M.H. the small angle X-ray scattering and analysis and A.S. and J.M.P. the electron microscopy and analysis. The manuscript was written by I-K.K., M.H., T.E., J.M.P., J.A.T. and A.E.T. with contributions from A.S. and I.R. All authors have given approval to the final version of the manuscript.

FUNDING

National Institutes of Health (NIH) [R01 ES012512 to A.E.T., R35 CA220430 to J.A.T., in part]; The Structural Cell Biology of DNA Repair Program [P01 CA092584 to A.E.T., T.E., J.A.T.]; V Foundation V Scholar Grant [V2018-25 to I.K.K.]; J.A.T. acknowledges the Cancer Prevention Research Institute of Texas (CPRIT) [RP180813]; Robert A. Welch Chemistry Chair; National Science and Engineering Research Council of Canada [RGPIN-2015-05776 to J.M.P.]; SIBYLS beamline at the Advanced Light Source was provided in part by the Offices of Science and Biological and Environmental Research, U.S. Department of Energy through DOE BER Integrated Diffraction Analysis Technologies (IDAT) program; NIGMS [P30 GM124169-01 ALS-ENABLE]. Funding for open access charge: NIH. Conflict of interest statement. None declared.

REFERENCES


