Crystallographic structure of the α-helical triple coiled-coil domain of avian reovirus S1133 fibre

Pablo Guardado-Calvo,1 Gavin C. Fox,2 Antonio L. Llamas-Saiz3 and Mark J. van Raaij1,4

1Departamento de Bioquímica e Biología Molecular, Facultade de Farmacia, Universidade de Santiago de Compostela, E-15782 Santiago de Compostela, Spain
2Spanish CRG beamline BM16, European Synchrotron Radiation Facility, 6 rue Jules Horowitz, F-38043 Grenoble, France
3Unidade de Raios X, Laboratorio Integral de Dinâmica e Estructura de Biomoléculas José R. Carracido, Edificio CACTUS, Campus Sur, Universidade de Santiago de Compostela, E-15782 Santiago de Compostela, Spain
4Instituto de Biología Molecular de Barcelona-CSIC, c/Josep-Samitier 1-5, E-08028 Barcelona, Spain

Avian reovirus fibre, a homo-trimer of the σC protein, is a minor component of the avian reovirus outer capsid. It is anchored via a short N-terminal sequence to the inner capsid λC pentamer, and its protruding globular C-terminal domain is responsible for primary host cell attachment. We have previously solved the structure of a receptor-binding fragment in which residues 160–191 form a triple β-spiral and 196–326 a β-barrel head domain. Here we have expressed, purified and crystallized a major σC fragment comprising residues 117–326. Its structure, which was solved by molecular replacement using the previously determined receptor-binding domain structure and refined to 1.75 Å (0.175 nm) resolution, reveals an α-helical triple coiled-coil connected to the previously solved structure by a zinc-ion-containing linker. The coiled-coil domain contains two chloride ion binding sites, as well as specific trimerization and registration sequences. The linker may act as a functionally important hinge.

Avian reovirus, a member of the Orthoreovirus genus of the Reoviridae family, is an important poultry pathogen (Songserm et al., 2003; Benavente & Martinez-Costas, 2007). This non-enveloped double-stranded RNA virus is composed of a double concentric icosahedral capsid (Spandidos & Graham, 1997), enclosing ten double-stranded RNA genome segments and several copies of the RNA polymerase complex (Spandidos & Graham, 1976), enclosing ten double-stranded RNA segments and several copies of the RNA polymerase complex (Spandidos & Graham, 1976), enclosing ten double-stranded RNA segments. This non-enveloped double-stranded RNA virus is composed of a double concentric icosahedral capsid (Spandidos & Graham, 1997), enclosing ten double-stranded RNA genome segments and several copies of the RNA polymerase complex (Spandidos & Graham, 1976), enclosing ten double-stranded RNA genome segments. Its structure revealed two triple β-spiral repeats (amino acids 160–191) and an eight-stranded circular β-barrel head domain (Guardado-Calvo et al., 2005). Here we report the crystallization and structure solution of an extended recombinant avian reovirus S1133 σC fragment containing amino acids 117–326, refined against data collected to 1.75 Å (0.175 nm) resolution. Our results show the shaft domain has a mixed α-helical coiled-coil and triple β-spiral domain, separated by a flexible linker containing a divalent zinc cation.

For expression, Escherichia coli strain JM109 (DE3) transformed with the plasmid pET28-sigmaC117-326 was grown as described (van Raaij et al., 2005). Cells harvested from 4 l of culture were resuspended in 40 ml cold resuspension buffer (10 mM Tris/HCl 8.0, 300 mM sodium chloride), frozen at −20 °C and lysed by a double pass through an Avestin C5 emulsifier (Avestin). After removing insoluble material, 3 ml Ni-nitriloacetic acid resin (Qiagen) was added. The suspension was incubated for 1 h on ice and poured into an empty column. The resin...
was washed with resuspension buffer; elution was performed with a step gradient of imidazole pH 7 in resuspension buffer. His-T7-tagged σC117–326 eluted at imidazole concentrations of 200–1000 mM, was dialysed overnight against TE buffer (10 mM Tris/HCl pH 8.5, 1 mM EDTA pH 8.0) and incubated with 0.2 μg ml⁻¹ trypsin for 45 min at 37 °C. PMSF was added to 1 mM to stop the reaction and the protein was applied onto a 6 ml Uno-Q column (Bio-Rad). σC117–326 eluted at the start of a linear 0–1 M sodium chloride gradient in TE buffer. The protein was concentrated to 11.5 mg ml⁻¹ using Centricron concentrators (Millipore).

Crystallization took place by vapour diffusion in sitting-drop plates with 0.8 ml reservoirs and drops of 5 μl protein solution mixed with 5 μl reservoir solution. Bar-shaped crystals with hexagonal cross-sections belonging to the trigonal space group P321 appeared after 1 day in solutions equilibrated against 0.6–0.75 M ammonium sulphate, 0.1 M Tris/HCl pH 8.4, 25 % glycerol and 50 mM zinc sulphate. Rectangular prism with parallelogram base-shaped crystals of the monoclinic space group C2 were harvested from a solution containing P321 crystals, which had been opened, left open for a while and had then been reclosed, thus partially drying out. Crystals were mounted in cryo-loops and kept at 100 K during data collection. Complete datasets were collected from a crystal of each crystal form, to 2.3 Å (0.23 nm) for the P321 crystal form and to 1.75 Å (0.175 nm) for the C2 crystal form; see Table 1. Reflections were integrated and scaled with the program HKL2000 (Otwinowski & Minor, 1997) and further processed using programs from the Collaborative Computational Project Number 4 (1994).

The structure of the C2 crystal form was solved by molecular replacement using protein structure database entry 2BSF (Guardado-Calvo et al., 2005) as a search model in the program MOLEREP (Vagin & Teplyakov, 2000). Three monomers were located in the asymmetric unit, forming the biological trimer. Automatic model building using ARP-WARP (Morris et al., 2003) combined with manual adjustment using the programs O (Jones et al., 1991) and COOT (Emsley & Cowtan, 2004) led to a final model containing 620 residues (amino acids 117–326 for chains A and B; 117–326 for chain C) and the solvent molecules specified in Table 1. Refinement was done using the REFMAC program (Murshudov et al., 1997); reflections selected in thin resolution shells were set apart for calculation of R-free (Table 1). Water molecules were built using ARP (Lamzin & Wilson, 1993). Model validation was performed with WHATCHECK (FRIEND, 1990), MolProbity (Davis et al., 2007) and MOLEMAN (Kleywegt et al., 2001). The most complete chain of this model, chain C, was used to solve the structure of the P321 crystal form by molecular replacement. The coordinates have been deposited in the protein structure database (http://www.rcsb.org) under accession code 2VRS for the C2 crystal form and 2JJL for the P321 crystal form; structure factors are also available. In the P321 crystal, the asymmetrical unit consists of a single monomer and the biological trimer is formed by crystallographic symmetry. When the P321 monomer is superimposed on monomers A, B and C of the C2 trimer, root mean square differences (r.m.s.d.) of only 0.4–0.6 Å are observed. If instead the trimeric biological units are superimposed, they superimpose equally well, with an r.m.s.d. of under 0.7 Å, revealing no significant differences. Therefore only the C2 structure will be discussed further. The refined structure shows good correspondence to the data, has good geometry and no residues in very unfavourable regions (>99.8 %) of the Ramachandran plot (Table 1).

Electron micrographs revealed σC115–326 to be a globular protein containing an 80 Å long stalk and 50 Å wide at the widest point of the head domain; while structure prediction with the program COILS, a window size of 28 residues and taking 0.15 as probability cut-off, suggests amino acids 53–106 and 126–154 are in a coiled-coil conformation (Costas, 2004; Lupas et al., 1991). The crystal structure of σC117–326 fragment, with a stalk 90 Å long and 55 Å wide at the head domain, largely substantiates these data (Fig. 1). When the overall structure is contemplated, a clear division between shaft (amino acids 117–191) and head domains (residues 196–326) is observed. The shaft domain can be further subdivided (Fig. 1b) into an α-helical triple-coiled-coil (amino acids 117–154), a linker region (residues 155–159) and two repeats of a triple β-spiral (Mitra et al., 2002). The head and triple β-spiral domains have been described before (Guardado-Calvo et al., 2005).

The α-helical region is an uninterrupted coiled-coil structure 53 Å (5.3 nm) long (Fig. 1b). It is composed of 5.5 heptad repeats, with a pitch length of 166 Å (16.6 nm), a coiled-coil radius of 6.2 Å (0.62 nm) and an average number of 3.6 residues per turn (Strelkov & Burkhard, 2002), which is comparable with average values for trimeric coiled-coils (Fraser & MacRae, 1973; Tao et al., 1997). Analysis of buried residues within the coiled-coil, at positions d and a in the heptad repeat (Leu152, Val149, Leu145, Ile142, Asn138, Ile135, Leu131, Ile128, Asn124, Val121 and Leu117) reveals all five residues at position a, four are valines and three isoleucines, while the six residues at the d position are not (four leucines and two asparagines). It has been proposed that coiled-coil structures with β-branched residues at position a and leucines at position d are more stable as dimers rather than trimers (Harbury et al., 1993), but the parallel trimeric coiled-coils of fibrin (Tao et al., 1997) and human mannose-binding protein (Sheriff et al., 1994) also contain many β-branched residues at position a and leucines at position d. In the other section of predicted coiled-coil region (residues 53–106), of the seven residues at position a, four are valines and three leucines; while of the eight residues at position d, there are four leucines, two isoleucines, a valine and a methionine, i.e. both β-branched amino acids and leucines at positions a and d.
The structure reveals a trimerization motif (Meier et al., 2002; Kammerer et al., 2005), composed of a network of interhelical salt bridges formed by contacts between Arg148 at position g of one chain and Glu153 at position e of the adjacent chain (Fig. 2a). The potential of the trimerization motif to dominate the effect of the hydrophobic core residues has been already demonstrated in short coiled-coils (Burkhard et al., 2000; Kammerer et al., 2004).

Interestingly, analysis of the sequence reveals another putative trimerization motif containing Arg36 and Glu41, suggesting the triple-helical coiled-coil extends to the N terminus further than the predicted residue 53. In addition to hydrophobic interactions occurring in the coiled-coil core and polar interactions at the surface, the structure reveals two asparagines situated at position d (Asn124 and Asn138), stabilized by central chloride ions (Fig. 2b). The electron density features were identified as chlorides based on crystallographic evidence (refinement as chlorides leads to good agreement with the crystallographic data and temperature factors very similar to the surrounding residues), suitability of the coordination distances and similarity to other viral coiled-coils (see below). On the other side of the chloride ions are the hydrophobic side chains of Val121 and Ile135, respectively. It has been postulated that these central polar interactions favour correct ‘registration’ (Oakley & Kim, 1998); asparagines at position d also favour the trimeric state (Tripet et al., 2000). A central chloride ion coordinated by asparagine residues appears to be a quite common feature in parallel trimeric coiled-coils of viral fusion proteins, fibres and bacterial adhesins (see for example Fass et al., 1996; Tao et al., 1997; Malashkevich et al., 1999; Renard et al., 2005; Table 1.

### Table 1. Data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>C2</th>
<th>P21</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cell parameters</strong></td>
<td>a=142.8 Å, b=54.0 Å, c=121.3 Å, β=101.8°</td>
<td>a=b=77.7 Å, c=121.4 Å</td>
</tr>
<tr>
<td><strong>Beamline</strong></td>
<td>Spanish beamline BM16 (European Synchrotron Radiation Facility)</td>
<td>Spanish beamline BM16 (European Synchrotron Radiation Facility)</td>
</tr>
<tr>
<td><strong>Wavelength</strong></td>
<td>0.8266 Å</td>
<td>0.9840 Å</td>
</tr>
<tr>
<td><strong>Detector</strong></td>
<td>ADSC Q210r CCD</td>
<td>ADSC Q210r CCD</td>
</tr>
<tr>
<td><strong>Crystal-to-detector distance</strong></td>
<td>178.5 mm</td>
<td>178.0 mm</td>
</tr>
<tr>
<td><strong>Observed reflections‡</strong></td>
<td>90954 (9036)</td>
<td>19318 (1858)</td>
</tr>
<tr>
<td><strong>Resolution range</strong></td>
<td>20.00–1.75 Å (1.81–1.75 Å)</td>
<td>50.00–2.30 Å (2.38–2.30 Å)</td>
</tr>
<tr>
<td><strong>Wilson B</strong></td>
<td>21.7 Å²</td>
<td>44.1 Å²</td>
</tr>
<tr>
<td><strong>Multiplicity</strong></td>
<td>3.3 (3.2)</td>
<td>20.0 (20.3)</td>
</tr>
<tr>
<td><strong>Completeness</strong></td>
<td>0.992 (0.992)</td>
<td>0.990 (0.987)</td>
</tr>
<tr>
<td><strong>Rmerge</strong></td>
<td>0.058 (0.384)</td>
<td>0.068 (0.428)</td>
</tr>
<tr>
<td><strong>Refinement†</strong></td>
<td>89213 (15852)</td>
<td>17714 (3116)</td>
</tr>
<tr>
<td><strong>Reflections used for R-free</strong></td>
<td>1738 (329)</td>
<td>1514 (268)</td>
</tr>
<tr>
<td><strong>Resolution range</strong></td>
<td>18.43–1.75 Å (1.87–1.75 Å)</td>
<td>25.00–2.30 Å (2.46–2.30 Å)</td>
</tr>
<tr>
<td><strong>R-factor‡</strong></td>
<td>0.180 (0.234)</td>
<td>0.181 (0.216)</td>
</tr>
<tr>
<td><strong>R-free§</strong></td>
<td>0.218 (0.278)</td>
<td>0.220 (0.301)</td>
</tr>
<tr>
<td><strong>r.m.s.d. (bonds / angles)</strong></td>
<td>0.017 Å / 1.6</td>
<td>0.014 Å / 1.5</td>
</tr>
<tr>
<td><strong>Final model</strong></td>
<td>4649 / 2 / 7 / 5 / 18 / 975 / glycerol / water atoms</td>
<td>1551 / 2 / 3 / 5 / 0 / 202</td>
</tr>
<tr>
<td><strong>Average B protein / Cl⁻ / Zn²⁺ / SO₄²⁻ / glycerol / water</strong></td>
<td>24.6 Å² / 26.8 Å² / 28.7 Å² / 45.7 Å² / glycerol / water</td>
<td>37.5 Å² / 49.1 Å² / 37.8 Å² / 69.1 Å² / glycerol / water</td>
</tr>
<tr>
<td><strong>Ramachandran statistics</strong></td>
<td>97.7% / 100.0%</td>
<td>97.1% / 100.0%</td>
</tr>
</tbody>
</table>

*Results from SCALEPACK (Otwinowski & Minor, 1997).
†One Ångstrom (Å) is 0.1 nm.
‡No σ cut-off or other restrictions were used for inclusion of reflections.
§Values in parentheses are for the highest resolution bin, where applicable.
††Results from REFMAC (Murshudov et al., 1997).
¶R-factor = Σ|Fobs(hkl)| – Fcalc(hkl)| / Σ|Fobs(hkl)|.
#According to Brunger (1997).
**According to the program MolProbity (Davis et al., 2007). The percentages indicated are for residues in favoured and total allowed regions, respectively.

The structure reveals a trimerization motif (Meier et al., 2002; Kammerer et al., 2005), composed of a network of interhelical salt bridges formed by contacts between Arg148 at position g of one chain and Glu153 at position e of the adjacent chain (Fig. 2a). The potential of the trimerization motif to dominate the effect of the hydrophobic core residues has been already demonstrated in short coiled-coils (Burkhard et al., 2000; Kammerer et al., 2004). Interestingly, analysis of the sequence reveals another putative trimerization motif containing Arg36 and Glu41, suggesting the triple-helical coiled-coil extends to the N terminus further than the predicted residue 53. In addition to hydrophobic interactions occurring in the coiled-coil core and polar interactions at the surface, the structure reveals two asparagines situated at position d (Asn124 and Asn138), stabilized by central chloride ions (Fig. 2b). The electron density features were identified as chlorides based on crystallographic evidence (refinement as chlorides leads to good agreement with the crystallographic data and temperature factors very similar to the surrounding residues), suitability of the coordination distances and similarity to other viral coiled-coils (see below). On the other side of the chloride ions are the hydrophobic side chains of Val121 and Ile135, respectively. It has been postulated that these central polar interactions favour correct ‘registration’ (Oakley & Kim, 1998); asparagines at position d also favour the trimeric state (Tripet et al., 2000). A central chloride ion coordinated by asparagine residues appears to be a quite common feature in parallel trimeric coiled-coils of viral fusion proteins, fibres and bacterial adhesins (see for example Fass et al., 1996; Tao et al., 1997; Malashkevich et al., 1999; Renard et al., 2005;
Conners et al., 2008), and it has been used in de novo design of trimeric coiled-coils (Lumb & Kim, 1995). As discussed before (Guardado-Calvo et al., 2005), C folding may begin at its monomeric C-terminal β-barrel. Interaction among three β-barrels could then lead to trimer formation. The stalk region would then ‘zip up’, starting with the short triple β-spiral, to form the intact fibre. The head and/or stalk domains could act as signals to ensure correct registration. Alternatively, trimerization could start at the registration domains in the coiled-coil domain identified here.

Amino acids 155–159 link the α-helical coiled-coil to the triple β-spiral (Fig. 2c). Thr155 is in an extended conformation, with residues 155–159 forming two nested β-turns (Ala156 and Ser157 form a type I, His158 and Gly159 a type II β-turn; Richardson, 1981). The three His158 bind a divalent zinc cation through their NE2 atoms (the other six zinc ions in the structural model are on the protein surface and are not likely to have a biological role). The fourth ligand in the distorted-tetragonal coordination environment is a water molecule. Zinc often contributes to catalytic sites and a coordination sphere similar to the one observed here is present in carbonic anhydrase (Auld, 2005). However, as zinc was added in the crystallization buffer, we cannot be sure that σC contains zinc in its natural state. However, it is tempting to speculate that zinc, or perhaps another divalent metal cation, has a structural role in stabilizing the extended σC conformation in the intermediate subviral particle obtained upon reovirion partial uncoating, when outer capsid proteins σB and μB are lost (Benavente & Martinez-Costas, 2007). If σC is folded away in the intact reovirion (as its analogue σ1 in mammalian reovirus is proposed to be; Guglielmi et al., 2006), its zinc-binding motif may be disordered and function as a hinge region. The presence of the metal ion may be dependent on the protonation state of its histidine residues, and thus on pH. A similar hinge region may be present around His110, which is near to its trypsin cleavage site (see above). This would be consistent with the COILS prediction, where there is a gap in likely coiled-coil formation from amino acid 106 onwards (see above). Residues up to around amino acid 30 probably interact with λC at the pentameric vertices of the avian reovirus core, and are likely to be unstructured in the isolated protein. It is currently not clear how σC incorporates into the virus; apart from a small stub of density emerging from the λC pentamer identified by electron microscopy, no electron density has been assigned

Fig. 1. Overall structure of avian reovirus S1133 fibre σC117–326. (a) Stereoview of C-αs connected by lines. The monomers are coloured blue, yellow and red, and some of the residues of the yellow monomer are numbered. Green balls, chloride anions; grey ball, zinc cation. (b) Ribbon diagram; secondary structure is shown as corkscrews (α-helices) or arrows (β-strands). The boundaries between the α-helical, hinge and β triple spiral-domains of the shaft domain and head domain are indicated. Figures were prepared using PyMOL (DeLano Scientific; http://www.pymol.org).
to σC (Zhang et al., 2005). This may be due to σC adopting different possible conformations and/or to σC not following the icosahedral symmetry imposed in the electron microscopy reconstruction.

In summary, we have for the first time experimentally determined the structure of the α-helical coiled-coil region of a reovirus fibre and shown that the linker between the α-helical and β-structured parts of the σC protein may bind a divalent zinc cation. Our structural data suggest the coiled-coil is important for in-register trimerization of the avian reovirus fibre and suggest the presence of hinge regions around residues 110 and 158, which may be important for receptor-binding or subsequent approach of the infectious viral particle to the cell. They may also be important for the accommodation of the σC trimer in the virus particle.

Acknowledgements

We thank Patricia Ferraces-Casais and Rebeca Menaya-Vargas for technical assistance, Laurent Terradot-Piot and the Partnership for Structural Biology (Grenoble, France) for use of crystallization facilities and Javier Benavente, José Martínez-Costas and David Auld for discussions. This research was sponsored by research grants BFU2005-02974 and BFU2005-24982-E and by a pre-doctoral FPU fellowship to P.G.-C. from the Spanish Ministry of Education and Science. This work was also supported by funds from the European Commission under contracts ERAS-CT-2003-980409 (as part of the European Science Foundation EUROCORES Programme EuroSCOPE) and NMP4-CT-2006-033256 (BeNatural coordinated project).

References


