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

Expression, purification and preliminary X-ray analysis of proliferating cell nuclear antigen from the archaeon *Thermococcus thioreducens*

Miranda L. Byrne-Steele[a] and Joseph D. Ng[a,b]*

[a]Laboratory for Structural Biology, Department of Biological Sciences, University of Alabama in Huntsville, Huntsville, AL 35899, USA, and [b]ExtremoZyme Inc., HudsonAlpha Institute for Biotechnology, 601 Genome Way, Huntsville, AL 35806, USA

Correspondence e-mail: ngj@uah.edu

Received 23 June 2009
Accepted 30 July 2009

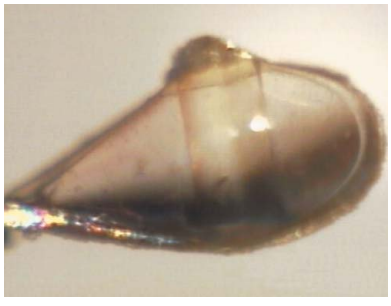
Proliferating cell nuclear antigen (PCNA) is a DNA sliding clamp which confers processivity on replicative DNA polymerases. PCNA also acts as a sliding platform that enables the association of many DNA-processing proteins with DNA in a non-sequence-specific manner. In this investigation, the PCNA from the hyperthermophilic archaeon *Thermococcus thioreducens* (TtPCNA) was cloned, overexpressed in *Escherichia coli* and purified to greater than 90% homogeneity. TtPCNA crystals were obtained by sitting-drop vapor-diffusion methods and the best ordered crystal diffracted to 1.86 Å resolution using synchrotron radiation. The crystals belonged to the hexagonal space group $P6_3$, with unit-cell parameters $a = b = 89.0$, $c = 62.8$ Å. Crystals of TtPCNA proved to be amenable to complete X-ray analysis and future structure determination.

1. Introduction

DNA sliding clamps are a central component of the replisome and are necessary for the maintenance of hereditary information and its transfer from generation to generation. PCNA is a member of the sliding-clamp family of proteins and acts as a processivity factor during DNA replication by tethering the replicative polymerases to the genomic template. However, PCNA also interacts with a wide variety of DNA-processing proteins which are involved in many cellular processes beyond DNA replication. These interacting partners associate with the DNA sliding clamp through their PCNA-interacting protein (PIP) motif and bind near the interdomain-connecting loop (IDCL) region of PCNA (Maga & Hubscher, 2003; Warbrick, 2000). Therefore, PCNA can also be regarded as a sliding platform that enables the association of many DNA-processing proteins with DNA in a non-sequence-specific manner (Warbrick, 2000).

All kingdoms of life maintain a toroidal-shaped DNA sliding clamp with pseudo-sixfold symmetry; however, the multimeric association of the monomeric subunits differs depending upon the kingdom from which the clamp originates. The bacterial clamp, or the DNA polymerase III β subunit, consists of a homodimeric assembly in which each subunit contains three topologically similar domains (Kong *et al.*, 1992; Argiriadi *et al.*, 2006). In contrast, eukaryotic and archaeal PCNAs assemble into a trimeric rather than a dimeric ring, in which each subunit contains two rather than three similar domains (Chapados *et al.*, 2004; Krishna *et al.*, 1994; Matsumiya *et al.*, 2001). While most eukaryotic and archaeal rings are homotrimeric, the situation is further complicated in archaea; the functional form of the PCNA from the crenarchaeaon *Sulfolobus solfactaricus* is a heterotrimer composed of three distinct PCNA homologues (Williams *et al.*, 2006; Hlinkova *et al.*, 2008; Dionne *et al.*, 2003).

Thermococcus thioreducens is a hyperthermophilic sulfur-reducing euryarchaeote with an optimal growth range of 356–358 K isolated from the Rainbow hydrothermal vent site on the Mid-Atlantic Ridge (Pikuta *et al.*, 2007). Genomic data indicate the presence of only one PCNA homolog, which is consistent with other euryarchaeal genomes. Here, we report the cloning, recombinant expression and crystallization of *T. thioreducens* PCNA, referred to in the following as TtPCNA. TtPCNA consists of 249 amino acids, resulting in a



calculated molecular weight of 28.0 kDa. The preliminary results described here will allow the future structure determination of TtPCNA by X-ray crystallography for structure–function studies.

2. Materials and methods

2.1. Cloning

The *pcna* gene (accession code EF058196) was PCR-amplified (Mullis & Faloona, 1987) from *T. thioreducens* genomic DNA using the forward primer 5′-tttgtttaactttaagaaggagatatacatATGCCGTTC-GAGATAGTTTTT-3′ and the reverse primer 5′-cttcctttcgggctttgtta-gcagccggatccTTAATCCTCTACACGCGGTGC-3′ (Operon, Huntsville, Alabama, USA). The primers were designed from exact alignment of the first and last 21 nucleotides of the PCNA open reading frame (upper-case letters). The oligonucleotide primers were also preceded by linker sequences corresponding to 30 and 32 nucleotides of the expression plasmid vector pET3a (Novagen) containing *Nde*I and *Bam*HI insertion sites, respectively (lower-case letters). The PCR amplification was performed in a 50 µl volume containing 1 µl *Pfu* Ultra II fusion HS DNA polymerase (Stratagene, La Jolla, California, USA), 2 µl each of both forward and reverse primers at 25 pmol µl^{-1} and 1 µl genomic DNA at approximately 100 ng µl^{-1}. The reaction was performed using 30 cycles of 368 K denaturation for 20 s, 328 K annealing for 20 s and 345 K extension for 90 s.

In vivo homologous recombination was used to subclone the amplification product into a specially prepared blunt-ended propagating plasmid vector. This method is based on the ability of many *Escherichia coli* strains (including the RecA-deficient ones used in cloning) to perform *in vivo* intermolecular recombination between DNA fragments sharing homologous sequences at their ends (Bubeck *et al.*, 1993; Jones & Howard, 1991; Marsic *et al.*, 2008; Oliner *et al.*, 1993). Gene fragments can be efficiently inserted into a linearized target plasmid vector without restriction digest, ligation or other enzymatic manipulation. 2 µl of the amplified product was mixed with 50 ng of the prepared linearized pET3a and used for transformation. The DNA mixture was gently mixed with 50 µl of *E. coli* strain NovaBlue competent cells. After 15 min incubation on ice, cells were heat-shocked at 315 K for 60 s. After resting for 2 min on ice, 150 µl Luria–Bertani (LB) medium was added to the transformed cells, incubated for 1 h at 310 K and spread on LB–agar plates containing 100 mg l^{-1} ampicillin. After 16 h incubation at 310 K, colonies were picked and grown in 5–10 ml LB containing 100 mg l^{-1} ampicillin at 310 K and shaken at 250 rev min^{-1} for 12 h. Plasmids were purified using EZNA plasmid Miniprepkit II (Omeg Bio-Tek, Doraville, Georgia, USA). The accuracy of the sequence was verified by sequencing by Functional Biosciences (Madison, Wisconsin, USA) using T7 sequencing primers flanking the cloning site.

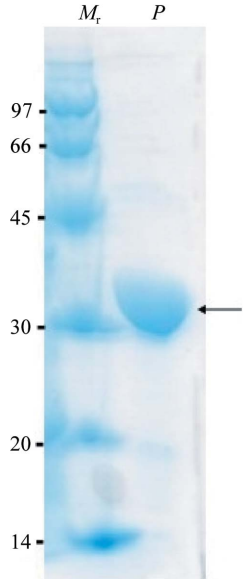


Figure 1
SDS–PAGE analysis of purified recombinant TtPCNA. Approximately 25 µg TtPCNA was analyzed on 12% polyacrylamide gel, which was stained with Coomassie Brilliant Blue and visualized by white-light illumination. The purified TtPCNA is indicated by an arrow in lane *P* and can be compared with the standard molecular-weight markers (kDa) in lane M_r. The recombinant protein was purposely overloaded to reveal any contaminations and its band nearly corresponds to the predicted size of 28 kDa with a purity of greater than 90% homogeneity.

Table 1
Data-collection statistics.

Space group	$P6_3$
Radiation	SER-CAT 22ID, APS
Unit-cell parameters (Å, °)	$a = b = 89.0$, $c = 62.8$, $\alpha = \beta = 90.0$, $\gamma = 120$
Temperature (K)	100
Wavelength (Å)	1.0
Oscillation range (°)	1.0
Crystal-to-detector distance (mm)	250
No. of frames	360
Resolution range (Å)	50.0–1.86 (1.93–1.86)
Total reflections	498170
Unique reflections	46243
Average redundancy	10.8
Average $I/\sigma(I)$	36.9 (3.9)
Overall B factor (Å^2)	27.7
Completeness (%)	98.8 (88.8)
$R_{\rm merge}$† (%)	5.6 (26.4)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of reflection hkl and $I_i(hkl)$ is the ith measurement of the intensity of reflection hkl.

2.2. Recombinant expression

Purified plasmid was transformed into competent *E. coli* BL21(DE3)-Rosetta cells (Novagen, Madison, Wisconsin, USA) as described above. Transformed cells were mixed with 150 µl LB medium, incubated for 1 h at 310 K and spread on LB–agar plates containing 100 mg l^{-1} ampicillin and 35 mg l^{-1} chloramphenicol. Overnight colonies were then used to inoculate 20 ml LB medium containing 50 mg l^{-1} ampicillin and 35 mg l^{-1} chloramphenicol. The cells were grown overnight at 310 K and shaken at 250 rev min^{-1}. The overnight culture was added equally to 2 × 2 l LB medium containing the same antibiotic concentrations and grown at 310 K while agitating at 250 rev min^{-1} until the optical density at 600 nm reached 0.6. Cells were induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 m*M* and incubated at 310 K with continuous shaking at 250 rev min^{-1} for 6 h. Cells containing recombinant protein were harvested by centrifugation at 3500*g* for 20 min at 277 K. Glycerol stocks of expressing cells were made for future use by adding 100 µl glycerol to 900 µl uninduced culture followed by a quick freeze in liquid nitrogen and storage at 193 K.

2.3. Purification

All purification steps were performed at 277 K. Harvested cells were resuspended in buffer *A* (50 m*M* Tris pH 8.0, 25 m*M* NaCl, 1 m*M* EDTA) and lysed *via* sonication with six cycles and 45 pulses using a Branson Sonifier 250 (VWR Scientific, West Chester, Pennsylvania, USA). Cell debris was removed by centrifugation at 12 000*g* for 20 min. The supernatant was heated for 30 min at 350 K and the

precipitate was further removed by centrifugation at 12 000*g* for 20 min. The supernatant was loaded onto a HiTrap Q Sepharose anion-exchange column (GE Healthcare, UK) pre-equilibrated with buffer *A*. Protein was eluted from the column with a 0.05–1 *M* NaCl gradient in buffer *B* (50 m*M* Tris pH 8.0, 1 *M* NaCl, 1 m*M* EDTA) using an ÄKTA Explorer FPLC system (Amersham-Pharmacia, USA). Fractions containing the recombinant protein were pooled and concentrated to 2 ml using an Amicon Ultra centrifugal filter device (Millipore, USA). The protein was applied onto a Sephacryl S-200 gel-filtration column (Pharmacia, USA) pre-equilibrated with buffer *C* (50 m*M* Tris pH 8.0, 50 m*M* NaCl, 1 m*M* EDTA) and eluted with the same buffer. Fractions corresponding to the principal peak were collected and determined to be >90% homogenous by SDS–PAGE analysis (Fig. 1). The protein was concentrated to 15 mg ml^{-1} as determined by Bradford analysis (Bradford, 1976) and utilized for subsequent crystallization screening.

2.4. Crystallization

The protein was screened for crystallization by sitting-drop vapor-diffusion experiments in a 96-well format Intelli-Plates (Art Robbins, Sunnyvale, California, USA) against Crystal Screen I and II (Hampton Research, Aliso Viejo, California, USA) sparse-matrix screens (Jancarik & Kim, 1991). The final diffraction-quality crystals were grown after two weeks at 295 K by sitting-drop vapor diffusion using 1 µl protein in buffer *C* with 2 µl reservoir solution against 75 µl reservoir solution containing 10%(*w*/*v*) PEG 6000 and 2.0 *M* NaCl (Fig. 1).

2.5. X-ray data collection

Crystals were transferred to a cryoprotectant solution composed of reservoir solution supplemented with 20% glycerol for 5 s and directly flash-cooled in liquid nitrogen. X-ray diffraction data were collected from a single crystal using a MAR 300 CCD on the South East Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. All frames were collected at 100 K using a crystal-to-detector distance of 250 mm and a 1° oscillation angle. Diffraction data were processed and scaled with *HKL*-2000 (Otwinowski & Minor, 1997). Data-collection statistics are summarized in Table 1.

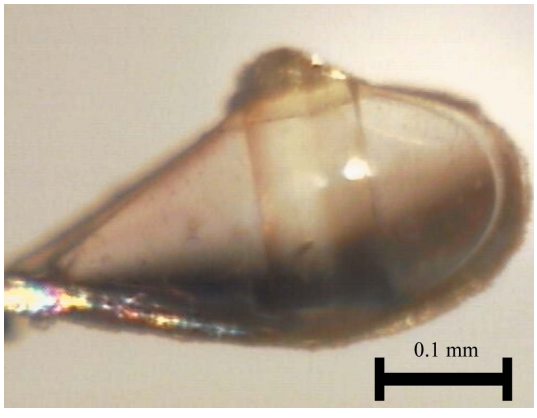


Figure 2
Single crystal of recombinant TtPCNA protein mounted on a cryoloop. The crystal was obtained from sitting-drop vapor-diffusion equilibration against 10%(*w*/*v*) PEG 6000 and 2.0 *M* NaCl. TtPCNA is shown mounted on a 0.3 mm diameter synthetic loop (Hampton Research) after cryogenic preparations. The longest dimensions of the crystals were measured to be 0.3 mm as visualized by a vertical CCD camera installed in the SER-CAT beamline hutch.

3. Results and discussion

The expression of recombinant TtPCNA did not show any significant adverse effects in the *E. coli* expression host as high-yield production of TtPCNA was possible. Over 50 mg recombinant protein can be overproduced in 1 l growth medium without any elaborate fermentation approaches. The intrinsic thermophilic resistant properties of TtPCNA provided an easy means of purifying the protein without any affinity tags that may interfere with subsequent structural studies. The employment of heat treatment during the first selection of the purification gave rise to more than 70% homogeneity based on SDS–PAGE analysis (data not shown). A protein purity of more than 90% homogeneity can be obtained by two chromatography steps after the heat selection treatment as determined by SDS–PAGE analysis (Fig. 1). The estimated molecular weight of the purified TtPCNA, as measured from the SDS–PAGE gel, was near 28 kDa, which is

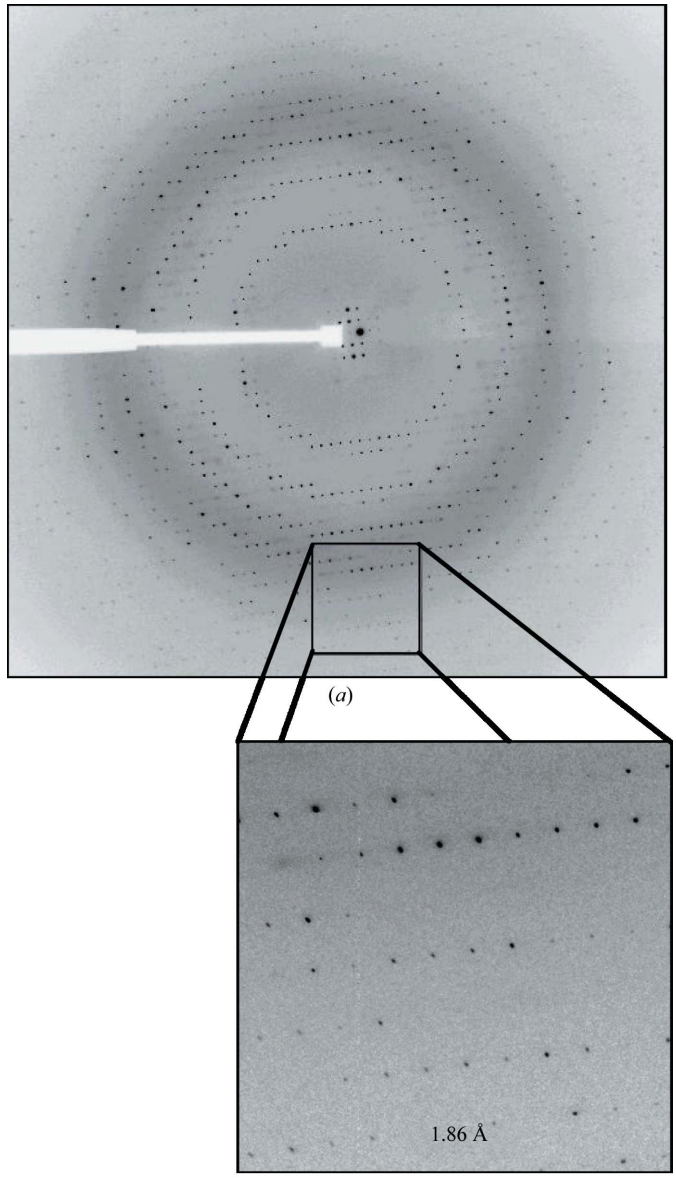


Figure 3
Representative X-ray diffraction image of TtPCNA. A typical 1° oscillation image obtained during data collection from a TtPCNA crystal is shown in (*a*). Reflections were measured to 1.86 Å at the outer edge with $I/\sigma(I)$ greater than 3, as demonstrated in (*b*).

consistent with the calculated molecular weight derived from the protein's primary sequence.

Prismatic TtPCNA crystals were obtained directly from the sparse-matrix screen without further optimization of the crystal-growth conditions. Crystals had typical dimensions of about 0.3 × 0.1 × 0.1 mm with a hexagonal habit (Fig. 2). Interestingly, TtPCNA was crystallized against a reservoir containing 2.0 *M* NaCl and 10% PEG 6000. Although it is common for PCNA crystallization conditions to contain ammonium sulfate (AS; Argiriadi *et al.*, 2006; Freudenthal *et al.*, 2009; Matsumiya *et al.*, 2001; Nishida *et al.*, 2006), there are a number of examples in which the primary precipitating agent lacks AS but usually contains either a high concentration of an alternative salt and/or an increased percentage of PEG (or a similar substance; Dore *et al.*, 2006; Williams *et al.*, 2006). Significantly, two other crystallization conditions were identified for TtPCNA that produced diffracting crystals. One of these conditions consisted of 2.0 *M* AS and 5%(v/*v*) 2-propanol, while the other did not contain any AS and included 0.02 *M* calcium chloride, 0.1 *M* sodium acetate pH 4.6 and 30%(*v*/*v*) 2-methyl-2,4-pentanediol (data not shown). The crystals obtained using these conditions did not diffract to the same resolution as those obtained using the TtPCNA condition presented here, even upon optimization attempts. However, these results demonstrate the ability of TtPCNA to crystallize in a traditional AS-type condition and an alternative condition that lacks AS without an improvement in resolution limit.

A complete diffraction data set for TtPCNA was collected to 1.85 Å resolution from a single crystal. The reflection spots were compact, had low mosaic spread and were easily indexed and scaled (Fig. 3). Data analysis revealed that the TtPCNA crystals belonged to the hexagonal space group $P6_3$, with unit-cell parameters $a = b = 89.0$, $c = 62.8$ Å. One molecule was assumed to be present in the asymmetric unit, resulting in a Matthews coefficient of 2.56 $Å^3\ Da^{-1}$, corresponding to a solvent content of 52.02% (Matthews, 1968). An initial solution for the structure of TtPCNA was obtained by molecular replacement using *MOLREP* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994) and PCNA from *Pyrococcus furiosus* (PDB code 1ge8; Matsumiya *et al.*, 2001), which shares 82% sequence identity with TtPCNA, as a search model. The overall correlation coefficient (CC) and *R* factor for the correct unrefined solution were 0.67 and 0.51, respectively, using data in the resolution range 48.6–3.7 Å. Examination of the best solution revealed good crystal packing and no clashes between symmetry-related molecules. Refinement of TtPCNA is in progress and details of the structure will be reported elsewhere.

MLB-S was supported by a fellowship from the Alabama EPSCoR Graduate Research Scholars Program (24229), which is funded by the Alabama State Legislature through the Alabama Commission on Higher Education. Additional funding was provided through a National Science Foundation EPSCoR grant (EPS-0447675). X-ray data were collected on Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. We thank Ronny C. Hughes and Damien Marsic for their technical assistance, and Edward J. Meehan for his support and helpful discussions.

References

Argiriadi, M. A., Goedken, E. R., Bruck, I., O'Donnell, M. & Kuriyan, J. (2006). *BMC Struct. Biol.* **6**, 2.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Bubeck, P., Winkler, M. & Bautsch, W. (1993). *Nucleic Acids Res.* **21**, 3601–3602.
Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J., Yelent, B., Shen, B. & Tainer, J. A. (2004). *Cell*, **116**, 39–50.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dionne, I., Nookala, R. K., Jackson, S. P., Doherty, A. J. & Bell, S. D. (2003). *Mol. Cell*, **11**, 275–282.
Dore, A. S., Kilkenny, M. L., Jones, S. A., Oliver, A. W., Roe, S. M., Bell, S. D. & Pearl, L. H. (2006). *Nucleic Acids Res.* **34**, 4515–4526.
Freudenthal, B. D., Gakhar, L., Ramaswamy, S. & Washington, M. T. (2009). *Acta Cryst.* D**65**, 560–566.
Hlinkova, V., Xing, G., Bauer, J., Shin, Y. J., Dionne, I., Rajashankar, K. R., Bell, S. D. & Ling, H. (2008). *Acta Cryst.* D**64**, 941–949.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jones, D. H. & Howard, B. H. (1991). *Biotechniques*, **10**, 62–66.
Kong, X. P., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992). *Cell*, **69**, 425–437.
Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M. & Kuriyan, J. (1994). *Cell*, **79**, 1233–1243.
Maga, G. & Hubscher, U. (2003). *J. Cell Sci.* **116**, 3051–3060.
Marsic, D., Hughes, R. C., Byrne-Steele, M. L. & Ng, J. D. (2008). *BMC Biotechnol.* **8**, 44.
Matsumiya, S., Ishino, Y. & Morikawa, K. (2001). *Protein Sci.* **10**, 17–23.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mullis, K. B. & Faloona, F. A. (1987). *Methods Enzymol.* **155**, 335–350.
Nishida, H., Matsumiya, S., Tsuchiya, D., Ishino, Y. & Morikawa, K. (2006). *Acta Cryst.* F**62**, 253–256.
Oliner, J. D., Kinzler, K. W. & Vogelstein, B. (1993). *Nucleic Acids Res.* **21**, 5192–5197.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pikuta, E. V., Marsic, D., Itoh, T., Bej, A. K., Tang, J., Whitman, W. B., Ng, J. D., Garriott, O. K. & Hoover, R. B. (2007). *Int. J. Syst. Evol. Microbiol.* **57**, 1612–1618.
Warbrick, E. (2000). *Bioessays*, **22**, 997–1006.
Williams, G. J., Johnson, K., Rudolf, J., McMahon, S. A., Carter, L., Oke, M., Liu, H., Taylor, G. L., White, M. F. & Naismith, J. H. (2006). *Acta Cryst.* F**62**, 944–948.