Universiteit
Leiden
The Netherlands

Structure-based insights into the repair of UV-damaged DNA

Meulenbroek, E.M.

Citation
Meulenbroek, E. M. (2012, October 9). *Structure-based insights into the repair of UV-damaged DNA*. Retrieved from https://hdl.handle.net/1887/19938

Version:	Corrected Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/19938

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/19938 holds various files of this Leiden University dissertation.

Author: Meulenbroek, Elisabeth Maria
Title: Structure-based insights into the repair of UV-damaged DNA
Issue Date: 2012-10-09

2

Purification and crystallization of Cockayne Syndrome protein A

Cockayne Syndrome protein A is one of the main components in mammalian Transcription-Coupled Repair. Here the overproduction, purification, and crystallization of human Cockayne Syndrome protein A in complex with its interacting partner DNA Damage Binding protein 1 are reported. The complex was coproduced in insect cells, copurified and crystallized using sitting drops with PEG3350 and sodium citrate as crystallizing agents. The crystals have unit cell parameters a=b=142.03 Å, c=250.19 Å and diffract to 2.9 Å on beamline ID14-1 at the European Synchrotron Radiation Facility.

published in: E.M. Meulenbroek, N.S. Pannu, *Overproduction, purification, crystallization and preliminary X-ray diffraction analysis of Cockayne syndrome protein A in complex with DNA damage-binding protein 1.*, Acta Cryst. **F68**, 45-48 (2012)

2.1 Introduction

The human genome is constantly subject to damaging agents such as ultra-violet irradiation. One of the major pathways to remove DNA lesions is Nucleotide Excision Repair (NER). It is responsible for removing a wide variety of chemically and structurally diverse lesions. A subpathway of NER is Transcription-Coupled NER (TC-NER), which removes lesions blocking transcription. Most proteins are shared between Global Genome NER (GG-NER) and TC-NER, but two proteins are unique to TC-NER: Cockayne Syndrome Protein A and B (CSA and CSB). As can be read in Chapter 1, these proteins were proposed to have the following functions (Fousteri *et al*., 2006): CSB attracts NER factors and chromatin remodelers to the site of the lesion and it uses its DNA-dependent ATPase activity to remodel the RNA polymerase - DNA interface to prevent RNA polymerase hiding the damaged site from the NER factors. CSA is found in a CSA-DDB1 E3-ubiquitin-ligase/CSN complex, which is inactive for E3-ubiquitin-ligase activity upon recruitment, and might prevent degradation of the stalled RNA polymerase and/ or other TCR factors in early stages of the repair or it might be involved in degradation of CSB at a later stage of the repair process. The degradation of CSB could be important for recovery of transcription after TCR is finished (Groisman *et al*., 2006).

The biological importance of CSA and CSB can be seen since mutations in either protein can cause the recessive human disorder Cockayne Syndrome. This disease, named after the London physician Edward Alfred Cockayne (1880-1956), is characterized by neurologic abnormality, growth retardation, abnormal sensitivity to sunlight and premature aging (Nance & Berry, 1992). Eighteen different mutations in CSA that can cause Cockayne Syndrome (Laugel *et al*., 2010) have been reported.

Cockayne Syndrome protein A is found in the cell in complex with DDB1, which is a multifunctional protein that links several different substrate adaptors like CSA to the Cul4A-Roc1 complex hence constituting the E3-ubiquitin ligase complex (recently reviewed in Iovine *et al*., 2011). Several crystal structures of DDB1 have been reported, the first of which was reported in Li *et al*., 2006. However, to date no structural information is available on CSA.

Determination of the structure of Cockayne Syndrome protein A will provide insight into the important DNA repair mechanism TC-NER and into the molecular basis by which the disease-causing mutations in CSA cause Cockayne Syndrome. To this end, we report the overproduction, purification, and crystallization of CSA in complex with DDB1.

2.2 Materials and Methods

2.2.1 Cloning and overproduction

The open reading frame (ORF; amino acids 1-396) for human CSA (ERCC8; OMIM 609412) was amplified by PCR from a clone in a pFastbac vector kindly made available by Wim Vermeulen (Erasmus Medical Center, The Netherlands) and cloned into the pETM-series of vectors (Dümmler *et al*., 2005) and into pET52b (Invitro-

gen). Overproduction in *E.coli* strains BL21 Rosetta, RIL, RP, pLysS and pLysS star was attempted at 277, 293 and 310 K with 0.1 – 1 mM IPTG.

Overproduction in *E.coli* was concluded to be unsuccessful, after which overproduction of CSA together with its interaction partner DDB1 was attempted in Sf9 insect cells. For this, a vector with the ORF (amino acids 1-1140) for human DDB1 (OMIM 600045) was kindly made available by Andrea Scrima and Nicolas Thomä (Friedrich Miescher Institute, Switzerland). The sequence of a N-terminal 6x-His-tag and a thrombin cleavage site were inserted in front of the ORF and together it was placed behind the p10 promoter in a pFastbac Dual vector (Invitrogen) using *Xma*I and *Acc*65I. The resulting DDB1 molecule had the following residues extra at its N-terminus: MHHHHHHRRLVPRGSGGR.

CSA was amplified from the pET52b construct with a C-terminal 10x His-tag and thrombin cleavage site using the following primers: 5'-TTT CAC GGT CCG GGG ATG CTG GGG TTT TTG TCC GCA CG-3' and 5'-AGT AGT CGA CGT TAA TTA GTG GTG GTG ATG GTG ATG ATG GTG-3' and cloned into the same pFastbac Dual vector as DDB1 behind the PolH promoter using the restriction enzymes *Rsr*II and *Sal*I. The resulting protein misses its C-terminal glycine and has the following residues extra at its C-terminus: LALVPRGSSAHHHHHHHHHH. The construct was verified by sequencing (Baseclear, Netherlands).

The pFastbac Dual with CSA and DDB1 was transformed into the strain DH10EMBacY (Berger *et al*., 2004), kindly provided by Imre Berger (EMBL, France). Recombinant bacmid was isolated and transfected to *Sf9* insect cells using Fugene HD (Roche) following the manufacturer's instructions. Recombinant baculovirus was produced and insect cells were infected at a density of 1.5-2 x 10^6 cells/ml using 2 ml 4^{th} generation virus per 100 ml cell culture in suspension. The cells were harvested 66 hours post-infection.

2.2.2 Purification

All purification steps were executed at 277 K. Harvested cells were resuspended in lysis buffer containing 50 mM Tris pH 8, 20 mM imidazole, 200 mM NaCl, 0,1 % Triton-X-100, 5 mM β-mercaptoethanol and Complete mini EDTA-free protease inhibitor cocktail (Roche). Cells were lysed by sonification and the lysate was centrifuged in an ultracentrifuge at 30.000 rpm (61.700 g) for 30 min at 277 K in a Beckman Coulter 70.1 Ti preparative rotor. The supernatant was loaded on a His-Trap HP column (GE healthcare) using an Äkta Express (GE healthcare) and washed with 20 column volumes of Ni buffer A containing 20 mM Tris pH 8, 30 mM imidazole, 200 mM NaCl and 5 mM β-mercaptoethanol. The CSA-DBB1 complex was eluted with a gradient of 24 column volumes to 30 % Ni buffer B and then 28 column volumes to 100 % Ni buffer B containing 20 mM Tris pH 8, 330 mM imidazole, 200 mM NaCl and 5 mM β-mercaptoethanol.

The Ni column fractions containing the CSA-DBB1 complex were diluted three times with 20 mM HEPES pH 7.2 and loaded on a Hitrap Q HP column (GE healthcare). The column was washed with Q buffer A containing 20 mM HEPES pH 7.2, 100 mM NaCl and 5 mM β-mercaptoethanol. The protein complex was eluted by

a gradient of 25 column volumes to 100 % Q column B containing 20 mM HEPES pH 7.2, 1 M NaCl and 5 mM β-mercaptoethanol. Next, the CSA-DDB1 containing fractions were loaded on a Superdex 200 gel filtration column (GE healthcare) equilibrated with 20 mM HEPES pH 7.2, 200 mM NaCl and 5 mM DTT and the column was run with a flow rate of 0.3 ml/min. Protein purity was assessed with 10 % SDS PAGE stained with Coomassie Blue (Simply Blue Safe stain, Invitrogen) and by silver staining (Silver Stain Plus, Biorad).

2.2.3 Crystallization

CSA-DDB1 was concentrated with a 10 kDa MWCO centrifugal filter unit (Millipore) to a concentration of 5-8 mg/ml. Crystallization conditions were screened using sitting-drop vapour-diffusion using the commercial screens JCSG+ and PACT (Qiagen) at 293 K with a drop size of 0.8 μl. A Genesis RS200 (Tecan) was used for pipetting the reservoir solution (75 μl) and an Oryx6 (Douglas Instruments) was used for pipetting the drops. After two days, small needle-like crystals appeared in conditions 71, 83, and 95 from PACT, which contain 0.2 M sodium citrate, 20% (w/v) PEG 3350 and 0.1 M Bis Tris propane 6.5, 7.5 or 8.5. The crystals were verified to be protein using a fluorescence microscope with filter U-MWU2 (Olympus).

These crystals were optimized at 293 K and the best condition was found to be 0.2 M sodium citrate, 24% PEG 3350, 0.1 M Bis-Tris propane pH 8.0. Based on this condition, an additive screen (Hampton Research) was performed following the manufacturer's instructions. 3% glycerol was found to cause a significant improvement in size and morphology of the crystals. After additional optimization, addition of 5-7% glycerol turned out to be optimal and drops of size between 2 and 4 μl using a protein solution to crystallant solution volume ratio of 2:1 or 3:1.

2.2.4 X-ray diffraction analysis

Crystals were caught in cryoloops and soaked in a solution containing mother liquor and 10–12 % glycerol before flash-cooling them in liquid nitrogen. X-ray diffraction experiments were done on ID14-1 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. 180 images were collected with an oscillation angle of 1.0° and the exposure time was 13 s per frame on ID14-1 at 0.9334 Å at 100 K. Images were processed with *iMosflm* (Leslie, 1999). Scaling and merging was done with *SCALA* from the *CCP4* suite (Winn *et al*., 2011).

2.3 Results and discussion

Overproduction of CSA in *E.coli* was attempted with several tags, but only an MBP-fusion protein was soluble. Purification on amylose resin (New England Biolabs) gave protein of good purity. Cleavage of the MBP-tag with TEV protease, however, led to precipitation of CSA. Gel filtration on the MBP-fusion protein showed that the fusion protein was present as a high-molecular weight entity, presumably a soluble aggregate.

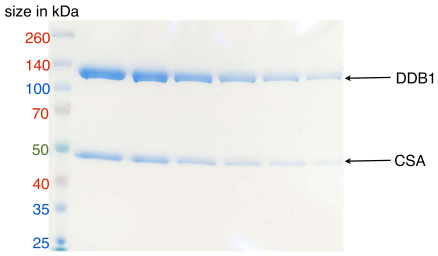

Figure 2.1: Coomassie-stained SDS PAGE gel showing CSA and DDB1 at the end of the purification procedure.

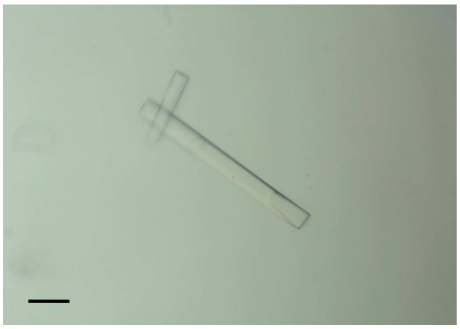

Figure 2.2: Crystals of CSA in complex with DDB1. The scale bar is 100 μm in length.

Coproduction of CSA and its interacting partner DDB1 was then attempted in insect cells, because the structurally and functionally related complex DDB1-DDB2 has been coproduced and purified from insect cells (Scrima *et al.*, 2008). Coproduction of CSA and DDB1 in *Sf9* cells yielded a soluble protein complex and a large excess of DDB1 (more than three times excess). The large excess of DDB1 could be removed by using a customized gradient on the Ni column, since DDB1 alone eluted in the first part of the gradient (0-30 % B), while the CSA-DDB1 complex eluted in the second part of the gradient (30-100 % B). CSA copurified with DDB1 as a 1:1 complex as judged from SDS PAGE (Figure 2.1) throughout the entire purification and no free CSA was obtained. The complex eluted as a single peak from the gel filtration around 170 kDa. This indeed approximates the size of a 1:1 complex of CSA with DDB1, whose theoretical sizes are 46 and 129 kDa in our constructs. Western blotting with antibodies against DDB1 and CSA and mass-spectrometry on proteolytic fragments confirmed the identity of both proteins.

Crystallization conditions were found in a pH anion cation screen (PACT) and optimized by a grid screen around the condition varying PEG concentration and pH. The initial crystals were around 0.09 by 0.01 by 0.01 mm and diffracted to around 8 Å at the ESRF. An additive screen was performed and 3 % glycerol was found to improve crystal size and quality significantly. The optimal glycerol con-

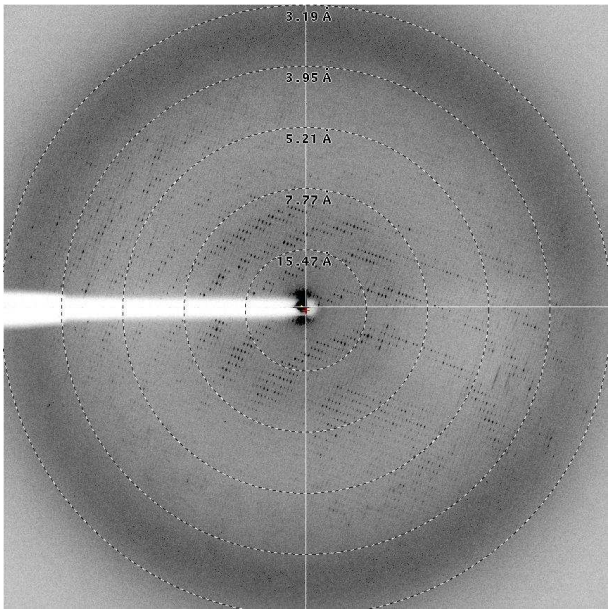

Figure 2.3: A diffraction image of a crystal of the CSA-DDB1 complex, in which the anisotropic nature of the diffraction can be seen.

Table 2.1: Data collection statistics.

Wavelength (Å)	0.9334
Space group	$P3_1$
Resolution (Å)	2.92 (3.08-2.92)[a]
No. measured reflections	610127 (49472)
No. unique reflections	119918 (15498)
Completeness (%)	97.8 (86.4)
Multiplicity	5.1 (3.2)
Mean I/σ(I)	4.9 (1.1)
R_{merge}	0.369 (1.108)
R_{pim}	0.181 (0.708)

[a] Values in parentheses are for the highest resolution shell

centration was found to be 5–7 % and crystals grown in this condition grew to 0.5 by 0.07 by 0.07 mm and diffracted to around 3 Å (Figure 2.2). Silver stained gels of washed crystals confirmed that the crystals contained both DDB1 and CSA.

Processing of the diffraction data showed that the crystals belonged to space group $P6_222$ or $P6_422$ according to Pointless (Evans, 2006) with unit cell parameters a=b=142.03 Å, c= 250.19 Å. However, the L and H twinning tests output by CTRUNCATE (Padilla & Yeates, 2003) indicated that the crystal was twinned. Many crystals were tested and all were perfectly twinned. The correct space group will be determined through molecular replacement and refinement by testing all possible subgroups and enantiomorphs of $P3_1$ (see Table 2.1 for data collection statistics for the crystal in the lowest possible subgroup $P3_1$).

The diffraction was anisotropic (Figure 2.3) and the anisotropic diffraction server (Strong *et al*., 2006) indeed showed that F/σ>3 at 2.9 Å in the direction of c*, but 3.3 Å in a* and b*. We chose a diffraction cutoff of 2.92 Å for our data, because data in this region is still useful though its I/σ is low owing to the anisotropy. It should be noted that the data can be better than the I/σ suggests, because the possible non-crystallographic symmetry (NCS) increases the signal to noise ratio. For example, the I/σ of the data merged in $P6_422$ is 1.9 in the highest resolution shell of 3.08-2.92 Å. The calculated Matthews coefficient is 2.08 $Å^3Da^{-1}$ assuming four DDB1-CSA complexes in the ASU.