Abhimanyu K. Singh, Rosa
Menéndez-Conejero, Carmen
San Martín and Mark J. van
Raaij*

Departamento de Estructura de Macromolé-
culas, Centro Nacional de Biotecnología
(CNB-CSIC), Calle Darwin 3, 28049 Madrid,
Spain

Correspondence e-mail: m.j.vanraaij@cnb.csic.es

Received 13 September 2013
Accepted 23 October 2013

Crystallization of the C-terminal domain of the fibre protein from snake adenovirus 1, an atadenovirus

Adenovirus fibre proteins play an important role in determining viral tropism. The C-terminal domain of the fibre protein from snake adenovirus type 1, a member of the *Atadenovirus* genus, has been expressed, purified and crystallized. Crystals were obtained belonging to space groups $P2_12_12_1$ (two different forms), $I2_13$ and $F23$. The best of these diffracted synchrotron radiation to a resolution of 1.4 Å. As the protein lacks methionines or cysteines, site-directed mutagenesis was performed to change two leucine residues to methionines. Crystals of selenomethionine-derivatized crystals of the $I2_13$ form were also obtained and a multi-wavelength anomalous dispersion data set was collected.

1. Introduction

First discovered in 1953, adenoviruses (family *Adenoviridae*) have been studied to understand gene expression in mammals (Philipson, 1995) and in recent times have attained importance as potential vectors for gene therapy (McConnell & Imperiale, 2004). Adenoviruses are characterized by 70–90 nm icosahedral, non-enveloped virions which enclose a non-segmented, linear double-stranded DNA. There are five established genera, *Mastadenovirus*, *Aviadenovirus*, *Atadenovirus*, *Siadenovirus* and *Ichtadenovirus* (Harrach *et al.*, 2011), classified based on phylogeny and the arrangement of genes within their genomes. Mastadenoviruses target mammalian species, while aviadenoviruses infect birds. Siadenoviruses have been isolated from various vertebrate animals, including frogs, turtles and birds. Ichtadenoviruses have only one confirmed member so far, infecting sturgeon. Atadenoviruses have been found in reptilian hosts as well as in ruminants and birds (Benkő & Harrach, 1998). The name *Atadenovirus* comes from the high AT content of the genomes of the first atadenoviruses identified.

The adenovirus capsid is a large macromolecular complex of 150 MDa in molecular weight and contains nearly one million amino acids (Nemerow *et al.*, 2012). The main structural proteins are the hexon, of which 12 trimers are present on each of the 20 facets of the icosahedron, the penton base, of which a pentamer forms each of the 12 vertices, and the trimeric fibre protein, which protrudes from each of the vertices. Some other proteins are also present both on the outside surface of the capsid as well as inside to stabilize the complex (San Martín, 2012). The conserved N-terminal domain of the fibre is inserted into the penton base, while the following shaft domain (van Raaij, Mittraki *et al.*, 1999) projects the C-terminal fibre head away from the capsid. The C-terminal domain recognizes cell-surface receptors, establishing the primary virus–host interaction. This is then followed by a secondary interaction, followed by internalization (Wickham *et al.*, 1993). As the C-terminal domains of known fibre proteins are important for host-cell recognition, characterization of these domains from different species may lead to the development of specific adenovirus-based vector systems (Nicklin *et al.*, 2005).

Snake adenovirus 1 (SnAdV1) was first isolated from the corn snake (*Elaphe guttata*; Juhasz & Ahne, 1993) and initial genome characterization indicated that it belongs to the *Atadenovirus* genus (Farkas *et al.*, 2002). A later, more detailed, genomic analysis revealed that SnAdV1 possesses an equilibrated base composition with 49.8%

© 2013 International Union of Crystallography
All rights reserved

AT content (Farkas *et al.*, 2008), rather than that of around 65% typical for ruminant adenoviruses (Das *et al.*, 2006). However, phylogenetic analysis clearly groups SnAdV1 with other members of the genus (Farkas *et al.*, 2008). Here, we report the expression, purification, crystallization and preliminary X-ray diffraction data analysis of the C-terminal domain of the fibre protein from SnAdV1, the first *Atadenovirus* protein that has been crystallized. By using site-directed mutagenesis and expression in a methionine-auxotroph bacterial strain, we also obtained selenomethionine-containing derivative crystals as a first step towards phase determination.

2. Methods

2.1. Viral DNA isolation

The virus was propagated in iguana heart cells (IgH-2; ATCC reference CCL-108; Sorensen & Mesner, 2005) and purified in double caesium chloride gradients as described by Menéndez-Conejero (2013). The final virus concentration was 3.2×10^{10} viral particles per millilitre of 0.15 M sodium chloride buffered with sodium 4-(2-hydroxyethyl)-1-piperazineethanesulfonate (HEPES) pH 7.8 (as quantified by absorbance measurements; Maizel *et al.*, 1968). Proteinase K (New England Biolabs, Ipswich, Massachusetts, USA) was added to a final concentration of 1 mg ml^{-1} to 0.15 ml virus solution. The tube was incubated overnight at 310 K. The next day, 0.15 ml water was added and DNA was extracted with 0.3 ml phenol:chloroform:isoamyl alcohol [25:24:1(v:v:v)] saturated with 10 mM Tris-HCl pH 8.0, 1 mM EDTA (Sigma-Aldrich, St Louis, Missouri, USA) and then with chloroform:isoamyl alcohol [24:1(v:v)]. DNA was precipitated by addition of sodium chloride to a final concentration of 0.4 M and absolute ethanol [5:2(v:v) ethanol:sample ratio] and incubation overnight at 253 K. Following centrifugation (10 min at 16 100g), the DNA pellet was washed with 0.3 ml 70%(v/v) ethanol. Approximately 0.6 µg of viral DNA was obtained.

2.2. Expression vector construction

Coding sequences for SnAdV1 fibre-protein constructs were amplified by the polymerase chain reaction (PCR) using forward primers including a *Bam*HI restriction site and a reverse primer with a *Hind*III restriction site. Amplified PCR products were cloned into expression vector pET-28c(+) (Novagen, Merck, Darmstadt, Germany) previously digested with the same restriction enzymes. The sequences of the inserts of pET-28c-SnAdV1f171 and pET-28c-SnAdV1f234 were verified (Secugen SL, Madrid, Spain).

2.3. Proteomic analysis

Purified SnAdV1 virions were dissolved in 8 M urea, 25 mM ammonium bicarbonate, reduced and alkylated with iodoacetamide (50 mM final concentration). The urea concentration was reduced to 2 M by dilution with 25 mM ammonium bicarbonate and the sample was digested with trypsin (Roche Diagnostics GmbH, Mannheim, Germany) in a 25:1(w:w) protein:trypsin ratio overnight at 310 K. Digested samples were dried in a SpeedVac and stored at 253 K. Coupled liquid chromatography/mass spectrometric analysis was performed as described by Menéndez-Conejero (2013).

2.4. Site-directed mutagenesis

A single set of primers (5'-GAC GAG GGA ATT **ATG** ACT **ATG** GAG ATT TCT CGC-3' and 5'-GCG AGA AAT CTC **CAT** AGT **CAT** AAT TCC CTC GTC-3') was designed containing the desired leucine-to-methionine changes (shown in bold). A 50 µl reaction

mixture containing 100 ng pET-28c-SnAdV1f234 plasmid and *Pfu* Turbo DNA polymerase (Agilent Technologies, Santa Clara, California, USA) was subjected to 16 cycles of PCR with a final extension step of 7 min. For digestion of parental plasmid DNA, the PCR product was treated with *Dpn*I (Agilent Technologies). 40 units of *Dpn*I were used, where a unit is defined as the amount required to digest 1 µg of dam methylated pBR322 DNA in 1 h at 310 K. In our case, the reaction was incubated for 3 h at 310 K. The introduced mutations were confirmed by sequencing (Secugen SL, Madrid, Spain); the resulting vector is named pET-28c-SnAdV1f234mut.

2.5. Protein expression and purification

Three different protein constructs were crystallized: one starting at residue 234, which led to three different native crystal forms (native 1, 2 and 3), and one starting at residue 171 (which led to native crystal form 4), plus a selenomethionine derivative starting at residue 234 in which two leucine residues (322 and 324) were mutated to methionines. In this section, the expression and purification of all three constructs are described.

For purification of the construct starting at residue 234, *Escherichia coli* BL21(DE3) cells (Novagen, Merck, Darmstadt, Germany) transformed with pET-28c-SnAdV1f234 were grown in 3 l Luria-Bertani medium at 310 K until the optical density at 600 nm reached 0.6–0.8. The cultures were then cooled on ice for 15–30 min and induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Following growth at 289 K overnight, the cells were collected by centrifugation at 5000g for 15 min and resuspended in lysis buffer consisting of 50 mM Tris-HCl pH 8.0, 300 mM sodium chloride, 10 mM imidazole, 10%(v/v) glycerol. Lysis was carried out in a French press (three passes) and the cell debris was removed by centrifugation at 20 000g for 45 min. The supernatant was incubated with 1.5 ml nickel-nitrilotriacetic acid (Ni-NTA) resin (Jena Bioscience GmbH, Jena, Germany) at room temperature for 1 h and then loaded onto an empty column, followed by washing twice with ten column volumes of buffer A [50 mM Tris-HCl pH 8.0, 300 mM sodium chloride, 20 mM imidazole, 10%(v/v) glycerol]. Step-gradient elution of the bound protein was performed using imidazole (0.05, 0.20, 0.30, 0.4 and 1 M in buffer A) and the fractions were checked by SDS-PAGE. Eluates of 0.2–1 M imidazole in buffer A were pooled together and dialysed against 10 mM Tris-HCl pH 8.5 for strong anion-exchange chromatography using a Resource Q column (GE Healthcare Biosciences, Uppsala, Sweden). The protein was eluted with a linear gradient of 0–1 M sodium chloride in 10 mM Tris-HCl pH 8.5. The collected peak fractions were desalted and concentrated to 10 mg ml^{-1} using Amicon Ultra concentrators (Millipore Iberica, Madrid, Spain), applying three washes with 10 mM Tris-HCl pH 8.5. The sample was stored at 277 K prior to crystallization trials.

For purification of the construct starting at residue 171, BL21(DE3) cells were transformed with pET-28c-SnAdV1f171 followed by the same procedure as for the fragment starting at residue 234, except that after NTA-affinity chromatography fractions eluting at between 0.1 and 1 M imidazole were pooled. After anion-exchange chromatography, the protein was desalted as above and concentrated to 13 mg ml^{-1} in 10 mM Tris-HCl pH 8.5.

Selenomethionine-derivatized protein (from the construct starting at residue 234 in which leucines 322 and 324 were mutated to methionine) was prepared by transforming *E. coli* B834(DE3) cells (Novagen, Merck, Darmstadt, Germany) with pET-28c-SnAdV1f234mut. A 1.5 l culture was grown overnight at 289 K in SelenoMet medium (Molecular Dimensions, Newmarket, England) according to the manufacturer's instructions. The same cell disruption

and protein purification procedures were followed as performed in the case of the native protein construct starting at residue 234, except for the addition of 10 mM β -mercaptoethanol to all buffers. The obtained protein was desalted and concentrated to 16 mg ml⁻¹, keeping β -mercaptoethanol in the final buffer (10 mM Tris-HCl pH 8.5, 10 mM β -mercaptoethanol).

2.6. Crystallization

Initial crystallization trials for the native protein construct starting at residue 234 were carried out using sitting-drop vapour diffusion. Commercially available kits were employed at 277 and 294 K using either robotic or manual setups. In robotic setups, a Genesis RSP 150 workstation (Tecan, Männedorf, Switzerland) was used, with MRC 96-well sitting-drop crystallization plates (Molecular Dimensions, Newmarket, England). Reservoir volumes were 50 μ l and drops were prepared by mixing 0.2 μ l protein solution with 0.2 μ l reservoir solution. For manual setups, CompactClover 4-Chamber plates (Jena Bioscience, Jena, Germany) were used, with reservoir volumes of 0.15 ml and drops prepared by mixing 1 μ l protein solution with 1 μ l reservoir solution. The crystals used for data collection were always those from manual setups. Several hundred conditions were tried, of which five conditions belonging to the Crystallization Cryo Kit for Proteins (Sigma-Aldrich, St Louis, Missouri, USA) produced crystals in a time period of 15–45 d. Condition 39, consisting of 1.7 M ammonium sulfate, 0.085 M HEPES sodium salt pH 7.5, 1.7% (v/v) polyethylene glycol (PEG) 400, 15% (v/v) glycerol, was further selected for optimization in sitting-drop vapour-diffusion crystallization experiments (manual setups as described above) and was used for crystallizing selenomethionine-derivatized protein.

Crystals of fragment 171–345 of the fibre protein were obtained after about 1 year in 1.5 M ammonium sulfate, 0.075 M Tris-HCl pH 8.5, 25% (v/v) glycerol (condition 4 of the Crystal Cryo Kit, without optimization) at 294 K using the manual setup as described above.

1 MKKIKRSAAD PDPVYPFGDE VPIPLPPFLV PGGGLTDDGL SLAVQTVDPDL
51 NVTLGGVGLK IGDGLSVVDG KLTSEAKIVA DPPLQSGDT LSLSTDSSMM
101 VLPSSGQLTIN NLPSISVTSS GVGLVSPNAP LQLMSNGALQ LSVGGGLTVG
151 AQGSLQISTG VGVNVNAAGV LESYPLPLV WDYSSKSLTL DIGPGLTVVN
201 GKLVIGATF SNQMSRMAPA PRADLQNSI EPLPSPSKT SLDIAEELQN
251 DKGVSAFAQA REEELGAFTK RTLFAYSGDG LTGPFFKAPAS AELSSFLTAH
301 PKGRWLIAPF LGTGIVSVDE GILTLEISRS LPEVSGSSS TSCLKVISIYF
351 MDLFFPVFFI DRASHAPARR SNNSRQLFHS KQRLFLKVKD FKRSWYSSL
401 FTLINLNQIE CPELS

1 MKKIKRSAAD PDPVYPFGDE VPIPLPPFLV PGGGLTDDGL SLAVQTVDPDL
51 NVTLGGVGLK IGDGLSVVDG KLTSEAKIVA DPPLQSGDT LSLSTDSSMM
101 VLPSSGQLTIN NLPSISVTSS GVGLVSPNAP LQLMSNGALQ LSVGGGLTVG
151 AQGSLQISTG VGVNVNAAGV LESYPLPLV WDYSSKSLTL DIGPGLTVVN
201 GKLVIGATF SNQMSRMAPA PRADLQNSI EPLPSPSKT SLDIAEELQN
251 DKGVSAFAQA REEELGAFTK RTLFAYSGDG LTGPFFKAPAS AELSSFLTAH
301 PKGRWLIAPF LGTGIVSVDE GILTLEISRS LPEVSGSSS YLTEK

Figure 1

The snake adenovirus 1 fibre protein. (a) Mass-spectrometric analysis of SnAdV1 fibre protein in purified virions. Peptides identified by *Mascot* in the database sequence (coverage 28%; top) and in the newly determined sequence (coverage 38%; bottom) are shaded grey; the C-terminal peptide identified that is specific to the new sequence is also underlined. (b) Proposed domain structure. The construct start residues for which protein could be obtained are underlined.

2.7. Crystallographic data collection and processing

Crystals were mounted using LithoLoops (Molecular Dimensions, Newmarket, England) or MicroMounts (Mitegen, Ithaca, New York, USA) and vitrified in liquid nitrogen for data collection at 100 K at the European Synchrotron Radiation Facility (ESRF), Grenoble, France or the Diamond Light Source (DLS), Harwell, England.

Crystallographic data collected at the ESRF were integrated using *iMosflm* (Battye *et al.*, 2011) and reduced using *POINTLESS*, *SCALA* and *TRUNCATE* (Evans, 2011), all of which are part of *CCP4* (Winn *et al.*, 2011). Data collected at the DLS were integrated using *xia2* (Winter *et al.*, 2013), which incorporates *XDS* (Kabsch, 2010) and *AIMLESS* (Evans & Murshudov, 2013). *SCALEIT* (Howell & Smith, 1992; Winn *et al.*, 2011) was used to compare native data with derivative data.

3. Results and discussion

Genomic analysis of an adenovirus isolated from a corn snake (*Elaphe guttata*; GenBank entry DQ1064141.1; Farkas *et al.*, 2002; Benkő *et al.*, 2002) suggested that it contains a fibre protein of 415 amino acids in length (Fig. 1a). DNA fragments corresponding to several constructs (two of them starting at amino acids 171 and 234) were amplified by PCR and cloned into an expression vector as described in §2. We assumed that a C-terminal globular head domain might start at residue 234 owing to a short sequence located there (234-PSPP-237) that we think may form a turn structure. Larger and smaller constructs starting at amino acids 1, 82, 171, 293 and 343 were also designed (Fig. 1b). The N-terminal 6 $\times$ His and T7 tag (M GSSHH HHHHS SGLVP RGS HM ASMTG GQQMG RGS) of the vector was retained and fused to the N-terminus of the target proteins to facilitate their purification using metal-affinity chromatography. For the constructs starting at amino acids 293 and 343 no protein was obtained, while proteins starting at residues 1 and 82 were less soluble and did not crystallize. Therefore, only constructs starting at residues 171 and 234 are described in detail.

The recombinant plasmids pET-28c-SnAdV1f171 and pET-28c-SnAdV1f234 were analysed by DNA sequencing, which revealed a major discrepancy between the database sequence and our cloned sequences. We observed an insertion of two bases, which results in a change of the five amino acids beyond position 339 and a termination codon (TGA) after position 345. As the results are from two independent PCR reactions, we believe that they are a reflection of a real difference in the template sequence and are not the result of a PCR artefact. This also explains why we did not observe protein production for constructs starting at residues 293 or 343, because they are likely to be too small to be stable in bacterial expression and would not have been detected by the protein electrophoresis method that we used.

In a proteomic analysis of SnAdV1 virions, peptides corresponding to residues 340–415 of the database sequence are absent, but a peptide including the new sequence was found which confirmed that the cloned gene corresponds to the fibre protein expressed and incorporated into the virions (Fig. 1a). At this point we do not know whether the virus that we cultivated is a variant or whether the sequence submitted to GenBank (entry DQ1064141.1) contains a sequence error.

After transformation of expression vectors into *E. coli* BL21(DE3), soluble protein was obtained for both the 171–345 and 234–345 protein constructs. Proteins were purified by nickel-affinity and ion-exchange chromatography. In strong anion-exchange chromatography, the proteins eluted between 0.2 and 0.4 M sodium

Figure 2

Denaturing gel electrophoresis of purified proteins. Lanes are marked *M* (molecular-weight markers, with the size indicated in kDa), *U* (unheated sample) and *B* (sample heated to 368–373 K for at least 5 min). (a) Purified 234–345 construct protein. (b) Purified 171–345 construct protein.

chloride in two separate peaks. Usually, the first peak was more prominent and was used for crystallization, but in one case crystals were also obtained from the second peak (native form 1, isomorphous crystals obtained from both peaks, pooled independently). As observed for C-terminal domains for other adenovirus fibres (van Raaij, Louis *et al.*, 1999; Singh *et al.*, 2013), it is not clear to us how the

protein in the peaks differs, although charge must somehow be involved. The electrophoretic behaviour under denaturing conditions (observed size of around 16 kDa; Fig. 2a) is consistent with the tagged protein construct starting at amino acid 234 ending at residue 345 (expected monomer size of 15 kDa), but not including 70 extra residues as would have been expected from the database sequence (expected monomer size of 24 kDa). The protein construct starting at amino acid 171 has an observed monomer size of around 26 kDa (Fig. 2b), where 22 kDa would have been expected if it ends at residue 345 or 31 kDa if it ends at residue 415. While for the 234–345 construct pure protein could be obtained, the 171–345 construct contained some impurities as evident from the SDS–PAGE gel (Fig. 2). For both constructs, the protein remains trimeric unless heated to at least 368 K for a few minutes before denaturing SDS–PAGE. Note the trimer band of the 171–345 construct entered only a few millimetres into the gel, possibly because of the protein being in a slowly migrating ‘hydra’ form owing to unfolding of the shaft part in SDS, as previously observed for the human adenovirus type 2 fibre protein (Mitraki *et al.*, 1999).

Yields of the 234–345 construct were typically around 5 mg of purified and concentrated protein per litre of bacterial culture, while for the 171–345 construct around 6 mg per litre of bacterial culture could be obtained, of which around half was lost upon concentration, indicating that the protein has a tendency to aggregate.

Crystals of different forms were obtained for the 234–345 construct in the same condition [1.7 M ammonium sulfate, 0.085 M HEPES sodium salt pH 7.5, 1.7% (v/v) PEG 400, 15% (v/v) glycerol]. The first crystal form obtained presented elongated crystals growing in clusters (Fig. 3a), while a second crystal form grew in the shape of diamonds (Fig. 3b). Both of these crystal forms appeared at 294 K after 15–20 d in independent experiments. At 277 K and after around 45 d, a third

Figure 3

Crystals of snake adenovirus 1 fibre-protein constructs. (a) Native crystal form 1 (234–345 construct). (b) Native crystal form 2 (234–345 construct). (c) Native crystal form 3 (234–345 construct). One of the crystals has grown from or around a fibre unintentionally present in the drop. (d) Native crystal form 4 (171–345 construct). (e) Selenomethionine-derivatized crystal of form 3 (234–345 construct). The bars represent 0.5 mm.

Table 1
Crystallographic data measured from native crystals.
Values in parentheses are for the highest resolution bin.

Data collection	Native 1	Native 2	Native 3	Native 4
Construct (residues)	234–345	234–345	234–345	171–345
Synchrotron, beamline	ESRF, ID14-4†	ESRF, ID29‡	ESRF, ID14-1	DLS, I02
Detector	ADSC Q315r	Pilatus 6M	ADSC Q210	Pilatus 6M
Crystal-to-detector distance (mm)	386	344	141	271
Wavelength (Å)	0.9393	0.9999	0.9334	0.9797
No. of images	117, 282§	3600	100	400
Oscillation range (°)	1.5, 0.4§	0.1	0.5	0.15
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>I</i> 23 or <i>I</i> 2 ₁ 3	<i>F</i> 23
Unit-cell parameters (Å)	<i>a</i> = 79.6, <i>b</i> = 122.3, <i>c</i> = 133.7	<i>a</i> = <i>b</i> = 96.8, <i>c</i> = 153.3	<i>a</i> = <i>b</i> = <i>c</i> = 149.6	<i>a</i> = <i>b</i> = <i>c</i> = 121.5
Resolution range (Å)	30.0–2.70 (2.85–2.70)	70.0–1.90 (2.00–1.90)	20.0–1.70 (1.79–1.70)	43.0–1.35 (1.38–1.35)
Reflections	36484 (5266)	113650 (16374)	60984 (8834)	33985 (2424)
Multiplicity	9.3 (9.0)	11.9 (12.1)	6.2 (6.2)	5.8 (2.4)
Completeness (%)	99.7 (99.8)	99.9 (99.9)	99.9 (100.0)	99.5 (95.3)
Mean <i>I</i> (<i>hkl</i>)	10.1 (2.8)	19.6 (6.5)	17.3 (4.2)	18.5 (2.2)
<i>R</i> _{merge} ¶ (%)	16.7 (87.9)	7.2 (39.0)	6.5 (40.0)	4.8 (37.3)
CC Imean	0.993 (0.794)	0.999 (0.953)	0.999 (0.896)	0.998 (0.732)
CC anom	−0.206 (0.021)	−0.254 (−0.046)	−0.067 (0.010)	0.053 (0.046)
Wilson <i>B</i> (Å ²)	55.1	25.1	16.5	15.6

† McCarthy *et al.* (2009). ‡ de Sanctis *et al.* (2012). § Data were collected in two different runs. ¶ $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th measurement of the same reflection and $\langle I(hkl) \rangle$ is the mean observed intensity for that reflection.

Table 2
Crystallographic data measured from a selenomethionine-derivative crystal.
Values in parentheses are for the highest resolution bin.

	Inflection point	Peak	Remote
Construct (residues)	234–345		
Synchrotron, beamline	ESRF, ID14-4†		
Detector	ADSC Q315r		
Space group	<i>I</i> 23 or <i>I</i> 2 ₁ 3		
Wavelength (Å)	0.9793	0.9791	0.9768
Distance (mm)	279	279	211
No. of images	200	200	200
Oscillation range (°)	0.5	0.5	0.5
Unit-cell parameters‡ (Å)	<i>a</i> = <i>b</i> = <i>c</i> = 149.50	<i>a</i> = <i>b</i> = <i>c</i> = 149.45	<i>a</i> = <i>b</i> = <i>c</i> = 149.54
Resolution range (Å)	30.0–1.90 (2.00–1.90)	30.0–1.90 (2.00–1.90)	30.0–1.60 (1.69–1.60)
Reflections	43740 (6346)	43706 (6352)	72960 (10590)
Multiplicity	12.2 (11.7)	12.2 (11.7)	12.2 (12.2)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Mean <i>I</i> (<i>hkl</i>)	17.5 (8.9)	16.2 (8.5)	18.2 (6.2)
<i>R</i> _{merge} § (%)	10.4 (25.4)	11.3 (25.9)	8.8 (40.2)
CC Imean	0.998 (0.981)	0.998 (0.981)	0.999 (0.957)
CC anom	0.500 (0.227)	0.646 (0.403)	0.498 (0.127)
Wilson <i>B</i> (Å ²)	14.9	14.7	13.9

† McCarthy *et al.* (2009). ‡ Although the three data sets were collected from the same crystal, the unit-cell parameters were refined independently for all three data sets. § $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the *i*th measurement of the same reflection and $\langle I(hkl) \rangle$ is the mean observed intensity for that reflection.

crystal form was obtained. These crystals looked like dodecahedra, although they were flatter in the third dimension (Fig. 3c). The crystallization solution could be directly vitrified, as it contained sufficient glycerol and/or PEG 400, and data were collected using synchrotron radiation (for native data-processing statistics see Table 1). It was found that crystal form 1 belonged to space group *P*2₁2₁2₁ and diffracted X-rays to around 2.7 Å resolution. Crystal form 2 also belonged to the orthorhombic space group *P*2₁2₁2₁, but with different unit-cell parameters (cell edges *a* and *b* are very similar but not identical). These crystals diffracted X-rays to around 1.9 Å resolution. The volume of the asymmetric unit is almost the same for both *P*2₁2₁2₁ crystal forms (3.3×10^5 and 3.6×10^5 Å³, respectively), and is expected to contain two, three or four protein trimers, leading to calculated solvent contents of between 85 and 32%. Crystal form 3

belonged to space group *I*23 or *I*2₁3 and diffracted X-rays to around 1.7 Å resolution. In this case, the volume of the asymmetric unit is 1.4×10^5 Å³ and may be expected to contain between three and five protein monomers, leading to predicted solvent contents of between 34 and 60%.

Crystals were also obtained for the 171–345 construct (Fig. 3d), but only after about 1 year, and after the decision was made to generate the double methionine mutant of the 234–345 construct (see below). These crystals belonged to space group *F*23 and diffracted X-rays to 1.35 Å resolution. In this case, the volume of the asymmetric unit is 3.7×10^4 Å³ and can only contain one protein monomer, with a predicted solvent content of 21%. The very low calculated solvent content suggests that the protein sample may have been partially proteolysed before crystallization, leading to a higher actual solvent content.

Although the crystals diffracted well, attempts to introduce heavy atoms for phasing purposes were unsuccessful. The protein (corrected sequence) does not contain cysteine residues for easy derivatization with mercury compounds, although other compounds were also tried. The protein also does not contain methionines beyond residue 217; therefore, we chose to introduce them by site-directed mutagenesis. Leucines were chosen for mutation to methionine as their side-chain physicochemical properties are roughly equivalent (Bordo & Argos, 1991) and their hydrophobicity values are similar (Finney *et al.*, 1980; Guy, 1985). As there were two leucines close together in sequence at positions 322 and 324, a single-step site-directed mutagenesis was performed, changing both of them simultaneously to methionine. The selenomethionine-modified mutant protein could be purified using a very similar protocol to the native, with a yield of 3 mg purified and concentrated protein per litre of bacterial culture.

After crystallization, cubic-shaped crystals were obtained (Fig. 3e) in 1.2 *M* ammonium sulfate, 0.085 *M* HEPES sodium salt pH 7.5, 1.7% (v/v) PEG 400, 15% (v/v) glycerol. The crystals belonged to the same space group as native form 3 and are apparently isomorphous, based on very similar unit-cell parameters and low *R* factors when the derivative data are compared with the native data (29.1% for the native data set scaled to the high-energy remote wavelength data set between 20.0 and 1.7 Å resolution). Data-processing statistics are

summarized in Table 2. As it is not possible to discriminate between $I23$ and $I2_13$ using only diffraction data, structure solution in both space groups will have to be attempted. Using the obtained structure for the derivatized construct starting at residue 234, it should also be possible to solve the structures of the other crystal forms. Structure solution, description and analysis will be published in a separate paper.

For the data sets that we collected, it may be argued that somewhat higher resolution data could have been obtained, as the correlation coefficients and signal-to-noise values are still relatively elevated at the high-resolution limit (Evans & Murshudov, 2013). However, we have chosen cutoffs on the conservative side, either in data collection (*i.e.* crystal-to-detector distance), to avoid spot overlap, or in integration, to ensure that only highly complete high-quality data are included in structure solution and refinement.

4. Conclusion

We present crystallization of a C-terminal domain of the fibre protein of snake adenovirus 1, the first atadenovirus for which this has been reported. We also present crystals of a selenomethionine derivative obtained by site-directed mutagenesis. The high anomalous signal obtained should allow structure solution. Together with the recently reported data on a *Siadenovirus* fibre (turkey adenovirus 3; Singh *et al.*, 2013), this should lead to a deeper understanding of animal adenovirus fibre structure and may lead to applications in biotechnology and novel vaccination and gene-therapy vectors.

We thank Hassan Belrhali (EMBL/ESRF beamline BM14), Andrés Palencia (ESRF beamline ID14-EH1), Philippe Carpentier and Andrew McCarthy (ESRF beamline ID14-EH4), Daniele de Sanctis (ESRF beamline ID29) and James Sandy (DLS beamline I02) for valuable help with synchrotron data collection, Rachel E. Marschang (Hohenheim University, Germany) for IgH-2 cells and SnAdV1 virus seeds, and Alberto Paradela (CNB-CSIC) for expert mass-spectrometric analyses. We thank the European Synchrotron Radiation Facility (proposal Nos. MX1364 and MX1477) and the Diamond Light Source (proposal No. MX3808) for access, which contributed to the results presented here. This research was sponsored by grants BFU2011-24843 (MJvR), BFU2010-16382 (CSM), DE2009-0019 (CSM) and the BioFiViNet network (FIS2011-16090-E) from the Spanish Ministry of Economy and Competitiveness. AKS and RM-C were recipients of pre-doctoral fellowships from La Caixa and the Instituto de Salud Carlos III of Spain, respectively.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Benkő, M., Élő, P., Ursu, K., Ahne, W., LaPatra, S. E., Thomson, D. & Harrach, B. (2002). *J. Virol.* **76**, 10056–10059.
- Benkő, M. & Harrach, B. (1998). *Arch. Virol.* **143**, 829–837.
- Bordo, D. & Argos, P. (1991). *J. Mol. Biol.* **217**, 721–729.
- Das, S., Paul, S. & Dutta, C. (2006). *Virus Res.* **117**, 227–236.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.
- Farkas, S. L., Benkő, M., Élő, P., Ursu, K., Dán, A., Ahne, W. & Harrach, B. (2002). *J. Gen. Virol.* **83**, 2403–2410.
- Farkas, S. L., Harrach, B. & Benkő, M. (2008). *Virus Res.* **132**, 132–139.
- Finney, J. L., Gellatly, B. J., Golton, I. C. & Goodfellow, J. (1980). *Biophys. J.* **32**, 17–33.
- Guy, H. R. (1985). *Biophys. J.* **47**, 61–70.
- Harrach, B., Benkő, M., Both, G. W., Brown, M., Davison, A. J., Echavarría, M., Hess, M., Jones, M. S., Kajon, A., Lehmkuhl, H. D., Mautner, V., Mittal, S. K. & Wadell, G. (2011). *Virus Taxonomy: Classification and Nomenclature of Viruses. Ninth Report of the International Committee on Taxonomy of Viruses*, edited by A. M. Q. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz, pp. 125–141. San Diego: Elsevier.
- Howell, P. L. & Smith, G. D. (1992). *J. Appl. Cryst.* **25**, 81–86.
- Juhász, A. & Ahne, W. (1993). *Arch. Virol.* **130**, 429–439.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Maizel, J. V., White, D. O. & Scharff, M. D. (1968). *Virology*, **36**, 115–125.
- McCarthy, A. A., Brockhauser, S., Nurizzo, D., Theveneau, P., Mairs, T., Spruce, D., Guijarro, M., Lesourd, M., Ravelli, R. B. G. & McSweeney, S. (2009). *J. Synchrotron Rad.* **16**, 803–812.
- McConnell, M. J. & Imperiale, M. J. (2004). *Hum. Gene Ther.* **15**, 1022–1033.
- Menéndez-Conejero, R. (2013). PhD thesis, Universidad Autónoma de Madrid, Spain.
- Mitraki, A., Barge, A., Chroboczek, J., Andrieu, J.-P., Gagnon, J. & Ruigrok, R. W. H. (1999). *Eur. J. Biochem.* **264**, 599–606.
- Nemerow, G. R., Stewart, P. L. & Reddy, V. S. (2012). *Curr. Opin. Virol.* **2**, 115–121.
- Nicklin, S. A., Wu, E., Nemerow, G. R. & Baker, A. H. (2005). *Mol. Ther.* **12**, 384–393.
- Philipson, L. (1995). *Curr. Top. Microbiol. Immunol.* **199**, 1–24.
- Raaij, M. J. van, Louis, N., Chroboczek, J. & Cusack, S. (1999). *Virology*, **262**, 333–343.
- Raaij, M. J. van, Mitraki, A., Lavigne, G. & Cusack, S. (1999). *Nature (London)*, **401**, 935–938.
- Sanctis, D. de *et al.* (2012). *J. Synchrotron Rad.* **19**, 455–461.
- San Martín, C. (2012). *Viruses*, **4**, 847–877.
- Singh, A. K., Ballmann, M. Z., Benkő, M., Harrach, B. & van Raaij, M. J. (2013). *Acta Cryst.* **F69**, 1135–1139.
- Sorensen, E. B. & Mesner, P. W. (2005). *Comp. Biochem. Physiol. B*, **140**, 163–170.
- Wickham, T. J., Mathias, P., Cheresch, D. A. & Nemerow, G. R. (1993). *Cell*, **73**, 309–319.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Winter, G., Lobley, C. M. C. & Prince, S. M. (2013). *Acta Cryst.* **D69**, 1260–1273.