Supplementary Information

Tsodikov et al. "Structural Basis for the Recruitment of ERCC1-XPF to Nucleotide Excision Repair Complexes by XPA"

Materials and Methods

Protein expression and purification. All proteins were expressed in BL21(DE3) *E. coli* (Stratagene). The cells were grown to OD_{600} = 0.5 at 37 ^{o}C, then cooled down to 22 ^{o}C and induced with 0.5 mM of IPTG at 22 ^{o}C for 15 hours. XPA proteins were purified by Ni^{2+} chromatography using a HiTrap Ni^{2+}-chelating column (Amersham-Pharmacia) following manufacturer's instructions. These proteins was dialysed overnight and the His-tag was concomitantly cleaved in a buffer containing 30mM Tris pH 8.0, 400 mM NaCl, 2 mM beta-mercaptoethanol and Prescission protease in approximately a 1:100 molar ratio to XPA. The XPA and ERCC1 proteins were further purified individually on S100 HiPrep (Amersham-Pharmacia) column. In order to obtain homogeneous XPA-ERCC1 complexes, purified XPA and ERCC1 were combined in excess of ERCC1 (for XPA_{1-273}, XPA_{59-273}, XPA_{59-219}) or of XPA (for XPA_{59-93} and XPA_{67-80}) and the same gel filtration purification step was repeated. The peak corresponding to XPA-ERCC1 was well separated from the excess ERCC1 or XPA for all XPA constructs. This separation or the shapes of the peaks was not affected in the salt concentration range of 50-400 mM NaCl. For the XPA_{59-219} fragment containing the Zn^{2+}-binding domain of XPA (Fig. 1AB in the Supplementary Information), an elemental analysis performed on the corresponding XPA-ERCC1 complex indicated that 98 % of XPA_{59-219} contains a Zn^{2+}

atom and we conclude that the structured part of XPA remained properly folded in association with ERCC1.

Labeled proteins for NMR studies were produced in M9 minimal media containing ^{15}N-labeled NH_4Cl and ^{13}C-labeled glucose as the sole sources of nitrogen and carbon, respectively. A perdeuterated, ^{15}N-labeled ERCC1 sample was prepared in media containing 100% D_2O with 100% deuterated glucose and ^{15}N-labeled NH_4Cl.

Analysis of equilibrium sedimentation experiments. Absorbance measured in the experiment,

$$A_{280}(r) = A_{280}(a)\exp\left(\frac{\omega^2 M(1-\rho v)(r^2-a^2)}{2RT}\right) \quad (1)$$

was analyzed using equation (1) in which A_{280} is the absorbance at 280 nm, r and a are an arbitrary and a reference radial distances, $\omega = 2\pi f$ (where f = 40,000 min^{-1}) is the angular velocity of the rotor, ρ = 1 g/mL is the density of water, M is the molecular weight of the sedimented species, R is the Boltzmann constant and T = 277 K is the absolute temperature.

The nonlinear regression fitting of the data to Eq. 1 to determine M was performed using SigmaPlot 9.0 (SSP). Data were fit (solid and dashed lines) using a single-component, non-interacting model and assuming partial specific volumes v = 0.75 mL/g for both samples. The fit (shown Fig. 1C in Supplementary Information) yields

molecular weights of (15.0 ± 1.0) kDa and (19.4 ± 1.2) kDa for free $ERCC1_{92-214}$ and $ERCC1_{92-214}$ -XPA_{59-93} complex respectively.

Protein crystallization and data collection. The complex of $ERCC1_{96-214}$-XPA_{67-80} was concentrated to 9 mg/ml using Amicon (Millipore) concentrator with a 5 kDa molecular weight cut-off in 30 mM Tris pH 8.0, 200 mM NaCl, 2 mM beta-mercaptoethanol and 0.1 mM EDTA. Crystals were grown by vapor diffusion in hanging drops containing 1 μL of the protein solution and 1 μL of the reservoir solution (100 mM Tris pH 8.5, 2 M ammonium dihydrogen phosphate, 10 % glycerol) at 21.5 ^{o}C. Single cubic crystals of XPA-ERCC1 complex grew in 1-2 weeks, reaching a size of 0.15-0.20 mm in each dimension. The crystals diffracted to 4.0 Å resolution using a rotating anode X-ray source at Harvard-Armenise X-ray facility. I/σ(I) decreases sharply (R_{merge} increases sharply) with increasing resolution at 4 Å. As a result, higher resolution shells contain no data useful for structure refinement.

Nuclease assay. A substrate consisting of a 12 base pair stem with a 22 nucleotide loop (5'-GCCAGCGCTCGGT_{22}CCGAGCGCTGGC) was 5' end labeled using a T4 polynucleotide kinase and [γ-^{32}P]-ATP. The DNA substrate (100 fmol; at a final concentration of 6.7 nM) was suspended in nuclease buffer (25 mM TrisHCl pH 8.0, 40 mM NaCl, 10% glycerol, 0.5 mM β-mercaptoethanol, 0.1 mg/ml BSA) containing 0.4 mM $MnCl_2$ then incubated with ERCC1-XPF (100 or 400 fmol, corresponding to 6.7 nM or 26.8 nM) in the presence of 92 μM XPA_{67-80} or XPA_{67-80}-F75A. The final reaction volume was 15 μl. These DNA cleavage reactions were incubated at 30 °C for 15 min

then stopped by adding 10 μl of loading dye (90% formamide/10 mM EDTA) and heating at 95 °C for 5 min. Samples were analyzed by 15 % denaturing PAGE (0.5x TBE) and the reaction products were visualized using a PhosphorImager (Typhoon 9400; Amersham Biosciences).

Competitive binding equilibrium titrations. The equilibrium titrations were performed in a binding buffer containing 20 mM Tris, pH 8.0, 20 mM NaCl, 2 mM beta-mercaptoethanol. The concentrations of single-stranded 5' 6-FAM-labeled DNA (5'-CCG GTG GCC AGC GCT CGG CG$(T)_{20}$) and $ERCC1_{96\text{-}214}$ were 50 nM and 2.33 μM, respectively, the concentration of the XPA peptide was varied from 0 to 30 μM, as shown in Fig. 3 in the Supplementary Information. Increasing XPA peptide concentration above 30 μM caused an increase in fluorescence anisotropy and quenching of fluorescence due to nonspecific interactions with the peptide, observed in the presence or absence of ERCC1 (data not shown).

The simplest binding model used in the data analysis consisted of two competitive equilibria, binding of XPA peptide and of the single-stranded 6FAM-labeled 40-mer DNA to the central domain of ERCC1. This model is consistent with the fact that at sufficiently high XPA concentration, the fluorescence anisotropy signal approaches that of unbound DNA. Therefore the two binding equilibria could be written as:

$$E + D \xleftrightarrow{K_{DNA}} ED \tag{2a}$$

$$E + X \xleftrightarrow{K_{XPA}} EX \tag{2b}$$

where E, D, X represent the ERCC1, DNA and XPA species, respectively. ED and EX represent ERCC1-DNA and ERCC1-XPA complexes, and K_{DNA} and K_{XPA} are the observed equilibrium association constants, defined in terms of the equilibrium concentrations of the species from Eqs. (2a) and (2b) as:

$$K_{DNA} = \frac{[ED]}{[E][D]} \quad \text{and} \tag{3a}$$

$$K_{XPA} = \frac{[EX]}{[E][X]}. \tag{3b}$$

The total concentrations of reaction species according to conservation of material are represented as:

$$[X]_{tot} = [X] + [EX] = [X] + K_{XPA}[E][X], \tag{4a}$$

$$[E]_{tot} = [E] + [EX] + [ED] \approx [E] + K_{XPA}[E][X]. \tag{4b}$$

$$[D]_{tot} = [D] + K_{DNA}[E][D] \tag{4c}$$

The approximation made in Eq. (4b) is due to the large excess of ERCC1 over DNA at the experimental conditions. It should be noted that in our analysis we assume that each single-stranded 40-mer DNA oligonucleotide contains only one binding site for ERCC1. Because the central domain of ERCC1 occludes 10-15 nucleotides upon binding DNA (Tsodikov et al, 2005), with a reasonable approximation, a maximum of 2-3 ERCC1

molecules could simultaneously bind to one DNA molecule without significant cooperativity. In this analysis, we use the value of $(K_{DNA})^{-1}$ = 1.5 μM, obtained from the direct titration of ERCC1 at constant concentration (50 nM) of the same DNA oligomer, using the same single site approximation (Tsodikov et al, 2005). Therefore the value of the equilibrium binding constant for XPA-ERCC1 complex formation, K_{XPA}, should be unaffected by the stoichiometry of ERCC1-DNA binding.

The system of Eqs. 4a, 4b and 4c yields expressions for [E], [X] and [D], which are then used to determined the fraction of bound DNA species, *f*, as follows :

$$[X] = \frac{K_{XPA}([X]_{tot} - [E]_{tot}) - 1 + \sqrt{(K_{XPA}([X]_{tot} - [E]_{tot}) - 1)^2 + 4K_{XPA}[X]_{tot}}}{2K_{XPA}}, \quad (5)$$

$$[E] = \frac{[E]_{tot}}{(1 + K_{XPA}[X])}, \quad (6)$$

$$[D] = \frac{[D]_{tot}}{(1 + K_{DNA}[E])}, \quad (7)$$

And, finally

$$f = \frac{K_{DNA}[E][D]}{[D]_{tot}} \quad (8)$$

The observed fluorescence anisotropy, *r*, is then given by

$$r = r_0 + (r_{\max} - r_0) f \,, \tag{9}$$

where r_0 *and* $r_{\max}$ are fluorescent anisotropy values of unbound and fully bound DNA, respectively. We used SigmaPlot 9.0 (company) to perform non-linear regression analysis of our data using Eq 9 in order to obtain the best-fit value for K_{XPA}.

NER assay. HeLa cell extracts and plasmid containing 1,3-intrastrand cisplatin adduct were prepared as described previously (Hoy et al, 1985). Hela cell extract (2μl) or XP-A (XP2OS) cell extract (3μl), 2 μl of 5x repair buffer (200 mM Hepes-KOH, 25 mM $MgCl_2$, 110 mM phosphocreatine (di-Tris salt, Sigma), 10 mM ATP, 2.5 mM DTT and 1.8 mg/ml BSA, adjusted to pH 7.8), 0.2 μl 2.5 mg/ml creatine phosphokinase (rabbit muscle CPK, Sigma) and either purified XPA peptide, XPA protein (WT, F75A, G73Δ or G73Δ/G74Δ) or NaCl (final NaCl concentration was 70 mM) in a total volume of 10 μl were pre-incubated at 30°C for 10 min. One μL of a covalently-closed circular DNA plasmid (50 ng) containing the 1,3-intrastrand cisplatin crosslink was added before incubating the mixture at 30 °C for 45 min. After placing the samples on ice, 0.5 μl of 1 μM 35-mer oligonucleotide (5'-GGGGGAAGAGTGCACAGAAGAAGACCTGGTCGACCp-3') was added and the mixtures heated at 95 °C for 5 min. The samples were allowed to cool down at room temperature for 15 min to allow the DNA to anneal. One μL of a Sequenase/[α-^{32}P]-dCTP mix (0.5 units of Sequenase and 2.5 μCi of [α-^{32}P]-dCTP per reaction) was added before incubating at 37 °C for 3 min, 1.2 μl of dNTP mix (100 μM of each dATP, dTTP,

dGTP; 50 μM dCTP) was added and the mixture incubated for another 12 min. The reactions were stopped by adding 8 μl of loading dye (90% formamide/10 mM EDTA) and heating at 95 °C for 5 min. The samples were run on a 20% sequencing gel (0.5x TBE) at 45 W for 2.5 hrs. A low molecular weight DNA marker (New England Biolabs) was used as a reference after end-labeling the DNA with [α-^{32}P]-dCTP and Klenow fragment polymerase. The reactions products were visualized using a PhosphorImager (Typhoon 9400, Amersham Biosciences). Slight variation in band intensities in band intensities (e.g. higher intensities in Fig. 4A, lane 3 in the main text) is due to experimental variability in the amount of material loaded in different lanes of the gel.

References

Hoy CA, Thompson LH, Salazar EP, Stewart SA (1985) Different genetic alterations underlie dual hypersensitivity of CHO mutant UV-1 to DNA methylating and cross-linking agents. *Somat Cell Mol Genet* **11**(6)**:** 523-532

Tsodikov OV, Enzlin JH, Scharer OD, Ellenberger T (2005) Crystal structure and DNA binding functions of ERCC1, a subunit of the DNA structure-specific endonuclease XPF-ERCC1. *Proc Natl Acad Sci U S A* **102**(32)**:** 11236-11241

Figure Legends

Supplementary Figure 1. XPA_{59-213} forms a stable 1:1 complex with the central domain of ERCC1 . (**A**) The complex of XPA_{59-213} with $ERCC1_{92-214}$ was separated by gel filtration chromatography (S-100 column; Pharmacia) from an excess of $ERCC1_{92-214}$. The XPA-ERCC1 complex (peak 1) elutes before the unbound ERCC1 (peak2). (**B**) SDS-PAGE analysis of fractions corresponding to peaks 1 and 2 from the gel filtration experiment demonstrates the

presence in peak 1 of XPA and ERCC1 at an approximately equimolar ratio of the two proteins. (**C**) Sedimentation equilibrium analysis of $ERCC1_{92-214}$ (the central domain of ERCC1) alone and in complex with XPA_{59-93}. The natural logarithm of the absorbance at 280 nm is plotted *versus* the square of the relative radial position. The data for unbound $ERCC1_{92-214}$ (open circles) and its complex with XPA_{59-93} (open triangles) are plotted in a similar matter. The curves represent the best fit of these data to Eq. 1 (Materials and Methods), yielding molecular weights of 15 kDa for unbound $ERCC1_{92-214}$ and 19.4 kDa for $ERCC1_{92-214}$- XPA_{59-93}, in good agreement with the prediction for a 1:1 binding stoichiometry.

Supplementary Figure 2. The XPA_{67-80} peptide binds specifically to $ERCC1_{92-214}$. The XPA_{67-80} peptide labeled with 6FAM at the N-terminus (10 nM peptide) was mixed with $ERCC1_{92-214}$ at the concentrations shown. Peptide binding was recorded as an increase in the anisotropic fluorescence emission (calculated by the methods described in the Supplementary Methods). The binding isotherm of ERCC1 to the XPA_{67-80} peptide was well fit by a simple hyperbolic 1:1 isotherm, yielding a K_D of 0.78 ± 0.06 μM.

Supplementary Figure 3. The XPA_{67-80} peptide is a competitive inhibitor of single-stranded DNA binding to $ERCC1_{92-214}$. Binding of a 6FAM-labeled single-stranded DNA 40mer to the central domain of ERCC1 was monitored by fluorescence polarization of the labeled DNA. DNA binding activity is plotted as a function of added XPA_{67-80} inhibitor. The theoretical curve given by Eq. 9 (solid line) corresponds to the best fit of

the data to the dissociation equilibrium constant of 300 nM for the ERCC1-XPA complex.

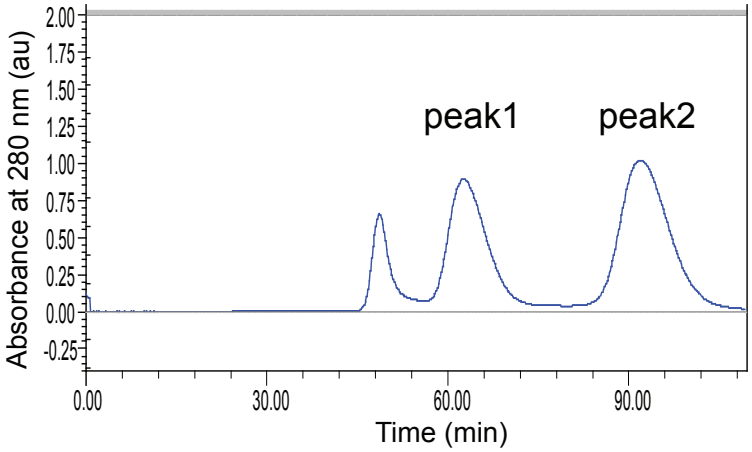
Table 1. Data collection and NMR and X-ray structure determination statistics.

X-ray diffraction and refinement without using NMR data	
Space group	$I4_132$
Unit cell parameters	a = b = c = 128.6 Å
Number of XPA-ERCC1 complexes per a.u.	1
Resolution	40 – 4.1 Å (4.3-4.1 Å)[a]
I/σ	16.8 (6.0)
Redundancy	10.5 (10.5)
R_{merge}	0.15 (0.41)
Number of unique reflections	1538
R/R_{free} without XPA, prior to refinement using NMR data	0.33/0.38
NMR and refinement using X-ray data	
Total NOE Distance Restraints	109
Intermolecular (ERCC1—XPA)	18
Intramolecular (XPA)	91
Intraresidue	61
Interresidue	30
Hydrogen Bond Restraints	1
Dihedral Angle Restraints	0
<RMSD> from mean structure (XPA 70-77) backbone/heavy atom (Å)	0.19/0.44
Ramachandran Plot (% residues)	
Most Favorable Region	77.6
Additionally Allowed Region	18.7
Generously Allowed Region	3.7
Disallowed Region	0

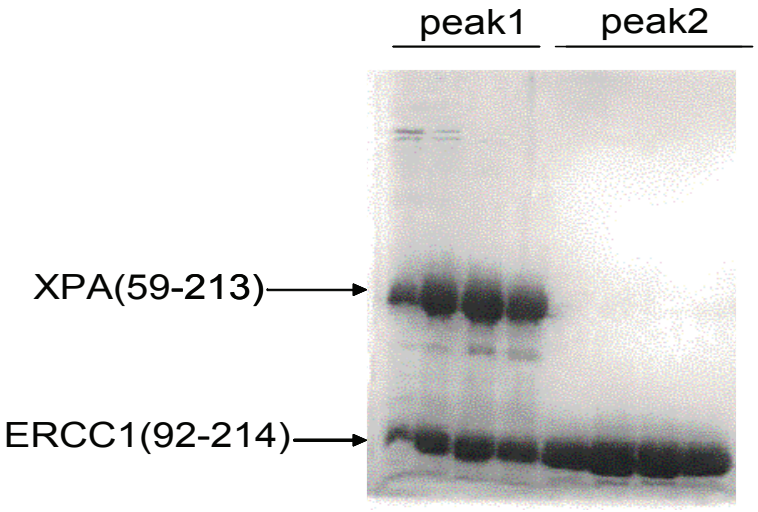
[a] Data in parentheses indicate the highest resolution shell

Tsodikov et al., 2007 Supporting Information, Figure 1

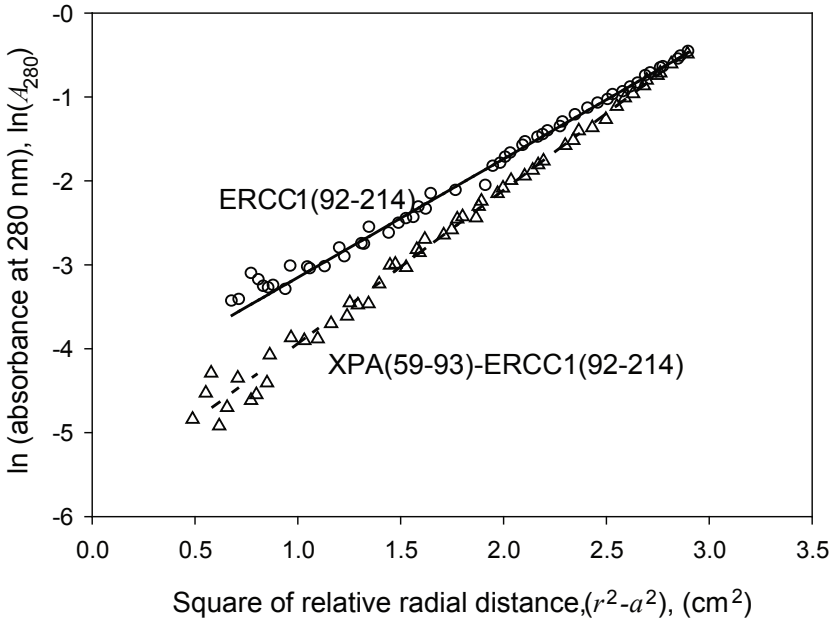
A



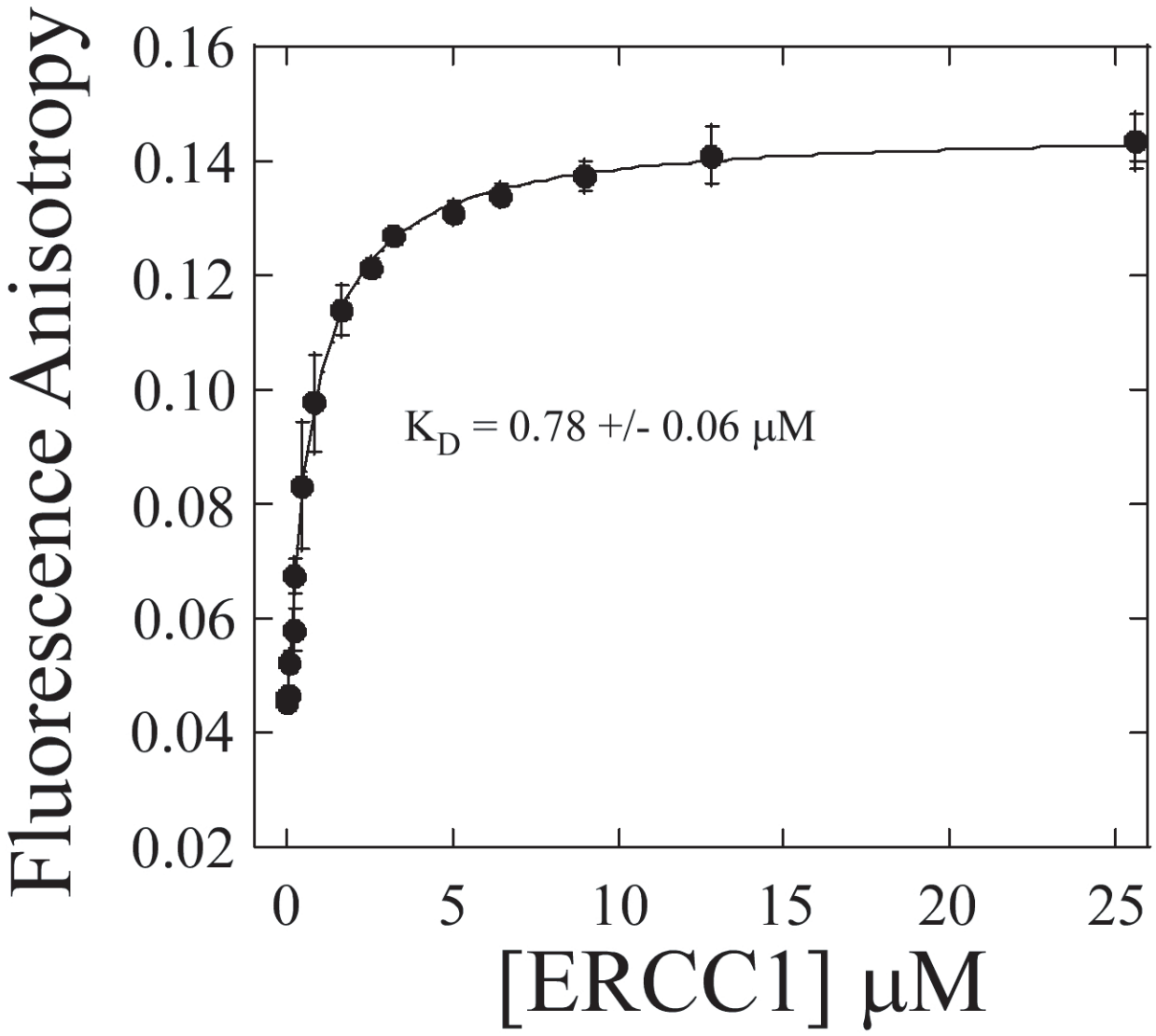
B



C



Tsodikov et al., 2007 Supporting Information, Figure 2



Tsodikov et al., 2007 Supporting Information, Figure 3