Molecular Modeling of Human Complement Component C4 and Its Fragments by X-ray and Neutron Solution Scattering†

Stephen J. Perkins,*,‡ Adam S. Nealis,‡ and Robert B. Sim§

Department of Biochemistry and Chemistry and Department of Protein and Molecular Biology, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, U.K., and MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

Received April 11, 1989; Revised Manuscript Received September 14, 1989

ABSTRACT: The solution structures of human complement component C4 and five derived fragments, C4u, C4(a+b), C4b, C4c, and C4d, were analyzed by synchrotron X-ray and neutron scattering. The X-ray radii of gyration R_G for C4, C4u, and C4(a+b) in H_2O buffers are similar at 5.23–5.28 nm, and likewise the cross-sectional radii of gyration R_{XS} are similar at 2.48–2.52 nm. Molecular mass calculations using X-rays and neutrons show unexpectedly that C4c is dimeric; however, all the other forms are monomeric. $C4c_2$ has an X-ray R_G of 5.18 nm and an R_{XS} of 2.89 nm. Neutron contrast variation gives R_G values at infinite contrast of 4.87–4.93 nm for C4 and C4u, 4.79 nm for C4b, 4.94 nm for $C4c_2$, and 2.69 nm for C4d. The R_{XS} values at infinite contrast are 2.23–2.25 nm for C4 and C4u, 1.89 nm for C4b, and 2.62 nm for $C4c_2$. These data show that a large conformational change occurs on going from C4 to C4b, but not on going from C4 to C4u, and this is attributed to the presence of the C4a moiety in C4u. Comparisons of the C4 and C4u scattering curves show that these are very similar out to a nominal resolution of 4 nm. Scattering-curve models were developed to account for the neutron scattering curves of C4, $C4c_2$, and C4d in ^2H_2O buffers. The C4c monomer could be represented by a lamellar ellipsoid of size 8 nm × 2 nm × 18 nm. C4d was found to be 4 nm × 2 nm × 9 nm. The combination of these structures gave good accounts of the neutron data for C4, C4b, and $C4c_2$ to resolutions of 5–6 nm. The C4 model was obtained by placing the long axis of C4d parallel to that of C4c such that the cross section is extended. C4b was best modeled by repositioning C4d relative to C4c such that this cross section becomes more compact. The C4 and C4b models are compared with possible structures for the C1 component of complement to show the importance of the surface accessibility of the protease domains and short consensus repeat domains in C1 for C4 activation.

C4 is the second most abundant component of the complement cascade of immune defense and exists in a number of polymorphic forms. It is composed of three disulfide-linked polypeptide chains α, β, and γ (Figure 1); it has four sites for attachment of N-linked oligosaccharides (Belt et al., 1984; Chan & Atkinson, 1985) and a molecular mass (M_r) of 197000. It is involved in the generation of the C3 convertase of the classical pathway in a three-stage process (Reid, 1983, 1986). Initially $C\bar{1}s$ in activated C1 cleaves C4 into the minor fragment C4a and the major fragment C4b (Figure 1; Table I). A thiolester bond in C4b is thereby exposed, which will bind covalently to receptive surfaces or react with water to form fluid-phase C4b. The second and third stages involve the binding of C2 to C4b in the presence of Mg^{2+} and then the reaction of activated C1 to remove the noncatalytic domain C2b from C4b·C2 to leave the C4b·C2a complex, which is the C3 convertase. The C3 convertase is regulated by the enzyme factor I, which employs the C4b-binding protein (C4BP) as cofactor. When C4b binds to C4BP, C4b is cleaved by factor I into the inactive forms C4c and C4d (Figure 1). Another form of C4 is generated by chaotrope or methylamine treatment of C4 to give C4u [sometimes named C4(H_2O) or C4i], in which the thiolester bond at the reactive site is cleaved. C3, the most abundant complement component, is highly homologous to C4 and has similar forms and fragments, termed C3u, C3a, C3b, C3c, and C3dg.

Structural studies of C4 and its fragments are important for a functional understanding of the physiological roles of C4 and its subunits. To date, such work has been based on electron microscopy, sedimentation studies, and solution scattering. Electron microscopy on C4b showed a relatively compact structure of approximate dimensions 9 nm × 17 nm (Dahlbäck et al., 1983). These, however, differ from the results of another study on C4 and C4b, and also C3 and C3b, where dimensions of 12.5 nm × 7.5 nm × 6.5 nm have been reported (Smith et al., 1984). Sedimentation coefficients $s^0_{20,w}$ show that C3 is similar to C4 and likewise C3b to C4b (Müller-Eberhard et al., 1966; Pâques, 1980; Reboul et al., 1980). However, the $s^0_{20,w}$ of C3c is different from that of C4c, and no explanation of this has been offered (Pâques, 1980). Preliminary scattering work has suggested that C3 and C4, while similar to one another, are conformationally different from C3u and C4u (Österberg et al., 1984, 1988). A more detailed study suggested, however, that C3 and C3u are of very similar structure (Perkins & Sim, 1986). These observations thus merit further investigation.

A systematic study of C4 and its fragments was therefore undertaken to define their overall shapes, the existence of conformational changes on C4 activation and inactivation, and the arrangement of C4c and C4d within C4b. Solution scattering is a powerful method of choice for C4 (Perkins, 1988a,b). This is based on solution measurements in conditions close to physiological, and as such has an advantage over electron microscopy, which can be prone to artifacts caused

† The Wellcome Trust and the Science and Engineering Research Council are gratefully thanked for support. A preliminary account of this work was given in the 1987/1988 Annual Report of the SRS at Daresbury.

* To whom correspondence should be addressed at the Department of Biochemistry and Chemistry.

‡ Royal Free Hospital School of Medicine.

§ University of Oxford.

0006-2960/90/0429-1167$02.50/0

Table I: Composition Data for Human Complement Component C4 and Its Fragments

	C4, C4u, C4(a+b)	C4b	C4c	C4d	C4a
amino acid residues					
total	1758	1634	1254	380	77
M_r	188200	179500	138900	40500	8700
volume (nm^3)	243	232	180	52	11
carbohydrate residues					
total	47	47	35	12	0
total volume[a] (nm^3)	253	241	187	55	11
$\bar{v}$[a] (mL/g)	0.735	0.736	0.737	0.731	0.724
$\sum b/M_r$ in H_2O (fm)	0.2276	0.2277	0.2267	0.2310	0.2243
A_{280}[b] (1%, 1 cm)	9.01	9.31	8.03	13.63	2.64
total M_r					
from sequence	197100	188400	145500	42900	8700
neutron data[c] C4	217000 (1)	172000 (2)	262000 (1)	43000 (1)	
M_r ratio					
from sequence	1.000	0.956	0.738	0.218	0.044
X-ray data[d]	1.00		1.53 (4)		
	1.00		1.63 (3)		
match point (% ^2H_2O)					
calculated[a]	40.7	40.6	40.5	41.1	43.1
neutron data C4	41.2 (7)	41.1 (15)	41.5 (6)	41.8 (4)	
	42.1	41.6	41.7		
neutron data C4u	40.5 (12)				
	41.1				

[a] The dry volume for match-point predictions and modeling calculations is from Chothia (1975) and Perkins (1986). Match points are calculated by using the assumption of the nonexchange of 10% of the main-chain peptide hydrogens (Perkins, 1986). Partial specific volumes correspond to hydrated volumes. These were obtained from the consensus volumes of Perkins (1986) and are close to the classical values of Cohn and Edsall (1943). [b] Calculated from the content of Tyr, Cys, and Trp residues in the sequence by the Wetlaufer (1962) procedure and rescaled by using a factor of 1.03 (Perkins, 1986). [c] The number of data points used in the experimental determinations of M_r and in four contrasts for the match points are in parentheses. [d] Two separate measurement sessions were employed; data were normalized to the M_r of C4, C4u, and C4(a+b) as unity.

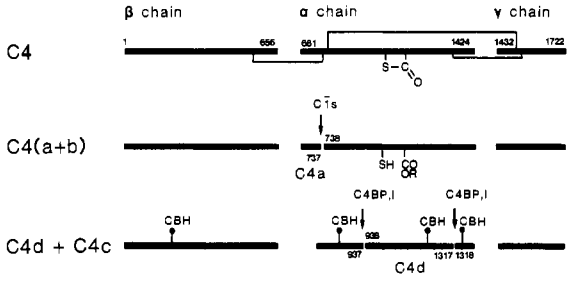

FIGURE 1: Summary of the primary structure of C4. The relationship of C4a, C4b, C4c, and C4d to the parent molecule C4 is indicated in terms of the polypeptide sequence (Belt et al., 1984). The inactive form of C4 is C4u [sometimes called C4(H_2O) or C4i] where the internal thiolester is hydrolyzed. Activated $C\bar{1}s$ is responsible for the cleavage of C4 into C4a and C4b. C4d is formed by the interaction with C4BP and factor I. Of the six potential carbohydrate attachment sites (CBH), one on the β chain and three on the α chain are occupied (Chan & Atkinson, 1985). The disulfide bridges are shown by thin lines only in C4 for clarity.

by beam damage, the need to work in vacuo, and the use of stains. Scattering enables molecular masses to be measured as a control. Conformational changes in solution can be detected in some circumstances. Furthermore, the modeling of the scattering curves can be used for quantitative comparisons with results from other structural techniques. Even though a unique structure determination cannot be made (since the macromolecules are viewed in random orientations), the analyses are much constrained by these independent data and are able to rule out structural models that are incompatible with the scattering data. The experiments show that the structure of C4 can be described in terms of two major domains.

Materials and Methods

(a) Protein Purification and Preparations of C4 Fragments. The materials used in the protein preparations are those as specified in Perkins and Sim (1986). Hemolytically active C4 was prepared from fresh human serum as described by Campbell et al. (1980). Some of the C4 used was made by other methods [Law et al., 1983; Sim et al., 1986; the L001 monoclonal antibody affinity method of Hsiung et al. (1987)] and was kindly donated by Dr. S. K. A. Law, Dr. E. Sim, and Dr. A. W. Dodds, respectively. No significant differences were seen in the C4 prepared by different methods, except that the material prepared by monoclonal antibody affinity chromatography was unsuitable for prolonged manipulation at ambient temperature (15–20 °C), as it became proteolyzed. This is likely to be due to coelution of traces of plasminogen/plasmin from the antibody column. $C\bar{1}s$, C4BP, and factor I were prepared from human serum or plasma by established methods (Sim, 1981; Sim & Sim, 1983; Hsiung et al., 1982; in that order).

The inactive form of C4, C4u, in which the thiolester is hydrolyzed, was prepared by incubating C4 in 100 mM methylamine (Law et al., 1983) at pH 9.0 for 1 h at 37 °C. Some C4u was kindly donated by Dr. L. P. Chung and Dr. K. B. M. Reid (Perkins et al., 1986). Hemolytically active C4 was cleaved to form the fragments C4a and C4b by digestion with $C\bar{1}s$. C4 (0.5–1 mg/mL) in 10 mM potassium phosphate, 140 mM NaCl, and 1 mM EDTA, pH 7.4, was incubated with $C\bar{1}s$ at a molar ratio of $C\bar{1}s$ to C4 of 1:100 for 1 h at 37 °C. The reaction was stopped by the addition of 0.5 mM diisopropyl fluorophosphate. After cleavage, C4a and C4b remain noncovalently associated, and this material is referred to as C4(a+b). C4b was separated from C4a by gel filtration on a column (90 cm × 1.5 cm diameter) of Ultrogel AcA 34 in 12.4 mM sodium phosphate, 2 M NaCl, and 0.5 mM EDTA, pH 7.0 (Tack et al., 1979).

C4b was digested with factor I in the presence of C4BP to produce C4c and C4d. C4b (0.2–0.7 mg/mL) in 12.4 mM sodium phosphate, 0.5 mM EDTA, and 0.5 mM diisopropyl fluorophosphate at pH 7.0 was incubated with C4BP (1.4%

w/w) and factor I (0.05% w/w) for 7 h at 37 °C. C4c and C4d were separated from each other by chromatography on Matrix Gel Blue A, as previously described for C3d and C3c (Perkins & Sim, 1986).

(*b*) *Sample Preparation for Scattering Experiments and Analyses.* As described for C3 and its fragments (Perkins & Sim, 1986), C4, C4b, C4(a+b), C4u, C4c, and C4d were initially concentrated by ion exchange on DEAE-Sephacel and then gel-filtered on a Sephacryl S300/Ultrogel AcA 34 column equilibrated in 12.4 mM sodium phosphate, 200 mM NaCl, and 0.5 mM EDTA, pH 7.0. The gel filtration step removes contamination by covalent or noncovalent protein aggregates and eliminates traces of C4BP, factor I, or $C\bar{1}s$. A final concentration step was done by using B15 Minicon concentrators. Sample volume reduction using B15 Minicons did not cause detectable aggregation or protein loss up to concentrations of about 6 mg/mL for C4, C4u, C4(a+b), and C4c. However, C4d tended to bind to the membrane and could not be concentrated above 2.5 mg/mL by this method. Protein samples were examined before and/or after scattering experiments by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (Laemmli, 1970). Samples were prepared for electrophoresis, with and without reduction of disulfide bridges, as described by Fairbanks et al. (1971). C4 activity was determined by hemolytic assay (Gigli et al., 1977), and the internal thiolester of C4 was assayed by the autolytic cleavage reaction (Sim & Sim, 1981).

(*c*) *X-ray and Neutron Data Collection.* Synchrotron X-ray scattering experiments were carried out on the solution camera at station 7.3 (Nave et al., 1985) at the SRS Daresbury, U.K. The synchrotron operated at 1.8 GeV with currents between 75 and 300 mA (data collection times of 6–10 min) or at 2.0 GeV with currents between 190 and 200 mA (data collection times of 3 min). Most data were collected by using a sample-to-detector distance of 2.180–2.210 m and a wavelength of 0.1608 nm in five sessions. A 500-channel linear detector gave a Q range of 0.1–1.3 nm^{-1} (where $Q = 4\pi \sin \theta/\lambda$ and 2θ is the scattering angle). The detector response to normalize spectral intensities was measured by using a uniform radioactive source. The Q range was calibrated by using wet, slightly stretched rat-tail collagen (spacing 67 nm), dry white adductor muscle from the North American clam *Mercenaria mercenaria* (spacing 14.4 nm), and the position of the attenuated main beam. Background subtraction errors were minimized by measuring the samples alternately with the buffers, using the same cell (mica windows of 10–15-μm thickness), and counting samples and buffers for equal times. Data were reduced by using the program SWANAL.

Neutron scattering experiments were carried out on Instrument D17 at the Institut-Laue-Langevin, Grenoble. Guinier data were collected by using a sample-to-detector distance of 3.46 m and wavelengths between 1.30 and 1.60 nm. Curves at larger Q were measured with a sample-to-detector distance of 1.40 m, detector–main beam angles of 0° and 20°, and wavelengths of 0.99–1.00 nm. The Q ranges are therefore 0.022–0.526 and 0.99–1.61 nm^{-1} and out to $Q = 0.36$ nm^{-1} in that order. Data reduction in Grenoble using the programs RNILS, SPOLLY, and RGUIM followed standard procedures (Ghosh, 1981). A cadmium background is first subtracted from each run. Next, the buffer background is subtracted from that of the sample, and the result is normalized by using a water run from which an empty cell background has been subtracted.

Buffers for the X-ray and neutron samples were 200 mM NaCl, 12.4 mM sodium phosphate, and 0.5 mM EDTA, pH 7.0. Some data on C4u were measured in 250 mM NaCl, 50 mM 6-aminohexanoic acid, 5 mM EDTA, and 50 mM Tris-HCl at pH 7.4, with no observable differences (Perkins et al., 1986). Neutron samples in 0%, 70%, 80%, and 100% ^2H_2O buffers were dialyzed with four changes over a 24–48-h period at 6 °C. X-ray data were recorded at 6–10 °C. Neutron data were recorded at 20 °C; some on C4u measured at 10 °C showed no differences. Sample concentrations were derived from the absorption measured at 280 nm in 1 (0% ^2H_2O) or 2 mm (70%, 80%, and 100% ^2H_2O) thick quartz Hellma cells, which in the case of neutron scattering were the actual cells used in data collection. The A_{280} coefficient for C4 has been reported as 8.3 or 10.0 (Nagasawa & Stroud, 1977; Reboul et al., 1980; Isenman & Kells, 1982). In this study, A_{280} coefficients were calculated from sequence data (Table I).

(*d*) *Experimental Data Analysis.* Guinier analyses of the scattering curves at low Q give the radius of gyration R_G and the forward scattered intensity $I(0)$ (Glatter & Kratky, 1982):

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

R_G values measure the degree of particle elongation. Molecular masses can be deduced from $I(0)/c$ values (c is the sample concentration in milligrams per milliliter) either as relative values from the X-ray data or as absolute values from neutron data by referencing $I(0)$ to the incoherent scattering of water as a standard (Kratky, 1963; Jacrot & Zaccai, 1981). Neutron data collection in H_2O buffers minimizes systematic errors in M_r arising from ^1H–^2H exchange or variation in the partial specific volume $\bar{v}$. Sample concentrations c were obtained from optical density measurements at 280 nm. For elongated macromolecules, the corresponding cross-sectional radius of gyration R_{XS} and the cross-sectional intensity at zero angle $I(Q)Q_{Q\to 0}$ are obtained from curve analyses in a Q range larger than that above from

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

The length L of the macromolecule can be calculated by combining the neutron R_G and R_{XS} data if the macromolecular shape corresponds to an elliptical cylinder:

$$L^2 = 12(R_G^2 - R_{XS}^2)$$

Another determination of L is available from the intensities $I(Q)$:

$$L = \pi I(0)/[I(Q)Q]_{Q\to 0}$$

Stuhrmann analyses of the R_G^2 and R_{XS}^2 values as a function of the reciprocal solute–solvent contrast $\Delta\rho^{-1}$ (Ibel & Stuhrmann, 1975) report on the internal structure of the macromolecule, where the simplified expressions (Perkins, 1988a,b)

$$R_G^2 = R_{G\text{-}C}^2 + \alpha_G\Delta\rho^{-1}$$

$$R_{XS}^2 = R_{XS\text{-}C}^2 + \alpha_{XS}\Delta\rho^{-1}$$

where $R_{G\text{-}C}$ and $R_{XS\text{-}C}$ are the radii of gyration at infinite contrast and the terms α_G and α_{XS} are a measure of the radial inhomogeneity of scattering density can be used. Use of the Parallel Axes theorem gives the separation Δ between two subunits 1 and 2 once these have been incorporated in a complex 3

$$\Delta^2 = (R_{G3}^2 - f_1R_{G1}^2 - f_2R_{G2}^2)/f_1f_2$$

where f_1 and f_2 are the volume fractions of subunits 1 and 2 in the complex 3, with the corresponding R_G values denoted by the subscripts. This assumes that no detectable confor-

mational change has occurred in subunits 1 or 2 when complex 3 is formed. For consistency, the final Guinier analyses of X-ray and neutron data were processed in London by using an interactive program SCTPL2. Stuhrmann plots were analyzed by using a weighted two-parameter least-squares procedure to determine the R_C and α values.

(*e*) *Scattering Curve Simulations.* Scattering curve modeling was carried out by using overlapping Debye spheres to follow standard procedures (Glatter & Kratky, 1982) in application to synchrotron X-ray and neutron scattering data (Perkins & Weiss, 1983; Perkins, 1985; Perkins & Sim, 1986). The total volumes (Table I) of C4, C4b, C4c, and C4d were calculated from sequences (Belt et al., 1984; Chan & Atkinson, 1985) by using the crystallographic dry volumes for amino acids and carbohydrates (Chothia, 1975; Perkins, 1986). Neutron data and their modeling correspond largely to the unhydrated glycoprotein volume, since the hydration shell that surrounds the macromolecule is invisible by the usual neutron contrast variation techniques (Perkins, 1986). This means that partial specific volumes $\bar{v}$ are larger if they are obtained from neutron match points instead of by densitometry. The final Debye models were generated by using spheres of this total dry volume, with the sphere volume set to be that of the cube of side 0.8 nm, as previously in Perkins and Sim (1986). For curve fitting, neutron curves in 100% ^2H_2O were used, since these had better counting statistics, correspond to the dry particle volume (which is well-defined from the sequence), and are closer to the infinite contrast condition (at a high negative solute–solvent contrast) than the X-ray curves (see Figure 3). The simulated neutron curves were corrected for wavelength spread and beam divergence (Cusack, 1981; Perkins & Weiss, 1983). The difference between negative and positive contrast conditions for C4 was tested and not found to be significant. Where required for X-ray curve simulations, hydrated Debye sphere models were used. These are based on 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).

Results and Discussion

(*a*) *Synchrotron X-ray Scattering Curves.* X-ray scattering curves for C4, C4u, C4(a+b), and C4c were obtained in the concentration range 2.0–5.6 mg/mL. Scattering curves measured by synchrotron radiation were prone to small buffer subtraction errors. Curves were accepted on the criteria of linear overall R_G and cross-sectional R_{XS} Guinier plots and their reproducibility in repeated measurements. Cross-sectional R_{XS} analyses can be carried out for particles whose diameter-to-length ratio is as low as 1:2 (Glatter & Kratky, 1982), as exemplified in several studies of immunoglobulins. Representative X-ray Guinier R_G analyses in Figure 2a for C4 and C4u show very similar R_G values. The plots are linear to a good approximation out to a QR_G of 1.5. Comparison of these R_G values with those measured to a QR_G of only 1.0 showed good agreements, even though the statistical precision is correspondingly reduced. Likewise, the analyses for C4, C4u, and C4(a+b) in Figure 2b show that these gave very similar R_{XS} values. In graphs of appearance similar to those of Figures 3 and 5 in Perkins and Sim (1986) for C3, no concentration dependence of the R_G and R_{XS} data was observed. The mean R_G values of 5.23–5.29 nm and R_{XS} values of 2.47–2.52 nm are summarized in Table II.

Comparisons of the C4, C4u, and C4(a+b) data of Table II with those for the homologues C3, C3u, and C3(a+b) in Table II of Perkins and Sim (1986) show that the R_G and R_{XS} values are in good agreement. However the X-ray scattering

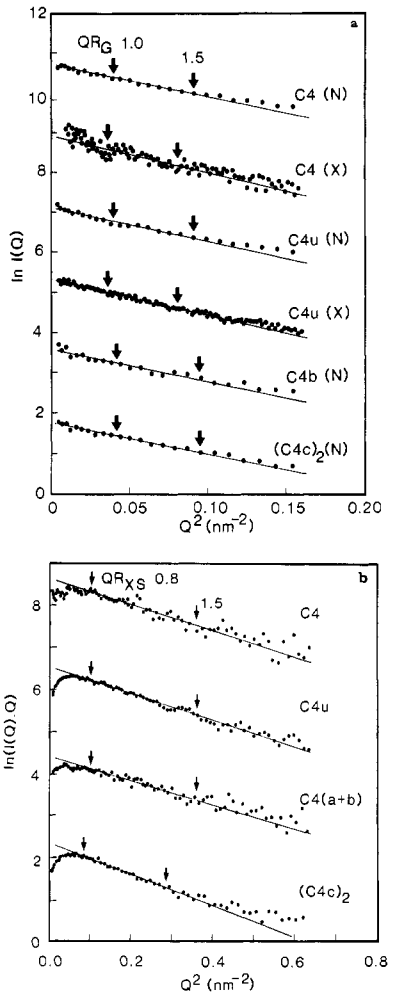

FIGURE 2: Guinier R_G and R_{XS} analyses of C4 and its fragments. Data were obtained by neutron scattering (N) or X-ray scattering (X). (a) The QR_G values of 1.0 and 1.5 are arrowed in each spectrum. These are based on an R_G of 5.2 nm for C4 and C4u by X-rays and of 5.0 nm for C4 and C4u by neutrons in ^2H_2O buffers. The neutron R_G for C4b is 4.9 nm; the X-ray R_G for $(C4c)_2$ is 5.2 nm. (b) The cross-sectional analyses using X-rays are shown for C4, C4u, and C4(a+b), where the QR_{XS} values of 0.8 and 1.5 correspond to an R_{XS} of 2.5 nm, and for $(C4c)_2$ where the R_{XS} is 2.8 nm. In (a) and (b), the plots are displaced arbitrarily for clarity.

data for C4c give a mean R_G of 5.18 nm and R_{XS} of 2.89 nm, which differ from the R_G of 4.7 nm and R_{XS} of 2.0 nm for C3c. Calculation of the relative molecular masses from $I(0)/c$ data gave values of 1.53–1.63 for C4c in place of the expected value of 0.738 predicted from the sequence (Table I). This shows that C4c is dimeric in solution, unlike C3c, and it is now denoted $C4c_2$. Analysis of nonreduced samples on SDS-PAGE demonstrated that the C4c monomers were not covalently associated.

The degree of particle elongation can be calculated from the elongation ratio R_G/R_0, where R_0 is the R_G of the sphere having the same volume as the macromolecule in question (Kratky, 1963). For C4, C4u, and C4b, similar values in the range 1.72–1.77 are obtained. These are significantly larger than the average value of 1.28 ± 0.10 for 47 globular proteins whose M_r values ranged between 10^3 and 10^6 (Perkins, 1988a). C4 is thus significantly elongated in comparison to more conventional globular proteins. For $C4c_2$, a ratio of 1.50 is calculated, which is less than that of 1.73 for C3c, showing that the dimer is more compact in structure than the monomer.

(*b*) *Neutron Scattering Curves: I(0) and* $[I(Q)Q]_{Q\to 0}$ *Analyses.* To verify the synchrotron X-ray data, neutron data were recorded for C4, C4u, C4b, $C4c_2$, and C4d. The back-

Table II: Summary of Shape Data for C4 and Its Fragments

radii of gyration (nm)	C4	C4u	C4(a+b)	C4b	$C4c_2$	C4d
X-ray data						
R_G[a]	5.23 (4)	5.28 (6)	5.29 (1)		5.18 (8)	
R_{XS}	2.48	2.52	2.47		2.89	
neutron data						
R_G	4.93 ± 0.04 (7)	4.87 ± 0.03 (12)		4.79 ± 0.03 (15)	4.94 ± 0.03 (6)	2.69 ± 0.01 (4)
R_{XS}	2.23 ± 0.07	2.25 ± 0.03		1.89 ± 0.03	2.62 ± 0.05	1.13[b]
α_G ($\times 10^{-5}$)	2 ± 10	2 ± 8		15 ± 8	8 ± 9	−3 ± 8
α_{XS} ($\times 10^{-5}$)	18 ± 9	30 ± 4		15 ± 4	13 ± 9	
computer models						
R_G	4.97	4.97		4.87, 4.90[c]	5.05	2.83

[a] The terms in parentheses relate to the number of data points used in each determination. The neutron R_G and R_{XS} values are those measured at infinite contrast, while the X-ray data correspond to positive solute–solvent contrasts. The Guinier Q ranges for R_G fits were taken out to $Q = 0.3$ nm^{-1} and between Q values of 0.3 and 0.6 nm^{-1} for the R_{XS} fits (Perkins & Sim, 1986). In the case of C4d, the Q range extended to 0.5 nm^{-1} for the R_G value and was between 0.5 and 1.4 nm^{-1} for the R_{XS} value. [b] Only one value in 100% ^2H_2O was measured. [c] These R_G values correspond to models I and II of Figure 5b, respectively.

ground levels were notably more reproducible from run to run, and there is now no risk of radiation damage. In addition, contrast variation data can be recorded. These examine the internal structure (as well as its external form) in terms of the arrangement of carbohydrates and hydrophilic amino acid residues (weighted average scattering density 52% ^2H_2O) relative to the hydrophobic amino acid residues (average scattering density 33% ^2H_2O). Sample concentrations ranged between 1.9 and 4.9 mg/mL. Contrast variation was achieved by using 0%, 70%, 80%, and 100% ^2H_2O buffers. Good neutron Guinier plots for C4, C4u, C4b, and $C4c_2$ in 100% ^2H_2O buffers are shown in Figure 2a. No differences were seen between the R_G values of C4 and C4u. No concentration dependence of the R_G or $I(0)/c$ values were observed (data not shown).

Comparison of the neutron absolute M_r calculations from $I(0)/c$ data in H_2O buffers with the sequence values show good agreement to within 10% if $C4c_2$ is dimeric (Table I), in confirmation of the X-ray calculations. Match points for C4 and its fragments were determined where the values of $[I(0)]^{1/2}$ or $[I(Q)Q]_{Q\to 0}^{1/2}$ (after normalization for concentration, sample thickness, and sample transmission) were plotted as a function of percentage ^2H_2O. The data for C4 and C4u were combined since these were indistinguishable. In all cases, excellent linear relationships were found, similar to those for C3 in Figure 6 of Perkins and Sim (1986). By linear regression analyses, the match points at zero intensities were 40.5–41.8% ^2H_2O for the $I(0)$ analyses and 41.1–42.1% ^2H_2O for the $I(Q)Q_{Q\to 0}$ analyses (Table I). These agree with theoretical match points of 40.5–41.1% ^2H_2O from sequence data based on the dry residue volumes of amino acids and carbohydrates (Table I). In conclusion, the M_r and match-point calculations show that monodisperse solutions of C4 and its fragments have been characterized by neutron scattering.

(c) *Neutron Scattering Curves: R_G and R_{XS} Analyses.* Structural data were obtained from the Stuhrmann plots of R_G^2 and R_{XS}^2 against the reciprocal solute–solvent contrast difference $\Delta\rho^{-1}$. These give R_C, the R_G or R_{XS} at infinite contrast, and α, which monitors the internal structure. The results of weighted linear least-squares fitting to the data of Figure 3 are summarized in Table II. As for the X-ray data, the $R_{G\text{-}C}$ and $R_{XS\text{-}C}$ values for C4 and C4u are closely similar. While the $R_{G\text{-}C}$ value for C4b is similar to that for C4 and C4u, the $R_{XS\text{-}C}$ value is reduced from 2.23–2.25 to 1.89 nm. This difference has also been noted for C3b (Perkins & Sim, 1986) and was attributed to a conformational change in C3b relative to C3. The $R_{G\text{-}C}$ and $R_{XS\text{-}C}$ data for $C4c_2$ give 4.94 and 2.62 nm, respectively, which differ from those of 4.76 and 2.02 nm for C3c (Perkins & Sim, 1986).

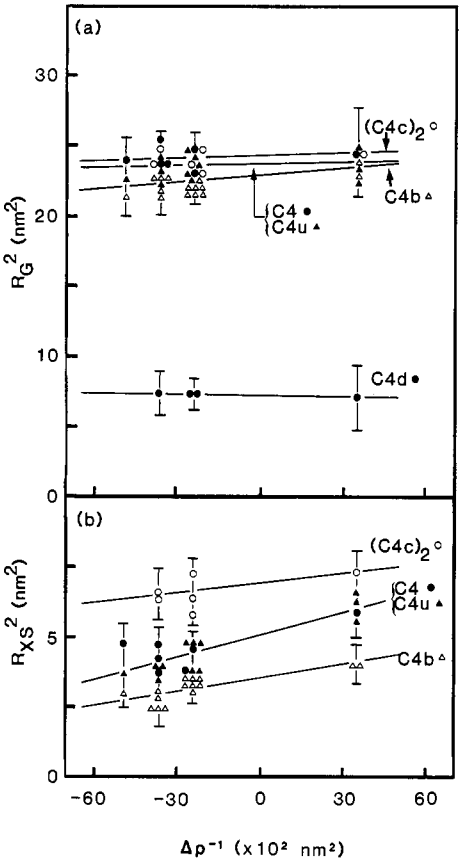

FIGURE 3: Neutron Stuhrmann plots of R_G^2 and R_{XS}^2 against the inverse solute–solvent contrast difference $\Delta\rho^{-1}$. Symbols: C4, ●; C4u, ▲; C4b, △; $C4c_2$, ○; C4d, ●. (a) The R_G^2 values for C4, C4u, $C4c_2$, C4b, and C4d are shown. Those for C4 and C4u are combined. The four straight-line fits were derived from a weighted least-squares procedure, and the R_C and α values are summarized in Table II. The averaged line for C4 and C4u has an R_C of 4.89 ± 0.04 nm and a slope α of (2 ± 10) × 10^{-5}. (b) The R_{XS}^2 values for C4, C4u, $C4c_2$, and C4b are shown. Those for C4 and C4u were combined to give an averaged line of R_C of 2.24 ± 0.05 nm and a slope α of (26 ± 5) × 10^{-5}.

The Stuhrmann α values are either zero or positive within error, although these are not as accurate as the R_C determinations. Table II shows that α_G ranges between -3×10^{-5} and 15×10^{-5} and α_{XS} ranges between 13×10^{-5} and 30×10^{-5}. The positive α values show that the hydrophilic residues lies closer to the glycoprotein surface than the hydrophobic ones, which is as expected from the solubility of C4 and its fragments in water. The X-ray R_G and R_{XS} values (Table II) are larger than the neutron R_C values. This confirms the positive α

values, since X-ray data are obtained in positive solute–solvent contrasts, which correspond to neutron data obtained in H_2O buffers (Figure 3).

In independent investigations, the limited scattering data of Österberg et al. (1984, 1988) suggested that C4 (R_G of 4.25 nm by X-rays; $R_{G\text{-}C}$ of 4.6 nm by neutrons) and C4u (R_G of 5.2 nm; $R_{G\text{-}C}$ of 5.3 nm) have different structures. While good agreement within error is found for C4u, this is not so for C4. In addition, an abnormal negative value of -100×10^{-5} was reported for the Stuhrmann α of C4. Such an α can be simulated only if C4 has a hydrophilic central core and a hydrophobic surface, which is implausible. The most likely explanation of the differences with the present data is the limited number of scattering curves used [e.g., Figure 2 of Österberg et al. (1988) vs Figure 3 of the present study] and possibly the much longer counting times that were employed for the C4 samples since C4 is labile (Tack et al., 1981).

(*d*) *Domain Structure of C4.* The neutron Guinier analyses lead to the length L of C4 and its fragments (Materials and Methods). The ratios of $I(0)$ to $I(Q)Q_{Q\to 0}$ give L values of 15.9, 16.3, 16.6, and 14.9 nm for C4, C4u, C4b, and $C4c_2$ in that order. The neutron $R_{G\text{-}C}$ and $R_{XS\text{-}C}$ values lead to similar L values of 15.2, 15.0, 15.3, and 14.5 nm. C4 on cleavage gives C4b, which in turn gives two large structures, C4c and C4d. Even though L values are similar for C4, C4u, and C4b, Table II shows that the structure of C4 and C4u differs the most from C4b (which is only 5% less in mass) in that its R_{XS} is reduced by 0.35 nm. The removal of C4a has thus caused the cross section of C4 and C4u to become more compact in C4b. The distances between the centers of the C4c and C4d structures within C4, C4u, and C4b can be derived from the Parallel Axes theorem if C4c and C4d do not change in shape on separation. The distance between C4c and C4d in C4 and C4u is 5.2 ± 0.6 nm if C4a is neglected and the R_G of C4c is taken as that of C3c. This is reduced by about 0.6 nm to 4.6 ± 0.4 nm in C4b. This indicates the extent of the conformational rearrangement in C4b. By the same reasoning, the distance between the centers of the C4c monomers in $C4c_2$ is 2.6 ± 0.6 nm if each monomer has an $R_{G\text{-}C}$ value of 4.76 ± 0.05 nm as observed for C3c (Perkins & Sim, 1986).

(*e*) *Molecular Modeling of C4 by Debye Simulations.* The use of Debye sphere models in curve simulations extends the interpretation of the structural data to higher resolutions of 4–5 nm. The R_G and R_{XS} analyses were based only on data out to $Q = 0.6$ nm^{-1}. However, experimental X-ray data were collected to $Q = 1.3$ nm^{-1}, while the neutron curves in 100% ^2H_2O were mostly measured out to 1.6 nm^{-1}. Previously, the homologous protein C3 was modeled as a single ellipsoid, which was then subdivided into the C3c and C3d domains (Perkins & Sim, 1986). The reverse and more flexible strategy is employed here, in which models for C4c and C4d were first developed and then combined to give C4 and C4b. This enables the assessment of the different conformations of C4 and C4b.

The modeling of C4c and $C4c_2$ is summarized in parts a and b of Figure 4. The monomer fit is based on a C4c model of 10 × 2 × 22 spheres from which the corners have been removed to give 360 spheres and a volume of 184 nm^3, which is close to the dry volume of 187 nm^3 (Table I). Its size is 8 nm × 2 nm × 18 nm, which is compatible with electron microscopy (Dahlbäck et al., 1983). Its R_G is 4.84 nm, in good agreement with the $R_{G\text{-}C}$ of 4.76 ± 0.05 nm for C3c (Perkins & Sim, 1986). The model leads to curves I of parts a and b of Figure 4, which agrees well with the neutron data for C3c in Figure 4a. The dimer was formed from the face-to-face

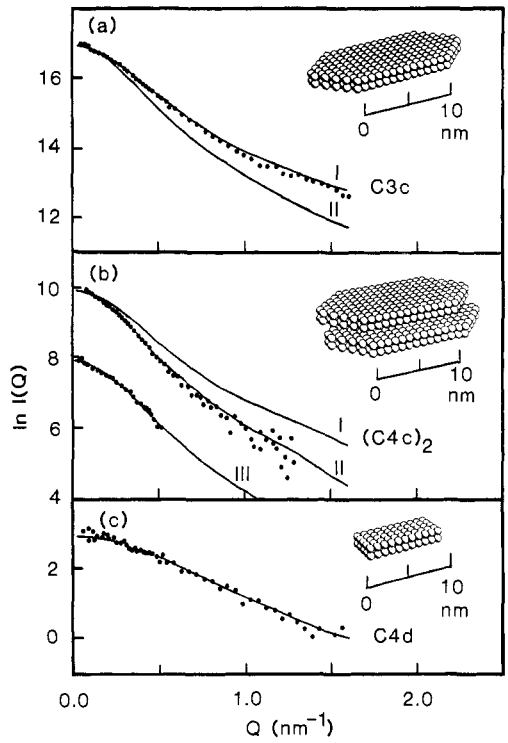

FIGURE 4: Comparisons of experimental and simulated scattering curves for the C4c and C4d fragments of C4. The neutron data were measured in 100% ^2H_2O buffers. The Debye models are shown as insets. (a) The experimental neutron curve (●) for C3c (Perkins & Sim, 1986) is compared with simulated curves for the monomer (I) and the dimer (II) of C4c to show the structural homology between C3c and C4c. (b) The experimental X-ray curve (●) for $C4c_2$ is also compared with the curve simulations for the monomer (I) and the dimer (II). Neutron data for $C4c_2$ are compared with the dimer simulation in III. (c) Experimental neutron data (●) for C4d are compared with the simulation from the Debye model as shown.

positioning of two C4c monomers in a range of models that were displaced laterally. A three-sphere displacement (2.4 nm) is in agreement with the separation of 2.6 ± 0.6 nm from the neutron R_C data and corresponds to curves II and III in parts a and b of Figure 4. Curve II gives a good account of the X-ray scattering curve for $C4c_2$ in Figure 4b and curve III likewise for a partial neutron curve III for $C4c_2$ that was available. In conclusion, the monomeric and dimeric forms of C3c and C4c can be readily distinguished. The success of the curve fits implies that the structure of C4c is closely similar to that of C3c.

The C4d model of Figure 4c is based on 5 × 2 × 12 spheres totaling 120, leads to a volume of 61 nm^3 in agreement with the sequence volume of 55 nm^3, and has an R_G of 2.83 nm in agreement with the experimental $R_{G\text{-}C}$ of 2.69 nm (Table II). This model gives the best account of the C4d neutron curve and has dimensions of 4 nm × 2 nm × 9 nm.

Models for C4 and C4u were now constructed from those for C4c and C4d (Figure 5a). The C4c model is again based on an array of 10 × 2 × 22 spheres; however, 8 spheres were added to each corner to lead to a total of 502 spheres and a volume of 257 nm^3, to agree with the total sequence volume of 253 nm^3 (Table I). This allows for the addition of C4a to C4c at an unknown location. Trial-and-error searches in which C4d was positioned onto C4c in a range of structures showed that only a limited number of structures gave good fits. The similar best fits in Figure 5a for C4 (curve I) and C4u (curve II) employed an arrangement in which the long edges of C4c and C4d are in contact with one another, and the lamella planes of C4c and C4d are perpendicular. These fits also indicate that both structures are similar, which was also shown

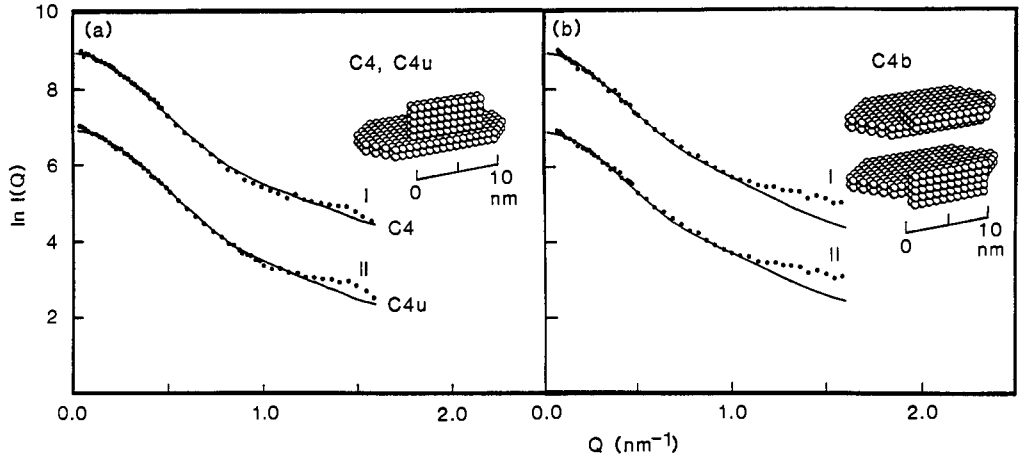

FIGURE 5: Comparisons of the neutron scattering curves for C4, C4u, and C4b with Debye models. The models are indicative of the extent of the conformational change between C4 and C4b; however, their accuracy should not be overrated. The neutron data were measured in 100% ^2H_2O buffers. (a) The C4 Debye model is constructed from joining together the individual models for C4c and C4d (Figure 4), as in the inset. The curve simulation is compared with the neutron data (●) for C4 (curve I) and C4u (curve II) to show that the experimental curve is well explained up to $Q = 1.3$ nm^{-1}. Similar curve fits were obtained for the experimental neutron curves for C3 and C3u from Perkins and Sim (1986) (not shown). (b) The same two-domain approach for C4b, starting from C4c and C4d, shows that two alternative, more compact arrangements of C4c and C4d within C4b lead to equivalent good curve fits up to $Q = 1.1$ nm^{-1} (curves I and II). These models are shown as insets to the right.

by further data for C4 and C4u collected out to a larger Q value of 3.6 nm^{-1} (2-nm resolution; not shown). The curve fitting shows that C4 (and C3 also) can be represented as the sum of at least two distinct protein domains.

The dimensions of the models show that C4 and its larger fragments are well approximated by lamellar objects. They are compatible with the globular images of dimensions 9 nm × 17 nm seen in micrographs of C4b (Dahlbäck et al., 1983). The solution data, however, stress that the shapes are oblate rather than prolate, which is not clear from electron microscopy. The smallest C4 fragment, C4d, is, however, close to a prolate ellipsoid.

The conformational change found on going from C4 to C4b was investigated by further trial-and-error simulations (Figure 5b). The best curve fits extended out to a Q of 1.1 nm^{-1} and correspond to a model in which C4d is positioned on the long edge of C4c. This repositioning of the C4d domain accounts for the reduced R_{XS} values of C4b relative to those of C4 and C4u. Even though the difference between the scattering curves of C4 and C4b is small, it has been reproducible whenever measured both for C4 vs C4b and C3 vs C3b. However, while these models indicate the extent of conformational rearrangement between C4 and C4b, the accuracy of the models should not be overrated.

Conclusions

The multifunctional role of C4 is linked with the specific roles of the C4a, C4c, and C4d structures within C4. Analyses based on distinct C4c and C4d domains offer satisfactory explanations of the scattering curves and support the domain structure originally proposed for C3 (Perkins & Sim, 1986). C4 is recognized by C1s to be activated to C4b. C4b binds to C2 to form the C3 convertase. C4b also binds to the C4b-binding protein (C4BP) or to the CR1 membrane protein as cofactors for the factor I mediated breakdown of C4b into C4c and C4d. The common element of these binding interactions may be the "short consensus repeat" (SCR) domain that is found in C1s, C2, C4BP, and CR1 and which has been implicated in the binding to C4 (Reid et al., 1986). Since the binding site on C4 for C4BP is in the C4c region (Ferreira & Nussenzweig, 1979) and use is made of the C4c–C4BP interaction in affinity purification of C4BP, it is possible that the C4c structure interacts with other SCR-containing proteins. In C3, most SCR-containing proteins also interact with the C3c region, although CR2 and possibly factor H interact with C3dg (Lambris, 1988). C4d possesses the thiolester bond and is implicated in the covalent binding of C4b to acceptor surfaces. It is notable that the polymorphism of C4 and the different chemical reactivities of the C4 isotypes are primarily located in the C4d structure (Chakravarti et al., 1987).

C4c is unexpectedly found to be dimeric, unlike the other fragments and forms of C3 or C4. This result accounts for the notable difference in the $s^0_{20,w}$ coefficients of C3c (6.7 S) and C4c (5.5 S) that was reported by Pâques (1980) but not explained. The significance of dimerization might be to provide a mechanism for C4c to withdraw from interactions with its cofactors after it has been generated by factor I. If this is true, dimerization would conceal the SCR binding sites on C4c to permit the release of the SCR domains for renewed activation of the cascade.

Solution scattering indicates that C4 and C4u have similar structures, while a conformational change was detected between C4 and C4b. C4 and C4b appeared to be indistinguishable in the electron microscope (Smith et al., 1984), most probably due to the nonphysiological conditions of the measurements. The removal of C4a from C4 appears to be the key to this structural difference, since the R_{XS} of C4(a+b) in Table II (in which C4a is cleaved but not removed) is similar to those of C4 and C4u. This implies that C4a may be located between the C4c and C4d domains in the C4 structure. Interestingly, Isenman and Kells (1982) used circular dichroism and UV absorption difference spectroscopy to detect changes in localized environments within both C4u and C4b relative to that of C4. Both C4u and C4b have hydrolyzed thiolester bonds. It appears likely that the spectroscopic method is sensitive to small conformational rearrangements within the C4d domain after thiolester hydrolysis. Solution scattering, however, reports on the relative dispositions of the C4c and C4d domains and shows that C4b is more compact than C4, C4u, and C4(a+b). The different functional properties of C4u and C4b are then accounted for by the movement of the C4d domain relative to the C4c domain in C4b, but not in C4u.

The C4 structure is largely lamellar with dimensions of up to 8 nm × 18 nm. C4 is cleaved by activated C1s within C1

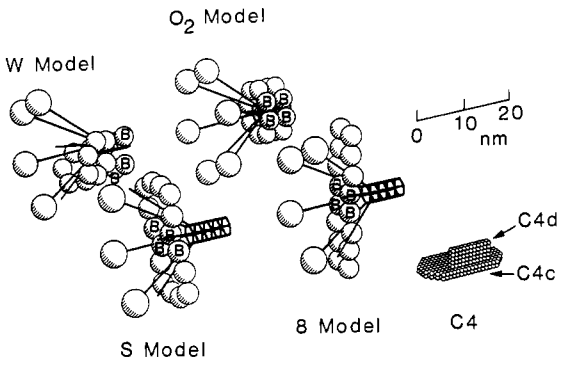

FIGURE 6: Comparisons of the Debye sphere model for C4 with four possible models for the C1 component of complement. All are drawn to the same scale and are viewed in perspective. Some of the C1q heads are removed for clarity in the W, S, and 8 models. The C1 models follow proposals of a W model (Perkins, 1985, 1989), an O_2 model (Cooper, 1985), a S model (Schumaker et al., 1986), and an 8 model (Colomb et al., 1984). The sphere models for C1 are derived from hydrodynamic simulations (Perkins, 1989). C1 is a complex of tetrameric $C1r_2C1s_2$ bound to hexameric C1q. The W model places each monomer of C1r and C1s on four adjacent arms of C1q [see Figure 8 of Perkins (1989)]. The O_2 model folds each C1r-C1s pair into a complete ring, and the two rings are then located on the base of C1q. The S model interweaves the $C1r_2C1s_2$ structure in and out of the C1q arms such that four arms are enclosed within the tetramer. The 8 model is similar to the S model except that two C1q arms are enclosed, not four. The protease activation domains of C1r and C1s that activate C4 to C4b are denoted by B. They are surface-exposed in the W and O_2 models, but are partially or fully buried in the C1q cone of arms in the S and 8 models. The two SCR domains are adjacent to the protease domains and can be approximated by two further spheres as shown (Perkins & Nealis, 1989). The SCR domains are again surface-exposed in the W and O_2 models, but not in the S and 8 models.

to form C4b. C1 is formed from the association of $C1r_2C1s_2$ with C1q, and four possible models have been proposed for its structure (Cooper, 1985; Schumaker et al., 1986; Colomb et al., 1984; Perkins, 1985, 1989). Hydrodynamic models for these from Perkins (1989) are summarized in Figure 6, together with the present C4 structure. The C1-mediated cleavage of C4 may involve the interaction of the C4c domain of C4 with C1s in C1 via the SCR domains of C1s, so that the serine protease domain in C1s can cleave C4a from C4c. Only two of the four C1 models, the W model and the O_2 model, have SCR domains and the serine protease domain in C1s that are solvent-exposed. Figure 6 shows that this surface accessibility would be crucial for the binding and proteolysis of C4, since C4 is large. The SCR and the protease domains are partially or fully buried within the C1q arms in the S model and the 8 model, and this is not supportive of these models.

ACKNOWLEDGMENTS

We thank B. Moffat for expert biochemical assistance, and Dr. C. Nave, Dr. J. Torbet, and Dr. D. L. Worcester for generous instrumental support at the SRS at Daresbury and the ILL in Grenoble.

Registry No. C4, 80295-48-3; C4u, 124379-52-8; C4(a+b), 124379-51-7; C4b, 80295-50-7; C4c, 80295-51-8; C4d, 80295-52-9.

REFERENCES

Belt, K. T., Carroll, M. C., & Porter, R. R. (1984) *Cell 36*, 907–914.

Campbell, R. D., Dodds, A. W., & Porter, R. R. (1980) *Biochem. J. 189*, 67–80.

Chakravarti, D. N., Campbell, R. D., & Porter, R. R. (1987) *Mol. Immunol. 24*, 1187–1197.

Chan, A. C., & Atkinson, J. P. (1985) *J. Immunol. 134*, 1790–1798.

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

Cohn, E. J., & Edsall, J. T. (1943) in *Proteins, amino acids and peptides*, pp 155–176 and 370–381, Van Nostrand Reinhold, New York.

Colomb, M. G., Arlaud, G. J., & Villiers, C. L. (1984) *Philos. Trans. R. Soc. London B306*, 283–292.

Cooper, N. R. (1985) *Adv. Immunol. 37*, 151–216.

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

Dahlbäck, B., Smith, C. A., & Müller-Eberhard, H. J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3461–3465.

Fairbanks, G., Steck, T. L., & Wallach, D. F. H. (1971) *Biochemistry 10*, 2602–2617.

Ferreira, A., & Nussenzweig, V. (1979) *J. Immunol. 122*, 490–493.

Ghosh, R. E. (1981) ILL Internal Publication 81GH29T.

Gigli, I., von Zabern, I., & Porter, R. R. (1977) *Biochem. J. 165*, 439–446.

Glatter, O., & Kratky, O., Eds. (1982) *Small-angle X-ray scattering*, Academic Press, New York.

Hsiung, L.-M., Barclay, A. N., Brandon, M. R., Sim, E., & Porter, R. R. (1982) *Biochem. J. 203*, 293–298.

Hsiung, L.-M., Mason, D. W., & Dodds, A. W. (1987) *Mol. Immunol. 24*, 91–96.

Ibel, K. (1976) *J. Appl. Crystallogr. 9*, 269–309.

Ibel, K., & Stuhrmann, H. B. (1975) *J. Mol. Biol. 93*, 255–266.

Isenman, D. E., & Kells, D. I. C. (1982) *Biochemistry 21*, 1109–1117.

Jacrot, B., & Zaccai, G. (1981) *Biopolymers 20*, 2413–2426.

Kratky, O. (1963) *Prog. Biophys. Chem. 13*, 105–173.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Lambris, J. D. (1988) *Immunol. Today 9*, 387–393.

Law, S.-K. A., Minich, T. M., & Levine, R. P. (1983) *Biochemistry 23*, 3267–3272.

Müller-Eberhard, H. J., Polley, M. J., & Calcott, M. A. (1966) *J. Exp. Med.*, 359–380.

Nagasawa, S., & Stroud, R. M. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2998–3001.

Nave, C., Helliwell, J. R., Moore, P. R., Thompson, A. W., Worgan, J. S., Greenall, R. J., Miller, A., Burley, S. K., Bradshaw, J., Pigram, W. J., Fuller, W., Siddons, D. P., Deutsch, M., & Tregear, R. T. (1985) *J. Appl. Crystallogr. 18*, 396–403.

Österberg, R., Eggertsen, G., Lundwall, A., & Sjoquist, J. (1984) *Int. J. Biol. Macromol. 6*, 195–198.

Österberg, R., Malmensten, B., Nilsson, U., Eggertsen, G., & Kjems, J. (1988) *Int. J. Biol. Macromol. 10*, 15–20.

Pâques, E. P. (1980) *Hoppe-Seyler's Z. Physiol. Chem. 361*, 445–456.

Perkins, S. J. (1985) *Biochem. J. 228*, 13–26.

Perkins, S. J. (1986) *Eur. J. Biochem. 157*, 169–180.

Perkins, S. J. (1988a) *New Compr. Biochem. 18B* (Part II), 143–264.

Perkins, S. J. (1988b) *Biochem. J. 254*, 313–327.

Perkins, S. J. (1989) *Behring Inst. Mitt. 84*, 129–141.

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., & Nealis, A. S. (1989) *Biochem. J. 263*, 463–469.

Perkins, S. J., Chung, L. P., & Reid, K. B. M. (1986) *Biochem. J. 233*, 799–807.

Reboul, A., Thielens, N., Villiers, M. B., & Colomb, M. G. (1980) *FEBS Lett. 155*, 118–121.

Reid, K. B. M. (1983) *Biochem. Soc. Trans. 11*, 1–12.

Reid, K. B. M. (1986) *Essays Biochem. 22*, 27–68.
Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T., & Tack, B. F. (1986) *Immunol. Today 7*, 230–234.
Schumaker, V. N., Hanson, D. C., Kilchherr, E., Phillips, M. L., & Poon, P. H. (1986) *Mol. Immunol. 23*, 557–565.
Sim, E., & Sim, R. B. (1983) *Biochem. J. 210*, 567–576.
Sim, E., Dodds, A. W., Wood, M., & Sim, R. B. (1986) *Biochem. Soc. Trans. 14*, 77–78.
Sim, R. B. (1981) *Methods Enzymol. 80*, 26–42.
Sim, R. B., & Sim, E. (1981) *Biochem. J. 193*, 129–141.
Smith, C. A., Vogel, C. W., & Müller-Eberhard, H. J. (1984) *J. Exp. Med. 159*, 324–329.
Tack, B. F., Morris, S. C., & Prahl, J. W. (1979) *Biochemistry 18*, 1497–1505.
Tack, B. F., Janatova, J., Thomas, M. L., Harrison, R. A., & Hammer, C. H. (1981) *Methods Enzymol. 80*, 64–101.
Wetlaufer, A. B. (1962) *Adv. Protein Chem. 17*, 303–390.

Structural Homologies of Component C5 of Human Complement with Components C3 and C4 by Neutron Scattering†

Stephen J. Perkins,*,‡ Kathryn F. Smith,‡ Adam S. Nealis,‡ Peter J. Lachmann,§ and Richard A. Harrison§

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

Received April 11, 1989; Revised Manuscript Received September 14, 1989

ABSTRACT: The complement component C5 is one of a family of structurally related plasma proteins that includes components C3 and C4. Activation of C5 is the initial step in the formation of the membrane attack complex of complement. Analysis of the solution structure of C5 and comparisons with similar analyses of the structures of C3 and C4 are reported here. Neutron solution scattering gave an M_r for C5 of 201 000, which demonstrates that C5 is monomeric in solution. The radius of gyration R_G of C5 at infinite contrast is 4.87 nm and corresponds to an elongated structure. The longest length of C5 was determined to be at least 15–16 nm from three calculations on the basis of the R_G, the scattering intensity at zero angle $I(0)$, and the indirect transformation of the scattering curve into real space. Comparison of the R_G and contrast variation data and indirect transformations of the scattering curves for C3, C4, and C5 show that these have very similar structures. Comparisons of the C5 scattering curve with Debye small-sphere models previously employed for C4 and C3 show that good curve fits could be obtained. Unlike previous studies that have suggested significant differences, these experiments indicate that, while C5 differs from C3 and C4 in its activation and inactivation pathways, significant structural homology exists between the native proteins, as might be predicted from their high (and similar) sequence homology.

Complement, a multicomponent cascade system found in plasma, plays a major role in the humoral immune system (Reid, 1986). Activation of the early components of the classical and alternative pathways by proteolytic cleavages generates C5 convertase activity. The convertases activate C5 to form C5b and C5a. The generation of C5b is the initial step in the formation of the membrane attack complex, which is assembled from the complement components C6, C7, C8, and C9 in association with C5b. Insertion of the complex into membranes leads to the formation of transmembrane channels in the target cell and ultimately to cell death due to osmotic imbalance.

The primary sequences of the complement components C3, C4, and C5 show significant homologies (Belt et al., 1984; de Bruijn & Fey, 1985; Wetsel et al., 1987, 1988). C5, like C3, is composed of two disulfide-linked polypeptide chains α and β. In contrast, C4 has three chains, α, β, and γ. Activation of C5 is brought about by cleavage of the peptide bond between residues 74 and 75 of the α chain, releasing C5a (Nilsson et al., 1975). C5a is the most potent of the three anaphylatoxins C3a, C4a, and C5a formed by the complement cascade. Despite the sequence homology, the genes for C3, C4, and C5 are found on different chromosomes (Carroll et al., 1984; Whitehead et al., 1982; Wetsel et al., 1988). While not as abundant as C3 and C4, C5 at 0.15 mg/mL is still one of the more abundant complement components in plasma. C3 and C4, but not C5, have a highly reactive internal β-cysteinyl-γ-glutamyl thiolester bond in a region of high sequence homology. During activation, this enables C3b and C4b to form covalent bonds to target surfaces (Law et al., 1980). Since there is no thiolester bond in C5, C5b is unable to form a covalent bond to a surface. Thus, C5b in the presence of C6 forms a fluid-phase C5b6 bimolecular complex, and this in association with C7 expresses a membrane-binding site (Podack et al., 1978).

In this study, comparisons between the solution structures of C3, C4, and C5 are made. Electron microscopy of C5 depicts a multilobal, irregular ultrastructure of dimensions of 10.4 nm × 14.0 nm × 16.8 nm (DiScipio et al., 1983); different

† We thank the Wellcome Trust for support, and the Science and Engineering Research Council for access to the facilities at SRS Daresbury and ILL Grenoble.
* To whom correspondence should be addressed at the Department of Biochemistry and Chemistry.
‡ Royal Free Hospital School of Medicine.
§ Medical Research Council Centre.

0006-2960/90/0429-1175$02.50/0