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

Crystallization of the C-terminal globular domain of avian reovirus fibre

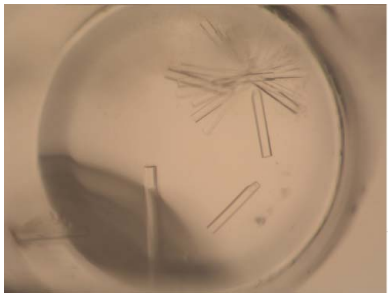
Mark J. van Raaij,[a,b]* X. Lois Hermo Parrado,[a] Pablo Guardado Calvo,[a] Gavin C. Fox,[c] Antonio L. Llamas-Saiz,[b] Celina Costas,[a]‡ José Martínez-Costas[a] and Javier Benavente[a]

[a]Departamento de Bioquímica y Biología Molecular, Facultad de Farmacia, Universidad de Santiago de Compostela, Campus Sur, E-15782 Santiago de Compostela, Spain, [b]Unidad de Difracción de Rayos X, Laboratorio Integral de Dinámica y Estructura de Biomoléculas José R. Carracido, Edificio CACTUS, Universidad de Santiago de Compostela, Campus Sur, E-15782 Santiago de Compostela, Spain, and [c]Spanish CRG Beamline BM16, European Synchrotron Radiation Facility (ESRF), 6 Rue Jules Horowitz, BP 220, F-38043 Grenoble, France

‡ Current address: Cancer Research UK Molecular Oncology Unit, Institute of Cancer, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, England.

Correspondence e-mail: vanraaij@usc.es

Received 1 April 2005
Accepted 27 May 2005
Online 15 June 2005



Avian reovirus fibre, a homotrimer of the σC protein, is responsible for primary host-cell attachment. Using the protease trypsin, a C-terminal σC fragment containing amino acids 156–326 has been generated which was subsequently purified and crystallized. Two different crystal forms were obtained, one grown in the absence of divalent cations and belonging to space group $P6_322$ (unit-cell parameters $a = 75.6$, $c = 243.1$ Å) and one grown in the presence of either zinc or cadmium sulfate and belonging to space group $P321$ (unit-cell parameters $a = 74.7$, $c = 74.5$ Å and $a = 73.1$, $c = 69.9$ Å for the Zn^{II}- and Cd^{II}-grown crystals, respectively). The first crystal form diffracted synchrotron radiation to 3.0 Å resolution and the second form to 2.2–2.3 Å. Its closest related structure, the C-terminal fragment of mammalian reovirus fibre, has only 18% sequence identity and molecular-replacement attempts were unsuccessful. Therefore, a search is under way for suitable heavy-atom derivatives and attempts are being made to grow protein crystals containing selenomethionine instead of methionine.

1. Introduction

Avian reovirus, a non-enveloped virus and a member of the *Orthoreovirus* genus of the *Reoviridae* family, is an important poultry pathogen (Rosenberger *et al.*, 1989). The virus has a dsRNA genome consisting of ten segments encased by a double concentric icosahedral capsid (Spandidos & Graham, 1976). The inner capsid is made up of the λA, λC and σA proteins, which enclose up to 12 copies of the RNA-dependent RNA polymerase complex (consisting of λB and μA), together forming the 'core'. Proteins μB, σB and σC are components of the outer shell (Martínez-Costas *et al.*, 1997). The virus also encodes several non-structural proteins, involved in cell fusion (p10), virus assembly (muNS and σNS) or as yet unknown functions (p17) (Varela *et al.*, 1996; Bodelón *et al.*, 2001; Touris-Otero *et al.*, 2004; Costas *et al.*, 2005).

The minor outer capsid protein, σC, is the avian reovirus fibre, the avian counterpart of mammalian reovirus σ1. σC is an elongated homotrimer and is responsible for primary host-cell attachment (Grande *et al.*, 2000, 2002). Avian reovirus σC has 20% sequence identity to mammalian reovirus σ1. The most similar region is at the N-terminus, which is presumably implicated in the interaction with λC (the equivalent of mammalian reovirus λ2). The region of σC comprising residues 50–156 is predicted to form a triple coiled-coil structure (Costas Iglesias, 2004), while the C-terminal part (residues 156–326) probably forms a globular receptor-binding domain.

The mammalian reovirus receptor has recently been determined to be a junction adhesion molecule, JAM (Barton *et al.*, 2001). However, σC probably binds a different as yet unknown receptor, since the C-terminal region of the two proteins have only 18% sequence identity and since mammalian reoviruses do not attach to avian cells (Barton *et al.*, 2001). The structure of a C-terminal fragment of mammalian reovirus σ1 containing the receptor-binding domain and part of the elongated shaft has been solved (Chappell *et al.*, 2002). The receptor-binding domain has a β-barrel fold, while the part of the elongated shaft contains several triple β-spiral repeats comparable to those of adenovirus fibre (van Raaij *et al.*, 1999).

Using proteolysis, we have now identified a crystallizable C-terminal fragment of σC. Resolution of its structure should hopefully shed light on the differences between the fibres of avian and mammalian reoviruses and provide clues about receptor binding.

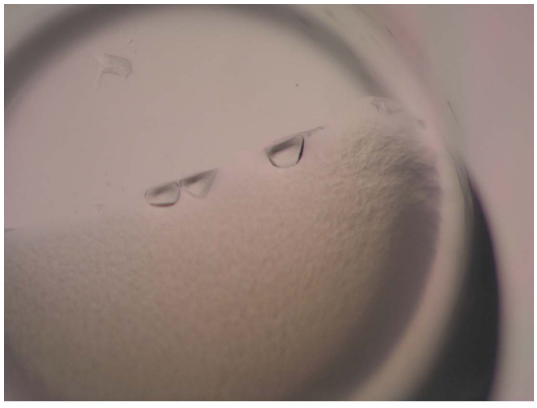
2. Methods

A DNA-fragment encoding σC residues 117–326 was produced by the polymerase chain reaction and cloned into the expression vector pET28c+ (Novagen, Darmstadt, Germany). The resultant plasmid pET28-σC117-326 encodes the mentioned σC fragment fused to an N-terminal purification tag containing six consecutive histidine residues and a T7 tag (the N-terminal 11 residues of the bacteriophage T7 gene 10 product). The sequence of the insert was confirmed by DNA-sequence analysis (Sistemas Genómicos, Valencia, Spain).

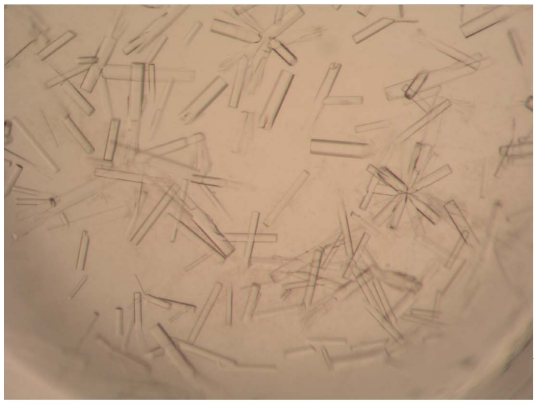
For expression, *Escherichia coli* strain BL21(DE3) was freshly transformed with the plasmid and four 1 l cultures were grown aerobically at 310 K to an optical density of 0.6–0.8 measured at 600 nm. The cultures were cooled to below 298 K, after which expression was induced by adding 1 m*M* isopropyl-thio-β-D-galactopyranoside and was allowed to continue for 4–5 h at 298 K. Harvested cells were resuspended in 40 ml cold resuspension buffer [4.29 m*M* disodium hydrogen phosphate, 1.47 m*M* potassium dihydrogen phosphate, 2.7 m*M* potassium chloride, 137 m*M* sodium chloride, 0.1%(*w*/*v*) Tween-20] and frozen at 253 K. Bacteria were lysed by a double pass through an emulsifier (Avestin emulsifier C5, Avestin Europe GmbH, Mannheim, Germany). After removing insoluble material, 3 ml Ni–NTA resin (Qiagen, Valencia CA, USA) was added. The suspension was incubated for 1 h at 277 K and poured into an empty column. The resin was washed with PBS buffer (resuspension buffer without Tween-20) and elution was performed with a step gradient of imidazole pH 7 in PBS buffer (10, 20, 50, 100, 200, 500 m*M* imidazole). His-T7-tagged σC eluted at high imidazole concentrations (200–500 m*M*), consistent with the protein being a trimer and containing three 6×His tags. This procedure was repeated up to four times until no more His-T7-tagged σC could be captured from the solution.

The partially purified protein was incubated with 1 mg ml^{-1} trypsin for 30 min at 310 K, precipitated by adding 1.5 volumes of saturated ammonium sulfate, redissolved in TE buffer (10 m*M* Tris–HCl pH 8.5, 1 m*M* EDTA), dialysed against the same buffer and applied onto a 1 ml Uno-Q column (Biorad, Barcelona, Spain). Pure σC eluted right at the start of a linear 0–1 *M* sodium chloride gradient in TE buffer. The protein was concentrated to between 15 and 20 mg ml^{-1} using Centricon concentrators (Millipore, Madrid, Spain), incorporating three washes with TE buffer to eliminate small-molecule impurities.

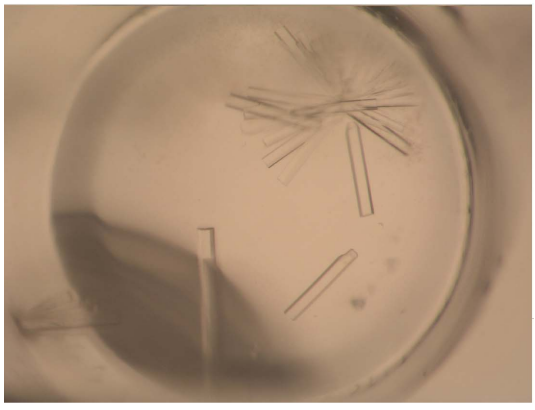
Crystallization took place by vapour diffusion in sitting-drop CompactClover plates (Jena Biosciences, Jena, Germany), with 0.1–0.15 ml reservoirs and drops of 2–5 µl protein solution mixed with 2–5 µl reservoir solution. For N-terminal sequence analysis, protein from the crystallization drop was taken up in 60 m*M* Tris–HCl pH 6.6, 2%(*w*/*v*) sodium dodecylsulfate, 5%(*v*/*v*) β-mercaptoethanol, 10%(*v*/*v*) glycerol, 0.05% bromophenol blue and heated and unheated samples were analysed by SDS–PAGE, transferred to polyvinylidene fluoride membrane and subjected to Edman degradation. Crystallographic data were obtained using *MOSFLM* (Powell, 1999; Leslie, 1999) and scaled using *SCALA* (Collaborative Computational Project, Number 4, 1994).



(*a*)



(*b*)



(*c*)

Figure 1
Crystals of the avian reovirus fibre C-terminal globular domain. (*a*) Irregular prisms belonging to space group $P6_322$. (*b*) Bar-shaped crystals of space group *P*321 grown in the presence of zinc sulfate. (*c*) Bar-shaped crystals of space group *P*321 grown in the presence of cadmium sulfate. The crystals are up to 0.15 mm wide (*a*), 0.1 mm long (*b*) and 0.4 mm long (*c*), respectively.

3. Results and discussion

Previously, σC had been cloned and expressed using the pIL plasmid (Grande *et al.*, 2000, 2002); however, we subsequently found it difficult to reliably obtain significant quantities of soluble and correctly folded protein. To overcome this problem, the protein was cloned into the expression vector pET21d+ (Novagen, Darmstadt, Germany), which expressed soluble trimeric σC with an N-terminal T7 tag, but still in relatively low yield. Trypsin digestion generated a major C-terminal stable fragment of this protein containing residues 115–326 (Costas Iglesias, 2004). Electron microscopy showed this

Table 1
Crystallographic data statistics.

Values in parentheses are for the highest resolution bin.

	Form I	Form II (Zn^{2+})	Form II (Cd^{2+})
Crystal habit	Irregular prisms/lenses	Hexagonal/trigonal bars	Hexagonal/trigonal bars
Crystal size (mm)	0.15 × 0.15 × 0.05	0.025 × 0.025 × 0.1	0.1 × 0.1 × 0.4
Likely space group	$P6_322$	$P321$	$P321$
Unit-cell parameters (Å)	a = 75.6, c = 243.1	a = 74.7, c = 74.5	a = 73.1, c = 69.9
No. of molecules per AU†	2	1	1
No. of molecules per cell†	24	6	6
Matthews coefficient† ($Å^3$ Da^{-1})	2.7	3.3	2.9
Beamline (at ESRF)	ID23-1	Spanish CRG BM16	Spanish CRG BM16
Wavelength (Å)	0.976	1.282	0.980
Detector	MAR Mosaic	165 mm MAR CCD	165 mm MAR CCD
No. of observed reflections‡	8895 (1256)	12298 (1724)	9252 (736)
Resolution range (Å)	30.0–3.0 (3.16–3.00)	20.0–2.2 (2.32–2.20)	30.0–2.3 (2.42–2.32)
Multiplicity	8.8 (9.2)	10.5 (10.6)	10.6 (9.8)
Completeness (%)	99.8 (100.0)	98.3 (97.4)	95.8 (65.5)
$\langle I/\sigma(I)\rangle$	8.2 (2.9)	9.1 (2.8)	8.0 (2.5)
R_{sym}§ (%)	7.7 (26.2)	6.8 (26.5)	6.7 (28.3)

† Estimation based on likely solvent content. ‡ No σ cutoff or other restrictions were used for inclusion of observed reflections. § $R_{sym} = \sum_h \sum_i |I_{hi} - \langle I_h\rangle| / \sum_h \sum_i |I_{hi}|$, where I_{hi} is the intensity of the ith measurement of the same reflection and $\langle I_h\rangle$ is the mean observed intensity of that reflection.

fragment to be globular, with a short vestige of stalk (results not shown).

We have now cloned a very similar fragment (amino acids 117–326, see §2) into the related expression vector pET28c+ and obtained soluble trimeric protein with an N-terminal His-T7 tag. Using low concentrations of trypsin, the expression tag could be removed and the protein was further purified using strong anion chromatography (results not shown). However, extensive trials failed to yield crystals. Therefore, we treated the protein with high concentrations of trypsin (1 mg ml^{-1}, see §2) and obtained a smaller fragment which could be crystallized (Fig. 1). Although we have not yet obtained electron-microscope images for this fragment, it is likely that the protease treatment removes the short remainder of the stalk, leaving the globular domain. It is also likely that this domain retains receptor-binding activity, although this has not yet been tested. Yields of purified concentrated protein were between 5 and 10 mg from 4 l culture (assuming the absorbance at 280 nm of a 1 mg ml^{-1} σC solution to be 1.0).

Crystals were obtained using various conditions, but only those grown using 1.5 *M* ammonium sulfate, 0.1 *M* Tris–HCl pH 8.5, 12% glycerol as reservoir solution yielded interpretable X-ray diffraction patterns, to around 3.0 Å (crystal form I). Optimization of the conditions using divalent-cation additives yielded other related crystal forms that diffracted X-rays to higher resolution (2.2–2.3 Å; crystal forms II and III grown in the presence of 10 m*M* zinc chloride and cadmium chloride, respectively). Crystallographic data were collected at the European Synchrotron Radiation Facility (ESRF) on the Spanish CRG beamline BM16 and on beamline ID23-1. Data from crystals of form I can be scaled in the hexagonal space group $P6_322$ and show systematic absences along the reciprocal axis l consistent with this space group. These crystals are most likely to contain two independent monomers in each asymmetric unit (each forming a trimer arising from crystallographic symmetry). Although the resolution is limited, the data quality is good to 3.0 Å (see Table 1). The crystals of form II belong to the trigonal space group $P321$ (no systematic absences), with one monomer in the asymmetric unit, and diffract X-rays to 2.2–2.3 Å resolution. The length of the c axis varies between crystals, a property that may make different crystals non-isomorphous and structure solution more difficult. This property appears to be correlated with the nature of the divalent cation used in the crystallization medium (Table 1), although we have not yet tested enough crystals to be certain.

SDS–PAGE combined with N-terminal sequence analysis showed that the crystals contained amino acids 156–326 of σC, although we cannot exclude the removal of a few residues from the C-terminus. Taking advantage of the extraordinary stability of the σC trimer, we compared the electrophoretic mobility of heated and unheated samples of dissolved crystals in the presence of sodium dodecyl-sulfate. The results showed that the σC 'head' domain, similar to the full-length protein, retains its trimeric structure when the sample is not boiled.

We were not successful in molecular-replacement attempts using the equivalent domain of the mammalian reovirus fibre (PDB code 1kke; Chappell *et al.*, 2002), presumably owing to the low sequence identity (18%). Similarly, zinc SAD and MAD structure-solution attempts were not successful, possibly because the crystals do not contain ordered zinc ions or because there is only one ordered zinc ion per σC trimer on the threefold crystallographic symmetry axis, as has been reported for the structure of the receptor-binding domain of the bacteriophage short tail fibre (PDB code 1ocy; Thomassen *et al.*, 2003). We are currently trying to obtain heavy-atom derivatives, which will hopefully allow us to solve the structure. Alternatively, we will attempt to grow crystals of protein containing selenomethionine instead of methionine. The structure will reveal whether the avian reovirus σC head domain has structural homology to that of mammalian reovirus σ1, despite its low sequence homology. We also hope to obtain other information from the structure, such as which regions may be involved in receptor binding.

We thank Salvador Blanco Turnes for technical assistance, Stefan Miller (Profos AG, Regensburg, Germany) for providing crystallization plates, Javier Varela (CSIC-CIB, Madrid, Spain) for N-terminal sequence analysis and Dave Hall (ESRF, France) for help with data collection. This research was sponsored by research grants BMC2002-02436 (MJvR) and BMC2001-2839 (JB) from the Spanish Ministry of Science and Technology and PGIDIT03PXIC20307PN (MJvR) and PGIDIT02PXIC20301PN (JB) from the Xunta de Galicia.

References

Barton, E. S., Forrest, J. C., Connolly, J. L., Liu, Y., Schnell, F. J., Nusrat, A., Parkos, C. A. & Dermody, T. S. (2001). *Cell*, **104**, 441–451.

Bodelón, G., Labrada, L., Martinez-Costas, J. & Benavente, J. (2001). *Virology*, **290**, 181–191.
Chappell, J. D., Prota, A. E., Dermody, T. S. & Stehle, T. (2002). *EMBO J.* **21**, 1–11.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Costas Iglesias, C. (2004). PhD thesis, Universidad de Santiago de Compostela, Spain.
Costas, C., Martínez-Costas, J., Bodelón, G. & Benavente, J. (2005). *J. Virol.* **79**, 2141–2150.
Grande, A., Costas, C. & Benavente, J. (2002). *J. Gen. Virol.* **83**, 131–139.
Grande, A., Rodriguez, E., Costas, C., Everitt, E. & Benavente, J. (2000). *Virology*, **274**, 367–377.
Leslie, A. G. W. (1999). *Acta Cryst.* D**55**, 1696–1702.
Martínez-Costas, J., Grande, A., Varela, R., García-Martínez, C. & Benavente, J. (1997). *J. Virol.* **71**, 59–64.
Powell, H. R. (1999). *Acta Cryst.* D**55**, 1690–1695.
Raaij, M. J. van, Mitraki, A., Lavigne, G. & Cusack, S. (1999). *Nature (London)*, **401**, 935–938.
Rosenberger, J. K., Sterner, F. J., Botts, S., Lee, K. P. & Margolin, A. (1989). *Avian Dis.* **33**, 535–544.
Spandidos, D. A. & Graham, A. F. (1976). *J. Virol.* **19**, 968–976.
Thomassen, E., Gielen, G., Schutz, M., Schoehn, G., Abrahams, J. P., Miller, S. & van Raaij, M. J. (2003). *J. Mol Biol.* **331**, 361–373.
Touris-Otero, F., Martínez-Costas, J., Vakharia, V. N. & Benavente, J. (2004). *Virology*, **319**, 94–106.
Varela, R., Martinez-Costas, J., Mallo, M. & Benavente, J. (1996). *J. Virol.* **70**, 2974–2981.