Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, purification, crystallization and preliminary X-ray studies of RFC boxes II–VIII of replication factor C from *Methanococcus jannaschii*

Ick Lee, Neratur K. Lokanath, Kyeongsik Min, Sung Chul Ha, Dong Young Kim and Kyeong Kyu Kim*

Department of Molecular Cell Biology, Center for Molecular Medicine, SBRI, Sungkyunkwan University School of Medicine, Suwon 440-746, South Korea

Correspondence e-mail: kkim@med.skku.ac.kr

Received 2 October 2001
Accepted 17 December 2001

Replication factor C (RFC) is the accessory protein required to load the proliferating cell nuclear antigen (PCNA) onto DNA in replication process. RFC is composed of several subunits and each subunit contains the highly conserved sequences RFC boxes II–VIII. RFC boxes II–VIII of the large subunit of replication factor C from *Methanococcus jannaschii* has been overexpressed in *Escherichia coli*, purified and crystallized at 295 K using ammonium sulfate as precipitant. Crystals belong to the space group *R*32, with unit-cell parameters $a = b = 238.23$ (5), $c = 73.17$ (12) Å. Native data were collected at 100 K to a resolution of 3.2 Å using a synchrotron-radiation source.

1. Introduction

Replication factor C (RFC) is a eukaryotic heteropentameric protein required for DNA replication and repair processes by loading proliferating cell nuclear antigen (PCNA) onto DNA in an ATP-dependent manner (Tsurimoto & Stillman, 1991). Upon association of PCNA, a mobile clamp which can slide along DNA chain, with either DNA polymerase δ or ε (Hindges & Hübscher, 1997), DNA synthesis commences with high-level processivity (Podust *et al.*, 1992). Prior to the loading of PCNA to DNA, RFC binds to DNA in order to insert DNA into the channel of PCNA as a clamp loader. Human RFC (hRFC), first identified as an essential factor required for simian virus 40 DNA replication *in vitro*, consists of four small RFC proteins and a single large RFC protein (Tsurimoto & Stillman, 1989). Each of the subunits possesses regions of highly conserved sequences (termed RFC boxes II–VIII), which share homology to the prokaryotic clamp loaders (Cullmann *et al.*, 1995). The boxes are between three and 16 amino acids in length and are clustered within the N-terminal portion of the small subunits, but are centrally located in the large subunit.

Archaea, the third domain of life (Woese & Fox, 1977), have been proposed to possess DNA replication that appears to be similar to that of eukarya (Bernander, 2000). The existence in archaeal cells of proteins similar to the essential factors for eukaryotic DNA replication heightens our expectation that unravelling the archaeal mechanism will contribute to the understanding of the replication mechanism in eukarya (Cann & Ishino, 1999). An archaeal clamp loader, replication factor C (RFC), consists of two proteins, a small subunit (RFCS) and a large subunit (RFCL), whose sequences are both highly homologous to those of the eukaryotic RFC components. Biochemical characterizations of archaeal RFCs from *Sulfolobus solfataricus* (Pisani *et al.*, 2000) and *Pyrococcus furiosus* (Cann *et al.*, 2001) proved that the role of RFC in the replication process is conserved in archaea and eukarya. In addition, a recent structural study of the small subunit of RFC from *P. furiosus* suggested that there are functional and structural homologies between archaeal RFC and prokaryotic clamp loaders (Oyama *et al.*, 2001). In order to understand the general mechanism of clamp loading at atomic resolution, we purified the most conserved domains, RFC boxes II–VIII, of the large subunit of replication factor C from *Methanococcus jannaschii* and report here the crystallization conditions and preliminary X-ray data of its crystals.

2. Bacterial expression and purification

The gene encoding RFC boxes II–VIII (amino-acid residues numbered 1–208) in the large subunits of replication factor C (MJ0884) was amplified by the polymerase chain reaction (PCR) using *M. jannaschii* genomic DNA as a template. The PCR product was inserted into *Xcm*I-digested pBluescript (Stratagene) and digested with restriction enzymes *Nde*I and *Bam*HI, since direct restriction of PCR product with *Nde*I and *Bam*HI was not successful. The DNA fragment was inserted into an *Nde*I/*Bam*HI-restricted pET22b vector (Novagen). The sequence identity of RFC boxes II–VIII was confirmed by nucleotide sequencing in an

ABI377 DNA sequencer. *E. coli* ER2566 cells transformed with pET22b-MjRFC plasmid were grown at 310 K in LB medium containing 5 μg ml^{-1} ampicillin. RFC boxes II–VIII were induced with 0.5 m*M* isopropyl-β-D-thiogalactoside (IPTG) for 6 h. The cells were harvested at 277 K by centrifugation at 5000*g* for 6 min, suspended in buffer *A* (25 m*M* Tris–HCl pH 7.5, 10 m*M* $MgCl_2$, 1 m*M* EDTA and 1 m*M* phenylmethylsulfonyl fluoride) and lysed using an ultrasonic processor. The crude lysate was centrifuged at 15 000*g* for 40 min at 277 K and the supernatant was heated and kept between 353 and 358 K for 5 min in order to denature contaminating heat-labile proteins. The supernatant was applied onto a 5 ml HiTrap Q column (Amersham-Pharmacia) equilibrated with buffer *A*. The fractions containing RFC boxes II–VIII were eluted with a linear gradient of 0–0.35 *M* NaCl. RFC boxes II–VIII was dialyzed against buffer *B* (10 m*M* Tris–HCl pH 8.9, 10 m*M* $MgCl_2$, 1 m*M* EDTA and 1 m*M* phenylmethylsulfonyl fluoride) and subjected to a 5 ml HiTrap SP column (Amersham-Pharmacia), employing a linear gradient of 0–0.25 *M* NaCl in buffer *B*. The fractions containing the protein were pooled and concentrated by ultrafiltration (Amicon, YM10) and loaded onto a gel-filtration column (Superdex 75, Amersham-Pharmacia) equilibrated with buffer *A*. The purified protein was homogenous on SDS–PAGE and native PAGE.

3. Crystallization

The protein was prepared at a concentration of 12 mg ml^{-1} in buffer *A*. Crystallization was performed by the hanging-drop vapour-diffusion method with the drops equilibrated against 500 μl of reservoir solution at 295 K. Initial crystallization conditions were tested using the Crystal Screen I, Crystal Screen II, MemFac Screen and Natrix Screen solutions (Hampton Research). Rod-shaped crystals were obtained using ammonium sulfate as precipitant. Further refinement of this condition gave crystals suitable for X-ray experiments. The best crystals were obtained from sitting drops prepared by mixing 1 μl of protein solution and 1 μl reservoir solution in microbridges (Hampton Research). Crystals grew to final dimensions of 0.2 × 0.2 × 0.6 mm within a week under the optimized reservoir condition of 50 m*M* HEPES pH 7.5, 10 m*M* $MgCl_2$ and 1.6 *M* ammonium sulfate at 295 K (Fig. 1).

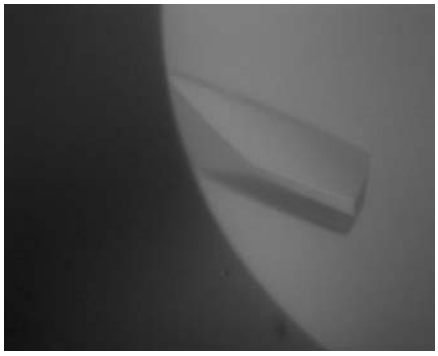

Figure 1
Crystal of RFC boxes II–VIII grown in ammonium sulfate as precipitant. Crystal dimensions are approximately 0.2 × 0.2 × 0.6 mm.

4. X-ray data collection and processing

Before flash-freezing the crystal in a nitrogen-gas stream at 100 K, it was dipped into a solution containing 30%(*v*/*v*) glycerol in addition to the reservoir solution. Native data from the crystal grown in ammonium sulfate were collected at 100 K using a Weissenberg camera for macromolecular crystallography at the BL-6B experimental station of the Photon Factory, Tsukuba, Japan (Sakabe, 1991). The synchrotron X-ray wavelength was 1.000 Å and a 0.2 mm collimator was used. The image plate (40 × 80 cm, Fuji BASIII) was placed at a distance of 573.0 mm from the crystal. The oscillation range per image plate was 2.5° with no overlap. The diffraction patterns recorded on the image plates were digitized by the off-line scanner IPR4080 (Rigaku). The raw data were processed and scaled using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The crystals belonged to the rhombohedral space group *R*32, with unit-cell parameters $a = b = 238.23$ (5), $c = 73.17$ (12) Å. A native data set was collected at 3.2 Å resolution with an $R_{\rm merge}$ of 5.2%; data statistics are summarized in Table 1. The data were strongly anisotropic; the eigenvalues of the anisotropic distribution of the structure factor were 0.7336:0.7289:1.0 as estimated by *SFCHECK* (Vaguine *et al.*, 1999). Since diffraction data were anisotropic along the *c* axis (Fig. 2), the low completeness in the last resolution shell was expected (Table 1). Anisotropic diffraction patterns are usually observed from proteins of elongated shape, such as potassium channel (Doyle *et al.*, 1998) or serine chemotaxis receptor (Kim *et al.*, 1999). Therefore, it could be expected that RFC boxes II–VIII might have an elongated structure. Assuming a molecular weight of 23 690 Da and two, three or four molecules in the asymmetric unit, the values of the crystal packing parameter $V_{\rm M}$ are 4.87, 3.25 or 2.43 Å^3 Da^{-1}, respectively (Matthews, 1968). However, twofold non-crystallographic symmetry axes were found at $\omega = 90$, $\varphi = 30 + (60n)^\circ$ (where *n* is an integer), with a peak intensity of 78% of the origin peak, *via* a self-rotation search using the program *POLARRFN* from the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). Therefore, it is assumed that there are two molecules in an asymmetric unit and that the solvent content is 75% by volume. The crystal structure determination of the *M. jannaschii* RFC

Table 1
X-ray data-collection statistics .

Values in square brackets refer to the highest resolution shell.

Space group	*R*32
Unit-cell parameters (Å, °)	$a = b = 238.23$ (05), $c = 73.17$ (12), $\alpha = \beta = 90$, $\gamma = 120^\circ$
Resolution range (Å)	100.0–3.2 [3.31–3.20]
Unique reflection	11414 [690]
Redundancy	9 [1.7]
$R_{\rm merge}$† (%)	5.2 [30.2]
Completeness (%)	86.5 [52.8]
Average $I/\sigma(I)$	15.5 [1.6]

† $R_{\rm merge} = \sum_{hkl}[\sum_i(|I_{hkl,i} - \langle I_{hkl,i}\rangle|)]/\sum_{hkl,i}\langle I_{hkl}\rangle$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices *h*, *k* and *l*, and $\langle I_{hkl}\rangle$ is the average intensity of that reflection.

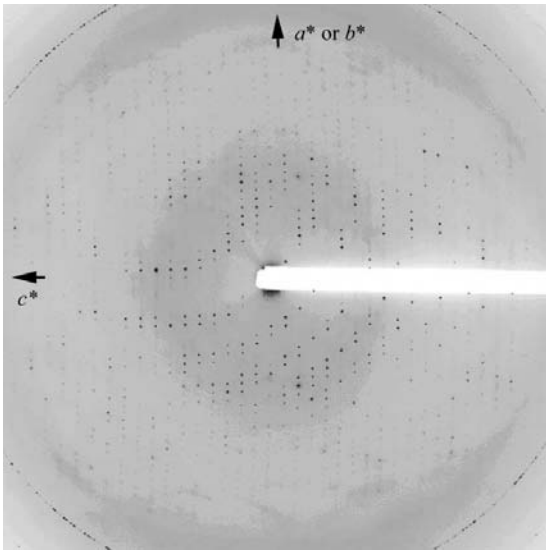

Figure 2
X-ray diffraction photograph of RFC boxes II–VIII crystal with 2.5° oscillation range. *a** or *b** and *c** axes are indicated.

boxes II–VIII is now under way by MAD methods.

We thank Professor N. Sakabe and beamline scientists for their assistance during data collection at BL-6B, Photon Factory, Japan. This work was supported by grant number FG-3-5-01 of 21C Frontier Functional Human Genome Project from the Ministry of Science and Technology of Korea.

References

Bernander, R. (2000). *Trends Microbiol.* **8**, 278–283.

Cann, I. K. O. & Ishino, Y. (1999). *Genetics*, **152**, 1249–1267.

Cann, I. K. O., Ishino, S., Yuasa, M., Daiyasu, H., Toh, H. & Ishino, Y. (2001). *J. Bacteriol.* **83**, 2614–2623.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cullmann, G., Fien, K., Kobayashi, R. & Stillman, B. (1995). *Mol. Cell. Biol.* **15**, 4661–4671.

Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. & MacKinnon, R. (1998). *Science*, **280**, 69–77.

Hindges, R. & Hübscher, U. (1997). *J. Biol. Chem.* **378**, 345–362.

Kim, K. K., Yokota, H. & Kim, S.-H. (1999). *Nature (London)*, **400**, 787–792.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Oyama, T., Ishino, Y., Cann, I. K. O., Ishino, S. & Morikawa, K. (2001). *Mol. Cell*, **8**, 455–463.

Pisani, M. N., De Felice, M., Carpentieri, F. & Rossi, M. (2000). *J. Mol. Biol.* **301**, 61–73.

Podust, L. M., Georgaki, A., Strack, B. & Hübscher, U. (1992). *Nucleic Acids Res.* **22**, 2970–2975.

Sakabe, N. (1991). *Nucl. Instrum. Methods A*, **303**, 448–463.

Tsurimoto, T. & Stillman, B. (1989). *Mol. Cell. Biol.* **9**, 609–619.

Tsurimoto, T. & Stillman, B. (1991). *J. Biol. Chem.* **266**, 1950–1960.

Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.

Woese, C. R. & Fox, G. E. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 5088–5090.