Sequence-specific ^1H -NMR assignments and folding topology of human CD59

C.M. FLETCHER,¹ R.A. HARRISON,² P.J. LACHMANN,² AND D. NEUHAUS¹

¹MRC Laboratory of Molecular Biology, Cambridge, United Kingdom

²MRC Molecular Immunopathology Unit, Cambridge, United Kingdom

(RECEIVED July 2, 1993; ACCEPTED August 26, 1993)

Abstract

CD59 is a recently discovered cell-surface glycoprotein that restricts lysis by homologous complement and has limited sequence similarity to snake venom neurotoxins. This paper describes the first results of a two-dimensional NMR study of CD59 prepared from human urine. Nearly complete ^1H -NMR assignments were obtained for the 77 amino acid residues and partial assignments for the N-glycan and the glycosylphosphatidylinositol (GPI) anchor. These results together confirm that the C-terminal residue of the mature protein is Asn 77 and that the urine-derived form retains the nonlipid part of the GPI anchor. The data further indicate that the GPI anchor and possibly the N-glycan are structurally inhomogeneous and suggest that the phospholipid present in the intact GPI anchor was removed by phosphatidylinositol-specific phospholipase-D. The folding topology of the protein was determined from NOE enhancements and slowly exchanging backbone amide protons and consists primarily of five extended strands (denoted β_1 – β_5 in sequence order), arranged into separate two-stranded (β_1 and β_2) and three-stranded (β_3 – β_5) antiparallel β -sheets. The same folding topology is found in all of the snake venom neurotoxins whose structures have been determined. The region between the β_4 and β_5 strands has helical character, a feature that is not present in the neurotoxins but that is seen in the topologically similar wheat germ agglutinin.

Keywords: assignment; complement; folding topology; glycoprotein; homologous restriction; NMR spectroscopy; toxin-agglutinin motif

CD59¹ is a cell-surface glycoprotein that protects plasma and vascular endothelial cells from complement-mediated lysis (Sugita et al., 1988; Davies et al., 1989; Holguin et al., 1989; Okada et al., 1989a; Hamilton et al., 1990; Rollins & Sims, 1990; Whitlow et al., 1990). It does so by binding to C8 and/or C9 in the partially assembled membrane attack complex and thereby preventing further re-

cruitment and polymerization of C9 (Meri et al., 1990; Rollins et al., 1991). In vitro experiments have also implicated CD59 in T cell activation in various ways (Groux et al., 1989; Korty et al., 1991; Deckert et al., 1992; Hahn et al., 1992), but it has yet to be demonstrated that these functions of CD59 are significant in vivo.

The CD59 gene codes for a 128-residue polypeptide that is subsequently modified by removal of a signal peptide and of a C-terminal peptide; the C-terminal cleavage occurs during attachment of the GPI membrane anchor (Davies et al., 1989; Okada et al., 1989b; Stefanova et al., 1989). The nascent protein is further modified by the addition of an N-glycan to Asn 18 (Davies et al., 1989; Stefanova et al., 1989; Ninomiya et al., 1992). A preliminary report (Tomita et al., 1991) indicated that Asn 77 is the C-terminal residue of CD59 isolated from urine and that this soluble form of the protein is derived from membrane-bound CD59 by removal of the lipid part of the GPI anchor.

Whereas the amino acid sequence shows no similarity to other complement components and control proteins,

Reprint requests to: D. Neuhaus, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, United Kingdom.

Abbreviations: GPI, glycosylphosphatidylinositol; WGA, wheat germ agglutinin; AChR, acetylcholine receptor; NOE, nuclear Overhauser effect; 3D, three-dimensional; 2Q, double-quantum correlation; COSY, correlated spectroscopy; DQF-COSY, double-quantum-filtered correlated spectroscopy; PE-COSY, primitive exclusive correlated spectroscopy; TOCSY, total correlated spectroscopy; RELAY, relayed correlated spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; ROESY, rotating-frame Overhauser effect spectroscopy; HSQC, heteronuclear single-quantum coherence; HMQC, heteronuclear multiple-quantum coherence; TPPI, time-proportional phase incrementation; CYCLOPS, cyclically ordered phase sequence; TSP, [2,3- $^2\text{H}_4$]3-trimethylsilylpropionate.

¹In the scientific literature, CD59 has also been referred to as P-18, MACIF, M1RL, HRF-20, 1F5 antigen, MEM-43 antigen, H19, and protectin.

it has limited sequence homology to murine LY-6 antigen and the urokinase plasminogen activator receptor (Davies et al., 1989; Stefanova et al., 1989). Motif pattern searching using residues that are conserved among these three proteins revealed further similarities with the snake venom neurotoxin family of proteins (Harrison, 1992). The 3D structures of a number of these snake venom neurotoxins have been determined using X-ray crystallography and NMR (Bourne et al., 1985; Love and Stroud, 1986; Labhardt et al., 1988; Betzel et al., 1991; Oswald et al., 1991; le Du et al., 1992; Le Goas et al., 1992; Yu et al., 1993; and references therein). These structures display a common folding topology, the so-called toxin-agglutinin motif (Drenth et al., 1980), which is also found in snake venom cardiotoxins (Rees et al., 1990; Steinmetz et al., 1988; and references therein) and the domains of WGA (Wright, 1987, and references therein). Harrison (1992) observed that the 10 cysteine residues of CD59 are highly conserved in the snake venom neurotoxins and that their disulfide bonding pattern is similar or identical (the preliminary report of Tomita et al. [1991] describes the CD59 disulfide pattern with one ambiguity regarding the bonding of the two adjacent cysteines 63 and 64). These observations led Harrison (1992) to suggest a possible evolutionary relationship between CD59 and the snake venom neurotoxins.

In this paper we present NMR results obtained from samples of CD59 purified from human urine. For the protein component, we have obtained almost complete sequence-specific ^1H resonance assignments and have used these to characterize the secondary structure and folding

topology via NOE enhancements, slowly exchanging amide protons, and some preliminary data from structure calculations. We are currently extending these results to enable calculation of the high-resolution 3D structure of the protein. In addition, we present here a limited number of ^1H resonance assignments for the carbohydrate components of CD59 and discuss the information that the NMR data provide concerning the composition of these groups.

Results

Protein assignments

Protein assignments were obtained using standard methods (Wüthrich, 1986). Firstly, amino acid spin systems were identified as far as possible using through-bond connectivities, including information from a ^{13}C - ^1H HMQC spectrum optimized for observation of one-bond couplings (this spectrum was recorded at natural abundance of ^{13}C) as well as standard homonuclear experiments. Secondly, these spin systems were fitted to the known sequence by observation of sequential NOE connectivities (Fig. 1). Cases in which amino acid spin systems could not be distinguished unambiguously from carbohydrate signals using chemical shifts and spin-coupling topology alone were resolved during the sequential assignment stage, since only the amino acids showed a significant number of NOE connectivities to other amino acids. Our assignment of the protein NMR spectrum indicates that Asn 77 is the C-terminal residue, in agreement with the

Fig. 1. Sequential and medium-range NOE connectivities and slowly exchanging backbone amide protons. NOE connectivities are specified using the notation for the corresponding proton-proton distances (Wüthrich, 1986). For proline residues the sequential connectivities d_{Nd} , $d_{\alpha\delta}$, and $d_{\beta\delta}$ are substituted for d_{NN} , $d_{\alpha\text{N}}$, and $d_{\beta\text{N}}$, respectively. The height of each shaded bar is approximately proportional to the strength of the corresponding NOE enhancement. Connectivities corresponding to positions marked by an asterisk (for sequential connections) or a dotted line (medium-range connections) could not be identified unambiguously due to chemical shift degeneracy. Slowly exchanging backbone amide protons are indicated by an "x" in the row labeled "Exchange".

preliminary report of Noda et al. (1991). The protein assignments are listed in Table 1, and only those that proved difficult or unusual are described below.

Identification of the single glycine spin system Gly 32 was made difficult by the presence of carbohydrate signals. Firstly, the (H^α , H^α) crosspeak in COSY was completely obscured by signals from the non-anomeric ring protons of sugar residues. Secondly, assignment of the characteristic (NH , $H^\alpha + H^\alpha$) remote connectivity in the 2Q spectrum was complicated by signals in the same region from the ethanolamine and sialic acid groups. However, the present assignment may be considered secure on the basis of sequential NOE connectivities $d_{\alpha N}$ to Leu 33 and the presence of (H^α , C^α) crosspeaks at an appropriate ^{13}C chemical shift in the ^{13}C - 1H HMQC spectrum.

Among the methyl group-containing amino acids, the assignment of alanine and threonine spin systems was initially complicated by a signal that we have now assigned to a fucose residue (Fuc 1). Although the degenerate $C^\delta H_3$ resonances of Leu 1 occur in a particularly crowded region, their assignment is supported by the observation of crosspeaks at appropriate chemical shifts in the ^{13}C - 1H HMQC spectrum (Fig. 2).

Among amino acids with AMX spin systems, the NH signal of Phe 47 has not been located in any of the spectra (not even in NOESY spectra with jump-return excitation), possibly due to line broadening caused by rapid exchange with the solvent. The remainder of this spin system is firmly assigned on the basis of sequential NOE connectivities and intraresidue NOE connectivities from the H^α and H^β protons to the aromatic ring protons.

Particular difficulties were encountered in assigning residues 19–24. In the case of Cys 19, these difficulties resulted from coincidence of the H^α resonance with the H_2O signal at 42 °C, the temperature at which TOCSY spectra gave better results, and probably from a small value of $^3J_{NH\alpha}$. As a result, the only through-bond correlations involving the NH proton that we observed were a very weak (NH , H^α) crosspeak in the 2Q spectrum and a very weak (NH , H^β) crosspeak in a TOCSY spectrum. In the case of Ser 20 also, difficulties arose from coincidence of the H^α resonance with the H_2O signal at 42 °C. The weak NH signals of both Ser 20 and Ser 21 caused problems, and in the latter case it proved difficult to distinguish the (H^α , H^β) crosspeaks from carbohydrate signals. However, for both serine residues (H^β , C^β) crosspeaks consistent with our assignment were found in the ^{13}C - 1H HMQC spectrum. The assignments of Asp 24 are very tentative due to overlap of the putative NH resonance with that of its sequential neighbor Ala 25 at both 27 °C and 42 °C. As a result of these difficulties, the assignments of Asp 22 and Phe 23 rely heavily upon the assignment of the ring protons of Phe 23 obtained by eliminating the other phenylalanine and tyrosine residues.

The ^{13}C - 1H HMQC spectrum proved useful in two further ways. Assignment of the aromatic ring proton sig-

Fig. 2. Methyl group region of the ^{13}C - 1H HMQC spectrum optimized for observation of carbon-proton one-bond couplings and recorded at 42 °C. The signals labeled in the figure are from the methyl groups of amino acid and fucose residues in the major form of the molecule. The remaining signals are probably from the same groups in one or more degraded or minor forms. One of the $C^\delta H_3$ signals of each of L54, L68, and L27 resonate at higher field in the proton dimension and do not appear on this plot.

nals of tyrosine, tryptophan, and histidine to specific positions in the rings was made as usual by observation of the relative strength of intraresidue NOE enhancements. It proved possible to corroborate these assignments by the observation of crosspeaks in the ^{13}C - 1H HMQC spectrum at the characteristic ^{13}C chemical shifts for these positions. Similarly, the assignment of proline H^δ signals, initially based upon intraresidue and sequential NOE connectivities, was supported by the presence of (H^δ , C^δ) crosspeaks at appropriate ^{13}C chemical shifts.

Protein secondary structure and folding topology

Regions of regular secondary structure were identified by the observation of characteristic sequential and medium-range NOE connectivities (Wüthrich, 1986; Figs. 1, 3). Extended strands were identified by the presence of strong $d_{\alpha N}$ sequential connectivities and very weak or absent d_{NN} sequential connectivities. In the region between residues 41 and 55 we observed strong d_{NN} sequential connectivities and a small number of $d_{\alpha N}(i, i + 3)$, $d_{\alpha\beta}(i, i + 3)$, and $d_{\alpha N}(i, i + 4)$ connectivities, all characteristic

Table 1. Protein assignments at 42 °C and pH 5.8^a

Residue	NH	α	β	γ	δ	Others
Leu 1	—	4.08	1.68, 1.52	1.52	0.72, 0.72	
Gln 2	8.86	5.54	1.75, 1.75	2.00, 2.27		NH ^{δ21} 7.11; NH ^{ε22} 6.70
Cys 3	8.35	5.02	3.03, 2.25			
Tyr 4	9.11	4.98	2.68, 2.54		6.82, 6.82	C ^ε H 6.74, 6.74
Asn 5	8.79	5.45	2.76, 2.33		7.20, 6.91	
Cys 6	8.32	5.49	3.70, 3.38			
Pro 7	—	4.33	2.43, 2.07	2.18, ?	4.07, 3.92	
Asn 8	7.55	5.27	2.83, 2.56		NH ^{δ21} 7.57 NH ^{δ22} 6.95	
Pro 9	—	4.37	1.51, 0.96	1.89, 1.59	3.67–3.73	
Thr 10	8.31	4.53	4.28	1.24		
Ala 11	8.38	4.10	1.41			
Asp 12	8.08	4.66	2.59, 2.38			
Cys 13	8.60	4.37	3.23, 2.81			
Lys 14	8.68	4.44	2.10, 1.63	1.38, 1.21	1.52, 1.52	C ^ε H ₂ 2.92, 2.92
Thr 15	7.22	4.32	4.32	1.33		
Ala 16	8.51	4.55	1.06			
Val 17	8.34	4.59	2.09	0.94, 0.89		
Asn 18	8.42	4.80	2.64–2.74		8.58 ^b	
Cys 19	8.79	4.57	3.28–3.37			
Ser 20	8.96	4.55	4.32, 4.15			
Ser 21	8.61	4.24	3.99–4.05			
Asp 22	8.44	4.26	2.13, 1.88			
Phe 23	8.00	4.51	3.57, 3.04	7.29, 7.29		C ^ε H 7.14, 7.14; C ^ε H 7.14 ^c
Asp 24	8.43 ^c	4.64 ^c	3.08 ^c , 2.61 ^c			
Ala 25	8.43	4.73	1.18			
Cys 26	9.15	5.55	3.15, 2.98			
Leu 27	9.04	5.30	1.33, 0.92	0.82	0.38, –0.77	
Ile 28	9.10	5.67	1.87	C ^{γ1} H ₂ 1.65 C ^{γ1} H ₂ 1.12 C ^{γ2} H ₃ 1.00	0.71	
Thr 29	9.00	5.27	3.78	0.96		
Lys 30	8.48	5.15	1.84, 1.55	1.36, 1.28	1.62, 1.62	C ^ε H ₂ 2.90, 2.90
Ala 31	9.05	5.00	1.38			
Gly 32	8.81	4.08, 3.64				
Leu 33	8.84	4.32	1.64–1.71	1.70	0.94, 0.87	
Gln 34	8.00	4.47	2.42, 1.68	2.28–2.33		N ^ε H ₂ 7.56 ^c , 6.69 ^c
Val 35	7.83	4.84	1.97	0.92, 0.82		
Tyr 36	9.08	4.93	2.69, 2.43		6.90, 6.90	C ^ε H 6.73, 6.73
Asn 37	8.84	5.66	2.81, 2.47		?, ?	
Lys 38	8.40	5.51	2.52, 1.48	1.34, 1.76	2.00 ^c , 2.00 ^c	C ^ε H ₂ 3.10 ^c , 3.10 ^c
Cys 39	7.94	4.79	2.20, 2.20			
Trp 40	9.70	5.18	2.91, 2.72		C ^{δ1} H 7.50	N ^{ε1} H 8.78; C ^{ε3} H 8.19; C ^{ε2} H 7.11; C ^{ε3} H 7.17; C ^{ε2} H 6.94
Lys 41	6.85	5.18	1.51, 1.08	1.24–1.31	1.27, 1.27	C ^ε H ₂ 2.51, 2.51
Phe 42	11.05	4.37	3.14, 3.05		7.26, 7.26	C ^ε H 7.37, 7.37; C ^ε H 7.32
Glu 43	9.06	4.04	1.93, 1.71	1.61, 1.61 ^c		
His 44	7.76	5.30	3.26, 2.62		C ^{δ2} H 6.61	C ^{ε1} H 7.36 ^d
Cys 45	7.84	4.48	3.48, 3.09			
Asn 46	7.63	5.01	2.95, 2.74		NH ^{δ21} 8.59 NH ^{δ22} 7.19	
Phe 47	?	3.77	3.44, 3.14		7.35, 7.35	C ^ε H 7.59, 7.59; C ^ε H 7.59 ^c
Asn 48	8.85	4.27	2.88, 2.88		NH ^{δ21} 7.64 NH ^{δ22} 6.91	
Asp 49	8.68	4.46	2.68, 2.31			
Val 50	7.97	3.76	1.46	0.63, 0.52		
Thr 51	7.77	3.77	4.03	0.65		
Thr 52	7.77	3.98	4.18	1.26		
Arg 53	7.69	4.15	1.97–2.03	1.82, 1.62	2.82, 2.70	
Leu 54	7.87	4.28	1.27, 0.91	1.46	0.37, –0.05	
Arg 55	7.56	3.80	2.08, 1.91	1.55, 1.55	3.23, 3.23	N ^ε H 7.12

(continued)

Table 1. Continued

Residue	NH	α	β	γ	δ	Others
Glu 56	7.32	4.76	1.87, 1.13	2.23, 1.90 ^c		
Asn 57	8.50	4.71	2.79–2.88		NH ^{δ21} 7.57 ^c NH ^{δ22} 6.95 ^c	
Glu 58	7.92	4.72	2.01, 2.09	2.29, 2.29 ^c		
Leu 59	7.94	5.24	2.06, 1.49	1.49	0.76, 0.51	
Thr 60	8.41	4.77	4.17	1.32		
Tyr 61	8.63	5.60	2.86, 2.74		6.73, 6.73	C ^{ϵ} H 6.64, 6.64
Tyr 62	8.76	4.53	2.89, 2.61		6.85, 6.85	C ^{ϵ} H 6.76, 6.76
Cys 63	8.20	5.84	3.56, 2.99			
Cys 64	8.92	4.93	3.46, 3.32			
Lys 65	9.00	4.80	2.16, 1.67	1.58, 1.28	1.44, 1.44	C ^{ϵ} H ₂ 2.61, 2.48
Lys 66	7.31	4.68	1.83, 1.69	1.50–1.55	1.79, 1.79	C ^{ϵ} H ₂ 3.09, 3.09
Asp 67	8.25	4.95	2.50, 2.39			
Leu 68	9.35	3.35	1.95, 1.46	1.15	0.72, –0.07	
Cys 69	7.78	4.91	3.72, 3.46			
Asn 70	9.12	4.65	2.78, 2.30		?, ?	
Phe 71	6.91	4.80	3.35, 2.97		7.47, 7.47	C ^{ϵ} H, C ^{ϵ} H 7.38–7.41
Asn 72	8.23	4.18	2.72, 2.72		7.35, 7.38 ^c	
Glu 73	9.09	4.06	2.11–2.17	2.50, 2.33		
Gln 74	7.66	4.18	1.40–1.51	2.10, 2.10 ^c		N ^{ϵ} H ₂ 7.44, 6.61
Leu 75	7.37	4.20	1.50–1.60	1.43	0.82, 0.74	
Glu 76	7.72	4.27	2.07, 1.93	2.22, 2.22		
Asn 77	8.07	4.68	2.80, 2.70		7.45 ^c , 6.83 ^c	

^a Chemical shifts are accurate to ± 0.01 ppm except for values in italics, which are accurate to ± 0.03 ppm (the latter indicates cases in which the chemical shift had to be measured in the F_1 dimension where resolution is necessarily lower). For some CH₂ resonances, a chemical shift range is specified; this refers to the whole range covered by a strongly coupled pattern due to both signals. For the side chain amide protons of Asn and Gln, the designation NH ^{δ 21} (Asn) or NH ^{ϵ 21} (Gln) indicates the proton that is *cis* to C ^{β} (Asn) or C ^{γ} (Gln), whereas NH ^{δ 22} or NH ^{ϵ 22} indicates the proton that is *trans*.

^b Only one N ^{δ} H resonance is observed due to the presence of the *N*-glycan.

^c Tentative assignment.

^d Resonance observed at 27 °C.

of helical turns. However, due to severe overlap of α H and NH chemical shifts in this region, many potential connectivities of these types could not be unambiguously assigned. Medium-range connectivities $d_{\alpha N}(i, i+2)$ and $d_{NN}(i, i+2)$ were also observed in several positions outside the helical region and presumably indicate loop or turn structures. No attempt was made to classify tight turns at this stage.

The folding topology of the β -strands (Fig. 3) was determined primarily from strong interstrand $d_{\alpha\alpha}$ and $d_{\alpha N}$ NOE connectivities. The alignment of the strands was clearly established by the $d_{\alpha\alpha}$ connectivities that were observed between the following pairs of residues: 2 and 18, 4 and 16, 27 and 63, 29 and 61, 30 and 35, and 31 and 59. Corroborative evidence for this topology came from the presence of slowly exchanging backbone amide protons in suitable positions to form interstrand, β -sheet-type hydrogen bonds between residues 3 and 17, 25 and 40, 26 and 64, 27 and 38, and 28 and 62. It should be emphasized that this pairing of hydrogen bond donors and acceptors is speculative; definite identification of the hydrogen bonding network awaits determination of the tertiary structure at high resolution. Preliminary calculations of the tertiary

structure were carried out using only those NOE enhancements that could be assigned on grounds of chemical shift alone; note that no disulfide bonds were included in these calculations, since we hope to determine the disulfide bonding pattern from the calculated structures at a later stage. The initial structures support the folding topology just described.

Carbohydrate assignments

A preliminary report has been published (Noda et al., 1991) describing the results of chemical analysis of CD59 purified from urine and presenting a partial description of the composition and structure of the two carbohydrate groups. Figure 4 shows model structures for the two carbohydrate groups of CD59 based upon this report and upon analogy with structures in the literature. We have obtained NMR assignments consistent with these structures as described below and listed in Table 2. This process was greatly facilitated by the well-dispersed signals from amide protons in the carbohydrate groups (Fig. 5), as described by Hard et al. (1991).

Fig. 3. Folding topology of CD59 in solution. The symbols β_1 – β_5 denote extended strands for which approximate residue numbers are given; the arrangement of these into two- and three-stranded sheets does not necessarily imply the presence of regular β -sheet hydrogen bonding between the strands. The region labeled “ α ”, corresponding to residues 41–55, has helical character but does not appear to be a single regular helix (see text). The C-terminal 12 residues appear to have no regular secondary structure and are not shown.

Firstly, a number of spin systems corresponding to aldohexose amino sugars were identified by connectivities to the amide NH resonance in DQF-COSY, TOCSY, and NOESY spectra. It was possible to assign one of these sets of signals specifically to the residue of the *N*-glycan that is directly linked to the protein, by observing the unique CONH.C¹H.C²H.CONH partial spin system of the *N*-glycosidic linkage and NOE connectivities to the protein side chain protons of Asn 18. None of the other spin systems of this type were assigned to specific positions in the carbohydrate chains; these signals could be from any of the remaining glcNAc residues in the *N*-glycan or from the galNAc residue in the GPI anchor (Fig. 4). Secondly, we assigned signals to the ethanolamine group that links the GPI anchor to the C-terminus of the protein, on the basis of its unique CONH.CH₂.CH₂.O spin system. NOE connectivities from the ethanolamine amide NH proton to the H ^{α} and H ^{β} protons of Asn 77 help to confirm that Asn 77 is the C-terminal residue of the urine-derived protein. Thirdly, we have assigned signals to the C⁵H.C⁶H₃ portion of the fucose residue. Although this spin system is not distinguishable from alanine or threonine residues on the basis of chemical shifts and spin-coupling topology, the signal shows no sequential NOE connectivities to indicate that it is part of an amino acid and is so intense that it is unlikely to result from a degraded or minor form of the protein. All other major crosspeaks connecting methyl groups to methine protons in this region of the spectrum are accounted for by sequence-specific resonance assignments for the protein portion of CD59.

In addition, we observed signals that we have assigned to sialic acid residues, a species that is not mentioned in

Fig. 4. Model structures for the carbohydrate groups of CD59, based upon the chemical analysis carried out by Noda et al. (1991) and upon analogy with other structures in the literature. **A:** *N*-glycan found in a variety of glycoproteins (Kornfeld & Kornfeld, 1985); the absence of the GlcNAc residue marked “ $\pm$ ” is a common variation. The Sia-Gal glycosidic linkages also vary among the reported structures of this type and are either α 2,3 or α 2,6. Noda et al. (1991) stated that the *N*-glycan of CD59 purified from urine is of the complex type, containing two residues of galactose and one of fucose, and has a typical biantennary structure; the structure shown here fits this description. **B:** GPI anchor of Thy-1 antigen (Homans et al., 1988), omitting the phosphate group attached to the inositol (see text). The GPI anchor of CD59 has the same oligosaccharide core and also contains two ethanolamine groups (Noda et al., 1991). In membrane-bound CD59 there is a phospholipid attached by its phosphate group to C¹ of the inositol moiety. There is strong indirect evidence to indicate that CD59 on erythrocytes has an additional single fatty acid attached elsewhere on the inositol ring (Ratnoff et al., 1992), but the NMR data contain no evidence for its presence in the urine-derived form of CD59. **C:** Structures of individual sugar monomers. In all cases the β anomeric configuration is shown and the ring carbons are numbered clockwise from the indicated starting point. In GlcNH₂, GlcNAc, and GalNAc the NH₂ or NAc group is attached to C² in place of the —OH group.

Table 2. Carbohydrate assignments at 42 °C and pH 5.8^a

Residue	Chemical shifts					
	C ¹ H	C ² H	CONH	CH ₃	Others	
Hex 1 ^b	4.93	3.87	7.81	1.91	3.71	
Hex 2 ^c	4.52	3.95	8.40	2.08	3.82	
Hex 3 ^c		4.07	8.32	2.09	3.93, 4.20	
Hex 4 ^c	4.57	4.06	8.51	2.10 ^d	3.94, 4.20	
Hex 5 ^c	4.57	4.05	8.48	2.07 ^d		
Hex 6 ^e	4.60	3.75	8.30	2.10		
Hex 7 ^e	4.60	3.80	8.10	2.07	3.60	
	CONH·CH ₂ ·CH ₂ ·O		CONH·CH ₂ ·CH ₂ ·O		CONH·CH ₂ ·CH ₂ ·O	
Eth 1 ^f	8.02		3.57, 3.40		3.95, 3.95	
Eth 2 ^f	7.96		3.50, 3.37		3.98, 3.98	
	C ⁵ H		C ⁶ H ₃		Other	
Fuc 1	4.09		1.20		3.80 ^d	
	C ³ Heq	C ³ Hax	C ⁴ H	C ⁵ H	CONH	CH ₃
Sia 1	2.78	1.81	3.72	3.86	7.84	2.06
Sia 2	2.69	1.72	3.69	3.82	7.91	2.06

^a Chemical shifts are accurate to ±0.01 ppm except for values in italics, which are accurate to ±0.03 ppm (the latter indicates cases in which the chemical shift had to be measured in the F₁ dimension where resolution is necessarily lower).

^b *N*-acetylated aldohexose amino sugar, directly linked to N18.

^c Other *N*-acetylated aldohexose amino sugar.

^d Tentative assignment.

^e Several overlapped cases of ^c.

^f Ethanolamine group, directly linked to C-terminus of protein.

the preliminary report of Noda et al. (1991) but that has been observed in urine-derived CD59 (Ninomiya et al., 1992). These assignments were based upon the observation of the C³H₂·C⁴H·C⁵H spin system in DQF-COSY and TOCSY spectra and upon appropriate connectivities to an amide proton signal in DQF-COSY, TOCSY, and NOESY spectra. The chemical shifts of the signals are all within 0.2 ppm of the expected values for sialic acid in an α-glycosidic linkage to Gal (Vliegthart et al., 1981). The extremely sharp lines observed for all the sialic acid signals is consistent with the presumed chain-terminal position of these residues and their consequent expected high mobility. We could not locate the C⁶H resonance, despite its presumably large coupling to C⁵H (³J_{5,6} is approximately 10 Hz according to Vliegthart et al. [1981]); this might be caused by signal overlap of C⁶H with C⁴H.

No signals attributable to the lipid fraction of the GPI anchor were observed in the NMR spectra, presumably indicating that this portion of the molecule is cleaved off before or during excretion. This is consistent with previous studies of CD59 isolated from urine (Davies et al., 1989; Tomita et al., 1991).

In general, the signals from carbohydrate groups have small linewidths, and the intensities of NOESY crosspeaks

between the protons of the carbohydrate groups are low. We observed no NOE connectivities between the protein and the carbohydrate groups with the exception of the residues Eth 1 and Hex 1, which are directly adjacent to the protein. These data suggest that the carbohydrate groups are highly mobile and that their conformation is largely independent of the protein structure.

Several of the carbohydrate spin systems showed clear evidence of structural inhomogeneity. For example, for each signal of the major ethanolamine spin system, we observed a minor companion signal with a slightly different chemical shift and a lower intensity; we refer to the major and minor sets of peaks as Eth 1 and Eth 2, respectively. A similar doubling was observed for the sialic acid and possibly for other cases among the aldohexose amino sugars. We note in this context that the GPI anchor of Thy-1 antigen was found to have a variable composition (Homans et al., 1988). The doubling of sialic acid and aldohexose amino sugar signals could conceivably result from the same residue occurring at different positions in the structure, as there is insufficient difference in the intensities of the pairs of signals to distinguish inhomogeneity from multiple occurrence. However, evidence for inhomogeneity (in either or both of the carbohydrate groups) comes from the number of aldohexose amino

Fig. 5. Carbohydrate assignments using amide NH signals of acetyl-amide groups. This figure shows sections of (A) TOCSY and (B) NOESY spectra of CD59 in H₂O, recorded at 42 °C. Mixing times were 70 ms and 250 ms, respectively, and the NOESY spectrum was acquired with jump-return excitation. Signals from the carbohydrate groups are labeled in the figure corresponding to the assignments listed in Table 2.

sugars that are observed (at least nine, based upon a count of the amide NH signals in TOCSY spectra), which is too many to correspond to the likely structures of the carbohydrate groups (four or five glcNAc residues in the *N*-glycan and one galNAc residue in the GPI anchor).

The work of Noda et al. (1991) strongly suggests that the intact GPI anchor contains a total of three phosphate groups, two as ethanolamine phosphate and one in a phospholipid, as found in other GPI anchors from eukaryotic species (Homans et al., 1988; Roberts et al., 1988a,b). The ³¹P-¹H HSQC spectrum contains two groups of signals in the ³¹P dimension (Fig. 6), and we tentatively assigned these groups to two of the three phosphorus atoms in the structure using chemical shift arguments as described in the legend to Figure 6. We are confident that signals from the third phosphorus atom (in the phosphate group attached to the inositol moiety) are not present and overlapping the two groups of signals, because these groups each have the same integrated intensity in the 1D ³¹P-NMR spectrum.

This assignment of the ³¹P-¹H HSQC spectrum has two implications. Firstly, the absence of the third phos-

phate group in our sample suggests that the phospholipid moiety has been removed from CD59 found in urine by phosphatidylinositol-specific phospholipase-D, since this enzyme (unlike other candidate phospholipases such as phosphatidylinositol-specific phospholipase-C) cleaves off the phosphate part of the phospholipid along with the di-(acyl/alkyl)-glycerol moiety. Secondly, the number of distinct ³¹P signals within each group indicates that at least four structurally different versions of the GPI anchor are present in our sample. The nature of these structural differences remains unknown due to the difficulty of obtaining further carbohydrate assignments in the face of extreme signal overlap and a limited knowledge of the covalent structure.

Discussion

Structure and function of CD59

As described, the secondary structure of CD59 consists primarily of extended strands arranged in an antiparallel manner. We have also identified a region of helical

Fig. 6. Two regions from the ^{31}P - ^1H HSQC spectrum optimized for observation of long-range couplings and recorded at 42 °C. The remaining parts of the spectrum (not shown here) contained no signals. The ^{31}P dimension was calibrated indirectly relative to P_i (≈ 0.0 ppm) in a ^{31}P detected 1D spectrum, and only positive contour levels are shown. The following assignments were made using the assumption that the GPI anchor of CD59 has the structure shown in Figure 4; note, however, that the GPI anchor is inhomogeneous (see text). The group of signals in the upper panel shows possible correlations to C^2H of a mannose residue at 4.5 ppm and to the $\text{CONH}_2\text{CH}_2\text{CH}_2\text{O}$ resonance of the peripheral ethanolamine moiety at 4.2 ppm, and was therefore assigned to the corresponding phosphate group. The group of signals in the lower panel shows possible correlations to the $\text{CONHCH}_2\text{CH}_2\text{O}$ resonance of the assigned ethanolamine moiety at 4.0 ppm and to C^6H of a mannose residue at 4.1 ppm, and was therefore assigned to the phosphate group nearest to the protein in the covalent structure. For further discussion of the assignment and interpretation of this spectrum, see text.

character, residues 41–55, but this is not well defined by the NOE connectivities that have been assigned thus far. It is not clear whether this region forms a single helix or a less regular structure, but we note that the apparent lack of a detectable backbone NH signal for residue 47 suggests that this proton is unlikely to be forming a stable hydrogen bond as it would do in a regular α -helix including residue 43. The C-terminal region of the sequence (residues 66–77) cannot be described as having any regular secondary structure at this stage, and so the nature of the hydrogen bonds formed by the slowly exchanging backbone amide protons of residues 68–70 remains unclear.

We expect the structural information presented here to illuminate functional studies of CD59, such as experiments designed to define the C8/C9 binding site. It has been shown that a fragment of CD59 consisting of three peptides, residues 1–14, 15–31, and 39–41, linked together by disulfide bonds, retains the ability to bind the C5b-8 complex (V. Ferriani, pers. comm.). In the light of the structure, we suggest the possibility of the C8 binding site being located in the loop between residues 7 and 14, on

the basis that it is accessible to the ligand (C8) and might retain its native conformation in the fragment due to preservation of the two-stranded β -sheet and/or the presence of the Cys 6–Cys 13 disulfide bond. In this context, it is also of interest to compare the sequence of CD59 with that of its homologues from other species. Partial N-terminal sequences have been published for rat and sheep homologues of CD59 (Hughes et al., 1992; van den Berg et al., 1993), but this comparison is made complicated by the fact that human CD59 specifically restricts lysis by human and primate complement (Rollins et al., 1991), whereas the rat and sheep homologues are not specific and inhibit complement from a variety of species, including humans (Hughes et al., 1992; van den Berg et al., 1993). In fact, the conserved residues are spread evenly throughout the three sequences, with no large region of either similarity or difference, and thus a satisfactory interpretation of this sequence comparison awaits more detailed knowledge of the structure of CD59 and of the C8/C9 binding site.

Comparison of CD59 with other proteins

The folding topology of CD59 is very similar to the homologous snake venom neurotoxins and to the other members of the toxin-agglutinin family, the snake venom cardiotoxins and WGA. All of the known structures of these proteins (see introduction for references) contain five extended strands with the connecting topology shown in Figure 3. Furthermore, the division of the β -sheet into separate two- and three-stranded units is clearly seen in the neurotoxin and cardiotoxin structures (it is difficult to make this distinction in WGA as the secondary structure is less regular). A difference arises in the presence of the helical region between the fourth and fifth extended strands. This feature is not observed in any of the neurotoxin or cardiotoxin structures but is seen in WGA.

This structural similarity between CD59 and the snake venom neurotoxins lends support to the idea of an evolutionary relationship between the two, although we cannot rule out the possibility of convergent evolution to the same stable fold. There is no obvious functional similarity between CD59 and the neurotoxins, which bind to nicotinic acetylcholine receptor molecules on postsynaptic membranes as competitive inhibitors of acetylcholine (Dufton & Hider, 1983). However, it is conceivable that a common molecular mechanism is responsible both for this binding and for the association of CD59 with C8/C9. The available evidence suggests that this is unlikely. In the case of α -bungarotoxin, chemical labeling experiments and structure and sequence comparisons with other neurotoxins indicate that the whole of one face of this molecule is involved in binding to nicotinic AChR, with the so-called “loop 2” (approximately equivalent to residues 24–40 in CD59) making important specific contacts (Love & Stroud, 1986). Significantly, a large part of this loop

(residues 32–38) is missing from the peptide fragment of CD59 that retained the ability to bind C8 (see above). If, as seems likely, no common mechanism exists to explain the structural similarity between CD59 and the neurotoxins, then we would conclude that this is an example of the same protein fold acting as a stable scaffold on which are placed different functional residues.

Very little is known about the mechanism of action of the snake venom cardiotoxins (Dufton & Hider, 1988), so the significance, if any, of their structural resemblance to CD59 remains unknown. In the light of the structural similarity of CD59 and WGA, it is intriguing that WGA causes proliferation of human T cells in vitro (Brown et al., 1976), as does CD59 (Deckert et al., 1992). However, it seems unlikely that this superficial functional similarity between CD59 and WGA is based upon the same fundamental molecular mechanism. WGA seems to have its effect by binding to GlcNAc on the cell surface (Brown et al., 1976), whereas CD59 has been shown to bind to the T cell surface protein CD2 (Hahn et al., 1992) and this is a plausible mechanism for its own effects on T cell proliferation.

Impact of the carbohydrate NMR signals on analysis of the protein structure

Previous studies of intact glycoproteins by NMR have either been limited to glycosylated peptides (e.g., a 22-residue peptide from serum IgM [Wormald et al., 1991]) or, in the case of larger proteins, have involved only partial assignment of the protein spectrum (e.g., the 42-kDa protein peroxidase C from horseradish [Veitch et al., 1992]) or have concentrated on the carbohydrate component (e.g., the 35-kDa protein phosphatase [Brockbank & Vogel, 1990]). To the best of our knowledge, the present study represents the first case in which almost complete assignments and detailed structural information have been obtained for the protein component of an intact glycoprotein of this size. With this in mind, it is interesting to consider the effect of the presence of the carbohydrate groups upon the NMR analysis of the protein component.

In the event, CD59 isolated from urine has proved to be an excellent subject for NMR study. The protein resonances are well dispersed and give high intensity crosspeaks in NOESY spectra. Of great importance in obtaining the protein assignments was the fact that the signals from the carbohydrate groups are limited to a few regions of the spectra. Crosspeaks involving the amide and methyl groups of the carbohydrate proved fairly simple to distinguish from protein signals, and in fact provided nearly all of our carbohydrate assignments. The few difficulties that these signals presented result from overlap with a small number of protein signals and from t_1 noise generated by the particularly sharp CH_3 resonances of the *N*-acetylated sugars. While the non-anomeric ring proton

resonances of sugar residues dominate the square region at 3–4 ppm in both frequency dimensions, this region is crucial for only a few protein signals. Furthermore, the general absence of significant numbers of NOE enhancements between carbohydrate and protein and between different spin systems within the carbohydrate means that the observation of NOE connectivities to other spin systems is a useful criterion for distinguishing amino acid signals from those of carbohydrate residues in ambiguous cases. In summary, the features of the carbohydrate spectra that were crucial to our success in analyzing the protein signals were, firstly, low chemical shift dispersion and, secondly, a scarcity of interresidue NOE enhancements.

Thus, in the case of CD59, the presence of the carbohydrate groups has been shown to have little adverse effect on the NMR study of the protein component and, as a result, we expect to obtain a high-resolution structure of the protein component based upon the NMR data described in this paper. As regards the general applicability of this approach, we would expect it to be successful in other cases of moderately glycosylated proteins in which the carbohydrate groups are flexible and independent of the protein, thus giving carbohydrate spectra with the same crucial features as those of CD59. However, we would expect greater difficulties for glycoproteins in which the carbohydrate interacts significantly with the protein, as has been observed in a human Fc fragment (Deisenhofer, 1981), or for significantly more heavily glycosylated proteins.

Materials and methods

Preparation of NMR samples

CD59 was purified from human urine as described (Davies et al., 1989) and either lyophilized immediately (D_2O sample) or dialyzed exhaustively against H_2O and concentrated using an Amicon ultrafiltration cell with PM10 membrane (molecular weight cutoff 10,000) (H_2O sample). Both samples were then dialyzed as above, lyophilized, and taken up in a buffer solution (5 mM KH_2PO_4 , 5 mM NaCl, 1 mM NaN_3 , and 67 μM TSP in either pure D_2O [D_2O sample] or 15% D_2O /85% H_2O [H_2O sample]). The resulting solutions were re-lyophilized and taken up in the appropriate solvent to a final protein concentration of 2–3 mM. The final concentrations of other solutes in the samples were as follows: 25 mM KH_2PO_4 , 25 mM NaCl, 5 mM NaN_3 , and 340 μM TSP (D_2O sample); 17 mM KH_2PO_4 , 17 mM NaCl, 3 mM NaN_3 , and 220 μM TSP (H_2O sample). The pH of both samples was measured at 5.8 (not corrected for deuterium isotope effects).

NMR measurements

NMR measurements were carried out at temperatures of 27 °C and 42 °C using a Bruker AMX500 spectrometer.

Two-dimensional spectra were acquired using TPPI in t_1 to achieve F_1 quadrature detection (Bodenhausen et al., 1980; Marion & Wüthrich, 1983) and using true quadrature in F_2 with simultaneous acquisition of complex data points in t_2 . With the exceptions noted below, the spectral width in both dimensions was 8 kHz, the filter width was 12 kHz, and the data size was 2,048 (complex) $\times$ 512 (real) in the time domain and 4,096 (complex) $\times$ 512 (complex) in the frequency domain. For 2Q spectra in H_2O , the spectral width in F_1 was 16 kHz. For 2Q and PE-COSY spectra in D_2O , the spectral widths in F_1 were reduced to 9 kHz and 4 kHz, respectively, and TPPI was not employed, in order to achieve greater resolution in F_1 . For PE-COSY spectra the data sizes were 4,096 (complex) $\times$ 700 (real) or 512 (real) in the time domain and 8,192 (complex) $\times$ 1,024 (complex) in the frequency domain. In the ^{31}P (F_1) dimension of the ^{31}P - 1H HSQC experiment, the spectral width was 750 Hz. In the ^{13}C (F_1) dimension of the ^{13}C - 1H HMQC experiment, the spectral width was 12.5 kHz, allowing the aromatic and anomeric resonances to fold.

Except in the case of jump-return NOESY experiments, water suppression was achieved using presaturation during the relaxation delay (and during the mixing time in the case of NOESY), with an irradiation field of the same frequency as and phase coherent with the hard pulses (Zuiderweg et al., 1986). Spectra were acquired using 64 transients per t_1 increment (homonuclear experiments; one TOCSY experiment was acquired with 224 transients per t_1 increment), 384 transients per t_1 increment (^{13}C - 1H HMQC), or 192 transients per t_1 increment (^{31}P - 1H HSQC).

NOESY spectra were acquired using mixing times of 20, 36, 65, 117, 200, and 250 ms. Jump-return NOESY spectra (Plateau & Guéron, 1982) were recorded using the same protocol as NOESY, except that presaturation was not applied and the final 90° observe pulse of NOESY was replaced by two 45° pulses of mutually opposite phase, separated by a delay of 100 μ s. For NOESY experiments, CYCLOPS was only carried out on the last pulse (or on both 45° pulses in the case of jump-return NOESY). ROESY spectra were acquired with a mixing time of 100 ms and a spin-lock field strength of 5–6 kHz. In order to avoid offset effects in ROESY experiments, a 90° pulse was applied immediately before and after the spin-locking period (Griesinger & Ernst, 1987).

TOCSY experiments employed the DIPSI-2 mixing sequence (Shaka et al., 1988), with two z -filters flanking the spin-locking period to achieve minimal phase distortion (Rance, 1987). Mixing times of approximately 35 ms (D_2O sample) and 70 ms (H_2O sample) were used with spin-lock field strengths of 5–6 kHz. The two z -filters employed delays of 0.5, 0.7, 1.1, and 1.3 ms (first z -filter) and 1.5, 1.7, 2.1, and 2.3 ms (second z -filter). In order to limit the total length of the TOCSY experiments, variation of the delays in the two z -filters was carried out in

parallel and was also combined with the CYCLOPS phase cycle on all the hard pulses and the spin-lock.

For RELAY spectra the mixing time was 30 ms, and for 2Q spectra the preparation time for evolution of double-quantum coherence was 30 ms. A ^{13}C - 1H correlation spectrum was obtained at natural abundance using the HMQC pulse sequence (Müller, 1979) with double-difference phase cycle (Cavanagh & Keeler, 1988). The fixed delay (Δ in the notation of Cavanagh & Keeler) was set to 3.3 ms for optimal evolution of antiphase coherence under carbon-proton one-bond couplings. A ^{31}P - 1H correlation spectrum was obtained using the HSQC pulse-sequence (Bodenhausen & Ruben, 1980; Bax et al., 1990). A fixed delay (Δ in the notation of Bax et al.) of 30 ms was used in this experiment, to give optimal magnetization transfer via proton-phosphorus three-bond couplings. All heteronuclear correlation spectra were decoupled in both frequency dimensions.

Processing of NMR data

Data processing was carried out using an ASPECT X-32 computer running the program UXNMR. Time-domain data were multiplied in both dimensions by phase-shifted sine-bell or squared sine-bell functions. F_2 baseline errors were corrected by subtracting a polynomial of order 1, 2, or 3. Spectra were referenced relative to internal TSP.

Slowly exchanging amide protons

Slowly exchanging amide protons were identified on the basis of crosspeaks in a NOESY spectrum obtained within a week after exchange into D_2O at pH 5.8.

Note added in proof

Since this manuscript was submitted, a paper has been published describing the NMR-derived 3D structure of the protein component of another glycosylated protein, human T lymphocyte glycoprotein CD2 (Withka et al., 1993). This is similar to CD59 in that both proteins are moderately glycosylated, contain predominantly β -sheet, and give good quality homonuclear and natural abundance ^{13}C - 1H correlation spectra. A manuscript describing the 3D structure of CD59 is currently in preparation.

Acknowledgments

We are grateful to Christine Willers and Rodney Oldroyd for technical assistance in purification of the protein and preparation of the NMR samples and to Dr. Alexandra Davies for allowing us to read her CD59 review article prior to its publication. During the course of this work, C.M. Fletcher was supported by an MRC research studentship.

References

- Bax, A., Ikura, M., Kay, L.E., Torchia, D.A., & Tschudin, R. (1990). Comparison of different modes of two-dimensional reverse-correlation NMR for the study of proteins. *J. Magn. Reson.* 86, 304–318.
- Betz, C., Lange, G., Pal, G.P., Wilson, K.S., Maelicke, A., & Saenger, W. (1991). The refined crystal structure of α -cobratoxin from *Naja naja siamensis* at 2.4 Å resolution. *J. Biol. Chem.* 266, 21530–21536.
- Bodenhausen, G. & Ruben, D.J. (1980). Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* 69, 185–189.
- Bodenhausen, G., Vold, R.I., & Vold, R.R. (1980). Multiple quantum spin-echo spectroscopy. *J. Magn. Reson.* 37, 93–106.
- Bourne, P.E., Sato, A., Corfield, P.W.R., Rosen, L.S., Birken, S., & Low, B.W. (1985). Erabutoxin b: Initial protein refinement and sequence analysis at 0.140 nm resolution. *Eur. J. Biochem.* 153, 521–527.
- Brockbank, R.L. & Vogel, H.J. (1990). Structure of the oligosaccharide of hen phosphatidyl as determined by two-dimensional ^1H NMR of the intact glycoprotein. *Biochemistry* 29, 5574–5583.
- Brown, J.M., Leon, M.A., & Lightboy, J.J. (1976). Isolation of a human lymphocyte mitogen from wheat germ with N-acetyl-D-glucosamine specificity. *J. Immunol.* 117, 1976–1980.
- Cavanagh, J. & Keeler, J. (1988). Improvement of carbon-13 satellite spectra by double difference spectroscopy. *J. Magn. Reson.* 77, 356–362.
- Davies, A., Simmons, D.L., Hale, G., Harrison, R.A., Tighe, H., Lachmann, P.J., & Waldmann, H. (1989). CD59, an LY-6-like protein expressed in human lymphoid cells, regulates the action of the complement membrane attack complex on homologous cells. *J. Exp. Med.* 170, 637–654.
- Deckert, M., Kubar, J., & Bernard, A. (1992). CD58 and CD59 molecules exhibit potentializing effects in T cell adhesion and activation. *J. Immunol.* 148, 672–677.
- Deisenhofer, J. (1981). Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9 and 2.8 Å resolution. *Biochemistry* 20, 2361–2370.
- Drenth, J., Low, B.W., Richardson, J.S., & Wright, C.S. (1980). The toxin-agglutinin fold: A new group of small protein structures organized around a four-disulphide core. *J. Biol. Chem.* 255, 2652–2655.
- Dufton, H.J. & Hider, R.C. (1983). Conformational properties of the neurotoxins and cytotoxins isolated from elapid snake venoms. *CRC Crit. Rev. Biochem.* 14, 113–171.
- Dufton, H.J. & Hider, R.C. (1988). Structure and pharmacology of elapid cytotoxins. *Pharmacol. Ther.* 36, 1–40.
- Griesinger, C. & Ernst, R.R. (1987). Frequency offset effects and their elimination in NMR rotating-frame cross-relaxation spectroscopy. *J. Magn. Reson.* 75, 261–271.
- Groux, H., Huet, S., Aubrit, F., Tran, H.C., Boumsell, L., & Bernard, A. (1989). A 19 kDa human erythrocyte molecule H19 is involved in rosettes, present on nucleated cells and required for T cell activation. Comparison of the roles of H19 and LFA-3 molecules in T cell activation. *J. Immunol.* 142, 3013–3020.
- Hahn, W.C., Menu, E., Bothwell, A.L.M., Sims, P.J., & Bierer, B.E. (1992). Overlapping but non-identical binding sites on CD2 for CD58 and a second ligand CD59. *Science* 256, 1805–1807.
- Hamilton, K.K., Zhao, J., Rollins, S., Stewart, B.H., & Sims, P.J. (1990). Regulatory control of the terminal complement proteins at the surface of human endothelial cells: Neutralization of a C5b-9 inhibitor by antibody to CD59. *Blood* 76, 2572–2577.
- Hard, K., Spronk, B.A., Hokke, C.H., Kamerling, J.P., & Vliegthart, J.F. (1991). The potency of amide protons for assignments of NMR spectra of carbohydrate chains of glycoproteins, recorded in $^2\text{H}_2\text{O}$ solutions. *FEBS Lett.* 287, 108–112.
- Harrison, R.A. (1992). Motif pattern searching reveals similarities between CD59, a complement regulatory protein, and snake venom neurotoxins. *Program of the 6th Annual Symposium of The Protein Society, San Diego, July 1992*, p. 90 [Abstr.].
- Holguin, M.H., Fredrick, L.R., Bershaw, N.J., Wilcox, A., & Parker, C.J. (1989). Isolation and characterization of a membrane protein from normal human erythrocyte that inhibits reactive lysis of the erythrocytes of paroxysmal nocturnal haemoglobinuria. *J. Clin. Invest.* 84, 7–17.
- Homans, S.W., Ferguson, M.A., Dwek, R.A., Rademacher, T.W., Anand, R., & Williams, A.F. (1988). Complete structure of the glycosylphosphatidylinositol membrane anchor of rat brain Thy-1 glycoprotein. *Nature* 333, 269–272.
- Hughes, T.R., Piddlesden, S.J., Williams, J.D., Harrison, R.A., & Morgan, B.P. (1992). Isolation and characterization of a membrane protein from rat erythrocytes which inhibits lysis by the membrane attack complex of rat complement. *Biochem. J.* 284, 169–176.
- Kornfeld, R. & Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 54, 631–664.
- Korty, P.E., Brando, C., & Shevach, E.M. (1991). CD59 functions as a signal-transducing molecule for human T cell activation. *J. Immunol.* 146, 4092–4098.
- Labhardt, A.M., Hunziker-kwik, E., & Wüthrich, K. (1988). Secondary structure determination for α -neurotoxin from *Dendroaspis polyplepis polyplepis* based on sequence-specific ^1H -nuclear magnetic resonance assignments. *Eur. J. Biochem.* 177, 295–305.
- le Du, M.H., Marchot, P., Bougis, P.E., & Fontecilla-Camps, J.C. (1992). 1.9 Å resolution structure of fasciculin 1, an anti-acetylcholinesterase toxin from green mamba snake venom. *J. Biol. Chem.* 267, 22122–22130.
- Le Goas, R., LaPlante, S.R., Mikou, A., Delsuc, M., Guittet, E., Robin, M., Charpentier, I., & Lallemand, J. (1992). α -Cobratoxin: Proton NMR assignments and solution structure. *Biochemistry* 31, 4867–4875.
- Love, R.A. & Stroud, R.M. (1986). The crystal structure of α -bungarotoxin at 2.5 Å resolution: Relation to solution structure and binding to acetylcholine receptor. *Protein Eng.* 1, 37–46.
- Marion, D. & Wüthrich, K. (1983). Application of phase-sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^1H - ^1H spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* 113, 967–974.
- Meri, S., Morgan, B.P., Davies, A., Daniels, R.H., Olavesen, M.G., Waldmann, H., & Lachmann, P.J. (1990). Human protectin (CD59), an 18,000–20,000 MW complement lysis restricting factor, inhibits C5b-8 catalysed insertion of C9 into lipid bilayers. *Immunology* 71, 1–9.
- Müller, L. (1979). Sensitivity-enhanced detection of weak nuclei using heteronuclear multiple quantum coherence. *J. Am. Chem. Soc.* 101, 4481–4484.
- Ninomiya, H., Stewart, B.H., Rollins, S.A., Zhao, J., Bothwell, A.L.M., & Sims, P.J. (1992). Contribution of the N-linked carbohydrate of erythrocyte antigen CD59 to its complement-inhibitory activity. *J. Biol. Chem.* 267, 8404–8410.
- Noda, K., Nakano, Y., Tobe, T., Choi-Miura, N., Tomita, M., Endo, T., & Kobata, A. (1991). Characterization of a glycosylphosphatidylinositol anchor and an N-glycosidic oligosaccharide chain of MACIF (CD59). *Complement Inflammation* 8, 199 [Abstr.].
- Okada, N., Harada, R., Fujita, T., & Okada, H. (1989a). A novel membrane glycoprotein capable of inhibiting membrane attack by homologous complement. *Int. Immunol.* 1, 205–208.
- Okada, H., Nagami, Y., Takanashi, K., Okada, N., Hideshima, T., Takizawa, H., & Kondo, J. (1989b). 20 kDa homologous restriction factor of complement resembles T cell activating protein. *Biochem. Biophys. Res. Commun.* 162, 1553–1559.
- Oswald, R.E., Sutcliffe, M.J., Bamberger, M., Loring, R.H., Braswell, E., & Dodson, C.M. (1991). Solution structure of neuronal bungarotoxin determined by two-dimensional NMR spectroscopy: Sequence-specific assignments, secondary structure and dimer formation. *Biochemistry* 30, 4901–4909.
- Plateau, P. & Guéron, M. (1982). Exchangeable-proton NMR without baseline distortion using new strong-pulse sequences. *J. Am. Chem. Soc.* 104, 7310–7311.
- Rance, M. (1987). Improved techniques for homonuclear rotating-frame and isotropic mixing experiments. *J. Magn. Reson.* 74, 557–564.
- Ratnoff, W.D., Knez, J.J., Prince, G.M., Okada, H., Lachmann, P.J., & Medif, M.E. (1992). Structural properties of the glycoplasmanin-inositol anchor phospholipid of the complement membrane attack complex inhibitor CD59. *Clin. Exp. Immunol.* 87, 415–421.
- Rees, B., Bilwes, A., Samama, J.P., & Moras, D. (1990). Cardiotoxin VII_4 from *Naja mossambica mossambica*: The refined crystal structure. *J. Mol. Biol.* 214, 281–297.

- Roberts, W.L., Myher, J.J., Kuksis, A., Low, M.G., & Rosenberry, T.L. (1988a). Lipid analysis of the glycoinositol phospholipid membrane anchor of human erythrocyte acetylcholinesterase. *J. Biol. Chem.* 263, 18766–18775.
- Roberts, W.L., Santikarn, S., Reinhold, V.N., & Rosenberry, T.L. (1988b). Structural characterization of the glycoinositol phospholipid membrane anchor of human erythrocyte acetylcholinesterase by fast atom bombardment mass spectrometry. *J. Biol. Chem.* 263, 18776–18784.
- Rollins, S.A. & Sims, P.J. (1990). The complement-inhibitory activity of CD59 resides in its capacity to block incorporation of C9 into membrane C5b-9. *J. Immunol.* 144, 3478–3483.
- Rollins, S.A., Zhao, J., Ninomiya, H., & Sims, P.J. (1991). Inhibition of homologous complement by CD59 is mediated by a species-selective recognition conferred through binding to C8 within C5b-8 or C9 within C5b-9. *J. Immunol.* 146, 2345–2351.
- Shaka, A.J., Lee, C.J., & Pines, A. (1988). Iterative schemes for bilinear operators: Application to spin decoupling. *J. Magn. Reson.* 77, 274–293.
- Stefanova, I., Hilgert, I., Kristofova, H., Brown, R., Low, M.G., & Horejsi, V. (1989). Characterization of a broadly-expressed human leucocyte surface antigen MEM-43 anchored in membrane through phosphatidylinositol. *Mol. Immunol.* 26, 153–161.
- Steinmetz, W.E., Bougis, P.E., Rochat, H., Redwine, D., Braun, W., & Wüthrich, K. (1988). ^1H nuclear magnetic resonance studies of the three-dimensional structure of the cardiotoxin CTXIIb from *Naja mossaambica mossaambica* in aqueous solution and comparison with the crystal structures of homologous toxins. *Eur. J. Biochem.* 172, 101–116.
- Sugita, Y., Nakano, Y., & Tomita, M. (1988). Isolation from human erythrocytes of a new membrane protein which inhibits the formation of complement transmembrane channels. *J. Biochem. (Tokyo)* 104, 633–637.
- Tomita, M., Tobe, T., Choi-Miura, N., Nakano, Y., Kusano, M., & Oda, E. (1991). Structural analysis of MACIF (CD59). *Complement Inflammation* 8, 233 [Abstr.].
- van den Berg, C.W., Harrison, R.A., & Morgan, B.P. (1993). The sheep analogue of human CD59: Purification and characterization of its complement-inhibitory activity. *Immunology* 78, 349–357.
- Veitch, N.C., Williams, R.J.P., Bray, R.C., Burke, J.F., Sanders, S.A., Thornley, R.N.F., & Smith, A.T. (1992). Structural studies by proton-NMR spectroscopy of plant horseradish peroxidase C, the wild-type recombinant protein from *Escherichia coli* and two protein variants, Phe 41 $\rightarrow$ Val and Arg 38 $\rightarrow$ Lys. *Eur. J. Biochem.* 207, 521–531.
- Vliegthart, J.F.G., Dorland, L., van Halbeek, H., & Haverkamp, J. (1981). NMR spectroscopy of sialic acids. *Cell Biol. Monogr.* 10, 127–172.
- Whitlow, M.B., Iida, K., Stefanova, I., Bernard, A., & Nussenzweig, V. (1990). H19, a surface membrane molecule involved in T cell activation, inhibits channel formation by human complement. *Cell. Immunol.* 126, 176–184.
- Withka, J.M., Wyss, D.F., Wagner, G., Arulanandam, A.R.N., Reinherz, E.L., & Recny, M.A. (1993). Structure of the glycosylated adhesion domain of human T lymphocyte glycoprotein CD2. *Structure* 1, 69–81.
- Wormald, M.R., Wooten, E.W., Bazzo, R., Edge, C.J., Feinstein, A., Rademacher, T.W., & Dwek, R.A. (1991). The conformational effects of N-glycosylation on the tailpiece from serum IgM. *FEBS Lett.* 198, 131–139.
- Wright, C.S. (1987). Refinement of the crystal structure of wheat germ agglutinin isolectin 2 at 1.8 Å resolution. *J. Mol. Biol.* 194, 501–529.
- Wüthrich, K. (1986). *NMR of Proteins & Nucleic Acids*. John Wiley, New York.
- Yu, C., Bhaskaran, R., Chuang, L., & Yang, C. (1993). Solution conformation of cobratoxin: A nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing study. *Biochemistry* 32, 2131–2136.
- Zuiderweg, E.R.P., Hallenga, K., & Olejniczak, E.T. (1986). Improvement of 2D NOE spectra of biomolecules in H_2O solution by coherent suppression of the solvent resonance. *J. Magn. Reson.* 70, 336–343.