Acta Crystallographica Section F
**Structural Biology and Crystallization Communications**
ISSN 1744-3091

# Cloning, expression, purification, crystallization and preliminary X-ray studies of a secreted lectin (Rv1419) from *Mycobacterium tuberculosis*

**Dhabaleswar Patra, R. Srikalaivani, Ashish Misra, D. D. Singh, M. Selvaraj and M. Vijayan***

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Correspondence e-mail: mv@mbu.iisc.ernet.in

Received 15 September 2010
Accepted 21 October 2010

A secreted lectin, Rv1419, from *Mycobacterium tuberculosis* has been cloned, expressed, purified and crystallized and the crystals have been characterized. This represents the first X-ray investigation of a lectin or lectin-like molecule from the pathogen. The cubic crystals contain one molecule in the asymmetric unit. Sequence comparisons indicate that the lectin has a $\beta$-trefoil fold and belongs to a well characterized family of carbohydrate-binding modules. Structural analysis of the crystals is in progress.

## 1. Introduction

Lectins are sugar-binding proteins that specifically recognize and bind diverse carbohydrate structures without modifying them (Vijayan & Chandra, 1999; Loris, 2002; Sharon, 2007). Although originally discovered in plants, they are now known to exist in all forms of life. They mediate a variety of biological processes such as symbiosis, host–microbe interactions, cell–cell adhesion, innate immunity and mitogenesis. We have been following a long-range programme on the structure and interactions of plant lectins (Banerjee *et al.*, 1994; Sankaranarayanan *et al.*, 1996; Vijayan & Chandra, 1999; Ramachandraiah *et al.*, 2003; Jeyaprakash *et al.*, 2004; Singh *et al.*, 2005; Natchiar *et al.*, 2007; Kulkarni *et al.*, 2007; Sharma *et al.*, 2009). The present communication represents the first attempt to extend this work to mycobacterial lectins as part of another ongoing progamme on the structural biology of mycobacterial proteins (Vijayan, 2005; Krishna *et al.*, 2007; Selvaraj *et al.*, 2007; Roy *et al.*, 2008; Kaushal *et al.*, 2008; Prabu *et al.*, 2009; Chetnani *et al.*, 2010).

Studies of microbial lectins have been much less extensive compared with those of plant and animal lectins, although their role in facilitating the attachment of microbes to their hosts is now well recognized (Bewley, 2006). This role is particularly important in relation to host–pathogen interactions. The most thoroughly studied microbial lectin is undoubtedly influenza virus haemagglutinin (Wiley *et al.*, 1981; Skehel & Wiley, 2000). The cholera toxin with a lectin component has also been extensively studied (Zhang *et al.*, 1995; Merritt *et al.*, 1998). The crystal structures of a number of other microbial lectins or toxins which bind carbohydrates have also been reported (Swaminathan & Eswaramoorthy, 2000; Emsley *et al.*, 2000; Tanaka *et al.*, 2009). However, no structure of a mycobacterial lectin has been reported to date. In fact, reports of other studies of such lectins have been few and far between (Nogueira *et al.*, 2010; Locht *et al.*, 2006). Recently, 11 lectins or lectin-like molecules were identified through a bioinformatic analysis (Singh *et al.*, 2007). Of these 11, Rv1419 is the smallest and has been shown to cause agglutination of rabbit erythrocytes, a characteristic feature of lectins, after 16 h of incubation (Nogueira *et al.*, 2010). Sequence analysis revealed that the bulk of the protein is homologous to the $\beta$-trefoil lectin domain of type II ribosome-inactivating proteins (RIPs), which in turn form part of carbohydrate-binding module 13 (CBM13; http://www.cazy.org) as categorized by Boraston *et al.* (2004). In addition, the lectin contains a 33-amino-acid signal sequence at the N-terminus according to the *SignalP* server (Emanuelsson *et al.*, 2007; Nogueira *et al.*, 2010). In

order to understand the structural basis of the function of the lectin, Rv1419 has been cloned and expressed. The purified protein has been crystallized and preliminary X-ray studies of the crystals have been carried out.

## 2. Materials and methods

### 2.1. Cloning, expression and purification

The *Mycobacterium tuberculosis Rv1419* gene of length 474 bp coding for a protein of 157 amino acids was PCR-amplified using *Pfu* DNA polymerase with *M. tuberculosis* H37Rv genomic DNA as template with forward primer 5′-GATCGCTAGCATGGGTGAATTACGGTTG-3′ and reverse primer 5′-TATCTCGAGCGGCACGCTATCCCA-3′. The PCR involved 30 cycles, with initial denaturation at 368 K for 45 s, annealing at 329.8 K for 1 min and extension at 345 K for 1 min. The PCR products were then digested with the restriction enzymes *Nhe*I and *Xho*I and cloned into the pET21b vector with restriction sites appropriate for the two enzymes to obtain C-terminally hexahistidine-tagged Rv1419 protein. The identity of the clone was verified using DNA sequencing. The construct thus obtained, pET21bRv1419 (-*Nhe*I-*Xho*I) with LEHHHHHH at the C-terminus of Rv1419, was used for overproduction of the protein.

The clone was transformed into *Escherichia coli* BL21 (DE3) strain, grown in 4 l LB medium to an $A_{595}$ of ∼0.6 and induced with

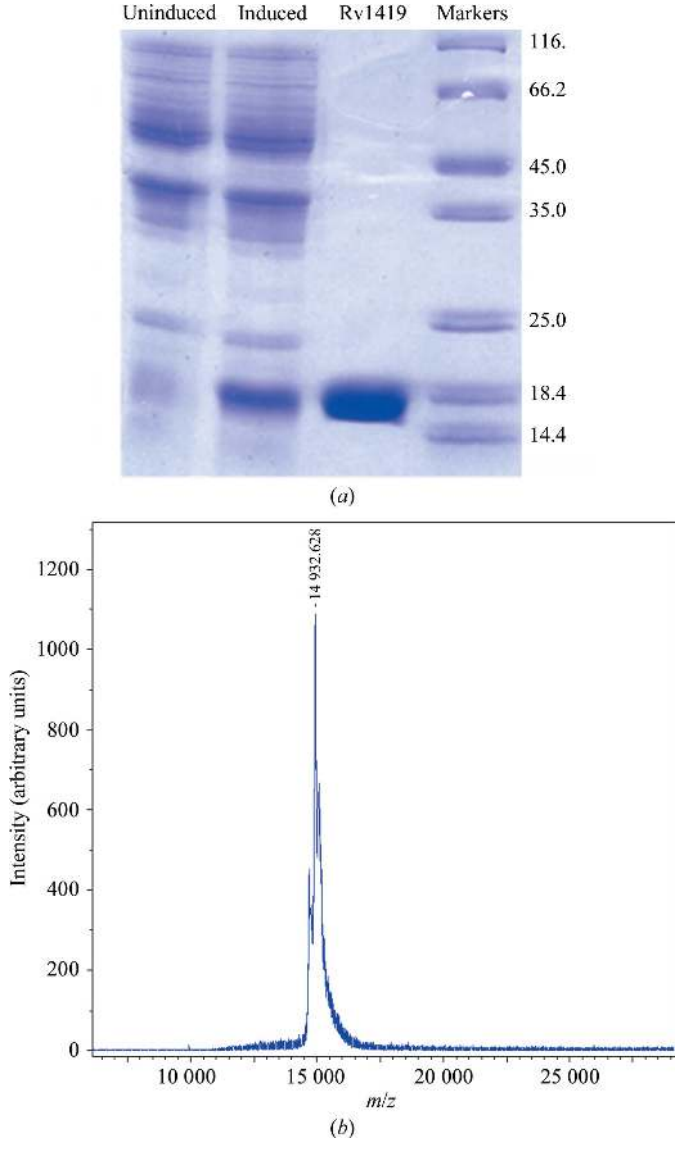


**Figure 1**
(*a*) Electrophoretic profile of Rv1419 and (*b*) mass-spectrometric analysis of the purified Rv1419 protein.

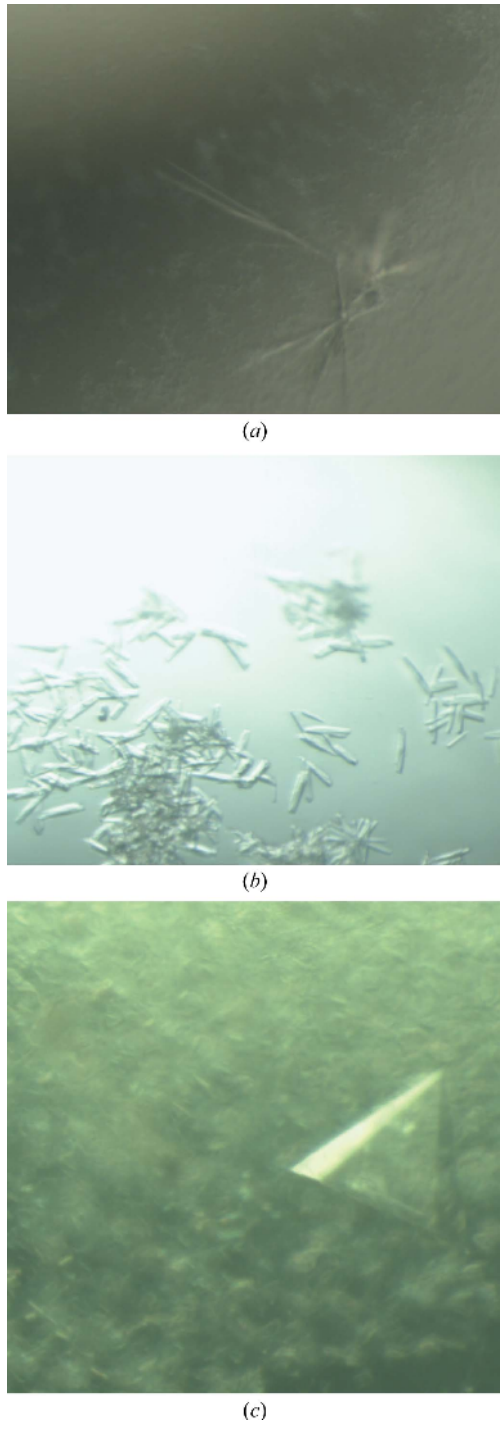


**Figure 2**
(*a*) Crystals grown from a solution consisting of 0.2 *M* magnesium chloride, 30%(*v*/*v*) 2-propanol, 0.1 *M* Na HEPES pH 7.5, (*b*) crystals obtained on addition of dextran sulfate and (*c*) crystals obtained on further addition of β-mercaptoethanol (approximate dimensions 0.1 × 0.1 × 0.1 mm).

1 m*M* IPTG for 6 h at 310 K. The cells were harvested by centrifugation at 6000*g* for 15 min, resuspended in a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 300 m*M* NaCl, 10 m*M* imidazole, 10 m*M* $\beta$-mercaptoethanol and 10%(*v*/*v*) glycerol and lysed by sonication. The lysate was then centrifuged at 15 000*g* for 30 min. The supernatant was loaded onto an Ni–NTA column equilibrated with the above-mentioned buffer, washed with the same buffer containing 15 m*M* imidazole and eluted with the same buffer containing an imidazole gradient from 15 to 300 m*M*. Fractions containing the desired protein were pooled and concentrated using a Microsep (5 kDa cutoff, Millipore) followed by gel filtration (Bio-gel P100) to obtain the homogeneous protein, which was dialysed against phosphate-buffered saline (150 m*M* NaCl, 20 m*M* $Na_2HPO_4$ and 4 m*M* $NaH_2PO_4$ pH 7.5). The protein purity was confirmed using SDS–PAGE (Fig. 1*a*). The mass spectrum of the purified protein was recorded using matrix-assisted laser desorption/ionization–time-of-flight mass spectrometry (MALDI–TOF; Fig. 1*b*). The presence of the His tag was confirmed by western blotting using an anti-His antibody. The purified protein was concentrated to 8 mg ml$^{-1}$ for crystallization experiments.

### 2.2. Crystallization

Initial crystallization attempts were performed by the hanging-drop vapour-diffusion method using Crystal Screen and Crystal Screen 2 from Hampton Research. Typically, for each condition 2 µl protein solution in phosphate-buffered saline (150 m*M* NaCl, 20 m*M* $Na_2HPO_4$ and 4 m*M* $NaH_2PO_4$ pH 7.5) was mixed with an equal volume of precipitant solution to form a hanging drop. The volume of the well solution was 400 µl. After 4 d, needle-shaped crystals appeared in Crystal Screen condition No. 12, consisting of 0.2 *M* magnesium chloride, 0.1 *M* Na HEPES buffer pH 7.5 and 30%(*v*/*v*) 2-propanol (Fig. 2*a*). The addition of 3%(*w*/*v*) dextran sulfate to the protein solution resulted in better crystals, but they were small and diffracted poorly (Fig. 2*b*). The subsequent addition of 15 m*M* $\beta$-mercaptoethanol to the solution resulted in crystals that could be used for diffraction studies (Fig. 2*c*).

### 2.3. X-ray data collection and processing

X-ray data were collected from a crystal of dimensions of around 0.10 × 0.10 × 0.10 mm. Before mounting, the crystals were transferred into freshly prepared well solution containing 20%(*v*/*v*) ethylene glycol as a cryoprotectant and soaked for 1 min. Diffraction data were collected at 100 K and were processed using *MOSFLM* (Leslie, 1992) and scaled using *SCALA* (Collaborative Computational Project, Number 4, 1994). The crystal parameters and data-processing statistics are summarized in Table 1.

## 3. Results and discussion

*M. tuberculosis* lectin Rv1419 with a C-terminal His tag has been purified to homogeneity (Fig. 1*a*) and has a molecular weight of about 15 kDa (Fig. 1*b*). The purified protein does not contain the signal peptide. The protein has been crystallized and the crystals have been characterized using X-ray diffraction. The cubic crystals contained one molecule in the asymmetric unit, with a solvent content of 65% (Matthews, 1968).

Sequence alignment (http://www.ebi.ac.uk/Tools/emboss/align/) of Rv1419 with the lectin B chains of type II RIPs of known structure [PDB codes 1rzo (A. G. Gabdoulkhakov, Y. Savochkina, N. Konareva, R. Krauspenhaar, S. Stoeva, S. V. Nikonov, W. Voelter, C. Betzel & A. M. Mikhailov, unpublished work), 1pc8 (Mishra *et al.*, 2004), 1ce7 (Krauspenhaar *et al.*, 1999), 1hwm (Pascal *et al.*, 2001), 2aai (Rutenber *et al.*, 1991), 1abr (Tahirov *et al.*, 1995) and 2vlc (Azzi *et al.*, 2009)] indicated sequence identity ranging from 19.7 to 26.5%. The highest identity (26.5%) was exhibited by the B chain of abrin (Tahirov *et al.*, 1995), which was used as a search model in molecular-replacement calculations using *Phaser* (McCoy *et al.*, 2007). The results were barely acceptable. A *BLASTP* (Altschul *et al.*, 1990) search of proteins for which three-dimensional structures are known picked out the xylan-binding domain of endo-1,4-$\beta$-xylanase from *Streptomyces olivavaceoviridis* (PDB code 1xyf; Fujimoto *et al.*, 2000), again a member of the CBM13 family, as a homologue with 24.8% sequence identity. X-ray analysis of the crystals using this xylanase as the search model is in progress.

**Table 1**
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

| | |
|---|---|
| No. of crystals | 1 |
| X-ray generator | Bruker Microstar Ultra II Cu $K\alpha$ rotating anode |
| Wavelength (Å) | 1.5418 |
| Detector | MAR 345 |
| Crystal-to-detector distance (mm) | 200 |
| Rotation range per image (°) | 1 |
| Total rotation range (°) | 50 |
| Exposure time per image (s) | 900 |
| Resolution range (Å) | 40.00–2.70 (2.85–2.70) |
| Space group | *F*23 |
| Unit-cell parameters (Å) | $a = b = c = 136.17$ |
| Mosaicity (°) | 0.65 |
| Total no. of measured intensities | 35355 (5156) |
| Unique reflections | 5862 (858) |
| Multiplicity | 6.0 (6.0) |
| Mean $I/\sigma(I)$ | 20.1 (3.4) |
| Completeness (%) | 99.9 (100) |
| $R_{\rm merge}$† (%) | 7.0 (50.9) |
| $R_{\rm meas}$‡ (%) | 7.6 (55.7) |
| Overall *B* factor from Wilson plot (Å$^2$) | 69.7 |

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th observation of reflection *hkl* and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations *i* of reflection *hkl*. ‡ $R_{\rm meas} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ are the observed intensities, $\langle I(hkl)\rangle$ are the average intensities and *N* is the multiplicity of reflection *hkl*.

The data used in the present work were collected at the X-ray Facility for Structural Biology at the Institute, supported by the Department of Science and Technology (DST). MV is a DAE Homi Bhabha Professor. DP and MS are CSIR Research Fellows. Financial support from the DST is acknowledged.

## References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Azzi, A., Wang, T., Zhu, D.-W., Zou, Y.-S., Liu, W.-Y. & Lin, S.-X. (2009). *Proteins*, **74**, 250–255.
Banerjee, R., Mande, S. C., Ganesh, V., Das, K., Dhanaraj, V., Mohanta, S. K., Suguna, K., Surolia, A. & Vijayan, M. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 227–231.
Bewley, C. A. (2006). *Protein–Carbohydrate Interactions in Infectious Diseases.* Cambridge: Royal Society of Chemistry.
Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2004). *Biochem. J.* **382**, 768–781.
Chetnani, B., Kumar, P., Surolia, A. & Vijayan, M. (2010). *J. Mol. Biol.* **400**, 171–185.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Emanuelsson, O., Brunak, S., Heinje, G. V. & Nielsen, H. (2007). *Nature Protoc.* **2**, 953–971.

Emsley, P., Fotinou, C., Black, I., Fairweather, N. F., Charles, I. G., Watts, C., Hewitt, E. & Isaacs, N. W. (2000). *J. Biol. Chem.* **275**, 8889–8894.
Fujimoto, Z., Kuno, A., Kaneko, S., Yoshida, S., Kobayashi, H., Kusakabe, I. & Mizuno, H. (2000). *J. Mol. Biol.* **300**, 575–585.
Jeyaprakash, A. A., Srivastav, A., Surolia, A. & Vijayan, M. (2004). *J. Mol. Biol.* **338**, 757–770.
Kaushal, P. S., Talawar, R. K., Krishna, P. D. V., Varshney, U. & Vijayan, M. (2008). *Acta Cryst.* D**64**, 551–560.
Krauspenhaar, R., Eschenburg, S., Perbandt, M., Kornilov, V., Konareva, N., Mikailova, I., Stoeva, S., Wacker, R., Maier, T., Singh, T. P., Mikhailov, A., Voelter, W. & Betzel, C. (1999). *Biochem. Biophys. Res. Commun.* **257**, 418–424.
Krishna, R., Prabu, J. R., Manjunath, G. P., Datta, S., Chandra, N. R., Muniyappa, K. & Vijayan, M. (2007). *J. Mol. Biol.* **367**, 1130–1144.
Kulkarni, K. A., Katiyar, S., Surolia, A., Vijayan, M. & Suguna, K. (2007). *Proteins*, **68**, 762–769.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Locht, C., Hougardy, J. M., Rouanet, C., Place, S. & Mascart, F. (2006). *Tuberculosis*, **86**, 303–309.
Loris, R. (2002). *Biochim. Biophys. Acta*, **1572**, 198–208.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K. & Hol, W. G. (1998). *J. Mol. Biol.* **282**, 1043–1059.
Mishra, V., Ethayathulla, A. S., Sharma, R. S., Yadav, S., Krauspenhaar, R., Betzel, C., Babu, C. R. & Singh, T. P. (2004). *Acta Cryst.* D**60**, 2295–2304.
Natchiar, S. K., Suguna, K., Surolia, A. & Vijayan, M. (2007). *Crystallogr. Rev.* **13**, 3–28.
Nogueira, L., Cardoso, F. C., Mattos, A. M., Bordignon, J., Figueiredo, C. P., Dahlstrom, P., Frota, C. C., Duarte-dos-Santos, C. N., Chalhoub, M., Cavada, B. S., Teixeira, H. C., Oliveira, S. C., Barral-Netto, M. & Bafica, A. (2010). *Eur. J. Immunol.* **40**, 744–753.
Pascal, J. M., Day, P. J., Monzingo, A. F., Ernst, S. R., Robertus, J. D., Iglesias, R., Perez, Y., Ferreras, J. M., Citores, L. & Girbes, T. (2001). *Proteins*, **43**, 319–326.
Prabu, R., Thamotharan, S., Khanduja, J. S., Chandra, N. R., Muniyappa, K. & Vijayan, M. (2009). *Biochim. Biophys. Acta*, **1794**, 1001–1009.
Ramachandraiah, G., Chandra, N. R., Surolia, A. & Vijayan, M. (2003). *Glycobiology*, **13**, 765–775.
Roy, S., Saraswathi, R., Chatterji, D. & Vijayan, M. (2008). *J. Mol. Biol.* **375**, 948–959.
Rutenber, E., Katzin, B. J., Ernst, S., Collins, E. J., Mlsna, D., Ready, M. P. & Robertus, J. D. (1991). *Proteins*, **10**, 240–250.
Sankaranarayanan, R., Sekar, K., Benerjee, R., Sharma, V., Surolia, A. & Vijayan, M. (1996). *Nature Struct. Biol.* **3**, 596–603.
Selvaraj, M., Roy, S., Singh, N. S., Sangeetha, R., Varshney, U. & Vijayan, M. (2007). *J. Mol. Biol.* **372**, 186–193.
Sharma, A., Sekar, K. & Vijayan, M. (2009). *Proteins*, **77**, 760–777.
Sharon, N. (2007). *J. Biol. Chem.* **282**, 2753–2764.
Singh, D. D., Chandran, D., Jeyakani, J. & Chandra, N. (2007). *Protein Pept. Lett.* **14**, 683–691.
Singh, D. D., Saikrishnan, K., Kumar, P., Sekar, K., Surolia, A. & Vijayan, M. (2005). *Glycobiology*, **15**, 1025–1032.
Skehel, J. J. & Wiley, D. C. (2000). *Annu. Rev. Biochem.* **69**, 531–569.
Swaminathan, S. & Eswaramoorthy, S. (2000). *Nature Struct. Biol.* **7**, 683–699.
Tahirov, T. H., Lu, T.-H., Liaw, Y.-C., Chen, Y.-L. & Lin, J.-Y. (1995). *J. Mol. Biol.* **250**, 354–367.
Tanaka, H., Chiba, H., Inokoshi, J., Kuno, A., Sugai, T., Takahashi, A., Ito, Y., Tsunoda, M., Suzuki, K., Takénaka, A., Sekiguchi, T., Umeyama, H., Hirabayashi, J. & Omura, S. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 15633–15638.
Vijayan, M. (2005). *Tuberculosis*, **85**, 357–366.
Vijayan, M. & Chandra, N. (1999). *Curr. Opin. Struct. Biol.* **9**, 707–714.
Wiley, D. C., Wilson, A. I. & Skehel, J. J. (1981). *Nature (London)*, **289**, 373–378.
Zhang, R. G., Westbrook, M. L., Westbrook, E. M., Scott, D. L, Otwinowski, Z., Maulik, P. R., Reed, R. A. & Shipley, G. G. (1995). *J. Mol. Biol.* **251**, 550–562.