

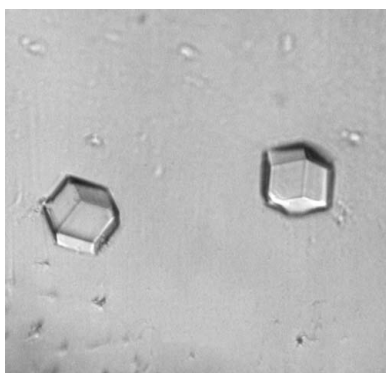
**Stéfano Ricagno,‡ Bruno Coutard,‡ Sacha Grisel, Nicolas Brémont, Karen Dalle, Fabienne Tocque, Valérie Campanacci, Julie Lichère, Violaine Lantez, Claire Debarnot, Christian Cambillau, Bruno Canard and Marie-Pierre Egloff\***

Centre National de la Recherche Scientifique and Universités d'Aix-Marseille I et II, UMR 6098, Architecture et Fonction des Macromolécules Biologiques, Ecole Supérieure d'Ingénieurs de Luminy-Case 925, 163 Avenue de Luminy, 13288 Marseille CEDEX 9, France

‡ These authors made equal contributions to this work.

Correspondence e-mail: marie-pierre.egloff@afmb.univ-mrs.fr

Received 17 January 2006  
Accepted 13 March 2006



© 2006 International Union of Crystallography  
All rights reserved

## Crystallization and preliminary X-ray diffraction analysis of Nsp15 from SARS coronavirus

The non-structural protein Nsp15 from the aetiological agent of SARS (severe acute respiratory syndrome) has recently been characterized as a uridine-specific endoribonuclease. This enzyme plays an essential role in viral replication and transcription since a mutation in the related H229E human coronavirus nsp15 gene can abolish viral RNA synthesis. SARS full-length Nsp15 (346 amino acids) has been cloned and expressed in *Escherichia coli* with an N-terminal hexahistidine tag and has been purified to homogeneity. The protein was subsequently crystallized using PEG 8000 or 10 000 as precipitants. Small cubic crystals of the apoenzyme were obtained from 100 nl nanodrops. They belong to space group  $P4_132$  or  $P4_332$ , with unit-cell parameters  $a = b = c = 166.8 \text{ \AA}$ . Diffraction data were collected to a maximum resolution of  $2.7 \text{ \AA}$ .

### 1. Introduction

An outbreak of severe acute respiratory syndrome (SARS) initiated suddenly in China in 2003 and rapidly spread worldwide, causing more than 8000 fatalities. The aetiological agent of SARS was characterized as a coronavirus named SARS-CoV. This virus has a  $\sim 30$  kbp long single-strand RNA genome with a 5' cap and 3' poly(A) tail (Rota *et al.*, 2003).

The genome of SARS-CoV codes for two large polyproteins, 1a and 1ab, comprising the replicase components. These polypeptides are autoproteolytically processed by two viral proteases, yielding 16 non-structural proteins (Nsps) which, together with a number of processing intermediates and host proteins, are thought to form a huge protein complex responsible for viral replication and transcription. Despite the growing interest in coronaviruses, the biochemical and structural characterization of the non-structural proteins still lags behind. In the last two years, despite tremendous efforts worldwide, only three X-ray structures of Nsps were reported: those of Nsp5, the main protease, of Nsp9, a novel RNA-binding protein, and of the complex Nsp7–Nsp8 (Anand *et al.*, 2003; Egloff *et al.*, 2004; Sutton *et al.*, 2004; Zhai *et al.*, 2005).

A number of unusual activities of coronavirus Nsps have been identified that differ from those found in other RNA viruses and Nsp15 is a striking example. Comparative genomics has made it possible to identify similarity between Nsp15 and only one other protein, XendoU, an endoribonuclease from *Xenopus laevis* studied by Caffarelli *et al.* (1997). XendoU is involved in the processing of U16 small nucleolar RNA from its host pre-mRNA. It is poly(U)-specific and requires  $\text{Mn}^{2+}$  as a cofactor (Caffarelli *et al.*, 1997; Laneve *et al.*, 2003). It has been suggested, therefore, that Nsp15 might be an endoribonuclease (Snijder *et al.*, 2003), which is unique to nidoviruses within the RNA viral world.

The biochemical characterization of SARS Nsp15 has recently been reported and its endoribonucleolytic activity as well as its specificity have been determined (Bhardwaj *et al.*, 2004; Ivanov *et al.*, 2004). The main activity of Nsp15 is to cleave 5' to uridylates in double-stranded RNA, which requires the presence of divalent cations. Like other endoribonucleases, Nsp15 uses  $\text{Mn}^{2+}$  as cofactor; in the presence of  $\text{Mg}^{2+}$  or other metal cations lower or no activity is detected. Nsp15 has been shown to be highly specific for RNA. DNA

is not cleaved and does not inhibit RNA digestion, thus indicating that it does not bind to Nsp15 (Bhardwaj *et al.*, 2004). Single-stranded RNA is also cleaved, but less specifically and efficiently than double-stranded RNA (Ivanov *et al.*, 2004). Electron microscopy as well as other biophysical techniques have revealed that Nsp15 monomers form hexameric structures (Guarino *et al.*, 2005).

RNases are divided into two classes (Saida *et al.*, 2003). The first group of enzymes is characterized by the presence of a metal ion in the active site. In the second group, the reaction is performed using a histidine as a general base and a histidine or a glutamic acid as a general acid (Irie *et al.*, 1997; Nishikawa *et al.*, 1987; Steyaert *et al.*, 1990; Thompson *et al.*, 1994). The reaction generates a product carrying a cyclic 2'-3' phosphodiester terminus. Recent work (Gioia *et al.*, 2005) has suggested that XendoU and the viral endoribonucleases related to it do not belong to either of the previously described groups. Instead, this work reveals the existence of a new class of RNA-processing enzymes that require a metal ion for activity (mainly  $Mn^{2+}$ ), produce cyclic 2'-3' phosphodiester termini and have two histidine residues that are essential for RNase activity.

The biochemical novelty of Nsp15 and the great medical importance of SARS-CoV prompted us to initiate the crystallographic study of this enzyme. Full-length Nsp15 consists of a polypeptide chain of 346 amino acids and has a molecular weight of 38.5 kDa. In the present communication, we report the cloning, expression, purification and crystallization of Nsp15 from SARS-CoV and preliminary diffraction experiments on Nsp15 crystals.

## 2. Expression and purification

The SARS-CoV (Frankfurt-1 strain; NCBI accession No. AY291315) Nsp15 full coding sequence (amino acids 6430–6775 of polyprotein 1ab) was amplified by PCR from cDNA kindly provided by J. Dobbe and E. J. Snijder (Drosten *et al.*, 2003) using two primers containing the attB sites of the Gateway recombination system (Invitrogen). A sequence encoding a hexahistidine tag was added at the 5' end of the gene. The cDNA was cloned in the pDest14 plasmid according to Gateway cloning technology (Invitrogen). Multiparameter expression screening in *Escherichia coli* was performed using a sparse matrix (Abergel *et al.*, 2003), followed by the dot blot detection procedure (Vincentelli *et al.*, 2005). The optimized expression condition used *E. coli* strain C41(DE3) (Avidis) transformed with the pLysS plasmid (Novagen) and grown in Turbo Broth media (Athena Enzymes) at 298 K. Overnight induction was performed with 0.5 mM IPTG when  $OD_{600}$  reached 0.5. Expression yielded about 1.5 mg of soluble Nsp15 per litre of culture. The cells were harvested by centrifugation and resuspended in 10 ml of buffer A (50 mM Tris pH 8, 150 mM NaCl, 10 mM imidazole, 1 mM PMSF, 0.2 mg ml<sup>-1</sup> lysozyme, 0.1 µg ml<sup>-1</sup> DNase, 20 mM MgSO<sub>4</sub> and protease-cocktail inhibitor; Sigma) per  $OD_{600}$  unit and per litre of culture. Cells were lysed by sonication and centrifuged (20 000g) for 40 min at 277 K. The purification consisted of two steps. Firstly, IMAC (immobilized metal-ion affinity chromatography) was performed on a Hisprep column (Amersham). The protein was eluted with 50 mM Tris-HCl, 150 mM NaCl, 250 mM imidazole pH 8.0. Size-exclusion chromatography was performed using a Superdex 200 HiLoad 26/60 (Amersham) in 10 mM Tris, 50 mM NaCl pH 7.5. Nsp15 was eluted with a retention volume corresponding to a molecular weight of about 250 kDa, corresponding to a hexameric association, as also observed by Guarino *et al.* (2005). Protein purity was analyzed by SDS-PAGE. The recombinant protein was concentrated using Amicon Ultra (5000 kDa molecular-weight cutoff).

## 3. Crystallization

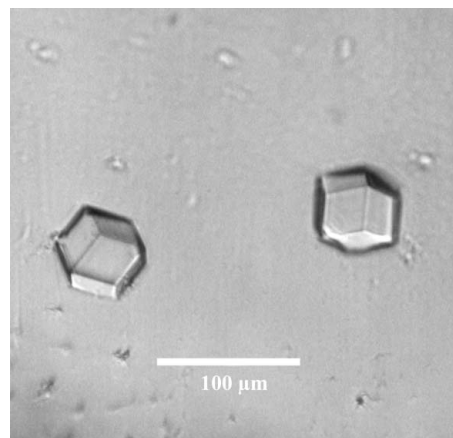
The protein was unstable in the buffer used for size-exclusion chromatography and could not be concentrated beyond 2–3.5 mg ml<sup>-1</sup>. Moreover, the protein had to be crystallized soon after purification to prevent precipitation. Microdialysis experiments indicated that the protein was more stable (i) at pH values between 6.5 and 9 and (ii) when the sodium chloride concentration was higher than 300 mM. However, crystallization failed under these conditions and the sodium chloride concentration had to be decreased, the final protein buffer being 50 mM NaCl, 10 mM Tris pH 7.5.

Crystallization trials were performed using the sitting-drop technique at 291 K using a nanodrop-dispensing robot (Cartesian Inc.). Wizard Screens 1 and 2 (Emerald Biostructures), Structure Screens 1 and 2 and Stura Footprint Screen (Molecular Dimensions Ltd) were used for initial screening. Drops of 300, 200 and 100 nl protein solution (1.0–3.5 mg ml<sup>-1</sup>) were mixed with 100 nl mother liquor (the reservoir volume was 150 µl). Three conditions from the Stura Footprint Screen and one from the Wizard Screens yielded crystal hits. After optimization, two of these conditions led to small cubic crystals [10–20% (w/v) polyethylene glycol 10 000, 0.2 M imidazole malate pH 8.0–9.0 and 15–25% (w/v) polyethylene glycol 8000, 0.1 M CHES pH 9.0–10.0]. The crystals grew in 2–4 d to maximum dimensions of 0.05 × 0.05 × 0.05 mm (Fig. 1).

## 4. Preliminary crystallographic analysis

The crystals of Nsp15 were flash-frozen in liquid nitrogen at 110 K. Glycerol was added to the mother liquor as a cryoprotectant to a final concentration of 33%. Diffraction data for the native enzyme were collected at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) beamline ID14-3 using a MAR 165 CCD detector. The data were processed using *MOSFLM/SCALA* (Collaborative Computational Project, Number 4, 1994).

Nsp15 crystals diffracted X-rays to a resolution limit of 2.7 Å. Crystals presented cubic symmetry, belonging to one of the two enantiomorphic space groups  $P4_132$  or  $P4_332$ , with unit-cell parameters  $a = b = c = 166.8$  Å (Table 1). Calculation of the Matthews coefficient indicated that the asymmetric unit may contain either one or two molecules of Nsp15, which would correspond to a calculated solvent content of 75 or 51%, respectively (Matthews, 1968). No peak was identified when searching for a twofold axis in the self-rotation map calculated using the program *GLRF* (Tong & Rossmann, 1997), thus suggesting that the asymmetric unit contains one molecule. No



**Figure 1**  
Crystals of Nsp15 from SARS-CoV crystallized as the apoenzyme.

**Table 1**

X-ray diffraction data statistics.

Values in parentheses are for the highest resolution shell (2.85–2.70 Å).

Data collection	ID14-3, ESRF
Wavelength (nm)	0.931
Resolution (Å)	59–2.70
Total observations	268080
Unique observations	39242
Completeness (%)	100 (100)
Redundancy	11.9 (12.2)
Average $I/\sigma(I)$	14.5 (4.4)
Wilson $B$ factor	50.2
$R_{\text{merge}}^{\dagger}$ (%)	17.4 (60.5)

$$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I - \langle I \rangle|}{\sum_{hkl} \sum_i I}$$

structure with sufficient sequence similarity to Nsp15 is currently available; structure determination using selenomethionine-substituted protein for MAD/SAD phasing is thus necessary and has been initiated.

We would like to acknowledge E. J. Snijder and J. Ziebuhr for providing us with the nsp3 cDNA. SR is supported by the VIZIER European Programme (LSHG-CT-2004-511960). BC, CG, NB, KD, FT, JL and VL were initially supported by the SPINE European Project (QLG2-CT-2002-00988). This work is supported by the Euro-Asian SARS-DTV Network (SP22-CT-2004-511064) from the European Commission Specific Research and Technological Development Programme 'Integrating and Strengthening the European Research Area', as well as the French Genopole programme and the Conseil General of the Bouches-du-Rhone.

## References

Abergel, C., Coutard, B., Byrne, D., Chenivresse, S., Claude, J. B., Deregnacourt, C., Fricaux, T., Giancesini-Boutreux, C., Jeudy, S., Lebrun, R., Maza, C., Notredame, C., Poirot, O., Suhre, K., Varagnol, M. & Claverie, J. M. (2003). *J. Struct. Funct. Genomics*, **4**, 141–157.

Anand, K., Ziebuhr, J., Wadhvani, P., Mesters, J. R. & Hilgenfeld, R. (2003). *Science*, **300**, 1763–1767.

Bhardwaj, K., Guarino, L. & Kao, C. C. (2004). *J. Virol.* **78**, 12218–12224.

Caffarelli, E., Maggi, L., Fatica, A., Jiricny, J. & Bozzoni, I. (1997). *Biochem. Biophys. Res. Commun.* **233**, 514–517.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.

Drosten, C. *et al.* (2003). *N. Engl. J. Med.* **348**, 1967–1976.

Egloff, M. P., Ferron, F., Campanacci, V., Longhi, S., Rancurel, C., Dutartre, H., Snijder, E. J., Gorbalenya, A. E., Cambillau, C. & Canard, B. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 3792–3796.

Gioia, U., Laneve, P., Dlakic, M., Arceci, M., Bozzoni, I. & Caffarelli, E. (2005). *J. Biol. Chem.* **280**, 18996–19002.

Guarino, L. A., Bhardwaj, K., Dong, W., Sun, J., Holzenburg, A. & Kao, C. (2005). *J. Mol. Biol.* **353**, 1106–1117.

Irie, M., Ohgi, K., Iwama, M., Koizumi, M., Sasayama, E., Harada, K., Yano, Y., Udagawa, J. & Kawasaki, M. (1997). *J. Biochem.* **121**, 849–853.

Ivanov, K. A., Hertzog, T., Rozanov, M., Bayer, S., Thiel, V., Gorbalenya, A. E. & Ziebuhr, J. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 12694–12699.

Laneve, P., Altieri, F., Fiori, M. E., Scaloni, A., Bozzoni, I. & Caffarelli, E. (2003). *J. Biol. Chem.* **278**, 13026–13032.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Nishikawa, S., Morioka, H., Kim, H. J., Fuchimura, K., Tanaka, T., Uesugi, S., Hakoshima, T., Tomita, K., Ohtsuka, E. & Ikehara, M. (1987). *Biochemistry*, **26**, 8620–8624.

Rota, P. A. *et al.* (2003). *Science*, **300**, 1394–1399.

Saida, F., Uzan, M. & Bontems, F. (2003). *Nucleic Acids Res.* **31**, 2751–2758.

Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L., Guan, Y., Rozanov, M., Spaan, W. J. & Gorbalenya, A. E. (2003). *J. Mol. Biol.* **331**, 991–1004.

Steyaert, J., Hallenga, K., Wyns, L. & Stanssens, P. (1990). *Biochemistry*, **29**, 9064–9072.

Sutton, G., Fry, E., Carter, L., Sainsbury, S., Walter, T., Nettleship, J., Berrow, N., Owens, R., Gilbert, R., Davidson, A., Siddell, S., Poon, L. L., Diprose, J., Alderton, D., Walsh, M., Grimes, J. M. & Stuart, D. I. (2004). *Structure*, **12**, 341–353.

Thompson, J. E., Venegas, F. D. & Raines, R. T. (1994). *Biochemistry*, **33**, 7408–7414.

Tong, L. & Rossmann, M. G. (1997). *Methods Enzymol.* **276**, 594–611.

Vincentelli, R., Cnaan, S., Offant, J., Cambillau, C. & Bignon, C. (2005). *Anal. Biochem.* **346**, 77–84.

Zhai, Y., Sun, F., Li, X., Pang, H., Xu, X., Bartlam, M. & Rao, Z. (2005). *Nature Struct. Mol. Biol.* **12**, 980–986.