Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Overproduction, purification, crystallization and preliminary X-ray diffraction analysis of Cockayne syndrome protein A in complex with DNA damage-binding protein 1

Elisabeth M. Meulenbroek and Navraj S. Pannu*

Biophysical Structural Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

Correspondence e-mail: raj@chem.leidenuniv.nl

Received 12 August 2011
Accepted 31 October 2011

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 PEG 3350 and sodium citrate as crystallizing agents. The crystals had unit-cell parameters $a = b = 142.03$, $c = 250.19$ Å and diffracted to 2.9 Å resolution on beamline ID14-1 at the European Synchrotron Radiation Facility.

1. Introduction

The human genome is constantly subjected to damaging agents such as ultraviolet irradiation. One of the major pathways for removal of DNA lesions is nucleotide excision repair (NER), which is responsible for the removal of 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 proteins A and B (CSA and CSB). These proteins have been proposed to have the following functions (Fousteri *et al.*, 2006). When RNA polymerase II encounters a DNA lesion during transcription, it becomes stalled and this stabilizes its interaction with CSB. This leads to recruitment of the core NER factors, a chromatin remodeller (HAT p300) as well as CSA in complex with DNA damage-binding protein 1 (DDB1) and E3-ubiquitin-ligase/CSN, which in turn recruits a chromatin remodeller (HMGN1), a scaffolding factor for protein-complex formation (XAB2) and a transcription cleavage factor (TFIIS). The recruitment of the CSA–DDB1–E3-ubiquitin-ligase/CSN complex, which is inactive for E3-ubiquitin-ligase activity upon recruitment, might be to prevent the degradation of the stalled RNA polymerase and/or other TC-NER factors in the early stages of repair, or it might be involved in the degradation of CSB at a later stage of the repair process. The degradation of CSB may be important for recovery of transcription after TC-NER is completed (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). 18 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 such as 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 by 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. Materials and methods

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 (Invitrogen). Overproduction in *Escherichia coli* strains BL21 Rosetta, RIL, RP, pLysS and pLysS Star was attempted at 277, 293 and 310 K with 0.1–1 m*M* 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 for an N-terminal 6×His tag and a thrombin cleavage site were inserted in front of the ORF and this sequence 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 extra residues at its N-terminus: MHHHHHHRRLVPRGSGGR.

CSA was amplified from the pET52b construct with a C-terminal 10×His tag and thrombin cleavage site using the 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 lacked the C-terminal glycine and has the following extra residues at its C-terminus: LALVPRGSSAHH-HHHHHHHH. 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). The recombinant bacmid was isolated and transfected into 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 × 10^6 cells ml^{-1} using 2 ml fourth-generation virus per 100 ml cell culture in suspension. The cells were harvested 66 h post-infection.

2.2. Purification

All purification steps were executed at 277 K. Harvested cells were resuspended in lysis buffer consisting of 50 m*M* Tris pH 8, 20 m*M* imidazole, 200 m*M* NaCl, 0.1% Triton X-100, 5 m*M* 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 rev min^{-1} (61 700*g*) for 30 min at 277 K in a Beckman Coulter 70.1 Ti preparative rotor. The supernatant was loaded onto a HisTrap HP column (GE Healthcare) using an ÄKTAexpress system (GE Healthcare) and washed with 20 column volumes of Ni buffer *A* (20 m*M* Tris pH 8, 30 m*M* imidazole, 200 m*M* NaCl, 5 m*M* 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* (20 m*M* Tris pH 8, 330 m*M* imidazole, 200 m*M* NaCl, 5 m*M* β-mercaptoethanol).

The Ni column fractions containing the CSA–DBB1 complex were diluted three times with 20 m*M* HEPES pH 7.2 and loaded onto a HiTrap Q HP column (GE Healthcare). The column was washed with Q buffer *A* (20 m*M* HEPES pH 7.2, 100 m*M* NaCl, 5 m*M* β-mercaptoethanol). The protein complex was eluted with a gradient of 25 column volumes to 100% Q buffer *B* (20 m*M* HEPES pH 7.2, 1 *M* NaCl, 5 m*M* β-mercaptoethanol). The fractions containing CSA–DDB1 were loaded onto a Superdex 200 gel-filtration column (GE Healthcare) equilibrated with 20 m*M* HEPES pH 7.2, 200 m*M* NaCl, 5 m*M* DTT and the column was run at a flow rate of 0.3 ml min^{-1}. The protein purity was assessed on 10% SDS–PAGE stained with Coomassie Blue (Simply Blue Safe stain, Invitrogen) and with silver staining (Silver Stain Plus, Bio-Rad).

2.3. Crystallization

CSA–DDB1 was concentrated to 5–8 mg ml^{-1} using a 10 kDa molecular-weight cutoff centrifugal filter unit (Millipore). Crystallization conditions were screened by 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 robot (Tecan) was used to pipette the reservoir solution (75 µl) and an Oryx6 robot (Douglas Instruments) was used to pipette the drops. After 2 d, small needle-like crystals appeared in condition Nos. 71, 83 and 95 from PACT, which consisted of 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 a U-MWU2 filter (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, Additive Screen (Hampton Research) was used following the manufacturer's instructions. 3% glycerol was found to cause a significant improvement in the size and morphology of the crystals. After additional optimization, addition of 5–7% glycerol turned out to be optimal with drops of between 2 and 4 µl using a 2:1 or 3:1 volume ratio of protein solution to crystallant solution.

2.4. X-ray diffraction analysis

Crystals were picked up in cryoloops and soaked in a solution consisting of mother liquor and 10–12% glycerol before flash-cooling them in liquid nitrogen. X-ray diffraction experiments were performed on ID14-1 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. 180 images were collected with an oscillation angle of 1.0° and an exposure time of 13 s per frame on ID14-1 at a wavelength of 0.9334 Å at 100 K. Images were processed with *iMOSFLM* (Battye *et al.*, 2011). Scaling and merging was performed with *SCALA* from the *CCP*4 suite (Winn *et al.*, 2011).

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. However, cleavage of the MBP tag with TEV protease led to precipitation of CSA. Gel filtration of the MBP-fusion protein showed that the fusion protein was present as a high-molecular-weight entity, presumably a soluble aggregate.

Coproduction of CSA and its interacting partner DDB1 was then attempted in insect cells, as the structurally and functionally related

complex DDB1–DDB2 has been coproduced and purified from insect cells (Scrima *et al.*, 2008). The coproduction of CSA and DDB1 in Sf9 cells yielded a soluble protein complex and a large excess of DDB1 (more than a threefold 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% buffer *B*) while the CSA–DDB1 complex eluted in the second part of the gradient (30–100% buffer *B*). CSA copurified with DDB1 as a 1:1 complex as judged from SDS–PAGE throughout the entire purification; no free CSA was obtained. The complex eluted as a single peak from the gel filtration at around 170 kDa. This indeed approximates the size of a 1:1 complex of CSA and DDB1, the theoretical sizes of which are 46 and 129 kDa in our constructs. Western blotting with antibodies against DDB1 and CSA and mass spectrometry of 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 varying PEG concentration and pH. The initial crystals were around 0.09 × 0.01 × 0.01 mm in size and diffracted to around 8 Å resolution at the ESRF. An additive screen was performed and 3% glycerol was found to improve the crystal size and quality significantly. The optimal glycerol concentration was found to be 5–7% and crystals grown in this condition grew to 0.5 × 0.07 × 0.07 mm in size and diffracted to around 3 Å resolution (Fig. 1). Silver-stained gels of washed crystals confirmed that the crystals contained both DDB1 and CSA.

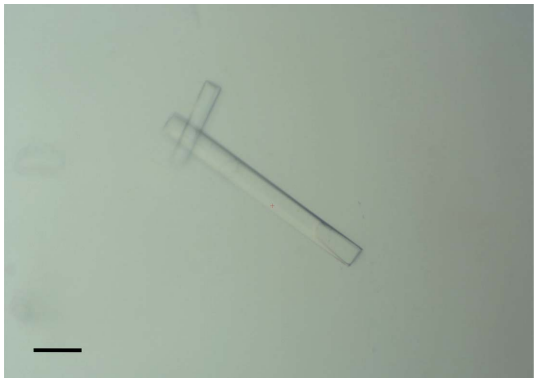

Figure 1
Crystals of CSA in complex with DDB1. The scale bar is 100 µm in length.

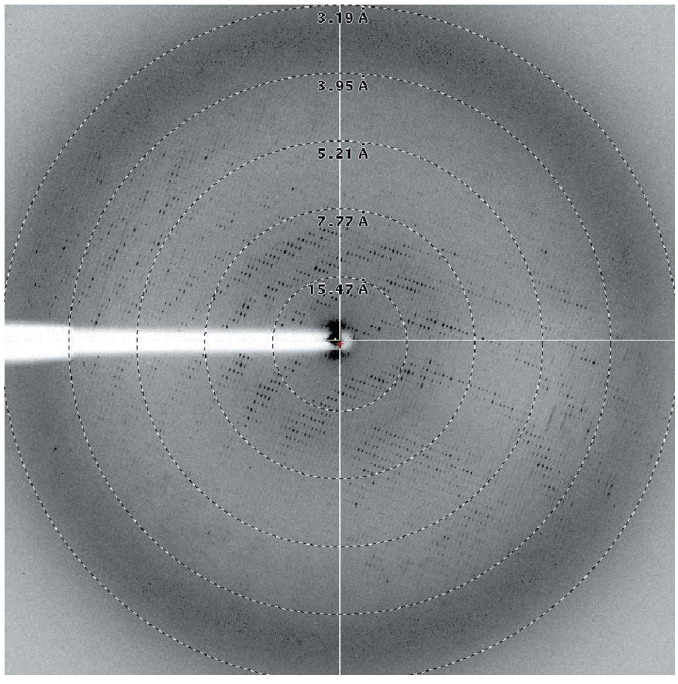

Figure 2
A diffraction image of a crystal of the CSA–DDB1 complex in which the anisotropic nature of the diffraction can be seen.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	0.9334
Space group	$P3_1$
Resolution (Å)	2.92 (3.08–2.92)
No. of measured reflections	610127 (49472)
No. of unique reflections	119918 (15498)
Completeness (%)	97.8 (86.4)
Multiplicity	5.1 (3.2)
Mean $I/\sigma(I)$	4.9 (1.1)
$R_{\rm merge}$	0.369 (1.108)
$R_{\rm p.i.m.}$	0.181 (0.708)

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 $a = b = 142.03$, $c = 250.19$ Å (see Table 1 for data-collection statistics). 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 $P6_422$ (see Table 1 for data-collection statistics for the crystal in the lowest possible subgroup $P3_1$).

The diffraction was anisotropic (Fig. 2) and the Diffraction Anisotropy Server (Strong *et al.*, 2006) indeed showed that $F/\sigma(F) > 3$ at 2.9 Å in the direction of c^* but at 3.3 Å in a^* and b^*. We chose a diffraction cutoff of 2.92 Å for our data because data in this region are still useful although the $I/\sigma(I)$ is low owing to the anisotropy. It should be noted that the data can be better than the $I/\sigma(I)$ suggest because the possible NCS increases the signal-to-noise ratio. For example, the $I/\sigma(I)$ of the data merged in $P6_422$ is 1.9 in the highest resolution shell (3.08–2.92 Å).

The calculated Matthews coefficient is 2.08 Å^3 Da^{-1} assuming the presence of four DDB1–CSA complexes in the asymmetric unit.

The authors would like to thank Bobby Florea for help with mass spectrometry, Patrick Voskamp and Hans den Dulk for help and advice in protein overproduction in insect cells and Igor Nederlof for help in using the crystallization robot. Furthermore, we would like to thank Wim Vermeulen for providing us with a pFastbac with the cDNA of CSA, Imre Berger for providing us with DH10EMBacY, and Andrea Scrima and Nicolas Thomä for providing a vector with the cDNA of DDB1. We acknowledge the European Synchrotron Radiation Facility for the provision of synchrotron-radiation facilities, and Tobias Klar and Mats Okvist for their help in using beamline ID14-1. EMM and NSP were supported by grants from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (grant Nos. 021.002.024 and 700.55.425).

References

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Berger, I., Fitzgerald, D. J. & Richmond, T. J. (2004). *Nature Biotechnol.* **22**, 1583–1587.

Dümmler, A., Lawrence, A. M. & de Marco, A. (2005). *Microb. Cell Fact.* **4**, 34.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Fousteri, M., Vermeulen, W., van Zeeland, A. A. & Mullenders, L. H. F. (2006). *Mol. Cell*, **23**, 471–482.
Groisman, R., Kuraoka, I., Chevallier, O., Gaye, N., Magnaldo, T., Tanaka, K., Kisselev, A. F., Harel-Bellan, A. & Nakatani, Y. (2006). *Genes Dev.* **20**, 1429–1434.
Iovine, B., Ianella, M. L. & Bevilacqua, M. A. (2011). *Int. J. Biochem. Cell Biol.* **43**, 1664–1667.
Laugel, V. *et al.* (2010). *Hum. Mutat.* **31**, 113–126.
Li, T., Chen, X., Garbutt, K. C., Zhou, P. & Zheng, N. (2006). *Cell*, **124**, 105–117.
Nance, M. A. & Berry, S. A. (1992). *Am. J. Med. Genet.* **42**, 68–84.
Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst*. D**59**, 1124–1130.
Scrima, A., Konícková, R., Czyzewski, B. K., Kawasaki, Y., Jeffrey, P. D., Groisman, R., Nakatani, Y., Iwai, S., Pavletich, N. P. & Thomä, N. H. (2008). *Cell*, **135**, 1213–1223.
Strong, M., Sawaya, M. R., Wang, S., Philips, M., Cascio, D. & Eisenberg, D. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 8060–8065.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.