Crystallization of the RNA guanylyltransferase of *Chlorella* virus PBCV-1

Aidan J. Doherty,[a] Kjell Håkansson,[a] C. Kiong Ho,[b] Stewart Shuman[b] and Dale B. Wigley[a]* at [a]*Laboratory of Molecular Biophysics, University of Oxford, Rex Richards Building, South Parks Road, Oxford OX1 3QU, England, and* [b]*Molecular Biology Program, Sloan-Kettering Institute, New York, NY 10021, USA. E-mail: wigley@biop.ox.ac.uk*

Abstract

RNA guanylyltransferase, or capping enzyme (E.C. 2.7.7.50) catalyzes the transfer of GMP from GTP to diphosphate-terminated RNA to form the cap structure GpppN. *Chlorella* virus capping enzyme expressed in *E. coli* has been purified, treated with GTP and crystallized. X-ray diffraction data have been collected from these crystals as well as for a mercury derivative obtained by soaking the crystals in thimerosal. Selenomethionine RNA guanylyltransferase was purified and crystallized in a similar fashion. The space group is *C*2221 and the cell parameters are $a = 93.3$, $b = 214.9$, $c = 105.8$ Å. Two Hg atoms and two subsets of Se atoms have been localized using difference Patterson and Fourier methods, suggesting that there are two molecules per asymmetric unit.

1. Introduction

Eucaryotic mRNA's contain a 5'-terminal cap structure, m^7GpppN which is synthesized by three enzymes. RNA triphosphatase (E.C. 3.1.3.33) cleaves the 5'-triphosphate terminus leaving a diphosphate end, which is then guanylated by the action of RNA guanylyltransferase (E.C. 2.7.7.50) that transfers GMP from GTP to form GpppN. Finally, RNA (guanine-7-)-methyltransferase (E.C. 2.1.1.56) methylates at the N-7 position of guanine (for a review see Shuman, 1995). This modification is believed to play an important role in RNA synthesis and function, by facilitating mRNA processing and translation and may also protect mRNA from nucleolytic degradation. The transguanylylation reaction is a two-step process involving an enzyme-guanylate intermediate. Studies of these mRNA capping enzymes have revealed that the reaction proceeds *via* a covalent GMP–enzyme intermediate in which the GMP moiety is attached to the protein *via* a lysine residue (Shuman & Hurwitz, 1981),

(i) $E + pppG \longleftrightarrow EpG + PPi$

(ii) $EpG + ppRNA \longleftrightarrow GpppRNA + E.$
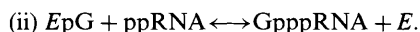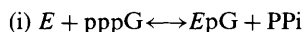
The first step (i) involves the nucleophilic attack by a lysine residue on the α-phosphorus of GTP resulting in the release of inorganic pyrophosphate (PPi) and the formation of a covalent intermediate (EpG) consisting of GMP linked to the ε amino group of the active-site lysine of the enzyme *via* a phosphoramidate bond. In the second step (ii) the enzyme transfers GMP to the terminal 5' phosphate of the RNA.

The lysine residue that is involved in this mechanism is part of a conserved K*X*DG motif which is the same as that found at the active site of DNA and RNA ligases (Lindahl & Barnes, 1992; Cong & Shuman, 1993). Furthermore, there are several other sequence motifs that are also conserved between DNA ligases, RNA ligases, tRNA ligases, and guanylyltransferases, with a similar spacing between. These similarities have led to

the suggestion that nucleotidyl transfer by all of these enzymes must share a common mechanism and that the enzymes are likely to have a similar structure (Shuman & Schwer, 1995), a proposal which is supported by biochemical data of mutant enzymes (Kodama, Barnes & Lindahl, 1991; Cong & Shuman, 1993, 1995; Shuman & Ru, 1995; Fresco & Buratowski, 1994; Shuman, Liu & Schwer, 1994; Heaphy, Singh & Gait, 1987) and more recently by the structure of T7 DNA ligase (Subramanya, Doherty, Ashford & Wigley, 1996). Examination of the positions of these motifs within the DNA ligase structure indicates that not only are they clustered around the ATP binding site, but also that they form the sides of the groove between the two domains of the protein.

To investigate the potential similarities/differences between RNA guanylyltransferases and ATP-dependent DNA ligases we have crystallized the newly discovered capping enzyme from *Paramecium bursaria Chlorella* virus PBCV-1 (Li *et al.*, 1995). PBCV-1 is a large dsDNA virus which replicates in a unicellular eucaryotic *Chlorella*-like green algae (Van Etten, Burbank, Xia & Meints, 1983). The sequence of an open reading frame (A103R) was found to have a high degree of homology with the two yeast guanylyltransferases. A103R encodes a polypeptide with a molecular weight of 38 kDa and contains an active-site lysine (K82) as well as another five motifs conserved across the nucleotidyltransferases. This enzyme has recently been overexpressed and purified in
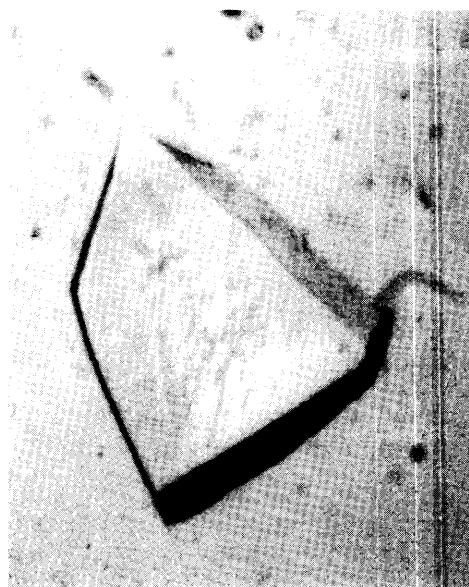

Fig. 1. A crystal of *Paramecium bursaria Chlorella* virus PBCV-1 RNA guanylyltransferase.

large quantities from *E.coli* and shown to have guanylyl-transferase activity (Ho, Van Etten & Shuman, 1996). In this paper we report an alternative protein purification protocol and the subsequent crystallization and preliminary crystallographic characterization of this enzyme. This is the first report of the crystallization of any member of the RNA guanylyltransferase class of enzymes.

2. Methods

2.1. Purification of A103R RNA guanylyltransferase

21 cultures of Luria broth containing $100\,\mu g\,ml^{-1}$ ampicillin and $50\,\mu g\,ml^{-1}$ chloramphenicol were inoculated with a 5 ml culture of B834(DE3)[plysS] [pA103R] and grown at 310 K until the A_{600} reached 0.5–0.6. The cultures were induced by the addition of 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG). Growth was continued for a further 3–4 h before harvesting the cells by centrifugation at 5000g. The cell pellets were stored at 253 K until required. The pellets were lysed by sonication of a 10%(w/v) cell suspension in buffer A [50 mM Tris–HCl, pH 7.5, 2 mM EDTA, 4 mM dithiothreitol (DTT)]. After removal of the cell debris, by centrifugation at 20000g, the supernatant was applied to a Heparin–Sepharose column (20 ml) equilibrated with buffer A. After washing the column with a further two column volumes of buffer A, the protein was step-eluted by washing with 10% buffer B (buffer A containing 2 M sodium chloride). Sodium dodecyl sulfate

polyacrylamide gel electrophoresis (SDS–PAGE) confirmed that this peak contained the semi-purified enzyme at high concentrations. The protein was diluted with buffer to reduce the conductivity to less than that of 50 mM NaCl and this was loaded onto a Blue-Sepharose column which had been pre-equilibrated in buffer A. The protein was step-eluted with 15% buffer B. Purity was again monitored by electrophoresing samples on a 12% SDS polyacrylamide gel. Trace contaminants were removed by gel filtration on a Superdex S75 column. The total yield was 15–20 mg of enzyme per litre of bacterial culture. Selenomethionine protein was produced using a defined medium where methionine was replaced by selenomethionine (LeMaster & Richards, 1985) and purified in the same way.

2.2. Protein crystallization

The eluate from the gel filtration (0.3 mg ml^{-1} protein) was incubated with GTP (5 mM) in the presence of 10 mM $MgCl_2$ for 2 h at room temperature followed by addition of NaCl to 0.5 M. The protein was then concentrated to 10 mg ml^{-1} using Centricon 30 concentrators (Amicon). Crystals were obtained by vapour diffusion at 293 K on a small scale in hanging drops (2 + 2 µl) over a 1 ml well solution of 100 mM Tris–HCl pH 7.5, 200 mM NaCl, 200 mM $(NH_4)_2SO_4$ and methoxypolyethylene glycol with a molecular weight of 5000 at a concentration of 34%(w/v). The crystals were harvested and soaked in the same solution.
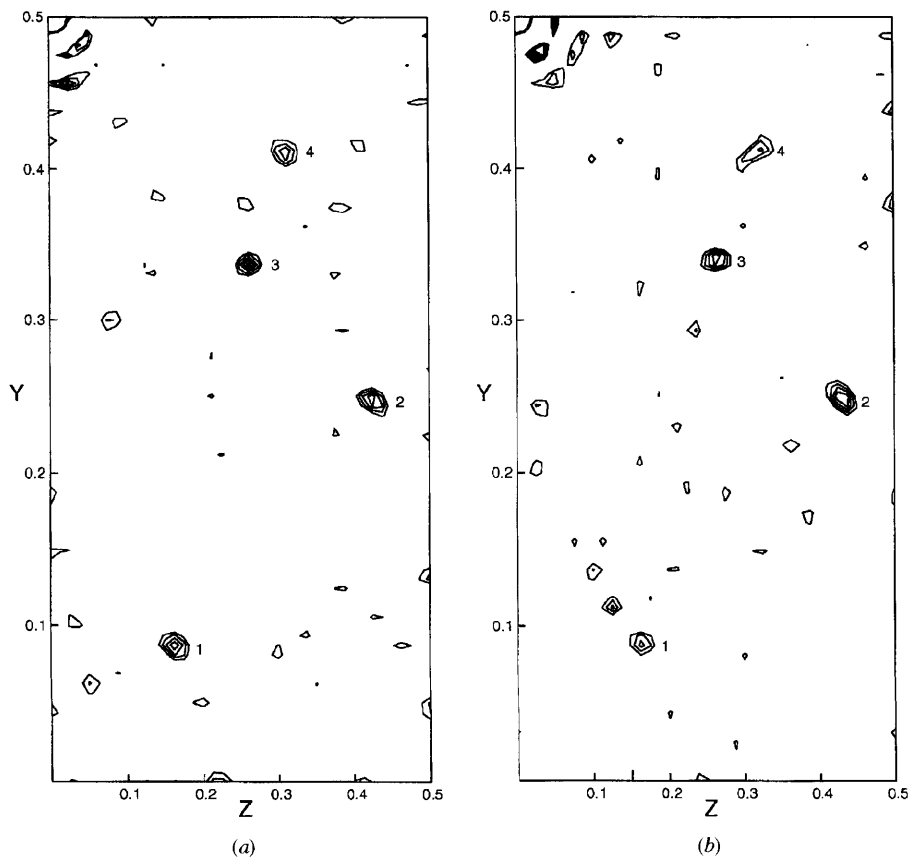

Fig. 2. (a) Difference Patterson function and (b) anomalous Patterson function of a thimerosal derivatized crystal. Shown is the Harker section $x = \frac{1}{2}$. There are two genuine Harker peaks (1,4) and two cross peaks (2,3) which are attributable to two Hg atoms.

2.3. *Data collection and preliminary crystallographic characterization*

X-ray data have been collected to 2.5 Å from a single crystal flash-frozen at 100 K in harvesting solution containing 20% glycerol using synchrotron radiation at Daresbury Laboratory beam station 7.2 with $\lambda = 1.48$ Å and a MAR image-plate detector. In addition, data were collected on selenomethionine and mercury derivatives with a Rigaku rotating-anode X-ray generator. The data were processed using *DENZO* (Otwinowski, 1993) and scaled together with *SCALEPACK*. Subsequent calculations were performed with the *CCP4* program package (Collaborative Computational Project, Number 4, 1994).

3. Results

Chlorella virus guanylyltransferase crystals grew as wedge-shaped plates appearing over a 24–48 h period. They reached full size after 4–8 d and grew to a maximum size of $0.5 \times 0.3 \times 0.15$ mm (Fig. 1). Although the crystals were reasonably stable on exposure to X-rays they were mounted in loops and flash frozen at 100 K in 20% glycerol. Data were collected at the Synchrotron Radiation Source, Daresbury, UK, and are complete to 97.5% and with $R_{merge} = 6.2\%$ for data between 20 and 2.5 Å resolution. The crystals are orthorhombic and belong to the space group $C2221$ with unit-cell dimensions of $a = 93.3$, $b = 214.9$, $c = 105.8$ Å. Assuming that there are two monomers in the asymmetric unit, these crystals have a calculated V_m of 3.5 Å^3 Da^{-1} (Matthews, 1968) and an estimated solvent content of 65%. Data to 2.8 Å were collected using a rotating anode source on a derivative produced by soaking the crystals in 1 mM thimerosal for 2 h followed by back transfer to the mother liquor for 0.5 h. Two Hg-atom positions were identified by difference Patterson analysis (Fig. 2). Selenomethionine-substituted protein was crystallized and data have been collected to 2.7 Å resolution. The Se-atom positions have been determined using difference Fourier techniques and phases calculated from the mercury derivative and the native data. There are two discrete subsets of heavy-atom positions which indicates that there are two molecules in the asymmetric unit.

References

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

Cong, P. & Shuman, S. (1993). *J. Biol. Chem.* 268, 7256–7260.

Cong, P. & Shuman, S. (1995). *Mol. Cell. Biol.* 15, 6222–6231.

Fresco, L. D. & Buratowski, S. (1994). *Proc. Natl Acad. Sci. USA*, 91, 6624–6628.

Heaphy, S., Singh, M. & Gait, M. J. (1987). *Biochemistry*, 26, 1688–1696.

Ho, C. K., Van Etten, J. L. & Shuman, S. (1996). *Virology*, 70, 6658–6664.

Kodama, K. I., Barnes, D. E. & Lindahl, T. (1991). *Nucleic Acids Res.* 19, 6093–6099.

LeMaster, D. M. & Richards, R. M. (1985). *Biochemistry*, 24, 7263–7268.

Li, Y., Lu, Z., Burbank, D. E., Kutish, G. F., Rock, D. L. & Vanetten, J. L. (1995). *Virology*, 212, 134–150.

Lindahl, T. & Barnes, D. E. (1992). *Annu. Rev. Biochem.* 61, 251–281.

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

Otwinowski, Z. (1993). *Data Collection and Processing*, pp. 56–62. Warrington: Daresbury Laboratory.

Shuman, S. (1995). *Prog. Nucleic Acid Res. Mol. Biol.* 50, 101–129.

Shuman, S. & Hurwitz, J. (1981). *Proc. Natl Acad. Sci. USA*, 78, 187–191.

Shuman, S., Liu, Y. & Schwer, B. (1994). *Proc. Natl Acad. Sci. USA*, 91, 12046–12050.

Shuman, S. & Ru, X. M. (1995). *Virology*, 211, 73–83.

Shuman, S. & Schwer, B. (1995). *Mol. Microbiol.* 17, 405–410.

Subramanya, H. S., Doherty, A. J., Ashford, S. R. & Wigley, D. B. (1996). *Cell*, 85, 607–615.

Van Etten, J. L., Burbank, D. E., Xia, Y. & Meints, R. H. (1983). *Virology*, 126, 117–125.