Crystallization and X-ray analysis of the transcription-activator protein C1 of bacteriophage P22 in complex with the P_{RE} promoter element

Avishek Mondal, Rajagopal Chattopadhyaya, Ajit Bikram Datta and Pradeep Parrack*

Received 17 June 2015

Accepted 21 August 2015

Edited by M. S. Weiss, Helmholtz-Zentrum Berlin für Materialien und Energie, Germany

Keywords: lysis–lysogeny choice; lambda CII; P22 C1; phage P22.

Department of Biochemistry, Bose Institute, P-1/12, CIT Scheme VIIM, Kolkata 700 054, India. *Correspondence e-mail: parrack@gmail.com

The transcription-activator protein C1 of the temperate phage P22 of *Salmonella typhimurium* plays a key role in the lytic *versus* lysogenic switch of the phage. A homotetramer of 92-residue polypeptides, C1 binds to an approximate direct repeat similar to the transcription activator CII of coliphage λ . Despite this and several other similarities, including 57% sequence identity to coliphage CII, many biochemical observations on P22 C1 cannot be explained based on the structure of CII. To understand the molecular basis of these differences, C1 was overexpressed and purified and subjected to crystallization trials. Although no successful hits were obtained for the apoprotein, crystals could be obtained when the protein was subjected to crystallization trials in complex with a 23-mer promoter DNA fragment (P_{RE}). These crystals diffracted very well at the home source, allowing the collection of a 2.2 Å resolution data set. The C1–DNA crystals belonged to space group $P2_1$, with unit-cell parameters $a = 87.27$, $b = 93.58$, $c = 111.16$ Å, $\beta = 94.51^\circ$. Solvent-content analysis suggests that the asymmetric unit contains three tetramer–DNA complexes. The three-dimensional structure is expected to shed light on the mechanism of activation by C1 and the molecular basis of its specificity.

1. Introduction

Temperate bacteriophages must decide on their preferred mode of development, lytic or lysogenic, in response to various environmental conditions. The best studied such virus is coliphage λ , in which the phage protein CII (λ CII) plays a crucial role in this decision making (Ptashne, 2004; reviewed in Echols, 1980; Herskowitz & Hagen, 1980; Wulff & Rosenberg, 1983). A transcription activator that binds direct-repeat sequences in DNA (Ho *et al.*, 1983), λ CII exists as a homotetramer in the native state (Ho *et al.*, 1982, 1983). Elucidation of the three-dimensional structure of this protein has revealed its unusual quaternary structure that allows the helix–turn–helix (HTH) motif of two of its four subunits to interact with neighbouring major grooves in the operator region of DNA (Jain *et al.*, 2005; Datta *et al.*, 2005).

Several CII-like proteins exist in other lambdoid phages that infect *Escherichia coli* (Grosschedl & Schwarz, 1979; Ljungquist *et al.*, 1984; Clark *et al.*, 2001; Casjens *et al.*, 2004) or other bacteria (Winston & Botstein, 1981*b*; Ho *et al.*, 1986; Sato *et al.*, 2003) and carry out similar decision-making functions. Other than λ CII, only two have been studied in depth at the protein level: the C1 protein of phage P22 that infects *Salmonella typhimurium* (P22 C1) and the C protein of phage P2 (P2 C) that infects *E. coli*. Both P22 C1 and P2 C recognize direct repeat sequences in DNA (Ho *et al.*, 1992; Ljungquist *et*

© 2015 International Union of Crystallography

al., 1984), which is a hallmark of these proteins. However, P2 C binds DNA as a dimer and exhibits little homology to λ CII (Massad *et al.*, 2010). On the other hand, P22 C1 is reported to be a tetramer (Ho *et al.*, 1992) and has 57% sequence identity to λ CII (Fig. 1). The significant similarities between λ CII (4×97 amino acids) and P22 C1 (4×92 amino acids) in terms of their sequence, size and function as well as their recognition sequence (TTGC-N6-TTGC/T; Ho *et al.*, 1992) strongly indicates that their three-dimensional structures are likely to be similar. Nevertheless, the following experimental observations appear to suggest otherwise, emphasizing the importance of independently determining the three-dimensional structure of C1.

In both λ CII and P22 C1, the tetramer direct-repeat motifs TTGC/T are crucial for DNA–protein interactions (Ho *et al.*, 1992). A homology model of the P22 C1–DNA complex, based on the X-ray structure of the λ CII–DNA complex, does not show any difference in DNA–protein interactions between the two systems. Therefore, it may be expected that the structures of λ CII and P22 C1 are largely identical. Nevertheless, each of the proteins is found to act selectively on its own cognate promoter without any cross-reactivity between the two (Gemski *et al.*, 1972; Botstein & Herskowitz, 1974; Gough & Tokuno, 1975; Hilliker & Botstein, 1976; Hilliker *et al.*, 1978; Winston & Botstein, 1981a; Wiggins & Hilliker, 1985; Wulff & Mahoney, 1987), a result that cannot be explained on the basis of a simple extrapolation of the three-dimensional structure of λ CII to P22 C1.

Interestingly, the binding site of C1 at the -35 region of P_{RE} (which contains the approximate direct repeat) overlaps with the coding region for C1 itself. Retallack *et al.* (1993) reported a so-called ‘operator-sensitive mutation’ *osi*, a single base-pair mutation in this region that changes both the gene sequence of *cI* and the target site for the protein at the -35 region. This results in the substitution of threonine by alanine in the N-terminal region of C1 (T6A). While wild-type (wt) C1 activates both wt and *osi* promoters, the *osi* mutant protein only acts at the *osi* promoter (Retallack *et al.*, 1993). A simple extrapolation of the structure of λ CII to P22 fails to explain this result, since in the former structure the N-terminal region of the protein is located away from the DNA (Jain *et al.*, 2005; Datta *et al.*, 2005).

Thus, it appears that in spite of their significant similarities, subtle structural differences exist between λ CII and P22 C1. An answer to the above issues can only be provided by solving

the three-dimensional structure of P22 C1. In this communication, we report the crystallization of P22 C1 in complex with the P_{RE} promoter element containing its cognate recognition sequence.

2. Materials and methods

2.1. Cloning and expression of C1

A 276 bp fragment from the P22 genome containing the *cI* gene was cloned into the NdeI/BamHI sites of pET-28a vector (Novagen). The presence of the correct insert was confirmed by sequencing. The recombinant protein was expressed in *E. coli* BL21 (DE3) cells transformed with the plasmid. Cells were grown at 310 K in Circlegrow medium (MP Biomedicals) containing $50 \mu\text{g ml}^{-1}$ kanamycin until the A_{590} reached 0.6. The cultures were then induced with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) for 16 h at 288 K. The cells were harvested by centrifugation and stored at 203 K until purification was carried out.

2.2. Protein purification

The cell pellet was resuspended in lysis buffer [20 mM Tris–HCl, 0.5 M NaCl, 2 mM MgCl_2 , 10 mM imidazole, 1 mM PMSF, 5% (v/v) glycerol pH 8.0] and lysed by sonication, followed by centrifugation to remove cell debris. The clarified cell lysate was filtered using a $0.22 \mu\text{m}$ membrane filter (Milipore) and loaded onto a 5 ml HisTrap HP affinity column (GE Healthcare) previously equilibrated with lysis buffer. The recombinant C1 protein was eluted by applying a linear gradient of imidazole (20–500 mM) in 20 mM Tris pH 8.0, 300 mM NaCl, 5% glycerol. The purified protein was dialysed overnight against 20 mM Tris pH 8.0, 300 mM NaCl, 5% glycerol. The six-histidine tag was removed from the purified recombinant C1 by thrombin cleavage ($100 \mu\text{g}$ thrombin per 100 mg protein) overnight at 4°C . The C1 protein was further purified using benzamide Sepharose (to remove thrombin) followed by size-exclusion chromatography using a HiLoad 16/600 Superdex 75 column previously calibrated with a range of molecular-weight standards (Bio-Rad): albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa) and ribonuclease (13.7 kDa). The majority of the protein eluting as a tetramer corresponding to a molecular weight of 44 kDa was collected, leaving out the small amounts of the dimeric protein that were present (Fig. 2a). The purity of the eluted protein

Figure 1

Similarity in the sequences of the C1 and CII proteins. Sequence alignment was carried out using *ClustalW* (Thompson *et al.*, 1994). The sequences which form the four helices ($\alpha 1$ – $\alpha 4$) of one subunit of homotetrameric CII are shown in different colours. Helices $\alpha 1$ and $\alpha 3$ contain residues 26–45, which form the HTH motif.

was further judged by SDS-PAGE followed by Coomassie staining (Fig. 2*a*, inset). Purified C1 was buffer-exchanged into 20 mM Tris pH 8.0, 300 mM NaCl, concentrated to $\sim 20 \text{ mg ml}^{-1}$ by centrifugal filtration (Millipore) and stored at -80°C after flash-cooling in liquid nitrogen. Macromolecule-production information is summarized in Table 1.

2.3. *In vitro* transcription assay

The activity of the recombinant C1 preparations was checked by *in vitro* transcription assays before attempting crystallization trials. The reactions were performed in a 20 μl mixture consisting of $1\times$ transcription buffer (40 mM Tris-HCl, 0.1 M potassium glutamate, 1 mM DTT, 20 mM MgCl_2 pH 8.0), 10 nM template DNA (P22 P_{RE}) and 60 nM sigma-saturated *E. coli* RNA polymerase (Epicentre, USA). 1 μM of C1 tetramers was added. Reactions were first incubated at 37°C for 20 min to form an open complex and transcription was initiated by addition of the nucleotide mixture [0.1 mM each of ATP, GTP and CTP, 0.01 mM UTP (all ribonucleotides were from Amersham Biosciences), 5 μCi [$\alpha\text{-}^{32}\text{P}$]-UTP (BRIT, India), 1 μg heparin]. After 20 min, the reactions were terminated by the addition of 5 μl formamide stop dye (90% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene-cyanol) and the transcripts were resolved by electrophoresis in a 10% (w/v) polyacrylamide-7 M urea gel. The amount of the transcribed RNA from the gel was quantified and scanned using a Typhoon phosphorimager (Fig. 2*b*).

2.4. Preparation of DNA oligomers

HPLC-purified oligonucleotides containing the P_{RE} promoter sequence (each 23 nucleotides in length; sequences are shown in Fig. 2*c*) were purchased from Sigma. Complementary DNA strands were dissolved in 20 mM Tris pH 8, 100 mM NaCl, 1 mM EDTA pH 8, mixed in a 1:1 stoichiometric ratio and annealed by heating the mixture to 363 K for 10 min followed by slow cooling to 277 K. In the case of the C1-DNA complex, DNA was designed with T-overhangs at both ends, in contrast to the majority of DNA-protein complex crystals, where the DNA contains either zero or one complementary overhanging bases at the ends (Fig. 2*c*).

2.5. Crystallization

Crystallization trials for the apoprotein as well as the DNA complex were initially carried out at 295 K in 24-well plates using the hanging-drop vapour-diffusion method with a sparse-matrix screen (Jancarik & Kim, 1991) and similar commercially available versions such as Crystal Screen, Crystal Screen 2 (Hampton Research), Natrix and JCSG-*plus* (Molecular Dimensions). Diffraction-quality crystals could not be obtained with any of these screens. Finally, C1-DNA complex cocrystals were obtained by vapour diffusion by mixing the duplex DNA and P22 C1 (at a molar ratio of 2.5:1 duplex DNA:C1 monomer), with the final concentration of DNA being 0.25 mM after mixing. The mixture was incubated on ice for 2 h, and 1 μl of this mixture was mixed with 1 μl 0.1 M HEPES pH 7.0, 10% PEG 6000 from JCSG-*plus*

Table 1

Composition of P22 C1 and expression strategy.

Restriction sites in the forward and reverse primers are underlined.

Source organism	Phage P22
DNA source	Genomic DNA
Forward primer	GAATTCCATATGGAAGCTCACAAGCACTCGCAAGA
Reverse primer	CGCGGATCCTCAGGCCTCAAACTGTTCCC
Cloning and expression vector	pET-28a
Expression host	<i>E. coli</i> BL21 (DE3)
Complete amino-acid sequence of the construct produced†	GSHMELTSTRKKANAITSSILNRIAIRGQRKVAD- ALGINESQISRWKGDFIPKMGMLLAVLEWGV- DEELAEALAKKVAHLLTKEKPQDCGNSFEA

† C1 in pET-28a after removing the His tag.

Table 2

Crystallization conditions.

Method	Vapour diffusion
Plate type	96-well
Temperature (K)	293
Protein concentration (mg ml^{-1})	2–4
Buffer composition of protein solution	20 mM Tris pH 8.0, 300 mM NaCl
Composition of reservoir solution	0.1 M HEPES pH 7.0, 10% PEG 6000, 0.01 M MgCl_2 , 0.08 M magnesium acetate
Volume and ratio of drop	2 μl , 1:1
Volume of reservoir (μl)	500

Figure 2

(a) Elution profile of C1 after removing the histidine tag on a Superdex 75 gel-filtration column calibrated with molecular-weight standards (labelled in kDa). T and D denote peaks corresponding to the tetramer and the dimer, respectively. The SDS-PAGE of peak T along with molecular-weight standards is shown in the inset. Lane 1, His-tagged C1 (control); lane 2, C1 from peak T. (b) *In vitro* transcription assay of purified C1 protein with its cognate P_{RE} promoter. Template DNA alone (lane 1) and in the presence of 1 μM C1 tetramer (lane 2) were used. The 155 nt runoff transcript induced by C1 is indicated by an arrowhead. (c) Various 23-mer DNA elements derived from the natural P22 P_{RE} promoter sequence or the -35 region containing the C1 recognition site (two underlined direct-repeat half-sites; Ho *et al.*, 1983) used in our crystallization trials and binding studies. (i) TT-overhang DNA. Crystallization of this DNA with C1 produced a sharp-edged plate-shaped crystal that diffracted to 1.68 Å resolution. (ii) AT-overhang DNA. Crystallization using this DNA produced microcrystals (less than 20 μm in size). (iii) Blunt-ended DNA, which did not yield any crystals.

supplemented with 0.01 M MgCl₂ and 0.08 M magnesium acetate, placed in a 20°C incubator and set up for crystallization as detailed in Table 2. Crystals that grew to final dimensions of 100 × 100 × 50 µm were obtained after several months (a smaller crystal is shown in Fig. 3*a*). Although many different DNA constructs varying in length and the nature of the terminus (blunt end *versus* overhang) were screened, diffraction-quality crystals were only obtained using the T-overhang 23-mer DNA. SDS-PAGE (Fig. 3*b*, right) and 1.0% (w/v) agarose gel (Fig. 3*b*, left) analysis of these crystals

confirmed that they indeed contained the expected complex formed by C1 and the DNA.

2.6. Data collection and processing

For data collection at cryogenic temperature, crystals were soaked in a cryoprotectant solution consisting of the artificial mother liquor supplemented with 20% (v/v) ethylene glycol and were flash-cooled in liquid nitrogen. X-ray diffraction data were collected using a Rigaku RU-300 generator equipped

Figure 3

(*a*) A plate-shaped crystal of the C1-DNA complex obtained using the hanging-drop vapour-diffusion method measuring about 50 µm on each side. X-ray data were collected from another larger crystal from the same drop. The crystals grew along with a heavy brown precipitate. (*b*) Authenticity of the protein-DNA crystal. Two crystals of the C1-DNA complex were washed, dissolved and run on (left) 1% agarose gel stained with ethidium bromide, showing the corresponding 23-mer DNA, and (right) 15% SDS-PAGE stained with Coomassie Blue, showing the results from the initial 2 min crystal wash (lane 1) and the crystal subsequently dissolved in 5 µl water warmed to about 90°C (lane 2). (*c*) A typical X-ray diffraction pattern from a crystal of the C1-DNA complex, annotated with resolution arcs (from the centre, 5, 3, 2 and 1.8 Å). The inset shows two diffraction spots at an even higher resolution of beyond 1.8 Å.

with a Cu anode and Blue optics and an R-AXIS IV⁺⁺ image-plate detector at 110 K in steps of 1° over a total rotation of 160° with an exposure time of 10 min per frame. Although we processed the data at 2.2 Å resolution to achieve a reasonable $I/\sigma(I)$ and completeness, some diffraction spots could be observed beyond 1.8 Å resolution (1.68 Å was the data-collection limit, which corresponded to the corner of the square R-AXIS IV⁺⁺ image plate; Fig. 3c). Diffraction images were indexed, integrated and scaled using *d*TREK* (Pflugrath, 1999). The file containing the intensities was imported into the *CCP4* suite (Winn *et al.*, 2011) for further processing.

From a measurement of the density of some smaller crystals from the same hanging drop (ascertained to be 1.08 g cm⁻³) and subsequent molecular-replacement calculations within the *CCP4* package (Winn *et al.*, 2011), we concluded that the asymmetric unit most probably contained three C1–DNA complexes. This corresponds to a V_M of 2.69 Å³ Da⁻¹ and a solvent content of 54.29% using a molecular mass of 56 kDa for the tetramer–DNA complex. The data-collection statistics are shown in Table 3.

3. Results and discussion

The C1 protein could be successfully crystallized as a complex with DNA. The crystal belonged to the monoclinic space group *P2*₁, with unit-cell parameters $a = 87.27$, $b = 93.58$, $c = 111.16$ Å, $\beta = 94.51^\circ$ (Table 3). However, the protein alone failed to crystallize under similar conditions. Moreover, the C1–DNA complex crystals grew at a very slow rate, appearing only after 3–4 months. They displayed extreme sensitivity to minor fluctuations in the environmental conditions, hampering their reproducibility. Considering C1 to bind DNA as a tetramer (Ho *et al.*, 1992), we carried out a Matthews coefficient analysis of our crystal (Matthews, 1968). It appeared that the asymmetric unit most probably contains three C1–DNA complexes ($V_M = 2.69$ Å³ Da⁻¹ with a solvent content of 54.29%). However, four such tetrameric complexes within the asymmetric unit ($V_M = 2.02$ Å³ Da⁻¹ with a solvent content of 39.06%) could not be ruled out.

The self-rotation function also failed to provide an unambiguous assignment of the number of molecules in the asymmetric unit, although it revealed the presence of two twofold axes (corresponding to $\theta = 26^\circ$ and $\theta = 52^\circ$ at $\chi = 180^\circ$) which could indicate the partial twofold axes in the DNA-binding region of C1, as observed in the λ CII structure. It also showed a fivefold axis characteristic of duplex DNA.

We further attempted to carry out molecular replacement in order to solve the phases using both the structures of λ CII alone (PDB entry 1xwr; Datta *et al.*, 2005) and the λ CII–DNA complex (1zs4; Jain *et al.*, 2005). To our dismay, none of the trials led to any obvious acceptable solution. The most probable solution, which had three C1 tetramers in the asymmetric unit, was obtained with PDB entry 1zs4 as the input model using data in the resolution range 8–3 Å for the rotation function and data in the resolution range 20–3 Å for the translation search with *MOLREP* (Vagin & Teplyakov, 2010). This solution was further refined by rigid-body and restrained

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	RU-300, Blue optics
Wavelength (Å)	1.54178
Temperature (K)	110
Exposure time per image (s)	300
Space group	<i>P2</i> ₁
Unit-cell parameters (Å, °)	$a = 87.27$, $b = 93.58$, $c = 111.16$, $\beta = 94.51$
Mosaicity (°)	1.19
Resolution range (Å)	39.45–2.20 (2.28–2.20)
Total No. of reflections	284086†
No. of unique reflections	89045
Completeness (%)	98.3 (96.4)
Average multiplicity	3.19 (3.12)
$\langle I/\sigma(I) \rangle$	6.2 (1.7)
R_{merge}	0.110 (0.474)
R_{meas}	0.132 (0.569)

† There were 449915 reflections accepted and 155374 unique reflections for the 1.68 Å resolution data; the data set was 76.6% complete to this resolution, with $R_{\text{merge}} = 0.135$ and $R_{\text{meas}} = 0.163$.

refinement using *REFMAC5* (Murshudov *et al.*, 2011). However, despite several trials the *R* factor could not be reduced beyond 45.67%, thereby raising doubts about the validity of the obtained solution. Given this difficulty in obtaining a molecular-replacement solution, we verified that the crystal did not suffer from any twinning artifacts. In view of the above, it appears that P22 C1, despite its high sequence similarity to λ CII, might have major structural differences from the latter. Efforts are in progress to obtain suitable heavy-atom derivatives to solve the ambiguity of the phases in the present solution.

Acknowledgements

The authors would like to thank Professor P. Chakrabarti of the Department of Biochemistry, Bose Institute for allowing use of his computer and various programs for data analysis and refinement and Ms Jesmita Dhar for assistance in generating the homology model for the P22 C1–DNA complex. This work was supported by the Institutional Programme II of the Bose Institute. AM was the recipient of a fellowship from CSIR, India.

References

- Botstein, D. & Herskowitz, I. (1974). *Nature (London)*, **251**, 584–589.
- Casjens, S., Winn-Stapley, D. A., Gilcrease, E. B., Morona, R., Kühlewein, C., Chua, J. E. H., Manning, P. A., Inwood, W. & Clark, A. J. (2004). *J. Mol. Biol.* **339**, 379–394.
- Clark, A. J., Inwood, W., Cloutier, T. & Dhillon, T. S. (2001). *J. Mol. Biol.* **311**, 657–679.
- Datta, A. B., Panjekar, S., Weiss, M. S., Chakrabarti, P. & Parrack, P. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 11242–11247.
- Echols, H. (1980). *Molecular Genetics of Development*, edited by T. Leighton & W. F. Loomis, pp. 1–16. New York: Academic Press.
- Gemski, P. Jr, Baron, L. S. & Yamamoto, N. (1972). *Proc. Natl Acad. Sci. USA*, **69**, 3110–3114.
- Gough, M. & Tokuno, I. (1975). *Mol. Gen. Genet.* **138**, 71–79.
- Grosschedl, R. & Schwarz, E. (1979). *Nucleic Acids Res.* **6**, 867–881.
- Herskowitz, I. & Hagen, D. (1980). *Annu. Rev. Genet.* **14**, 399–445.
- Hilliker, S. & Botstein, D. (1976). *J. Mol. Biol.* **106**, 537–566.

- Hilliker, S., Gottesman, M. & Adhya, S. (1978). *Virology*, **86**, 37–47.
- Ho, Y.-S., Lewis, M. & Rosenberg, M. (1982). *J. Biol. Chem.* **257**, 9128–9134.
- Ho, Y.-S., Pfarr, D., Strickler, J. & Rosenberg, M. (1992). *J. Biol. Chem.* **267**, 14388–14397.
- Ho, Y.-S., Wulff, D. & Rosenberg, M. (1983). *Nature (London)*, **304**, 703–708.
- Ho, Y.-S., Wulff, D. & Rosenberg, M. (1986). *Regulation of Gene Expression*, edited by I. Booth & C. Higgins, pp. 79–103. Cambridge University Press.
- Jain, D., Kim, Y., Maxwell, K. L., Beasley, S., Zhang, R., Gussin, G. N., Edwards, A. M. & Darst, S. A. (2005). *Mol. Cell*, **19**, 259–269.
- Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
- Ljungquist, E., Kockum, K. & Bertani, L. E. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 3988–3992.
- Massad, T., Skaar, K., Nilsson, H., Damberg, P., Henriksson-Peltola, P., Haggård-Ljungquist, E., Högbom, M. & Stenmark, P. (2010). *Nucleic Acids Res.* **38**, 7778–7790.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D* **67**, 355–367.
- Pflugrath, J. W. (1999). *Acta Cryst. D* **55**, 1718–1725.
- Ptashne, M. (2004). *A Genetic Switch: Phage Lambda Revisited*, 3rd ed. New York: Cold Spring Harbor Laboratory Press.
- Retallack, D. M., Johnson, L. L., Ziegler, S. F., Strauch, M. A. & Friedman, D. I. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 9562–9565.
- Sato, T., Shimizu, T., Watarai, M., Kobayashi, M., Kano, S., Hamabata, T., Takeda, Y. & Yamasaki, S. (2003). *J. Bacteriol.* **185**, 3966–3971.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D* **66**, 22–25.
- Wiggins, B. A. & Hilliker, S. (1985). *J. Virol.* **56**, 1030–1033.
- Winn, M. D *et al.* (2011). *Acta Cryst. D* **67**, 235–242.
- Winston, F. & Botstein, D. (1981*a*). *J. Mol. Biol.* **152**, 209–232.
- Winston, F. & Botstein, D. (1981*b*). *J. Mol. Biol.* **152**, 233–245.
- Wulff, D. L. & Mahoney, E. (1987). *Genetics*, **115**, 597–604.
- Wulff, D. L. & Rosenberg, M. (1983). *Lambda II*, edited by R. W. Hendrix, J. W. Roberts, F. W. Stahl & R. A. Weisberg, pp. 53–73. New York: Cold Spring Harbor Laboratory Press.