Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Production, crystallization and preliminary X-ray crystallographic studies of the bacteriophage φ12 packaging motor

Erika J. Mancini,[a] Denis E. Kainov,[b] Hui Wei,[c] Paul Gottlieb,[c] Roman Tuma,[b] Dennis H. Bamford,[b] David I. Stuart[a] and Jonathan M. Grimes[a]*

[a]Division of Structural Biology, The Henry Wellcome Building for Genomic Medicine, Oxford University, Roosevelt Drive, Oxford OX3 7BN, England, [b]Institute of Biotechnology and Department of Biosciences, Vikki Biocenter, PO Box 56 (Viikinkaari 5), 00014 University of Helsinki, Finland, and [c]Department of Microbiology and Immunology, The Sophie Davis School of Biomedical Education, The City College of New York, NY 10031, USA

Correspondence e-mail: jonathan@strubi.ox.ac.uk

Received 21 November 2003
Accepted 14 January 2004

The hexameric ATPase P4 from bacteriophage φ12 is responsible for packaging single-stranded genomic precursors into the viral procapsid. P4 was overexpressed in *Escherichia coli* and purified. Crystals of native and selenomethionine-derivatized P4 have been obtained that belong to space group *I*222, with half a hexamer in the asymmetric unit and unit-cell parameters a = 105.0, b = 130.5, c = 158.9 Å. A second crystal form of different morphology can occur in the same crystallization drop. The second form belongs to space group *P*1, with four hexamers in the asymmetric unit and unit-cell parameters $a = 114.9, b = 125.6, c = 153.9$ Å, $\alpha = 90.1, \beta = 91.6, \gamma = 90.4^\circ$. Synchrotron X-ray diffraction data have been collected for the *I*222 and *P*1 crystal forms to 2.0 and 2.5 Å resolution, respectively.

1. Introduction

Motor proteins convert the energy from nucleoside triphosphate (NTP) hydrolysis to mechanical work. The mechanism of energy conversion in such motors is the subject of intense study. NTP-dependent translocation is utilized in many tailed DNA viruses for genome encapsidation or 'packaging' (Catalano *et al.*, 1995; Simpson *et al.*, 2000; Baumann & Black, 2003; Moore & Prevelige, 2002). Emerging structural and biochemical evidence suggests that these translocation motors are specialized macromolecular machines composed of several connector proteins and one or more NTPases. This complexity has made it difficult to verify packaging models. In contrast, the double-stranded RNA (dsRNA) bacteriophage φ6 from the *Cystoviridae* family packages single-stranded genomic precursors into procapsid particles using a single protein species, P4, a hexameric NTPase with a subunit molecular weight of 32 kDa (Juuti *et al.*, 1998). Cryo-electron microscopy revealed the P4 hexamers to be located at the 12 fivefold positions on the particle surface (de Haas *et al.*, 1999). The production of stable P4-deficient particles allowed the role of P4 in the φ6 life cycle to be examined directly (Paatero *et al.*, 1998). Particles devoid of P4 showed no detectable packaging activity. In contrast, particles containing reduced amounts of P4 (statistically one fivefold position occupied per particle) packaged the ssRNA segments and replicated them to a double-stranded form but were not able to carry out the plus-strand synthesis reaction (Pirttimaa *et al.*, 2002). Given these results, a model has been proposed where only one out of 12 PC-associated hexamers operates as the packaging portal, while the others are required for efficient transcription (Pirttimaa *et al.*, 2002; Poranen *et al.*, 2004).

Both structural and functional comparisons suggest that cystoviral P4s are related to a number of translocation enzymes, such as hexameric helicases, that couple NTP hydrolysis with unidirectional DNA translocation (Kainov *et al.*, 2003). The structure of the hexameric helicase gp4 from bacteriophage T7 suggested a mechanism for the coordination of NTP binding and hydrolysis around the helicase ring and how this could be coupled to DNA translocation (Singleton *et al.*, 2000). However, this is merely a working hypothesis based on a conformation of the protein that appears to be caught at one particular stage in the catalytic cycle and evidence from enzyme kinetics and low-resolution techniques suggests that other conformations exist (Hingorani *et al.*, 1997; Kim *et al.*, 2002).

In order to understand the mechanism of ssRNA translocation into the procapsids of cystoviruses, we have initiated structural studies of P4 proteins. However, crystallographic analysis of P4 proteins from bacteriophages φ6, φ8 and φ13 has been fraught with problems (Mancini *et al.*, 2003). We report here the cloning and purification of P4 from bacteriophage φ12, a recently sequenced member of the *Cystoviridae* family (Gottlieb *et al.*, 2002), and demonstrate that this protein yields crystals suitable for high-resolution structure analysis.

2. Construction of the plasmid

To construct a plasmid for φ12 P4 protein expression, the nucleotide P4 gene was PCR-amplified from the pP12L1 (Gottlieb *et al.*, 2002) template using the Expand High Fidelity

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

PCR System (Roche Molecular Biochemicals, Indianapolis, IN, USA), with the oligonucleotides 5′-GACGTCCTA**CATATGAT**CCATCTGTACGACGC-3′ and 5′-TACA**GGATCC**TTAGTTGGAGGTGACGAA-3′ as upstream and downstream primers, respectively. The PCR fragment was then digested with *Nde*I and *Bam*HI (underlined sites in the primer sequences), gel-purified using a low-melting-point agarose and ligated into pET-32a (Novagene) vector. The pET-32a vector possesses *Nde*I sites at positions 346 and 691; complete *Nde*I digestion removes the *trx* A region. The vector was therefore cut with *Nde*I and *Bam*HI and the large vector fragment isolated from low-melting-point agarose after electrophoresis. The ligation products were used for transformation of *Escherichia coli* XL1-Blue supercompetent cells {*recA*1, *end*1, *gyrA*96, *thi*1, *hsdR*17, *supE*44, *relA*1, [F′, *proAB*, *lacI*^{q}, *ZDM15*, Tn*10* (Tet^r)]} (Stratagene, La Jolla, CA, USA). The resulting recombinant plasmid pPG27 was used for protein expression in *E. coli* strain B834 (DE3; Novagen).

3. Protein expression and purification

To achieve expression of soluble φ12 P4, starter cultures of B834 (DE3/ pPG27) in LB medium containing 150 µg ml^{-1} ampicillin were grown at 310 K with shaking until the absorbance at 540 nm (A_{540}) reached 0.5, diluted 50-fold into 2 l of the same medium and grown at 310 K to an A_{540} of 1.0. The cultures were chilled on ice and induced with 1 m*M* isopropyl β-D-thiogalactopyranoside (IPTG). Induced cells were transferred to 290 K and shaken for 14 h. All the following steps were performed at 277 K. Bacteria were collected by centrifugation and resuspended in 30 ml TNM buffer (20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl, 7.5 m*M* $MgCl_2$). The suspension was passed at ∼105 MPa through a precooled French pressure cell. The lysate was centrifuged at 120 000*g* for 1 h. Supernatant fractions were loaded onto a HiTrap heparin agarose HP column (Amersham Biosciences). The P4-containing flowthrough fraction from the heparin column was applied onto an anion-exchange column (HiTrap Q-Sepharose HP; Amersham Biosciences). Proteins were eluted with a linear 0.1–1 *M* NaCl gradient buffered with 20 m*M* Tris–HCl pH 8.0 containing 7.5 m*M* $MgCl_2$. Fractions containing P4 were pooled, filtered, concentrated and injected onto a gel-filtration column (Superdex-200 26/60, Amersham Biosciences). Purified protein was stored in 20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl and 7.5 m*M* $MgCl_2$ at 277 K.

P4 contains five methionines in 331 residues, which makes the production of selenomethionyl protein a useful tool for phasing. To obtain the selenomethionyl protein, cells were grown in LB media with 150 µg ml^{-1} ampicillin at 310 K to an A_{540} of 0.6, collected by centrifugation and resuspended in the same volume of MOPS minimal medium containing 50 µg ml^{-1} L-selenomethionine (Nanduri *et al.*, 2002). The cultures were induced by adding 1 m*M* IPTG and incubated for 14 h at 290 K. Protein purification was performed as for the native protein.

4. Crystallization

The purified protein was concentrated to 10 mg ml^{-1} by centrifugation using a Centricon 10 microconcentrator (Amicon). Crystals were obtained by sitting-drop vapour diffusion at 293 K. 1 µl protein was mixed with an equal volume of precipitant solution composed of 10% PEG 1500 in 100 m*M* sodium acetate pH 4.8. Crystals of two different morphologies, rhombs and thin plates, appeared in the same drop within hours of setting up and grew within 2 d to approximate dimensions of 0.3 × 0.3 × 0.2 and 0.4 × 0.4 × 0.05 mm, respectively. Addition of 30 m*M* ADP to the precipitant solution pushed the crystallization equilibrium towards the rhombic morphology. Crystals of selenomethionyl P4 of the rhombic morphology were obtained under the same conditions as the native crystals, but the sitting drop contained 2 µl protein, 0.1 µl 20 m*M* dithiothreitol (DTT) and 2 µl well solution.

Table 1
Native data-set statistics for different crystal forms.

Values in parentheses refer to the highest resolution shells.

X-ray source	ESRF ID-14 EH2	ESRF ID-14 EH2
Wavelength (Å)	0.9326	0.9326
Space group	*I*222	*P*1
Unit-cell parameters		
a	105.0	114.9
b (Å)	130.5	125.6
c (Å)	158.9	153.9
α (°)	90.0	90.1
β (°)	90.0	91.6
γ (°)	90.0	90.4
Resolution range (Å)	20.0–2.0	20.0–2.5
Observations	2815332	2075996
Unique observations	68669	291757
$I/\sigma(I)$	31.2 (2.4)	18.7 (3.9)
$R_{\rm merge}$† (%)	9.4 (87.0)	6.6 (31.8)
No. of hexamers in AU	0.5	4

† $R_{\rm merge} = \sum|I - \langle I\rangle|/\sum\langle I\rangle$.

5. X-ray analysis

Diffraction data from native P4 crystals were collected at beamline ID-14 EH2, ESRF, Grenoble (wavelength 0.9326 Å) and measured using an ADSC Q4 CCD detector. Data were processed and scaled using the *HKL*2000 suite of programs (Otwinowski & Minor, 1997). Details of the statistics of the native data for the two different crystal forms are summarized in Table 1.

The rhombic morphology crystals were cryoprotected by transferring them with a loop into a well containing reservoir solution made up to 20% trehalose and 15% sucrose

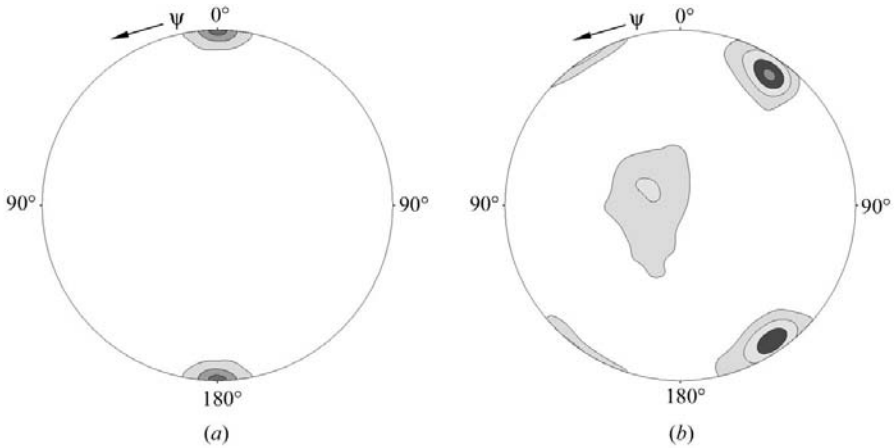

Figure 1
$\kappa = 60°$ sections of self-rotation functions calculated for both crystal forms. Rotation functions were calculated in *CNS* and the angles are defined according to the convention of the program (Brünger *et al.*, 1998). (*a*) The $\kappa = 60°$ section for the *I*222 crystal form, calculated using data in the resolution range 20–2 Å. The sixfold peak is 42% of the origin peak and is aligned with one of the crystallographic twofold axes. (*b*) The $\kappa = 60°$ section for the *P*1 crystal form, calculated using data in the resolution range 20–2.5 Å, showing the location of the two sixfold peaks (height 27% of the origin peak). Figure drawn with *GROPAT* (R. Esnouf & D. I. Stuart, unpublished program).

for a few seconds prior to freezing in a nitrogen-gas stream. This cryoprotectant is in routine use in Oxford and for these crystals yielded higher resolution diffraction than was observed with other cryoprotectants, such as glycerol. Analysis of the diffraction data revealed that the crystals belonged to space group *I*222, with unit-cell parameters $a = 105.0$, $b = 130.5$, $c = 158.9$ Å, $\alpha = \beta = \gamma = 90^\circ$. These unit-cell parameters were consistent with the presence of three subunits of P4 in the crystallographic asymmetric unit, assuming a solvent content of ∼50%. The data are reliable to 2.0 Å resolution [$I/\sigma(I) = 2.4$ at that resolution], despite the rather poor $R_{\rm merge}$ values, which reflects the very high redundancy of the data (∼40). A self-rotation function carried out using *CNS* (Brünger *et al.*, 1998) produced a strong (6.6σ) peak at $\psi = 0$, $\varphi = 0$, $\kappa = 60^\circ$, consistent with a molecular sixfold axis aligned with a crystallographic twofold axis, suggesting half a P4 hexameric ring in the asymmetric unit, related to the other half by crystallographic twofold symmetry (Fig. 1*a*). The thin-plate morphology crystals were cryoprotected by transfer into a drop containing mother liquor supplemented with 30% glycerol prior to freezing. This second crystal form belongs to space group *P*1, with unit-cell parameters $a = 114.9$, $b = 125.6$, $c = 153.9$ Å, $\alpha = 90.1$, $\beta = 91.6$, $\gamma = 90.4^\circ$, and can accommodate 24 subunits of P4 in the unit cell with a solvent content of 54%. Analysis of the self-rotation function revealed two strong peaks (4.4σ) at $\psi = 36$, $\varphi = 164$, $\kappa = 60^\circ$ and $\psi = 145$, $\varphi = 169$, $\kappa = 60^\circ$, consistent with the presence of four P4 hexameric rings arranged in pairs (possibly dodecamers; Mancini *et al.*, 2003) with coincident sixfold axes (Fig. 1*b*). Despite the similarity in unit-cell parameters, the molecular sixfold axes are disposed differently in the two crystal forms, presumably reflecting quite different packing arrangements. MAD data collection was carried out at beamline BM14 using a MAR CCD detector. Data were measured at three wavelengths from a single selenomethionyl crystal of the *I*222 form to a resolution of better than 2.5 Å. Structure determination is in progress and statistics of the MAD data collection will be presented subsequently.

The authors are grateful to Geoff Sutton, Martin Walsh at the MRC UK beamline BM14 and the staff of ESRF at beamline ID-14 EH2 for assistance with X-ray data collection. This work was supported by the Human Frontier Science Programme, the Academy of Finland ['Finnish Centre of Excellence Program 2000–2005', grants 172621 and 1206926 (to RT) and 1202855 and 1202108 (to DHB)] and the National Science Foundation Career Award (MCB9984310 to PG). EJM was the recipient of an EMBO long-term fellowship (ALTF 192-2000), DEK is a fellow of the National Graduate School in Informational and Structural Biology, JMG is supported by the Royal Society and DIS is supported by the UK Medical Research Council.

References

Baumann, R. G. & Black, L. W. (2003). *J. Biol. Chem.* **278**, 4618–4627.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Catalano, C. E., Cue, D. & Feiss, M. (1995). *Mol. Microbiol.* **16**, 1075–1086.

Gottlieb, P., Potgieter, C., Wei, H. & Toporovsky, I. (2002). *Virology*, **295**, 266–271.

Haas, F. de, Paatero, A. O., Mindich, L., Bamford, D. H. & Fuller, S. D. (1999). *J. Mol. Biol.* **294**, 357–372.

Hingorani, M. M., Washington, M. T., Moore, K. C. & Patel, S. S. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 5012–5017.

Juuti, J. T., Bamford, D. H., Tuma, R. & Thomas, G. J. Jr (1998). *J. Mol. Biol.* **279**, 347–359.

Kainov, D. E., Pirttimaa, M. J., Tuma, R., Butcher, S. J., Thomas, G. J. Jr, Bamford, D. H. & Makeyev, E. V. (2003). *J. Biol. Chem.* **278**, 48084–48091.

Kim, D. E., Narayan, M. & Patel, S. S. (2002). *J. Mol. Biol.* **321**, 807–819.

Mancini, E. J., Grimes, J. M., Malby, R., Sutton, G. C., Kainov, D. E., Juuti, J. T., Makeyev, E. V., Tuma, R., Bamford, D. H. & Stuart, D. I. (2003). *Acta Cryst.* D**59**, 2337–2341.

Moore, S. D. & Prevelige, P. E. Jr (2002). *J. Mol. Biol.* **315**, 975–994.

Nanduri, B., Byrd, A. K., Eoff, R. L., Tackett, A. J. & Raney, K. D. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 14722–14727.

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

Paatero, A. O., Mindich, L. & Bamford, D. H. (1998). *J. Virol.* **72**, 10058–10065.

Pirttimaa, M. J., Paatero, A. O., Frilander, M. J. & Bamford, D. H. (2002). *J. Virol.* **76**, 10122–10127.

Poranen, M. M., Pirttimaa, M. J. & Bamford, D. H. (2004). In *Viral Genome Packaging*, edited by C. Catalano. Georgetown, TX USA: Landes Bioscience. In the press.

Simpson, A. A., Tao, Y., Leiman, P. G., Badasso, M. O., He, Y., Jardine, P. J., Olson, N. H., Morais, M. C., Grimes, S., Anderson, D. L., Baker, T. S. & Rossmann, M. G. (2000). *Nature* (*London*), **408**, 745–750.

Singleton, M. R., Sawaya, M. R., Ellenberger, T. & Wigley, D. B. (2000). *Cell*, **101**, 589–600.