protein structure communications

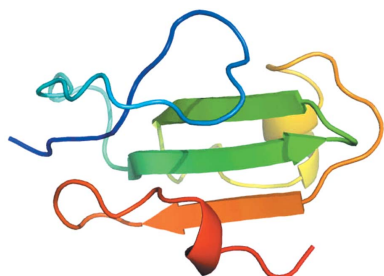# High-resolution structures of bacterially expressed soluble human CD59

Kirstin J. Leath,[a] Steven Johnson,[a] Pietro Roversi,[a] Timothy R. Hughes,[b] Richard A. G. Smith,[c] Lloyd Mackenzie,[d] B. Paul Morgan[b] and Susan M. Lea[a]*

[a]Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, England, [b]Department of Medical Biochemistry and Immunology, School of Medicine, Cardiff University, Cardiff CF14 4XN, Wales, [c]Department of Medicine, University of Cambridge, Level 5 Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, England, and [d]Inflazyme Pharmaceuticals, 425-5600 Parkwood Way, Richmond, British Columbia, V6V 2M2, Canada

Correspondence e-mail:
susan.lea@path.ox.ac.uk

CD59 is a membrane-bound glycoprotein that protects host cells from lysis by inhibiting the terminal pathway of complement, preventing the formation and insertion of the membrane attack complex (MAC). Crystals of bacterially expressed and nonglycosylated recombinant soluble human CD59 have been obtained from three crystallization conditions, each of which gave rise to a distinct crystal form. Each crystal form led to a crystal structure at high resolution (1.15, 1.35 and 1.8 Å). In one of these structures the electron-density map shows an as yet unidentified small molecule in the predicted C8/C9-binding site. The presence/absence of this ligand is linked to alternate conformations of the amino acids implicated in C8/C9 binding.

1. Introduction

The complement system straddles the innate and adaptive immune system and is a complex biochemical chain reaction leading to opsonization, cell lysis and inflammation. The terminal complement pathway results in assembly of the membrane attack complex, formed by complement proteins C5, C6, C7, C8 and C9, and its insertion through the membrane of the target cell to form a lytic pore. While an important weapon in the immunological arsenal, the nonspecific nature of complement action renders it a double-edged sword. Host tissues must express a number of membrane-bound and fluid-phase complement regulators in order to protect themselves from complement attack. Characterized human complement regulators include CD55, CD46, CD35 and CD59. CD55, CD46 and CD35 all inhibit complement by restricting C3 cleavage and deposition; CD59, however, protects self-cells from complement-mediated lysis by inhibiting the formation and insertion through the membrane of the MAC. In so doing, CD59 protects its host cell from bystander complement attack.

CD59 is a 77-amino-acid glycosylphosphatidylinositol (GPI) anchored membrane-bound regulator of the terminal complement pathway. Found on the surface of a wide range of cells, CD59 binds to C5b-8 and C5b-9, preventing further C9 attachment and inhibiting the final stages of assembly of the membrane attack complex (MAC) on the host cell (Farkas *et al.*, 2002; Ninomiya & Sims, 1992; Rollins & Sims, 1990; Meri *et al.*, 1990; Chang *et al.*, 1993). Two NMR structures of mammalian-expressed CD59 are currently available (Kieffer *et al.*, 1994; Fletcher *et al.*, 1994), revealing it to be a globular β-stranded domain (SCOP classification snake toxin-like, 57301). CD59 has been implicated in providing protection for cancerous tumour cells during monoclonal antibody immunotherapy (Chen *et al.*, 2000; Gelderman *et al.*, 2004) and has been found to be upregulated in some cancers (Caragine *et al.*, 1998; Fishelson *et al.*, 2003; Murray *et al.*, 2000). A more comprehensive molecular understanding of the CD59 binding site and of its mode of action in preventing complement attack could lead to drugs to increase the effectiveness of monoclonal antibody immunotherapy. It would also have the potential to lead to therapeutics for autoimmune and inflammatory diseases (Morgan &

Harris, 2003), as well as drugs designed to minimize the rejection of xenotransplanted tissue (Chen *et al.*, 2001). As an example of the potential of CD59 to be an effective biopharmaceutical agent in its own right, the human protein used in these studies has been employed in a membrane-targeted form to reverse the phenotype of paroxysmal nocturnal haemoglobinuria (PNH), in which CD59 is missing from the surface of blood cells (Hill *et al.*, 2006). Here, we present three high-resolution crystal structures of recombinantly produced human CD59.

2. Materials and methods

2.1. Protein production

Mature human CD59 was expressed in *Escherichia coli* and refolded using a large-scale variation of the method previously described for rat CD59 (Fraser *et al.*, 2003) and outlined for the human protein (Hill *et al.*, 2006). The coding region of the mature protein [(M)L1–N77 with a C-terminal additional cysteine] was expressed from the plasmid pET59-06 which utilizes a T7-RNA polymerase promoter and this was used to transform an *E. coli* K12-derived strain (HAMS113, Inflazyme Pharmaceuticals). An overnight culture was used to inoculate 5 l cultures of vegetable peptone broth containing 50 µg ml^{-1} kanamycin; these were grown to an $OD_{600 nm}$ of approximately 6 before being induced with 1 mM IPTG and grown for a further 3 h at 310 K. Cells were isolated by centrifugation and washed in 0.1 M NaCl, 50 mM Tris–HCl, 1 mM EDTA pH 8.0. The pellet (150 ml per litre of culture equivalent containing approximately 100 mg sCD59 per litre) was suspended in this buffer by blending and homogenized by passing it three times through an Emulsifex C5 homogeniser at 83–100 kPa. The inclusion bodies were isolated by centrifugation at 16 000g at 277 K and washed twice in 2%(w/v) sodium deoxycholate, 1 mM EDTA, 50 mM Tris–HCl pH 8.0. The washed pellets were resolubilized in 8 M urea, 5 mM dithiothreitol, 1 mM EDTA, 50 mM Tris–HCl pH 8.0 for 2 h at ambient temperature, clarified by centrifugation and the supernatant added slowly to a 50× volume of 0.5 M L-arginine, 50 mM Tris–HCl, 1 mM EDTA, 1 mM L-cystine, 3 mM L-cysteine pH 9.0 initially at ambient temperature (for between 20 and 25 min) and then held static at 277 K for 2–3 weeks. This prolonged period is required for correct formation of the five disulfide bonds in the construct. The product was concentrated 30-fold by ultrafiltration using a 5 kDa cutoff membrane and then diafiltered against ten volumes of PBS pH 7.3 at 277 K. The retentate was then mixed 1:1.5(v:v) with 3.8 M ammonium sulfate, 0.1 M sodium phosphate pH 6.5 and centrifuged at 16 000g for 15 min at 283 K. The supernatant was filtered using a 0.22 µm filter, subjected to the above ultrafiltration and diafiltration steps again and finally concentrated using an ultrafiltration step. sCD59 was finally prepared for crystallization at 10 mg ml^{-1} in 150 mM NaCl Tris–HCl pH 7.5 buffer.

2.2. Crystallization

Initial screening was carried out using Molecular Dimensions Structure Screens I and II with the protein at a concentration of approximately 10 mg ml^{-1}. Crystallization trials were set up with a Tecan Genesis 150 robot using the sitting-drop vapour-diffusion method. The crystallization sitting drop contained 0.2 µl protein solution and 0.2 µl mother liquor equilibrated against 100 µl mother-liquor solution. The crystallization trays were stored at 293 K.

Molecular Dimensions Structure Screen I yielded small crystals in two conditions, No. 43 (30% PEG 4000, 0.2 M ammonium sulfate) and No. 46 (20% PEG, 0.05 M potassium phosphate), after 4 d.

Optimization around these conditions yielded much larger crystals in three distinct crystal forms (Fig. 1): 28% PEG 4000, 0.25 M ammonium sulfate (Xtal1), 25% PEG 8000, 0.03–0.04 M potassium phosphate (Xtal2) and 23% PEG 8000, 20% ethanol, 0.03 M potassium phosphate (Xtal3). Prior to data collection, each crystal was briefly washed in a solution of mother liquor made up with an additional 30%(v/v) glycerol. Crystals were then mounted on a fibre loop and flash-cooled directly in liquid nitrogen.
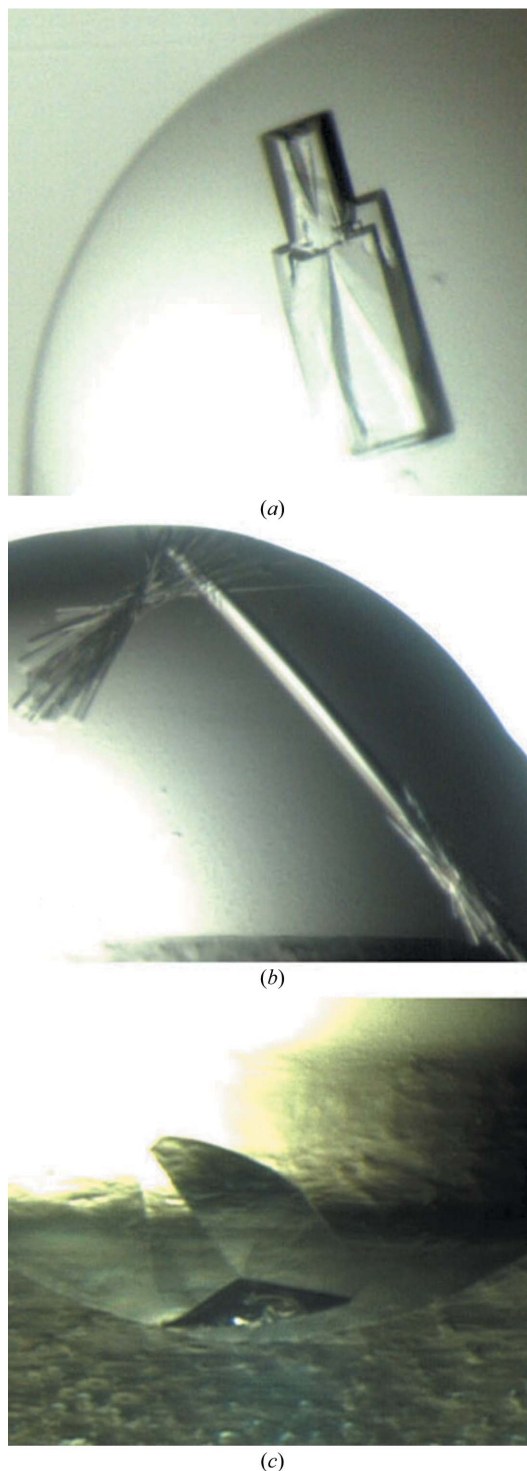

Figure 1
CD59 crystals grown under the conditions described in the text: (*a*) Xtal1, (*b*) Xtal2, (*c*) Xtal3.

Table 1
Data-collection statistics.

Data sets were collected at two wavelengths for Xtal1 and Xtal2: one at short wavelength for optimum resolution and one at a higher wavelength (potentially for sulfur phasing). Owing to misassignment of the space group at the beamline, these data sets were incomplete and had to be merged. The statistics given for Xtal1 and Xtal2 are for the merged data sets. Values in parentheses are for the highest resolution shell.

Crystal	Xtal1	Xtal2	Xtal3
Crystallization conditions	28% PEG 4000, 0.25 M ammonium sulfate	25% PEG 8000, 0.03 M potassium phosphate	23% PEG 8000, 20% ethanol, 0.03 M potassium phosphate
Beamline	ID29, ESRF	ID29, ESRF	ID14-2, ESRF
Wavelength (Å)	0.92, 1.24	0.82, 1.24	0.93
Space group	$P2_12_12_1$	$P6_1$	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 31.9$, $b = 45.9$, $c = 6.0$	$a = 60.8$, $b = 60.8$, $c = 107.5$	$a = 25.5$, $b = 27.9$, $c = 81.1$
Resolution range (Å)	25–1.10 (1.16–1.10)	38–1.8 (1.9–1.8)	40–1.34 (1.41–1.34)
No. of unique reflections	27926 (3751)	20934 (3066)	13368 (1881)
Multiplicity	12.0 (3.5)	14.3 (14.1)	5.2 (3.2)
Completeness (%)	99 (93.3)	100.0 (100.0)	99.0 (97.9)
$I/\sigma(I)$	20.7 (3.7)	29.0 (7.0)	16.3 (4.5)
R_{merge}	0.097 (0.235)	0.091 (0.324)	0.074 (0.220)
Solvent content (%)	33.92	41.78	21.7
Molecules per ASU	1	3	1

2.3. Data collection

High-resolution data were collected from all three crystal forms. The data sets from the first two crystals (1.15 and 1.8 Å) were collected at beamline ID29 at the ESRF using an ADSC Q210 two-dimensional detector. The data from the third crystal (grown in 20% ethanol) were collected at beamline ID14-2 at the ESRF (ADSC Q4 CCD detector) and yielded data to 1.34 Å. In each case, the data were collected in high- and low-resolution passes at 100 K. These were scaled and merged together to create a single data set that was complete over the full resolution range.

The data sets were integrated in *MOSFLM* (Leslie, 1992) and sorted and combined using *SORTMTZ* in *CCP*4 (Collaborative Computational Project, Number 4, 1994). The resultant data sets were then scaled in *SCALA* (Weiss, 2001). Any overloaded reflections were recovered using profile fitting in *SCALA*. Table 1 contains data statistics.

2.4. Structure solution and refinement

The structure of Xtal1 was determined with *MOLREP* (Vagin & Teplyakov, 1997) using the highest resolution data set and one of the available NMR CD59 structures (PDB code 1cdq; Fletcher *et al.*, 1994) as a search model. Rigid-body refinement against a truncated resolution data set (20–2 Å) in *BUSTER-TNT* (Blanc *et al.*, 2004) was followed by iterative cycles of full refinement with an incremental increase in the resolution range of the data to 1.15 Å. *ARP/wARP* (Morris *et al.*, 2003) was used for autobuilding. *Coot* (Emsley & Cowtan, 2004) was used to interactively adjust the model. Waters

were initially added using *autoBUSTER*. *Coot* was used to add further waters. The progress of refinement was monitored by calculation of R_{free} using 5% of the total independent reflections.

The two other crystal forms were solved by molecular replacement using *Phaser* (McCoy *et al.*, 2005) with the Xtal1 structure as a search model. *BUSTER-TNT* was again used to carry out iterative cycles of full refinement with an incremental increase in the resolution range of the data used (Xtal2, 1.8 Å; Xtal3, 1.35 Å). *Coot* was used to add unmodelled terminal residues to the structure and to interactively alter the model. Waters were once again initially added using *autoBUSTER* and *Coot* was used to add further waters. Refinement statistics are given in Table 2.

3. Results and discussion

Three crystal structures of human CD59 expressed and refolded from *E. coli* have been determined to resolutions of 1.15, 1.8 and 1.35 Å. Fig. 2 shows the highest resolution structure. As in the NMR structures, the protein is comprised of three β-sheets and an α-helix; in addition, there is a small α-helix present that is not described in either of the NMR structures or in the recently published 2.1 Å crystal
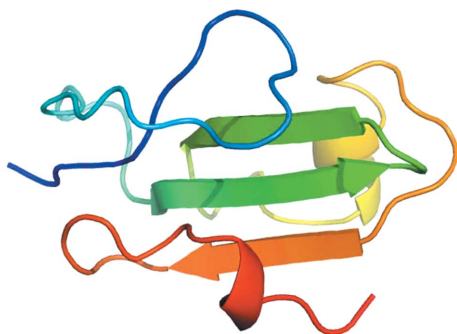

Figure 2
Cartoon representation showing the secondary structure of the highest resolution CD59 structure (Xtal1; N-terminus, blue; C-terminus, red). Graphics were generated using *PyMOL* (DeLano, 2002).
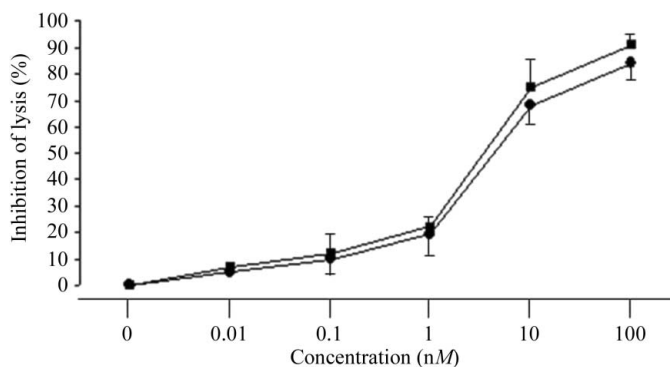

Figure 3
Graph showing a comparison of the recombinant soluble bacterially expressed CD59 with soluble CD59 expressed in CHO cells. Guinea pig erythrocytes (1% in VBS) were incubated with 10% NHS-C8 (15 min, 310 K) to form C5b-7 sites; they were then washed and resuspended in VBS (1%). C8 and excess C9 were added to give approximately 80–90% lysis. CD59 was added simultaneously. These were then incubated at 310 K for 30 min and the haemolysis was measured. Mammalian-derived CD59 is represented by circles and bacterially expressed CD59 is represented by squares. Each result is plotted as percentage inhibition of lysis and is a mean (± standard error) of three measurements.

Table 2
Refinement statistics.

Values in parentheses are for the highest resolution shell.

Crystal	Xtal1	Xtal2	Xtal3
PDB code	2j8b	2uwr	2ux2
Refinement resolution range (Å)	20–1.15 (1.18–1.15)	20–1.8 (1.85–1.8)	40–1.35 (1.44–1.35)
R factor (%)	16.36 (20.0)	18.62 (18.85)	20.21 (25.56)
R_{free} (%)	19.68 (20.0)	22.32 (26.28)	22.31 (30.22)
R.m.s.d. bonds (Å)	0.010	0.004	0.006
R.m.s.d. angles (°)	1.407	0.827	1.120
Ramachandran favoured (%)	98.7	96.0	97.4
Ramachandran generously allowed (%)	1.3	4	2.6
Ramachandran outliers (%)	0	0	0
No. of reflections	23463	20879	11863
No. of protein atoms	676	1917	680
No. of waters	117	213	93
No. of other atoms	0	0	8
Mean B, protein atoms ($\mathring{A}^2$)	11.1	17.2	17.7
Mean B, main chain ($\mathring{A}^2$)	9.9	14.2	13.9
Mean B, side chain ($\mathring{A}^2$)	12.3	20.1	21.3
Mean B, water atoms ($\mathring{A}^2$)	23.5	28.2	28.3

structure (Huang *et al.*, 2007). In the two lower resolution crystal structures the N-terminal loop is interrupted by two short β-sheets; however, at high resolution these sheets are not present. Two of the crystal forms belong to space group $P2_12_12_1$, while the third belongs to $P6_1$. The two $P2_12_12_1$ crystals (Xtal1 and Xtal3) have one molecule per asymmetric unit, while the $P6_1$ crystal (Xtal2) has three. Since the refolded protein has been demonstrated to have full biological activity (Fig. 3), it is unsurprising that the structure revealed is consistent with the earlier NMR structures of eukaryotically expressed CD59. *THESEUS* (Theobald & Wuttke, 2006) was used to analyse and compare the variation per residue between the new structures and the closest to mean representatives of the two independently determined NMR structures. Fig. 4 illustrates the outcome of this analysis and emphasizes the similarity between the crystal structures and NMR models. Both NMR models show greater variability per residue than the ensemble of crystal structures, especially in the N-terminal loop (residues 0–23).

Site-directed mutagenesis of human CD59 has implicated several residues as being in the C8/C9-binding site, including Asp22, Phe23, Trp40, Phe42, His44, Arg53, Leu54 and Glu56 (Bodian *et al.*, 1997; Tomlinson *et al.*, 1994; Huang *et al.*, 2005). Trp40 lies in the centre of a surface-exposed hydrophobic groove on the surface of CD59; Arg53 is one of several hydrophilic residues which together form a ridge running along one side of this hydrophobic groove. Mutating Trp40 to Glu results in a loss of complement inhibition by CD59 and prevents inhibitory monoclonal antibodies binding to CD59 (Bodian *et al.*, 1997). Arg53 has also been shown through mutation studies to be central to ligand binding (Huang *et al.*, 2006) as well as to the binding of monoclonal antibodies (Petranka *et al.*, 1996; Bodian *et al.*, 1997) used to inhibit CD59 complement regulatory action.

The electron-density map from Xtal3 revealed an unexpected ligand in the proposed C8/C9-binding site (Fig. 5; Huang *et al.*, 2005). The position of this ligand crudely approximates the position of peptides modelled into the human CD59 NMR structures (Huang *et al.*, 2007). The presence of this ligand correlates with changes in the conformation of the side chains of several amino acids surrounding the ligand, including Arg53 (Tomlinson *et al.*, 1994; Huang *et al.*, 2006; Bodian *et al.*, 1997). Overlaying all the crystal structures reveals the reorganization of several side chains (Glu43 and Arg53; Fig. 5) around the ligand. Reorganization was not observed in any of the other residues implicated in ligand binding by mutational studies (Huang *et al.*, 2007). Arg55 was seen in multiple conformations which may correlate to the open and closed conformations predicted

(Huang *et al.*, 2006); however, these varied conformations were observed independently of the presence of the ligand.

In Fig. 5 the side chain of Glu43 can be seen to 'swing out' from the binding site. This would provide more space in the binding site for the ligand. The most dramatic conformational change is seen in the side chain of Arg53. In the presence of the bound ligand, Arg53 is observed in two conformations. Approximately 75% lies in a position that could allow coordination of a ligand, while ~25% adopts the conformation seen in all the other 'empty' crystal forms (presumably owing to partial occupation of the binding site by the ligand). Since the primary difference between this crystal form and the others was the presence of 20% ethanol in the mother liquor, our initial interpretation was that ethanol was binding in a site designed to bind the planar peptide bond. Attempts to refine a single ethanol in the density revealed that this was not a satisfactory method to model this feature. Further consideration of the crystallization conditions reveals no other obvious candidate and, since the other crystal forms grown from protein produced in the same way do not have any ligand bound, we consider it most likely that the feature results from the superposition of densities arising from multiple partially occupied differently bound ethanol molecules within the crystal.

Although the bound ligand remains unidentified, the fact that it is observed in the putative C8/C9-binding site alongside conformational changes in key binding-site residues offers a tantalizing glimpse of the binding events that may occur during CD59 inhibition of MAC
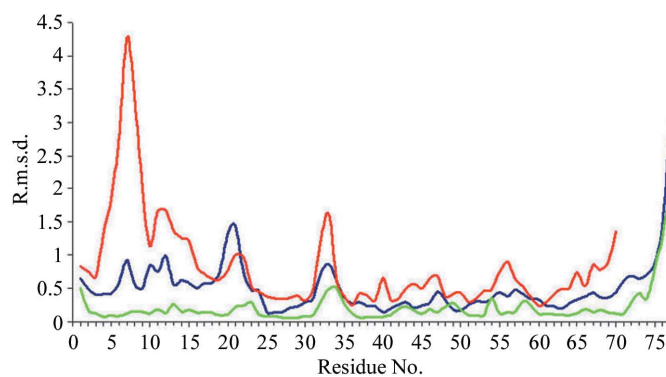

Figure 4
R.m.s.d. by residue for CD59 structures known to date generated using the program *THESEUS* (Theobald & Wuttke, 2006). Blue, 1cdq (Fletcher *et al.*, 1994; NMR; 20 structures); red, 1erh (Kieffer *et al.*, 1994; NMR; ten structures); green, X-ray structures (five structures).
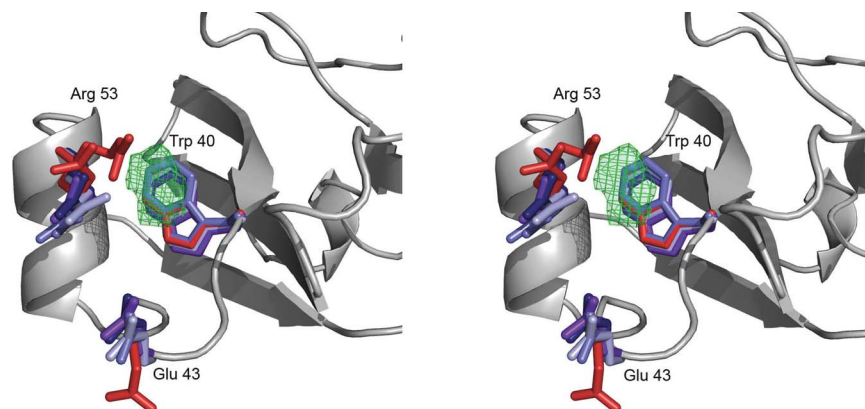
protein structure communications

Figure 5
Overlaying all crystal structures highlights the alternate conformations observed for the side chains of Glu43 and Arg53 in the presence of bound ligand (red, bound ligand; blue, no bound ligand). Also shown is Trp40, a residue heavily implicated in the C8/C9-binding site (Huang *et al.*, 2006; Bodian *et al.*, 1997). Graphics were generated using *PyMOL* (DeLano, 2002).

assembly and insertion. The biological activity and the essential identity to NMR-based structures of the CD59 used in this study also provide important validation for a molecule prepared by high-level expression in *E. coli* using a process amenable to industrial-scale operation. Such validation will assist the use of CD59 derivatives as therapeutic biopharmaceuticals (Hill *et al.*, 2006).

References

Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M. & Bricogne, G. (2004). *Acta Cryst.* D**60**, 2210–2221.
Bodian, D. L., Davis, S. J., Morgan, B. P. & Rushmere, N. K. (1997). *J. Exp. Med.* **185**, 507–516.
Caragine, T., Chen, S., Frey, A. & Tomlinson, S. (1998). *Mol. Immunol.* **35**, 337.
Chang, C. P., Husler, T. C., Zhao, J., Webster, J. L., Wiedmer, T. & Sims, P. J. (1993). *Blood*, **82**, A202.
Chen, S., Chen, G., Cai, C. C., Wang, X. M., Guo, H., Shen, S. Q., Wang, H., Wu, Y., Li, G. X., Huang, J., Li, W. X., Wei, Q. X. & Sun, F. Z. (2001). *Transplant. Proc.* **33**, 3851–3852.
Chen, S. H., Caragine, T., Cheung, N. K. V. & Tomlinson, S. (2000). *Cancer Res.* **60**, 3013–3018.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *PyMOL.* DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Farkas, I., Baranyi, L., Ishikawa, Y., Okada, N., Bohata, C., Budai, D., Fukuda, A., Imai, M. & Okada, H. (2002). *J. Physiol. (London)*, **539**, 537–545.
Fishelson, Z., Donin, N., Zell, S., Schultz, S. & Kirschfink, M. (2003). *Mol. Immunol.* **40**, 109–123.
Fletcher, C. M., Harrison, R. A., Lachmann, P. J. & Neuhaus, D. (1994). *Structure*, **2**, 185–199.
Fraser, D. A., Harris, C. L., Williams, A. S., Mizuno, M., Gallagher, S., Smith, R. A. G. & Morgan, B. P. (2003). *J. Biol. Chem.* **278**, 48921–48927.
Gelderman, K. A., Tomlinson, S., Ross, G. D. & Gorter, A. (2004). *Trends Immunol.* **25**, 158–164.
Hill, A., Ridley, S. H., Esser, D., Oldroyd, R. G., Cullen, M. J., Kareclas, P., Gallagher, S., Smith, G. P., Richards, S. J., White, J., Smith, R. A. G. & Hillmen, P. (2006). *Blood*, **107**, 2131–2137.
Huang, Y., Fadarovich, A., Tomlinson, S. & Davies, C. (2007). *Acta Cryst.* D**63**, 714–721.
Huang, Y., Qiao, F., Abagyan, R., Hazard, S. & Tomlinson, S. (2006). *J. Biol. Chem.* **281**, 27398–27404.
Huang, Y. X., Smith, C. A., Song, H. B., Morgan, B. P., Abagyan, R. & Tomlinson, S. (2005). *J. Biol. Chem.* **280**, 34073–34079.
Kieffer, B., Driscoll, P. C., Campbell, I. D., Willis, A. C., Vandermerwe, P. A. & Davis, S. J. (1994). *Biochemistry*, **33**, 4471–4482.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Meri, S., Morgan, B. P., Davies, A., Daniels, R. H., Olavesen, M. G., Waldmann, H. & Lachmann, P. J. (1990). *Immunology*, **71**, 1–9.
Morgan, B. P. & Harris, C. L. (2003). *Mol. Immunol.* **40**, 159–170.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.
Murray, K. P., Mathure, S., Kaul, R., Khan, S., Carson, L. F., Twiggs, L. B., Martens, M. G. & Kaul, A. (2000). *Gynecol. Oncol.* **76**, 176–182.
Ninomiya, H. & Sims, P. J. (1992). *J. Biol. Chem.* **267**, 13675–13680.
Petranka, J., Zhao, J., Norris, J., Tweedy, N. B., Ware, R. E., Sims, P. J. & Rosse, W. F. (1996). *Blood Cells Mol. Dis.* **22**, 281–296.
Rollins, S. A. & Sims, P. J. (1990). *J. Immunol.* **144**, 3478–3483.
Theobald, D. L. & Wuttke, D. S. (2006). *Bioinformatics*, **22**, 2171–2172.
Tomlinson, S., Whitlow, M. B. & Nussenzweig, V. (1994). *J. Immunol.* **152**, 1927–1934.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.