electronic reprint

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction analysis of Langat virus envelope protein domain III

Mark A. White, Deqian Liu, Michael R. Holbrook, Robert E. Shope, Alan D. T. Barrett and Robert O. Fox

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site provided that this cover page is retained. Republication of this article or its storage in electronic databases or the like is not permitted without prior permission in writing from the IUCr.

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction analysis of Langat virus envelope protein domain III

Mark A. White,[a] Deqian Liu,[a] Michael R. Holbrook,[b] Robert E. Shope,[b] Alan D. T. Barrett[b] and Robert O. Fox[a]*

[a]Sealy Center for Structural Biology, Department of Human Biological Chemistry and Genetics, University of Texas Medical Branch at Galveston, Galveston, Texas 77555-0647, USA, and [b]WHO Collaborating Center for Tropical Diseases and Department of Pathology, University of Texas Medical Branch at Galveston, Galveston, Texas, 77555-0366, USA

Correspondence e-mail: fox@xray.utmb.edu

Received 23 October 2002
Accepted 25 February 2003

The putative receptor-binding domain (domain III) of the flavivirus Langat envelope glycoprotein has been crystallized using the hanging-drop vapor-diffusion method at 277 K. Two distinct crystal morphologies were observed to grow under the same conditions. The crystal forms both belong to a trigonal space group, $P3_121$ or $P3_221$, with unit-cell parameters a = 80.93, c = 132.1 Å and a = 104.8, c = 219.5 Å for forms I and II, respectively. Complete data sets to 2.9 and 3.35 Å, respectively, have been collected at 100 K with Cu $K\alpha$ X-rays from a rotating-anode generator.

1. Introduction

The genus *Flavivirus*, family *Flaviviridae*, contains several major human pathogens including the mosquito-borne yellow fever, West Nile and dengue viruses and members of the tick-borne encephalitis (TBE) serogroup, including Central European TBE. Langat (LGT) virus is a naturally attenuated member of the TBE serogroup. The mosquito- and tick-borne pathogens share a highly similar icosahedral viral envelope which consists of 90 copies of a single multi-domain envelope (E) glycoprotein. The structure of a soluble form of the TBE-E glycoprotein (sE), domains I, II and III, has been solved by X-ray crystallography in its dimeric form (Rey *et al.*, 1995). The dimeric form has also been observed in dengue-2 virions using cryo-electron microscopy (Kuhn *et al.*, 2001). These sE domains have been shown to form homodimers on the viral surface (Allison *et al.*, 1995) and on liposomes (Stiasny *et al.*, 2002). Flaviviruses utilize a receptor-mediated endocytotic entry mechanism and undergo an acid-induced membrane-fusion event resulting in the release of their positive-strand RNA genome into the cytosolic compartment of the cell (Heinz & Allison, 2000). The sE component contains both the receptor-binding and the membrane-fusion activities. Through a series of mutagenic studies (Allison *et al.*, 2001), the fusion-peptide region has been isolated to the cd loop (residues 98–113) at the tip of domain II. The putative receptor-binding surface has been mapped onto domain III (Crill & Roehig, 2001; Mandl *et al.*, 2000). The acid-induced fusion mechanism has been shown to proceed through two stages, one reversible and the other irreversible. Acidification of the sE dimer in solution causes its disassociation, but on the surface of a lipsoome it forms trimers (Stiasny *et al.*, 2002). The full-length TBE-E protein has been shown to irreversibly trimerize upon acidification (Allison *et al.*, 1995). The structure of the Langat E protein domain III will serve as the basis for rational drug design of virus-entry inhibitors. Additionally, the pentameric arrangement of subunits in the crystals, as demonstrated by the self-rotation function, may shed light on the conformational transitions of the E protein responsible for membrane fusion.

2. Materials and methods

2.1. Purification

LGT-E-D3 plasmids, coding for residues 301–395 of the E protein, were produced using the pGEX-2T (Pharmacia) GST fusion construct (Bhardwaj *et al.*, 2001). The recombinant plasmid was transformed into *Escherichia coli* DH5α cells, which were grown in 20 l of 2×YT medium with 100 mg l^{-1} ampicillin at 310 K until an OD_{600} of 0.5–0.7 was reached. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 m*M* at 289 K and the cells were grown for an additional 12 h. Cells were harvested by centrifugation. The normal yield was 45 g of wet cells. The cell pellet was suspended in 950 ml buffer *A* (10 m*M* sodium phosphate, 3 m*M* potassium phosphate, 140 m*M* NaCl, 3 m*M* KCl pH 7.0) with 0.1% Tween 20 and sonicated on ice for 30 min with cycles of 30 s on and 30 s off. Cell debris was removed by centrifugation at 39 000*g* for 60 min. The supernatant (~1000 ml) was applied to a 100 ml GST 4B column, washed with buffer *A* containing 0.1% Tween 20 and digested with 100 ml digestion buffer (buffer *A* containing 1000 units of thrombin). The protein was eluted with buffer *A* and the GST fraction was eluted with 10 m*M* *R*-glutathione in 50 m*M* Tris–HCl pH 8.0. The protein pool was run through a Superdex 75

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

electronic reprint

column (prep-grade, 20/600) at a rate of 2 ml min^{-1} with 20 m*M* Tris–HCl, 20 m*M* NaCl pH 7.2 buffer. The peak pool was dialyzed with 3 × 500 ml sample buffer (10 m*M* Tris–HCl pH 7.2) and concentrated to a final concentration of 12–15 mg ml^{-1}.

2.2. Crystallization

Initial crystallization screening was performed using the Hampton Crystal Screen Kit I and the hanging-drop vapor-diffusion method at 277 K. 2 µl of 4 mg ml^{-1} protein solution was mixed with 2 µl reservoir solution (Cudney *et al.*, 1994; Jancarik & Kim, 1991). Several of the drops, including the solution yielding crystals, formed a slight precipitate upon addition of the reservoir solution. After 3–4 weeks, crystals appeared in Hampton Crystal Screen condition Nos. 32 and 47. Crystallization conditions were optimized to 1.8 *M* $(NH_4)_2SO_4$, 0.1 *M* sodium/potassium acetate pH 4.0. These conditions yielded crystals with two different morphologies after 6–8 weeks at 277 K. Form I crystals were thin rods with dimensions of about 0.2 × 0.05 mm (Fig. 1), while form II crystals were thin hexagonal plates of about 0.3 × 0.05 mm in size (Fig. 2).

2.3. Data collection and processing

A crystal was soaked in reservoir solutions containing increasing amounts of glycerol to a maximum of 30%(*v*/*v*) over a period of 12 h. The crystal was then flash-cooled in liquid nitrogen and mounted on the diffractometer in a nitrogen stream at 100 K. Cu $K\alpha$ X-ray diffraction data were collected on a MacScience DIP 2030H image-plate system mounted on a MacScience M06HF rotating-anode generator running at 50 kV and 90 mA and equipped with Bruker Goëbel optics. A native form I data set was collected to 2.9 Å resolution. The data were indexed, integrated and scaled using *DENZO* and *SCALEPACK* from the *HKL* package (Otwinowski, 1993; Otwinowski & Minor, 1997). There were 78 457 measurements, which reduced to 11 128 unique reflections with an average $R_{\rm merge}$ of 10.4% on intensities. The data set was 95.1% complete. A form II data set was collected to 3.35 Å resolution and processed in the same way as the form I data. The high-resolution edge was limited to 3.35 Å by the ability to resolve the long unit-cell axis on the 30 cm diameter detector face and by the small crystal size. There were 103 545 measurements, which reduced to 20 638 unique reflections with an average $R_{\rm merge}$ of 12.7% on intensities. The data set was 99.5% complete. Data-collection statistics are summarized in Table 1.

3. Results and discussion

After purification by GST affinity and Superdex 75 size-exclusion chromatography, a single band was observed on a 15% SDS–PAGE gel corresponding to the 10.4 kDa molecular weight of Langat virus E protein domain III (E-D3). The molecular weight was later confirmed by electrospray mass spectrometry to within 1 amu. Crystallization trials took place using Hampton Research Crystal Screen kits. Only two conditions, Nos. 32 and 47 from kit I, produced crystals. A fine screen around this condition of 1.4–1.8 *M* ammonium sulfate, 100 m*M* sodium/potassium acetate and pH 4.0–4.8 produced diffraction-quality crystals.

Table 1
Summary of data-collection and processing statistics.

	Form I	Form II
Resolution (Å)	2.9	3.35
No. observations	78457	103545
No. unique reflections	11128	20638
Data completeness (%)	95.1 (97.5)	99.5 (99.3)
Average $I/\sigma(I)$	14.9 (6.9)	12.7 (7.1)
Redundancy	7.1 (6.7)	5.0 (4.6)
$R_{\rm merge}$ (%)	10.4 (29.9)	13.2 (29.1)
Space group	$P3_121$ or $P3_221$	$P3_121$ or $P3_221$
Unit-cell parameters (Å)		
a	80.93	104.8
c	132.1	219.5

The best conditions for crystal growth were 1.8 *M* ammonium sulfate, 100 m*M* sodium/potassium acetate pH 4.0. Two distinct crystal morphologies were observed to grow in the same crystallization drops. Form I grew as narrow hexagonal rods (Fig. 1), while form II grew as thin hexagonal plates (Fig. 2). These crystals were soaked in reservoir solutions containing up to 30% glycerol prior to cryocooling and collection of X-ray diffraction data.

Both crystal forms initially produced crystals that diffracted to 3.8 Å resolution. Crystals subsequently obtained from the same well diffracted to better than 3.0 Å. It is possible that the internal disulfide bond was not fully oxidized and that exposure of the wells to fresh air allowed complete oxidation of the disulfide bonds and improved crystal order, resulting in improved diffraction resolution. Complete high-redundancy native data sets have been collected from both forms. Both crystal forms belong to either space group $P3_121$ or $P3_221$, with unit-cell parameters a = 80.93, c = 132.1 Å for form I and a = 104.8, c = 219.5 Å for form II.

The Matthews coefficients (Matthews, 1968) $V_{\rm M}$ of forms I and II are 12.1 and 33.2 Å^3 Da^{-1}, respectively. These coefficients present a large range for the possible number of protomers in the asymmetric unit (a.u.). Given a possible solvent range of 40–70% in the crystals, there could be between three and six monomers in the a.u. of form I and between eight and 16 monomers in the a.u. of form II. The protein was observed to be monomeric in solution. The program *GLRF* (Tong & Rossmann, 1997) was used to produce self-rotation maps and locked rotation searches. Form I self-rotation maps show a peak along c^* for $\kappa = 72^\circ$ and an array of peaks in the a^*b^* plane in the $\kappa = 180^\circ$ map consistent with 5/2 non-crystallographic symmetry. The form II self-rotation maps also show peaks in the $\kappa = 72^\circ$ map 39° from c^* and peaks in the $\kappa = 180^\circ$ map orthogonal to the fivefold axes, again

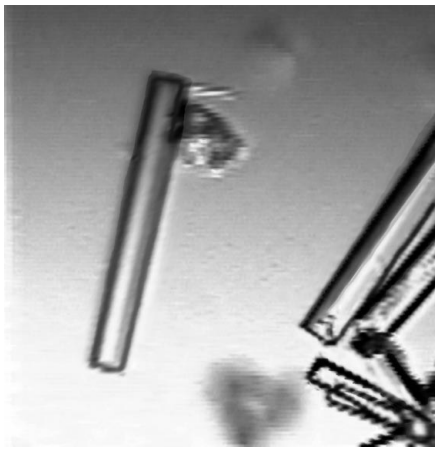

Figure 1
A typical form I crystal. The crystal is a hexagonal rod approximately 0.20 mm long and 0.05 mm in diameter.

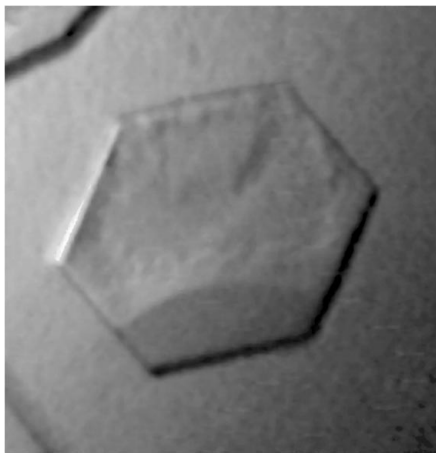

Figure 2
The trigonal morphology of the form II crystal is apparent. This crystal is a thin plate approximately 0.30 mm in diameter and 0.05 mm thick.

electronic reprint

consistent with 5/2 non-crystallographic symmetry. The self-rotation functions are consistent with five monomers in the asymmetric unit for form I and ten copies per asymmetric unit in form II. Molecular replacement has been attempted with form I using the *EPMR* (Kissinger *et al.*, 1999) program with the E-D3 model from the TBE structure as the search model, but without success. The large number of molecules in the asymmetric unit has prompted us to use MIR to solve the structure.

This work was supported by DARPA and CDC grants to RES, a Robert A. Welch Foundation grant to ROF (H-1345) and by a Sealy and Smith Foundation grant to the Sealy Center for Structural Biology. MRH was supported by The James W. McLaughlin Fellowship Fund and an NIH T32 training grant AI07526. We wish to thank S. Bhardwaj and S. J. Watowich for kindly providing the LGT-E-D3 expression plasmid and B. Xu for expression and purification of protein.

References

Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W. & Heinz, F. X. (2001). *J. Virol.* **75**, 4268–4275.

Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W., Kunz, C. & Heinz, F. X. (1995). *J. Virol.* **69**, 695–700.

Bhardwaj, S., Holbrook, M., Shope, R. E., Barrett, A. D. T. & Watowich, S. J. (2001). *J. Virol.* **75**, 4002–4007.

Crill, W. D. & Roehig, J. T. (2001). *J. Virol.* **75**, 7769–7773.

Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.

Heinz, F. X. & Allison, S. L. (2000). *Adv. Virus Res.* **55**, 231–269.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Kissinger, C. R., Gelhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.

Kuhn, R. R., Zhang, W., Rossman, M. G., Pletnev, S. V., Corver, J., Lenches, E., Jones, C. T., Mukhopadhday, S., Chipman, P. R., Strauss, E. G., Baker, T. S. & Strauss, J. H. (2001). *Cell*, **108**, 717–725.

Mandl, C. W., Allison, S. L., Holzmann, H., Meixner, T. & Heinz, F. X. (2000). *J. Virol.* **74**, 9601–9609.

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

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C. & Harrison, S. C. (1995). *Nature (London)*, **375**, 291–298.

Stiasny, K., Allison, S. L., Mandl, C. W. & Heinz, F. X. (2002). *J. Virol.* **75**, 7392–7398.

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

electronic reprint