Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and initial X-ray analysis of the head–tail connector of bacteriophage φ29

Mohammed O. Badasso,[a] Petr G. Leiman,[b] Yizhi Tao,[b]† Yongning He,[b] Douglas H. Ohlendorf,[c] Michael G. Rossmann[b]* and Dwight Anderson[a]

[a]Department of Microbiology and Oral Science, University of Minnesota, Minneapolis, MN 55455, USA, [b]Department of Biological Sciences, Purdue University, West Lafayette, IN 47907-1392, USA, and [c]Department of Biochemistry, University of Minnesota, Minneapolis, MN 55455, USA

† Current address: Department of Molecular and Cell Biology, Harvard University, Cambridge, MA 02138, USA.

Correspondence e-mail: mgr@indiana.bio.purdue.edu

Received 4 April 2000
Accepted 26 June 2000

The head–tail connector of bacteriophage φ29, an oligomer of gene product 10 (gp10), was crystallized into various forms. The most useful of these were an orthorhombic $P22_12_1$ form (unit-cell parameters a = 143.0, b = 157.0, c = 245.2 Å), a monoclinic $C2$ form (a = 160.7, b = 143.6, c = 221.0 Å, β = 97.8°) and another monoclinic $C2$ form (a = 177.0, b = 169.1, c = 185.2 Å, β = 114.1°). Frozen crystals diffracted to about 3.2 Å resolution. There is one connector per crystallographic asymmetric unit in each case. Rotation functions show the connector to be a dodecamer. Translation functions readily determined the position of the 12-fold axis in each unit cell. The structure is being determined by 12-fold electron-density averaging within each crystal and by averaging between the various crystal forms.

1. Introduction

The *Bacillus subtilis* bacteriophage φ29 is a small double-stranded DNA (~19 kbp) virus with a prolate head and complex structure (Anderson *et al.*, 1966). Because DNA packaging *in vitro* rivals *in vivo* assembly in efficiency, φ29 is a leading system for study of the DNA packaging mechanism. The φ29 prohead, a precursor capsid into which DNA is packaged, consists of only four structural proteins. These are the head–tail connector (gene product 10; gp10), scaffolding protein (gp7), major capsid protein (gp8) and head fibers (gp8.5) (Méndez *et al.*, 1971). A cyclic hexamer of a 174-base φ29-encoded RNA, prohead RNA (pRNA), is bound to the connector and is needed for DNA packaging (Guo *et al.*, 1998; Hendrix, 1998; Zhang *et al.*, 1998). The connector, pRNA and the ATPase gp16 comprise the DNA-packaging machine; analysis of the structure of these components is needed in order to determine the packaging mechanism.

The φ29 head–tail connector is a cone-shaped oligomer of the 35.9 kDa gp10 (Carazo *et al.*, 1986; Tsuprun *et al.*, 1994; Müller *et al.*, 1997; Valpuesta *et al.*, 1999). The connector occupies the vertex at the base of both the prohead and mature head, which are extended icosahedrons having an elongation number of Q = 5 and 235 subunits (Tao *et al.*, 1998). The connector has a diameter ranging from 140 to 80 Å, with an axial channel of 60–35 Å that is large enough to accommodate DNA passage (Tao *et al.*, 1998). Free connectors wrap supercoiled DNA in a single turn around their circumference, restraining a negative supercoil, and wrapping of DNA by the connector of the prohead may provide a stable initiation complex for packaging (Turnquist *et al.*, 1992). In addition to its DNA-translocation function, the connector functions in the initiation of prohead assembly (Guo *et al.*, 1991; Anderson & Reilly, 1993). The three-dimensional X-ray structure analysis of the connector protein is important in understanding these functions.

Connectors of double-stranded DNA phages, such as T3 (Carazo *et al.*, 1996), T4 (Driedonks *et al.*, 1981), T7 (Kocsis *et al.*, 1985), λ (Kochan *et al.*, 1984), P22 (Bazinet & King, 1985) and φ29 (Carrascosa *et al.*, 1982; Guasch *et al.*, 1998; Valle *et al.*, 1999), have all shown 12-fold symmetry, although 13-fold symmetry has also been reported (Dube *et al.*, 1993; Tsuprun *et al.*, 1994). These investigations have been based on negative staining, cryo-electron microscopy and image processing, as well as atomic force microscopy. It is clear that φ29 connector interactions with proteins gp8 of the head and gp11 of the neck, pRNA, the DNA-translocating ATPase gp16 and DNA would be better understood if high-resolution X-ray data were available. In this paper, we describe the crystallization conditions that yielded well ordered crystals of φ29 connectors suitable for X-ray analysis and the preliminary crystallographic characterization of these crystals.

2. Materials and methods

2.1. Protein expression and purification

The connectors of bacteriophage φ29 were assembled in *Escherichia coli* (pPC28D1) that expresses gp10 (García *et al.*, 1984; Ibanez *et*

Table 1
φ29 head–tail connector crystal data.

Type	1	2	3
Space group	$P22_12_1$	$C2$	$C2$
Unit-cell dimensions (Å, °)	a = 143.0, b = 157.6, c = 245.2	a = 160.7, b = 143.6, c = 221.0, β = 97.8	a = 177.0, b = 169.1, c = 185.2, β = 114.1
V_M (Å^3 Da^{-1})	3.2	2.9	2.9
No. of connectors per crystallographic asymmetric unit	1	1	1
Resolution limit† (Å)	3.5 (3.60)	4.3 (4.40)	3.2 (3.25)
$R_{\rm merge}$‡ (%)	11.2 (31.3)	7.4 (25.9)	6.5 (24.0)
Completeness‡ (%)	90.6 (71.7)	99.2 (99.5)	99.0 (99.0)
Redundancy	2.53	4.55	3.77
Oscillation angle (°)	1.0	1.5	1.0
Exposure time (s)	20	25	30
Crystal-to-detector distance (mm)	250	420	250
Detector	ADSC Quatum 4 2K × 2K CCD	SBC 3K × 3K CCD	ADSC Quatum 4 2K × 2K CCD
Source	APS 14-BM-C BioCARS	APS 19-ID SBC	APS 14-BM-C BioCARS
Wavelength (Å)	1.00000	0.97935	1.00000
Crystallization conditions	30–35% MPD, 0.2 *M* $MgCl_2$, 0.1 *M* HEPES pH 7.5	30–40% MPD, 0.2 *M* $MgCl_2$, 0.2 *M* NaCl, 0.1 *M* Tris–HCl pH 8.0	36–40% MPD, 0.05 *M* $CaCl_2$, 0.1 *M* Tris–HCl pH 8.0
Max. crystal size (mm)	0.3	0.5	0.3

† Values in parentheses indicate the inside resolution of the outermost shell. ‡ Values in parentheses indicate results for the outermost resolution shell.

al., 1984) and purified with several changes to the methods described. Cells were grown overnight at 303 K in LB medium containing ampicillin (50 µg ml^{-1}), diluted tenfold in the same medium and grown to 5×10^8 cells ml^{-1}. The culture was induced at 315 K and cells were pelleted by centrifugation and resuspended in 50 m*M* Tris–HCl pH 7.7, 0.3 *M* KCl (buffer *A*). The cells were lysed in a French press at 83 MPa, treated with DNase I (0.1 mg ml^{-1}) and RNase A (0.2 mg ml^{-1}) and incubated for 1 h at 277 K. The lysate was clarified by centrifugation and the gp10 constituted about one-third of the total protein of the supernatant by SDS–PAGE.

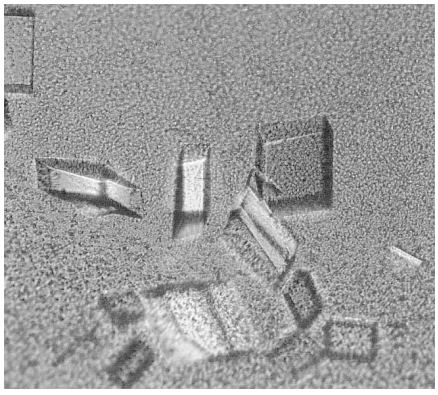

Figure 1
Monoclinic type 3 crystals of bacteriophage φ29 connectors obtained by the vapor-diffusion method using 30–40% MPD in 0.1 *M* Tris–HCl pH 8.0 containing 0.05 *M* $CaCl_2$. The length of one edge of the top right crystal is 0.1 mm.

The protein solution was brought to 50% ammonium sulfate, kept at 277 K overnight and the precipitate was recovered by centrifugation. The pellets were resuspended to about 10 mg ml^{-1} protein in 50 m*M* Tris–HCl pH 7.7, 0.1 *M* NH_4Cl (buffer *B*) and the buffer was changed by centrifugation in Centricon-30 concentrators (Amicon, Beverly, MA, USA). The dialyzed sample was concentrated in an Amicon stirred cell with a PM30 membrane and subjected to ion-exchange chromatography using FPLC (Waters 650E Advanced Protein Purification System). The sample was loaded on a DEAE–Sepharose ion-exchange column (Pharmacia Biotech, Sweden) equilibrated in the same buffer and eluted with a NH_4Cl gradient (0–1.0 *M*) in buffer *B*. The connectors were eluted at 0.35 *M* NH_4Cl. Fractions judged by SDS–PAGE were concentrated in an Amicon stirred cell with a PM30 membrane and applied to a Superdex-200 column (Pharmacia, Biotech, Sweden) equilibrated with buffer *B*. Fractions containing homogenous gp10 were concentrated for crystallization using Centricon-30 concentrators.

2.2. Crystallization

Well ordered connector crystals suitable for X-ray analysis (Fig. 1) of up to 0.5 mm in their maximum dimension were obtained using the vapor-diffusion method. Connector protein at 7–10 mg ml^{-1} in 5 m*M* Tris–HCl pH 7.5 containing 0.4 *M* NaCl was diluted with an equal volume of 0.1 *M* HEPES pH 7.5 containing 0.2 *M* $MgCl_2$ or with an equal volume of Tris–HCl pH 8.0 containing either 0.2 *M* $MgCl_2$ and 0.2 *M* NaCl or 0.05 *M* $CaCl_2$. The diluted protein was then equilibrated against a well solution containing the same buffer and 30–40% MPD. The type 1 orthorhombic crystal form ($P22_12_1$) grew using 30–35% MPD in 0.1 *M* HEPES pH 7.5 containing 0.2 *M* $MgCl_2$. The type 2 and type 3 monoclinic crystals ($C2$) grew in MPD precipitant at 30–40% in 0.1 *M* Tris–HCl pH 8.0 containing either 0.2 *M* $MgCl_2$ and 0.2 *M* NaCl or 0.05 *M* $CaCl_2$, respectively. The crystals appeared after 4–6 weeks and grew to a size of about 0.3–0.5 mm at room temperature (Table 1). Complete data sets were collected from single frozen crystals.

2.3. Synchrotron X-ray data collection and reduction

A variety of different crystal forms were characterized at the Cornell High Energy Synchrotron Source (CHESS) beamline F1, at the Advanced Photon Source BioCARS beamline 14-BM-C and Structural Biology Center beamline 19-ID. The three most useful data sets are described in Table 1. The crystals were flash-frozen in a liquid-nitrogen stream while being protected with a cryosolution of 30–40% MPD in 0.1 *M* HEPES pH 7.5 or 0.1 *M* Tris–HCl pH 8.0 containing 0.2 *M* $MgCl_2$, 0.2 *M* $MgCl_2$ and 0.2 *M* NaCl or 0.05 *M* $CaCl_2$. All data sets were processed using the programs *DENZO* and *SCALEPACK* (Gewirth, 1996) (Table 1).

3. Results and discussion

3.1. Purification, crystallization and data collection

Pure crystalline connector solutions were achieved using a FPLC system that included DEAE–Sepharose followed by gel-filtration on Superdex-200. Most contaminants of the expressed protein solutions were removed during the initial ion-exchange batch chromatography and pure connectors were sieved at the later gel-filtration step. Samples used in crystallization experiments showed a single band by SDS–PAGE. Higher molecular-weight contaminants present in previous purifications (García *et al.*, 1984; Turnquist *et al.*, 1992) were not found. Thus, the earlier main obstacle in obtaining well ordered diffracting crystals of connectors may have been the lack of homogeneity of the purified samples. Connector assemblies were found to be stable in high-salt solutions containing 0.2–0.4 *M* NaCl. Additives such

as $MgCl_2$ and $CaCl_2$ assisted in crystallization and connector protein concentration for crystallization varied from 5 to 10 mg ml^{-1}.

3.2. Structure determination

Rotation-function results clearly showed a 12-fold axis perpendicular to the *a* axis in type 1 crystals and perpendicular to the unique *b* axis in type 2 and 3 crystals. We show (Fig. 2) here only the result for type 3 crystals, but the other rotation functions were essentially of the same quality. The 24 peaks in a great circle in the $\kappa = 180^\circ$ rotation-function section is a consequence of the product of the 12-fold axis perpendicular to a crystallographic twofold axis. The rotation function also showed a significant peak at $\psi = 18.4$, $\varphi = 180$, $\kappa = 30^\circ$ (Fig. 2). The 12-fold symmetry of the φ29 connector is consistent with the results reported by Guasch *et al.* (1998). The Matthews coefficient (Table 1) shows that there is one connector per crystallographic asymmetric unit in each crystal form, assuming a molecular weight of 36 kDa per monomer or 432 kDa per connector assembly.

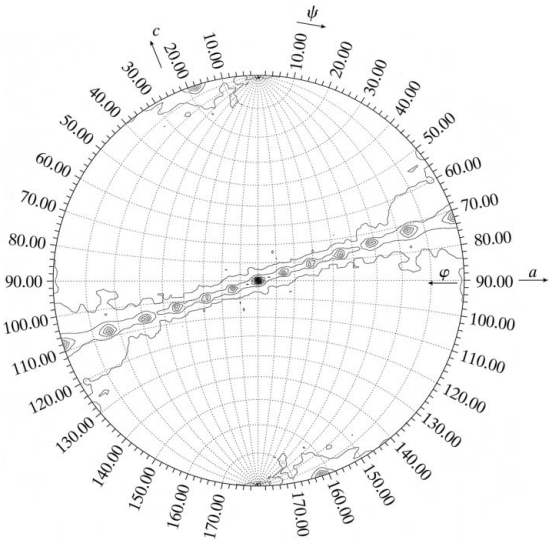

Figure 2
Stereographic projections of the $\kappa = 180^\circ$ section self-rotation function for type 3 crystals calculated with the general locked rotation function (Tong & Rossmann, 1997). Resolution of the data was between 10 and 5 Å. The integration radius was 80.0 Å. The 12-fold axis of the particle is perpendicular to the *b* axis. Contours are at about one standard deviation interval. The 24 peaks around the great circle are perpendicular to the NCS 12-fold axis and are a consequence of a product between the NCS axis and the crystallographic twofold axis. The peak near the *c* axis is a sub-operator of the 12-fold axis.

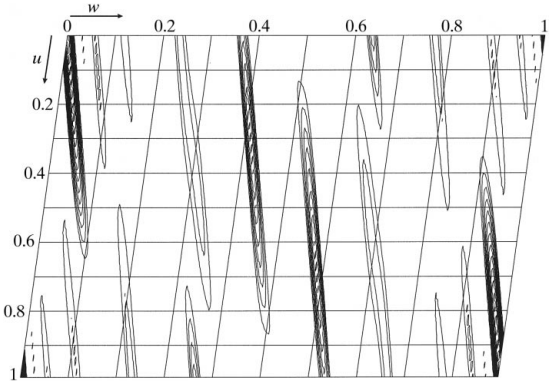

Figure 3
The translation function $T(\mathbf{S})$ for the type 2 crystal form. Data were limited to 4.0 Å resolution. The radius of integration was 80.0 Å. Shown is the section $y = 0$, thus giving the distance between NCS axes related by twofold axes in the C2 space group. The $T(\mathbf{S})$ streak passing through the origin is a result of the symmetry in the self-Patterson, whereas the major other streak is the result of the NCS symmetry between twofold-related connector molecules.

The position of the connector 12-fold axis in the type 1 orthorhombic form could be found readily from a Patterson function because the non-crystallographic symmetry (NCS) 12-fold axis was parallel to the crystallographic *b* axis. The resultant very large Patterson peak located the NCS axis running parallel to *b* through the point at $x = 0.0635$, $z = 0.1250$ in the cell. The position of the NCS 12-fold axes in the two monoclinic unit cells was established by a translation function (Argos & Rossmann, 1980) of the form

$$T(\mathbf{S}) = \sum_{\mathbf{h}} \sum_{\mathbf{p}} \{F_{\mathbf{h}}^2 F_{\mathbf{p}}^2 G_{\mathbf{hp}}^2 \times \cos[2\pi(\mathbf{h}+\mathbf{p})\mathbf{S}]\},$$

where **h** and **p** are Miller indices, **S** is a position in the Patterson map and $G_{\mathbf{hp}}$ is a diffraction function dependent on the direction of the NCS axis and the radius of integration (Rossmann & Blow, 1962). The cross vectors in a Patterson of a structure containing a 12-fold NCS axis perpendicular to the crystallographic twofold axis will contain a NCS twofold at a distance **S** from the origin. The $T(\mathbf{S})$ function measures the agreement of features in the Patterson after rotation about the NCS axis. Hence, the $T(\mathbf{S})$ function will be large when the point **S** is situated on the NCS axis in the Patterson at a distance of twice the distance from the origin as in real space (Fig. 3).

Attempts are currently in progress to obtain suitable phasing starting points with reasonable assumptions about the molecular envelope and connector structure. The initial electron-density maps are being averaged both within and between each of the three crystal forms while extending the resolution in suitably small increments (Arnold *et al.*, 1987).

We thank Charlene Peterson, Cheryl Towell and Sharon Wilder for help in preparation of the manuscript. We also acknowledge the provision of beam time at CHESS and the BioCARS and the Structural Biology Center beamlines of the Advanced Photon Source. This work was supported by National Institutes of Health grant DE03606 to DA and by National Science Foundation grant MCB9603571 to MGR.

References

Anderson, D. L., Hickiman, D. D. & Keilly, B. E. (1966). *J. Bacteriol.* **91**, 2081–2089.
Anderson, D. L. & Reilly, B. E. (1993). *Bacillus subtilis and other Gram-Positive Bacteria: Biochemistry, Physiology and Molecular Genetics*, edited by A. L. Sonenshein, J. A. Hock & R. Losick, pp. 859. Washington, DC: American Society for Microbiology.
Argos, P. & Rossmann, M. G. (1980). *Theory and Practice of Direct Methods in Crystallography*, edited by M. F. C. Ladd & R. A. Palmer, pp. 361–417. New York: Plenum.
Arnold, E., Vriend, G., Luo, M., Griffith, J. P., Kamer, G., Erickson, J. W., Johnson, J. E. & Rossmann, M. G. (1987). *Acta Cryst.* A**43**, 346–361.
Bazinet, C. & King, J. (1985). *Annu. Rev. Microbiol.* **39**, 109–129.
Carazo, J. M., Donate, L. E., Herranz, L., Secilla, J. P. & Carrascosa, J. L. (1986). *J. Mol. Biol.* **192**, 853–867.
Carazo, J. M., Fujisawa, H., Nakasu, S. & Carrascosa, J. L. (1996). *J. Ultrastruct. Mol. Struct. Res.* **94**, 105–113.
Carrascosa, J. L., Vinuela, E., Garcia, N. & Santisteban, A. (1982). *J. Mol. Biol.* **154**, 311–324.
Driedonks, R. A., Engel, A., TenHeggeler, B. & Van Driel, R. (1981). *J. Mol. Biol.* **152**, 641–662.
Dube, P., Tavares, P., Lurz, R. & van Heel, M. (1993). *EMBO J.* **12**, 1303–1309.
García, J. A., Mendez, E. & Salas, M. (1984). *Gene*, **30**, 87–98.
Gewirth, D. (1996). *The HKL Manual. A Description of the Programs DENZO, XDISPLAYF and SCALEPACK.* New Haven: Yale University.
Guasch, A., Pous, J., Párraga, A., Valpuesta, J. M., Carrascosa, J. L. & Coll, M. (1998). *J. Mol. Biol.* **281**, 219–225.
Guo, P., Erickson, S., Xu, W., Olson, N., Baker, T. S. & Anderson, D. (1991). *Virology*, **183**, 366–373.
Guo, P., Zhang, C., Chen, C., Garver, K. & Trottier, M. (1998). *Mol. Cell.* **2**, 149–155.

Hendrix, R. W. (1998). *Cell*, **94**, 147–150.
Ibanez, C., Garcia, J. A., Carrascosa, J. L. & Salas, M. (1984). *Nucleic Acids Res.* **12**, 2351–2365.
Kochan, J., Carrascosa, J. L. & Murialdo, H. (1984). *J. Mol. Biol.* **174**, 433–447.
Kocsis, E., Cerritelli, M., Trus, B., Cheng, N. & Steven, A. C. (1985). *Ultramicroscopy*, **60**, 219–228.
Méndez, E., Ramírez, G., Salas, M. & Viñuela, E. (1971). *Virology*, **45**, 567–576.
Müller, D. J., Engel, A., Carrascosa, J. L. & Vélez, M. (1997). *EMBO J.* **16**, 2547–2553.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Tao, Y., Olson, N. H., Xu, W., Anderson, D. L., Rossmann, M. G. & Baker, T. S. (1998). *Cell*, **95**, 431–437.
Tong, L. & Rossmann, M. G. (1997). *Methods Enzymol.* **276**, 594–611.
Tsuprun, V., Anderson, D. & Egelman, E. H. (1994). *Biophys. J.* **66**, 2139–2150.
Turnquist, S., Simon, M., Egelman, E. & Anderson, D. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 10479–10483.
Valle, M., Kremer, L., Martînez-A, C., Roncal, F., Valpuesta, J. M., Albar, J. P. & Carrascosa, J. L. (1999). *J. Mol. Biol.* **288**, 899–909.
Valpuesta, J. M., Fernandez, J. J., Carazo, J. M. & Carrascosa, J. L. (1999). *Structure*, **7**, 289–296.
Zhang, F., Lemieux, S., Wu, X., St-Arnaud, D., McMurray, C. T., Major, F. & Anderson, D. (1998). *Mol. Cell.* **2**, 141–147.