Jesús S. Carrasco-Miranda,^a
Cesar S. Cardona-Felix,^b
Alonso A. Lopez-Zavala,^a
Enrique de-la-Re-Vega,^c Eugenio
De la Mora,^d Enrique Rudiño-
Piñera,^d Rogerio R. Sotelo-
Mundo^{a*} and Luis G. Brieba^{b*}

^aCentro de Investigación en Alimentación y Desarrollo, A.C. (CIAD), Carretera a Ejido La Victoria Km 0.6, Apartado Postal 1735, Hermosillo, Sonora 83304, Mexico, ^bLaboratorio Nacional de Genómica para la Biodiversidad (LANGEBO), Centro de Investigación y Estudios Avanzados-Unidad Irapuato, Km 9.6 Libramiento Norte Carretera Irapuato-León, Apartado Postal 629, Irapuato, Guanajuato 36500, Mexico, ^cDepartamento de Investigaciones Científicas y Tecnológicas (DICTUS), Universidad de Sonora, Blvd. Luis Encinas y Rosales S/N, Col. Centro, Hermosillo, Sonora 83000, Mexico, and ^dDepartamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología (IBT), Universidad Nacional Autónoma de México (UNAM), Av. Universidad 2001, Col. Chamilpa, Cuernavaca, Morelos 62210, Mexico

Correspondence e-mail: rrs@ciad.mx,
lgbrieba@ira.cinvestav.mx

Received 18 June 2012
Accepted 25 September 2012

© 2012 International Union of Crystallography
All rights reserved

Crystallization and X-ray diffraction studies of crustacean proliferating cell nuclear antigen

Proliferating cell nuclear antigen (PCNA), a member of the sliding clamp family of proteins, interacts specifically with DNA replication and repair proteins through a small peptide motif called the PCNA-interacting protein or PIP box. PCNA is recognized as one of the key proteins involved in DNA metabolism. In the present study, the recombinant PCNA from *Litopenaeus vannamei* (LvPCNA) was heterologously overexpressed and purified using metal ion-affinity chromatography. Crystals suitable for diffraction grew overnight using the hanging-drop vapour-diffusion method. LvPCNA crystals belong to space group C2 with unit-cell parameters $a = 144.6$, $b = 83.4$, $c = 74.3$ Å, $\beta = 117.6^\circ$. One data set was processed to 3 Å resolution, with an overall R_{meas} of 0.09 and a completeness of 93.3%. Initial phases were obtained by molecular replacement using a homology model of LvPCNA as the search model. Refinement and structural analysis are underway. This report is the first successful crystallographic analysis of a marine crustacean decapod shrimp (*L. vannamei*) proliferating cell nuclear antigen.

1. Introduction

Proliferating cell nuclear antigen (PCNA) acts as a scaffold protein in DNA replication, DNA repair, chromatin remodelling and cell-cycle progression. In most organisms PCNA is a homotrimer in which its three subunits adopt a ring-shaped structure; the characteristic ring structure of PCNA is highly conserved among protozoa, humans, yeast and plants (Chia *et al.*, 2010; Krishna *et al.*, 1994; Matsumiya *et al.*, 2001; Cardona-Felix *et al.*, 2011; Strzalka *et al.*, 2012). However, PCNA can assemble as a heterotrimer in the crenarchaeon *Sulfolobus solfataricus* and in *Arabidopsis* (Dionne *et al.*, 2006; Strzalka *et al.*, 2009), or as a non-canonical elliptic ring-like heterotetrameric structure in *Sulfolobus tokodaii* (Kawai *et al.*, 2011). The bacterial homologue of PCNA is dubbed the β -clamp and assembles as a homodimer in which each monomer contains three domains of identical topology (Kong *et al.*, 1992).

Each monomer of eukaryotic PCNA is formed by two domains joined by an inter-domain connecting loop (IDCL) (Krishna *et al.*, 1994), while the homotrimer forms a doughnut-shaped structure. The α -helices are situated towards the inner face of the ring, which contacts DNA. Therefore, the arrangement of the α -helices in each monomer leads to a pseudo-sixfold symmetry in a trimer composed of 12 α -helices (Ivanov *et al.*, 2006). The internal face of the ring has positively charged amino acids that allow interactions with the sugar-phosphate backbone of double-stranded DNA, with an inside diameter of approximately 34 Å sufficient to encircle a B-DNA and to slide freely on it (McNally *et al.*, 2010).

Most PCNA-interacting proteins bind PCNA via a short sequence motif called a PIP box, which has a consensus sequence QXX(M/L/I)XX(F/Y)(F/Y) (Moldovan *et al.*, 2007). Recently it has been reported that several proteins like human AlkB homologue 2, TFIIS-L, TFII-L, type topoisomerase II and RAD51B interact with a post-translationally modified PCNA via a novel PCNA-interacting motif (APIM) with an apparent consensus sequence MD(L/R)W(L/V/I)₂(K/R). The APIM sequence is present in about 200 nuclear proteins, identified by informatics analysis, and it has been suggested that it mediates PCNA binding with proteins involved in DNA repair

and cell-cycle control during genotoxic stress (Gilljam *et al.*, 2009). PCNA interacts with multiple protein partners mainly through its IDCL. Alternatively, PCNA binds to its interacting proteins through other regions, one of them being its C-terminus, and other regions of interaction with other proteins have been demonstrated as shown in the PCNA–DNA ligase and PCNA–PolB complexes from *Pyrococcus furiosus* (Vijayakumar *et al.*, 2006; Mayanagi *et al.*, 2009, 2011).

Additionally, it has been observed that PCNA can be used as an accessory replication factor of pathogens like simian virus 40 for its genome replication (Tsurimoto & Stillman, 1991). The archetypical role of PCNA is associated with the interaction with DNA polymerases that allows the latter to tightly bind DNA while allowing their lateral movement and thus increases their DNA synthesis processivity (Bravo *et al.*, 1987).

We have recently reported the cDNA sequence of the shrimp *Litopenaeus vannamei* PCNA (LvPCNA) (de-la-Re-Vega *et al.*, 2011), and others and ourselves have reported its gene expression in crustaceans under viral infection (de-la-Re-Vega *et al.*, 2011; Li, 2010; Wu *et al.*, 2008; Xie *et al.*, 2008). The white spot syndrome virus (WSSV) is a DNA virus that affects shrimp aquaculture around the world (Escobedo-Bonilla *et al.*, 2008). It is suggested that shrimp PCNA is necessary for viral replication, owing to an interaction with the WSSV DNA polymerase (de-la-Re-Vega *et al.*, 2011). However, specific interactions between LvPCNA and viral DNA polymerase remain to be corroborated. As a first step towards a structural understanding of the putative interaction between LvPCNA and viral DNA polymerase, herein we report the crystallization and diffraction of the first crustacean recombinant *L. vannamei* PCNA.

2. Materials and methods

2.1. LvPCNA over-expression and purification

Escherichia coli BL21-SI (Novagen), previously transformed with pG-KJE8 chaperone plasmid (Takara), was transformed with a synthetic construct containing the optimized open reading frame (ORF) of shrimp PCNA (GenBank: AEJ89927.1). An N-terminal tag containing nine histidine residues was added to the PCNA amino acid sequence to facilitate heterologous purification and a prescission protease recognition site (LEVLFQ/GP) was inserted between the histidine tag and the first methionine of LvPCNA to remove the histidine tag after purification. This construct has a theoretical molecular weight of approximately 31 kDa. A single colony was used to inoculate 100 ml of LB medium without NaCl, supplemented with 100 µg ml⁻¹ ampicillin and 20 µg ml⁻¹ chloramphenicol. This initial culture was grown overnight and was used to inoculate a 3 l culture medium. The culture was grown at 310 K until it reached an OD₆₀₀ of 0.6.

Heterologous protein expression (LvPCNA and chaperones) was induced with 0.4 mM IPTG, 5 ng ml⁻¹ tetracycline and 0.5 mg ml⁻¹ L-arabinose. The culture was incubated at 310 K for 4 h with shaking at 225 rev min⁻¹. The cell culture was centrifuged at 5000g and the pellet was dissolved in 50 ml of lysis buffer [20 mM Na₂HPO₄ pH 7.5, 0.5 M NaCl, 1 mM DTT, 0.5 mM phenylmethylsulfonyl fluoride (PMSF) and 0.1 mg ml⁻¹ lysozyme]. Cells were disrupted by sonication and centrifuged at 35 000g for 30 min at 277 K. The supernatant was filtered through a 0.45 µm pore-size filter and passed through a 5 ml HisTrap Fast Flow column (GE Healthcare) previously equilibrated with buffer A (20 mM Na₂HPO₄ pH 7.5, 0.5 M NaCl). The column was washed with ten volumes of buffer A and the elution of LvPCNA was carried out with a linear gradient of imidazole from 0 to 0.5 M in running buffer A. LvPCNA eluted at 130–200 mM of

Figure 1

Chromatogram of purified LvPCNA loaded onto a Superdex 200 column. The y axis corresponds to the absorbance and the x axis to the elution volume. In the left inset we observe the exclusion profile of the molecular-weight standards; thyroglobulin (670 kDa), γ -globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and vitamin B12 (1.35 kDa) in this column are observed in the inset. A regression analysis of the V_e/V_o for LvPCNA (1.378) in comparison to that of the molecular-weight standards predicts a molecular mass for LvPCNA of 110 kDa, indicating that the quaternary structure of LvPCNA corresponds to a trimer. The right inset shows the four eluted fractions of purified LvPCNA loaded onto 12% SDS-PAGE and Coomassie-blue stained.

imidazole in a final volume of 15 ml. The eluate was immediately dialysed in 2 l of dialysis buffer (20 mM Na₂HPO₄ pH 7.5, 200 mM NaCl, 5 mM DTT, 2 mM EDTA) at 277 K for 12 h. A second purification step was performed using a Superdex 200 HR10/30 gel filtration column (GE Healthcare) previously calibrated with molecular-weight standards thyroglobulin (670 kDa), γ -globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa) and vitamin B12 (1.35 kDa) in a buffer containing 20 mM Na₂HPO₄ pH 7.5, 200 mM NaCl. The LvPCNA was approximately 98% pure with an estimated molecular mass of 110 kDa that indicates its trimeric oligomeric state (Fig. 1).

2.2. Protein crystallization

LvPCNA was concentrated by ultrafiltration using a 10 kDa centrifugal filter (Millipore). Protein crystallization screening was performed using the hanging-drop vapour-diffusion method by mixing 1 µl of LvPCNA protein solution at a concentration of 10 mg ml⁻¹ in 20 mM Na₂HPO₄ pH 7.5, 200 mM NaCl and 2 mM EDTA with 1 µl of each of the 50 reservoir solutions from the Crystal Screen kit (Hampton Research). Thin crystals appeared overnight in a condition containing 200 mM CaCl₂, 100 mM sodium HEPES pH 7.5 and 28% (v/v) PEG 400 (solution 14 from Hampton Research Crystal Screen) (Fig. 2a). The condition was optimized varying each component (CaCl₂ 0.2–0.35 M, HEPES pH 7.2–7.7 and PEG 400 26–32%). The best crystals were obtained by seeding in a condition containing 300 mM CaCl₂, 100 mM sodium HEPES pH 7.5 and 30% (v/v) PEG 400. The seeding was performed utilizing a cat whisker to transfer microcrystals to a drop containing 1 µl of reservoir solution and 1 µl of 10 mg ml⁻¹ LvPCNA. Hexagonal-shaped crystals suitable for X-ray diffraction grew after 24 h, with dimensions around 0.1 × 0.6 mm (Fig. 2b). Crystals were loop-mounted and cryoprotected by quick soaking in 190 mM CaCl₂, 95 mM sodium HEPES pH 7.5 and 26.6% (v/v) PEG 400 and flash-cooled in liquid nitrogen.

Figure 2

LvPCNA crystals were obtained by the hanging-drop vapour-diffusion method. (a) A cluster of amorphous crystals appeared in a condition containing 0.2 M CaCl₂, 0.1 M sodium HEPES pH 7.5 and 28% (v/v) PEG 400. (b) *LvPCNA* crystals after seeding and refinement, with dimensions of 0.1 × 0.07 mm; the initial condition contained 0.3 M CaCl₂, 0.1 M sodium HEPES pH 7.5 and 30% (v/v) PEG 400. These crystals were used for X-ray data collection.

2.3. X-ray data collection and crystallographic analysis

Diffraction data sets were collected on beamline X6A and data collection was carried out on beamline X4C of the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory (BNL), USA, using a MAR CCD 165 detector. A data set consisting of 280 images was collected at 0.9795 Å. The crystal-to-detector distance was kept at 230 mm with a $\Delta\varphi = 0.5^\circ$ and an exposure time of 30 s per image at 100 K (Fig. 3). The data set was integrated and scaled using *XDS* and *XSCALE*, respectively (Kabsch, 2010).

The *LvPCNA* crystal belonged to the *C2* space group, with unit-cell parameters $a = 144.6$, $b = 83.4$, $c = 74.3$ Å, $\beta = 117.6^\circ$. A total of 43 794 reflections were integrated to a resolution of 3.0 Å and were merged to obtain 14 802 unique reflections with an overall R_{meas} of 0.094 and a completeness of 93.3%. Matthews coefficient calculations indicated that there was space for three PCNA molecules in the

Table 1

Crystallographic data and data-collection statistics of shrimp PCNA.

Values in parentheses are for the highest-resolution shell.

X-ray source	NSLS beamline X4C
Wavelength (Å)	0.9795
Space group	<i>C2</i>
Unit-cell parameters (Å)	$a = 144.57$, $b = 83.38$, $c = 74.31$
Resolution range (Å)	40.0–3.0 (3.1–3.0)
Total reflections	43794 (3463)
Unique reflections	14802 (1128)
$R_{\text{meas}}^\dagger$	0.094 (0.415)
Completeness (%)	93.3 (97.0)
$\langle I/\sigma(I) \rangle$	10.69 (3.57)
Multiplicity	2.96 (3.10)

$^\dagger R_{\text{meas}}$ is a redundancy-independent version of R_{sym} . $R_{\text{meas}} = \sum_{hkl} \{ [N(hkl)/[N(hkl) - 1]]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl) \}$.

asymmetric unit (AU) ($V_M = 2.36$ Å³ Da^{−1}), consistent with the experimental data of the homotrimeric molecule found in size-exclusion chromatography (Fig. 1). Thus, a data set comprising 140° was collected, with the data-collection statistics listed in Table 1.

Initial phases of *LvPCNA* were determined by molecular replacement in *Phaser* (McCoy *et al.*, 2007) using a homology model of *LvPCNA* based on the three-dimensional structure of human PCNA (PDB entry 1vym; Kontopidis *et al.*, 2005), as previously reported (de-la-Re-Vega *et al.*, 2011). The best solution from *Phaser* presented a log-likelihood gain of 2849, an RFZ (rotation function Z score) of 9.7 and a TFZ (translation function Z score) of 41.6. The Z score is calculated from the loglikelihood gain observed for the solutions calculated in the experiment compared to the values obtained from random rotations (RFZ) or random translations (TFZ). The larger the values of the Z score, the more likely that a correct solution has been obtained in the molecular replacement. An analysis of the self-rotation function performed with data integrated in space group *C2* showed a significant peak at $\chi = 120^\circ$ (data not shown), consistent with a threefold non-crystallographic symmetry (NCS), as previously reported (Ivanov *et al.*, 2006), and also corroborated in this work, when a clear threefold NCS between the copies comprising the AU is found. Additionally, evidence of a trimeric state was found after the size-exclusion chromatography experiments were analysed (Fig. 1). That evidence and the calculation of a reasonable Matthews coefficient were consistent with the first electron-density map and the coordinates from the solution obtained in *Phaser*. Overall, most of the main-chain atoms were properly fitted after a visual inspection. The R_{work} and R_{free} after molecular replacement were 0.25 and 0.31,

Figure 3

X-ray diffraction pattern recorded at NSLS X4C from an *LvPCNA* crystal (0.5° oscillation) with visible diffraction to 3.0 Å resolution.

respectively. *Lv*PCNA refinement and rebuilding are underway using the programs *Phenix* (Adams *et al.*, 2010) and *Coot* (Emsley *et al.*, 2010), respectively.

JSC-M, CSC-F and AAL-Z were supported by PhD fellowships from CONACyT (Mexico's National Science and Research Council). RRS-M acknowledges financial support from CONACyT (grant Nos. CB-2009-131859 and E0007-2011-01-179940) and the Texas A&M-CONACyT Collaborative Grant Program (grant No. 2011-050). LGB acknowledges financial support from CONACyT (grant No. CB-2009-128647). We thank Dr Vivian Stojanoff, Dr John Schwanoff and the staff at BNL NSLS beamlines X4C and X6A for the data-collection facilities. Beamline X6A is funded by NIGMS (GM-0080) and the US Department of Energy (No. DE-AC02-98CH10886). The authors thank Felipe Isac-Martinez, Luis Leyva-Durán, Adalberto Murrieta-Valenciana, José L. Aguilar-Valenzuela and Martin Peralta-Contreras for their computational technical support.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Bravo, R., Frank, R., Blundell, P. A. & Macdonald-Bravo, H. (1987). *Nature (London)*, **326**, 515–517.
- Cardona-Felix, C. S., Lara-Gonzalez, S. & Briebe, L. G. (2011). *Acta Cryst.* **D67**, 497–505.
- Chia, N., Cann, I. & Olsen, G. J. (2010). *PLoS One*, **5**, e10866.
- de-la-Re-Vega, E., Muhlia-Almazan, A., Arvizu-Flores, A. A., Islas-Osuna, M. A., Yepiz-Plascencia, G., Briebe, L. G. & Sotelo-Mundo, R. R. (2011). *Results Immunol.* **1**, 24–30.
- Dionne, I., Nookala, R. K., Jackson, S. P., Doherty, A. J. & Bell, S. D. (2006). *Mol. Cell*, **11**, 275–282.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Escobedo-Bonilla, C. M., Alday-Sanz, V., Wille, M., Sorgeloos, P., Pensaert, M. B. & Nauwynck, H. J. (2008). *J. Fish Dis.* **31**, 1–18.
- Gilljam, K. M., Feyzi, E., Aas, P. A., Sousa, M. M., Müller, R., Vågbo, C. B., Catterall, T. C., Liabakk, N. B., Slupphaug, G., Drabløs, F., Krokan, H. E. & Otterlei, M. (2009). *J. Cell Biol.* **186**, 645–654.
- Ivanov, I., Chapados, B. R., McCammon, J. A. & Tainer, J. A. (2006). *Nucleic Acids Res.* **34**, 6023–6033.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kawai, A., Hashimoto, H., Higuchi, S., Tsunoda, M., Sato, M., Nakamura, K. T. & Miyamoto, S. (2011). *J. Struct. Biol.* **174**, 443–450.
- Kong, X. P., Onrust, R., O'Donnell, M. & Kuriyan, J. (1992). *Cell*, **69**, 425–437.
- Kontopidis, G., Wu, S. Y., Zheleva, D. I., Taylor, P., McInnes, C., Lane, D. P., Fischer, P. M. & Walkinshaw, M. D. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 1871–1876.
- Krishna, T. S. R., Kong, X. P., Gary, S., Burgers, P. M. & Kuriyan, J. (1994). *Cell*, **79**, 1233–1243.
- Li, P., Zha, J., Kong, Y., Chen, C., Sun, H., Song, D. & Zhou, K. (2010). *Comp. Biochem. Physiol. A*, **157**, 170–176.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- McNally, R., Bowman, G. D., Goedken, E. R., O'Donnell, M. & Kuriyan, J. (2010). *BMC Struct. Biol.* **10**, 3.
- Matsumiya, S., Ishino, Y. & Morikawa, K. (2001). *Protein Sci.* **10**, 17–23.
- Mayanagi, K., Kiyonari, S., Nishida, H., Saito, M., Kohda, D., Ishino, Y., Shirai, T. & Morikawa, K. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 1845–1849.
- Mayanagi, K., Kiyonari, S., Saito, M., Shirai, T., Ishino, Y. & Morikawa, K. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 4647–4652.
- Moldovan, G. L., Pfander, B. & Jentsch, S. (2007). *Cell*, **129**, 665–679.
- Strzalka, W., Labecki, P., Bartnicki, F., Aggarwal, C., Rapala-Kozik, M., Tani, C., Tanaka, K. & Gabrys, H. (2012). *J. Exp. Bot.* **63**, 2971–2983.
- Strzalka, W., Oyama, T., Tori, K. & Morikawa, K. (2009). *Protein Sci.* **18**, 1072–1080.
- Tsurimoto, T. & Stillman, B. (1991). *J. Biol. Chem.* **266**, 1950–1960.
- Vijayakumar, S., Chapados, B. R., Schmidt, K. H., Kolodner, R. D., Tainer, J. A. & Tomkinson, A. E. (2006). *Nucleic Acids Res.* **35**, 1624–1637.
- Wu, C., Söderhäll, I., Kim, Y. A., Liu, H. & Söderhäll, K. (2008). *Proteomics*, **8**, 4226–4235.
- Xie, Y., Wang, B., Li, F., Jiang, H. & Xiang, J. (2008). *Comput. Biochem. Physiol. B*, **151**, 225–229.