Nelsestuen, G. L., & Broderius, M. (1977) *Biochemistry 16*, 4172-4176.
Olsen, P. H. (1991) Ph.D. Dissertation, University of New Hampshire, Durham, NH.
Owen, W. G., Esmon, C. T., & Jackson, C. M. (1974) *J. Biol. Chem. 249*, 594-605.
Sakai, T., Lund-Hansen, T., Thim, L., & Kisiel, W. (1990) *J. Biol. Chem. 265*, 1890-1894.
Seegers, W. H., McCoy, L., Kipfer, R. K., & Murano, G. (1968) *Arch. Biochem. Biophys. 128*, 194-201.
Stearns, D. J., Kurosawa, S., Sims, P. J., Esmon, N. L., & Esmon, C. T. (1988) *J. Biol. Chem. 263*, 826-832.
Stearns, D. J., Kurosawa, & Esmon, C. T. (1989) *J. Biol. Chem. 264*, 3352-3356.
Suzuki, K., Hayashi, T., Nishioka, J., Kosaka, Y., Zushi, M., Honda, G., & Yamamoto, S. (1989) *J. Biol. Chem. 264*, 4872-4876.
Teller, D. C. (1976) *Nature 260*, 729-731.
Teller, D. C., Swanson, E., & De Haen, C. (1979) *Methods Enzymol. 61*, 103-124.
Van Holde, K. E., & Weischet, W. O. (1978) *Biopolymers 17*, 1387-1403.
Vigano-D'Angelo, S., Comp, P. C., Esmon, C. T., & D'Angelo, A. (1986) *J. Clin. Invest. 77*, 416-425.
Walker, F. J., Sexton, P. W., & Esmon, C. T. (1979) *Biochim. Biophys. Acta 571*, 333-342.
Weast, R. C., Ed. (1977) *Handbook of Chemistry and Physics*, 58th ed., Chemical Rubber Publishing Co., Cleveland, OH.
Williams, R. C., Jr. (1978) *Methods Enzymol. 48*, 185-191.
Winnard, P. W., Esmon, C. T., & Laue, T. M. (1989) *Arch. Biochem. Biophys. 269*, 339-344.
Yphantis, D. A. (1960) *Ann. N.Y. Acad. Sci. 88*, 586-601.
Yphantis, D. A. (1964) *Biochemistry 3*, 297-317.
Yphantis, D. A., & Waugh, D. F. (1956) *J. Phys. Chem. 60*, 623-629.

Molecular Modeling of the Domain Structure of C9 of Human Complement by Neutron and X-ray Solution Scattering

Kathryn F. Smith,[‡] Richard A. Harrison,[§] and Stephen J. Perkins*[,‡]

Department of Biochemistry and Chemistry, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, U.K., and MRC Molecular Immunopathology Unit, MRC Centre, Hills Road, Cambridge CB2 2QH, U.K.

Received August 1, 1991; Revised Manuscript Received October 11, 1991

ABSTRACT: C9 is the most abundant component of the membrane attack complex of the complement system of immune defense. This is a typical mosaic protein with thrombospondin (TSR) and low density lipoprotein receptor (LDLr) domains at its N-terminus and an epidermal growth factor-like (EGF) domain at its C-terminus. Between these lies a perforin-like sequence. In order to define the arrangement in solution of these four moieties in C9, high-flux neutron and synchrotron X-ray solution scattering studies were carried out. The neutron radius of gyration R_G at infinite contrast is 3.33 nm, and its cross-sectional R_G (R_{XS}) is 1.66 nm. Similar values were obtained by synchrotron X-ray scattering after allowance for radiation effects. Stuhrmann analyses showed that the neutron radial inhomogeneity of scattering density α is 35×10^{-5} from the R_G data and 16×10^{-5} from the R_{XS} data. These values are typical for soluble glycoproteins and show no evidence for the existence of any large hydrophobic surface patches on free C9 that might form contacts with lipids. Indirect transformation of the neutron and X-ray scattering curves into real space showed that C9 had a maximum dimension estimated as 12 ± 2 nm, and this suggests that the lengths of 7-8 nm deduced from previous electron microscopy studies in vacuo are underestimated. Molecular modeling of the C9 scattering curves utilized small spheres in the Debye equation, in which the analyses were constrained by the known volumes of the four moieties of C9 and the known sizes of the TSR and EGF-like domains. The most likely models for C9 suggest that these four regions of C9 are arranged in a V-shaped structure, with an angle of 10° between the two arms, each of length 11.1 nm. This structure has a more hydrophobic character between the two arms. The scattering model is fully consistent with hydrodynamic sedimentation data on C9. Similar V-shaped hydrodynamic models could be developed for C6, C7, C8, and C9 of complement. Such a compact structure is atypical of other multidomain complement proteins so far studied by solution scattering and is fully compatible with mechanisms in which C9 is postulated, on activation, to undergo a drastic unfolding of its domain structure and to expose a more hydrophobic surface which can be embedded into lipid bilayers.

The activation of the complement system in response to the challenge of foreign material in plasma provides a major non-adaptive immune defense mechanism for its host (Reid, 1986; Law & Reid, 1988). Initially, activation proceeds through a series of limited proteolytic steps in one of two largely independent pathways, the classical and alternative pathways. These merge in the lytic pathway, in which the generation of C3b constitutes the essential component of the C5 convertases, which generate C5b from C5. The sequential addition of the late complement components C6, C7, and C8 (a three-chain molecule constructed from C8α, C8β, and C8γ) and several

* To whom correspondence and requests for reprints should be addressed.
[‡] Royal Free Hospital School of Medicine.
[§] MRC Molecular Immunopathology Unit.

0006-2960/92/0431-754$03.00/0

molecules of C9 (n = 1–18) to C5b then leads to the formation of the membrane attack complex (MAC).[1] Concomitant with C9 recruitment, components of the MAC insert into the membrane. Electron micrographs show that the MAC appears as a ring-like or cylinder-like pore structure penetrating the membrane. This creates open lesions through which cellular components can leak, thereby causing cell death (Borsos et al., 1964; Tschopp et al., 1982, 1984; Dankert & Esser, 1985; Bhakdi & Tranum-Jensen, 1986). It is thus clear that C9 is the major protein involved in the terminal stage of complement-mediated cell lysis.

C9 consists of a 537-residue polypeptide chain to which are attached two N-linked oligosaccharides (DiScipio et al., 1984; Stanley et al., 1985; Kontermann & Rauterberg, 1989). C9 contains three cysteine-rich presumed independently folded domains, namely, a thrombospondin repeat (TSR) and a low density lipoprotein receptor (LDLr) domain at its N-terminus and an epidermal growth factor-like (EGF) domain at its C-terminus. The TSR and LDLr domains form a large part of the hydrophilic C9a fragment (M_r 28 000) formed by the cleavage of C9 into C9a and C9b by α-thrombin (Biesecker et al., 1982; DiScipio et al., 1984; Stanley et al., 1985; Shiver et al., 1986). While the central region of the C9 sequence has strong sequence similarities with perforin, a pore-forming protein secreted from cytotoxic T lymphocytes (Shinkai et al., 1988), and has a similar amino acid composition to the remainder to the remainder of the molecule, this central region appears to possess some specific hydrophobic properties. Both C9 and perforin form large aqueous channels in membranes (internal diameter 10 nm for C9 and 16 nm for perforin) with little ion selectivity (Young et al., 1986).

The polymerization of C9 and its insertion into lipid bilayers necessitates a dramatic structural change in C9 in which the N-terminal and C-terminal segments of C9 appear to unfold (Podack & Tschopp, 1982a; Tschopp et al., 1982; DiScipio & Hugli, 1985). Interaction of C9 with the C5b–8 complex appears to catalyze unfolding of C9 such that its length is doubled (Tschopp et al., 1984). The sequential association of between 12 and 18 of these extended molecules forms a circular tubular polymer of M_r in excess of 10^6. Electron micrographs of tubular poly-C9 suggest that this assembly has a length of 16–21 nm and a diameter of 15–22 nm, in which its length is doubled when compared with dimensions of 7–8 nm × 5–5.5 nm for a monomer of free C9. Two major hypotheses for the action of C9 on lipid bilayers currently exist. In one, it is proposed that C9 forms a stable hollow polymeric structure in the bilayer and that leakage of cellular components through this cause cell death (the "donut" hypothesis); in the other, that C9 insertion destabilizes the lipid bilayer surrounding the inserted polymer by a detergent-like mechanism (the "leaky patch" hypothesis) (Mayer, 1972; Esser et al., 1979).

Other constituents of the MAC (C6, C7, C8α, and C8β) have structural, functional, and antigenical similarities to C9. Sequences show that, while C6, C7, C8α, and C8β all contain domains similar to those found in C9, they also contain between one and six extra TSR, SCR, and FIM domains at their amino- and/or carboxy-terminal ends (SCR, short complement/consensus repeat; FIM, factor I module) (Haefliger et al., 1989; Chakravarti et al., 1989; DiScipio & Hugli, 1989). C9 is thus the archetypal protein structure whose molecular properties are likely to be present in other components of the MAC.

Small-angle scattering yields structural data on the arrangement of domains in multidomain proteins, especially if the analyses can be constrained by known volumes or known structures (Perkins, 1988a,b). Here, the structure of C9 and its overall dimensions could be defined under conditions close to physiological. Interpretations of the scattering data were constrained by analyses of the sequences for the four moieties in C9 (Perkins et al., 1989; Smith et al., 1991a; K. F. Smith and S. J. Perkins, unpublished results; Cooke et al., 1987). We show that the solution structure of free soluble C9 is more elongated than previously believed from electron microscopy studies carried out in vacuo and yet has a more compact structure compared to other mosaic proteins of the complement system. The scattering data were compatible with a limited range of arrangements of the three domains and perforin-like region in monomeric C9. While such a disposition is atypical of other complement components so far studied by solution scattering, it is compatible with proposed mechanisms for the drastic unfolding of C9 when free C9 becomes embedded into lipid bilayer. Together with similar studies on C5 of complement (Perkins et al., 1990a), and complementary hydrodynamic modeling studies of C6, C7, C8, and C9, the two principal protein structures involved in MAC formation have now been investigated by scattering. This opens the way for more detailed studies of the proteins and their complexes in the MAC.

Materials and Methods

(*a*) *Preparation of C9 for Solution Scattering.* Three different C9 preparations were used in the studies described in this paper. Two were from freshly drawn plasma, and the third was from outdated plasma. No differences could be detected in the C9 purified from either source. Plasma volumes ranged from 2–4 L. The initial stages of C9 preparation (Na_2SO_4 precipitation, lysine-Sepharose chromatography, DEAE Sephacel chromatography, and Sepharose CL 6B chromatography) were as described in Harrison and Lachmann (1986) for multicomponent fractionation, with the following modifications. The NaCl concentration in the DEAE-Sephacel equilibration buffer was 50 mM for two preparations and 100 mM for the third, this starting concentration being determined by the requirement for other, earlier-eluting, plasma proteins. The position at which C9 was eluted (10.5–13 mS at 2 °C) was not affected by alteration of the equilibration conditions. Hydroxyl-apatite chromatography was not used. Instead, the C9-containing pool from Sepharose CL 6B was then concentrated by ultrafiltration and by $(NH_4)_2SO_4$ precipitation (60%) before gel filtration on Sephadex G150SF in 20 mM sodium/potassium phosphate, 200 mM NaCl, and 1 mM EDTA, pH 7.0 (2-L column; protein applied at 10 mg/mL in up to 15 mL). This final gel filtration was performed immediately prior to dialysis of the protein against data collection buffers. Only the peak C9-containing fractions from this column were used in solution scattering studies. EDTA was maintained in all buffers used throughout the preparation as it inhibits the spontaneous polymerization of C9 (Tschopp, 1984). $(NH_4)_2SO_4$ precipitation in addition to ultrafiltration was used to concentrate the C9 prior to gel filtration on Sephadex G150 as this removed the bulk of the transthyretin, the major contaminant at this point. While residual transthyretin (a loosely associated tetramer of identical subunits of M_r 16 000) was largely resolved from C9 on Sephadex G150, a low (<5% by weight) level still remained in the purified C9. No high molecular weight polymers of C9

[1] Abbreviations: TSR, thrombospondin repeat; LDLr, low density lipoprotein receptor; EGF, epidermal growth factor, SCR, short complement/consensus repeat; FIM, factor I module; MAC, membrane attack complex.

could be detected in the purified protein.

Neutron scattering data were collected using buffers containing 0.2 M NaCl, 20 mM sodium/potassium phosphate, and 1 mM EDTA, pH 7.0, or 0.5 M NaCl, 20 mM sodium/potassium phosphate, and 1 mM EDTA, pH 7.0. The neutron samples of C9 were dialyzed at 6 °C into their buffers containing 0%, 80%, or 100% ^2H_2O for at least 36 h with four buffer changes. X-ray data were collected on the basis of C9 in buffers containing either 0.2 M NaCl, 0.02 M borate, 1 mM $CaCl_2$, and 5 mM EDTA, pH 7.5, or 0.15 M NaCl, 12.5 mM phosphate, and 0.5 mM EDTA, pH 7.0.

(*b*) *Solution Scattering Data Measurements on C9.* Neutron scattering data on C9 were principally obtained in two independent sessions on Instrument D17 at the Institut Laue Langevin, Grenoble, France. Two further D17 sessions were used to complete data collection on C9 in 100% ^2H_2O. Samples were measured at 20 °C in rectangular quartz Hellma cuvettes of path length 2 mm for samples in 80% and 100% ^2H_2O buffers and 1 mm for samples in H_2O buffers; these were also used for 280-nm absorbance measurements. Sample and buffer transmissions were measured relative to an empty cell transmission for use in data reduction and analyses. Sample–detector distances of 1.40 and 3.46 m, a wavelength λ of 1.106 nm, and a main beam detector angle of 0° and 20° were used. The resulting Q range was 0.05–3.25 nm^{-1} (where $Q = 4\pi \sin\theta/\lambda$ and 2θ is the scattering angle). Individual sample counting times typically ranged between 3 and 21 min for C9 concentrations c between 2.4 and 7.6 mg/mL in 100% ^2H_2O buffers and between 26 and 104 min for c between 2.3 and 6.8 mg/mL in 0% ^2H_2O buffers. Data reduction was performed using standard Grenoble software (RNILS, SPOLLY, RGUIM, and RPLOT; Ghosh, 1989). A cadmium background was first subtracted from each scattering curve. The buffer background run was subtracted from that of the sample run, and the result was normalized for the detector response by using a water run from which an empty cell background had been subtracted.

X-ray scattering data were obtained using the low-angle solution scattering camera at station 8.2 (Towns-Andrews et al., 1989) at the SRS Daresbury, Warrington, U.K. Experiments were performed with beam currents of 99–130 mA and a ring energy of 2.00 GeV in two sessions. Samples were measured for 10 min at concentrations between 2 and 23 mg/mL. The use of a 500-channel quadrant detector at sample-to-detector distances of 4.16 and 3.45 m resulted in Q ranges of 0.11–1.61 nm^{-1} and 0.08–2.40 nm^{-1}. The Q range was calibrated using fresh, wet, slightly stretched rat tail collagen, on the basis of a diffraction spacing of 67.0 nm. Samples were held in Perspex cells of sample volume 20 μL, contained within mica windows of thickness between 10 and 15 μm, in a brass temperature-controlled sample holder connected to a water bath at 20 °C. Buffers and samples were measured in alternation for equal times to minimize background subtraction errors; data were only accepted if the subsequent Guinier plots were linear and led to reproducible data in repeated measurements (Perkins et al., 1990b). Curves were recorded in 10 time frames as a control for time-dependent radiation damage effects. Data reduction was performed using the standard Daresbury software package OTOKO (P. Bendall, J. Bordas, M. H. C. Koch, and G. R. Mant, EMBL Hamburg and SERC Daresbury Laboratory, unpublished software). The raw scattering curves were normalized on the basis of an ion chamber monitor positioned after the sample and a detector response measured for 8 h using a uniform ^{55}Fe radioactive source. Reduced curves were calculated by subtraction of the buffer runs from those of the samples.

(*c*) *Solution Scattering Data Analyses.* At small Q values, analyses of the scattering curves by the Guinier equation gave the radius of gyration R_G (which is a measure of macromolecular elongation) and the forward scattered intensity at zero $QI(0)$ (Glatter & Kratky, 1982):

$$\ln I(Q) = \ln I(0) - R_G^2Q^2/3$$

Absolute molecular weights M_r were obtained from the neutron $I(0)/c$ values measured in H_2O buffers, in which systematic errors of measurements were minimized (Kratky, 1983; Jacrot & Zaccai, 1981). If one of the molecular dimensions is much larger than the other two, the radius of gyration of the cross section R_{XS} and the cross-sectional intensity at zero angle $[I(Q)Q]_{Q\to 0}$ were obtained in a larger Q range than that used in the R_G analyses (Pilz, 1982):

$$\ln [I(Q)Q] = \ln [I(Q)Q]_{Q\to 0} - R_{XS}^2Q^2/2$$

In neutron contrast variation studies, the matchpoint of C9 [at which C9 was invisible to neutrons, i.e., when $I(0)$ is zero] was obtained experimentally from a plot of $\sqrt{[I(0)/cT_st]}$ or $\sqrt{[[I(Q)Q]_{Q\to 0}/ctT_s]}$ against percent ^2H_2O (T_s, sample transmission; t, path length). This was also calculated from the amino acid and carbohydrate composition on the basis that the protein volume was unhydrated and there was 10% of nonexchange of the main-chain NH protons (Perkins, 1986). Since different constituents within C9 have different scattering densities, the R_G depended on the solvent scattering density. This dependence was described by a simplified form of the Stuhrmann equation (Ibel & Stuhrmann, 1976; Perkins, 1988a,b):

$$R_G^2 = R_C^2 + \alpha/\Delta\rho$$

where R_C is the R_G at infinite contrast, $\Delta\rho$ is the difference in scattering density ρ between the solute and solvent, and α is the radial inhomogeneity of scattering density fluctuations within the protein. α is positive if the outermost scattering densities are higher than those nearer the center. The subscripts G and XS correspond to the R_G and R_{XS} analyses, respectively.

Triaxial molecular dimensions were obtained from the Guinier and cross-sectional plots. If C9 was represented by an elliptical cylinder, its length L was given by $[12(R_G^2 - R_{XS}^2)]^{1/2}$ (Glatter & Kratky, 1982). From intensity ratios, $L = \pi I(0)/[I(Q)Q]_{Q\to 0}$ (Perkins et al., 1986). The two semiaxes A and B of the elliptical cylinder were calculated by combining the dry (neutrons) or hydrated (X-rays) volume V ($V = \pi ABL$) with the R_{XS}, where $R_{XS}^2 = (A^2 + B^2)/4$ for an elliptical cylinder. The hydrated volume was obtained on the basis of a hydration of 0.3 g of water/g of glycoprotein and 0.0245 nm^3 per water molecule (Perkins, 1986).

At the Royal Free, London, neutron and X-ray data analyses were performed jointly using a common interactive analysis and plotting program SCTPL3 (Perkins & Sim, 1986; A. S. Nealis and S. J. Perkins, unpublished software). Results were processed as a spreadsheet using Lotus 1-2-3 version 2.01 (Lotus Development, U.K.). Statistical analyses were performed using MINITAB version 6.1 (Ryan et al., 1985).

Indirect transformation of the scattering data $I(Q)$ into real space $P(r)$ was carried out using the ITP program (Glatter, 1982). This offered an alternative calculation of the R_G and L for C9 that is based on the whole scattering curve measured out to $Q = 3.25$ nm^{-1} (neutrons) or 1.61 or 2.40 nm^{-1} (X-rays). A total of eight splines (neutrons and X-rays) were used to fit the experimental curves (83 and 113 data points respec-

tively). The maximum dimension D_{max} of C9 for ITP was specified as 15 and 20 nm, respectively (in which D_{max} should be 50% greater than the anticipated value of L). The number of points used to define $P(r)$ in real space was 101. Criteria for an acceptable calculation of $P(r)$ were (i) $P(r)$ should exhibit positive values, (ii) the R_G from ITP should agree with the R_G from Guinier analyses, (iii) $P(r)$ is zero when r is zero, and (iv) $P(r)$ is stable and reproducible for different experimental $I(Q)$ curves when the number of splines and D_{max} is varied over a reasonable range. L was determined from $P(r)$ when this became zero at large r; errors in L can be significant as a result of the low intensity of $P(r)$ in this region.

(*d*) *Modeling of Scattering and Hydrodynamic Data for C6, C7, C8, and C9.* Calculations of scattering curves for C9 employed small Debye spheres to follow standard procedures (Glatter & Kratky, 1982) in application to neutron and synchrotron X-ray scattering data (Perkins & Weiss, 1983; Perkins & Sim, 1986). To constrain the simulations, the unhydrated volume (83.4 nm^3) of C9 was calculated from the composition using crystallographic data (Chothia, 1975; Perkins, 1986). This volume was arbitrarily subdivided into 168 cubes of side 0.792 nm. Overlapping spheres of the same volume as each cube was placed at the center of each cube. Neutron data correspond largely to this unhydrated volume, since the hydration shell that surrounds the macromolecule is invisible by neutron contrast variation (Perkins, 1986). This means that partial specific volumes $\bar{v}$ are larger if calculated from neutron matchpoints instead of by densitometric means. Neutron curve fitting was based on data in 0% ^2H_2O and in 80% and 100% ^2H_2O to correspond to high positive and high negative solute–solvent contrasts, respectively. A trial-and-error procedure was employed to test a wide range of different sphere models in order to determine which ones could be ruled out by the scattering data. Note that this method does not lead to the determination of a unique structure as the result of spherical averaging. Standard wavelength spread and beam divergence corrections were applied to the simulated neutron curves before these were compared with the experimental curves (Cusack, 1981; Perkins & Weiss, 1983). Since the C9 samples contained low amounts (>5%) of transthyretin (previously known as prealbumin), control calculations using the X-ray crystallographic coordinates (Blake et al., 1978; Brookhaven Protein Data Bank reference 2PAB) were carried out to show that its effects on the scattering curve simulations of C9 were negligible.

X-ray curve fitting for C9 (in a high positive solute–solvent contrast) was based on the hydrated volume. Here, the total volume of 110.5 nm^3 for the 168 Debye spheres was obtained using a hydration of 0.3 g of H_2O/g of glycoprotein and an electrostricted water molecule volume of 0.0245 nm^3 (in place of the free water volume of 0.0299 nm^3) (Perkins, 1986). The final Debye models were generated with each sphere volume set to be that of the cube of side 0.870 nm. No corrections were made for instrumental errors since these are negligible with synchrotron radiation.

Modeling of the sedimentation coefficient $s^\circ_{20,w}$ for C6, C7, C8, and C9 was based on the hydrated volume as in the X-ray curve fitting. The frictional coefficient f was calculated from

$$f = M_r(1 - \bar{v}\rho_{20,w})/N_a s^\circ_{20,w}$$

where $\rho_{20,w}$ is the density of water at 20 °C and N_a is Avogadro's constant. From sequences, the dry volumes of C6, C7, and C8 were computed to be 134.1, 119.8, and 193.1 nm^3 in that order, assuming that all the potential glycosylation sites were occupied with biantennary complex-type oligosaccharides. The calculated hydrated $\bar{v}$ of C9 is 0.722 mL/g (Perkins, 1986), and those for C6, C7, and C8 are 0.720, 0.719, and 0.721 mL/g. If a small number of hydrodynamic spheres were used, this hydration was replaced by a notional value of 0.39 g of H_2O/g of glycoprotein in order to compensate for the volume lost in the void spaces between the nonoverlapping spheres required for the simulations of f (Perkins, 1989). Alternatively, the hydrated X-ray model for C9 (168 nonoverlapping spheres of diameter 0.870 nm) was employed directly in the simulations. Calculations were performed by the modified Oseen tensor procedure using the program GENDIA (Garcia de la Torre & Bloomfield, 1977a,b).

Results

(*a*) *Neutron Scattering Studies of C9.* Neutron scattering was used to determine the overall solution structure of C9 and its internal structure from data in three contrasts at concentrations between 2.4 and 7.6 mg/mL. C9 samples were occasionally found to contain small amounts of nonspecific aggregates that resulted in nonlinear Guinier plots in the scattering analyses. For final studies, these were successfully eliminated by leaving the final gel-filtration stage until just prior to dialysis, and the dialysis was followed immediately by neutron data collection. The use of phosphate buffers in 0.5 M NaCl rather than 0.2 M NaCl and the use of 0% ^2H_2O rather than 100% ^2H_2O buffers also minimized aggregate formation. Figure 1a,b shows that the resulting neutron Guinier plots for C9 in 0.5 M NaCl were linear over a fully acceptable QR_G range of 0.6–1.5.

The Guinier intensity $[I(0)/c]$ and radius of gyration (R_G) parameters showed no dependence on C9 concentrations c between 2.2 and 7.7 mg/mL or buffer compositions in 0.2 M NaCl or 0.5 M NaCl (data not shown). The molecular weight M_r for C9 was calculated from the $I(0)/c$ values measured in H_2O buffers (Jacrot & Zaccai, 1981). Values of c were determined using an A_{280} coefficient of 8.37 (1%, 1 cm) calculated (Perkins, 1986) from the human sequence. The calculation assumed that there were two triantennary complex-type carbohydrate chains per C9 molecule (Konterman & Rauterberg, 1989). This composition resulted in a M_r of 66 400, which was comparable with the M_r of 71 000 from SDS–PAGE (Biesecker & Müller-Eberhard, 1980). The 8.6% carbohydrate content by weight was in good agreement with experimental values of 7.8–8.0% (Biesecker & Müller-Eberhard, 1980; DiScipio & Hugli, 1985). The predicted A_{280} value of 8.37 was, however, 11%–18% lower than literature values of 9.6 (Podack & Tschopp, 1982a) and 9.88 (Monahan et al., 1983) based on quantitative amino acid analyses. The total of scattering lengths $\sum b$ in C9 normalized by M_r gave a $\sum b/M_r$ of 0.2308 fm in H_2O solution, which is typical for globular proteins. From the neutron data, the M_r of C9 was determined to be 63 500 ± 3000 (five values), which was in excellent agreement with the value of 66 400 from the composition, showing that C9 was monomeric. Had the A_{280} value from amino acid analyses been used, the M_r would have been overestimated as 71 000–76 000.

The neutron contrast variation $I(0)/c$ data (Figure 2) resulted in a linear plot with a matchpoint of 42.9 ± 0.3% ^2H_2O. This was in good agreement with the prediction of 42.3% ^2H_2O from the amino acid and carbohydrate composition. In summary, the linear Guinier plots and the agreements with M_r and matchpoint data showed that the scattering curves corresponded to monodisperse preparations of C9, as required for structural analyses.

The Stuhrmann analysis of the dependence of the neutron R_G data in 0.2 M and 0.5 M NaCl on the solute–solvent contrast resulted in an R_{C-G} of 3.33 ± 0.03 nm at infinite

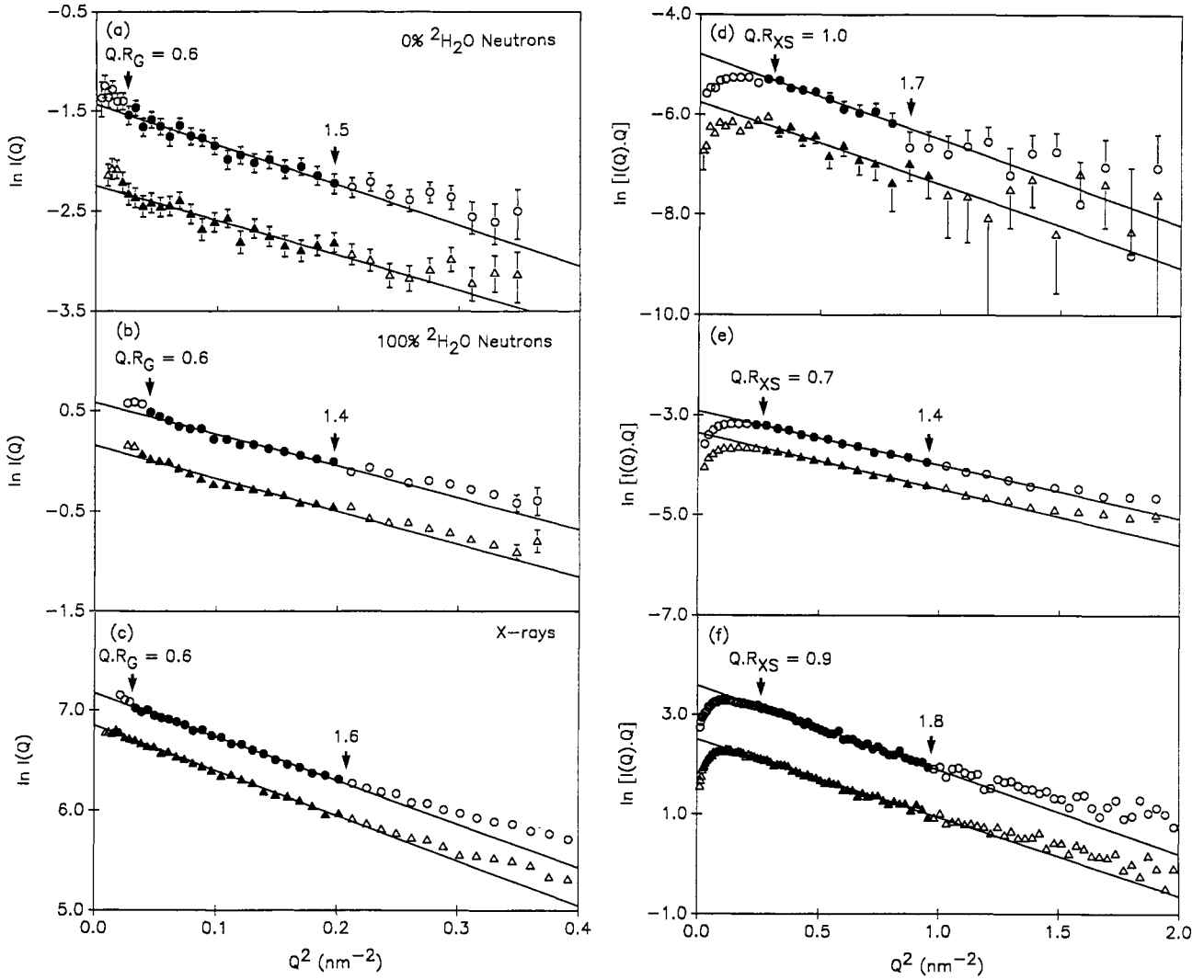

FIGURE 1: Guinier analyses of C9 by neutron and X-ray scattering. Filled circles correspond to the $I(Q)$ data employed in determining the R_G and R_{XS} values, and the straight lines correspond to the best fit through these points. The Q range used for the R_G values is 0.20–0.44 nm^{-1}, and that for the R_{XS} values is 0.49–0.97 nm^{-1}. The QR_G and QR_{XS} values are based on the R_G and R_{XS} values below. The neutron data were obtained in 0.5 M NaCl buffers, while the X-ray data were obtained in 0.2 M NaCl buffers (Materials and Methods). The X-ray curves in panels c and f are displaced on the verticl axis for clarity. (a and d) Neutron R_G and R_{XS} Guinier plots, respectively, for C9 in 0% ^2H_2O buffers, sample concentrations 6.9 and 4.6 mg/mL, and a mean R_G of 3.51 ± 0.18 nm (five values). (b and e) Neutron R_G and R_{XS} Guinier plots, respectively, for C9 in 100% ^2H_2O buffers, sample concentrations 7.3 and 4.7 mg/mL, and a mean R_G of 3.16 ± 0.10 nm (nine values). (c and f) X-ray R_G and R_{XS} Guinier plots respectively for C9 in H_2O buffers (after exposure times of 3 min); the sample concentration is 7.7 mg/mL, with a mean R_G of 3.66 ± 0.13 nm (eight values for c between 7.4 and 22.9 mg/mL).

contrast and a positive slope α_G of (35 ± 7) × 10^{-5} for C9 (Figure 2). Calculation of the elongation ratio or anisotropy R_{C-G}/R_0 (where R_0 is the R_G of the sphere with the same dry volume as C9) gave a value of 1.59 ± 0.01. Since the R_{C-G}/R_0 values for typical globular proteins are close to 1.28 (Perkins, 1988b), C9 possesses an extended solution structure. This structural anisotropy is similar to that found with the complement components C3, C4, and C5 but corresponds to a less elongated structure than those of C1̄ inhibitor, C1q, $C1r_2C1s_2$, C4bp, properdin, and factor H (Perkins et al., 1990c).

Since C9 is elongated, Guinier analyses of ln $[I(Q)Q]$ against Q^2 will provide the cross-sectional parameters $[I(Q)Q]_{Q\to 0}$ and R_{XS} (Pilz, 1982). As shown in Figure 1d,e, linear cross-sectional plots were readily obtained in an acceptable QR_{XS} range of 0.7–1.8. The data analysis of Figure 2 gave a matchpoint of 43.6 ± 0.5% ^2H_2O, in agreement with the expected value of 42.5% ^2H_2O. The cross-sectional radius of gyration at infinite contrast R_{C-XS} was 1.66 ± 0.02 nm, and the slope α_{XS} was again positive at (16 ± 2) × 10^{-5}.

Approximate molecular dimensions of C9 were determined assuming in the first instance that its shape is an elliptical cylinder. From the R_G and R_{XS} values at infinite contrast, the length L of C9 was determined as 10.0 ± 0.2 nm. The determination of L from the intensity ratio of 10.1 ± 0.8 nm (13 values) is in good agreement with this. Use of the dry volume of 83.4 nm^3 resulted in axial dimensions of 10.0 nm × 6.4 nm × 1.7 nm for C9. The cross section of C9 is therefore elongated also.

The values of α_G and α_{XS} can be compared with those published for other glycoproteins once these have been normalized on the basis of its proportionality to R_G^2 and R_{XS}^2. Using α_G values for lysozyme, myoglobin, trypsin, α_1-acid glycoprotein, α_1-antitrypsin, and C1 inhibitor (Stuhrmann & Fuess, 1976; Ibel & Stuhrmann, 1975; S. J. Perkins, unpublished data; Perkins et al., 1985, 1990d; Smith et al., 1990), α_G for C9 was predicted to be of the order (32 ± 17) × 10^{-5}. This was in good agreement with the observed α_G of (35 ± 7) × 10^{-5} (Figure 2). Using α_{XS} values for α_1-antitrypsin, C4, C4u, C4b, C5, and factor H (R_{XS-2}) (Smith et al., 1990; Perkins et al., 1990a,b, 1991a), the predicted value of α_{XS} for C9 would be close to (13 ± 5) × 10^{-5}. This was also in good agreement with the observed α_{XS} for C9 of (16 ± 2) × 10^{-5}.

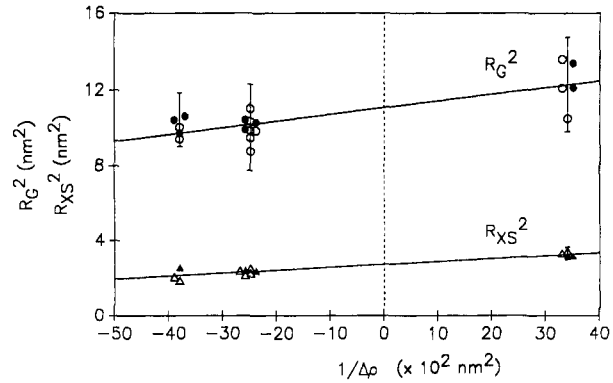

FIGURE 2: Neutron contrast variation analyses of C9. A concentration range of 7.6–2.4 mg/mL was used. Buffers were either 0.2 M NaCl or 0.5 M NaCl in 12 mM phosphate, pH 7.4. Statistical error bars were shown only when these were large enough to be visible. The matchpoint determination of C9 from a graph of $\sqrt{[I(0)/cT_s]}$ versus the volume percentage of ^2H_2O was based on the following mean values: 0% ^2H_2O, 0.285 ± 0.006 (five values); 80% ^2H_2O, −0.242 ± 0.003 (five values); 100% ^2H_2O, −0.385 ± 0.009 (nine values). The corresponding graph based on $\sqrt{[[I(Q)\cdot Q]_{Q\to 0}/cT_s]}$ data used the following values: 0% ^2H_2O, 0.052 ± 0.003 (four values); 80% ^2H_2O, −0.042 ± 0.001 (three values); 100% ^2H_2O, −0.067 ± 0.001 (six values). MINITAB regressions gave matchpoints of 42.9 ± 0.3 % ^2H_2O (19 points) and 43.8 ± 0.5 ^2H_2O (13 points), respectively. (b) Stuhrmann plot of R_G^2 against the reciprocal solute–solvent contrast difference $\Delta\rho^{-1}$. The R_G data correspond to buffers in 0.2 M NaCl (○) and 0.5 M NaCl (●); the R_{XS} data correspond likewise to 0.2 M NaCl (△) and 0.5 M NaCl (▲). MINITAB regressions gave an R_{G-C} of 3.33 ± 0.03 nm at zero $\Delta\rho^{-1}$ and a slope α_G of (35 ± 7) × 10^{-5} (19 points) and an R_{XS-C} of 1.66 ± 0.02 nm at zero $\Delta\rho^{-1}$ and a slope α_{XS} of (16 ± 2) × 10^{-5} (13 points).

[CĪ inhibitor (Perkins et al., 1990d) has an exceptionally high α_{XS}; however, this has an unusually high content (26% by mass) of extended O-linked and N-linked oligosaccharides located along the long axis of this structure.] These comparisons show that C9 contains a similar radial distribution of hydrophilic surface residues and hydrophobic core residues as that found in other soluble glycoproteins.

Use of ITP to convert the scattering curve $I(Q)$ in reciprocal space into its distance distribution function $P(r)$ in real space was applied to three full curves for C9 at 7.3, 4.7, and 2.2 mg/mL in 0.5 M NaCl buffers in 100% ^2H_2O. Using the ITP parameters specified in Figure 3a, the R_G value from ITP was determined as 3.32 nm, and the $I(0)$ value was 1.84 for the sample at 7.3 mg/mL. These were in excellent agreement with the values of 3.08 nm and 1.79, respectively, from the Guinier plots for that C9 sample. In this high negative solute–solvent contrast in 100% ^2H_2O, the three $P(r)$ curve showed that the most frequently occurring distance r within C9 was 3.00–3.15 nm at its maximum. The error bars in the calculation of $P(r)$ at large r showed that the maximum dimension of C9 was in a range between 10.5 and 14 nm. These dimensions were comparable to but larger than the two estimates of L of 10.1–10.2 nm above from the Guinier analyses. The difference was attributed to the difficulty in determining a precise cut-off when $P(r)$ approached zero at large r.

(b) Synchrotron X-ray Scattering Studies of C9. In order to complement the neutron data, synchrotron X-ray experiments (corresponding to high positive solute–solvent contrasts) were performed with C9 at concentrations of 1.8–22.9 mg/mL. However, irradiation caused substantial aggregate formation during the 10-min exposure time to X-rays. The R_G and R_{XS} Guinier analyses of 10 1-min time frames for four C9 samples are shown as a function of time in Figure 4. The mean R_G (4 values) increased from 3.51 ± 0.16 nm after 1 min to 4.61 ± 0.18 nm after 10 min, measured with beam currents of

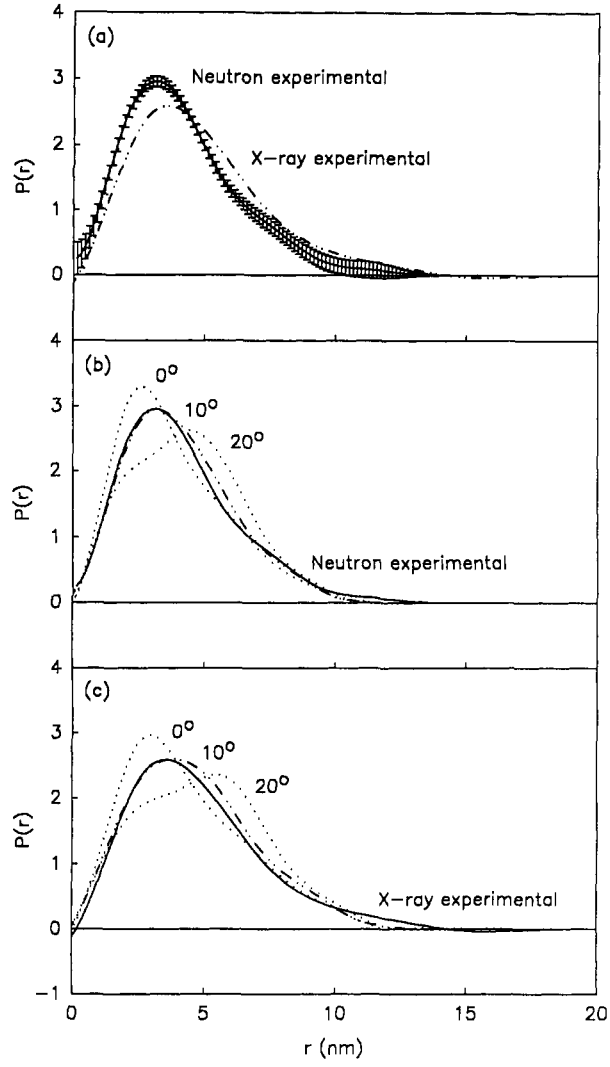

FIGURE 3: Distance distribution functions $P(r)$ calculations (Glatter, 1982) based on the neutron and X-ray $I(Q)$ data for C9. (a) The neutron $P(r)$ curve corresponds to the solid line with statistical error bars, using a C9 sample at 7.3 mg/mL in a 100% ^2H_2O buffer with 0.5 M NaCl. The neutron maximum length of C9 lies between 10.5 and 14 nm, based on a calculation with eight splines and D_{max} of 15 nm. The X-ray $P(r)$ curve is depicted by the dashed line; this corresponds to C9 at 7.4 mg/mL in a buffer with 0.2 M NaCl. The X-ray maximum length of C9 is estimated as 14 ± 2 nm. (b) Comparison of the neutron experimental $P(r)$ curve (solid line) with three simulated $P(r)$ curves for two-density V-shaped models of C9 in which the angle between the arms is set at 0°, 10°, and 20° (dotted or dashed lines), and the densities are set to correspond to negative solute–solvent contrasts. The maximum of $P(r)$ is located at r = 2.6, 3.2, and 4.5 nm for angles of 0°, 10°, and 20°, in that order. (c) Comparison of the X-ray experimental $P(r)$ curve (solid line) with the same models for C9 as in panel b, now hydrated and density weighted for positive solute–solvent contrasts. The maximum of $P(r)$ is located at r = 3.0, 3.6, and 5.4 nm for angles of 0°, 10°, and 20°, in that order.

112–130 mA. In another session, the mean R_G (eight values) rose from 3.64 ± 0.25 nm to 4.03 ± 0.22 nm after 10 min (beam current 99–105 mA). The nonuniform increase in R_G may be due to the different beam currents or sample buffers in use. Analyses of irradiated C9 samples by SDS–PAGE showed no difference from the starting material. This suggested that some form of noncovalent association is induced by the X-ray beam. Since the R_G and $I(0)$ values were constant in the first 3 min of exposure (Figure 4a,b), only the first three time frames were used for Guinier analyses. The R_G Guinier analyses of Figure 1c were linear. No concentration dependence was observed, and the final mean R_G value was 3.66 ± 0.13 nm. These data could be compared with the

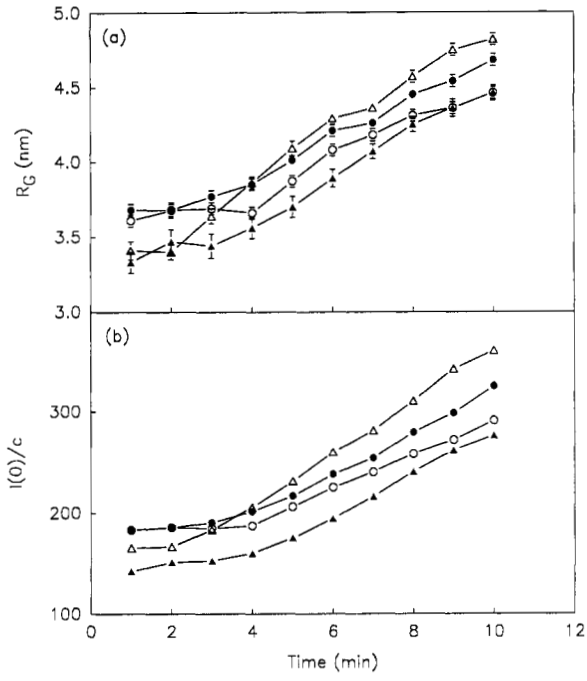

FIGURE 4: Synchrotron X-ray radiation damage effects observed in the R_G and R_{XS} Guinier analyses of C9. The R_G and $I(0)/c$ data from each of 10 consecutively recorded time slices for four C9 samples all at 7.4 mg/mL in 0.2 M NaCl, 0.02 M borate, 1 mM $CaCl_2$, and 5 mM EDTA, pH 7.5, are shown in panels a and b, respectively, and denoted by ○, ●, △, and ▲.

neutron R_G values in 0% ^2H_2O buffers, which also correspond to high positive solute–solvent contrasts. Good agreement was found with the mean neutron R_G of 3.51 ± 0.18 nm for C9 in H_2O (Figure 2).

The X-ray cross-sectional Guinier analyses were linear also (Figure 1f). The R_{XS} and $[I(Q)Q]_{Q\to 0}$ parameters were not affected over time (data not shown), where the mean R_{XS} (four values) was unchanged from 1.82 ± 0.08 nm after 1 min to 1.93 ± 0.07 nm after 10 min. The mean R_{XS} was 1.84 ± 0.07 nm (15 values, in excellent agreement with the neutron R_{XS} of 1.81 ± 0.03 nm in H_2O (Figure 2). The value of L from the intensities ratio was 12.3 ± 0.9 nm (eight values), while that from the mean R_G and R_{XS} data was 11.0 ± 0.7 nm. The molecular dimensions of C9 were determined as 11.0 nm × 7.1 nm × 1.8 nm if this was represented by an elliptical cylinder of hydrated volume 110.6 nm^3.

The ITP calculation of $P(r)$ in high positive solute–solvent contrast was noticeably different from that of the neutron curve (Figure 3a). The main envelope of the $P(r)$ curve was shifted to larger r, and the maximum in $P(r)$ was increased to 3.40–3.60 nm. The difference between the neutron and X-ray curves was attributed to the hydrophilic nature of C9. The more hydrophilic residues in C9 are positioned farther away from the center of scattering densities than the more hydrophobic ones. The higher scattering densities of the former makes a stronger contribution to the X-ray $P(r)$ at large r and a weaker one to the neutron $P(r)$ at large r. The ITP calculation also showed that the largest dimension of C9 was estimated to fall in a range of 14 ± 2 nm by X-rays, although again a precise cut-off cannot be determined. This dimension was comparable with that estimated from the neutron $P(r)$ curve. Since a uniform monolayer of H_2O surrounding a protein surface would increase D_{max} by twice its thickness of 0.36 nm (Perkins, 1989), such a typical hydration was not expected to influence the determination of D_{max}.

(*c*) *Scattering Curve Simulations for C9.* The use of small Debye spheres to model the full scattering curve in the Q range between 0.2–1.6 nm^{-1} enabled molecular models for C9 to be constructed. These permitted the assessment of possible arrangements of the TSR, LDLr, and EGF-like domains in C9.

Initial calculations were based on the assumption that C9 has a structure resembling an elliptical cylinder. Good fits for the neutron scattering curves in 0%, 80%, and 100% ^2H_2O buffers (not shown) were obtained with a sphere model based on overall dimensions of 10.9 × 6.8 × 1.4 nm and two internal scattering densities of 35% ^2H_2O (innermost 120 spheres) and 55% ^2H_2O (outermost 144 spheres). This was constructed as an array of 16 × 10 × 2 cubes of side 0.681 nm, from which 14 were removed at each corner to lead to a total of 264 spheres. This C9 model gave an R_G of 3.31 nm, which was close to the observed neutron R_{G-C} of 3.33 nm at infinite contrast, and a Stuhrmann α of 34 × 10^{-5}, which was in good agreement with the observed value of 35 × 10^{-5}. The dimensions of this model were close to those of the elliptical cylinder derived from the Guinier analyses above. The goodness-of-fit parameter R [defined by analogy to crystallography (Smith et al., 1990)] should generally be less than 0.02. Here, R was 0.009 (0% ^2H_2O, out to a Q of 0.09 nm^{-1}), 0.011 (80% ^2H_2O) and 0.011 (100% ^2H_2O) [where the poor counting statistics in 0% ^2H_2O precluded the inclusion of data beyond Q of 0.9 nm^{-1} in that contrast]. This model was however oblate in shape, and this may be unrealistic for a globular protein.

More satisfactory neutron curve fits were obtained by the use of V-shaped models for C9. Sphere models based on 168 cubes of side 0.792 nm were arranged into a two-armed structure, with each arm of length 10.3 nm, 11.1 nm or 11.9 nm (length 13, 14, or 15 spheres) and cross-sectional dimensions 2.4 nm × 2.4 nm (up to 3 × 3 spheres). The two identical arms were joined together at one end and separated by angles of 0°–20°. The internal distribution of hydrophilic and hydrophobic residues was considered (Figure 6a). At the center of this structure, 72 spheres were assigned as hydrophobic (with a matchpoint of 35% ^2H_2O); around the periphery of this model, the remaining 96 spheres were set as hydrophilic (with a matchpoint of 55% ^2H_2O) (Figure 6b). Using neutron data in 0%, 80%, and 100% ^2H_2O, the testing of these nine models showed that that of arm length 11.1 nm with an interarm angle of 10° (Figure 6a) gave the best curve fits. Angles of 0° (corresponding to an elliptical cylinder) and 20° clearly deviated from the experimental $P(r)$ and $I(Q)$ curves (Figures 3b and 5a). The length of the arms was in accord with the maximum distance calculated from $P(r)$ (Figure 3b). This model had a R_{G-C} of 3.37 nm, with individual R_G values of 3.53 (0% ^2H_2O), 3.14 (80% ^2H_2O), and 3.23 nm (100% ^2H_2O) and an α_G of 35 × 10^{-5}. These data agreed very well with the observed values reported in Figures 1 and 2. The use of different arrangements of hydrophobic and hydrophilic spheres within the C9 model resulted in poorer agreements with α_G, such as that in which an internal hydrophobic core was located along the central axis of each arm of C9. The goodness-of-fit parameter R was 0.009 (out to Q of 0.9 nm^{-1}), 0.007, and 0.007 in that order for the 0%, 80%, and 100% ^2H_2O curve fits. The quality of the curve fits was thus slightly improved over that for the oblate model. This was particularly so in the Q range above 1.0–1.1 nm^{-1} in 80% and 100% ^2H_2O buffers (Figure 5a–c); note that these buffer contrasts accentuate the hydrophobic structure of C9.

This neutron model was hydrated in order to test the synchrotron X-ray data. 168 cubes of side 0.870 nm were now used. Since the inclusion of both hydration and two densities (401 e/nm^3 and 433 e/nm^3 for the hydrophobic and hydro-

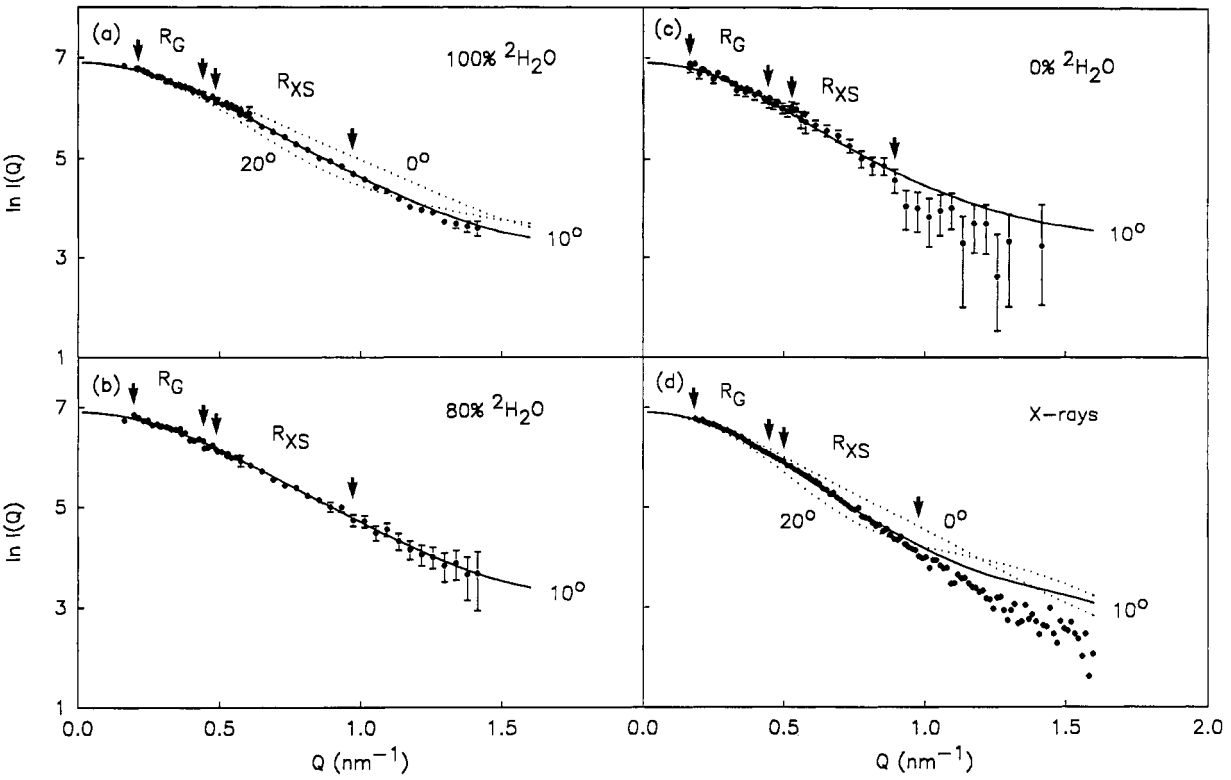

FIGURE 5: Comparisons of the curve simulations for C9 with experimental data. The C9 curve (solid line) calculated from the two-density V-shaped model with an interarm angle of 10° shown in Figure 6a (sphere diameter 0.792 nm) is compared with the experimental curves measured in (a) 100%, (b) 80%, (c) 0% ^2H_2O buffers on Instrument D17. In panel a, scattering curve simulations are shown for V-shaped models of C9 in which the interarm angle was set also as 0° and 20° (dotted lines). C9 samples were at concentrations of 7.3, 7.3, and 6.90 mg/mL in that order in panels a, b, and c. Guinier R_G and R_{XS} analyses were carried out in the Q ranges specified in Figure 1 as indicated by the arrows. Error bars due to data counting statistics are shown when these are large enough to be seen. In panel d, the curve fit to the X-ray data is shown (sample concentration 7.3 mg/mL), based on a hydrated one-density V-shaped model with interarm angles of 0°, 10°, and 20° to follow panel a.

philic components, relative to 334 e/nm^3 for the solvent) have similar effects on the scattering curves (Perkins et al., 1991b), only the hydration level of a single density C9 model was varied in the curve fits. The best curve fit for the V-shaped model corresponded to the neutron model of Figure 6a, with hydrated arms of length 12.2 nm and an interarm angle of 10°. Other models with angles of 0° and 20° gave large deviations from the experimental $I(Q)$ and $P(r)$ X-ray data (Figures 3c and 5d). The R_G and R_{XS} data were fully explained by this modeling. The R was, however, high at 0.070. This was attributed to the deviations seen beyond Q of 0.9 nm^{-1}, which was seen also with the neutron 0% ^2H_2O curve fit of Figure 5c, and this shows that the hydrophilic regions of C9 is not as well modeled as the hydrophobic regions.

(d) Simulations of Hydrodynamic Data for C6, C7, C8, and C9. For C9, sedimentation coefficients $s°_{20,w}$ of 4.5 S (Podack & Tschopp, 1982b; Tschopp et al., 1984) and 4.7 S (Tschopp, 1984) have been reported. These provided an independent measurement of the degree of elongation of the C9 structure in solution. From a hydrated volume of 110.5 nm^3, a frictional ratio f/f_o of 1.22 ± 0.05 was calculated. This is one of the lowest ratios found among the complement proteins (Perkins et al., 1990c).

Hydrodynamic spheres were used to calculate $s°_{20,w}$ values for C9. For example, simple models based on six spheres of diameter 3.34 nm arranged in a 3 × 2 array (dimensions 10.0 × 6.7 × 3.3 nm) resulted in predicted $s°_{20,w}$ values of 4.7–4.4 S as the angle between the two arms of length three spheres was increased from 0° to 20°. The best agreement was obtained with an angle of 10° between the arms (Figure 6c), which gave a $s°_{20,w}$ of 4.5 S. A typical error of ±0.2 S in the simulation (Perkins, 1989) corresponded to changes in the length of ±2 nm. In a second approach, the direct use of the final hydrated X-ray sphere model gave $s°_{20,w}$ values of 4.77, 4.49, and 4.40 S for angles of 0°, 10°, and 20°, in that order. Agreement with the experimental $s°_{20,w}$ of 4.5 S was excellent with the 10° X-ray model. In summary, both sets of calculations showed that the hydrodynamic length of C9 agrees well with the length of 11.1 nm determined by solution scattering modeling.

The availability of $s°_{20,w}$ values and sequences for the homologous proteins C6, C7, and C8 permit the six-sphere model for C9 to be compared with other proteins involved in MAC formation. The $s°_{20,w}$ for C6 was taken to be 5.7 ± 0.2 S, the mean of 5.5, 5.5, 5.6, 6.0, and 5.7 S (Podack et al., 1976; Podack & Tschopp, 1982, 1984; Priessner et al., 1985; DiScipio & Gagnon, 1982). That for C7 was taken as 5.4 ± 0.2 S (the mean of 5.3, 5.3, 5.1, 5.6, and 5.6 S). That for C8 was taken as 8.1 ± 0.1 S (the mean of 8.0 and 8.1 S). The frictional ratios f/f_o for C6, C7, and C8 were found to be 1.34 ± 0.05, 1.31 ± 0.05, and 1.19 ± 0.01. When these are compared with that for C9 (1.22 ± 0.05), the monomeric C6 and C7 structures are seen to be marginally more elongated than C9. This is as expected from the presence of six and five additional TSR, SCR, and FIM domains in C6 and C7, respectively, when compared with C9. The f/f_o of the C8$\alpha\gamma$–C8β trimeric complex is similar to C9. This suggests that the C8$\alpha\gamma$ and C8β subunits of C8 associate in a side-by-side interaction (rather than in an end-to-end one), in which the long axes of C8$\alpha\gamma$ and C8β are parallel and close to each other.

The $s°_{20,w}$ data on C6, C7, and C8 were quantified using hydrodynamic sphere models, starting from the six-sphere

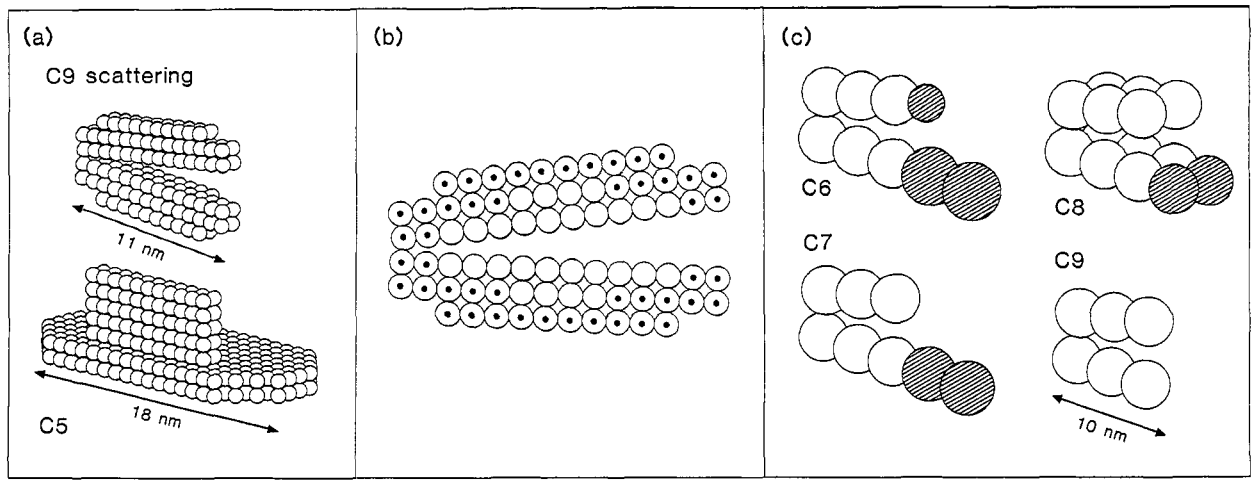

FIGURE 6: Sphere modeling of C9. (a) The scattering model of C9 (sphere diameter 0.792 nm) is shown in comparison with complement component C5 (sphere diameter 0.8 nm) (Perkins et al., 1990a). It should be noted that other scattering-equivalent models can offer equally satisfactory curve fits to the data and that the general form of these structures should not be overrated. (b) Face-on view of the two-density neutron model for C9 from panel a, in which the 72 hydrophobic and 96 hydrophilic spheres are denoted by open and spotted circles, respectively. (c) Comparison of possible hydrodynamic models for C6, C7, C8, and C9 of complement. The extra spheres relative to C9 are shaded for clarity. Sphere diameters: 3.34 nm for C9 (length 10.0 nm); 2.54 and 3.96 nm for the one and two additional spheres in C6, in that order; 3.70 nm for the two additional spheres in C7; 3.34 nm for the two additional spheres in C8 (which was constructed as a dimer of C9).

model for C9 (Figure 6c). It was assumed that all but one of the extra domains in C6, C7, and C8 relative to those found in C9 were most realistically positioned at one end of the V-shaped C9 model, i.e., at the C-terminus of C9. The additional TSR that is found at the N-terminus of C6 when compared to C7, C8α, C8β, and C9 was represented by a single sphere at the end of the second arm of the C9 model. The location of the C8γ subunit in C8 is unknown, and the volume of this was combined with those of the additional TSRs domains in C8 relative to C9. The volume of the additional spheres corresponded to the volume difference between C9 and each of C6, C7, and C8. Use of the models shown in Figure 6c resulted in calculated $s^{\circ}_{20,w}$ values of 5.6 S for C6, 5.2 S for C7, and 8.1 S for C8. These were in excellent agreement with the above experimental $s^{\circ}_{20,w}$ values. The length of the longest dimension of the C6 model is 18.0 nm; this is in satisfactory agreement with the lengths of 14.4 nm and 18 ± 2 nm for C6 that were deduced from electron microscopy (DiScipio & Hugli, 1989; Haefliger et al., 1989). The longest dimension of the C7 model is 17.5 nm; this is also close to the length of 15.1 nm for C7 from electron microscopy (DiScipio et al., 1988). These calculations show that the solution model for C9 is readily able to account for hydrodynamic data on C6, C7, and C8 in a manner that is consistent with sequence data and electron micrographs of C6 and C7.

Discussion

C9 is best visualized as a typical mosaic protein found in the complement system, but possessing unique structural features. Its two most distinctive properties are its ability to polymerize (either on its own or in association with C5b-8) and its ability to insert itself into and penetrate lipid bilayers (including cell membranes). Evidence from electron microscopy, tryptic cleavage experiments, and the analysis of monoclonal antibody binding sites shows that C9 undergoes a drastic unfolding when polymerized and inserted into membranes [summarized in Podack and Tschopp (1984) and Stanley et al. (1986)].

By the use of solution scattering techniques, free C9 was found to be fully monomeric in solution. That the A_{280} of C9 is 8.37 rather than 9.6–9.88 accounts fully for the M_r determination of C9 by neutron scattering. It should also be noted that the conventionally assumed M_r of C9 is reduced by about 11% to 66 400 when this is calculated from its sequence. In other words, these revised values may affect determinations of the stoichiometry of C9 within the MAC (generally estimated as 12–18 copies of C9).

Both the R_G and P(r) scattering calculations as well as the $s^{\circ}_{20,w}$ hydrodynamic calculation show that electron microscopy underestimated the length of free C9. This is now 11–12 nm. Electron micrographs of free C9 (Podack & Tschopp, 1982a; DiScipio & Hugli, 1985) had shown that C9 has dimensions of 8.0 × 5.5 nm or 7.0 × 5.0 (±0.7) nm. That these dimensions are less than those found here may result from artifacts caused by electron beam damage, the nonphysiological conditions of measurement in vacuo (using stains), or magnification errors. The dimensions of a tubular ring of poly-C9 are 15–16 nm in length with an internal diameter of 9–10 nm and an outer diameter of 18–21 nm (Podack & Tschopp, 1982a; DiScipio & Hugli, 1985), i.e., a monomer of C9 is increased in length from ca. 11–16 nm in poly-C9. The solution structure of C9 thus remains compatible with hypotheses that C9 unfolds to form a longer structure in poly-C9.

Of interest is that C9 has a relatively compact structure for a multidomain protein, as evidenced by the R_G/R_0 and f/f_0 ratios from scattering and hydrodynamic data. The 11–12-nm length of C9 from these solution analyses can be compared with the sizes of the TSR, LDLr, and perforin-like and EGF-like moieties in C9. The TSR is of approximate size 4 nm × 1.7 nm × 1.7 nm (Smith et al., 1991a), while the EGF-like domain may be approximately 3.5 nm × 2 nm × 2 nm (Cooke et al., 1987). If spherical (the most compact shape possible), the LDLr domain would be of diameter 2.2 nm, and the perforin-like region would be of diameter 4.6 nm (K. F. Smith and S. J. Perkins, unpublished results). The total length of a linear arrangement of these four structures will therefore be between 10.5 and 14.3 nm, depending on the orientation of the TSR and EGF-like domains. While it is possible that these four domains can be contained within an elliptical cylinder of dimensions 10.9 × 6.8 × 1.4 nm (Results), such a structure appears to be too thin. It is more likely that the LDLr domain and the perforin-like region are nonspherical in shape and that the long axes of the TSR and LDLr domains are colinear with that of C9. In such a case, the hypothesis

of a compact V-shaped structure for C9 becomes plausible, within which the perforin-like region could be readily located at the angle of the V, and the TSR, LDLr, and EGF-like domains would be placed in the order predicted from the C9 sequence at the ends of the arms.

The initial structural modeling of C6, C7, and C8 (Figure 6c) suggests that C6, C7, C8, and C9 possess similar structural morphologies of lengths between 11 and 18 nm. The additional length of 7 nm for C6 and C7 when compared to C9 can be attributed to the presence of two SCRs and two FIMs at the C-terminus of these components. Each SCR is of length 2.6–4.6 nm by various estimates (Perkins et al., 1991a); the lengthwise disposition of two adjacent SCRs as found in C1r, C1s, C4bp, and factor H will account for the increased length of C6 and C7 relative to C9. This means that, even though the length of the FIM is unknown, the two FIMs of C6 asnd C7 may be located in a compact conformation relative to the two SRCs within these proteins such that the overall length of C6 and C7 is not increased further.

In electron micrographs, a crevice on one face of C9 has been reported by DiScipio and Hugli (1985), even though these authors were not able to detect a bilobal V-shaped structure for C9. This would be consistent with the C9 model shown in Figure 6a. Only an elliptical cylinder and an V-shaped structure were considered in this study, where the latter offers obvious possibilities for a large conformational change. Some experimental evidence for a V-shaped structure was found in the analyses of Figures 3 and 5. Even though the scattering analyses were constrained by the volume calculated from the primary structure and were compatible with hydrodynamic modeling, it should be noted that other scattering-equivalent models, including more complex ones suggested by electron microscopy studies of C6 and C7, may offer equally satisfactory curve fits (e.g., V-shaped models with two arms of different lengths). Many other possible models would be eliminated by the size constraints inherent in both the known sequence and in the scattering curves themselves. Relatively few parameters are uniquely determined from solution scattering data, and this is due to the fact that the data are spherically averaged because of the random orientations of the scattering macromolecules.

Such V-shaped models for C9 facilitate discussion of the unfolding of C9 when this is polymerized or inserted into membranes (Stanley et al., 1986). The significance of the C9 structure is shown by comparisons with the sizes of other proteins with which it interacts, where C5 with a length of 18 nm (Perkins et al., 1990a) is almost twice the length of C9 at 11 nm (Figure 6a). By analogy with structural studies on the formation of C3b from C3 and C4b from C4, C5b will be of similar length to C5, which is 18 nm long (Perkins & Sim, 1986; Perkins et al., 1990b). This length is coincident with the overall length of C6 and C7. It is interesting to note from this study that C5b may act as an extended template for complex formation with C6. The length of C5b is sufficient that part of it may interact with the two SCR and FIM domains of length 7 nm at the C-terminus of C6 (DiScipio & Hugli, 1989; Haefliger et al., 1989) to leave a considerable segment of C5b which may interact with the C9-like segment of C6 in order to facilitate the unfolding of C6. The unfolded C6 structure would in turn be able to unfold sequentially the homologous proteins C7, C8, and *n* copies of C9, in order to lead to MAC formation and insertion into membranes. Electron microscopy data on complexes formed between C5b6 and C7 have been discussed by DiScipio et al. (1988), although these authors propose that the final form of these complexes possess highly extended structures. Solution scattering data on C6, C7, and C8 and their complexes with C5b in conditions close to physiological will extend structural knowledge of these complexes.

Neutron scattering probed the internal structure of C9 and interestingly showed that this had a radial distribution of hydrophilic and hydrophobic residues that is typical of soluble glycoproteins. This accounts for its solubility in aqueous media, i.e., the three cysteine-rich TSR, LDLr, and EGF-like domains are all in contact with the aqueous phase. This internal structure is similar to C5 (Perkins et al., 1990a), which also participates in the MAC after activation to C5b. The good curve fits for C9 in Figure 5 were obtained with the distribution of hydrophilic and hydrophobic residues as shown in Figure 6b. The implication of this distribution is that the unfolding of the C9 structure would expose the more hydrophobic region of this structure, thus facilitating contacts with lipids.

ACKNOWLEDGMENTS

We thank the Wellcome Trust for support, the Science and Engineering Research Council for use of the Synchrotron Research Source (Daresbury, U.K.), the Institut Laue Langevin (Grenoble, France), and Professor P. J. Lachmann, F. R. S., for his interest in this project. Mr. A. S. Nealis, Mr. D. Seilly, Dr. W. Bras, and Dr. A. de Geyer are gratefully thanked for support and assistance. This work was presented at the David Phillips Symposium on Protein Structure and Function, Oxford (July 1–3, 1991) and the XIVth International Complement Workshop, Cambridge (Sept 15–20, 1991) (Smith et al., 1991b).

Registry No. C9, 94949-12-9.

REFERENCES

Bhakdi, S., & Tranum-Jensen J. (1984) *J. Immunol. 136*, 2999–3005.

Biesecker, G., & Müller-Eberhard, H. J. (1980) *J. Immunol. 124*, 1291–1296.

Biesecker, G., Gerard, C., & Hugli, T. E. (1982) *J. Biol. Chem. 261*, 2584–2590.

Blake, C. C. F., Geisow, M. J., Oatley, S. J., Rerat, B., & Rerat, C. (1978) *J. Mol. Biol. 121*, 339–356.

Borsos, T., Dourmashkin, R. R., & Humphrey, J. H. (1964) *Nature (London) 202*, 251–252.

Chakravarti, D. N., Chakravarti, B., Parra, C. A., & Müller-Eberhard, H. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2799–2803.

Chothia, C. (1975) *Nature (London) 254*, 304–308.

Cooke, R. M., Wilkinson, A. J., Baron, M., Pastore, A., Tappin, M. J., Campbell, I. D., Gregory, H., & Sheard, B. (1987) *Nature (London) 327*, 339–341.

Cusack, S. (1981) *J. Mol. Biol. 145*, 539–541.

Dankert, J. R., & Esser, A. F. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2128–2132.

DiScipio, R. G., & Gagnon, J. (1982) *Mol. Immunol. 19*, 1425–1431.

DiScipio, R. G., & Hugli, T. E. (1985) *J. Biol. Chem. 260*, 14802–14809.

DiScipio, R. G., & Hugli, T. E. (1989) *J. Biol. Chem. 264*, 16197–16206.

DiScipio, R. G., Gehring, M. R., Podack, E. R., Kan, C. C., Hugli, T. E., & Fey, G. H. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7298–7302.

DiScipio, R. G., Chakravarti, D. N., Müller-Eberhard, H. J., & Fey, G. H. (1988) *J. Biol. Chem. 263*, 549–560.

Esser, A. F., Kolb, W. P., Podack, E. R., & Müller-Eberhard, H. J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1410–1414.
Garcia de la Torre, J., & Bloomfield, V. A. (1977a) *Biopolymers 16*, 1747–1761.
Garcia de la Torre, J., & Bloomfield, V. A. (1977b) *Biopolymers 16*, 1779–1793.
Ghosh, R. E. (1989) *Internal Publication 89GH02T*, Institut Laue Langevin, Grenoble, France.
Glatter, O. (1982) in *Small-Angle X-ray Scattering* (Glatter, O., & Kratky, O., Eds.) pp 119–196, Academic Press, London.
Glatter, O., & Kratky, O., Eds. (1982) *Small-Angle X-ray Scattering*, Academic Press, New York.
Haefliger, J.-A., Tschopp, J., Vial, N., & Jenne, D. E. (1989) *J. Biol. Chem. 264*, 18041–18051.
Harrison, R. A., & Lachmann, P. J. (1986) Complement Technology, in *Handbook of Experimental Immunology* (Weir, D. M., Herzenberg, L. A., Blackwell, C., & Herzenberg, L. A., Ed.) pp 39.1–39.49, Blackwell Scientific Publications, Edinburgh.
Ibel, K., & Stuhrmann, H. B. (1975) *J. Mol. Biol. 93*, 225–266.
Jacrot, B., & Zaccai, G. (1981) *Biopolymers 20*, 2413–2426.
Kontermann, R., & Rauterberg, E. W. (1989) *Mol. Immunol. 26*, 1125–1132.
Kratky, O. (1963) *Prog. Biophys. Chem. 13*, 105–173.
Law, S. K. A., & Reid, K. B. M. (1988) *Complement*, IRL Press.
Mayer, M. M. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 2954–2958.
Monahan, J. B., Stewart, J. L., & Sodetz, J. M. (1983) *J. Biol. Chem. 258*, 5056–5062.
Perkins, S. J. (1986) *Eur. J. Biochem. 157*, 169–180.
Perkins, S. J. (1988a) *Biochem. J. 254*, 313–327.
Perkins, S. J. (1988b) *New Compr. Biochem. 18B*, 143–264.
Perkins, S. J. (1989) in *Dynamic Properties of Biomolecular Assemblies* (Harding, S. E., & Rowe, A. J., Eds.) pp 226–245, Royal Society of Chemistry, London.
Perkins, S. J., & Weiss, H. (1983) *J. Mol. Biol. 168*, 847–866.
Perkins, S. J., & Sim, R. B. (1986) *Eur. J. Biochem. 157*, 155–168.
Perkins, S. J., Kerckaert, J. P., & Loucheux-Lefebvre, M. H. (1985) *Eur. J. Biochem. 147*, 525–531.
Perkins, S. J., Chung, L. P., & Reid, K. B. M. (1986) *Biochem. J. 233*, 799–807.
Perkins, S. J., Nealis, A. S., Haris, P. I., Chapman, D., Goundis, D., & Reid, K. B. M. (1989) *Biochemistry 28*, 7176–7182.
Perkins, S. J., Smith, K. F., Nealis, A. S., Lachmann, P. J., & Harrison, R. A. (1990a) *Biochemistry 29*, 1175–1180.
Perkins, S. J., Nealis, A. S., & Sim, R. B. (1990b) *Biochemistry 29*, 1167–1175.
Perkins, S. J., Smith K. F., & Nealis, A. S. (1990c) *Biochem. Soc. Trans. 18*, 1151–1154.
Perkins, S. J., Smith, K. F., Amatayakul, S., Ashford, D., Rademacher, T. W., Dwek, R. A., Lachmann, P. J., & Harrison, R. A. (1990d) *J. Mol. Biol. 214*, 751–763.
Perkins, S. J., Nealis, A. S., & Sim, R. B. (1991a) *Biochemistry 30*, 2847–2857.
Perkins, S. J., Nealis, A. S., Sutton, B. J., & Feinstein, A. (1991b) *J. Mol. Biol. 221*, 1345–1366.
Pilz, I. (1982) in *Small-Angle X-ray Scattering* (Glatter, O., & Kratky, O., Eds.) pp 239–293, Academic Press, London.
Podack, E. R., & Tschopp, J. (1982a) *Proc. Natl. Acad. Sci. U.S.A. 79*, 574–578.
Podack, E. R., & Tschopp, J. (1982b) *J. Biol. Chem. 257*, 15204–15212.
Podack, E. R., & Tschopp, J. (1984) *Mol. Immunol. 21*, 589–603.
Podack, E. R., Kolb, W. P., & Müller-Eberhard, H. J. (1976) *J. Immunol. 116*, 263–269.
Priessner, K. T., Podack, E. R., & Müller-Eberhard, H. J. (1985) *J. Immunol. 135*, 452–458.
Reid, K. B. M. (1986) *Essays Biochem. 27*, 27–68.
Ryan, B. F., Joiner, B. L., & Ryan, T. A. (1985) *Minitab Handbook*, 2nd ed., PWS-Kent Publishing Co., Boston.
Shinkai, Y., Takio, K., & Okumura, K. (1988) *Nature (London) 334*, 525–527.
Shiver, J. W., Dankert, J. R., Donovan, J. J., & Esser, A. F. (1986) *J. Biol. Chem. 261*, 9629–9636.
Smith, K. F., Harrison, R. A., & Perkins, S. J. (1990) *Biochem. J. 267*, 203–212.
Smith, K. F., Nolan, K. F., Reid, K. B. M., & Perkins, S. J. (1991a) *Biochemistry 30*, 8000–8008.
Smith, K. F., Harrison, R. A., & Perkins, S. J. (1991b) *Complement Inflammation 8*, 225–225 (abstract).
Stanley, K. K., Kocher, H.-P., Luzio, J. P., Jackson, P., & Tschopp, J. (1985) *EMBO J. 4*, 375–382.
Stanley, K. K., Page, M., Campbell, A. K., & Luzio, J. P. (1986) *Mol. Immunol. 23*, 451–458.
Stuhrmann, H. B., & Fuess, H. (1976) *Acta Crystallogr. A32*, 67–74.
Towns-Andrews, E., Berry, A., Bordas, J., Mant, G. R., Murray, P. K., Roberts, K., Sumner, I., Worgan, J. S., Lewis, R., & Gabriel, A. (1989) *Rev. Sci. Instrum. 60*, 2346–2349.
Tschopp, J. (1984) *J. Biol. Chem. 259*, 10569–10573.
Tschopp, J., Müller-Eberhard, H. J. & Podack, E. R. (1982) *Nature (London) 298*, 534–538.
Tschopp, J., Engel, A., & Podack, E. R. (1984) *J. Biol. Chem. 259*, 1922–1928.
Young, J. D.-E., Cohn, Z. A., & Podack, E. R. (1986) *Science 233*, 184–190.