Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic studies on the bacteriophage Φ6 RNA-dependent RNA polymerase

Sarah J. Butcher,[a] Eugeny V. Makeyev,[a] Jonathan M. Grimes,[b] David I. Stuart[b,c]* and Dennis H. Bamford[a]

[a]Department of Biosciences and Institute of Biotechnology, University of Helsinki, PL56 (Viikinkaari 5), FIN-00014 University of Helsinki, Finland, [b]Division of Structural Biology, The Henry Wellcome Building for Genomic Medicine, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, England, and [c]Oxford Centre for Molecular Sciences, New Chemistry Laboratory, South Parks Road, Oxford OX1 3QT, England

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

Received 8 April 2000
Accepted 28 July 2000

The RNA-dependent RNA polymerase (P2) from bacteriophage Φ6 has been cloned and the protein overexpressed in *Escherichia coli* to produce an active enzyme. A fully substituted selenomethionyl version of the protein has also been produced. Crystals of both proteins have been grown; most belong to the monoclinic space group $P2_1$, with unit-cell parameters a = 105.9, b = 94.0, c = 140.9 Å, $\beta = 101.4^\circ$, but some are trigonal (space group $P3_1$ or $P3_2$), with unit-cell parameters $a = b = 110.1$, $c = 159.4$ Å, $\gamma = 120^\circ$. Both crystal forms occur in the same crystallization drop and are morphologically indistinguishable. Native data sets have been collected from both types of crystals to better than 3 Å resolution.

1. Introduction

Double-stranded RNA (dsRNA) viruses infect a range of eukaryotic and prokaryotic hosts, causing a number of economically important infectious diseases (Fields *et al.*, 1996). A key component for the replication of most of these viruses is the viral polymerase, which carries out two essential reactions: firstly, upon entering the cell, it functions as a transcriptase, producing positive-sense messenger RNA (+ssRNA) from the virus genomic dsRNA. These +ssRNAs are also assembled into new polymerase complexes where, in the second reaction, they act as templates for the synthesis of complementary minus-strand RNA, resulting in complete dsRNA genomes. Each polymerase complex is made up of several proteins and the dsRNA genome, forming an icosahedrally symmetric particle or core that can subsequently mature into an infectious virion. For bacteriophage Φ6, only one of the polymerase complex proteins, P2, contains the characteristic polymerase sequence motifs (Bruenn, 1991, 1993; Koonin *et al.*, 1989), although standard sequence alignments cannot detect any sequence similarity. Although an atomic model of one dsRNA virus core has been determined (Grimes *et al.*, 1998), so far no atomic resolution structures are available for the actual polymerase subunit.

To understand the relationship of these RNA-dependent RNA polymerases to other polymerases and to facilitate structure-based design of anti-dsRNA-virus compounds, we have chosen to study the three-dimensional structure of bacteriophage Φ6 P2.

2. Protein expression and purification

P2 (M_r = 75 000) is one of the four protein components of the Φ6 polymerase complex. Enzymatically active P2 was isolated and purified from *E. coli* BL21(DE3/pEM2) as described previously (Makeyev & Bamford, 2000), except that the incubation temperature during induction was further optimized to 293 K, increasing the yield of soluble protein. The purified protein was concentrated to 18 mg ml^{-1} by ultrafiltration using a Centricon-10 microconcentrator (Amicon Inc). P2 contains 25 methionines (Makeyev & Bamford, 2000; Mindich *et al.*, 1988), so a selenomethionyl protein was also produced and purified for a multiwavelength anomalous diffraction experiment. The vector coding for P2 (pEM2) was used to transform the methionine auxotroph *E. coli* B834 (DE3). Cells were grown to a density of 10^9 cells ml^{-1} in 2 l LB medium, 150 µg ml^{-1} ampicillin at 310 K, pelleted, washed with 10 m*M* potassium phosphate pH 7.2, 140 m*M* NaCl and resuspended in 2 l M9 minimal medium supplemented with 150 µg ml^{-1} ampicillin, 3.35 µg ml^{-1} thiamine chloride and 50 µg ml^{-1} seleno-L-methionine (Sigma). To induce expression, IPTG was added to 1 m*M* and the culture was incubated at 296 K for 15 h. The cells were harvested and P2 was purified as for the native protein (Makeyev & Bamford, 2000). The purified protein was concentrated to 7.7 mg ml^{-1} and stored in 10 m*M* Tris pH 8, 100 m*M* NaCl, 0.02% NaN_3 at 277 K. This protein had replicase activity when assayed according to Makeyev & Bamford (2000) (data not shown).

3. Mass determination

Mass spectrometry was performed on both the native and selenomethionyl protein. Concentrated P2 protein was desalted by reverse-phase HPLC in a 0.21 × 10 cm TSK TMS 250 (C1, TosoHaas Corp., Japan) column equilibrated with 0.1% trifluoroacetic acid. Elution was performed with a linear gradient of acetonitrile (3–60% in 60 min) at a flow rate of 0.2 ml min^{-1} and was monitored at 214 nm, giving a single peak in each case. Electrospray mass spectrometry was performed on a Q-TOF mass spectrometer (Micromass Ltd, UK). The mass determined for the native protein was 74 809 Da, exactly the expected value (Fig. 1*a*). The selenomethionyl P2 yielded 75 982 Da compared with a theoretical value of 75 981 Da for P2 containing 25 selenomethionine residues (Fig. 1*b*), indicating that the protein was fully labelled.

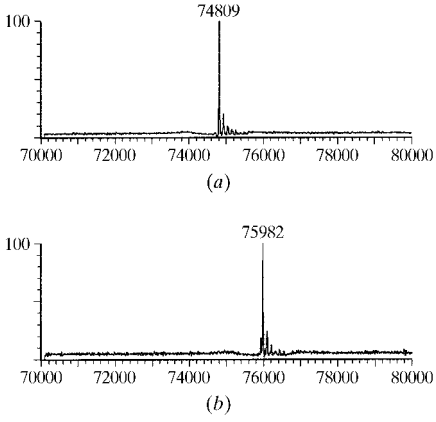

Figure 1
Electrospray ionization mass spectra of (*a*) native, (*b*) selenomethionyl P2.

4. Crystallization

Crystallization conditions were screened using the sitting-drop vapour-diffusion technique, a native protein concentration of 5 mg ml^{-1} and Hampton Crystal Screens I and II (Jancarik & Kim, 1991) at 277 and 293 K. Crystals were found in four conditions at 293 K and five conditions at 277 K, over the pH range 5.6–9 and in a number of buffers. The precipitant was 10–20% PEG with MW ranging from 4000 to 20 000. The largest crystals (with variable morphology, often resembling truncated rhombs with a wedge-shaped cross-section and maximum dimensions of approximately 0.4 × 0.3 × 0.4 mm) were obtained by sitting-drop vapour diffusion at 293 K with a 1:1 ratio of protein (5 mg ml^{-1}) to well solution. The well solution consisted of 10% PEG 8000, 0.1 *M* MES pH 6.7. Crystals started to appear within 30 min with freshly prepared protein, but took up to two weeks to appear as the protein aged. For data collection at cryogenic temperatures, individual crystals were transferred in a loop to a drop containing 10 µl 100% glycerol mixed with 20 µl of well reservoir for a few seconds prior to cooling in a stream of liquid nitrogen.

Crystals of selenomethionyl P2 were obtained at 293 K but only with PEG 20 000 at pH 7.3. The well solution contained between 10 and 15% PEG 20 000, 8 m*M* $MgCl_2$, 4.2% ethylene glycol, 100 m*M* HEPES pH 7.3. The sitting drop contained 2 µl protein, 0.2 µl 20 m*M* DTT and 2 µl well solution.

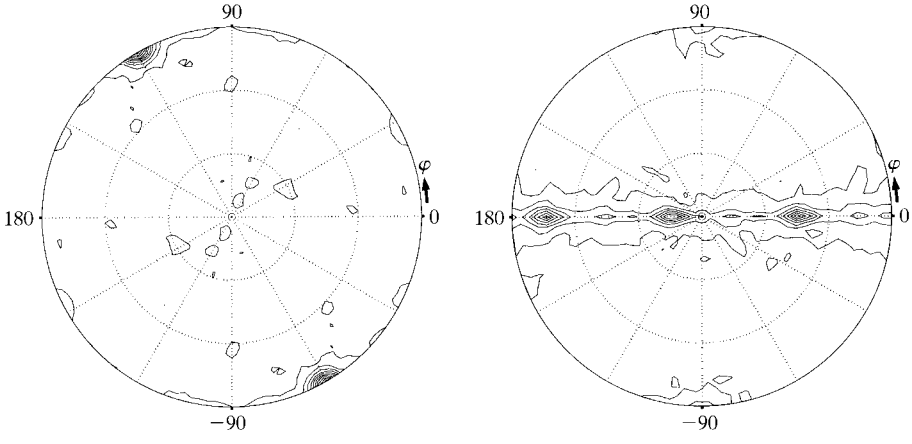

Figure 2
Sections of self-rotation functions calculated for both crystal forms. Rotation functions were calculated in *X-PLOR* (Brünger, 1992) and the angles are defined according to the convention of that program. φ is marked (note that the crystallographic *a* axis lies at $\varphi = 0°$) and ψ runs from zero (in the centre of the plot) outwards to 90°. (*a*) The $\kappa = 120°$ section for the $P2_1$ crystal form, calculated using data between 20 and 3 Å. The peak is 42% of the origin peak. The crystallographic twofold axis is perpendicular to the plane of the section. (*b*) The $\kappa = 180°$ section for the trigonal crystal form, calculated using data between 20 and 4 Å, showing the non-crystallographic peaks (height 27% of the origin peak). The crystallographic threefold axis is at $\varphi = 90$, $\psi = 90°$, *i.e.* in the plane of the paper. Figure drawn with *GROPAT* (R. Esnouf, E. Y. Jones & D. I. Stuart, unpublished program).

5. X-ray analysis

Two crystal forms have been characterized. The primary crystal form belongs to space group $P2_1$, with unit-cell parameters $a = 105.9$, $b = 94.0$, $c = 140.9$ Å, $\beta = 101.4°$. A data set has been collected from a cryo-cooled native crystal to minimum Bragg spacings of 2.7 Å at beamline ID-14 EH3, ESRF, Grenoble, wavelength 0.93 Å, using a MAR CCD detector. Data were processed and scaled using the *HKL*2000 suite of programs (Otwinowski & Minor, 1997). This data set is 99.5% complete, containing 72 535 unique reflections derived from 232 155 observations. $R_{\rm merge}$ was 7.1% and the overall value for $I/\sigma(I)$ is 13 [for the 2.8–2.7 Å shell, the data are 100% complete and $I/\sigma(I)$ is 1.3]. A self-rotation function calculated using *X-PLOR* (Brünger, 1992) showed a clear single peak (2.5σ in height; origin peak 5.9σ) on the $\kappa = 120°$ section, lying in the plane orthogonal to the crystallographic twofold axis, suggesting three copies in the asymmetric unit related by either a threefold rotation or threefold screw axis (calculated solvent content 59%) (Fig. 2*a*).

The second crystal form belongs to the trigonal space group $P3_1$ or $P3_2$, with unit-cell parameters $a = b = 110.1$, $c = 159.4$ Å, $\gamma = 120°$. One crystal of this space group from the selenomethionine-derivatized protein has been analysed and was used to collect a four-wavelength MAD data set to a resolution of better than 2.7 Å at the SBC ID19 beamline at the Advanced Photon Source, Argonne, USA. As an example, for one wavelength the data set (scaled to 2.7 Å resolution) is 99.4% complete and contains 58 954 unique reflections derived from 158 837 observations. $R_{\rm merge}$ was 7.5% and the overall value for $I/\sigma(I)$ was 13 [for the 2.8–2.7 Å shell, the data are 96% complete and $I/\sigma(I)$ is 1.7]. A self-rotation function revealed peaks on the $\kappa = 180°$ section orthogonal to the crystallographic threefold axis (Fig. 2*b*). This is consistent with two molecules in the asymmetric unit related to each other by a twofold rotation or screw axis orthogonal to the crystallographic threefold axis (calculated solvent content 66.7%). Structure determination is in progress.

We thank Sergey Korolev for help with data collection at the APS, Dr Javier Caldentey for helpful discussions, Jari Helin for mass-spectrometry analysis, and Jonathan Diprose and Geoff Sutton for help and support on data-collection trips to the APS and the ESRF. SJB thanks the organisers of the Advanced Training Course on the Expression, Purification and Crystallization of Proteins held at the EMBL Hamburg outstation in 1998 for advice. SJB is a Marie Curie Fellow and EVM is a fellow of the Centre for International Mobility. The work was supported by the Finnish Academy of Sciences (grants 16164298, 168694 to DB) and the European Union (grant BIO4-CT97-2364 to DB and DIS). JMG is funded by the Royal Society and DIS by the MRC.

References

Bruenn, J. A. (1991). *Nucleic Acids Res.* **19**, 217–226.
Bruenn, J. A. (1993). *Nucleic Acids Res.* **21**, 5667–5669.
Brünger, A. T. (1992). *X-PLOR Version* 3.1. *A Program for Crystallography and NMR*. New Haven: Yale University Press.
Fields, B.N., Knipe, D. M. & Howley, P.M. (1996). Editors. *Fields' Virology*. Philidelphia, New York: Lippincott–Raven.
Grimes, J. M., Burroughs, J. N., Gouet, P., Diprose, J. M., Malby, R., Mertens, P. P. C. & Stuart, D. I. (1998). *Nature (London)*, **395**, 470–478.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Koonin, E. V., Gorbalenya, A. E. & Chumakov, K. M. (1989). *FEBS Lett.* **252**, 42–6.
Makeyev, E. V. & Bamford, D. H. (2000). *EMBO J.* **19**, 124–133.
Mindich, L., Nemhauser, I., Gottlieb, P., Romantschuk, M., Carton, J., Frucht, S., Strassman, J., Bamford, D. H. & Kalkkinen, N. (1988). *J. Virol.* **62**, 1180–5.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.