Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A new conformation of the integrin-binding fragment of human VCAM-1 crystallizes in a highly hydrated packing arrangement

Paul Taylor, Morag Bilsland and Malcolm D. Walkinshaw*

Structural Biochemistry Group, Institute of Cell and Molecular Biology, The University of Edinburgh, Michael Swann Building, King's Buildings, Edinburgh EH9 3JR, Scotland

Correspondence e-mail: m.walkinshaw@ed.ac.uk

Received 23 February 2001
Accepted 4 July 2001

PDB Reference: VCAM integrin-binding fragment, 1ij9.

An X-ray crystal structure of two N-terminal integrin-binding IgSF domains of human VCAM-1 is reported. This new crystal form shows an unusual and highly hydrated packing arrangement in which over 80% of the crystal is occupied by solvent. The relative orientations of the two domains adopt a new intermediate conformation. The tilt angle between the two domains is 19.4°, compared with other related structures that have tilt angles ranging from 7.3 to 39.9°. An analysis of the torsion angles shows that residues Ile88, Tyr89, Ser90, Pro92 and Glu96 play a major role in defining the interdomain conformations.

1. Introduction

Vascular cell adhesion molecule 1 (VCAM-1/CDM8; Osborn *et al.*, 1989; INCAM-110; Rice *et al.*, 1990) is a cell-surface receptor molecule that represents an integrin-binding subgroup of the immunoglobulin superfamily (IgSF; Wang *et al.*, 1995). VCAM-1 preferentially binds the integrin heterodimer of the very late antigen (VLA-4) β1 family (Wang & Springer, 1998) containing α4 (CD49d) and β1 (CD29) subunits (Vonderheide & Springer, 1992). The VCAM-1–integrin interaction mediates intercellular adhesion of leukocytes to the blood-vessel wall (Osborn *et al.*, 1989) and also regulates trans-endothelial emigration at inflammatory sites (Wang & Springer, 1998). VCAM-1 may also act as a host receptor for viruses and parasites (Huber, 1994), as well as having important implications in tumour formation and metastasis (Rice *et al.*, 1990).

The extracellular region of VCAM-1 typically consists of seven contiguously arranged IgSF domains (VCAM-7D; Chothia & Jones, 1997) that adopt an extended rod-like structure (Osborn *et al.*, 1994). Ligand-binding surfaces specific to VLA-4 integrin are exposed on homologous IgSF domains 1 and 4 of VCAM-7D (Wang & Springer, 1998). X-ray crystal structures are available for a functional ligand-binding fragment of domains 1 and 2 of VCAM-1 (Jones *et al.*, 1995; Wang *et al.*, 1995, 1996) and ligand-binding domains of related immune-cell adhesion molecules of the subfamily, including ICAM-1 (Casasnovas *et al.*, 1998; Bella *et al.*, 1998), ICAM-2 (Casasnovas *et al.*, 1997), MAdCAM-1 (Tan *et al.*, 1998), CD2 (Jones *et al.*, 1992; Bodian *et al.*, 1994) and CD4 (Brady *et al.*, 1993; Ryu *et al.*, 1994). Here, we present a novel crystal form of VCAM-1 that adopts a different conformation and exists in an exceptionally highly hydrated packing arrangement.

2. Materials and methods

2.1. Crystallization

Purified protein was a gift from Dr P. Lake. Crystals were grown at room temperature using the hanging-drop method and could be obtained using either polyethylene glycol or ammonium sulfate as the precipitant. Large hexagonal crystals used for data collection were obtained from 10 µl hanging drops. These drops comprised 5 µl of protein solution (10 mg ml^{-1} in 50 m*M* HEPES buffer pH 7.5) added to a 5 µl drop of well solution [50 m*M* Tris buffer pH 7.1, 42%(*v*/*v*) saturated ammonium sulfate].

2.2. Structure determination

A summary of the crystallographic data is given in Table 1. Crystals of dimensions 1.0 × 1.0 × 0.8 mm were mounted in thin glass Lindemann tubes. X-ray data were collected at room temperature using a MAR image plate and processed using *DENZO* (Otwinowski & Minor, 1993). The structure was solved by molecular replacement using the program *AMoRe* (Navaza, 1994), with one domain of the VCAM structure as the model (Jones *et al.*, 1995). Positional and *B*-factor refinement were performed using *X-PLOR* Version 3.1 (Brünger, 1993); water molecules were placed on appropriate peaks in subsequent difference Fourier maps.

Table 1
Crystallographic data.

Unit-cell parameters (Å,°)	$a = b = 152.94$, $c = 45.98$, $\alpha = \beta = 90$, $\gamma = 120$
Space group	$P3_121$
Z (No. of molecules in the unit cell)	6
d_{min} (Å)	3.0 (3.05–3.0)
Total No. of observed reflections†	183518 (5762)
No. of unique reflections†	12530 (622)
Merging R factor† (%)	11.5 (59.6)
Multiplicity†	9.4 (9.3)
Completeness† (%)	99.9 (100)
% > $I/3\sigma(I)$†	74.8 (55.4)
Final R‡ (%)	20.8
Final R_{free}‡ (%)	23.8
R.m.s.d.	
Bonds (Å)	0.006
Angles (°)	1.31
Ramachandran plot	
Most favoured region (%)	82.7
Allowed (%)	16.2
Disallowed (%)	1.2
B values (Å^2)	
Average B value for protein (s.d.)	45.03 (20.81)
Average B value for main chain (s.d.)	42.00 (18.60)
Average B value for side chain (s.d.)	48.20 (22.47)
Average B value for the 50 waters (s.d.)	61.58 (19.56)

† Values in parentheses are for the highest resolution shell of data. ‡ R factors were calculated using all data in the resolution range 24.0–3.00 Å.

2.3. Density measurements

Aqueous stock solutions of Ficoll 400 (Pharmacia) were made up at 60, 50, 40, 30, 20 and 10%(*w*/*w*). Aliquots of approximately 40 µl were layered into a glass tube with an internal diameter of 2 mm. The gradient was calibrated using 5 µl drops of water-saturated chloroform/toluene mixtures. The calibration solutions were saturated with water and their density was measured using a 10 ml pycnometer to an estimated accuracy of 0.001 g ml^{-1}. The tube with the calibration droplets was spun at 4000*g* for 5 min.

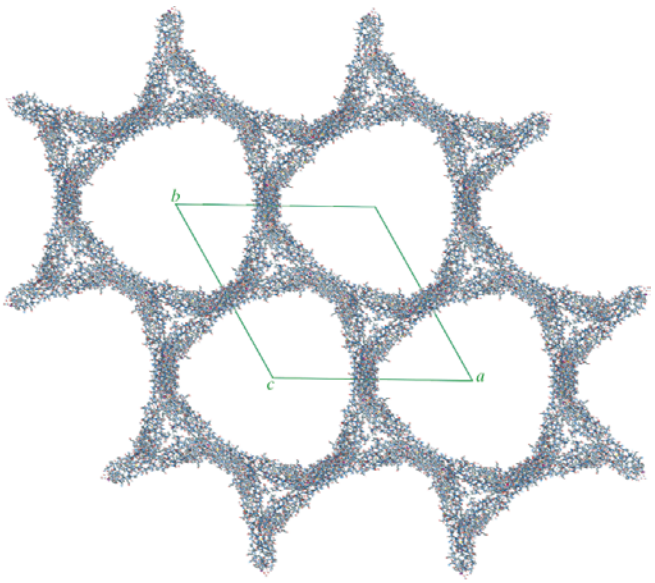

Figure 1
A highly hydrated packing arrangement in VCAM-1 shown in projection along *c*. The unit cell is shown.

3. Results and discussion

3.1. X-ray structure and solvent channels

Statistics for the refined structure are given in Table 1. The packing diagram for this form of VCAM (Fig. 1) shows a spectacularly large solvent channel with V_M = 3.44 Å^3 Da^{-1} and over 80% of the crystal being occupied by solvent. In order to show that lack of electron density in the crystal channels was indeed arising from solvent and not disordered VCAM, density measurements of the crystal using the Ficoll method were carried out.

The single crystal of VCAM which had been used for X-ray intensity measurements was introduced onto the top of the Ficoll gradient tube and within 30 s sank to a point corresponding to a crystal density (d) of 1.12 g cm^{-3}. After centrifugation for 5 min at 4000*g* the crystal sank to a point corresponding to 1.135 g cm^{-3}. The crystal continued to drop in the gradient tube until stabilizing after about 60 min with an apparent density of 1.17 g cm^{-3}. This increase of crystal density with time has been commonly observed (Dragovich *et al.*, 1996; Borgstahl *et al.*, 1996) and is attributed to the incorporation of Ficoll into the lattice. The number of molecules per unit cell (n) can be calculated from (Westbrook, 1985)

$$n = (NV/Mv_p)(\rho_c - \rho_s)/[(1/v_p - \rho_S)],$$

Table 2
Angles (°) that define interdomain motion of VCAM-1.

			Ile88		Tyr89		Ser90		Pro92		Glu96	
Name	Tilt	Skew	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ
VCAM (1ij9)	19.4	77.7	−112.3	162.0	−160.3	151.7	−128.5	120.0	−66.7	−45.5	−121.0	134.0
1vca (*A*)	39.9	85.5	−109.6	156.9	−164.0	170.1	−144.3	125.7	−77.2	−24.2	−117.3	131.7
1vca (*B*)	33.9	87.8	−115.9	150.9	−152.3	168.4	−145.5	121.9	−76.4	−22.0	−121.9	129.5
1vsc (*A*)	16.7	70.8	−131.4	153.2	−150.2	162.2	−146.0	129.1	−82.2	−8.6	−128.0	117.3
1vsc (*B*)	7.3	55.3	−127.2	149.3	−130.6	153.4	−149.6	119.4	−88.9	−14.2	−134.3	134.1

where V is the volume of the unit cell (930 600 Å^3) and M is the molecular weight (22 491 Da); ρ_c is the measured crystal density and ρ_s is the density of the total solvent compartment. The protein partial specific volume (ν_p) was taken as 0.737 cm^3 g^{-1} (Bode & Schirmer, 1985). The initial concentration of the hanging drop was 20% saturated ammonium sulfate (ρ_s = 1.117 g cm^{-3}). Protein crystals are modelled as a lattice of solvated protein surrounded by channels of free solvent (Westbrook, 1985). The calculation of n is very sensitive to the difference between ρ_c and ρ_s. With an initial measured crystal density of 1.12 g cm^{-3}, n is calculated to be 11.5, 6.0 and 0.3 for ρ_s values of 1.0 (water), 1.07 and 1.117 g cm^{-3} (20% saturated ammonium sulfate), respectively. The value of n = 6, required for space group $P3_121$, is consistent with the large solvent channels being partially filled with ammonium sulfate solution. This is supported by the crystallographic refinement, which gives no trace of connected electron density in the solvent channels.

3.2. Interdomain flexibility

IgSF domains of VCAM-1 and homologous members of the Ig subfamily are comprised of two antiparallel β-pleated sheets packed face-to-face (Jones, 1993). The common core structure is stabilized by interconnecting loops and disulfide bridges (Lesk & Chothia, 1982). A connecting linker peptide (residues 90–95) mediates the intramolecular dimerization of VCAM-1 Ig domains, with the last β-sheet of domain 1 joining to the first β-sheet of domain 2 (Jones, 1993) (Fig. 2). The linker peptide mediates interdomain flexibility at the hydro-

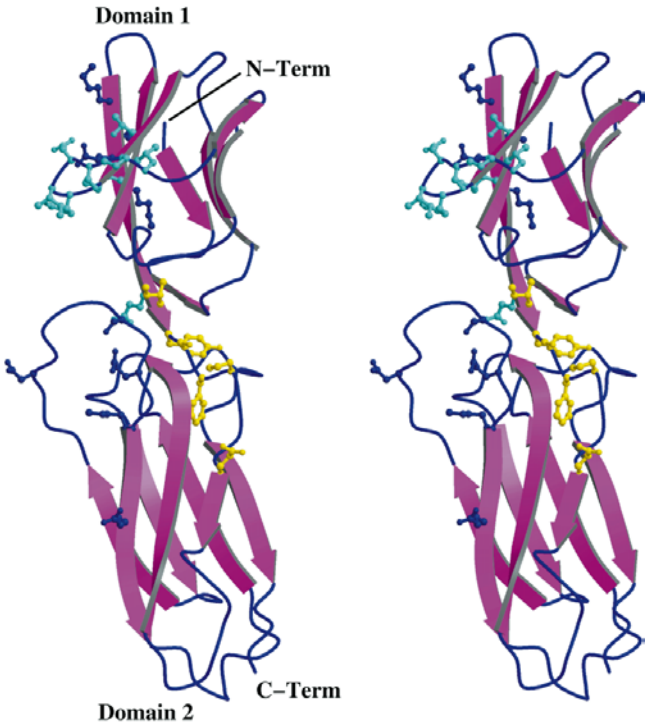

Figure 2
MOLSCRIPT representation of VCAM-1. Antiparallel β-strands of Ig domains 1 and 2 (magenta) are stabilized by interconnecting loops (blue). The structure of VCAM-1 is orientated to reveal the extensive integrin-binding surface. The key residues involved in basal levels of integrin binding (cyan; Gln38, Ile39, Asp40, Ser41, Pro42, Leu43, Leu70, Thr72, Glu87) map to the protruding *CD* loop and nearby *F* and *G* strands of domain 1. Residues that form specific $\alpha4\beta1$ and $\alpha4\beta7$ integrin contacts (purple; Lys79, Leu80, Glu81, Lys82, Asp143, Ser148, Glu150, Lys152 and Glu155) reside in domains 1 and domain 2, respectively. The linker peptide (blue) connects Ig domains 1 and 2. Key linker residues (yellow; Ile88, Tyr89, Ser90, Phe91, Pro92 and Glu96) proposed to mediate interdomain movement are shown.

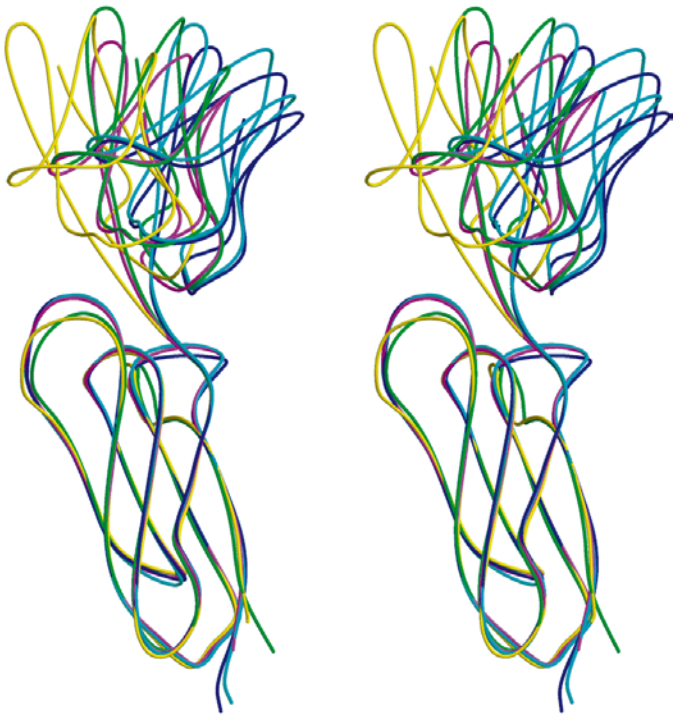

Figure 3
A *MOLSCRIPT* stereo overlay of the five known independent conformations of VCAM-1. The secondary structure of VCAM was defined using the program *DSSP*. Each of the five VCAM-1 monomers were superimposed with exclusion of residues 140–149 (*C′E* loop) and 176–185 (*FG* loop) to give a structural overlay of domain 2 backbone atoms. The domain 2 overlay of available VCAM-1 X-ray structures demonstrates the rigid-body motion of domain 1 relative to domain 2. Our structural form (magenta) represents an intermediate conformation of other VCAM-1 structures, 1vsc monomers *A* and *B* (green and yellow) and 1vca monomers *A* and *B* (blue and purple).

phobic interface of domains 1 and 2. The average *B* factor of 42.0 Å^2 (s.d. 18.6 Å^2) for main-chain atoms in this structure solved at room temperature is rather high (Table 1) and only 50 solvent molecules were located. Despite this, the linker region (residues 88–96) between the two domains is rather well defined and has an average main-chain *B* factor of 36.4 Å^2 (s.d. 4.51 Å^2) and slightly more mobile side chains with an average *B* factor of 40.4 Å^2 (s.d. 12.1 Å^2). All φ and ψ angles of this linker are found to lie within the allowed regions of the Ramachandran plot.

The analysis of dihedral angles measured from all available structural results (Table 2) suggests that interdomain flexibility is not uniquely determined by the φ,ψ angles of any one linker residue but is rather determined by changes in the conformation of some or all of the residues in the sequence Ile88–Glu96. In particular, Ile88, Tyr89, Ser90, Pro92 and Glu96 play a major role. The φ angle of Ser90 of -128.5° for the new VCAM-1 form varies considerably from the narrow range of values obtained for the other crystal forms (-144.3 to -149.6°). The φ and ψ rotation angles of Pro92 of -66.7 and -45.5°, respectively, both deviate from values obtained for the other four monomers (Table 2). Previous analyses attributed interdomain flexibility to changes in the conformation of the pivot residue Tyr89 (Wang *et al.*, 1995) or Ile88, Pro92 and Glu96 (Jones *et al.*, 1995).

A least-squares fit of main-chain atoms of domain 2 (residues 96–196) provides an optimized structural alignment for each of the five crystallographically independent VCAM-1 structures (Table 3; Fig. 3). The 'tilt' angle between the domains (Table 2) is calculated as the angle between the two lines defining the principal axes for each of the domains (Bork *et al.*, 1996). The tilt angle is essentially two-dimensional and, in association with the 'skew' angle, defines the orientation of the domains. The tilt and skew values obtained for the VCAM structure presented here (19.4°, 77.7°) lie between the tilt (7.3–39.9°) and skew (55.3–87.8°) angles measured for the other four monomers. The VCAM structure reported here adopts a unique conformation, as evident from the tilt and skew angles which demonstrate projection of domain 1 with respect to domain 2. This orientation provides an intermediate VCAM conformation that facilitates interaction of domain 1 with both the flexible *C′E* and *FG* loops of domain 2.

The torsion angles of the linker residues may in part be governed by the interdomain contacts. Indeed, the least-squares fit of the domains of each of the five structures shows that there are large differences in the conformations of the longer *FG* and *C′E* loops (Table 3). The inherently mobile *C′E* and *FG* loops located at the hydrophobic interface of Ig domains 1 and 2 are likely to change conformation in order to accommodate any changes in the relative positions of domains 1 and 2 (Chothia & Jones, 1997).

3.3. Integrin binding

The current model of integrin binding implicates domain 1 as the critical domain for binding, with domain 2 providing additional contact residues and conformational stabilization.

Table 3
Domain 2 pairwise superposition of VCAM structures.

Template	Superimposed molecule	R.m.s. deviation (Å)
VCAM (1ij9)	1vca (*A*)	0.394
VCAM (1ij9)	1vca (*B*)	0.443
VCAM (1ij9)	1vsc (*A*)	0.862
VCAM (1ij9)	1vsc (*B*)	0.858

Mutagenesis experiments revealed the general integrin-binding motif of VCAM-1 as the sequence Q38IDSPL located on the loop connecting β-strands *C*1 and *D*1 of the first domain of VCAM-1 (Fig. 2; Newham *et al.*, 1997). Asp40 is a key integrin-binding residue located at the tip of the *CD* loop. It provides a negative charge that may complete a divalent cation-binding site in the counter-receptor molecule (Jones *et al.*, 1995). Mutagenesis studies have implicated the sequence G64NEH exposed on the *EF* loop in α4 integrin binding to domain 1 of VCAM-1 (Osborn *et al.*, 1994). The *EF* loop of domain 1 lies in close proximity to the *CD* loop to form a general integrin-presenting surface (Wang *et al.*, 1996) positioned near the interface of domains 1 and 2 (Wang *et al.*, 1995).

Mutational analysis provides the first evidence for the existence of *C′E* loop–integrin interactions (Newham *et al.*, 1997) that support findings of VCAM-1 interdomain contacts. Residues of the *C′E* loop (Asp143, Ser148) and *E* β-strand (Glu150, Lys152 and Glu155) create a negatively charged surface required for specific binding to $\alpha4\beta7$ integrin. Sequence analysis suggests equivalent negatively charged residues are a common requirement for $\alpha4\beta7$ binding to VCAM-1 and MAdCAM-1 (Newham *et al.*, 1997). Specific integrin binding of $\alpha4\beta1$ to VCAM-1 maps to residues K(79)LEK that reside in the *G* β-strand of domain 1 (Kilger *et al.*, 1997) (Fig. 2).

The structural data presented here provides further insight into the mechanism of adhesive interactions in the Ig subfamily of integrin-binding proteins. The conformation of our X-ray crystal structure demonstrates interdomain motion stabilized by the extensive loops of domain 2.

We thank Dr P. Lake from Sandoz AG for providing protein.

References

Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M. & Rossmann, M. G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4140–4145.
Bode, W. & Schirmer, T. (1985). *Biol. Chem. Hoppe-Seyler*, **366**, 287–295.
Bodian, D. L., Jones, E. Y., Harlos, K., Stuart, D. I. & Davis, S. J. (1994). *Structure*, **2**, 755–766.
Borgstahl, G. E. O., Parge, H. E., Hickey, M. J., Johnson, M. J., Boissinot, M., Hallewell, R. A., Lepock, J. R., Cabelli, D. E. & Tainer, J. A. (1996). *Biochemistry*, **35**, 4287–4297.
Bork, P., Downing, A. K., Kieffer, B. & Campbell, I. D. (1996). *Q. Rev. Biophys.* **29**, 119–167.

Brady, R. L., Dodson, E. J., Dodson, G. G., Lange, G., Davis, S. J., Williams, A. F. & Barclay, A. N. (1993). *Science*, **260**, 979–983.
Brünger, A. T. (1993). *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*. New Haven & London: Yale University Press.
Casasnovas, J. M., Springer, T. A., Liu, J. H., Harrison, S. C. & Wang, J. H. (1997). *Nature (London)*, **387**, 312–315.
Casasnovas, J. M., Stehle, T., Liu, J. H., Wang, J. H. & Springer, T. A. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 4134–4139.
Chothia, C. & Jones, E. Y. (1997). *Annu. Rev. Biochem.* **66**, 823–862.
Dragovich, P. S., Barker, J. E., French, J., Imbacuan, M., Kalish, V. J., Kissinger, C. R., Knighton, D. R., Lewis, C. T., Moomaw, E. W., Parge, H. E., Pelletier, L. A. K., Prins, T. J., Showalter, R. E., Tatlock, J. H., Tucker, K. D. & Villafranca, J. E. (1996). *Abstr. Papers Am. Chem. Soc.* **211**, 206-MEDI.
Huber, S. A. (1994). *J. Virol.* **68**, 3453–3458.
Jones, E. Y. (1993). *Curr. Opin. Struct. Biol.* **3**, 846–852.
Jones, E. Y., Davis, S. J., Williams, A. F., Harlos, K. & Stuart, D. I. (1992). *Nature (London)*, **360**, 232–239.
Jones, E. Y., Harlos, K., Bottomley, M. J., Robinson, R. C., Driscoll, P. C., Edwards, R. M., Clements, J. M., Dudgeon, T. J. & Stuart, D. I.. (1995). *Nature (London)*, **373**, 539–544.
Kilger, G., Clements, J. & Holzmann, B. (1997). *Int. Immunol.* **9**, 219–226.
Lesk, A. M. & Chothia, C. (1982). *J. Mol. Biol.* **160**, 325–342.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Newham, P., Craig, S. E., Seddon, G. N., Schofield, N. R., Rees, A., Edwards, R. M., Jones, E. Y. & Humphries, M. J. (1997). *J. Biol. Chem.* **272**, 19429–19440.
Osborn, L., Hession, C., Tizard, R., Vassallo, C., Luhowskyj, S., Chirosso, G. & Lobb, R. (1989). *Cell*, **59**, 1203–1211.
Osborn, L., Vassallo, C., Browning, B. G., Tizard, R., Haskard, D. O., Benjamin, C. D., Dougas, I. & Kirchhausen, T. (1994). *J. Cell Biol.* **124**, 601–608.
Otwinowski, Z. & Minor, W. (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.
Rice, G. E., Munro, J. M. & Bevilacqua, M. P. (1990). *J. Exp. Med.* **171**, 1369–1374.
Ryu, S. E., Truneh, A., Sweet, R. W. & Hendrickson, W. A. (1994). *Structure*, **2**, 59–74.
Tan, K., Casasnovas, J. M., Liu, J. H., Briskin, M. J., Springer, T. A. & Wang, J. H. (1998). *Structure*, **6**, 793–801.
Vonderheide, R. H. & Springer, T. A. (1992). *J. Exp. Med.* **175**, 1433–1442.
Wang, J. H., Pepinsky, R. B., Stehle, T., Liu, J. H., Karpusas, M., Browning, B. & Osborn, L. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 5714–5718.
Wang, J. H. & Springer, T. A. (1998). *Immunol. Rev.* **163**, 197–215.
Wang, J. H., Stehle, T., Pepinsky, B., Liu, J. H., Karpusas, M. & Osborn, L. (1996). *Acta Cryst.* D**52**, 369–379.
Westbrook, E. M. (1985). *Methods Enzymol.* **114**, 187–196.