electronic reprint

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and molecular-replacement studies of a recombinant antigen-binding fragment complexed with single-stranded DNA

Season P. Prewitt, Andrey A. Komissarov, Susan L. Deutscher and John J. Tanner

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and molecular-replacement studies of a recombinant antigen-binding fragment complexed with single-stranded DNA

Season P. Prewitt,[a] Andrey A. Komissarov,[b] Susan L. Deutscher[b] and John J. Tanner[a,b]*

[a]Department of Chemistry, University of Missouri-Columbia, Columbia, MO 65211, USA, and [b]Department of Biochemistry, University of Missouri-Columbia, Columbia, MO 65211, USA

Correspondence e-mail: tannerjj@missouri.edu

Received 15 March 2000
Accepted 5 June 2000

Anti-DNA antibodies have been implicated in autoimmune diseases and also serve as models for understanding protein–DNA recognition. Crystals of a recombinant antigen-binding fragment (Fab) complexed with dT_5 have been obtained and initial phases have been determined using molecular replacement. The crystals diffract to 2.1 Å resolution and occupy space group $P6_522$, with unit-cell parameters a = 171.8, c = 144.6 Å; there are two Fabs per asymmetric unit. *X-PLOR* direct rotation-function calculations followed by Patterson correlation filtering were successful when using a Fab search model; however, they failed when using the individual variable and conserved domains of the Fab as search models. *AMoRe* successfully identified the correct solution in cases where *X-PLOR* failed.

1. Introduction

Anti-DNA antibodies (Abs) constitute an important and unique class of DNA-binding protein for which relatively little three-dimensional structural information is available. Two types of anti-DNA Abs, anti-single stranded (ss)DNA and anti-double stranded (ds)DNA, are of interest because they are produced *in vivo* in autoimmune disease patients (Stollar, 1981; Schwartz & Stollar, 1985). The presence of anti-dsDNA Abs in sera is diagnostic of the autoimmune disease systemic lupus erythematosus (SLE; Tan, 1989). The association of both anti-dsDNA and anti-ssDNA Abs in the pathogenesis of SLE has been suggested based on studies demonstrating immune-complex deposition in patient tissues and glomeruli (Pisetsky *et al.*, 1990; Murakami *et al.*, 1991; Winfield *et al.*, 1991; Swanson *et al.*, 1996). The resulting glomerulonephritis is the leading cause of mortality in SLE patients.

Despite the biomedical importance of anti-DNA Abs, the structural basis of Ab–DNA recognition has not been extensively characterized (Eilat & Anderson, 1994). As of January 2000, there were over 250 crystal structures of Ab fragments or intact Abs in the Protein Data Bank (PDB); however, the PDB includes only a few entries for anti-DNA Abs. Currently, there are crystal structures of two anti-ssDNA antigen-binding fragments (Fab), BV04-01 (Herron *et al.*, 1991) and HED10 (Cygler *et al.*, 1987), one anti-dsDNA Fab, Jel 72 (Mol, Muir, Lee *et al.*, 1994) and one Fab that binds triple-stranded DNA, Jel 318 (Mol, Muir, Cygler *et al.*, 1994). Structural information for Fab–DNA complexes is even more limited, with the only example being the 2.7 Å crystal structure of BV04-01/$d(pT)_3$ (Herron *et al.*, 1991). Furthermore, there are no crystal structures of recombinant anti-DNA Ab fragments in the PDB.

A more general motivation for studying anti-DNA Abs is that they serve as models for understanding protein–DNA

recognition. Abs that bind ssDNA are particularly useful in this regard because the three-dimensional structure database of protein–ssDNA complexes is quite small compared with that of protein–dsDNA complexes (Jones *et al.*, 1999; Nadassy *et al.*, 1999). Since only a few structures are available, for example, replication protein A (Bochkarev *et al.*, 1997) and helicases (Korolev *et al.*, 1997; Sawaya *et al.*, 1999; Soultanas *et al.*, 1999; Velankar *et al.*, 1999), it is impossible to draw statistically significant general conclusions about protein–ssDNA recognition in the manner that has been performed for protein–dsDNA recognition (Jones *et al.*, 1999; Nadassy *et al.*, 1999).

This paper reports the crystallization and molecular-replacement solution of a recombinant anti-ssDNA Fab complexed with oligothymidine. This Fab, denoted DNA-1, was isolated from a combinatorial phage display library of IgG fragments derived from the immunoglobulin repertoire of an autoimmune MRL/lpr mouse (Calcutt *et al.*, 1993). The VH of DNA-1 shares a high sequence homology to members of the VH 558 family that is often used in autoimmune mice. The DNA-binding characteristics of DNA-1 have been studied extensively. Fluorescence-quenching, equilibrium gel-filtration and radioimmunoassay experiments have shown that DNA-1 binds to oligo(dT) 15 nucleotides or greater in length with a K_d of 150 n*M*, although binding is also observed with oligo(dT) of length 3–5 nucleotides (Calcutt *et al.*, 1996; Deutscher *et al.*, 1996; Komissarov *et al.*, 1996; Komissarov, Marchbank, Calcutt *et al.*, 1997; Komissarov, Marchbank & Deutscher, 1997; Komissarov & Deutscher, 1999). The K_d values for polyU, polyG, poly(dC) and poly(dA) are in the range 1–10 m*M*, indicating that DNA-1 has a marked preference for oligo(dT). Mutational and thermodynamic analyses indicated that DNA-1/oligo(dT) complex formation is enthalpy driven and is primarily directed by the third heavy-chain complementarity-determining region (Komissarov, Marchbank, Calcutt *et al.*, 1997; Komissarov & Deutscher, 1999).

The present crystallographic work represents the first steps in determining the crystal structures of DNA-1 and mutant variants with and without bound ligand. These structures, in combination with mutagenesis and thermodynamics data, will contribute to a better understanding of Ab-DNA recognition, anti-DNA Ab pathogenicity and, more generally, protein–DNA recognition.

2. Materials and methods

Bacterial expression and purification of DNA-1 with a hexahistidine (His_6) tag engineered at the C-terminus of the heavy chain has been described previously (Calcutt *et al.*, 1996). Briefly, the purification involved cell disruption by sonication, centrifugation, batch affinity chromatography using a Ni–NTA resin (Qiagen) and two rounds of Resource S cation-exchange chromatography using an Akta FPLC. Preparation of DNA-1 without the His_6 tag has also been discussed (Komissarov *et al.*, 1996). Following the sonication and centrifugation steps, the untagged Fab was purified using a 10 ml goat IgG anti-mouse Fab affinity column followed by Mono-S cation-exchange chromatography using an FPLC. In both procedures, fractions from the different steps were analyzed by SDS–PAGE and the protein concentration was determined from UV absorption.

The purified proteins were dialyzed into 10–50 m*M* Tris, 50 m*M* NaCl pH 7.0 and concentrated to 16 mg ml^{-1} using a stirred ultrafiltration cell (Amicon). dT_5 was synthesized by Genosys Inc., HPLC purified, quantitated spectrophotometrically and dissolved in water to achieve a 9.4 m*M* stock solution. Prior to crystallization, the protein and DNA solutions were combined such that the final Fab concentration was approximately 14 mg ml^{-1} and the molar ratio of dT_5 to Fab was about 5–6. All crystallization experiments were performed using the sitting-drop method of vapor diffusion with drops formed by mixing 2 µl of the Fab/dT_5 solution with 2 µl of the reservoir. Crystal Screen 1 from Hampton Research was used to search for initial crystallization conditions.

X-ray diffraction data were recorded with an R-AXIS IV detector coupled to a Rigaku RU-H3R rotating-anode generator equipped with Osmic MaxFlux confocal optics. Preliminary unit-cell characterization was performed at room temperature using crystals mounted in glass capillaries. Data sets for structure determination were collected using crystals maintained at 173 K using an X-stream system. The data were processed with *HKL* (Otwinowski & Minor, 1997) and *d*TREK* as implemented in the *CrystalClear* software package (Pflugrath, 1999).

Molecular-replacement calculations were performed with the programs *X-PLOR* (Brünger, 1992) and *AMoRe* (Navaza, 1994). Since the success of Fab molecular replacement is critically dependent on starting with an approximately correct elbow angle, eight Fab search models encompassing an elbow-angle range of 138–177° were tested. The models were aligned to a common orientation before molecular replacement in order to facilitate comparison of the results. The individual CH1/CL and VH/VL domains of these Fab structures were also used as search models.

X-PLOR was used to calculate direct rotation functions (DRF) using a 5° Euler angle grid, followed by Patterson correlation (PC) filtering of the top 300 DRF peaks. The filtering calculations included optimization of the search-model orientation, followed by two-body optimization of the CH1/CL and VH/VL domains and lastly refinement of the orientations of the individual CH1, CL, VH and VL domains. The DRF and PC calculations were performed using all reflections in the 10–4 Å resolution range. Translation functions were calculated using 15–3 Å data and a fractional coordinate grid with spacing 0.008 × 0.008 × 0.01. The top translation-function solution was input to rigid-body refinement.

The stand-alone version of *AMoRe* was used in the fully automated mode. These calculations were performed using the strongest 95% of the reflections in the 10–4 Å resolution range. Rotation functions employed a 2° angular grid and the top 200 orientations were input to translation-function calculations. The translation-function calculations employed the centered overlap target function and a solution was

Table 1
Crystallization conditions and unit-cell parameters.

Form	His tag	Conditions	*a* (Å)	*b* (Å)	*c* (Å)	Space group/ crystal system
1	Yes	1.8–2.0 *M* $(NH_4)_2SO_4$, 0.1 *M* Na acetate pH 4.6–5.2	172	172	145	$P6_522$
2a	Yes	1.8–2.1 *M* $(NH_4)_2SO_4$, 0.1 *M* Na acetate pH 4.7–5.4	182	182	96	Hexagonal or trigonal†
2b	Yes	2% PEG 400, 2 *M* $(NH_4)_2SO_4$, 0.1 *M* Na HEPES pH 7.5	183	183	93	Hexagonal or trigonal†
3	Yes	25% PEG 3000 or PEG 4000, 0.2–0.3 *M* $(NH_4)_2SO_4$, 0.1 *M* Na acetate pH 4.6–5.0	61	114	155	Orthorhombic†
4	No	16–18% PEG 4000, 17–18% 2-propanol, 0.1 *M* Na citrate pH 5.9	117	117	70	Tetragonal or orthorhombic†

† Assignment of the crystal system is tentative pending collection of complete data sets and molecular-replacement calculations.

rejected if the distance between the centroids of any two search models was less than 20 Å. The top 100 translation-function solutions were input to rigid-body refinement.

3. Results and discussion

3.1. Crystallization and X-ray data collection

The His_6 tagged Fab/dT_5 complex was more amenable to crystallization than the untagged Fab/dT_5 complex, with crystals appearing at 293 K in four of the 50 Crystal Screen 1 conditions (Table 1; Fig. 1). The successful solutions included ammonium sulfate and PEG as precipitating agents and spanned the pH range 4.6–7.5. Forms 1 and 2a grew under identical conditions and sometimes appeared in the same droplet. Forms 1, 2a and 2b have hexagonal external morphologies, while form 3 crystals grew as rectangular bars. Autoindexing revealed that forms 1 and 2a are distinct crystal forms and that forms 2a and 2b are probably different manifestations of the same crystal form (Table 1). Screening of the untagged Fab/dT_5 complex yielded only a single crystal form (form 4), which appeared as rectangular blocks at both 293 and 277 K under relatively low ionic strength conditions involving PEG 4000 and 2-propanol near pH 6. Large crystals proved difficult to grow; however, crystals with dimensions of 0.2 × 0.2 × 0.2 mm could be obtained using macroseeding. Unfortunately, the diffraction was limited to 5 Å resolution with mosaicity greater than 1°.

Structure-determination efforts were focused on form 1 because it appeared to be the most promising crystal form based on its superior diffraction characteristics and the reproducibility of crystal growth. The largest crystals were hexagonal pyramids with a 0.2 mm base and a length of 0.4–0.5 mm (Fig. 1). Frozen crystals diffracted to better than 2.5 Å resolution using a rotating-anode source and diffraction to 2.1 Å resolution could be observed at NSLS beamline X8C. The space group is $P6_522$, with unit-cell parameters a = 171.8, c = 144.6 Å and two Fabs per asymmetric unit. The solvent content is approximately 60% (Matthews, 1968).

Crystals of form 1 were prepared for low-temperature X-ray data collection by replacing the mother liquor with solutions of the harvest buffer (2 *M* ammonium sulfate, 0.1 *M* sodium acetate pH 5.0) supplemented with increasing amounts of glycerol. The glycerol concentration was raised from zero to 30% in increments of 5% over a period of about 10 min; the crystal was then plunged into liquid nitrogen.

A data set was collected at 173 K using an R-AXIS IV detector and a rotating-anode source. The crystal-to-detector distance was 210 mm and the detector θ was zero. The data collection consisted of 134 frames, with an oscillation angle of 0.5° and an exposure time of 15 min. This data set is 98% complete to 2.5 Å resolution, with an $R_{\rm merge}$ on I of 0.076 and an $\langle I/\sigma(I)\rangle$ of 19. The 2.6–2.5 Å shell of data is 94% complete, with $R_{\rm merge}$ = 0.273 and $\langle I/\sigma(I)\rangle$ = 9. The average crystal mosaicity is 0.43° and the average multiplicity is 8.0.

3.2. Molecular replacement for form 1

Two of the eight Fab search models led to unambiguous cross rotation function solutions in *X-PLOR*. The clearest result was obtained with the anti-lysozyme Fab D44.1 (Braden *et al.*, 1994; PDB entry 1mlc), which has an elbow angle of 165°. The anti-lysozyme Fab D1.3 (Fischmann *et al.*, 1991), which has an elbow angle of 169°, produced similar results. The results of PC filtering for the 1mlc Fab search model are shown in Fig. 2(*a*). Two features are notable in Fig. 2(*a*). Firstly, the PC filtering indicates only one orientation having a PC value significantly above the noise level (7σ), which suggests the two Fab molecules in the asymmetric unit are related by a translation vector. Secondly, the correct solution identified by PC filtering is the eighth peak in the DRF peak list, which demonstrates the utility of post-processing the DRF with PC filtering. The Fab search model was oriented according to the 7σ PC solution and input to translation-function calculations using *X-PLOR*. The top translation-function solution with two Fabs per asymmetric unit had an R factor of 0.400 for 15–3.0 Å data after rigid-body refinement, which is indicative of a correct molecular-replacement solution.

In contrast to the results obtained for the Fab D44.1 search model, the CH1/CL and VH/VL domains of D44.1 failed to produce an unambiguous rotation-function solution in *X-PLOR*. The results of PC filtering for the domains are shown in Figs. 2(*b*) and 2(*c*). Whereas PC filtering with the Fab search model produced a clear solution with a PC value 7σ above the mean (Fig. 2*a*), none of the peaks in Figs. 2(*b*) and 2(*c*) can be convincingly distinguished from the noise. Interestingly, the correct orientations of the CH1/CL and VH/VL domains do in fact appear within the top 300 DRF peaks. To investigate the failure of PC refinement, the domains were

oriented according the 7σ solution and input to PC refinement. The correctly oriented CH1/CL and VH/VL domains registered PC values of 0.0376 and 0.0306, respectively. Note from Figs. 2(*b*) and 2(*c*) that these values were not significantly higher than those of incorrectly oriented domains. It appears that the individual domains had too little scattering power to be detected by PC refinement. On the other hand, one entire Fab molecule, which represents 4% of the protein in the unit cell, apparently was sufficient to register a significant PC signal.

Calculations using *AMoRe* were performed concurrently with the *X-PLOR* calculations. In contrast to the *X-PLOR* results, *AMoRe* calculations using the CH1/CL and VH/VL domains of D44.1 were successful. A clear solution, indicated by an *R* factor of 0.402 and a correlation coefficient of 0.510, was obtained in space group $P6_522$ with two CH1/CL domains and two VH/VL domains per asymmetric unit. Inspection of the *AMoRe* solution revealed that it was identical to the *X-PLOR* solution after accounting for crystallographic symmetry.

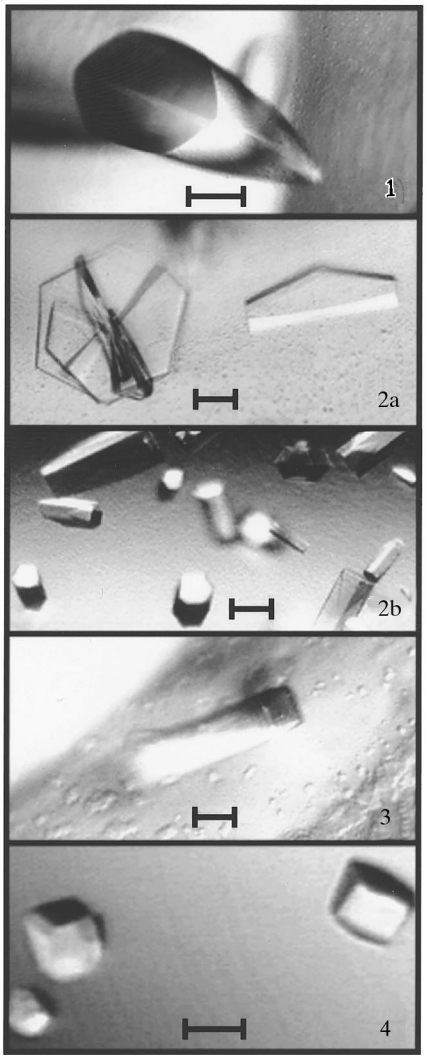

Figure 1
Crystal forms of DNA-1 grown in the presence of dT_5. The scales indicate 0.1 mm.

Thus, both the *X-PLOR* DRF/PC and *AMoRe* approaches could be used to solve the molecular-replacement problem. However, *X-PLOR* required Fab search models while *AMoRe* succeeded with smaller fragments. The fact that *AMoRe* succeeded where the DRF/PC approach failed is interesting because the DRF method has been shown to have a higher signal-to-noise ratio than other molecular-replacement methods, including *AMoRe* (Brünger, 1997). These results underscore the utility of using different programs and approaches to solve difficult molecular-replacement problems. Additionally, when relatively small fragments are employed as search models, automated approaches in which many orientations are input to translation functions, such as *AMoRe* and full six-dimensional searches (Sheriff *et al.*, 1999), are particularly useful.

The structure has been refined using *CNS* (Brunger *et al.*, 1998) to a working *R* factor of 0.218 and a test *R* factor (10% random set) of 0.270 for 29–2.5 Å data. Electron-density maps indicate the presence of DNA in the combining site, with the prominent Fab–DNA interaction being the stacking of Tyr side chains against thymine rings. Current efforts are being focused on refinement of the model and collection of higher resolution data using synchrotron radiation.

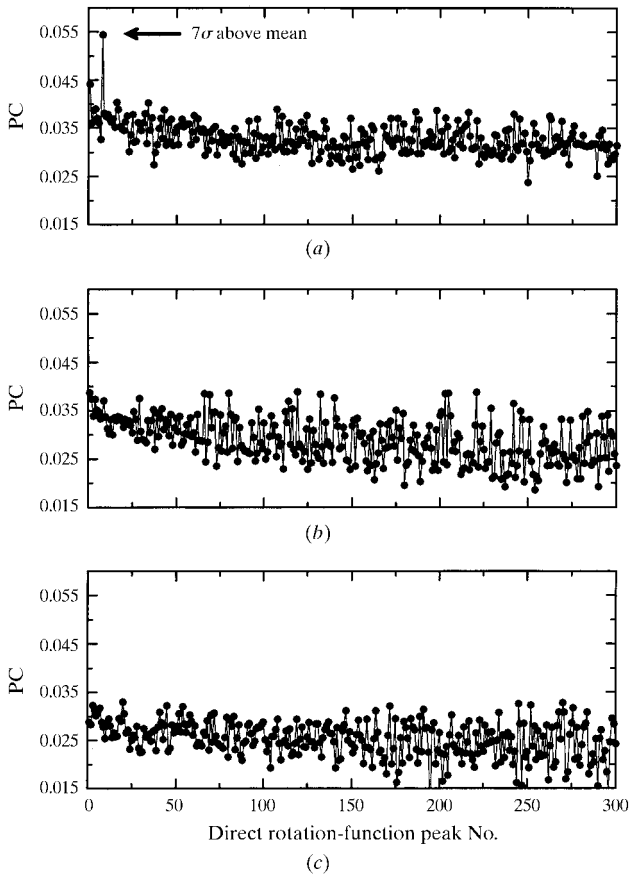

Figure 2
Results of Patterson correlation filtering of direct rotation function peaks for (*a*) a Fab search model, (*b*) a CH1/CL search model and (*c*) a VH/VL search model. The search models were derived from Protein Data Bank entry 1mlc (Braden *et al.*, 1994). Note that only the Fab search model produces a convincing cross rotation function solution.

We thank Dr Steven Sheriff for helpful discussions. This research was supported in part by grants from the University of Missouri Research Board to JJT and SLD. Part of this research was carried out at the National Synchrotron Light Source, Brookhaven National Laboratory, which is supported by the US Department of Energy Contract No. DEAC02-76CH00016, Division of Materials Sciences and Division of Chemical Sciences.

References

Bochkarev, A., Pfuetzner, R. A., Edwards, A. M. & Frappier, L. (1997). *Nature (London)*, **385**, 176–181.
Braden, B. C., Souchon, H., Eisele, J. L., Bentley, G. A., Bhat, T. N., Navaza, J. & Poljak, R. J. (1994). *J. Mol. Biol.* **243**, 767–781.
Brünger, A. T. (1992). *X-PLOR Version* 3.1. *A System for X-ray Crystallography and NMR.* New Haven: Yale University Press.
Brünger, A. T. (1997). *Methods Enzymol.* **276**, 558–580.
Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Calcutt, M. J., Komissarov, A. A., Marchbank, M. T. & Deutscher, S. L. (1996). *Gene*, **168**, 9–14.
Calcutt, M. J., Kremer, M. T., Giblin, M. F., Quinn, T. P. & Deutscher, S. L. (1993). *Gene*, **137**, 77–83.
Cygler, M., Boodhoo, A., Lee, J. S. & Anderson, W. F. (1987). *J. Biol. Chem.* **262**, 643–648.
Deutscher, S. L., Crider, M. E., Ringbauer, J., Komissarov, A. A. & Quinn, T. P. (1996). *Arch. Biochem. Biophys.* **1**, 207–213.
Eilat, D. & Anderson, W. F. (1994). *Mol. Immunol.* **31**, 1377–1390.
Fischmann, T. O., Bentley, G. A., Bhat, T. N., Boulot, G., Mariuzza, R. A., Phillips, S. E., Tello, D. & Poljak, R. J. (1991). *J. Biol. Chem.* **266**, 12915–12920.
Herron, J. N., He, X. M., Ballard, D. W., Blier, P. R., Pace, P. E., Bothwell, A. L. M., Voss, E. W. Jr & Edmundson, A. B. (1991). *Proteins Struct. Funct. Genet.* **11**, 159–175.
Jones, S., van Heyningen, P., Berman, H. M. & Thornton, J. M. (1999). *J. Mol. Biol.* **287**, 877–96.
Komissarov, A. A., Calcutt, M. J., Marchbank, M. T., Peletskaya, E. N. & Deutscher, S. L. (1996). *J. Biol. Chem.* **271**, 12241–12246.
Komissarov, A. A. & Deutscher, S. L. (1999). *Biochemistry*, **38**, 14631–14637.
Komissarov, A. A., Marchbank, M. T., Calcutt, M. J., Quinn, T. P. & Deutscher, S. L. (1997). *J. Biol. Chem.* **272**, 26864–26870.
Komissarov, A. A., Marchbank, M. T. & Deutscher, S. L. (1997). *Anal. Biochem.* **247**, 123–129.
Korolev, S., Hsieh, J., Gauss, G. H., Lohman, T. M. & Waksman, G. (1997). *Cell*, **90**, 635–647.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mol, C. D., Muir, A. K., Cygler, M., Lee, J. S. & Anderson, W. F. (1994). *J. Biol. Chem.* **269**, 3615–3622.
Mol, C. D., Muir, A. K., Lee, J. S. & Anderson, W. F. (1994). *J. Biol. Chem.* **269**, 3605–3614.
Murakami, H., Lam, Z., Furie, B. C., Reinhold, V. N., Asano, T. & Furie, B. (1991). *J. Biol. Chem.* **266**, 15414–15419.
Nadassy, K., Wodak, S. J. & Janin, J. (1999). *Biochemistry*, **38**, 1999–2017.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Pisetsky, D. S., Grudier, J. P. & Gilkeson, G. S. (1990). *Arthritis Rheum.* **33**, 153–159.
Sawaya, M. R., Guo, S., Tabor, S., Richardson, C. C. & Ellenberger, T. (1999). *Cell*, **99**, 167–177.
Schwartz, R. S. & Stollar, B. D. (1985). *J. Clin. Invest.* **75**, 321–332.
Sheriff, S., Klei, H. E. & Davis, M. E. (1999). *J. Appl. Cryst.* **32**, 98–101.
Soultanas, P., Dillingham, M. S., Velankar, S. S. & Wigley, D. B. (1999). *J. Mol. Biol.* **290**, 137–148.
Stollar, B. D. (1981). *Clin. Immunol. Allergy*, **1**, 243–260.
Swanson, P. C., Ackroyd, C. & Glick, G. D. (1996). *Biochemistry*, **35**, 1624–1633.
Tan, E. M. (1989). *Adv. Immunol.* **44**, 93–151.
Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. (1999). *Cell*, **97**, 75–84.
Winfield, J. B., Faiferman, I. & Koffler, D. (1991). *J. Clin. Invest.* **59**, 90–95.