Three-Dimensional Solution Structure of the Extracellular Region of the Complement Regulatory Protein CD59, a New Cell-Surface Protein Domain Related to Snake Venom Neurotoxins[†,‡]

Bruno Kieffer,[§,∥] Paul C. Driscoll,[*,§] Iain D. Campbell,[§] Antony C. Willis,[⊥] P. Anton van der Merwe,[#] and Simon J. Davis[#]

Department of Biochemistry and MRC Immunochemistry Unit, Rex Richards Building, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K., and MRC Cellular Immunology Unit, Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, U.K.

Received December 1, 1993; Revised Manuscript Received January 26, 1994[⊛]

ABSTRACT: The cell surface antigen CD59 is an inhibitor of complement-mediated lysis and a member of the Ly6 superfamily (Ly6SF) of cysteine-rich cell-surface molecules whose sequences are related to those of snake venom neurotoxins. The three-dimensional solution structure of a recombinant form of the extracellular region of the molecule (residues 1–70 of the mature protein; sCD59) has been solved by 2D NMR methods. sCD59 is a relatively flat, disk-shaped molecule consisting of a two-standed β-sheet finger loosely packed against a protein core formed by a three-stranded β-sheet and a short helix. Structure calculations allowed an unambiguous assignment of the disulfide-bonded cysteine pairs as 3–26, 6–13, 19–39, 45–63, and 64–69. The topology of sCD59 is similar to that of the snake venom neurotoxins and consistent with an evolutionary relationship existing between the Ly6SF and the neurotoxins.

The 20-kDa glycophosphatidylinositol- (GPI-)[1] anchored glycoprotein CD59, known previously as MEM-43 antigen, MIRL, H19, MACIF, HRF20, and protectin, is found on the surface of cells of both hemopoietic and nonhemopoietic origin [reviewed by Walsh *et al.* (1992)]. The function of CD59 was first suggested by the finding that the purified antigen inhibits complement-mediated lysis (Sugita *et al.*, 1988). CD59 inhibits complement activation by binding in a glycosylation-dependent manner (Ninomiya *et al.*, 1992) to C5b-C8 and/or C9 and preventing the formation of the membrane attack complex (Meri *et al.*, 1990; Ninomya & Sims, 1992). Comparisons of CD59 with the other complement inhibitors, decay accelerating factor and membrane cofactor protein, indicate that CD59 is the most potent inhibitor of complement-mediated lysis of human endothelial cells (Brooimans *et al.*, 1992). In addition to inhibiting complement and like other GPI-anchored molecules, CD59 can transmit activating signals to T cells (Korty *et al.*, 1991) and forms part of a signal-transducing complex on the surface of these cells (Štefanová & Hořejší, 1991; Štefanová *et al.*, 1991). It has been proposed that CD59 is also a second ligand for the human T lymphocyte adhesion molecule CD2 (Hahn *et al.*, 1992; Deckert *et al.*, 1992), but this is controversial (Arulanandam *et al.*, 1993; P. A. van der Merwe *et al.*, unpublished data).

CD59 belongs to a group of cell-surface molecules, the first members of which were the mouse Ly6A/E and Ly6C antigens implicated in T lymphocyte activation (LeClair *et al.*, 1986; Palfree *et al.*, 1988) and the squid glycoprotein of unknown function, Sgp-2 (Williams *et al.*, 1988). These molecules each consist of single, cysteine-rich domains of 70–92 amino acids attached to the cell surface by GPI anchors [reviewed by Williams (1991)]. More recently discovered members of what is now known as the Ly6 superfamily (Ly6SF; Williams, 1991) are the urokinase plasminogen activator receptor which has three Ly6SF domains (Roldan *et al.*, 1990; Palfree, 1991), the mouse thymocyte antigen ThB (Gumley *et al.*, 1992), and additional mouse Ly6 antigens (Fleming *et al.*, 1993). A second, distinct group of molecules, the snake venom neurotoxins, has been shown to be related to the Ly6SF on the basis of sequence comparisons (Fleming *et al.*, 1993), together with the *Xenopus* xenoxins of presently unknown function (Kolbe *et al.*, 1993).

In this study the three-dimensional structure of a recombinant form of the extracellular region of the CD59 molecule (sCD59) has been determined by 2D solution NMR methods. This prototypic Ly6SF structure provides both additional evidence for an evolutionary relationship between the Ly6SF and snake venom neurotoxins and a basis for future analysis of the structure–function relationships of CD59 and other members of the Ly6SF.

MATERIALS AND METHODS

Expression of sCD59. The polymerase chain reaction (PCR) was used to produce a truncated cDNA encoding sCD59 from a full-length cDNA (Davies *et al.*, 1989). The 5′ oligonucleotide was complementary to the CD59 leader-encoding sequence, and this also inserted, immediately

[†] This work was supported by the Oxford Centre for Molecular Sciences, funded by the SERC and MRC, and by the Human Frontier Science Program. P.C.D. is a Royal Society University Research Fellow.

[‡] The solution structure has been deposited in the Brookhaven Protein Data Bank under the file name 1ERH.

[*] Address correspondence to this author at the Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

[§] Department of Biochemistry.

[∥] Current address: Institut de Biologie Moleculaire et Cellulaire, Centre National de la Recherche Scientifique, 15, rue Descartes, 67084 Strasbourg Cedex, France.

[⊥] MRC Immunochemistry Unit.

[#] MRC Cellular Immunology Unit.

[⊛] Abstract published in *Advance ACS Abstracts,* April 1, 1994.

[1] Abbreviations: 2D, two dimensional; 3D, three dimensional; COSY, correlated spectroscopy; GPI, glycophosphatidylinositol; HOHAHA, homonuclear Hartmann–Hahn; Ly6SF, Ly6 superfamily; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; rmsd, root mean square deviation; PCR, polymerase chain reaction; sCD59, soluble form of the extracellular region of human CD59 (residues 1–70); CHO, Chinese hamster ovary.

0006-2960/94/0433-4471$04.50/0

upstream from the leader sequence, a *Hin*dIII site followed by 25 bp of the 5′ untranslated sequence from the region immediately upstream of the initiating codon of the rat CD4 cDNA sequence (Clark *et al.*, 1987). The 3′ oligonucleotide introduced a termination codon immediately after the codon for Asn-70 of the CD59 cDNA [the numbering is based on the sequence of the processed polypeptide (Štefanová *et al.*, 1989)], followed by an *Xba*I site to facilitate subcloning. The entire coding sequence of the PCR product was then checked by dideoxy sequencing.

To express sCD59, the *Hin*dIII–*Xba*I fragment was subcloned into the polylinker of pEE6.hcmv-GS (Bebbington & Hentschel, 1987; Davis *et al.*, 1990), and this was then transfected with calcium phosphate into Chinese hamster ovary (CHO) cells. After selection with 25 μM methionine sulfoximine, the highest expressing clone was chosen after comparison of clonal expression levels by metabolic labeling with [^{35}S]cysteine followed by gel electrophoresis and fluorography. For large-scale production the highest expressing cell line was grown in cell factories (Nunc). The sCD59 was purified from the spent tissue culture medium precleared by centrifugation at 10000*g* for 30 min by affinity chromatography using an antibody affinity column prepared with the YTH53.1 monoclonal antibody (Davies *et al.*, 1989) according to published methods (Arvieux & Williams, 1988), followed by gel filtration in 10 mM Hepes and 140 mM NaCl, pH 7.4, on a Superdex S-75 fast protein liquid chromatography system (Pharmacia) or Sephadex G-75 (Pharmacia).

NMR Spectroscopy. NMR spectroscopy was performed using the hybrid GE Omega/home-built or Bruker AM600 spectrometers of the Oxford Centre for Molecular Sciences. Each machine is equipped with a superconducting magnet (Oxford Instrument Co.) operating at ^{1}H frequencies of 500, 600, or 750 MHz and fitted with either a Bruker multinuclear (500- and 600-MHz machines) or home-built ^{1}H (750-MHz) probe.

All 2D spectra were recorded in a phase-sensitive manner with quadrature detection in the F_1 dimension obtained using hypercomplex data acquisition in the States–TPPI mode (Marion *et al.*, 1989a,b). Optimization of the spectral baseline properties in the F_2 and F_1 dimensions was achieved by appropriate selection of the preacquisition delay and initial t_1 value, respectively (Bax *et al.*, 1991). For spectra recorded with a jump–return read pulse each FID was convoluted in the time domain with a Gaussian window (Marion *et al.*, 1989a,b) prior to Fourier transformation. NMR data processing was performed using the software package FELIX 2.05 (Biosym Inc.).

The sequence-specific resonance assignment of the ^{1}H NMR spectrum of sCD59 was performed using standard procedures (Wagner & Wüthrich, 1982; Wüthrich, 1986) involving the analysis of 2D nuclear Overhauser effect spectroscopy (NOESY), homonuclear Hartmann–Hahn spectroscopy (HOHAHA), and correlated spectroscopy (COSY) data. The assignment of the spectrum of sCD59 was based on the analysis of the following spectra. Spectra recorded in a 90% H_2O/10% D_2O solution: 2D NOESY (Jeener, 1979; Kumar *et al.*, 1980) with a jump–return read pulse (Plateau & Guéron, 1982; Driscoll *et al.*, 1989) and mixing times of 200 ms (600 MHz, 25 °C) and 150 ms (600 MHz, 15, 35, and 42 °C); 2D HOHAHA (Braunschweiler & Ernst, 1983; Davis & Bax, 1985) with a jump–return read pulse (Bax 1989; Driscoll *et al.*, 1989) and mixing times of 40 ms (600 MHz, 15, 35 °C). Spectra recorded in a 99.9% D_2O solution (with low-power irradiation of the residual solvent peak): 2D NOESY with mixing times of 150 ms (750 MHz, 35 °C; 500 MHz, 25 °C; and 600 MHz, 42, 35 °C) and 50 ms (600 MHz, 35 °C); 2D HOHAHA (Davis & Bax, 1985; Bax, 1989) with mixing times of 40 ms (600 MHz, 25, 35, and 42 °C); 2D primitive-COSY (Marion & Bax, 1988) (600 MHz, 35 °C).

For each experiment typically 400 complex t_1 increments of 1K complex t_2 points were recorded. The sweep width was 13.4 ppm for H_2O experiments and 11.7 ppm for experiments in D_2O solution. The spectra were processed using a Gaussian/Lorentzian window function in the F_2 dimension before zero filling, Fourier transformation, and linear baseline correction. The t_1 interferograms were extended to 512 complex points using linear prediction followed by apodization with an 80° shifted sine-squared bell window function, zero filling, and Fourier transformation. For COSY experiments, 2K complex t_2 points were recorded for each increment, and skewed sine bell window functions were employed in the processing of both F_1 and F_2 dimensions.

Slowly exchanging NH protons were identified by recording NOESY spectra (600 MHz, mixing time 150 ms, 25 °C) of a sample of sCD59 freeze-dried from H_2O solution, 30 min after dissolution in 99.9% D_2O.

Vicinal $^3J_{H^NH^\alpha}$ spin–spin coupling constants were obtained from a series of COSY spectra recorded at 500 and 600 MHz at 25, 35, and 42 °C. A combination of low-power irradiation of the solvent resonance with the SCUBA pulse sequence (Brown *et al.*, 1987) was used to recover cross-peaks close to the water frequency. Theoretical cross-peak profiles with variable coupling constant, offset, and line-width parameters were fitted in a nonlinear least-squares manner to one-dimensional cross sections parallel to the $\omega_1 = 0$ axis of the 2D spectrum (Redfield & Dobson, 1990).

RESULTS

Expression of sCD59. In order to generate enough of the extracellular region of CD59 for structural studies, a soluble form of the molecule was prepared by heterologous expression in mammalian cells. To our knowledge the site of addition of the GPI anchor of CD59 has not thus far been determined. The GPI anchor of the related glycoprotein, Sgp-2, is known to be added to a highly conserved Asn residue equivalent to Asn-70 of the mature CD59 polypeptide, suggesting that this residue terminates the Ly6SF domain (Williams *et al.*, 1988; Williams, 1991). However, there is a very good consensus site for addition of a GPI anchor at Asn-77 of the mature protein sequence (Barclay *et al.*, 1992), in which case a seven amino acid stalk would link the Ly6SF domain of CD59 to the GPI anchor. Given these uncertainties, and because a compact protein was considered to be more amenable to structural analysis, the CD59 cDNA (Davies *et al.*, 1989) was modified by insertion of a stop codon after the codon for Asn-70 of the mature protein sequence. The glutamine synthetase-based gene expression system was used to generate CHO cell lines expressing sCD59 at levels up to ~2 mg/L (Bebbington & Hentschel, 1987; Davis *et al.*, 1990).

Soluble CD59 was purified from the supernatant of cultures of sCD59-secreting CHO cells by affinity chromatography using an affinity column prepared with the YTH53.1 monoclonal antibody (Davies *et al.*, 1989). Protein eluting from the affinity column separated into three fractions upon gel filtration (Figure 1A). Fraction I consisted of aggregated material and antibody eluting from the affinity column (data not shown). Protein in fractions II and III eluted well after deglycosylated human sCD2 (MW 21 000) used as a molecular weight marker, indicating that sCD59 eluting in fractions II

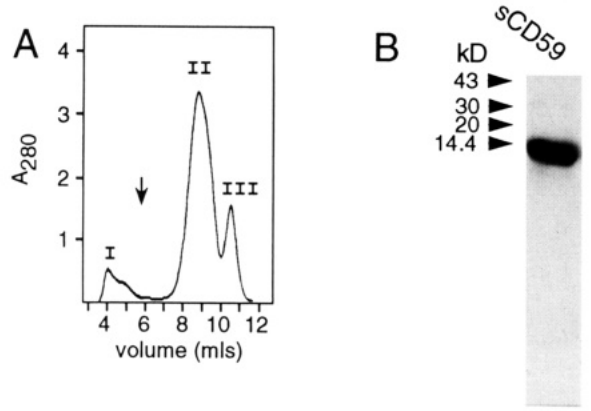

FIGURE 1: Preparation of recombinant sCD59 expressed in Chinese hamster ovary cells. (A) sCD59 purified by affinity chromatography was applied to a Superdex S-75 gel filtration column and eluted in three fractions (I, II, and III). The arrow indicates the position of elution of deglycosylated human soluble CD2 (21 kDa) run as a molecular weight marker. (B) Five micrograms of the fraction II protein used for the structural analysis was subjected to SDS–PAGE analysis under reducing conditions in a 20% gel which was then stained with Coomassie blue.

and III was monomeric. Fraction II protein migrated as a single band with an apparent molecular weight of 12 000–14 000 during SDS–PAGE under reducing conditions (Figure 1B), and the anti-CD59 monoclonal antibody-binding properties of this material were indistinguishable from those of human CD59 purified from urine (P. A. van der Merwe, B. P. Morgan, and S. J. Davis, unpublished data). SDS–PAGE analyses were consistent with fractions II and III containing glycosylated and unglycosylated sCD59, respectively, suggesting that glycosylation is not essential for the folding and secretion of sCD59 (data not shown). In order to restrict sample heterogeneity, only fraction II was used for data collection; there was insufficient fraction III protein for detailed analysis. N-Terminal sequencing of the protein used for the structural analysis gave the amino acid sequence LeuGln-CysTyrAsn, indicating that sCD59 signal-peptide processing in CHO cells is identical to that *in vivo* (data not shown; Štefanová *et al.*, 1989). Initial crystallization trials with sCD59 treated with endoglycosidase H to trim the glycosylation to single *N*-acetylglucosamine residues produced clusters of small, highly birefringent crystals, but these were unsuitable for crystallographic studies.

Sequence-Specific ^1H NMR Spectral Assignments and Secondary Structure. Samples of sCD59 yielded 2D ^1H NMR spectra of high quality, with a well-dispersed range of amino acid chemical shifts and with a number of additional cross-peaks due to the N-linked glycan attached at residue Asn-18. For each amino acid only a single spin system was observed, suggesting that the sample is homogeneous in the protein component and is not seriously affected by any heterogeneity in the *N*-glycan component of the glycoprotein. The sequence-specific assignment of the spectrum of sCD59 proceeded in a straightforward manner. Sequential $d_{\alpha N}(i,i+1)$ or $d_{NN}(i,i+1)$ NOEs were observed through the complete amino acid sequence. Figure 2A shows a portion of the 2D NOESY spectrum of sCD59 recorded at 600 MHz with a mixing time of 150 ms, together with the corresponding region from the SCUBA-COSY spectrum (Figure 2B). Annotations are included in Figure 2A to illustrate some of the sequential NOE connectivities observed. NOEs connecting proline residues (Pro-7 and Pro-9) were of the $d_{\alpha P\delta}(i,i+1)$ type, indicating *trans*-Xxx–Pro peptide bonds. A summary of the observed short- and medium-range NOEs is given in Figure

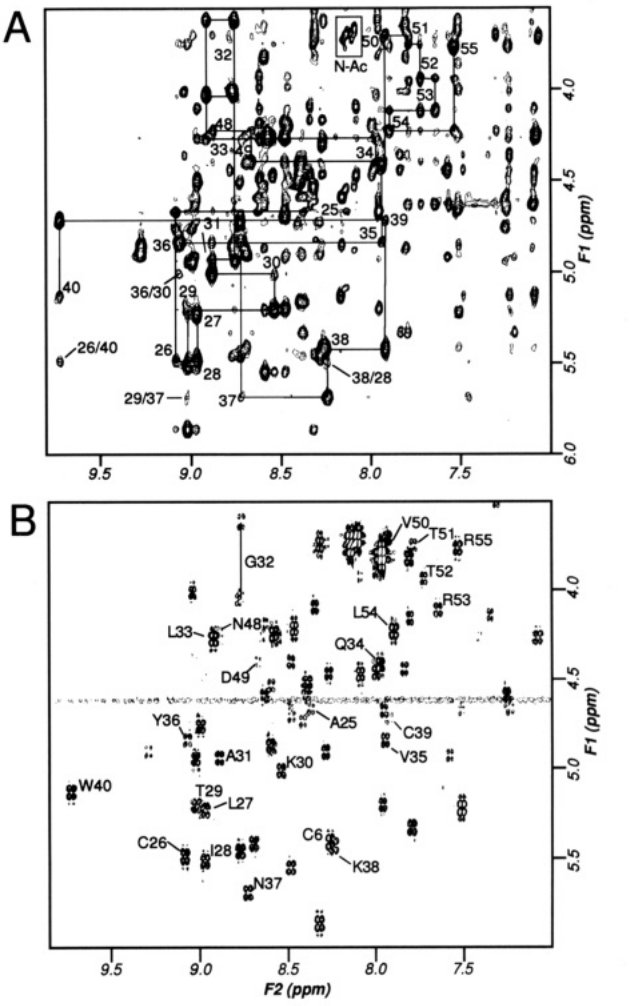

FIGURE 2: (A) A portion of the 600-MHz NOESY spectrum of sCD59 in 90% H_2O/10% D_2O at pH 5.5 and 308 K with a mixing time of 150 ms. The region shown is that containing cross-peaks between backbone amide NH protons and $C^\alpha H$ resonances. Two sets of sequential NOE connectivities are indicated by an unbroken line with the labels at the intraresidue NH–$C^\alpha H$ cross-peaks. The first set (residues 25–40) includes residues belonging to β-strands C and D. The interstrand cross-peaks are also indicated. The second set (residues 48–55) indicates the residues located in the helical part of the structure. The box in the upper part of the spectrum identifies cross-peaks between *N*-acetyl amide proton signals and anomeric resonances of the *N*-acetylglucosamine moieties of the N-linked glycan. (B) The fingerprint region of the SCUBA-COSY spectrum of sCD59 in 90% H_2O/10% D_2O at pH 5.5 and 308 K. The region shown corresponds to the same region as in (A).

3. Also shown are the residues for which slowly exchanging amide NH protons were observed and outlying values (either greater than 8 Hz or less than 6 Hz) of the $^3J_{H^NH^\alpha}$ spin–spin coupling constants were detected. The assignment of side-chain resonances was completed through the analysis of 2D ^1H HOHAHA and COSY experiments. The list of resonance assignments made for sCD59 is given in Table 1.

A complete analysis of the short-, medium-, and long-range backbone NOEs observed for sCD59 together with the pattern of slowly exchanging amide NH protons and $^3J_{H^NH^\alpha}$ spin–spin coupling constants was made to form an impression of the secondary structure elements in sCD59. The results of this analysis are shown in schematic form in Figure 4. sCD59 contains two elements of antiparallel β-sheet secondary structure: a two-stranded finger [residues 2–5 (strand A) and 15–18 (strand B)] linked by a loop including proline residues Pro-7 and Pro-9 and a triple-stranded β-sheet [residues 25–30 (strand C), 35–40 (strand D), and 61–64 (strand E)].

Table 1: Assignment of ^{1}H NMR Resonances for sCD59 at 308 K and pH 5.5[a]

residue	NH	αH	βH	γH	other
1 Leu		4.01	1.18/1.55		
2 Gln	8.77	5.46	1.68	1.95/2.62	
3 Cys	8.28	4.90	2.17/2.97		
4 Tyr	9.28	4.90	2.50/2.55		δ 6.71; ε 6.64
5 Asn	8.69	5.42	2.28/2.73	γNH$_2$ 6.74/7.10	
6 Cys	8.26	5.42	3.30/3.62		
7 Pro		4.25	1.98/2.38	2.10	δ 3.86/4.00
8 Asn	7.51	5.22	2.48/2.76	γNH$_2$ 6.93/7.57	
9 Pro		4.30	0.91/1.54	1.82	δ 3.62/3.67
10 Thr	8.27	4.47	4.15	1.15	
11 Ala	8.34	4.10	1.34		
12 Asp	8.08	4.47	2.35/2.53		
13 Cys	8.54	4.27	2.73/3.09		
14 Lys	8.57	4.25	1.50/1.97	1.00	δ 1.23/1.32; ε 2.75/2.80
15 Thr	7.08	4.26	4.26	1.26	
16 Ala	8.48	4.41	1.07		
17 Val	8.39	4.52	2.02	0.84/0.89	
18 Asn	8.41	4.74	2.61/2.64	8.59	
19 Cys	8.73	4.51	3.27		
20 Ser	8.96	4.49	4.07/4.27		
21 Ser	8.62	4.19	3.91/3.97		
22 Asp	8.46	4.22	1.85/2.08		
23 Phe	7.99	4.47	3.00/3.50		δ 7.30; ε 7.53
24 Asp	8.39	4.56	2.55/3.02		
25 Ala	8.37	4.67	1.09		
26 Cys	9.08	5.49	2.90/3.12		
27 Leu	8.97	5.23	0.87/1.27	0.80	δ 0.33/–0.81
28 Ile	8.96	5.51	1.89	0.99/0.99	γ$_2$ 0.93; δ 0.81
29 Thr	9.01	5.21	3.73	0.94	
30 Lys	8.54	5.01	1.66/1.83	1.29	δ 1.60; ε 2.87
31 Ala	8.88	4.94	1.32		
32 Gly	8.75	3.62/4.04			
33 Leu	8.91	4.28	1.60/1.65		δ 0.81/0.88
34 Gln	7.97	4.42	2.20/2.34	1.63	
35 Val	7.94	4.84	1.93	0.80/0.90	
36 Tyr	9.06	4.84	2.41/2.64		δ 6.85; ε 6.67
37 Asn	8.72	5.69	2.52/2.63	γNH$_2$ 6.24/7.46	
38 Lys	8.24	5.43	1.44	1.27	δ 1.27; ε 3.03
39 Cys	7.92	4.72	2.14		
40 Trp	9.72	5.13	2.70/2.88		δ$_1$ 7.49; ζ$_2$ 7.06; ζ$_3$ 7.11; η$_2$ 6.90; ε$_3$ 8.17; NHε 8.78
41 Lys	6.83	5.17	1.00	1.45	δ 1.15/1.23; ε 2.41
42 Phe	11.08	4.36	3.03/3.10		δ 7.23; ε 7.35; ζ 7.27
43 Glu	9.03	4.02	1.45/1.61		
44 His	7.79	5.33	2.57/3.26		δ$_2$ 6.64
45 Cys	7.84	4.45	3.08/3.49		
46 Asn	7.57	4.92	2.67/2.90	γNH$_2$ 7.20/8.38	
47 Phe	8.97	3.71	3.09/3.40		δ 7.30; ε 7.53
48 Asn	8.87	4.23	2.59/2.84	γNH$_2$ 6.90/7.63	
49 Asp	8.66	4.40	2.32/2.67		
50 Val	7.92	3.71	1.44	0.50/0.60	
51 Thr	7.79	3.75	3.95	0.63	
52 Thr	7.72	3.95	4.13	1.20	
53 Arg	7.64	4.12	1.93/1.98	1.56/1.76	δ 2.63/2.76
54 Leu	7.90	4.23	0.84	1.42	δ –0.08/0.33
55 Arg	7.53	3.77	2.01/2.83	1.49	NHε 7.10
56 Glu	7.23	4.64	2.02	1.79	
57 Asn	8.59	4.87	3.64/3.81	γNH$_2$ 6.68/7.10	
58 Glu	7.95	4.67	1.95/2.02	2.23	
59 Leu	7.96	5.20	1.98	1.47	δ 0.46/0.70
60 Thr	8.48	4.70	4.17	1.24	
61 Tyr	8.48	5.55	2.68/2.83		δ 6.68; ε 6.60
62 Tyr	8.59	4.54	2.76/2.96		δ 6.78; ε 6.95
63 Cys	8.32	5.87	2.97/3.50		
64 Cys	9.02	4.95	3.26/3.39		
65 Lys	8.99	4.76	1.63/2.11	1.53	δ 1.23/1.42; ε 2.44
66 Lys	7.25	4.59	1.61/1.72	1.41/1.46	δ 1.70; ε 3.00
67 Asp	8.16	4.86	2.33/2.44		
68 Leu	9.27	3.24	1.25/1.80	1.05	δ –0.20/0.66
69 Cys	7.51	4.67	3.31/3.67		
70 Asn	8.62	4.59	2.11/3.02		

[a] The chemical shifts are referenced with respect to 1,4-dioxane at 3.74 ppm and are accurate to ±0.02 ppm.

Figure 5 illustrates a number of the interstrand C$^\alpha$H–C$^\alpha$H NOEs observed for sCD59 that are characteristic of antiparallel β-sheet secondary structure. The C and D strands are connected by a four-residue turn between residues 31 and 34. The long loop connecting strands D and E contains an element of α-helical structure between residues 48 and 55

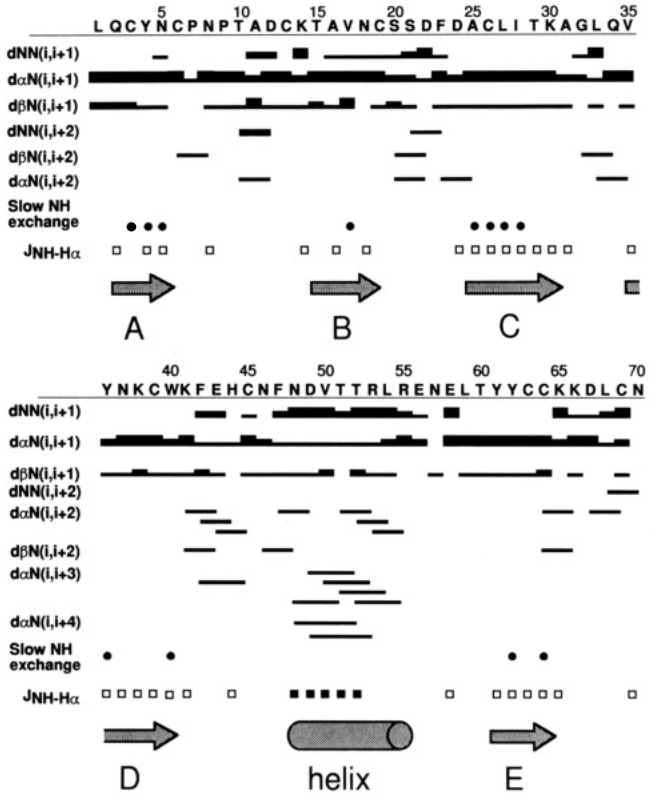

FIGURE 3: Overview of all the NMR data defining the secondary structure of sCD59. The sequential NOEs have been observed on NOESY spectra recorded at 278, 298, 308, and 315 K and pH 5.5. The closed circles indicate the amide protons which have a lifetime longer than 2 h in a H_2O/D_2O exchange experiment at 298 K. $^3J_{H^NH^\alpha}$ coupling constants are labeled with open squares when they are higher than 8 Hz and with closed squares when they are lower than 6 Hz as measured on a high-resolution COSY spectrum at 308 K. An interpretation of the secondary structure elements of sCD59 based on the experimental observations is shown in schematic form at the bottom of the figure.

[identified on the basis of the stretch of $d_{NN}(i,i+1)$, $d_{\alpha N}(i,i+2)$, $d_{\alpha N}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ short-range NOEs and low values of the $^3J_{H^NH^\alpha}$ coupling constants]. The helical element superposes one side of the triple-stranded β-sheet, shown schematically in Figure 4. The correspondence between the secondary structure analysis and the $C^\alpha H$ chemical shift index analysis (data not shown) is good but not perfect. Strand B is not clearly identified in the latter analysis and the end points of the other secondary structure elements differ slightly.

Three-Dimensional Structure Determination. Three-dimensional structures of sCD59 were calculated on the basis of the experimentally derived distance and torsion angle restraints using dynamical simulated annealing with the program X-PLOR 3.0 (Brünger, 1992). The protocol adopted is essentially that described in Brünger (1992) for the calculation using simulated annealing starting from a template structure with randomized backbone ϕ and ψ torsion angles. The procedure involves several picoseconds of high-temperature dynamics under conditions of low nonbonded interaction terms, during which the polypeptide chain rearranges to satisfy the restraint terms of the force field. The system is then cooled with concomitant increase in the nonbonded contact "repel" force constant, followed by conjugate gradient minimization. This procedure was used to generate a series of model structures with a wide range of conformations, only a subset of which are in reasonable agreement with the input restraints. Acceptable structures produced at this stage were then refined using further simulated annealing procedures to overcome local energy barriers to convergence. A geometric force field comprising very high bond and angle force constants (to maintain covalent geometry and chirality) and a simple quartic repel van der Waals term were used throughout the calculations. No explicit H-bond or electrostatic terms were used.

Disulfide Bonding. CD59 contains 10 Cys residues (Cys-3, -6, -13, -19, -26, -39, -45, -63, -64, and -69) which could potentially combine to form five disulfide bonds. During the initial structure calculation a strategy was used that made no assumption about the disulfide-bonding pattern in the molecule. Thus, in the first rounds of the calculation no explicit S–S bonding terms were included, and structures were obtained that satisfied the experimental restraints but contained no covalent disulfide bonds. The analysis of the structures obtained showed that only a very limited number of disulfide-bonding patterns could be accommodated given the topology of the molecule.

Two disulfide bonds, between Cys-6 and Cys-13 and between Cys-45 and Cys-63, could be identified unambiguously on the basis of these calculations. In each case the Cys residue side chain has no other Cys residue within bonding distance. The first of these disulfides links the two strands (A and B) of the N-terminal β-sheet finger, whereas the second links the long loop between β-strands D and E and the side of the triple-stranded β-sheet (strand E) that adjoins the helical segment within the loop. Of the other Cys residues, Cys-39, Cys-26, and Cys-69 are located side by side on the opposite side of the β-sheet, one in each strand and in positions immediately opposite Cys-19, Cys-3, and Cys-64, respectively. The pairing Cys-3, -26, Cys-6, -13, Cys-19, -39, Cys-45, -63, and Cys-64, -69 was finally adopted as the correct disulfide map for CD59 on the basis of stereochemical and steric considerations (Thornton, 1981). This information was incorporated into the final round of calculations by inclusion of an explicit S–S bonding term into the geometric force field during the refinement stages.

Attempts were made, however, to confirm this arrangement independently using traditional protein chemical methods. To determine whether all of the cysteines are disulfide bonded, sCD59 was labeled with [^{3}H]iodoacetic acid in 4 M guanidinium chloride with or without prior reduction with dithiothreitol. Chromatographic analysis indicated that only ~0.5% of the cysteines was accessible to the labeled alkylating agent under nonreducing conditions, clearly establishing that all of the cysteines are disulfide bonded (data not shown). Attempts to assign the disulfide-bonding pattern by enzymatic digestion and N-terminal sequencing were less successful. The ideal digestion regime in which trypsin and then α-chymotrypsin are used at pH 8 gave clear evidence of disulfide rearrangement. In an attempt to minimize disulfide interchange during the digestion, endoproteinase Glu-C and pepsin were used at pH 4 and 2, respectively, but under these conditions sCD59 was refractory to cleavage by these enzymes (data not shown). The arrangement of the disulfide bonds determined by the structure calculations is homologous to that determined for the related snake venom neurotoxins and cytotoxins (Endo & Tamiya, 1987; Dufton & Hider, 1988) and the same as that given in a preliminary report of the disulfide bonding pattern in CD59 by Tomita *et al.* (1991).

Structure Calculations. The calculation of the 3D solution structure of sCD59 was performed in an iterative manner (Kraulis *et al.*, 1989) with successive rounds of recalculation and subsequent incorporation of initially ambiguous distance restraints. The interproton distance restraints were classified into three classes on the basis of the cross-peak intensity in

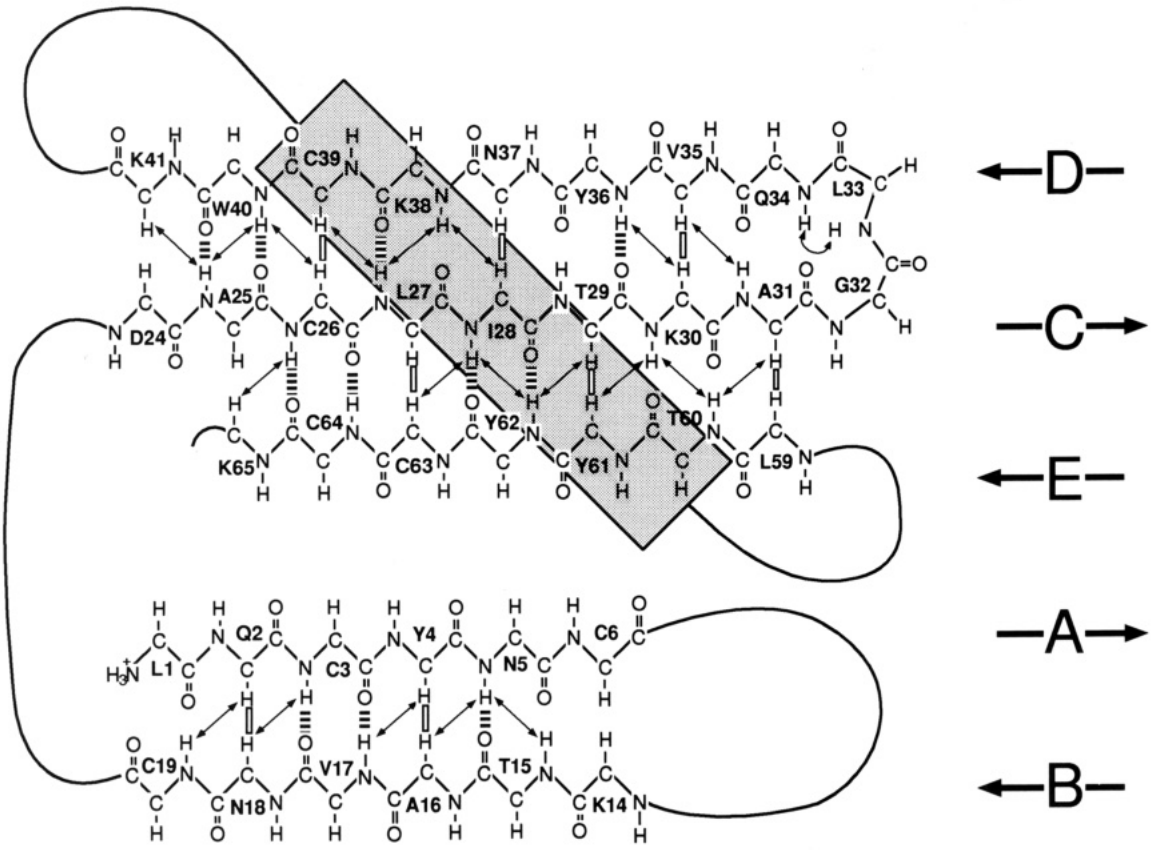

FIGURE 4: Secondary structure topology of sCD59 deduced from qualitative analysis of the NOE, $^3J_{H^NH^\alpha}$ coupling constant, and slowly exchanging amide proton data. The interstrand NOEs are shown by plain lines, and the $C^\alpha H$–$C^\alpha H$ connectivities shown in Figure 5 are indicated by open bars. The hydrogen bonds deduced from the NH exchange experiment are labeled with dashed bars. The gray rectangle signifies the presence of a helical region within the segment linking residues 41 and 59. The position and scale of the helix have not been respected for purposes of clarity.

150-ms NOESY spectra recorded at 600 MHz, 35 °C, and pH 5.5. Cross-peak intensities were quantified by counting contour levels. The intensities were interpreted in terms of conservative upper limit distance restraints using a distance/intensity calibration scale based on the intraresidue NH(i)-$C^\alpha H(i)$ cross-peak volumes for residues within regular secondary structure elements. Three distance classes were used in the structure calculations called strong, medium, and weak, corresponding to upper limit interproton distance restraint bounds of 2.7, 3.7, and 5.0 Å, respectively. The distance restraints were adjusted for groups with degenerate NMR resonances with the appropriate pseudoatom corrections (Wüthrich, 1986; Billeter *et al.*, 1982) combined with center averaging. An additional 0.5-Å correction was applied to distance restraints involving methyl groups (Wagner *et al.*, 1987). The lower limit distance bound is effectively the sum of the van der Waals radii. Eleven sets of H-bond distance restraints were incorporated on the basis of the observation of a slowly exchanging backbone amide NH resonance and the interpretation of the pattern of short-, medium-, and long-range backbone–backbone NOEs in terms of regular secondary structure elements, as shown in Figure 4. Three additional pairs of H-bond restraints were also incorporated, corresponding to H-bonds between the NH and O atoms of residues 38 and 27, 29 and 36, and 30 and 60, respectively. The interatomic distance restraints used were 1.7 Å < H–O < 2.2 Å and 1.8 Å < N–O < 3.3 Å. All the H-bond restraints were only included after rounds of structure calculations, performed without explicit H-bond restraints. The structures obtained without H-bond restraints retained the same overall polypeptide topology and allowed the unambiguous identification of antiparallel β-sheet secondary structure elements. The inclusion of H-bond distance restraints is required within the X-PLOR simulated annealing protocol to improve the local precision of the secondary structure, where this is known (the force field potential used in the simulated annealing does not include any explicit H-bond or electrostatic terms). Twenty-eight dihedral ϕ torsion angle restraints were included for residues with measured values of the $^3J_{H^NH^\alpha}$ coupling constant, obtained by line-shape fitting, at the extremes of the Karplus curve. An additional seven ϕ torsion angle restraints were included for residues 25, 27, 29, 31, 37, 38, and 39. For these residues, the SCUBA-COSY NH(i)–$C^\alpha H(i)$ cross-peaks partially overlapped and did not yield to the full line-shape analysis. In these cases the ϕ torsion angle restraints were included on the basis of qualitative analysis of the COSY cross-peak shape. The restraints were incorporated in a pseudo-square-well manner as follows: for $^3J_{H^NH^\alpha} > 8.0$ Hz, $-90° < \phi < -150°$; for $^3J_{H^NH^\alpha} < 6.0$ Hz, $-90° < \phi < -30°$. A total of 861 experimental restraints were used for the structure calculations, comprising 798 NOE-derived interproton distance restraints [253 intraresidue, 201 sequential, 95 medium range (i.e., for two residues numbered i,j: $1 < |i - j| < 5$) and 249 long range ($|i - j| > 5$)], 28 interatomic distance restraints for 14 H-bonds, and 35 ϕ torsion angle restraints; 222 of the interproton distance restraints (mostly intraresidue) may be considered redundant because the upper limits lie above accessible values for a polypeptide chain with idealized geometry. These restraints were included in the simulated annealing, however, as it has been observed that they aid the convergence properties of the calculations.

From a calculation involving 50 starting conformers, the 10 conformers ($\langle SA \rangle$) with the lowest values of the NOE (F_{NOE}) and constrained dihedral angle (F_{tor}) terms of the

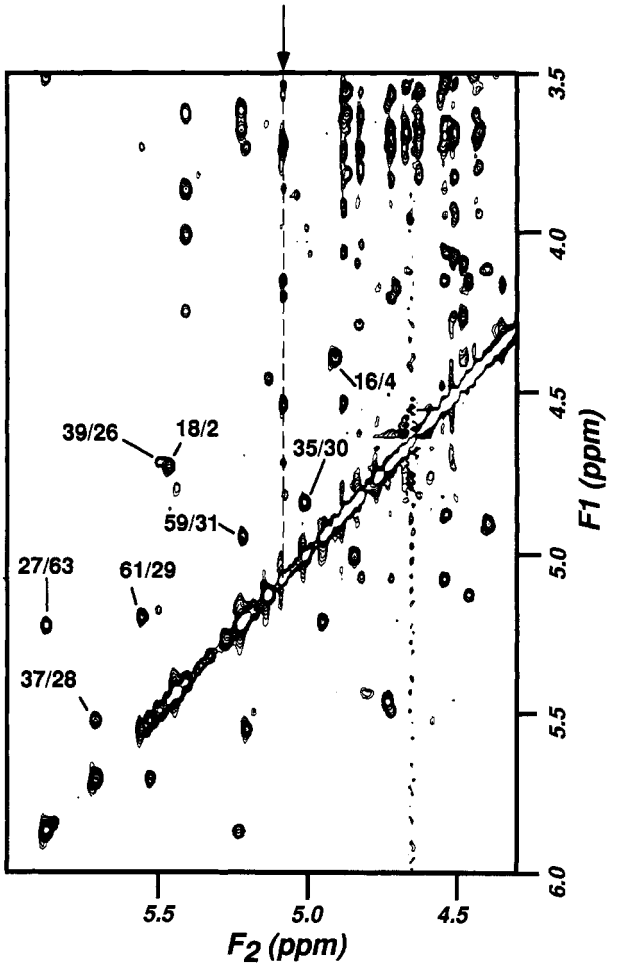

FIGURE 5: A portion of the 750-MHz 2D NOESY spectrum of sCD59 in D_2O at pD 5.5 (pH meter reading) and 308 K with a mixing time of 150 ms. The interstrand $C^{\alpha}H_i$–$C^{\alpha}H_j$ NOEs are labeled i/j. The dashed line connects a set of cross-peaks derived from an identified glycan sugar spin system.

force field target function were selected for further analysis. Each of these 10 structures satisfied the input experimental distance restraints with no single NOE violation greater than 0.4 Å and no single dihedral violation greater than 8°. These structures all demonstrated good covalent geometry and good nonbonded contacts as judged by the root mean square deviations (rmsd) from idealized geometry and low values for the Lennard-Jones van der Waals energy computed for each structure using the CHARMM force field (Brooks *et al.*, 1983). A geometrically averaged structure was obtained from coordinate averaging the pairwise best fit superposition of the 10 final structures. This average structure was regularized by 500 steps of restrained energy minimization to provide a representative structure $\overline{(SA)}_r$ which satisfies the experimental restraints as well as any of the individual structures from the final set. Table 2 lists the statistics of the structure calculations, including measures of the rmsd of the final structures from the experimental restraints and ideal geometry, as well as the final values of the restraint terms in the repel force field. The values of the Lennard-Jones van der Waals energy $E_{L\text{-}J}$ are also listed. The negative value of $E_{L\text{-}J}$ obtained for sCD59 indicates that the structures possess good nonbonded contacts, despite the absence of the Lennard-Jones potential from the force field used in the calculations.

A stereo representation of the best fit superposition of the 10 structures upon the mean coordinate positions (for residues 1–4 and 17–70) is shown in Figure 6. Orthogonal views of the superposition show that the structure has a disk-like shape with dimensions 30 Å × 30 Å × 15 Å. The center of the three-dimensional structure is composed of the three-stranded antiparallel β-sheet [residues 25–30 (strand C), 35–40 (strand D), and 61–64 (strand E); see Figure 3]. This sheet is overlaid on either side by the N-terminal β-sheet finger (residues 2–18) and the long loop between strands D and E (residues 41–60). The loop contains a short segment (approximately two turns) of helical structure (residues 48–55) which packs against one face of the three-stranded β-sheet, with hydrophobic residues Leu-27, Thr-29, Tyr-36, Trp-40, Val-50, Thr-51, Leu-54, and Tyr-61 buried in the interface and residues Asp-49, Thr-52, Arg-53, and Arg-55 exposed to the solvent. The triple-stranded β-sheet can be easily recognized in Figure 6, as well as the long loop with the short helix. The rmsd in atomic positions for the pairwise best fit superposition of each structure with the mean coordinate positions is 2.06 ± 0.39 Å for all atoms and 1.36 ± 0.36 Å for the backbone atoms (N, C^{α}, C, and O). The relative position of the N-terminal loop (residues 6–14) with respect to the rest of the sCD59 structure is not well-defined by the experimental restraints. For the well-defined part of the structure comprising residues 1–4 and 17–70 the corresponding rmsd figures are 1.68 ± 0.23 and 0.90 ± 0.13 Å, respectively. These values indicate that the position of the first β-sheet (strands A and B) finger is less well-defined than the center of the structure. Very few NOEs are observed between these two parts of the structure, which may be indicative of flexibility of the β-sheet finger relative to the rest of the molecule. Stereospecific constraints have not yet been included in the data set, which is reflected by the relatively high values of all-atoms rmsd. Further refinement of the structure is currently limited by crowding in the aliphatic region of the 2D spectrum due to the presence of a complex and potentially heterogeneous *N*-glycan moiety attached at residue Asn-18. The preparation of sCD59 with either a single small defined glycan or no glycosylation at Asn-18 has been initiated in order to address this issue.

The side-chain conformation of Asn-18 is not well-defined in the calculated structures, but the location of Asn-18 suggests that the N-linked glycan at this site may occlude one face of the CD59 disk-like structure (Figure 7). An analysis of the distribution of the surface electrostatic potential of sCD59 indicates that most of the charged residues are located at the perimeter of the domain (Figure 7). An unusual feature of the unglycosylated face of the molecule is a series of solvent-exposed Tyr residues (Tyr-4, Tyr-61, and Tyr-62) and a Phe residue (Phe-47) which form a relatively hydrophobic strip across the protein face (Figure 7, left view).

A comparison of the structure of sCD59 with other known three-dimensional protein folds reveals that the overall architecture and disulfide-bonding patterns of the snake venom neurotoxins are essentially identical to those found for sCD59. A ribbon diagram of the restrained minimized average structure of sCD59 made using the program MOLSCRIPT (Kraulis, 1991) illustrates the position of the secondary structure elements identified in this study, and this is compared with the 2.4-Å resolution X-ray structure of *Naja naja siamensis* α-cobratoxin (Brookhaven Protein Data Bank code 2CTX) (Betzel *et al.*, 1991) in the same orientation as the structure of sCD59 (Figure 8). The rmsd for the best fit superposition of 29 C^{α} atoms in the core regions of CD59 $\overline{(SA)}_r$ and 2CTX is 1.27 Å. An alignment of representative members of the Ly6SF and cytotoxin/neurotoxin family, showing the positions within the sequence of the secondary structure elements of sCD59 and α-cobratoxin, is given in

Table 2: Structural Statistics and Atomic Root Mean Square Deviations

structure[a]	⟨SA⟩	$(\overline{SA})_r$
rms deviations from exptl distance restraints (Å)		
all (826)[b]	0.042 ± 0.003	0.036
intraresidue (253)	0.033 ± 0.003	0.034
sequential ($\|i-j\| = 1$) (201)	0.038 ± 0.006	0.036
medium range ($2 \leq \|i-j\| \leq 4$) (95)	0.053 ± 0.007	0.053
long range ($5 \leq \|i-j\|$) (249)	0.043 ± 0.006	0.031
H-bond (28)	0.021 ± 0.008	0.026
rms deviations from exptl dihedral restraints (deg)		
ϕ torsion angle restraints (35)[b]	1.82 ± 0.28	2.02
residual terms of target function containing the repel potential[c] (kcal·mol^{-1})		
F_{NOE}	43.0 ± 6.0	32.0
F_{tor}	7.2 ± 2.1	8.7
F_{vdW}	42.9 ± 12.1	27.0
Lennard-Jones van der Waals energy[d] (kcal·mol^{-1})		
E_{L-J}	−137.1 ± 23.2	−166.4
deviations from idealized covalent geometry		
bonds (Å)	0.0044 ± 0.0005	0.004
angles (deg)	0.706 ± 0.043	0.633
impropers (deg)	0.601 ± 0.052	0.542

[a] ⟨SA⟩ represents the 10 individual converged sCD59 structures; $(\bar{S}A)_r$ is the restrained minimized average structure derived from the ⟨SA⟩ set by geometric averaging of the pairwise best fitted ⟨SA⟩ structures followed by 500 steps of restrained energy minimization. [b] The number of restraints is given in parentheses. [c] F_{NOE}, F_{tor}, and F_{vdW} represent the residual energies of the NOE restraint, dihedral torsion angle restraint, and quartic van der Waals repulsion terms of the simulated annealing target function, respectively. The final values of the target function force constants were k_{NOE} = 50 kcal·mol^{-1}·Å^{-2}, k_{tor} = 200 kcal·mol^{-1}·rad^{-2}, and k_{vdW} = 4.0 kcal·mol^{-1}·Å^{-4} (with van der Waals radii set to 0.8 times the standard values). [d] E_{L-J} is the Lennard-Jones van der Waals energy calculated with parameters taken from CHARMM version 19 (Brooks *et al.*, 1983). This term was not used in the target function of the structure calculations.

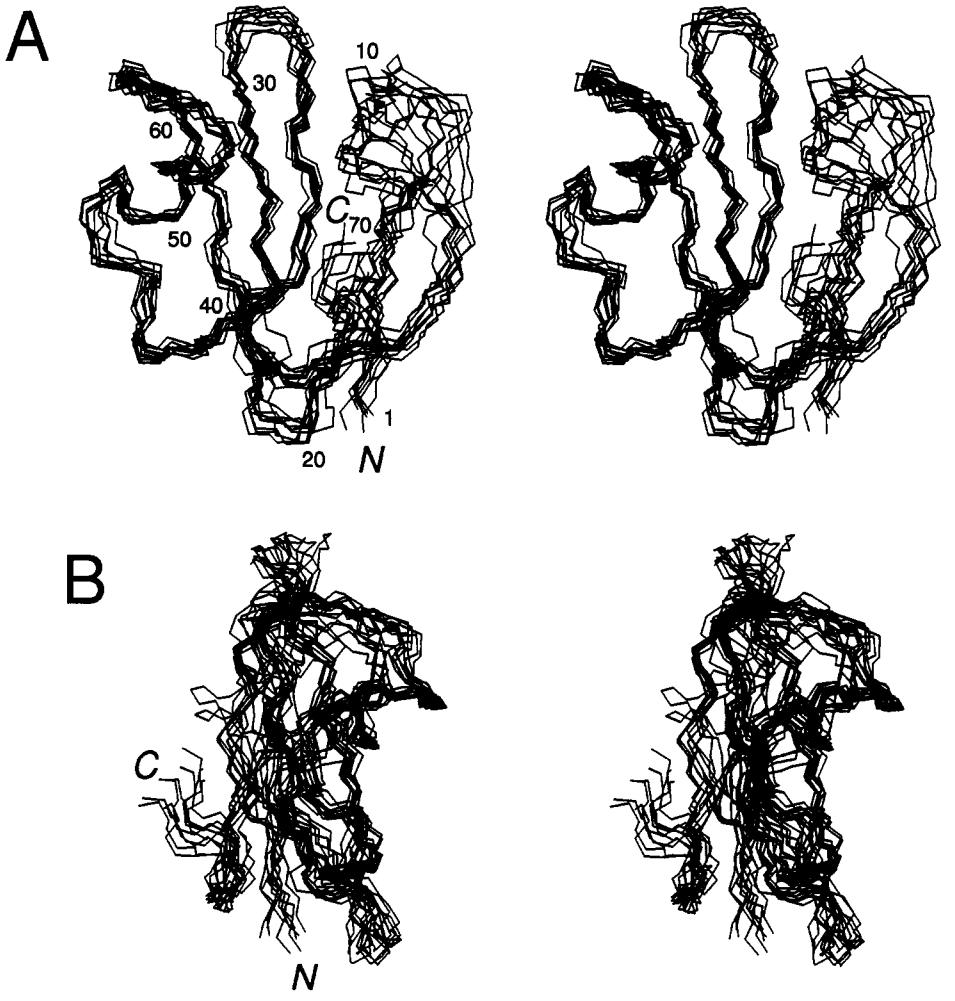

FIGURE 6: (A) Stereo image of the best fit superposition of the backbone atoms of the 10 ⟨SA⟩ structures of sCD59 computed from the NMR-derived experimental restraints. The structures are best fit over the backbone atoms of residues 1–4 and 17–70 on the coordinates of a geometric average of the 10 structures. (B) The same superposition as in (A) but rotated 90° about the vertical axis.

Figure 9. In addition to showing that the degree of similarity between the two groups of molecules is very limited, the alignments indicate that the variation in the sequences is greatest in the segments forming loops between the β-strands.

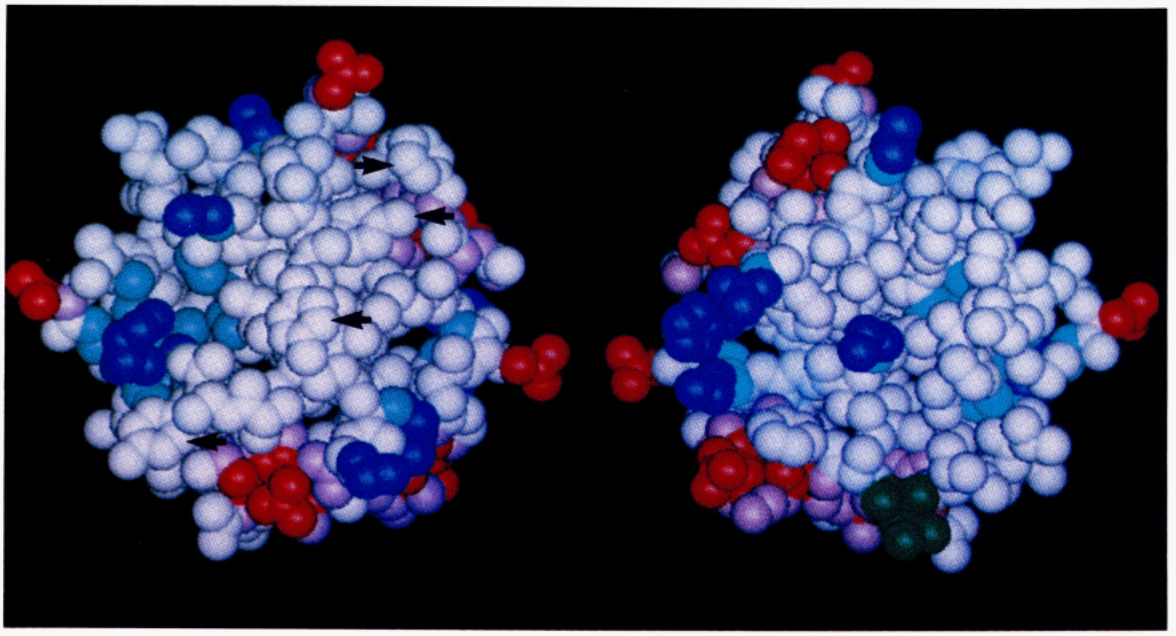

FIGURE 7: Surface electrostatic potential of sCD59. The two views of the structure, shown in CPK format, are related by a 180° rotation about the vertical axis; the right-hand view is similar to that in Figure 6A. The potential surfaces were computed with the program DELPHI (Gilson *et al.*, 1987) using full charges. Blue represents positive potential, red negative, and white neutral. The Asn-18 side chain that is glycosylated is colored green. The arrows in the left view point to Tyr-4, Tyr-61, Tyr-62, and Phe-47.

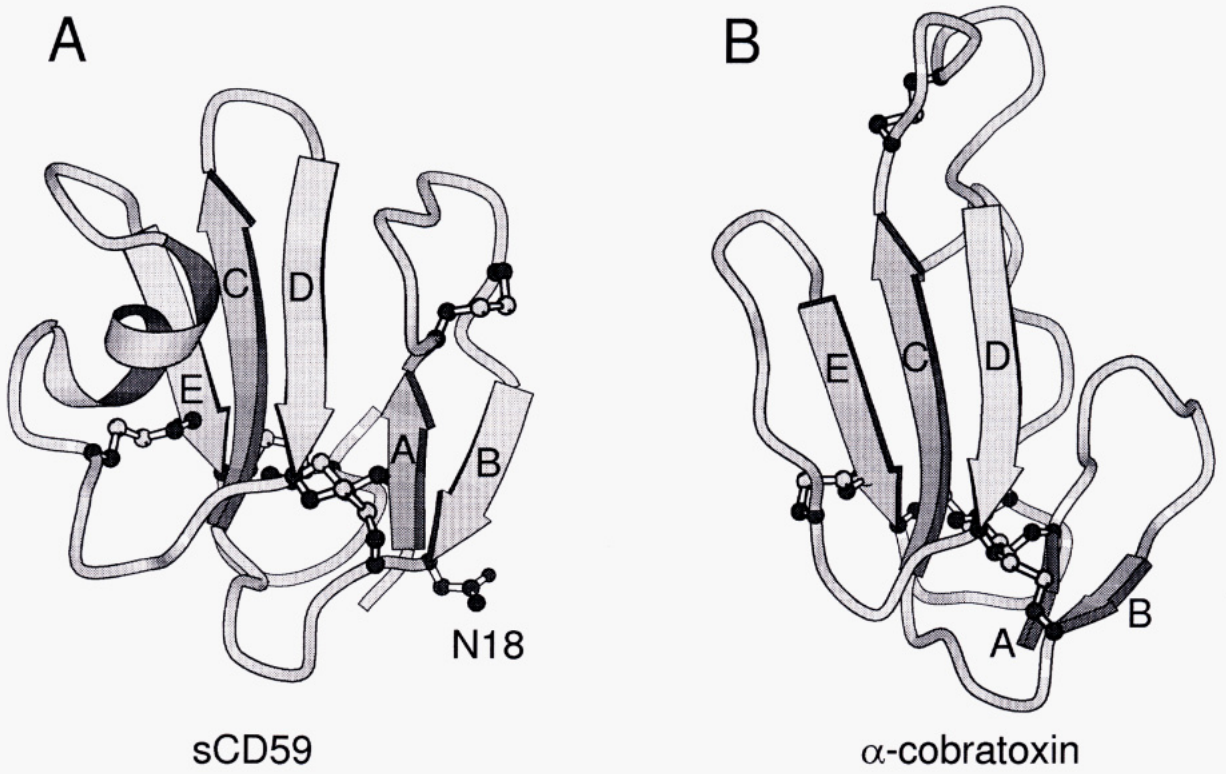

FIGURE 8: MOLSCRIPT pictures of (A) the restrained minimized average structure $(\overline{SA})_r$ of sCD59 and (B) the X-ray crystal structure of *N. naja siamensis* α-cobratoxin (PDB code 2CTX) showing the secondary structure elements and the positions of the disulfide bond connections. The structures are shown in the same orientation to emphasize the similarity of the overall topology, disulfide pattern, and core secondary structure; the orientation of sCD59 is similar to that in Figure 6A and Figure 7 (left). The side chains of the disulfide-bonded cysteine residues and the glycosylated Asn-18 residue of sCD59 are shown in "ball and stick" form.

DISCUSSION

The principal objectives of the present NMR analysis of the extracellular region of CD59 were to provide both a molecular basis for understanding the function of CD59 and a prototypic structure for Ly6SF domains. These objectives were achieved in obtaining the low-resolution solution structure described above. The structure consists of a central three-stranded β-sheet sandwiched between a long helix-containing loop on one side and an antiparallel β-sheet finger on the other. Together, these structural elements form a relatively flat, disk-like domain with dimensions of 30 Å × 30 Å × 15 Å. While the conformations of the central sheet and the long helix-containing loop are reasonably well-defined by the NMR data, the conformation of the end of the β-sheet finger and

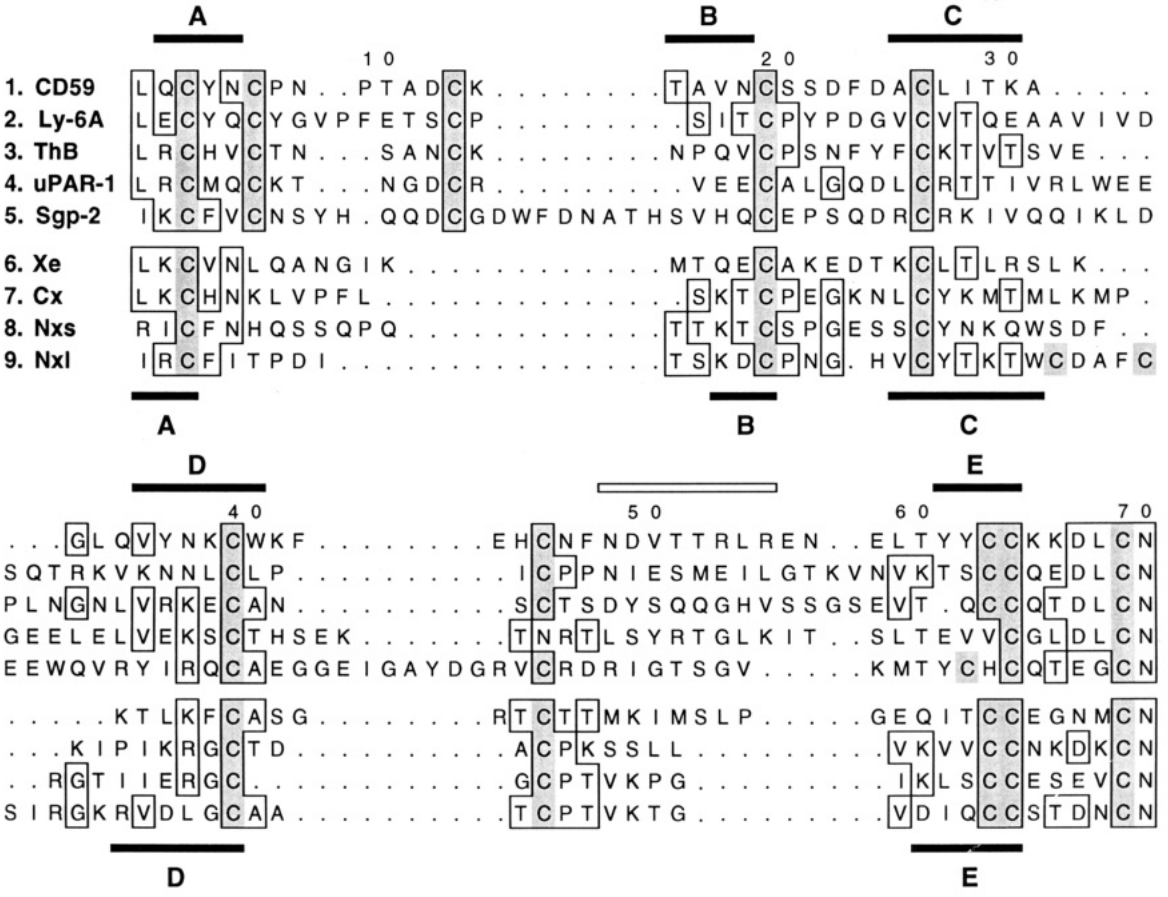

FIGURE 9: Sequence alignments and secondary structure. Sequences of representative members of the Ly6SF (sequences 1–5) and the cytotoxin/neurotoxin families (sequences 6–9) were aligned by eye. The sequences, with numbering based on the actual or predicted mature protein sequences and obtained from the SWISSPROT and PIR databases, are for CD59 (P13987) residues 1–70, Ly-6A (P05533) residues 1–79, ThB (A46528) residues 1–71, urokinase plasminogen activator receptor domain 1 (uPAR-1; A39743) residues 1–77, squid glycoprotein-2 (Sgp-2) residues 1–92 from Williams *et al.* (1988), *Xenopus laevis* xenoxin-1 (Xe) residues 1–66 from Kolbe *et al.* (1993), *Hemachatus haemachatus* cytotoxin-1 (Cx; P01471) residues 1–61, *Laticauda semifasciata* short neurotoxin-1 (Nxs, erabutoxin A; P01435) residues 1–61, and *N. naja siamensis* long neurotoxin (Nxl, α-cobratoxin; P01391) residues 1–63. Residues present at the same positions in three or more sequences and which include members of both the Ly6SF and the cytotoxin/neurotoxin families are boxed. Cysteine residues are shaded. The residues forming β-strands or a helix in CD59 and the α-cobratoxin (Nxl) sequences are highlighted with filled and open bars, respectively. CD59 residues are numbered above the sequence.

its position with respect to the core of the protein are less well-defined. This lack of definition appears to be due to the paucity of NOE restraints between the two structural elements and not due to any difficulties arising out of the possible oligosaccharide heterogeneity of sCD59 or the potential for *cis–trans* peptide bond isomerism of the two Xxx–Pro linkages in the β-sheet finger.

In its native state CD59 is a 20-kDa glycoprotein attached to the membrane by a GPI anchor. The structure of the extracellular region of CD59 appears to be unaffected by truncation at Asn-70 as kinetic analyses have indicated that the YTH53.1 (Davies *et al.*, 1989), MEM-43 (Štefanová *et al.*, 1989) and H19 (Groux *et al.*, 1989) antibody-binding properties of the recombinant molecule are indistinguishable from those of CD59 purified from human urine (P. A. van der Merwe, B. P. Morgan, and S. J. Davis, unpublished data). Interestingly, the kinetic study also indicated that the recombinant forms of sCD59 and CD59 purified from urine both fail to bind to CD2, a result which is consistent with the work of Arulanandam *et al.* (1993). The high level of conservation of the cysteines at the N- and C-termini of each Ly6SF sequence (Figure 9), which are likely to be organized into disulfide bonds as in CD59, suggests that in all cases Ly6SF domains are structurally delimited in a manner analogous to CD59. The Ly6SF domain of CD59 is either attached directly to the GPI anchor by analogy to Sgp-2 (Williams *et al.*, 1988) or attached via a short stalk, depending on whether Asn-70 or Asn-77 is the site of attachment of the GPI anchor. The uncertainty concerning the site of attachment of the GPI anchor makes it very difficult to predict the orientation of the extracellular Ly6SF domain of CD59 with respect to the cell surface.

The Ly6SF consists almost exclusively of molecules with single Ly6SF domains. It has been proposed that this is due to the absence of an intron between the Ly6SF domain and the GPI anchor attachment signal which would otherwise have allowed duplication and insertion of the gene segments encoding the extracellular region into preexisting genes (Williams, 1991). The enigmatic exception is the urokinase plasminogen activator receptor which has three Lys6SF domains. The present data suggest that relatively long linker regions separate domains 1 and 2 and domains 2 and 3 of the receptor (15 and 14 residues, respectively) and it is not obvious how these multiple Ly6SF domains are organized. One possibility is that the extra sequence forms C-terminal extensions similar to those in the long neurotoxins which place the N- and C-termini at opposite ends of the domain (Walkinshaw *et al.*, 1980; Nickitenko *et al.*, 1993).

Preliminary data suggest that the sCD59 used for the present study was fully glycosylated. While detailed characterization of the glycosylation of the sCD59 has thus far not been undertaken, it is anticipated that the oligosaccharides at this

site are predominantly of the complex type typically seen for recombinant glycoproteins expressed in CHO cells [reviewed by Ashford *et al.* (1993)]. Although their presence adds complexity to the 2D NMR spectra, the cross-peaks arising from glycan resonances could be readily distinguished from the sCD59 polypeptide cross-peaks because of their relatively narrow line widths. Figure 5 shows the identification of a particularly well-resolved set of glycan cross-peaks in the 2D D_2O NOESY spectrum of sCD59. The NH resonances of the *N*-acetyl groups from at least three *N*-acetylglucosamine residues, including two forming part of the pentasaccharide core common to all N-linked oligosaccharides (Kornfeld & Kornfeld, 1985), can also be distinguished in the fingerprint region of the spectrum in H_2O solution (see Figure 2). Moreover, the region of the spectrum between 4.9 and 4.5 ppm is populated by many sharp cross-peaks due to the connectivities among glycan proton resonances. As far as analysis of the glycan conformation was possible, only a single NOE was found that connected the glycan to the protein. This NOE is tentatively assigned to a connection between the *N*-acetyl NH resonance of the first core *N*-acetylglucosamine residue and the $C^{\alpha}H$ resonance of residue Val-17. The absence of an extensive network of glycan–protein NOE connectivities combined with the low level of chemical shift heterogeneity observed in the protein spectrum suggests that the oligosaccharide in sCD59 behaves as a flexible appendage with little evidence for a well-defined interaction with the protein. The N-linked glycan was not incorporated into the structure calculations. The solvent-exposed position of Asn-18 in the sCD59 structure, close to the N- and C-terminal residues, is indicated in Figures 7 and 8A. Analyses of the serum IgM tail-piece glycopeptide (Wormald *et al.*, 1991), the glycoprotein phosvitin (Brockbank & Vogel, 1990), and ribonuclease B (Berman *et al.*, 1981) also indicate that N-linked oligosaccharides can have considerable mobility and are largely unaffected by the presence of the protein.

It was proposed on the basis of sequence comparisons that the Ly6SF and the snake venom cytotoxin/neurotoxin family were evolutionarily related (Fleming *et al.*, 1993). The current structural analysis is consistent with this proposition. The nonneurotoxic cytotoxins, which cause muscle and nerve cell membrane depolarization, are polypeptides of 60–62 residues with four disulfides (Dufton & Hider, 1988). There are two classes of the nicotinic acetylcholine receptor-binding neurotoxins: short neurotoxins consist of 60–62 residues with four disulfides whereas long neurotoxins consist of 70–74 residues and have five disulfides (Endo & Tamiya, 1987). The conservation of both the cysteine spacing and amino acid sequences at the C-termini of the Ly6SF and neurotoxin/cytotoxin sequences implied that these two groups of molecules have similar structures (Fleming *et al.*, 1993). Numerous ^1H NMR and crystallographic studies of long neurotoxins (e.g., Love & Stroud, 1986; Betzel *et al.*, 1991; Le Goas *et al.*, 1992; Nickitenko *et al.*, 1993), short neurotoxins (e.g., Tsernoglou & Petsko, 1977; Smith *et al.*, 1988; Brown & Wüthrich, 1992), and a cytotoxin (Rees *et al.*, 1990) have been undertaken. The topologies of sCD59 and the α-cobratoxin are very similar, with each molecule consisting of two-stranded and three-stranded β-sheets (Figure 8). Three significant differences in the two structures are (1) sCD59 has an extension of the first β-sheet finger between strands A and B that includes two disulfide-bonded cysteines that are conserved within the Ly6SF and absent from the neurotoxin/cytotoxin sequences, (2) the loop between strands C and D is shorter in sCD59 than in the neurotoxin, and (3) the loop between strands D and E is significantly shorter in the neurotoxin and in sCD59 this loop has significant helical character. The squid glycoprotein Sgp-2, which is another putative member of the Ly6SF (Williams *et al.*, 1988), may represent a third structural class: the residue insertions between the third and fourth and the sixth and seventh cysteines of the Sgp-2 sequence, along with the displaced, presumably disulfide-bonded cysteine within β-strand E, may impose significantly different topology on this molecule.

In the short period since its discovery, considerable progress has been made in understanding the function of CD59. While it is clear that CD59 protects tissues from attack by the complement system, the molecular basis of this process remains to be elucidated. The unusually highly exposed Tyr and Phe residues forming the relatively hydrophobic strip across the unglycosylated face of CD59 may constitute a candidate binding site for C8 and/or C9 if this interaction has a hydrophobic component. The solution structure described herein provides a basis for testing this and other possible mechanisms of CD59 function.

ADDED IN PROOF

The inclusion of 40 mM iodoacetamide in protease digestion buffers to prevent disulfide bond rearrangement during conventional disulfide bond analysis has now allowed independent confirmation of the disulfide bond arrangement assigned on the basis of the structure calculations.

ACKNOWLEDGMENT

The CD59 cDNA and YTH53.1 hybridoma were obtained from Dr. David L. Simmons (Institute of Molecular Medicine, Oxford, U.K.) and Prof. Herman Waldmann (School of Pathology, Cambridge, U.K.), respectively. We thank Dr. Jonathan Boyd and Nick Soffe for recording the 750-MHz NMR spectra, Dr. A. Neil Barclay for help with the sequence analysis, Dr. Dale L. Bodian for assistance with the DELPHI calculation, and Elizabeth Davies for help with the purification of the sCD59.

REFERENCES

Arulanandam, A. R. N., Moingeon, P., Concino, M. F., Recny, M. A., Kato, K., Yagita, H., Koyasu, S., & Reinherz, E. L. (1993) *J. Exp. Med. 177*, 1439–1450.

Arvieux, J., & Williams, A. F. (1988) *Antibodies: A Practical Approach*, Chapter 5, pp 113–136, IRL Press, Oxford.

Ashford, D. A., Alafi, C. D., Gamble, V. M., Mackay, D. J. G., Rademacher, T. W., Williams, P. J., Dwek, R. A., Barclay, A. N., Davis, S. J., Somoza, C., Ward, H. A., & Williams, A. F. (1993) *J. Biol. Chem. 268*, 3260–3267.

Barclay, A. N., Beyers, A. B., Birkeland, M. L., Brown, M. H., Davis, S. J., Somoza, C., & Williams, A. F. (1992) *The Leukocyte Antigen Factsbook*, Academic Press, London.

Bax, A. (1989) *Methods Enzymol. 176*, 151–168.

Bax, A., Ikura, M., Kay, L. E., & Zhu, G. (1991) *J. Magn. Reson. 91*, 174–178.

Bebbington, C. R., & Hentschel, C. C. G. (1987) in *DNA Cloning: A Practical Approach* (Glover, D. M., Ed.) Vol. III, pp 163–188, IRL Press, Oxford.

Berman, E., Walters, D. E., & Allerhand, A. (1981) *J. Biol. Chem. 256*, 3853–3857.

Betzel, C., Lange, G., Pal, G. P., Wilson, K. S., Maelicke, A., & Saenger, W. (1991) *J. Biol. Chem. 266*, 21530–21536.

Billeter, M., Braun, W., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 321–346.

Braunschweiler, L. R., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Brockbank, R. L., & Vogel, H. J. (1990) *Biochemistry 29*, 5574–5583.
Brooimans, R. A., van Wieringen, P. A. M., van Es, L. A., & Daha, M. R. (1992) *Eur. J. Immunol. 22*, 3135–3140.
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.
Brown, L. R., & Wüthrich, K. (1992) *J. Mol. Biol. 227*, 1118–1135.
Brown, S. C., Weber, P. L., & Mueller, L. (1987) *J. Magn. Reson. 77*, 166–169.
Brünger, A. T. (1992) *X-PLOR Manual, Version 3.0*, Yale University, New Haven, CT.
Clark, S. J., Jefferies, W. A., Barclay, A. N., Gagnon, J., & Williams, A. F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1649–1653.
Davies, A., Simmons, D. L., Hale, G., Harrison, R. A., Tighe, H., Lachmann, P. J., & Waldmann, H. (1989) *J. Exp. Med. 170*, 637–654.
Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820–2821.
Davis, S. J., Ward, H. A., Puklavec, M. J., Willis, A. C., Williams, A. F., & Barclay, A. N. (1990) *J. Biol. Chem. 265*, 10410–10418.
Deckert, M., Kubar, J., Zoccola, D., Bernard-Pomier, G., Angelisova, P., Hořejší, V., & Bernard, A. (1992) *Eur. J. Immunol. 22*, 2943–2947.
Driscoll, P. C., Clore, G. M., Beress, L., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2178–2187.
Dufton, M. J., & Hider, R. C. (1988) *Pharmacol. Ther. 36*, 1–40.
Endo, T., & Tamiya, N. (1987) *Pharmacol. Ther. 34*, 403–451.
Fleming, T. J., O'hUigin, C., & Malek, T. R. (1993) *J. Immunol. 150*, 5379–5390.
Gilson, M., Sharp, K., & Honig, B. (1987) *J. Comput. Chem. 9*, 327–335.
Groux, H., Huet, S., Aubrit, F., Tran, H. C., Boumsell, L., & Bernard, A. (1989) *J. Immunol. 142*, 3013–3020.
Gumley, T. P., McKenzie, I. F. C., Kozak, C. A., & Sandrin, M. S. (1992) *J. Immunol. 149*, 2615–2618.
Hahn, W. C., Menu, E., Bothwell, A. L. M., Sims, P. J., & Bierer, B. E. (1992) *Science 256*, 1805–1807.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Kolbe, H. V. J., Huber, A., Cordier, P., Rasmussen, U. B., Bouchon, B., Jaquinod, M., Vlasak, R., Délot, E. C., & Kreil, G. (1993) *J. Biol. Chem. 268*, 16458–16464.
Kornfeld, S., & Kornfeld, R. (1985) *Annu. Rev. Biochem. 54*, 631–664.
Korty, P. E., Brando, C., & Shevach, E. M. (1991) *J. Immunol. 146*, 4092–4098.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.
Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., & Gronenborn, A. M. (1989) *Biochemistry 28*, 7241–7257.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.
LeClair, K. P., Palfree, R. G. E., Flood, P. M., Hämmerling, U., & Bothwell, A. (1986) *EMBO J. 5*, 3227–3234.
Le Goas, R., LaPlante, S. R., Mikou, A., Delsuc, M.-A., Guittet, E., Robin, M., Charpentier, I., & Lallemand, J.-Y. (1992) *Biochemistry 31*, 4867–4875.
Love, R. A., & Stroud, R. M. (1986) *Protein Eng. 1*, 37–46.
Marion, D., & Bax, A. (1988) *J. Magn. Reson. 80*, 528–533.
Marion, D., Ikura, M., & Bax, A. (1989a) *J. Magn. Reson. 84*, 425–430.
Marion, D., Ikura, M., Tschudin, R., & Bax, A. (1989b) *J. Magn. Reson. 85*, 393–399.
Meri, S., Morgan, B. P., Davies, A., Daniels, R. H., Olavesen, M. G., Waldmann, H., & Lachmann, P. J. (1990) *Immunology 71*, 1–9.
Nickitenko, A. V., Michailov, A. M., Betzel, C., & Wilson, K. S. (1993) *FEBS Lett. 320*, 111–117.
Ninomiya, H., & Sims, P. J. (1992) *J. Biol. Chem. 267*, 13675–13680.
Ninomiya, H., Stewart, B. H., Rollins, S. A., Zhao, J., Bothwell, A. L. M., & Sims, P. J. (1992) *J. Biol. Chem. 267*, 8404–8410.
Palfree, R. G. E. (1991) *Immunol. Today 12*, 170–171.
Palfree, R. G. E., Sirlin, S., Dumont, F. J., & Hämmerling, U. (1988) *J. Immunol. 140*, 305–310.
Plateau, P., & Guéron, M. E. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.
Redfield, C., & Dobson, C. M. (1990) *Biochemistry 29*, 7201–7214.
Rees, B., & Bilwes, A., Samama, J. P., & Moras, D. (1990) *J. Mol. Biol. 214*, 281–297.
Roldan, A. L., Cubellis, M. V., Masucci, M. T., Behrendt, N., Lund, L. R., Dano, K., Appella, E., & Blasi, F. (1990) *EMBO J. 9*, 467–474.
Smith, J. L., Corfield, P. W. R., Hendrickson, W. A., & Low, B. W. (1988) *Acta Crystallogr. A44*, 357–368.
Štefanová, I., & Hořejší, V. (1991) *J. Immunol. 147*, 1587–1592.
Štefanová, I., Hilgert, I., Krištofová, H., Brown, R., Low, M. G., & Hořejší, V. (1989) *Mol. Immunol. 26*, 153–161.
Štefanová, I., Hořejší, V., Ansotegui, I. J., Knapp, W., & Stockinger, H. (1991) *Science 254*, 1016–1019.
Sugita, Y., Nakano, Y., & Tomita, M. (1988) *J. Biochem. 104*, 633–637.
Sugita, Y., Tobe, T., Oda, E., Tomita, M., Yasukawa, K., Yamaji, N., Takemoto, T., Furuichi, K., Takayama, M., & Yano, S. (1989) *J. Biochem. 106*, 555–557.
Thornton, J. M. (1981) *J. Mol. Biol. 151*, 261–287.
Tomita, M., Tobe, T., Choi-Miura, N., Nakano, Y., Kusano, M., & Oda, E. (1991) Proceedings of the 14th International Complement Workshop, Cambridge, U.K., Abstract No. 284.
Tsernoglou, D., & Petsko, G. A. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 971–974.
Wagner, G., & Wüthrich, K. (1982) *J. Mol. Biol. 155*, 347–366.
Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.
Walkinshaw, M. D., Saenger, W., & Maelicke, A. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2400–2404.
Walsh, L. A., Tone, M., Thiru, S., & Waldmann, H. (1992) *Tissue Antigens 40*, 213–220.
Williams, A. F. (1991) *Cell Biol. Int. Rep. 15*, 769–777.
Williams, A. F., Tse, A. G.-D., & Gagnon, J. (1988) *Immunogenetics 27*, 265–272.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1992) *Biochemistry 31*, 1647–1651.
Wormald, M. R., Wooten, E. W., Bazzo, R., Edge, C. J., Feinstein, A., Rademacher, T. W., & Dwek, R. A. (1991) *Eur. J. Biochem. 198*, 131–139.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, Inc., New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949–961.