Abhimanyu K. Singh,^a Mónica Z. Ballmann,^b Mária Benkő,^b Balázs Harrach^b and Mark J. van Raaij^{a*}

^aDepartamento de Estructura de Macromoléculas, Centro Nacional de Biotecnología (CNB-CSIC), Calle Darwin 3, E-28049 Madrid, Spain, and ^bInstitute for Veterinary Medical Research, Centre for Agricultural Research, Hungarian Academy of Sciences, H-1143 Budapest, Hungary

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

Received 23 July 2013

Accepted 26 August 2013

Crystallization of the C-terminal head domain of the fibre protein from a siadenovirus, turkey adenovirus 3

Turkey adenovirus 3 belongs to the genus *Siadenovirus*. Its predicted fibre protein consists of an N-terminal virus-attachment domain, a central shaft domain and a head domain at the C-terminus. The head domain has little sequence identity to known adenovirus fibre head structures. Crystals of the fibre head domain consisting of amino acids 304–454 with an N-terminal purification tag were produced. Crystals of native and selenomethionine-derivatized protein belonged to space group *I*23 (unit-cell parameter 99 Å). They diffracted synchrotron radiation to 2.0 and 2.14 Å resolution, respectively, and are expected to contain one monomer in the asymmetric unit.

1. Introduction

The family *Adenoviridae* is divided into five genera: *Mastadenovirus*, *Aviadenovirus*, *Atadenovirus*, *Siadenovirus* and *Ichtadenovirus* (Harrach *et al.*, 2011). Mastadenoviruses have mammalian species, including humans, as hosts, while aviadenoviruses exclusively infect birds. Atadenoviruses occur in squamate reptile hosts, but also in ruminants and birds (Benkő & Harrach, 1998). The genus *Ichtadenovirus* has only one confirmed member, white sturgeon adenovirus 1. Siadenoviruses possess the shortest genomes in the *Adenoviridae* family at around 26 kb and have been found in various vertebrate animals, including one amphibian host, Sulawesi turtles and birds of several species. The genus *Siadenovirus* is named after a genus-specific open reading frame, the putative gene product of which has a high sequence similarity to bacterial sialidase proteins (Davison *et al.*, 2000).

The adenovirus capsid consists of three major structural proteins: the hexon, the penton base and the fibre (San Martín, 2012). The majority of the capsid surface is accounted for by the 240 trimeric hexon capsomeres, while the 12 vertices are occupied by a pentameric penton base each. The trimeric fibre proteins protrude from the penton bases. Normally one fibre per vertex occurs and it is the principal structural determinant of adenovirus tropism (Nicklin *et al.*, 2005; Nouredini & Curiel, 2005). The fibre facilitates primary host cell binding through its globular distal tip or head domain. Some distinctive features present in adenoviruses, such as high transgene load capacity, recombinant growth up to high titre and efficient transduction ability, qualify them as candidates to be used as vectors for gene and cancer therapy, as well as vaccination agents (McConnell & Imperiale, 2004; Bachtarzi *et al.*, 2008; Arnberg, 2009; Thacker *et al.*, 2009). Considering the importance of fibre head domains for receptor interaction, structure determination of these proteins could not only shed light on the diverse tropism shown by these viruses, but may also enable the development of vectors based on a rational engineering and targeting approach (Nicklin *et al.*, 2005). However, no structural data concerning any siadenovirus fibre protein have been published to date.

Accepted species in the genus *Siadenovirus* are *Frog adenovirus A* (FrAdV-A) represented by frog adenovirus 1 (FrAdV-1), *Raptor adenovirus A* (RAV-A) and *Turkey adenovirus A* (TAdV-A), including turkey haemorrhagic enteritis virus (THEV), officially named turkey adenovirus 3 (TAdV-3; Harrach *et al.*, 2011). Virulent TAdV-3 strains may cause a range of diseases including haemorrhagic enteritis in turkeys and marble spleen disease in pheasants (Benkő,

1-MATPGKRS~~AE~~ EPDQQT~~LKKS~~ KQSDQS~~QGLN~~ LAYPFDKITE FEATPPFIHV- 50
51-GQGLDISDLS LNMRI~~GKGLK~~ FENG~~NLVVSD~~ QQYNVTPPLI ADQSTLGLKY-100
101-NPDVLSLTHS GALTLP~~TIQH~~ PLQASAGKFE LALSSgLKSD DQGLTL~~DLDP~~-150
151-VFSTESSKFL LNCSLPLDKN SDKLTLKFGN GLGLNNDQLE NTMTYNLPLK-200
201-RDgTNVSLSF GTNFKILNEM LDLNLVAPMS NSAGGLALQF KSPLSADDGI-250
251-LSIKTDTsLG ITGNKLGIRL APNSGLQITP NGLAVSVNAV QILSSPLITA-300
301-ASIGPPTTMV TGTVSPGRAT NGQFVTKTAK VLRYKFVRWD ALLIIQFIDN-350
351-IGVMENPTFY RNKSIELRSA DFLSPMLNNT YIVPLNGGVR VESPTIPVQL-400
401-EVILENNSSF IQVGFVRLTV KNGNPHMIIQ CNPVPGNIKM IKIKSVMLFT-450
451-CLIG

Figure 1

Predicted domain organization of turkey adenovirus 3 (TAdV-3). (a) The amino-acid sequence in one-letter code. The predicted virus-binding tail (residues 1–45) is shown in italics, the shaft domain (46–303) in normal type and the expressed predicted head domain (304–454) in bold. Putative triple β -spiral repeats in the shaft domain (van Raaij, Mitraki *et al.*, 1999) are alternately underlined or not. (b) Schematic drawing of the fibre indicating the tail, shaft and head domains. The start residue (1), the residues where the predicted shaft and head regions start (46 and 304, respectively) and the expected location of the final residue (454) are indicated (in known adenovirus fibre head structures, the final β -strand is usually folded back towards the amino-terminus).

2008). Both disease forms are often fatal. Serologically indistinguishable strains of TAdV-3 are known to have variable levels of pathogenicity. Some of the less virulent strains are being used for vaccination against turkey haemorrhagic enteritis (Beach *et al.*, 2009a). A study comparing the genome sequence of several virulent and avirulent TAdV-3 strains has shown differences in terms of mutations in the fibre sequence (Beach *et al.*, 2009b).

The 454-residue fibre protein of TAdV-3 contains a predicted tail domain (residues 1–45), which is likely to interact with the penton base, and a predicted shaft domain with putative triple β -spiral repeats between residues 46 and 303 (Fig. 1; van Raaij, Mitraki *et al.*, 1999). The remaining part (comprising amino acids 304–454) is likely to form the head domain. Two of the aforementioned mutations in the fibre sequence are located in the predicted head domain, namely residues 354 and 376, which are methionines in the avirulent form and isoleucine and threonine, respectively, in a virulent strain. Here, we describe the crystallization of the C-terminal end of the TAdV-3 fibre (avirulent vaccine strain) and report the data collected from these crystals.

2. Methods

2.1. Construction of expression vectors

The terminal part of the TAdV-3 fibre gene was obtained by PCR from a commercially available turkey haemorrhagic enteritis vaccine (H.E. Vac, ARKO Laboratories Ltd, Jewell, USA) containing attenuated live virus. From the rehydrated ready-to-use vaccine solution, a tenfold dilution was prepared. In the PCR, 1 μ l of this dilution was used as template without any further treatment or exact determination of the viral DNA concentration. The primer sequences

to amplify the carboxy-terminus of the fibre gene encoding amino acids 304–454 were taken from the genome sequence of the Virginia strain of TAdV-3 recently deposited in GenBank (entry AY849321; Beach *et al.*, 2009b). To facilitate the molecular cloning, the oligonucleotides (forward primer 5'-GCGAATTCATGGCTAC-TCTTGG-3'; reverse primer 5'-CGAAGCTTTCAGCCTATCA-AAC-3') were synthesized with an eight-nucleotide-long 5' extension to allow cleavage by the restriction enzymes *Eco*RI and *Hind*III, respectively (restriction sites are shown in bold). The PCR was carried out in a 50 μ l reaction mixture using Phusion High-Fidelity DNA polymerase enzyme (Thermo Scientific, Waltham, Massachusetts, USA) according to the manufacturer's instructions. The two-step PCR programme, consisting of an initial denaturation at 367 K for 5 min, followed by 35 cycles of 367 K for 30 s and 345 K for 120 s, was run in a Tpersonal thermocycler (Biometra, Göttingen, Germany). The final extension step lasted 5 min. The PCR product was purified by agarose gel electrophoresis using the MEGAquick-spin Total Fragment DNA Purification Kit (iNtRON Biotechnology, Kyungki-Do, Korea). The PCR product and the plasmid DNA [pET28a(+)] expression vector from Novagen, Merck, Darmstadt, Germany] were subjected to double digestion with *Eco*RI and *Hind*III restriction enzymes (Fermentas, Thermo Scientific, Waltham, Massachusetts, USA), ligated with T4 DNA ligase (Thermo Scientific) and transformed into chemically competent *Escherichia coli* TOP10 cells (Life Technologies Corporation, Carlsbad, California, USA). The start codon of the construct was supplied by the pET28a(+) expression vector, whereas the stop codon was present in the inserted fibre gene fragment. The foreign protein expressed in this vector [called pET28a(+)-THEVfib304-454] has an N-terminal tag (M GSSHH HHHHS SGLVP RGS HM ASMTG GQQMG RGSEF) for easy purification by nickel-affinity chromatography.

Plasmid preparations were screened by restriction analysis and the construction was checked by DNA sequencing.

2.2. Protein expression and purification

For native fibre head production, *E. coli* BL21(DE3) cells (Novagen, Merck, Darmstadt, Germany) were transformed with pET28a(+)-THEVfib304-454 and four 800 ml cultures were grown in Luria–Bertani medium aerobically at 310 K to an optical density of 0.6–0.8 measured at 600 nm. The temperature was reduced by cooling on ice and the cultures were induced with 1 mM isopropyl β -D-1-thiogalactopyranoside followed by overnight growth at 289 K. Cells were harvested by centrifugation at 5000g and 277 K for 15 min and resuspended in lysis buffer consisting of 50 mM Tris–HCl pH 7.5, 700 mM sodium chloride, 20 mM imidazole, 10% (v/v) glycerol. Lysis was performed using a French press cell disruptor (three passes). Insoluble material was removed by centrifugation (20 000g, 45 min), 1.5 ml of nickel–nitrilotriacetic acid (Ni–NTA) resin was added to the supernatant and the sample was left for incubation with gentle shaking for 1 h. The suspension was poured into an empty column and washed twice with ten column volumes of buffer consisting of 50 mM Tris–HCl pH 7.5, 700 mM sodium chloride, 30 mM imidazole, 10% (v/v) glycerol. This was followed by step gradient elution with

imidazole (50, 200, 300, 400 and 1000 mM) in the same buffer (each step was with two column volumes of the respective elution buffer). The fractions containing imidazole at between 300 and 1000 mM were pooled and dialysed three times against 500 ml 25 mM MES [2-(*N*-morpholino)ethanesulfonic acid]–NaOH pH 6.0. The dialysed sample was loaded onto a 6 ml Resource S column (GE Healthcare Bio-Sciences, Uppsala, Sweden) and eluted with a 0–1 M sodium chloride gradient in the same buffer. The protein eluted at around 0.3–0.4 M sodium chloride. Fractions containing pure protein were pooled and concentrated to 11 mg ml^{−1} using Amicon Ultra concentrators (Millipore Iberica, Madrid, Spain), incorporating three washes with 25 mM MES pH 6.0 to remove sodium chloride and stored temporarily at 277 K prior to use in crystallization trials.

To prepare selenomethionine-derivatized protein, pET28a(+)-THEVfib304-454 was transformed into *E. coli* B834(DE3) (Novagen, Merck, Darmstadt, Germany) and two 800 ml cultures were grown overnight at 289 K in SelenoMet medium (Molecular Dimensions, Newmarket, Suffolk, England) following the manufacturer's instructions. The cell disruption and protein purification steps were the same as for the native protein except for the addition of 7–10 mM β -mercaptoethanol to all buffers to prevent oxidation of selenomethionine residues. The selenomethionine-derivatized protein was concentrated to 4 mg ml^{−1} for crystallization in the same way as the native protein, but including β -mercaptoethanol.

2.3. Crystallization, data collection and data processing

To obtain suitable crystallization conditions, commercial screens from several suppliers were tested at 277 and 294 K either by manual or robotic setups. A total of about 500 conditions were tried and crystals were obtained in at least three of those. Condition No. 33 from the Emerald BioSystems (Seattle, Washington, USA) Wizard II

Figure 2

Purification, crystallization and diffraction analysis of TAdV-3 fibre head. (a) Chromatogram of the final strong anion-exchange purification step. The pure protein eluted in several peaks; protein from each peak was kept separate. Protein from the highest three peaks was used for crystallization trials. The sodium chloride concentration gradient is indicated in grey. (b) Crystals belonging to space group *I*23. The scale bar measures approximately 0.5 mm. (c) Diffraction pattern of TAdV-3 fibre head crystals. The beam centre is indicated by a cross-hair and resolution rings (in Å) are shown.

Table 1

Crystallographic data statistics.

Values in parentheses are for the highest resolution bin.

	Native	Selenomethionine derivative
Synchrotron and beamline	ESRF ID14-4	ESRF ID14-4
Wavelength (Å)	0.9768	0.9791
Crystal-to-detector distance (mm)	263.0	329.6
Oscillation range (°)	1.0	1.0
Number of images	100	360
Unit-cell parameter ($a = b = c$) (Å)	98.57	98.75
Mosaicity (°)	0.5–0.6	0.2
Resolution range (Å)	30–2.0 (2.11–2.00)	30–2.14 (2.26–2.14)
Reflections	10930 (1553)	8945 (1293)
Multiplicity	11.7 (12.0)	42.4 (37.4)
Completeness (%)	100.0 (100.0)	100.0 (100.0)
Mean $I/\sigma(I)$	13.7 (1.8)	27.4 (4.9)
$R_{\text{merge}}^{\dagger}$ (%)	10.5 (145.4)	10.7 (104.8)
V_M (Å ³ Da ^{−1}) [‡]	1.94	1.94
Wilson B factor (Å ²)	38.7	44.0

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, 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. [‡] According to Matthews (1968).

kit (1.0 *M* diammonium phosphate, 0.1 *M* sodium citrate pH 5.5) was chosen for optimization in sitting-drop vapour-diffusion crystallization experiments in CombiClover 4 Chamber plates (Jena Bioscience, Jena, Germany). Reservoir volumes were 0.15 ml and drops were made by mixing 1 µl each of protein and reservoir solution. Cube-shaped crystals of up to 0.1 mm in each dimension were obtained for both the native and selenomethionine-derivatized protein within 4 d of setup at 294 K when reservoir solutions consisting of 0.5–1.0 *M* diammonium phosphate, 0.1 *M* sodium citrate pH 5.5–5.8, 0.2–0.3 *M* sodium chloride were used (Fig. 2a).

For data collection, crystals were transferred to a solution containing the same components as the reservoir solution plus 25% (v/v) glycerol and soaked for 30 s. Crystals were then mounted in LithoLoops (Molecular Dimensions, Newmarket, Suffolk, England) or MicroMounts (Mitegen, Ithaca, New York, USA) and vitrified at 100 K. X-ray diffraction data were collected on beamline ID14-4 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France) at 100 K using an ADSC Q315r CCD detector. Crystallographic data were integrated using *iMOSFLM* (Battye *et al.*, 2011) and reduced using *POINTLESS*, *SCALA* and *TRUNCATE* (Evans, 2011), all integrated in *CCP4* (Winn *et al.*, 2011). Data-processing statistics are summarized in Table 1.

3. Results and discussion

Protein sequence analysis of the TAdV-3 fibre suggests that, like other adenovirus fibres, it contains a short, highly charged, penton base attachment sequence and a shaft domain containing putative triple β -spiral repeats (van Raaij, Mitraki *et al.*, 1999; Fig. 1). The remaining carboxy-terminal sequence (amino acids 304–454) may form a head domain, although it shows only between 10 and 20% sequence identity when aligned with known adenovirus head domain structures. The highest sequence identity, 31 identical residues out of 152, is with canine adenovirus 2 fibre head (Seiradake *et al.*, 2006). An expression vector was constructed for the putative fibre head domain, fused to an amino-terminal purification tag containing six consecutive histidine residues. The vector could be used to produce milligram amounts of soluble protein, which was purified by nickel-affinity chromatography followed by strong cation-exchange chromatography. The protein was eluted from the nickel column at relatively

Table 2

Data quality as a function of resolution.

Resolution (Å)	Reflections	R_{merge}	Mean $I/\sigma(I)$	CC lmean	CC anom
Native					
6.32	376	0.030	35.9	0.999	0.122
4.47	646	0.061	35.8	0.997	−0.053
3.65	840	0.080	33.0	0.996	−0.218
3.16	968	0.072	28.2	0.997	−0.159
2.83	1112	0.102	19.3	0.997	−0.020
2.58	1199	0.205	11.9	0.988	−0.088
2.39	1323	0.391	6.9	0.960	−0.050
2.24	1392	0.607	4.5	0.890	−0.027
2.11	1521	1.054	2.6	0.737	−0.019
2.00	1553	1.519	1.8	0.506	0.010
Selenomethionine					
6.78	313	0.049	57.6	1.000	0.914
4.80	531	0.086	54.3	0.999	0.577
3.92	681	0.091	56.8	0.999	0.432
3.39	802	0.072	50.3	1.000	0.706
3.03	894	0.094	42.0	0.999	0.740
2.77	994	0.151	29.4	0.998	0.580
2.56	1078	0.244	19.9	0.997	0.459
2.40	1146	0.428	12.5	0.990	0.235
2.26	1213	0.650	8.6	0.976	0.105
2.14	1293	1.063	4.9	0.930	0.031

high imidazole concentrations, consistent with it being a stable trimer as expected and thus containing three six-histidine tags per protein unit.

In cation-exchange chromatography, the protein (predicted pI 9.9) eluted in several peaks around 0.3–0.5 *M* sodium chloride when a linear gradient was used (Fig. 2a), each of which contained pure or nearly pure fibre head protein. Protein from the different peaks was kept separate and, after concentration, used for crystallization. Isomorphous crystals for protein from each of two major peaks were obtained in conditions containing diammonium phosphate as a precipitant, sodium chloride as an additive and sodium citrate at pH 5.5–5.8 as a buffer. It appears that the protein exists in several states differing in charge, but it is not clear to us what causes these differences. We do note that the same was observed for HAdV-2 fibre head expressed in baculovirus-infected insect cells (van Raaij, Louis *et al.*, 1999).

Crystals were obtained for the protein containing the purification tag (Fig. 2b). The crystals were cubic in shape and belonged to space group *I*23 (unit-cell parameter 99 Å). They are expected to contain one protein monomer per asymmetric unit, leading to a solvent content of 37% (Table 1). The biologically relevant trimer is generated by the crystallographic threefold symmetry axis. Complete and highly redundant X-ray diffraction data sets were collected from a native crystal and a crystal derivatized with selenomethionine, the latter using a wavelength at which the anomalous signal of the Se atoms was maximized (Table 1; Fig. 2). The data used extended to 2.0 Å for the native data set (cutoff imposed by us) and 2.14 Å resolution for the derivative crystal (cutoff imposed by detector edge). Merging R factors for shells higher than 2.4 Å resolution were quite high; however, the correlation between independent measurements of the same reflection and mean $I/\sigma(I)$ in the highest resolution bins are still significant (Table 2; Evans & Murshudov, 2013). Structure solution and refinement will be performed at different resolutions to decide to which resolution data should be included. There was no clear difference in crystal or data quality between the native protein and the selenomethionine derivative.

4. Conclusion

We collected high-quality data from crystals of the fibre head of the avirulent variant of TAdV-3, both native and derivatized with sele-

nomethionine. This is the first siadenovirus for which crystals of the fibre head have been obtained and for which data have been collected. The structure should reveal similarities and differences between siadenoviral fibre heads and those of other adenovirus genera. Knowledge of the structure may also allow the design of chimeric adenoviruses incorporating the TAdV-3 fibre head to investigate their tropism, which in turn may lead to vaccination or gene-therapy vectors that may target specific cell types.

We thank Andrew McCarthy for valuable help with data collection at ESRF beamline ID14-EH4. This research was sponsored by grant No. BFU2011-24843 from the Spanish Ministry of Economy and Competitiveness and by grant OTKA K100163 from the Hungarian Scientific Research Fund. AKS was a recipient of a pre-doctoral fellowship from La Caixa.

References

- Arnberg, N. (2009). *Rev. Med. Virol.* **19**, 165–178.
- Bachtarzi, H., Stevenson, M. & Fisher, K. (2008). *Expert Opin. Drug Deliv.* **5**, 1231–1240.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Beach, N. M., Duncan, R. B., Larsen, C. T., Meng, X., Sriranganathan, N. & Pierson, F. W. (2009a). *Avian Dis.* **53**, 370–375.
- Beach, N. M., Duncan, R. B., Larsen, C. T., Meng, X., Sriranganathan, N. & Pierson, F. W. (2009b). *J. Gen. Virol.* **90**, 1978–1985.
- Benkő, M. (2008). *Encyclopedia of Virology*, edited by B. W. J. Mahy & M. H. V. van Regenmortel, pp. 24–29. Waltham: Academic Press.
- Benkő, M. & Harrach, B. (1998). *Arch. Virol.* **143**, 829–837.
- Davison, A. J., Wright, K. M. & Harrach, B. (2000). *J. Gen. Virol.* **81**, 2431–2439.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.
- Harrach, B., Benkő, M., Both, G. W., Brown, M., Davison, A. J., Echavarria, 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. King, M. J. Adams, E. B. Carstens & E. J. Lefkowitz, pp. 125–141. San Diego: Elsevier.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McConnell, M. J. & Imperiale, M. J. (2004). *Hum. Gene Ther.* **15**, 1022–1033.
- Nicklin, S. A., Wu, E., Nemerow, G. R. & Baker, A. H. (2005). *Mol. Ther.* **12**, 384–393.
- Noureddini, S. C. & Curiel, D. T. (2005). *Mol. Pharm.* **2**, 341–347.
- 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.
- San Martín, C. (2012). *Viruses*, **4**, 847–877.
- Seiradake, E., Lortat-Jacob, H., Billet, O., Kremer, E. J. & Cusack, S. (2006). *J. Biol. Chem.* **281**, 33704–33716.
- Thacker, E. E., Timares, L. & Matthews, Q. L. (2009). *Expert Rev. Vaccines*, **8**, 761–777.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.