Solution Structure Determination of Monomeric Human IgA2 by X-ray and Neutron Scattering, Analytical Ultracentrifugation and Constrained Modelling: A Comparison with Monomeric Huma...

8 AUTHORS, INCLUDING:

Solution Structure Determination of Monomeric Human IgA2 by X-ray and Neutron Scattering, Analytical Ultracentrifugation and Constrained Modelling: A Comparison with Monomeric Human IgA1

Patricia B. Furtado[1]†, Patrick W. Whitty[2]†, Alexis Robertson[1]†
Julian T. Eaton[1], Adel Almogren[2], Michael A. Kerr[2]‡, Jenny M. Woof[2]
and Stephen J. Perkins[1]*

[1]*Structural Immunology Group Department of Biochemistry and Molecular Biology, Royal Free and University College Medical School, University College London, Gower Street London WC1E 6BT, UK*

[2]*Division of Pathology and Neuroscience, University of Dundee Medical School Ninewells Hospital, Dundee DD1 9SY, UK*

Immunoglobulin A (IgA), the most abundant human immunoglobulin, mediates immune protection at mucosal surfaces as well as in plasma. It exists as two subclasses IgA1 and IgA2, and IgA2 is found in at least two allotypic forms, IgA2m(1) or IgA2m(2). Compared to IgA1, IgA2 has a much shorter hinge region, which joins the two Fab and one Fc fragments. In order to assess its solution structure, monomeric recombinant IgA2m(1) was studied by X-ray and neutron scattering. Its Guinier X-ray radius of gyration R_G is 5.18 nm and its neutron R_G is 5.03 nm, both of which are significantly smaller than those for monomeric IgA1 at 6.1–6.2 nm. The distance distribution function $P(r)$ for IgA2m(1) showed a broad peak with a subpeak and gave a maximum dimension of 17 nm, in contrast to the $P(r)$ curve for IgA1, which showed two distinct peaks and a maximum dimension of 21 nm. The sedimentation coefficients of IgA1 and IgA2m(1) were 6.2 S and 6.4 S, respectively. These data show that the solution structure of IgA2m(1) is significantly more compact than IgA1. The complete monomeric IgA2m(1) structure was modelled using molecular dynamics to generate random IgA2 hinge structures, to which homology models for the Fab and Fc fragments were connected to generate 10,000 full models. A total of 104 compact best-fit IgA2m(1) models gave good curve fits. These best-fit models were modified by linking the two Fab light chains with a disulphide bridge that is found in IgA2m(1), and subjecting these to energy refinement to optimise this linkage. The averaged solution structure of the arrangement of the Fab and Fc fragments in IgA2m(1) was found to be predominantly T-shaped and flexible, but also included Y-shaped structures. The IgA2 models show full steric access to the two FcαRI-binding sites at the Cα2–Cα3 interdomain region in the Fc fragment. Since previous scattering modelling had shown that IgA1 also possessed a flexible T-shaped solution structure, such a T-shape may be common to both IgA1 and IgA2. The final models suggest that the combination of the more compact IgA2m(1) and the more extended IgA1 structures will enable human IgA to access a broader range of antigens than either acting alone. The hinges of both IgA subclasses appear to

show reduced flexibility when compared to their equivalents in IgG, and this may be important for maintaining an extended IgA structure.

Corresponding author

Introduction

Immunoglobulin A (IgA) is the most abundant human immunoglobulin class found on mucosal surfaces and is the second most abundant in plasma.[1,2] The mucosal surfaces that secrete IgA are major potential sites of pathogenic invasion because of their large surface areas. IgA1 and IgA2, the two subclasses of IgA, can exist in different polymerisation states. Serum IgA is mainly monomeric, and 90% of this is IgA1. Secretory IgA is predominantly dimeric with some higher oligomers, and IgA1 and IgA2 are found in similar amounts. IgA dimerisation involves the formation of disulphide bridges between two IgA monomers and a 15 kDa J chain. Secretory IgA is formed when dimeric IgA binds to the polymeric immunoglobulin receptor on the basolateral surface of epithelial cells lining mucosal sites. The transcytosis of this complex across the cell and proteolysis of the receptor results in the formation of secretory IgA, in which the major receptor fragment is called secretory component and this is disulphide bridged with dimeric IgA. Secretory IgA serves as the first line of defence against invading pathogenic microorganisms, preventing their entry into the body.[3,4] Important effector functions of serum IgA are mediated by binding to the cell surface Fcα receptor (FcαRI; CD89), which binds to serum and secretory IgA of both subclasses to trigger clearance mechanisms.[2,5,6] Serum IgA appears to function as a second line of defence in order to eliminate pathogens that have penetrated the mucosal surface.[7] Interestingly, homologues of human IgA1 are found only in some higher apes, while human IgA2 can be considered to be more similar to the IgA found in all other species.[8,9]

The monomeric form of both IgA subclasses have a 12-domain structure that is composed of two light and two heavy chains (Figure 1). The light chain contains a V-type and a C1-type immunoglobulin domain (V_L and C_L). The heavy chain contains a V-type and three C1-type domains (V_H, C_H1, C_H2 and C_H3). An 18 residue tailpiece peptide is attached to the C terminus of the C_H3 domains with an *N*-linked glycosylation site.[10] IgA1 contains a 23 residue hinge between Cys220 in the C_H1 domain and Pro244 in the C_H2 domain with up to five *O*-linked glycosylation sites, whilst in IgA2 there are only ten residues between Cys220 and Pro244 and no *O*-glycosylation sites (the residue numbering for IgA2 follows that in our previous modelling study of IgA1[10]). IgA2 exists as one of the two allotypes, IgA2m(1) and IgA2m(2), with a third form IgA2(n) possibly representing another allotype.[11] In distinction to IgA1, all three IgA2 variants contain two additional *N*-linked glycosylation sites at Asn166 and Asn337 in the C_H1 and the C_H2 domains, respectively (Figure 1). The IgA2m(2) and IgA2(n) variants also contain a fifth N-linked glycosylation site at Asn211 in the C_H1 domain. The N-linked sugar moieties of IgA exhibit a marked degree of compositional heterogeneity.[12,13] The different hinge regions of IgA1 and IgA2 account for their different susceptibilities to IgA1 proteinases secreted by many important mucosal pathogens.[14,15] IgA1 but not IgA2 is reported to bind to the transferrin receptor (CD71) with a binding profile reminiscent of that of a human T-cell receptor that recognises only IgA1 O-linked sugar moieties.[16,17] Also, the mouse epithelial M-cell receptor binds human IgA2 but not IgA1.[18]

Until now, structural studies on IgA have not resulted in a full structural model for IgA2m(1). The crystal structure of the Fab fragment from murine IgA J539 confirmed the expected β-sheet structures in the V and C domains.[19] The human IgA1 Fc crystal structure was published[20] after we had first presented our modelling study of IgA2m(1),[21]† and resembles those determined for IgG Fc and IgE Fc.[22–25] Crystal structures for intact human IgG1 and mouse IgG1 and IgG2a provided single snapshot views of asymmetric T-shaped and Y-shaped structures, showing that the IgG hinge exists in multiple conformations.[24,25] Several electron microscopy studies of IgA1 dimers revealed double Y-shaped structures with variable inter-Fab fragment angles and extended hinges.[26] Electron microscopy studies of IgA2 monomers suggested that these were more rigidly-hinged.[27] Scattering and constrained modelling of human IgA1 showed that the hinges were extended and gave flexible T-shaped structures in solution.[10] Averaged solution scattering structures have been likewise determined for bovine IgG1/2 and a mouse IgG1 chimeric antibody.[28–30] Here, we have investigated the solution structure of monomeric IgA2 for comparison with IgA1 and IgG. We present X-ray and neutron-scattering data on IgA2m(1), compare these with those for IgA1,[10] and supplement these by analytical ultracentrifugation data on both IgA2m(1) and IgA1. We show that IgA2m(1) is significantly more compact
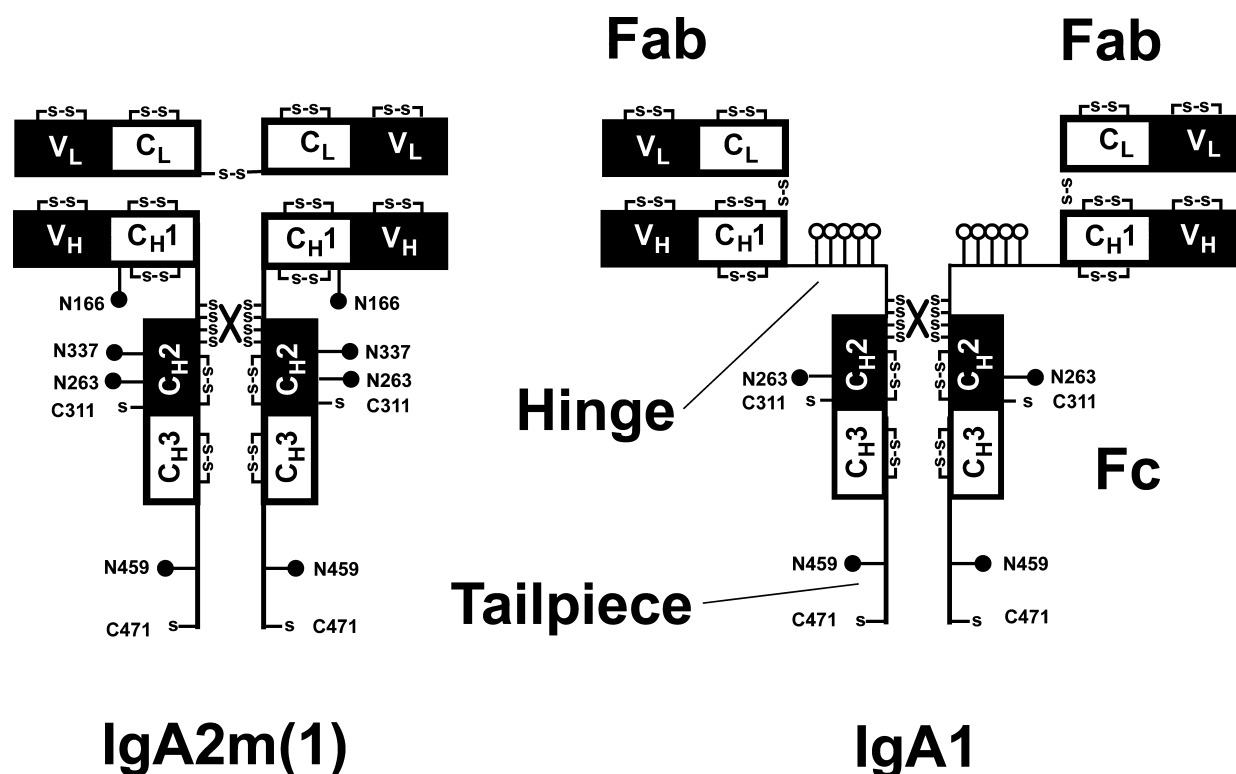
Figure 1. Schematic diagram of the Ig fold domains in human IgA1 and IgA2m(1). Each heavy chain contains the V_H, C_H1, C_H2 and C_H3 domains, and each light chain contains the V_L and C_L domains. The four-domain Fab fragments are linked to the four-domain Fc fragment by a 23 residue hinge in IgA1 and by a ten residue hinge in IgA2m(1). Each IgA heavy chain contains an 18 residue tailpiece at the C terminus. Each Ig fold has a conserved internal Cys–Cys disulphide bridge (S–S). Unlike IgA1, the two light chains of IgA2m(1) are linked by their C-terminal Cys214 residues. Both IgA molecules have several inter-heavy chain disulphide bridges. In IgA2m(1), Cys311 on the C_H2 domain and Cys471 in the tailpiece are shown free. IgA2m(1) has one N-linked oligosaccharide site at Asn166 on β-strand G of the C_H1 domain, two sites at Asn263 on β-strand B and Asn337 on β-strand G of the C_H2 domain, and one site at Asn459 on the tailpiece (●). IgA1 has up to five O-linked oligosaccharide sites in the hinge (○) that are not present in IgA2m(1), and two fewer N-linked oligosaccharide sites than in IgA2m(1).

than IgA1, its solution structure is T-shaped with possible contributions of Y-shapes, and both subclasses show reduced hinge flexibilities when compared to IgG, and discuss the functional implications of these results. We also discuss the functional implications of these results in terms of IgA polymerisation and the recent crystal structure for the IgA1 Fc fragment.[20]

Results and Discussion

X-ray and neutron Guinier analyses of monomeric IgA2m(1)

The solution arrangement of the Fab and Fc fragments of IgA2m(1) was first analysed using synchrotron X-ray scattering data $I(Q)$. Recombinant IgA2m(1) was used rather than serum IgA2 in order to enable the study of a pure monoclonal antibody in adequate quantities with a known antigenic specificity against the hapten NIP.[31] This IgA2 was a chimeric molecule composed of a heavy chain comprising the human α-chain C_H1, C_H2 and C_H3 domains and a murine V_H domain,

combined with a murine λ light chain with V_L and C_L domains (Figure 1). IgA2m(1) was studied at concentrations between 0.42 mg/ml and 1.12 mg/ml at the Synchrotron Radiation Source (SRS) and European Synchrotron Radiation Facility (ESRF). Time-frame analyses showed that the X-ray Guinier analyses for the SRS and ESRF data for IgA2m(1) were not completely resistant to X-ray-induced aggregation, as also observed for the Fc fragment of IgE,[32] but not for intact IgA1.[10] This aggregation was circumvented by calculating R_G values from the first time-frame only for the SRS and ESRF data. At the lowest Q values, Guinier analyses resulted in linear plots in three distinct regions of the $I(Q)$ curves from which the R_G, R_{XS-1} and R_{XS-2} values were obtained within satisfactory QR_G and QR_{XS} limits (Figure 2). IgA2m(1) exhibited an X-ray R_G value of $5.17(\pm 0.11)$ nm and $5.18(\pm 0.09)$ nm from the SRS and ESRF data analyses, respectively (Table 1). The anisotropy ratio R_G/R_O (where R_O is the R_G value of the sphere with the same volume as the hydrated glycoprotein) was calculated from the X-ray R_G value of IgA2m(1) to be 1.66. Comparison of this ratio to that of 1.99 determined for serum IgA1[10]
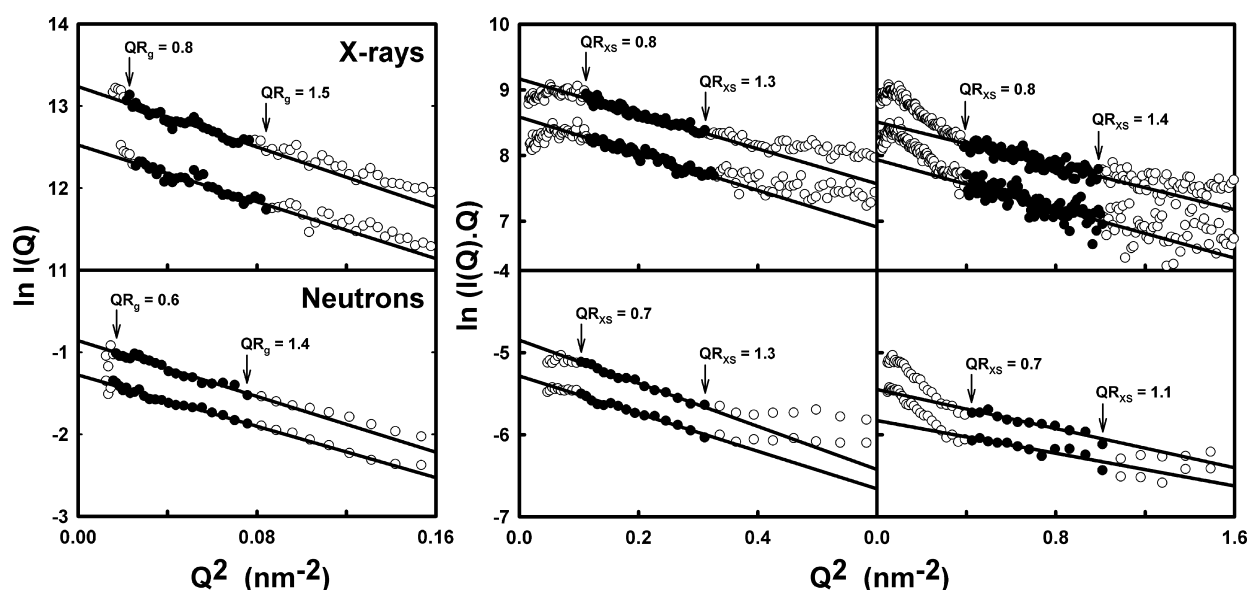
Figure 2. Guinier R_G and R_{XS} plots for recombinant IgA2m(1). These are shown for two different concentrations of 0.83 mg/ml and 0.42 mg/ml (X-rays) and 3.0 mg/ml and 2.0 mg/ml (neutrons). The filled circles between the arrowed QR_G and QR_{XS} ranges show the data points used to determine R_G, R_{XS-1} and R_{XS-2} values. The X-ray Q range used for the R_G values was 0.16–0.29 nm^{-1} for the SRS data, and 0.12–0.29 nm^{-1} for the ESRF data, and those for the R_{XS-1} and R_{XS-2} values were 0.33–0.51 nm^{-1} and 0.63–1.04 nm^{-1}, respectively, for both the SRS and ESRF data. The neutron Q range used for the R_G data was 0.13–0.27 nm^{-1}, and those for the R_{XS-1} and R_{XS-2} data were 0.32–0.55 nm^{-1} and 0.65–1.04 nm^{-1}, respectively.

showed that IgA2m(1) is less elongated than IgA1 in solution.

The X-ray cross-sectional Guinier analyses lead to a length determination L for IgA2m(1). From the SRS data, the R_{XS-1} value was 2.39($\pm$0.10) nm, whilst the R_{XS-2} value was 1.37($\pm$0.16) nm. From the ESRF data, the R_{XS-1} value was 2.47($\pm$0.09) nm, whilst the R_{XS-2} value was 1.47($\pm$0.08) nm. The combination of the R_G and R_{XS-1} values resulted in an overall length L for IgA2m(1) of 16.0 nm from both the SRS and ESRF data (Table 1), which is less than that of 20.1 nm for IgA1.

IgA2m(1) contains 11% (w/w) carbohydrate at eight sites (Figure 1). Carbohydrate has a higher scattering density than protein.[33] In order to confirm both the absence of aggregation in the final X-ray curves, and possible contrast-dependence effects from carbohydrate, neutron

scattering data was obtained for IgA2m(1) in 100% ^{2}H$_2$O buffers, which corresponds to a high negative protein–solvent scattering contrast. Neutron-scattering also provided complementary data on the hydration shell surrounding IgA2m(1), which is invisible by neutrons but not by X-rays.[34] Neutron data were collected at 2.0 mg/ml and 3.0 mg/ml at ISIS. The similarity of the neutron R_G value of 5.03($\pm$0.01) nm to the X-ray value (Table 1) confirmed that IgA2m(1) was stable during the first time-frame of X-ray irradiation. The slight decrease observed in the neutron R_G value when compared to the X-ray value is attributed to the invisible hydration shell.[34] The neutron Guinier $I(0)/c$ value was determined to be 0.14 $\pm$ 0.02 for IgA2m(1). The linear relationship between protein $I(0)/c$ values measured in ^{2}H$_2$O buffers and normalised against a standard deuterated polymer and the protein molecular mass M

Table 1. Summary of the scattering analyses of human IgA1 and IgA2m(1)

Protein	Data	Guinier analyses				$P(r)$ analyses			
		R_G (nm)	R_{XS-1} (nm)	R_{XS-2} (nm)	L (nm)	R_G (nm)	L (nm)	$M1$ (nm)	$M2$ (nm)
Recombinant IgA2m(1)	X-ray (SRS)	5.17 $\pm$ 0.11	2.39 $\pm$ 0.10	1.37 $\pm$ 0.16	16.0	5.08 $\pm$ 0.09	17.0	5.1 $\pm$ 0.9	n.a.
	X-ray (ESRF)	5.18 $\pm$ 0.09	2.47 $\pm$ 0.09	1.47 $\pm$ 0.08	16.0	5.21 $\pm$ 0.15	17.0	4.5 $\pm$ 0.3	6.8 $\pm$ 0.7
	Neutron	5.03 $\pm$ 0.01	2.21 $\pm$ 0.10	1.04 $\pm$ 0.06	15.5	5.12 $\pm$ 0.01	17.0	4.5	n.a.
Serum IgA1[a]	X-ray	6.20 $\pm$ 0.13	2.20 $\pm$ 0.26	1.56 $\pm$ 0.16	20.1	6.12	21.0	3.7	9.1
	Neutron	6.11 $\pm$ 0.18	2.17 $\pm$ 0.23	1.18 $\pm$ 0.12	19.8	–	–	–	–

n.a., not available.
[a] Taken from the work by Boehm *et al.*[10]

for LOQ ($M = I(0)/c \times 9 \times 10^5$) resulted in an estimated M value of 130,000 for IgA2m(1).[10] This measurement is less than but within the precision of the composition-derived value of 163,000 for IgA2m(1), showing that this was monomeric in solution.

The neutron length determination L of IgA2m(1) agreed with the X-ray data. The neutron R_{XS-1} value was 2.21(± 0.10) nm, and the neutron R_{XS-2} value was 1.04(± 0.06) nm. As also observed for IgA1,[13] the neutron R_{XS} values for IgA2m(1) were significantly less than their X-ray values. This is attributed to the hydration shell effect. The value of L calculated from the R_G and the R_{XS-1} values is 15.5 nm, which is within 0.5 nm of the X-ray value above, and is significantly less than the neutron L value of 19.8 nm for IgA1.

X-ray and neutron distance distribution analyses *P(r)*

Transformation of the observed curve $I(Q)$ into the distance distribution function $P(r)$ provided real space information on the IgA2m(1) structure. The maximum M in these curves corresponds to the most frequently occurring interatomic distance within the structure. The neutron and X-ray $P(r)$ curves for IgA2m(1) in Figure 3 showed a single peak. This was located at $r = 4.5$ nm (neutrons: Figure 3(a)), 5.1(± 0.9) nm (SRS X-rays) and 4.5(± 0.3) nm (ESRF X-rays: Figure 3(b)). In the latter, a marked inflexion could be observed at $r = 6.8(\pm 0.7)$ nm as the result of the higher instru-

mental quality of this data (Figure 3(b)). The R_G values determined from the $P(r)$ curves were 5.08(± 0.09) nm (SRS: ten curves), 5.21(± 0.15) nm (ESRF, seven curves) and 5.12(± 0.01) nm (ISIS, two curves), all of which agreed with the Guinier analyses (Table 1). The neutron and X-ray lengths L of IgA2m(1) were determined as 17.0 nm from the $P(r)$ curves where these intersect $P(r) = 0$, this being within 1.0–1.5 nm of the Guinier values (Table 1).

The observation of a single peak at 4.5 nm and a subpeak at 6.8 nm in the X-ray $P(r)$ curve of IgA2m(1) (Figure 3(b)) instead of the separate two peaks observed for PTerm455 and IgA1 at $r = 3.7$–4.0 nm and $r = 8.7$–9.1 nm (Figure 3(c) and (d)) shows that IgA2m(1) adopts a more compact shape than IgA1.[10] The IgA1 peak $M1$ at $r = 3.7$ nm can be assigned to the most commonly occurring distance in a single Fab or Fc fragment, given that each is approximately 8 nm long. As $M2$ is about half the overall length L for both IgA2m(1) and IgA1, this is assigned to the most common distance within the whole IgA structure. The existence of the $M2$ peak indicates that the Fab and Fc fragments are held apart in a relatively inflexible fashion as distinct structural entities by their hinge peptides.

Analytical ultracentrifugation analyses of monomeric IgA2m(1)

The molecular mass for the recombinant IgA2m(1) chimera was determined by sedimentation
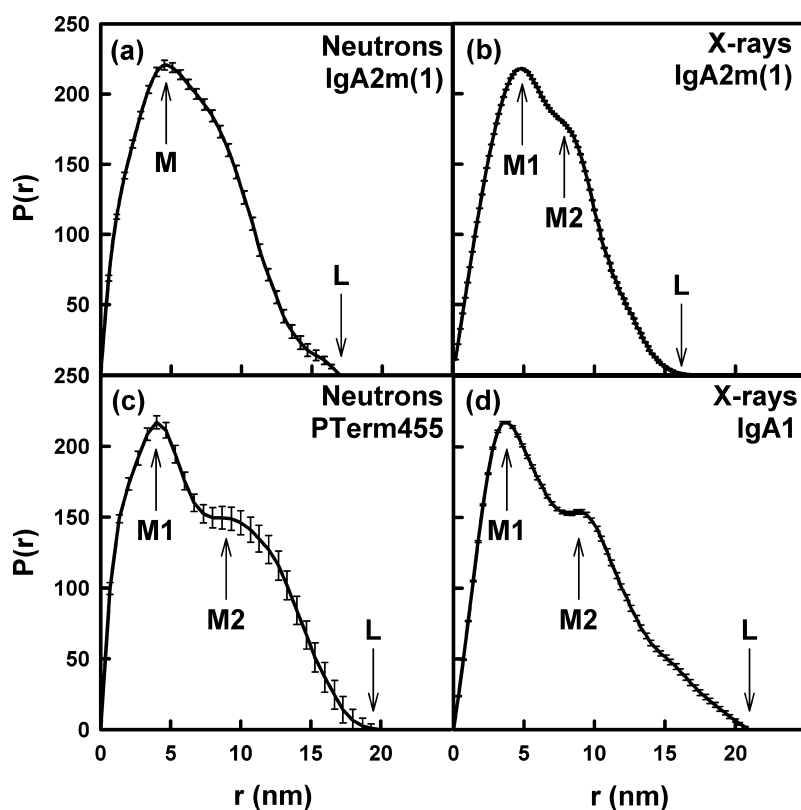
Figure 3. Distance distribution functions $P(r)$ for IgA2 and IgA1. (a) and (b) The $P(r)$ curves for recombinant IgA2m(1) using neutron and ESRF X-ray data, respectively. The single neutron peak M at 4.5 nm and the two X-ray peaks $M1$ and $M2$ at 4.5 nm and 6.8 nm are arrowed, and the maximum length of 17 nm is denoted by L. (c) The neutron $P(r)$ curve for the tailpiece deleted form of IgA1 (PTerm455) is shown, for which the positions of two maxima $M1$ and $M2$ occur at 4.0 nm and 8.7 nm respectively, and L is 19 nm. (d) The X-ray $P(r)$ curve for human IgA1 is shown, for which the peaks $M1$ and $M2$ occur at 3.7 nm and 9.1 nm, respectively, and the length L is 21 nm.

equilibrium experiments at concentrations between 0.13 mg/ml and 1.11 mg/ml (see Materials and Methods). Using both absorbance data monitored at 280 nm and interference data, the molecular mass of IgA2m(1) was determined from the data measured at three rotor speeds assuming that a single species was present (Figure 4). A slight concentration dependence in the molecular mass for IgA2m(1) was observed. Its molecular mass was determined as 161,000($\pm$6000) Da (interference) and 149,000($\pm$7000) Da (absorbance) at zero concentration. As these are within error of the molecular mass of 163,000 Da calculated from its sequence, IgA2m(1) was confirmed to be a monomer in solution.

The sedimentation coefficient $s_{20,w}^0$ monitors the extent of macromolecular elongation, being independent of the R_G measurements. Velocity experiments with IgA2m(1) were performed at six rotor speeds at concentrations between 0.28 mg/ml and 1.12 mg/ml (see Materials and Methods). Time-derivative analysis using the $g(s^*)$ method gave a mean $s_{20,w}^0$ value of 6.42($\pm$0.1) S from the peak position in Figure 5(a). The width of this peak leads to the diffusion coefficient. Hence, the molecular mass was determined from the Svedberg equation to be 147,000 under conditions in which time-broadening effects were negligible. To eliminate systematic errors due to diffusion-broadening, direct boundary fitting of the raw sedimentation scans was performed (Figure 5(c)). This gave a mean $s_{20,w}^0$ value of 6.57($\pm$0.1) S, which is slightly larger but within
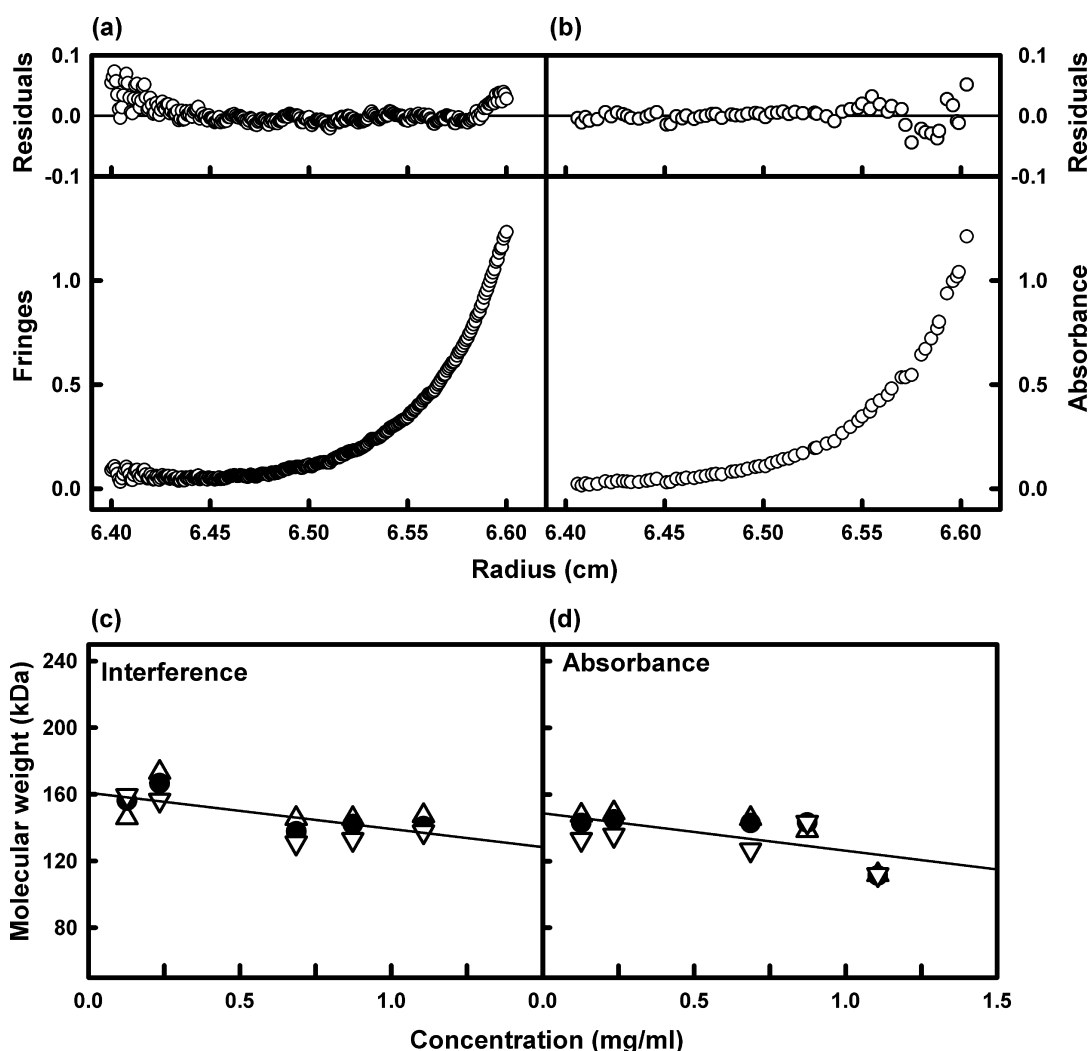
Figure 4. Sedimentation equilibrium data for IgA2m(1) by analytical ultracentrifugation. (a) and (b) Best-fit curves for IgA2m(1) at 0.24 mg/ml, both recorded at a rotor speed of 14,000 rpm using interference and absorbance optics, respectively. The fits were based on the assumption of a single species, and gave molecular masses of 162,000 Da and 154,000 Da, respectively. In the upper panels, the residuals of the curve fits are randomly distributed. (c) and (d) The dependence of the observed molecular mass values of IgA2m(1) at five concentrations based on the interference and absorbance data, respectively. Three rotor speeds were used: ($\triangle$) 11,000 rpm; ($\bullet$) 14,000 rpm; ($\triangledown$) 17,000 rpm. Regressions to zero concentrations of each set of 15 data points gave molecular mass values of 161,000($\pm$6000) Da and 149,000($\pm$7000) Da, respectively.
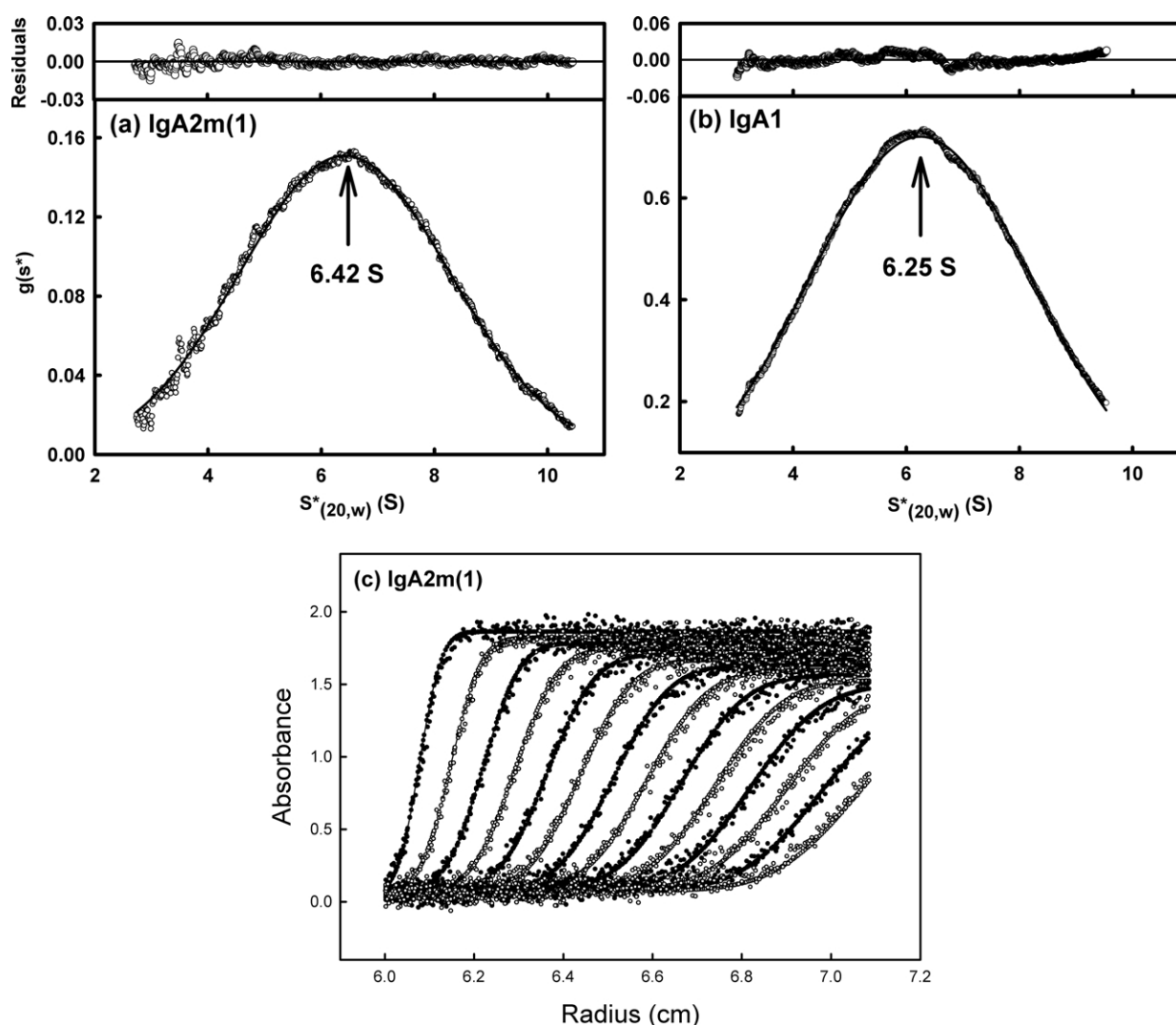
Figure 5. Sedimentation velocity analyses of IgA2m(1) and IgA1. (a) and (b) Fits of the $g(s^*)$ distribution to a Gaussian function based on the DCDT+ program. In (a), the IgA2m(1) fit as shown was based on ten scans recorded at 20,000 rpm at eight minutes intervals using absorbance optics at 280 nm at a sample concentration of 0.28 mg/ml in PBS. From this analysis, $s^0_{20,w}$ was determined to be 6.42 S, and the molecular mass was estimated to be 147,000($\pm$25,000) Da. In (b), the serum IgA1 fit as shown was based on 12 scans recorded at 20,000 rpm at a sample concentration of 1.12 mg/ml, yielding an $s^0_{20,w}$ value of 6.25 S in this case. (c) Fit of the 120 raw sedimentation scans for IgA2m(1) at 25,000 rpm based on the SEDFIT program. For clarity, only the fit at every fifth scan during the first 70 scans is shown. This analysis resulted in an $s^0_{20,w}$ value of 6.48 S.

error of the $g(s^*)$ analysis. The slight increase is as expected if diffusion had affected the $g(s^*)$ analyses.

The $s^0_{20,w}$ value of IgA1 was also determined. This was 6.20($\pm$0.2) S for human myeloma IgA1 between 0.70 mg/ml and 1.61 mg/ml, and 6.15($\pm$0.1) S for human serum IgA1 between 0.7 mg/ml and 1.5 mg/ml (Figure 5). Direct boundary fitting gave 6.33($\pm$0.1) S (not shown). These values agree well with those of 6.2 S for myeloma IgA1 Hui and 6.3 S for serum IgA1 measured on a Beckman model E[26] and are smaller than a value of 6.56 S for a human myeloma IgA1 protein.[35] The slightly smaller $s^0_{20,w}$ value for IgA1 compared to IgA2m(1) might reflect their molecular masses of 160,000 Da and 163,000 Da, respectively derived from their compositions.

However, unlike the R_G/R_O anisotropy ratios, which indicated a clear difference between IgA1 and IgA2m(1), the frictional coefficient ratios for IgA1 and IgA2m(1) were almost identical at 1.56 and 1.53, respectively. This outcome suggested that either IgA1 or IgA2m(1) or both possess unusual hydrodynamic properties.

Molecular dynamics search for an IgA2m(1) scattering structure

The constrained modelling analysis of IgA2m(1) differs from that for bovine IgG1 and IgG2, human IgA1 and murine Crry-Ig,[10,28,29] for reason of the unique short hinge in IgA2m(1). Its structure is also influenced by a disulphide bridge frequently found between the C termini of the two
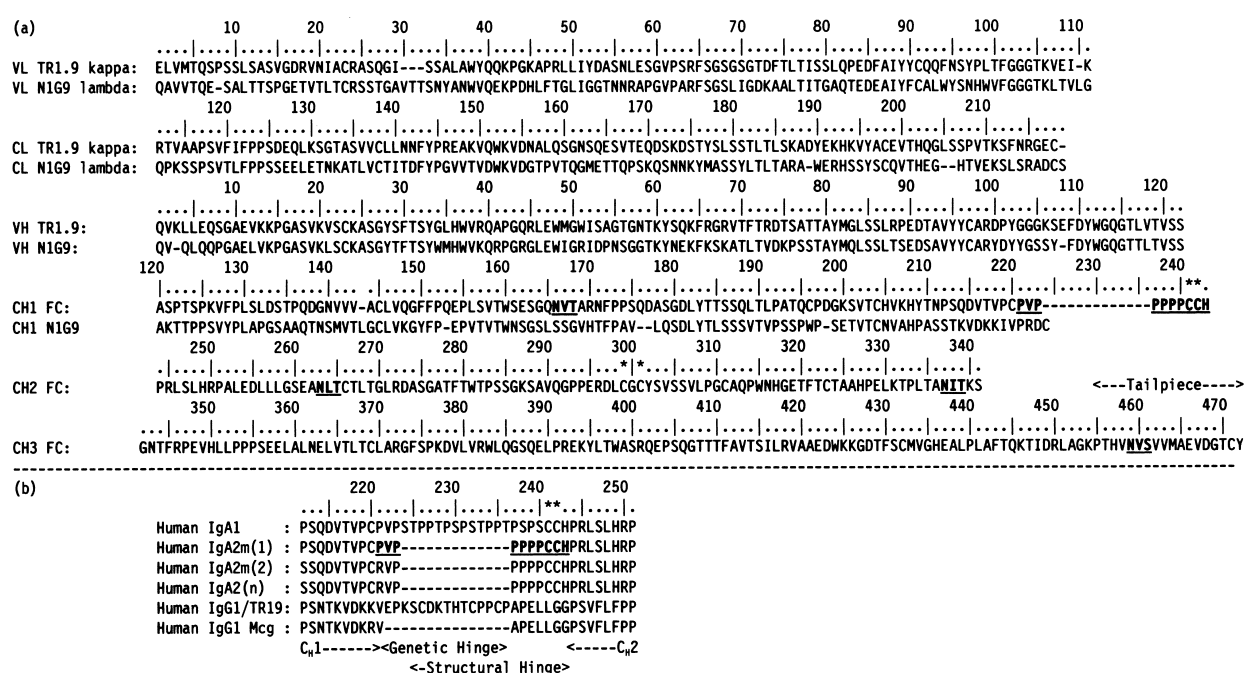
```
(a)                        10        20        30        40        50        60        70        80        90       100       110
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|.
VL TR1.9 kappa:  ELVMTQSPSSLSASVGDRVNIACRASQGI---SSALAWYQQKPGKAPRLLIYDASNLESGVPSRFSGSGSGTDFTLTISSLQPEDFAIYYCQQFNSYPLTFGGGTKVEI-K
VL N1G9 lambda:  QAVVTQE-SALTTSPGETVTLTCRSSTGAVTTSNYANWVQEKPDHLFTGLIGGTNNRAPGVPARFSGSLIGDKAALTITGAQTEDEAIYFCALWYSNHWVFGGGTKLTVLG
                          120       130       140       150       160       170       180       190       200       210
                        ...|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....
CL TR1.9 kappa:  RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC-
CL N1G9 lambda:  QPKSSPSVTLFPPSSEELETNKATLVCTITDFYPGVTVDWKVDGTPVTQGMETTQPSKQSNNKYMASSYLTLTARA-WERHSSYSCQVTHEG--HTVEKSLSRADCS
                         10        20        30        40        50        60        70        80        90       100       110       120
                        ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|..
VH TR1.9:        QVKLLEQSGAEVKKPGASVKVSCKASGYSFTSYGLHWVRQAPGQRLEWMGWISAGTGNTKYSQKFRGRVTFTRDTSATTAYMGLSSLRPEDTAVYYCARDPYGGGKSEFDYWGQGTLVTVSS
VH N1G9:         QV-QLQQPGAELVKPGASVKLSCKASGYTFTSYWMHWVKQRPGRGLEWIGRIDPNSGGTKYNEKFKSKATLTVDKPSSTAYMQLSSLTSEDSAVYYCARYDYYGSSY-FDYWGQGTTLTVSS
                   120       130       140       150       160       170       180       190       200       210       220       230       240
                        |....|....|....|....|... .|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|**.
CH1 FC:          ASPTSPKVFPLSLDSTPQDGNVVV-ACLVQGFFPQEPLSVTWSESGQNVTARNFPPSQDASGDLYTTSSQLTLPATQCPDGKSVTCHVKHYTNPSQDVTVPCPVP-------------PPPPCCH
CH1 N1G9         AKTTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFP-EPVTVTWNSGSLSSGVHTFPAV--LQSDLYTLSSSVTVPSSPWP-SETVTCNVAHPASSTKVDKKIVPRDC
                     250       260       270       280       290       300       310       320       330       340
                        .|....|....|....|....|....|....|....|....|.*|*....|....|....|....|....|....|....|....|....|....|....|....
CH2 FC:          PRLSLHRPALEDLLLGSEANLTCTLTGLRDASGATFTWTPSSGKSAVQGPPERDLCGCYSVSSVLPGCAQPWNHGETFTCTAAHPELKTPLTANITKS                  <---Tailpiece---->
                     350       360       370       380       390       400       410       420       430       440       450       460       470
                        ...|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....
CH3 FC:          GNTFRPEVHLLPPPSEELALNELVTLTCLARGFSPKDVLVRWLQGSQELPREKYLTWASRQEPSQGTTTFAVTSILRVAAEDWKKGDTFSCMVGHEALPLAFTQKTIDRLAGKPTHVNVSVVMAEVDGTCY
```

```
(b)                                    220       230       240       250
                                     ...|....|....|....|....|****..|....|.
                 Human IgA1     : PSQDVTVPCPVPSTPPTPSPSTPPTPSPSCCHPRLSLHRP
                 Human IgA2m(1) : PSQDVTVPCPVP-------------PPPPCCHPRLSLHRP
                 Human IgA2m(2) : SSQDVTVPCRVP-------------PPPPCCHPRLSLHRP
                 Human IgA2(n)  : SSQDVTVPCRVP-------------PPPPCCHPRLSLHRP
                 Human IgG1/TR19: PSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFPP
                 Human IgG1 Mcg : PSNTKVDKRV---------------APELLGGPSVFLFPP
                                  CH1------><Genetic Hinge>        <-----CH2
                                          <-Structural Hinge>
```

Figure 6. Sequence of human IgA2m(1) used for structural modelling. (a) The sequences of the light and heavy chains are shown. To follow Figure 1, the hinge peptide and four glycosylation sites are shown in bold and underlined, and Cys241, Cys242, Cys299 and Cys301 are asterisked. The sequence of the human IgG1 TR1.9 Fab fragment (PDB code 1vge) is compared with that of the murine N1G9 Fab fragment (PDB code 1ngq). The remaining sequence alignment and numbering follows that for IgA1. (b) Alignment for the hinge regions of IgA1, IgA2 and IgG1 and their flanking C_H1 and C_H2 domains. TR1.9 contains a normal IgG1 hinge peptide while that of IgG1 Mcg has had its genetic hinge deleted. These sequences are compared with those of the human IgA1 α-chain and the three variants of human IgA2. The IgA2m(1) hinge peptide of this study is shown in bold and underlined.

light chains[31,36,37] (Figure 1). Initially, this disulphide bridge was disregarded. IgA2m(1) models were first created with their Fab and Fc fragments in a wide unrestrained range of conformations (see Materials and Methods). Molecular dynamics simulations generated 10,000 random hinge conformations starting from an extended conformation of length 3.54 nm for the ten residue hinge peptide 221-PVPPPPPCCH-243 (Figure 6). Homology models for the Fab and Fc fragments of human IgA2 were created from those used for IgA1.[10] Finally, the Fab and Fc fragments were connected by two identical hinges to give a 2-fold symmetric model for IgA2m(1).

Figure 7(a) showed that the simulations had sampled the three-dimensional space between the Fab and Fc fragments in sufficient detail to enable any model compatible with the scattering data to be identified. Filters based on the X-ray and neutron-scattering data were now used to reject unsatisfactory models. The following results were obtained.

(1) The 5% filter for steric overlap eliminated the IgA2m(1) models in which the Fab and Fc fragments were sterically overlapped. The hydrated X-ray models needed at least 1800 ± 90 spheres, while the dry neutron models needed at least 1362 ± 68 spheres. These over-

lapping models correspond to the white circles at the bottom of the distribution in Figure 7(a). Only 2498 models (yellow circles) of the 10,000 models passed this filter, and they corresponded to only the longer Fab–Fab distances and Fab–Fc distances.

(2) The $\pm 5\%$ filter for the X-ray R_G value of 5.18 nm and the $\pm 10\%$ filter for the X-ray R_{XS-1} value of 2.47 nm (bottom of Table 2) together with the steric overlap filter gave 197 satisfactory hydrated models (green circles in Figure 7(b)). These corresponded to the more extended models with larger R_G values, for which the hinge length ranged between 1.2 nm and 3.2 nm.

(3) The same filters for the neutron R_G value of 5.03 nm and R_{XS-1} value of 2.21 nm (bottom of Table 2) together with the steric overlap filter gave 318 satisfactory dry models (green circles in Figure 7(c)). This distribution is similar to that obtained from the X-ray data. These models corresponded to hinge lengths between 1.3 nm and 3.5 nm. The combination of the filters in Figure 7(b) and (c) resulted in the 104 best-fit structures shown in Figure 7(a).

(4) The 10,000 models were sorted against the goodness-of-fit X-ray and neutron R-factors in Figure 7(d) and (e), respectively. The 104 best-fit structures of Figure 7(a) (green circles) mostly corresponded to the lowest R-factors. The 104
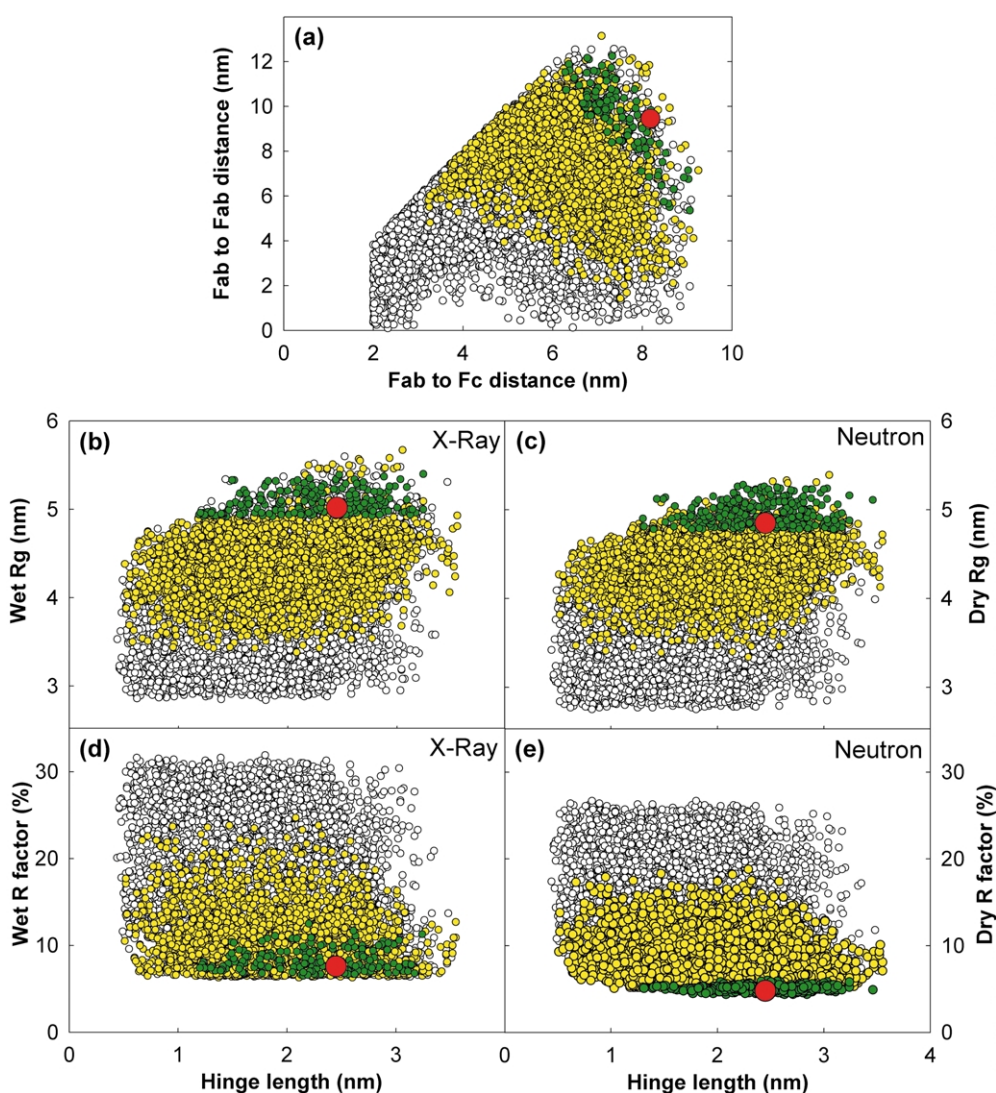
Figure 7. Outcome of the molecular dynamics searches of 10,000 hinge structures connecting the Fab and Fc fragments in IgA2m(1). In (a)−(e), yellow circles denote the 2498 models that passed the filters for the minimum number of spheres N required in each model. Green circles denote the 104 models that passed the filters for R_G, R_{XS-1} and N and also gave the lowest R-factors. The best-fit model is denoted by a red circle. Statistics are reported in Table 2. (a) Distribution of the 10,000 models generated by molecular dynamics simulation of IgA2m(1) hinge structures as a function of the distance between the centres of mass of the two Fab fragments and that between the Fab and Fc fragments. Note that, in here and in (b)−(e), the best-fit models in green are found as a single cluster, and not to several. (b) and (c) The X-ray and neutron radii of gyration R_G of the 10,000 models is shown as a function of hinge length for the hydrated and dry models, where the target experimental R_G values are 5.18 nm for hydrated IgA2m(1) (X-ray models) and 5.03 nm for dry IgA2m(1) (neutron models). The maximum length of the ten residue hinge is 3.54 nm. (d) and (e) The X-ray and neutron R-factors of the 10,000 models is shown as a function of hinge length.

neutron R-factors ranged between 4.4% and 5.8%, while some small spread between 6.5% and 12.5% was observed in the X-ray R-factors (last column of Table 2). Molecular views of the superimposed 104 best-fit structures are shown in Figure 8.

Refinement of a best-fit IgA2m(1) structure

The IgA2m(1) modelling was completed by connecting the Cys214−Cys214 inter-light chain disulphide bridge. In the 10,000 starting models, the α-carbon separations between the two Cys214 residues ranged between 0.1 nm and 9.1 nm. Initially, a selection of 50 best-fit models was made by filtering the 10,000 models for the X-ray R-factors calculated using data from the SRS, Daresbury. These 50 best-fit models were submitted to Crystallography and NMR System (CNS) energy minimisation to form the disulphide bridges (see Materials and Methods). This process resulted in the loss of about 75% of the models, either because 2-fold symmetry was lost, the

Table 2. Summary of the modelling searches for an IgA2m(1) solution structure

Modelling procedure	Filter	Models	Hinge length[a] (nm)	Data	Spheres	R_G (nm)	R_{XS-1} (nm)	R_{XS-2} (nm)	R-factor (%)
Molecular dynamics	None	10,000	0.44−3.55	X-ray	1029−1881	2.84−5.67	0.89−2.95	0.40−2.83	6.3−31.9
				Neutron	835−1370	2.75−5.39	0.04−2.72	0.78−2.51	4.4−26.7
Molecular dynamics	N, R_G and R_{XS-1}	104	1.31−3.24	X-ray	1734−1847	4.93−5.50	2.02−2.54	1.06−1.63	6.5−12.5
				Neutron	1295−1354	4.76−5.16	2.21−2.58	1.15−1.52	4.4−5.8
CNS	Best 15	15	1.34−3.03	X-ray	1772−1928	4.66−5.30	2.19−2.60	1.36−1.68	6.8−10.7
				Neutron	1320−1390	4.46−5.11	2.26−2.50	1.38−1.54	5.0−7.4
CNS	Final model	1	2.45	X-ray	1832	5.01	2.53	1.43	7.9
				Neutron	1354	4.85	2.44	1.40	5.3
Experimental				X-ray	1800	5.18	2.47	1.47	
				Neutron	1362	5.03	2.21	1.04	

[a] The hinge length is defined as the distance between the α-carbon atoms of Cys220 and Pro244 in the sequence 220-CPVPPPPP CCHP-244. Other details are reported in Figure 7.

inter-light chain disulphide was formed inside the Fc fragment, or the curve-fit deteriorated because the Fab and Fc fragments had moved too much. Comparison of the 50 models with the 104 best-fit structures from Figure 7(a) showed that 15 of these occurred in both sets of models (Table 2). The Cys214−Cys214 α-carbon separations in these 15 models ranged between 2.3 nm and 6.6 nm prior to the connection of the disulphide bridge (optimally this separation should be 0.6 nm). One of these 15 models was superior to the remainder when they were ranked in order of their ESRF X-ray R-factors (red circle in Figure 7). This model had a pre-CNS refinement Cys214 α-carbon separation of 2.7 nm, an X-ray R-factor of 7.6% and a neutron R-factor of 4.8%. After CNS refinement, the X-ray and neutron scattering curve fits showed good visual agreements in a large Q range, with only a slightly worsened X-ray R-factor

of 7.9% and a slightly worsened neutron R-factor of 5.3% (Figure 9). Another of these 15 models with a pre-refinement Cys214−Cys214 α-carbon separation of 3.5 nm has been published elsewhere.[21]

Two control calculations completed the analyses.

(1) A human IgA2m(1) model with a κ class light chain (Figure 6(a)) had been used for the modelling of IgA2m(1), which contained a human−mouse chimeric sequence with a λ class light chain. A CLUSTAL analysis showed that only five small insertions were needed to align the κ and λ light chains (dashes in Figure 6(a)). The crystal structure of the murine N1G9 Fab fragment (PDB code 1ngq[24]) comprises an appropriate set of V_H, C_H1, V_L and C_L domains for remodelling. Its C_H1 domain was superimposed with that in the human IgA2m(1) model. In
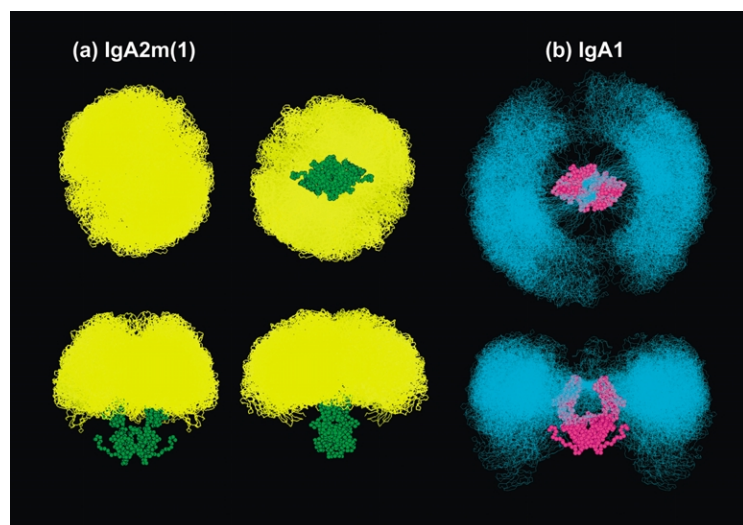

Figure 8. Superimposition of the best-fit models for IgA2m(1) and IgA1. The upper view looks down onto the Fc fragment (left and right images) or up from the bottom (middle image) of the Fc fragment, while the lower view is a face-on view (left and right images) or side-on view (middle image) of the α-carbon atoms in the Fc fragment (represented using green or purple spheres). (a) Four views of the 104 best-fit models for IgA2m(1) superimposed onto the Fc fragment. The 208 Fab fragments are shown by yellow α-carbon traces. Note that the Cys214−Cys214 light chain disulphide bridges are not connected in these models, as suggested in Figure 1. (b) The 104 best-fit models for the tailpiece-deleted PTerm455 mutant of IgA1 from the previous modelling[10] are shown for comparison purposes. The 208 Fab fragments are shown in cyan.
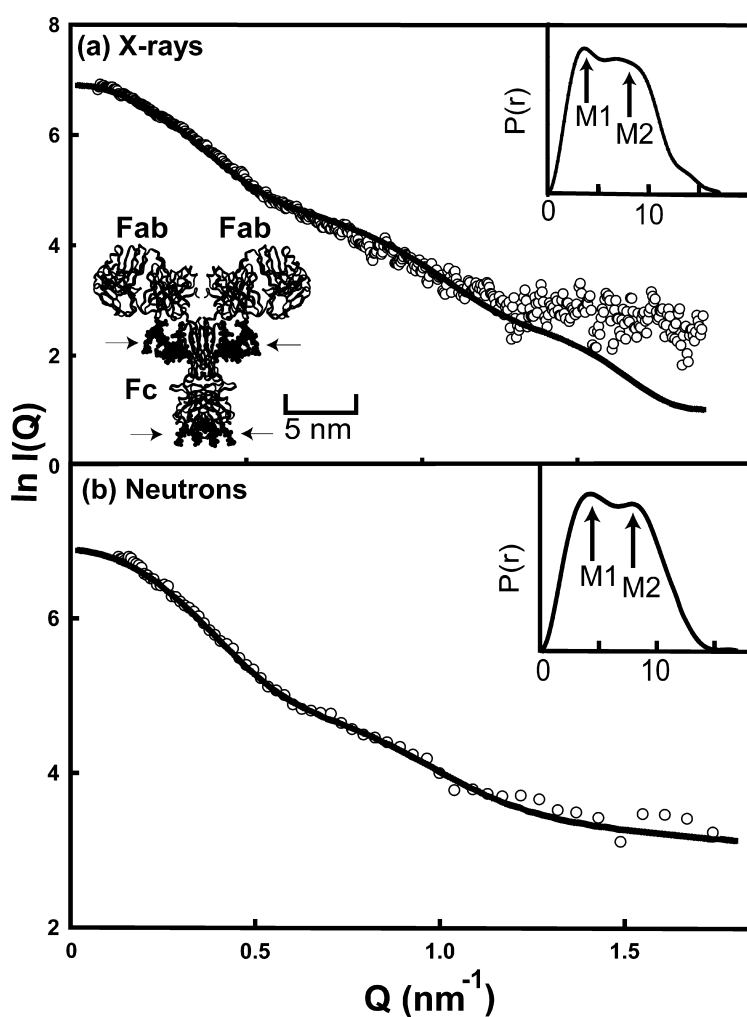
Figure 9. Final curve fits for (a) the X-ray model and (b) the neutron model of IgA2m(1). The modelled curves (continuous lines) are compared with ESRF X-ray data for IgA2m(1) and neutron data for IgA2m(1) in 100% ^2H_2O from Instrument LOQ (open circles). A flat background correction of 1.7% of log $I(0)$ was applied in order to allow for the uniform incoherent scattering arising from the proton content in the sample. Experimental error bars are shown when significant. The $P(r)$ curves calculated directly from the models are shown as insets. The α-carbon ribbon trace of the best-fit model is shown. Possible positions of the eight N-linked oligosaccharide chains shown in bold are arrowed, where six of these occur at the angle between the Fab and Fc fragments.

turn, the murine V domains were independently superimposed onto the human V domains independently of the C domains in order to preserve the human Fab elbow angle between the V and C domains. This resulted in only a slightly worsened X-ray R-factor of 8.9% and a similar neutron R-factor of 5.3%, thus showing that the curve fit is unaffected.

(2) The IgA2m(1) modelling of Figure 9 had been performed[21] prior to publication of the human IgA1 Fc crystal structure.[20] The Fc homology model in the IgA2m(1) model was replaced by this crystal structure. After the addition of the tailpieces and missing oligosaccharide chains, the resulting curve fits gave an identical X-ray R-factor of 7.9% and a slightly better neutron R-factor of 5.2%. This indicated that the IgA1 Fc crystal structure was structurally compatible with the Fc homology model used for the scattering curve fits.

Sedimentation coefficient modelling for IgA1 and IgA2m(1)

The calculated $s^0_{20,w}$ values for the 104 best-fit

IgA2m(1) models (Figure 8) ranged between 7.32 S and 7.82 S using the HYDRO program, and between 7.39 S and 7.93 S using the GENDIA program. These are higher than the experimental sedimentation coefficient $s^0_{20,w}$ value of 6.42 S for IgA2m(1). The mean predicted value from 82 best-fit IgA1 models was 6.92($\pm$0.07) S.[10] The final IgA1 model gave a predicted $s^0_{20,w}$ value of 6.81 S, which is higher than our experimental values of 6.15 S and 6.20 S, and those of 6.2–6.3 S and 6.56 S determined by others.[26,35] This meant that both the IgA2m(1) and IgA1 models were too small to replicate the experimental $s^0_{20,w}$ values, even though both structures had been hydrated by the standard amount of 0.3 g/g glycoprotein.

An investigation of the HYPRO hydration algorithm (see Materials and Methods) suggested that the IgA models were insufficiently hydrated. Many of the extra hydration spheres were found in the void spaces between the $V_L–V_H$ and the $C_L–C_H1$ domain pairs and the two $C_H1–C_H2$ domain pairs, and not on the protein surface. This discrepancy could be rectified by increasing the hydration in both the IgA1 and IgA2 models until agreement was obtained with the experimental $s^0_{20,w}$ values. The addition of a hydration volume

of 198 nm^3 instead of 67 nm^3 caused the $s^0_{20,w}$ value of IgA1 predicted by GENDIA to decrease to 6.41 S. The addition of 350 nm^3 instead of 67 nm^3 caused the predicted $s^0_{20,w}$ value of IgA2m(1) to decrease to 6.6 S. This outcome implied that both IgA1 and IgA2m(1) possess a high level of hydration.

Control calculations using GENDIA supported this interpretation. Four crystal structures for the FcαRI monomer (PDB codes 1ovz and 1ow0; hydration of 0.3 g/g of glycoprotein) gave a predicted $s^0_{20,w}$ value of 2.7($\pm$0.2) S, in satisfactory agreement with the experimental value of 2.36 S.[38] The IgA1 Fc fragment (PDB code 1ow0) gave an $s^0_{20,w}$ value of 3.9 S, also in agreement with the experimental value of 3.8 S.[38] That for rat Crry (PDB code 1ntj) gave an $s^0_{20,w}$ value of 2.51 S, also within error of the experimental value of 2.4 S.[29] The two structures for the mouse chimeric antibody Crry-Ig (PDB code 1ntl) gave $s^0_{20,w}$ values of 5.58 S and 5.73 S, also in satisfactory agreement with the experimental value of 5.4 S.[29]

Conclusions

Experimental structures of monomeric IgA1 and IgA2 and their functions

IgA1 and IgA2 have few biologically distinct properties. The main difference between the two subclasses is their susceptibilities to proteases, where IgA1 is specifically cleaved by certain bacterial proteases and this is explained by the exposed nature of the longer IgA1 hinge[39–41] (Figure 11). In contrast, the short hinge of human IgA2 not only lacks these protease cleavage sites, but is short enough to make IgA2 less susceptible to proteolysis in the harsh luminal environment following mucosal secretion. The present X-ray and neutron-scattering data have provided new insights into elongation and flexibility in the solution structures of monomeric IgA1 and IgA2m(1). The R_G values and the $P(r)$ curves show that IgA2m(1) has a more compact structure than that of IgA1. This result is readily attributable to the different lengths of the 23 residue hinge of IgA1 and the ten residue hinge of IgA2m(1). The $P(r)$ curves show that the hinges in both IgA subclasses are less flexible in IgG,[28] and this is discussed below. These results were determined from experimental data in phosphate-buffered saline that were acquired in a concentration range comparable to the total serum IgA, which is about 3 mg/ml, making the structures physiologically relevant. The same outcome was obtained from both X-ray and neutron data, indicating that the hydration shell surrounding IgA2 had little influence upon the observed structure. In contrast, it was interesting that the analytical ultracentrifugation data were not able to distinguish structurally between the two subclasses, and this

outcome may result from different levels of hydration in IgA1 and IgA2m(1).

The identification of structurally extended IgA1 molecules implies that IgA1 is capable of recognising antigenic epitopes that are spatially widely separated, while the identification of more compact IgA2m(1) molecules will correspond to the recognition of epitopes that are positioned more closely together. In combination, monomeric IgA1 and IgA2m(1) will provide a broad range of structural possibilities for high-avidity antigen recognition. Compared to the IgG subclasses and IgA2, IgA1 appears to possess the greatest capability to bind bivalently, and hence with greater avidity, to antigens spaced relatively far apart. Given that the IgA1 hinge is expected to be flexible, as for other antibodies, any IgA1 hinge flexibility will mean that IgA1 can bind efficiently to antigens at a wider range of separations. The increased possibility of high avidity binding to foreign antigens provided by the enhanced "antigenic reach" of IgA1 may explain why the evolution of this subclass has been favoured in humans and apes,[9] despite its susceptibility to proteolytic attack. This capability may be advantageous in the immune recognition of pathogens, especially if, as in some bacteria, they express repeated antigenic structural motifs far apart on their surface. The overall picture is, however, more complex, given that the IgA-enhanced pneumoccal adherence of bacteria utilises an IgA1 protease to target the IgA1 hinge in order to promote bacterial colonisation.[41] Likewise the concern about susceptibility to proteolysis
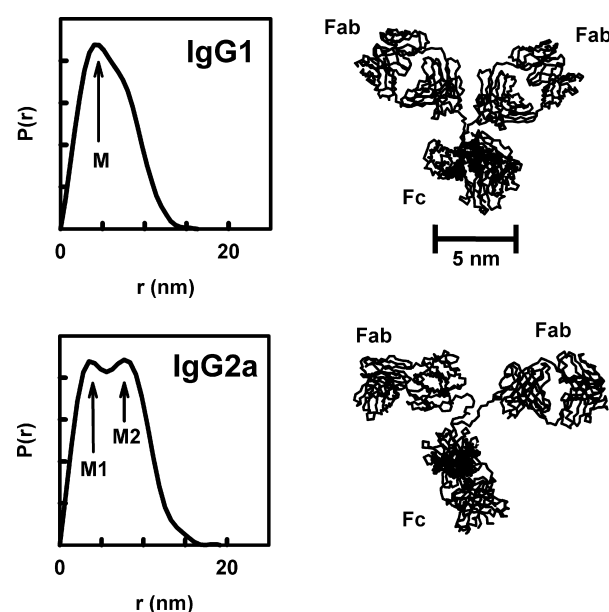

Figure 10. Theoretical distance distribution functions $P(r)$ for the crystal structures of murine IgG1 and IgG2a. The α-chain trace of each structure is shown to the right of each $P(r)$ curve. The $P(r)$ curves correspond to static Y-shaped IgG1 and T-shaped IgG2a structures. In solution, both are expected to be flexible and the $P(r)$ curve will change accordingly.

may not apply to secretory IgA, nor for IgA agglutinated on a bacterial surface.

The expected flexibility of the IgA hinge is qualified by the scattering evidence for relatively inflexible hinges in IgA that was obtained from the separate resolution of the *M*1 and *M*2 peaks in the X-ray $P(r)$ curves of IgA1 and IgA2 (Figure 3). In contrast, the neutron-scattering data for bovine IgG1/2 showed no sign of these two separate peaks,[28] nor were they seen in X-ray scattering data for human myeloma IgD (P.B.F. *et al.*, unpublished observations). The *M*2 peak would not have been detectable if the IgA hinge was flexible enough to exist in a range of both extended and compact (bent) conformations, as this would mean that the variable separations of the equally sized Fab and Fc fragments in IgA molecules would correspond to a broad spectrum of interatomic distances. This premise was examined further by theoretical calculations of the $P(r)$ curve[10] based on the intact mouse IgG1 and IgG2a crystal structures. This revealed two separate *M*1 and *M*2 peaks for the static T-shaped IgG2a structure, but less visible *M*1 and *M*2 peaks in the Y-shaped IgG1 structure (Figure 10). Hence, the observation of a single peak in the $P(r)$ curve for bovine IgG1/2 and human IgD means: (i) hinge flexibility has broadened the *M*2 peak beyond recognition; (ii) the solution structure is Y-shaped; or (iii) both explanations are correct. The extended IgA1 hinge seen by solution-scattering is consistent with electron microscopy observations of IgA1 dimers in which the hinges appear to be extended.[26] Other electron microscopy observations of monomeric IgA2 showed that its structure was less flexible than those found in a panel of IgG, IgE and IgM antibodies.[27] A functional rationale for this reduced hinge flexibility in IgA would be to ensure that the two Fab fragments are maintained in an extended conformation relative to the Fc fragment, while at the same time retaining enough residual flexibility to enable the two Fab fragments to access a wide range of conformations for optimal antigen binding. It should, however, be remembered that the higher degree of flexibility assigned to monomeric IgA1 compared to IgA2 might be altered upon IgA dimerisation or after secretory component binding.

Flexible T-shaped models for monomeric IgA

The modelling provides dimensions for the structure of IgA2m(1). In the 15 final IgA2m(1) models, the Fab-to-Fab centre-to-centre distance is 8.2 nm, which is much reduced compared to that of 16.9 nm for IgA1.[10] A T-shaped antibody has an inter-Fab angle of 180° and a Y-shaped antibody has one of 120°. The constrained scattering modelling showed that IgA2m(1) has a predominantly T-shaped arrangement in both the 104 best-fit models (Figure 8) and the final 15 disulphide-linked models. In contrast to this, good scattering fits for bovine IgG1/2 were obtained by fitting to
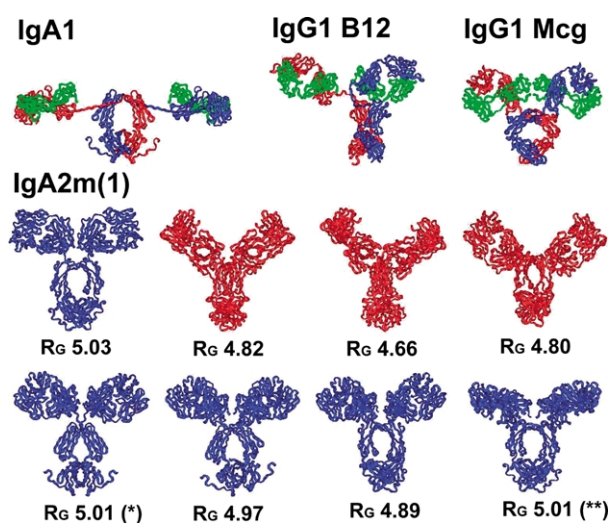

Figure 11. Comparison of the human IgA1 and IgA2m(1) structures determined by solution scattering modelling with the crystal structures of human IgG1 B12 and hinge-deleted IgG1 Mcg. In the top three images, the light chains are shown in green, and the heavy chains are shown in blue and red. For IgA2m(1), eight light chain disulphide-linked models are shown in blue if the model is T-shaped, and in red if Y-shaped, together with their calculated R_G values. For reason of the CNS refinement procedure for the disulphide bridge, these models have partially lost their 2-fold symmetry. The asterisk indicates the published IgA2m(1) model,[11] and the double asterisk indicates the model used for Figures 9 and 12.

a 2-fold symmetric Y-shaped arrangement similar to that seen in the crystal structure of IgG1 Kol.[28] By comparison, the six known crystal structures for intact human and mouse IgG antibodies also reveal both T-shaped and Y-shaped structures.[24,25] One of each of these types is shown in Figure 11. The most related one to IgA2m(1) is the hinge-deleted IgG1 Mcg structure, which has a 2-fold symmetric T-shaped structure that is stabilised by the same inter-light chain disulphide bridge found in IgA2m(1).[42] The hinge-deleted human IgG1 Dob also shows a 2-fold symmetric T-shaped structure, while the mouse IgG2a Mab231 shows an asymmetric distorted T-shaped structure. Three symmetric and asymmetric Y-shaped structures are represented by human IgG1 Kol, mouse IgG1 Mab61.1.3 and human IgG1 B12.

The IgA2m(1) modelling also provides insight into the hinge conformation in IgA. The IgA1 hinge has 23 residues with ten proline residues and is O-glycosylated (PVPSTPPTPSPSTPPTPSP SCCH), while the equivalent in IgA2m(1) has ten residues with six proline residues (PVPPPPPCCH). In the IgA1 hinge, the combination of its O-glycosylation and its proline residues (to reduce main-chain flexibility) with the absence of glycine residues (to act as main-chain conformational swivel points) is predicted to reduce the flexibility of this hinge compared to that in IgG. This prediction is based on Ramachandran $\phi-\psi$ plots for

Gly residues showing that these access 45% of the total rotational space about a peptide bond, hence leading to a high degree of conformational variability, while others such as Ala residues can only access 7.5% of this rotational space, and bulkier side-chains can access even less space.[43] Proline is a special case when only one ϕ value of 60° is allowed. In IgA2m(1), the combination of the Cys214−Cys214 inter-light chain disulphide bridge with the high abundance of proline residues, the absence of glycine residues, and the relatively short hinge length is predicted to reduce hinge flexibility compared to that in IgG. The existence of relatively inflexible hinges in IgA2 is not only indicated by the $P(r)$ curves (see above) but is also suggested from the curve fit modelling procedure, where only a single cluster of the more extended structures gives the best-fit structures. In comparison, the four crystal structures for human IgG1 and mouse IgG1 and IgG2a provided explanations for the more variable hinge conformations in these antibodies.[24,25] These hinges are relatively long with 17−23 residues, and contain glycine residues, relatively few proline residues, and no glycosylation.

Other implications of the IgA1 and IgA2m(1) structures

A speculative view of how fully extended monomeric extended T-shaped IgA1 and compact T-shaped IgA2 structures might interact with repeated epitopes on cell surfaces is shown in Figure 12. IgA also binds through its Fc fragment to the IgA receptor FcαRI. There are two FcαRI sites, one on each heavy chain, at a hydrophobic surface patch lying at the C_H2−C_H3 interface with contributions from both the C_H2 and C_H3 domains.[20,44,45] This site localisation is analogous to the Fc-binding site in IgG for bacterial proteins A
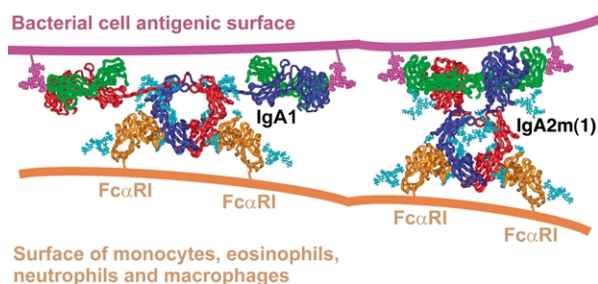

Figure 12. Simplified cartoon of the interaction of monomeric IgA1 and IgA2 with oligosaccharide epitopes on an antigenic bacterial surface (purple) to give an impression of the extended reach of IgA1 compared to IgA2. The simultaneous interaction with the FcαRI receptor on monocytes, eosinophils, neutrophils and macrophages (orange) is also shown. The IgA oligosaccharides are shown in light blue. The IgA structures are shown as T-shapes to follow the best-fit averaged solution structure determinations, however, the hinge region will show flexibility, and Y-shaped IgA structures are possible.

and G, rheumatoid factor, rat neonatal Fc receptor, and the Fc−III peptide.[20] Figure 12 shows that each Fc heavy chain site is accessible for interactions with FcαRI. This binding is compatible with the simultaneous binding of the same IgA molecule through its Fab fragments to its antigens, enabling the observed 2:1 FcαRI:IgA stoichiometry to take place.[20,38] In both IgA models, the location of the FcαRI site at the C_H2−C_H3 interface is well separated from the Fab fragments in their T-shaped arrangement, showing that the T-shaped structures are compatible with FcαRI receptor binding.

IgA2m(1) is mostly held intact only by non-covalent interactions between the V_H−V_L and C_H1−C_L domain pairs (Figure 1), because there are no heavy−light chain disulphide bridges.[46] In the presence of denaturants, dissociation into heavy and light chain dimers occurs.[47] Interestingly, minor forms of IgA2m(1) have been shown to possess heavy−light disulphide bridges.[11] At the top of the Fc fragment, four disulphide bridges can be potentially formed (Figure 1). The crystal structure of the IgA1 Fc fragment[20] did not reveal the Cys241−Cys241 bridge because Cys242 is the N-terminal residue of this recombinant protein. In the crystal structure, two disulphide bridges between two Cys242−Cys299 pairs were observed, and the two Cys301 side-chains were left unbound. It has been reported that there is less than one free −SH group in intact IgA1,[48] and this appears to be incompatible with the unbound Cys301 pair in the IgA1 Fc crystal structure. There is, however, the possibility that in IgA2m(1) an unbound Cys301 residue may bridge with the light chain Cys214, and this would offer an explanation of the minor forms with heavy−light bridges. Since the separation between the α-carbon atoms of Cys214 and Cys301 is 1.0 nm in the best-fit IgA2m(1) model, small conformational rearrangements in the hinge region would be sufficient for the formation of these minor heavy−light forms of IgA2m(1) if free Cys−SH residues exist.

IgA2m(1) is more heavily glycosylated (11% (w/w) carbohydrate) than either IgA1 (9.5% (w/w) carbohydrate), or IgG (2.9% (w/w) carbohydrate) (Figure 1). In the IgA2m(1) model, all four N-linked carbohydrate sites at Asn166, Asn262, Asn337 and Asn459 are solvent-exposed, unlike IgG where the single carbohydrate site at Asn297 is buried within the Fc fragment. The rotational orientation of the two Fab fragments is not constrained by the Cys214−Cys214 disulphide bridge, meaning that the position of the Asn166 oligosaccharide chain on the C_H1 domain relative to the Fc fragment remains indeterminate. The two oligosaccharide chains at Asn263 and Asn337 are on opposite sides of the Fc fragment, where Asn263 is near the C_H2−C_H3 junction, so steric access to the C_H2 surface will be affected by these. The IgA1 Fc crystal structure showed that the Asn263 oligosaccharide makes contact with the protein surface of both the C_H2 and C_H3 domains.[20]

Mutation of the attachment site at Asn263 to prevent glycan attachment did not reduce the affinity of the resultant protein for FcαRI.[13] This is as expected, since the other oligosaccharide at Asn337 is too far away to perturb the FcαRI-binding site.

Utility of constrained solution-scattering methods to study antibody conformations

The scattering modelling study of IgA2m(1) represents a third example of automated curve fit for antibodies based on randomising a hinge peptide conformation and constrained by known crystal structures.[10,29] This "random peptide" procedure is well suited for antibodies because this involves only variations in a single peptide conformation, and the outcome is sufficiently detailed to be useful. Solution-scattering does not yield a unique structure determination. However, if the scattering modelling identifies a single best-fit conformational family from randomised trials, this can be reasonably concluded to correspond to the averaged solution structure. Both IgA1 and IgA2m(1) gave a family of 104 related best-fit structures (Figure 8). The Crry-Ig chimeric antibody gave multiple structures that fitted the scattering data slightly less well.[29] With hindsight, we consider that these poorer Crry-Ig curve fits may result from excessive flexibility in the Crry-Ig hinge, meaning that all structures were possible, and it was not possible to represent this in terms of a single averaged conformational family. A single IgA2m(1) molecule in solution will be symmetric or asymmetric at any moment, as exemplified by the single snapshot views of the IgG crystal structures.[24,25] However, scattering only observes the averaged 2-fold symmetric solution structure. Consequently, an uncertainty in the modelling is the rotational orientation of the Fc fragment about its 2-fold symmetry axis relative to the pair of Fab fragments above it. This is illustrated by the different views of the Fc fragment when the Fab pairs are viewed face-on in the eight IgA2m(1) structures in Figure 11.

The constraints used for scattering modelling result in a medium-resolution structure determination with a precision of between 0.2 nm and 1 nm.[30] The actual precision depends on the protein of interest, and is assessed by the extent to which deviations from the best-fit structural models lead to worsened curve fits. The accuracy of the homology models used for the curve fits also requires consideration. After this study was completed,[21] the crystal structure of the IgA1 Fc fragment complex with the FcαRI receptor was published.[20] This provides a fortuitous independent measure of precision. The comparison of our modelled Fc fragment with the IgA1 Fc crystal structure confirmed the previous analysis that showed that the two IgA1 C_H2 domains had to be rotated towards each other to form the Cys241–Cys241 disulphide bridge.[10] The rmsd of the 200

α-carbon atoms in the C_H2 domains between the model and the crystal structure was 0.69 nm, in comparison to a rmsd of 0.24 nm when the 216 α-carbon atoms in the C_H3 domains were superimposed. The difference of 0.69 nm is within the estimated precision of the present scattering modelling of 1.0 nm. At this level of accuracy, domain arrangements can be modelled, but not the individual positions of the amino acid side-chains.

Materials and Methods

Preparation and composition of IgA2m(1) and IgA1

Recombinant anti-NIP IgA2m(1) was purified from the culture supernatant of transfected Chinese hamster ovary K1 cells on 4-hydroxy-3-iodo-5-nitrophenylacetyl (NIP)-Sepharose essentially as described[31] except that a batch washing procedure was used rather than column chromatography. Briefly, 5 ml of NIP-Sepharose was added to each litre of supernatant and the slurry was rotated overnight at 4 °C. The NIP-Sepharose was washed extensively in phosphate-buffered saline (PBS) (10 mM sodium phosphate (pH 7.2), 3 mM potassium phosphate, 136 mM NaCl, 3 mM KCl) using repeated rounds of sedimentation under gravity, removal of supernatant, and dispersion in fresh PBS. IgA2m(1) was eluted with 0.5 mM NIP and dialysed exhaustively against PBS, then concentrated in model 8003 stirred cell concentrators using PM10 membranes (Amicon Ltd, Stonehouse, UK). Alternatively, NIP was removed in a fast performance liquid chromatography (FPLC) purification step after concentration. Immediately prior to X-ray and neutron-scattering data collection, IgA2m(1) samples were applied to a Superdex 200 gel-filtration column using FPLC to remove non-specific aggregates. The buffer used for data collection was Dulbecco's PBS (12.5 mM sodium phosphate (pH 7.4), 140 mM NaCl) supplemented with EDTA and sodium azide as anti-bacterial preservatives (0.5 mM EDTA, 0.02% (w/v) NaN_3). X-ray experiments utilised the H_2O buffer eluate used in gel filtration, while neutron experiments involved dialysis at 6 °C into 100% ^2H_2O buffer for at least 36 hours with four buffer changes. Non-reducing SDS-PAGE confirmed that IgA2m(1) migrated principally as a light chain dimer and a heavy chain dimer together with some heavy–light chain dimers.[31,37]

For the initial data analyses of IgA2m(1), the α-chain amino acid sequence of IgA2m(1) for the C_H1, C_H2 and C_H3 domains was taken from SWISSPROT (database code P01877). The V_H, V_L and C_L sequences were initially taken from the crystal structure of human TR1.9 IgG1 Fab (PDB code 1vge), and included the human κ light chain. The V_H, V_L and C_L sequences for the final data analyses were taken from the crystal structure of murine N1G9 IgG1 Fab (PDB code 1ngq), for which the antigen is the hapten 4-hydroxy-3-nitrophenylacetate, and included the murine λ light chain. The mouse anti-NIP V_H sequence (SWISSPROT code J00529) is identical with the latter's V_H sequence. The IgA2m(1) carbohydrate composition was assumed to be similar to that for N-linked oligosaccharides in serum IgA1,[10] suggesting that the oligosaccharides at Asn169, Asn253, Asn327 and Asn449 were complex-type with a Man_3-$GlcNAc_2$ core and two NeuNAc.Gal.GlcNAc antennae. This chimeric murine–human composition resulted in a

calculated molecular mass of 163,420 Da, an unhydrated volume of 207.5 nm³, a hydrated volume of 274.2 nm³ (based on a hydration of 0.3 g H_2O/g of glycoprotein and an electrostricted volume of 0.0245 nm³ per bound water molecule) and an absorption coefficient at 280 nm (1%, 1 cm path length) of 14.8.[33]

Human myeloma IgA1 was purified from serum taken from myeloma patients using a combination of a modified thiophilic gel chromatography step to separate the immunoglobulins from other serum proteins in a single step together with jacalin-affinity chromatography, which is specific for the O-linked glycans in the IgA1 hinge.[37,49,50] Human serum IgA1 was purified by a similar route from normal individuals. Other details are as before.[10]

Synchrotron X-ray and neutron data collection

X-ray scattering data were obtained for recombinant IgA2m(1) in two sessions using the camera at Station 2.1 at the SRS at Daresbury Laboratory, Warrington, UK.[51] At the SRS, experiments utilised beam currents ranging from 118 mA to 179 mA with a ring energy of 2.0 GeV. Samples were measured at 15 °C in 13 runs at concentrations of 0.42 mg/ml, 0.50 mg/ml and 0.83 mg/ml for ten minutes using ten time-frames. Data acquisitions utilised an area detector with the main beam and beam stop located at one edge of this. Data reduction employed a detector sector between 30° and 150°, which resulted in 300 $I(Q)$ intensities for use in the Guinier analyses at low Q values (Q range of 0.08–1.01 nm⁻¹). A second data reduction used a detector sector between 60° and 120°, which resulted in 500 $I(Q)$ intensities in a fuller range of Q values (Q range of 0.10–2.25 nm⁻¹), and these were used both for the $P(r)$ calculations and the modelling fits. Other details including the data reduction procedure are as described.[10]

X-ray scattering data were also obtained in two sessions on the Beamline ID02 at the ESRF, Grenoble, France.[52] At ESRF, experiments were performed in single-bunch mode to reduce the incident X-ray flux. Storage ring currents ranged from 9 mA to 17 mA with a ring energy of 6.0 GeV. Samples were measured at 15 °C at five concentrations of 0.55, 0.56, 0.82, 0.84 and 1.12 mg/ml in ten time-frames, each exposure lasting for one second or two seconds in Perspex cells of path thickness 1 mm and using 25 μm thick mica windows. Slight radiation damage was observed when determining the R_G value from Guinier fits of the ten merged time-frames; accordingly, only the first time-frame was used to avoid its effect. The determinations of the R_{XS-1} and R_{XS-2} values were unaffected by this slight radiation damage, and all ten time-frames averaged together were used for these analyses. Data reduction included the calibration of detector intensities using a Lupolen sample, and the calibration of the Q range using a silver salt standard.

Neutron-scattering data for recombinant IgA2m(1) were obtained in three different sessions on Instrument LOQ at the pulsed neutron source ISIS at the Rutherford Appleton Laboratory, Didcot, UK.[53] The pulsed neutron beam was derived from proton beam currents of 167–200 μA. Data acquisitions were for fixed totals of 40×10^6–300×10^6 monitor counts in runs lasting 0.5–3.5 hours each for protein in two concentrations of 2.0 mg/ml and 3.0 mg/ml measured at 15 °C. Other

details including the data reduction procedure are as described.[10]

Analysis of reduced X-ray and neutron data

In a given solute–solvent contrast, the radius of gyration R_G is a measure of structural elongation if the internal inhomogeneity of scattering densities within the glycoprotein has no effect. Guinier analyses at low Q give the R_G, and the forward scattering at zero angle $I(0)$:[54]

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

This expression is valid in a QR_G range of up to 1.5. The relative $I(0)/c$ values ($c =$ sample concentration) for samples measured in the same buffer during a data session gives the relative molecular mass of the proteins when referenced against a suitable standard.[10] If the structure is elongated, the mean radius of gyration of cross-sectional structure R_{XS} and the mean cross-sectional intensity at zero angle $[I(Q)Q]_{Q\to0}$ can be obtained from:

$$\ln[I(Q)Q] = [I(Q)Q]_{Q\to0} - R_{XS}^2 Q^2 / 2$$

For immunoglobulins, the cross-sectional plot exhibits two regions, a steeper innermost one and a flatter outermost one,[55] and the two analyses are identified by R_{XS-1} and R_{XS-2}, respectively. The R_G and R_{XS-1} analyses lead to the triaxial dimensions of the macromolecule if the structure can be represented by an elliptical cylinder, where $L = \sqrt{12(R_G^2 - R_{XS-1}^2)}$ and L is its length.[54] The R_G and R_{XS} analyses were performed using an interactive graphics program SCTPL5 on a Silicon Graphics 4D35S Workstation (A. S. Nealis *et al.*, unpublished software) and a PERL script program SCTPL7 (J. T. Eaton & S. J. Perkins, unpublished software) on Silicon Graphics Indy and O2 Workstations.

Indirect transformation of the scattering data $I(Q)$ in reciprocal space into real space to give $P(r)$ was carried out using the GNOM program:[56]

$$P(r) = \frac{1}{2\pi^2} \int_0^\infty I(Q)Qr \sin(Qr) \mathrm{d}Q$$

$P(r)$ corresponds to the distribution of distances r between volume elements. This offers an alternative calculation of R_G and $I(0)$ which is now based on the full range of the scattering curve, and also gives the maximum dimension L. This was used with the IgA2m(1) X-ray curve in the Q range between 0.10 nm⁻¹ and 2.20 nm⁻¹ and the IgA2m(1) neutron curve in the Q range between 0.13 nm⁻¹ and 2.20 nm⁻¹. A range of D_{max} values was tested, and the final choice of D_{max} was based on three criteria: (i) $P(r)$ should exhibit positive values; (ii) the R_G from GNOM should agree with the R_G from Guinier analyses; (iii) the $P(r)$ curve should be stable as D_{max} was increased beyond the estimated macromolecular length.

Sedimentation equilibrium and sedimentation velocity data for IgA2m(1)

Analytical ultracentrifugation was performed at 20 °C on a Beckman XL-I instrument in which the IgA2m(1) concentration was monitored using its absorbance at 280 nm and its refractive index measured by interferometry. Sedimentation equilibrium data sets for samples of IgA2m(1) at concentrations of 0.13, 0.24, 0.69,

0.87 and 1.11 mg/ml were each acquired over 45 hours using six-sector cells with solution column heights of 2 mm at rotor speeds of 11,000, 14,000 and 17,000 rpm until equilibrium had been reached at each speed as shown by the perfect overlay of runs measured at five hours intervals. The molecular mass M was analysed on the basis of a single species using Beckman software provided as an addon to Microcal Origin Version 4.1, where the partial specific volume $\bar{v}$ for IgA2m(1) was calculated to be 0.722 ml/g from its sequence:[33]

$$c_r = c_{r_0} \exp[(\omega^2/2RT)M(1 - \bar{v}\rho)(r^2 - r_0^2)]$$

where c_r is the concentration at radius r, c_{r_0} is the concentration of the monomer at the reference radius r_0, ω is the angular velocity, R is the gas constant, T is the temperature (Kelvin), and ρ is the solvent density of the buffer. The buffer density was calculated as 1.00543 g/ml from the program SEDNTERP.[57]

Sedimentation velocity data for IgA2m(1) were acquired over 16 hours at rotor speeds of 15,000, 20,000, 25,000, 30,000, 35,000 and 42,000 rpm in two-sector cells with column heights of 12 mm. Concentrations were 0.28, 0.46, 0.52, 0.69, 1.05 and 1.12 mg/ml. For the $g(s^*)$ analyses, successive absorbance and interference scans were recorded at ten minute intervals, the shortest interval possible under standard measurement conditions, for which the rotor speed of 20,000 rpm was optimal for IgA2m(1). In time-derivative analyses, the subtraction of pairs of concentration scans *versus* radius in the cell eliminates systematic errors from baseline distortions in the cell windows and permits the averaging of many pairs of subtractions. The extrapolation of individual subtractions to the start time gives the $g(s^*)$ function, which was computed using DCDT+,[58] from which the sedimentation and diffusion coefficients (and hence the molecular mass) were determined from the peak position and width, respectively (Figure 5). Between eight and 12 scans were used in order to ensure that the maximum permissible measurable molecular mass calculated from these was over 16 times the expected value,[58] meaning that time-broadening effects were negligible. Increased numbers of scans were used to improve the signal–noise ratio to verify the sedimentation coefficient determination $s_{20,w}^0$. To avoid time-broadening, direct boundary fitting to 120 raw absorbance and interference scans was performed using SEDFIT.[59] Initially. the $g(s^*)$ distribution and a size-distribution analysis were performed to verify the absence of any other sedimentating species in the sample. Final analyses utilised the non-interacting species model based on a single species, which was optimised until the root-mean-square deviation (rmsd) values and visual appearance of the fits were satisfactory.

Homology modelling of the human IgA2m(1) α-chain domains

For the modelling, the chimera composition was taken from the human IgA2m(1) sequence,[10] with the replacement mouse V_H, V_L and C_L sequences being taken from the crystal structure of murine N1G9 IgG1 Fab (PDB code: 1ngq[60]), for which the antigen is the hapten 4-hydroxy-3-nitrophenylacetate, and included the murine λ light chain. The IgA2m(1) carbohydrate composition was assumed to be similar to that for the N-linked oligosaccharides in IgA1,[10–13] suggesting that the oligosaccharides at Asn169, Asn253, Asn327 and

Asn449 were complex-type with a $Man_3GlcNAc_2$ core and two NeuNAc.Gal.GlcNAc antennae.

The modelling of IgA2m(1) utilised the previous multiple sequence alignment of human α-chain sequences and averaged secondary structure and solvent accessibility predictions.[10] This choice was justified by the similarity of the molecular masses calculated from the mouse and human Fab fragment sequences of Figure 6(a) (murine 163,400 Da; human 163,600 Da) and also their dry volumes (murine 207.5 nm^3; human 207.8 nm^3).[33] The sequence numbering followed that for IgA1 (Figure 6(a)) for ease of comparison.[10] The homology model for the Fab and Fc fragments in IgA2m(1) was constructed from that created and validated for IgA1, including the folded-back structure for the tailpiece.[10] This utilised INSIGHT II 98.0 molecular graphics software with the BIOPOLYMER, DISCOVER, DISCOVER3, HOMOLOGY and DELPHI modules (Accelrys, San Diego, CA, USA) on Silicon Graphics O2 and OCTANE Workstations. The IgA2m(1) V_H, V_L and C_L domains were used unchanged from the crystal structure of the human TR1.9 IgG1 Fab fragment (PDB code 1vge[61]). The IgA2m(1) C_H1 model originated from the C_H1 domain in the crystal structure of murine IgA J539 Fab fragment (PDB code 2fbj[19]). The IgA2m(1) C_H2 and C_H3 models originated from the C_H2 and C_H3 domains in the crystal structure of human IgG1 Fc fragment (PDB code 1fc1-A[22]). A total of 13 amino acid residue changes were made in the IgA1 model to convert this to IgA2m(1), five in the IgA1 C_H1 domain, and eight in the IgA1 C_H2 domain (including the introduction of an extra N-linked oligosaccharide site at Asn337: Figure 1). The disulphide bridge connectivity in the Fc fragment of IgA2m(1) was retained from that used in the IgA1 model, in which the two C_H2 domains were rotated equally towards each other to enable four S–S bridges to be formed, namely the generally accepted Cys241–Cys241 bridge, together with an assumed Cys299–Cys299 bridge and two assumed Cys242–Cys301 bridges (asterisked in Figure 6). The stereochemical validity of the resulting IgA2m(1) homology model was confirmed using PROCHECK[62] and the Protein Data Bank deposition procedures.

Automated modelling of the IgA2m(1) solution structure

An automated search strategy was used to model the IgA2m(1) structure, in which a vertical 2-fold symmetry axis centred on the Fc fragment (Figure 1) defined the positions of both Fab fragments relative to the Fc fragment. Structurally, each model was defined by the distance between the centres of mass of the two Fab fragments and that between the Fab and Fc fragment. The hinge length of each model was defined as the distance between the α-carbon atoms of Cys220 and Pro240. The preparation of the models and their scattering curve fits was adapted from Method 2 used for the structure determination of IgA1,[10] described as follows. The homology models for the Fab and Fc fragments of IgA2m(1) were combined with 10,000 structures for the 12-residue hinge peptide 220-CPVPPPPPCCHP-244 (Figure 6(a)). This peptide was set up as a β-strand of maximum length 3.54 nm using BIOPOLYMER. A molecular dynamics simulation using DISCOVER3 created random hinge structures from this starting model. In order to generate a wide range of structures, the simulation was performed at a temperature of 773 K. First,

Table 3. Definition of IgA2m(1) rigid bodies used for final minimisation

Rigid body	Residues[a]	Corresponding to
1	Pro244 to Tyr472 (A and B)	Fc fragment
2	Gln1 to Cys220 (A) and Glu1 to Cys214 (L)	First Fab fragment
3	Gln1 to Cys220 (B) and Glu1 to Cys214 (M)	Second Fab fragment
4–13	Pro221 (A) to His243 (A)	Hinge peptide
14–23	Pro221 (B) to His243 (B)	Hinge peptide

[a] In the first Fab fragment, the heavy chain is denoted as A and the light chain is denoted as L. In the second Fab fragment, the chains are denoted as B and M, respectively.

the peptide structure was subjected to energy minimisation for 300 iterations. After a temperature equilibration step of 5000 fs, the simulation was run at 773 K for 1,000,000 fs, capturing models every 100 fs to produce 10,000 of these. Each of the 10,000 complete IgA2m(1) models was produced by a fragment assembly method. Holding the Fc fragment fixed, two copies of each hinge model were added to it by superimposing the four main-chain atoms of Pro240 in the hinge onto those of Pro240 in the Fc fragment. Likewise, the Fab fragments were now added to the hinge peptides by superimposing the main-chain atoms of Cys220 in the hinge and the fragment. The duplicate Pro240 and Cys220 residues were deleted from the model. N-Linked oligosaccharide structures were added to Asn166, Asn263, Asn337 and Asn459 to complete the glycosylated structure (Figure 6(a)).

At the end of the automated search for best-fit models, the 50 best-fit IgA2m(1) models were subjected to energy minimisation in order to form connected interchain disulphide bridges in these models. This procedure used the CNS software.[63] Initially, structure files were generated for each IgA2m(1) model defining the disulphide bonds within the IgA2m(1) molecule. These were defined as the inter-light chain bridge Cys214–Cys214 and the inter-heavy chain bridges Cys241–Cys241 and Cys299–Cys299. The two internal light chain disulphide bridges Cys23–Cys88 and Cys134–Cys194 and the six internal heavy chain disulphide bridges Cys23–Cys97 (V_H), Cys145–Cys204 and Cys196–Cys220 (C_H1), Cys242–Cys301 and Cys266–Cys323 (C_H2), and Cys369–Cys432 (C_H3) were also defined likewise for CNS (Figure 1). Subsequently, 30 rounds of ten cycles of Powell minimisation followed by 100 cycles of rigid body minimisation were carried out on each of the 50 best-fit models. The rigid bodies are defined in Table 3. For the two hinge peptides joining the Fab fragment to the Fc fragment, each residue was defined as an individual rigid body to allow complete freedom of movement of both the linkers and the adjoining Fab fragments in order to achieve the disulphide connections. At the end, each model was subjected to a final round of 250 cycles of Powell minimisation.

Debye scattering curve calculations from sphere models of IgA2m(1)

Each IgA2m(1) model was used to calculate X-ray and neutron-scattering curves for comparison with the experimental curves. Each set of atomic coordinates for a model was placed within a three-dimensional grid of cubes. A sphere of equal volume to the cube was placed at the centre of each cube if a user-specified cutoff for the minimum number of atoms contained within a cube was satisfied. For IgA2m(1), a cube side length of 0.534 nm in combination with a cutoff of four atoms con-

sistently produced sphere models within 2% of the total dry volume of 208 nm³ calculated from its composition. The optimal total of dry spheres for IgA2m(1) was 1362 (bottom row of Table 2). Since the hydration shell surrounding glycoproteins is detected by X-ray scattering, the sphere models were accordingly adapted by adding spheres to the surface of the dry models using HYPRO[64] based on 0.3 g H_2O/g of glycoprotein. The optimal total of hydrated spheres for IgA2m(1) is 1800 (bottom row of Table 2).

The X-ray and neutron-scattering curve $I(Q)$ was calculated assuming a uniform scattering density for the spheres using the Debye equation as adapted to spheres:[65]

$$\frac{I(Q)}{I(0)} = g(Q)\left(n^{-1} + 2n^{-2}\sum_{j=1}^{m} A_j \frac{\sin Qr_j}{Qr_j}\right)$$

$$g(Q) = (3(\sin QR - QR \cos QR))^2/Q^6R^6$$

where $g(Q)$ is the squared form factor for the sphere of radius r, n is the number of spheres filling the body, A_j is the number of distances r_j for that value of j, r_j is the distance between the spheres, and m is the number of different distances r_j. Other details, including those of calibration studies used to validate this approach, are given elsewhere.[10,34] X-ray curves were calculated from the hydrated sphere models without corrections for wavelength spread or beam divergence, as these are considered to be negligible for synchrotron X-ray data. Neutron curves were calculated from the dry sphere models and corrected using a 10% wavelength spread $\Delta\lambda/\lambda$ for a nominal λ of 0.6 nm and a beam divergence of 0.016 rad.[64] The number of spheres N in the dry and hydrated models after grid transformation was used to assess steric overlap between the Fab and Fc fragments, where models showing less than 95% of the optimal total were discarded (white circles in Figure 7). The models were assessed by calculation of the X-ray and neutron R_G and R_{XS-1} values in the same Q ranges used for the experimental Guinier fits, where values that exceeded specified ranges of ±5% resulted in the rejection of these models. Models that passed the filters were then ranked using a goodness-of-fit R-factor defined by analogy with protein crystallography and based on the experimental curves, in the Q range extending to 2.0 nm⁻¹ (denoted as R^{32}). A flat background was applied to the final neutron-scattering curve fits in Figure 9 to allow for incoherent scattering contributions from residual protons in the sample (J. T. E. & S. J. P., unpublished calculations). The sedimentation coefficient $s^0_{20,w}$ was calculated directly from the IgA2m(1) models derived from the 104 best-fit scattering curve fits using the program GENDIA.[66,67] For reason of computational memory limitations in GENDIA, a larger cube side

length of 0.834 nm and a cutoff of nine atoms was used in place of 0.534 nm during the creation of spheres, and the models were hydrated using HYPRO to give about 380 spheres from which the $s^0_{20,w}$ values were calculated.[64] The $s^0_{20,w}$ values were also calculated using the program HYDRO as a confirmation of the GENDIA analysis.[68] In the controls, the cube side was 0.480 nm and the cutoff was two atoms for the IgA1 Fc calculations, and 0.532 nm and three atoms for the FcαRI calculations.

Protein Data Bank accession numbers

The IgA2m(1) α-carbon co-ordinate model has been deposited in the Protein Data Bank with the accession number 1R70.

References

1. Kerr, M. A. (1990). The structure and function of human IgA. *Biochem. J.* **271**, 285–296.
2. van Egmond, M., Damen, C. A., van Spriel, A. B., Vidarsson, G., van Garderen, E. & van de Winkel, J. G. J. (2001). IgA and the IgA Fc receptor. *Trends Immunol.* **22**, 205–211.
3. Underdown, B. J. & Schiff, J. M. (1986). Immunoglobulin A: strategic defense initiative at the mucosal surface. *Annu. Rev. Immunol.* **4**, 389–417.
4. Mazanec, M. B., Nedrud, J. G., Kaetzel, C. S. & Lamm, M. E. (1993). A three-tiered view of the role of IgA in mucosal defense. *Immunol. Today,* **14**, 430–435.
5. Morton, H. C., van Egmond, M. & van de Winkel, J. G. J. (1996). Structure and function of human IgA Fc receptors (FcαR). *Crit. Rev. Immunol.* **16**, 423–440.
6. Kerr, M. A. & Woof, J. M. (1999). Fcα receptors. In *Mucosal Immunology* (Ogra, P. L., Mestecky, J., Lamm, M. E., Strober, W., McGhee, J. R. & Bienenstock, J., eds), 2nd edit., pp. 213–224, Academic Press Inc., San Diego.
7. van Egmond, M., van Garderen, E., van Spriel, A. B., Damen, C. A., van Amersfoort, E. S., van Zandbergen, G. *et al.* (2000). FcαRI-positive liver Kupffer cells: reappraisal of the function of immunoglobulin A in immunity. *Nature Med.* **6**, 680–685.
8. Russell, M. W., Kilian, M. & Lamm, M. E. (1999). Biological activities of IgA. In *Mucosal Immunology* (Ogra, P. L., Mestecky, J., Lamm, M. E., Strober, W., McGhee, J. R. & Bienenstock, J., eds), 2nd edit., pp. 225–240, Academic Press Inc., San Diego.
9. Sumiyama, K., Saitou, N. & Ueda, S. (2002). Adaptive evolution of the IgA hinge region in primates. *Mol. Biol. Evol.* **19**, 1093–1099.
10. Boehm, M. K., Woof, J. M., Kerr, M. A. & Perkins, S. J. (1999). The Fab and Fc fragments of IgA1 exhibit a different arrangement from that in IgG: a study by X-ray and neutron solution scattering and homology modelling. *J. Mol. Biol.* **286**, 1421–1447.
11. Chintalacharuvu, K. R., Raines, M. & Morrison, S. L. (1994). Divergence of human α-chain constant region gene sequences. A novel recombinant α2 gene. *J. Immunol.* **152**, 5299–5304.
12. Endo, T., Mestecky, J., Kulhavy, R. & Kobata, A. (1994). Carbohydrate heterogeneity of human myeloma proteins of the IgA1 and IgA2 subclasses. *Mol. Immunol.* **31**, 1415–1422.
13. Mattu, T. J., Pleass, R. J., Willis, A. C., Kilian, M., Wormald, M. R., Lellouch, A. C. *et al.* (1998). The glycosylation and structure of human serum IgA1, Fab and Fc regions and the role of N-glycosylation on Fcα receptor interactions. *J. Biol. Chem.* **273**, 2260–2272.
14. Senior, B. W., Loomes, L. M. & Kerr, M. A. (1991). Microbial IgA proteases and virulence. *Rev. Med. Microbiol.* **2**, 200–207.
15. Kilian, M., Reinholdt, J., Lomholt, H., Poulsen, K. & Frandsen, E. V. G. (1996). Biological significance of IgA1 proteases in bacterial colonization and pathogenesis: critical evaluation of experimental evidence. *APMIS,* **104**, 321–338.
16. Rudd, P. M., Fortune, F., Patel, T., Parekh, R. B., Dwek, R. A. & Lehner, T. (1994). A human T-cell receptor recognises O-linked sugars from the hinge region of human IgA1 and IgD. *Immunology,* **83**, 99–106.
17. Monteiro, R. C., Moura, I. C., Launay, P., Tsuge, T., Haddad, E., Benhamou, M. *et al.* (2002). Pathogenic significance of IgA receptor interactions in IgA nephropathy. *Trends Mol. Med.* **8**, 464–468.
18. Mantis, N. J., Cheung, M. C., Chintalacharuvu, K. R., Rey, J., Corthesy, B. & Neutra, M. R. (2002). Selective adherence of IgA to murine Peyer's patch M cells: evidence for a novel IgA receptor. *J. Immunol.* **169**, 1844–1851.
19. Suh, S. W., Bhat, T. N., Navia, M. A., Cohen, G. H., Rao, D. N., Rudikoff, S. & Davies, D. R. (1986). The galactan-binding immunoglobulin Fab J539. An X-ray diffraction study at 2.6 Å resolution. *Proteins: Struct. Funct. Genet.* **1**, 74–80.
20. Herr, A. B., Ballister, E. R. & Bjorkman, P. J. (2003). Insights into IgA-mediated immune responses from the crystal structures of human FcαRI and its complex with IgA1-Fc. *Nature,* **423**, 614–620.
21. Perkins, S. J., Furtado, P. B., Robertson, A., Whitty, P. W., Woof, J. M. & Kerr, M. A. (2002). Structure determinations of new forms of IgA antibodies by solution scattering. In *ISIS Experimental Report RB13103*, Rutherford Appleton Laboratory, Didcot, UK.
22. Deisenhofer, J. (1981). Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9 and 2.8 Å resolution. *Biochemistry,* **20**, 2361–2370.
23. Wan, T., Beavil, R. L., Fabiane, S. M., Beavil, A. J., Sohi, M. K., Keown, M. *et al.* (2002). The crystal structure of IgE Fc reveals an asymmetrically bent conformation. *Nature Immunol.* **3**, 681–686.
24. Saphire, E. O., Stanfield, R. L., Crispin, M. D. M., Parren, P. W. H. I., Rudd, P. M., Dwek, R. A. *et al.* (2002). Contrasting IgG structures reveal extreme asymmetry and flexibility. *J. Mol. Biol.* **319**, 9–18.
25. Harris, L. J., Larson, S. B. & McPherson, A. (1999).

Comparison of intact antibody structures and the implications for effector function. *Advan. Immunol.* **72**, 191–208.

26. Feinstein, A., Munn, E. A. & Richardson, N. E. (1971). The three-dimensional conformation of γM and γA globulin molecules. *Ann. N. Y. Acad. Sci.* **190**, 104–121.

27. Roux, K. H., Strelets, L., Brekke, O. H., Sandlie, I. & Michaelsen, T. E. (1998). Comparisons of the ability of human IgG3 hinge mutants, IgM, IgE and IgA2, to form small immune complexes: a role for flexibility and geometry. *J. Immunol.* **161**, 4083–4090.

28. Mayans, M. O., Coadwell, W. J., Beale, D., Symons, D. & Perkins, S. J. (1995). Demonstration by pulsed neutron scattering that the arrangement of the Fab and Fc fragments in the overall structures of bovine IgG1 and IgG2 in solution is similar. *Biochem. J.* **311**, 283–291.

29. Aslam, M., Guthridge, J. M., Hack, B. K., Quigg, R. J., Holers, V. M. & Perkins, S. J. (2003). The extended multidomain solution structures of the complement protein Crry and its chimeric conjugate Crry-Ig by scattering, analytical ultracentrifugation and constrained modelling: implications for function and therapy. *J. Mol. Biol.* **329**, 525–550.

30. Perkins, S. J., Ashton, A. W., Boehm, M. K. & Chamberlain, D. C. (1998). Molecular structures from low angle X-ray and neutron scattering studies. *Int. J. Biol. Macromol.* **22**, 1–16.

31. Morton, H. C., Atkin, J. D., Owens, R. J. & Woof, J. M. (1993). Purification and characterization of chimeric human IgA1 and IgA2 expressed in COS and Chinese hamster ovary cells. *J. Immunol.* **151**, 4743–4752.

32. Beavil, A. J., Young, R. J., Sutton, B. J. & Perkins, S. J. (1995). Bent domain structure of recombinant human IgE-Fc in solution by X-ray and neutron scattering in conjunction with an automated curve fitting procedure. *Biochemistry,* **34**, 14449–14461.

33. Perkins, S. J. (1986). Protein volumes and hydration effects: the calculation of partial specific volumes, neutron scattering matchpoints and 280 nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur. J. Biochem.* **157**, 169–180.

34. Perkins, S. J. (2001). X-ray and neutron scattering analyses of hydration shells: a molecular interpretation based on sequence predictions and modelling fits. *Biophys. Chem.* **93**, 129–139.

35. Björk, I. & Lindh, E. (1974). Gross conformation of human secretory immunoglobulin A and its component parts. *Eur. J. Biochem.* **45**, 135–145.

36. Chintalacharuvu, K. R. & Morrison, S. L. (1996). Residues critical for H-L disulphide bond formation in human IgA1 and IgA2. *J. Immunol.* **157**, 3443–3449.

37. Almogren, A. (2003). Structural and functional analysis of different forms of IgA1 and IgA2 purified from human serum or colostrum. PhD thesis, University of Dundee.

38. Herr, A. B., White, C. L., Milburn, C., Wu, C. & Bjorkman, P. J. (2003). Bivalent binding of IgA1 to FcαRI suggests a mechanism for cytokine activation of IgA phagocytosis. *J. Mol. Biol.* **327**, 645–657.

39. Almogren, A., Senior, B. W., Loomes, L. M. & Kerr, M. A. (2003). Structural and functional consequences of cleavage of human secretory and human serum immunoglobulin A1 by proteinases from *Proteus mirabilis* and *Neisseria meningitides.* *Infect. Immun.* **71**, 3349–3356.

40. Batten, M. R., Senior, B. W., Kilian, M. & Woof, J. M. (2003). Amino acid sequence requirements in the hinge of human immunoglobulin A1 (IgA1) for cleavage by streptococcal IgA1 proteases. *Infect. Immun.* **71**, 1462–1469.

41. Weiser, J. N., Bae, D., Fasching, C., Scamurra, R. W., Ratner, A. J. & Janoff, E. N. (2003). Antibody-enhanced pneumococcal adherence requires IgA1 protease. *Proc. Natl Acad. Sci. USA,* **100**, 4215–4220.

42. Guddat, L. W., Herron, J. N. & Edmundson, A. B. (1993). Three-dimensional structure of a human immunoglobulin with a hinge deletion. *Proc. Natl Acad. Sci. USA,* **90**, 4271–4275.

43. Creighton, T. E. (1993). *Proteins, Structures and Molecular Properties,* 2nd edit., pp. 172–176, W. H. Freeman & Co, New York.

44. Carayannopoulos, L., Hexham, J. M. & Capra, J. D. (1996). Localisation of the binding site for the monocyte immunoglobulin (Ig) A-Fc receptor (CD89) to the domain boundary between Cα2 and Cα3 in human IgA1. *J. Expt. Med.* **183**, 1579–1586.

45. Pleass, R. J., Dunlop, J. I., Anderson, C. M. & Woof, J. M. (1999). Identification of residues in the C_H2/C_H3 domain interface of IgA essential for interaction with the human Fcα receptor (FcαRI) CD89. *J. Biol. Chem.* **274**, 23508–23514.

46. van Loghem, E. & Biewenga, J. (1983). Allotypic and isotypic aspects of immunoglobulin A. *Mol. Immunol.* **20**, 1001–1007.

47. Jerry, L. M. & Kunkel, H. G. (1974). Special characteristics of the IgA2 subclass. In *The Immunoglobulin A System* (Mestecky, J. & Lawton, A. R., eds), pp. 151–160, Plenum Press, New York.

48. Biewenga, J. & van Run, P. E. M. (1992). Effects of limited reduction on disulphide bonds in human IgA1 and IgA1 fragments. *Mol. Immunol.* **29**, 327–334.

49. Loomes, L. M., Stewart, W. W., Mazengera, R. L., Senior, B. W. & Kerr, M. A. (1991). Purification and characterisation of human immunoglobulin IgA1 and IgA2 isotypes from serum. *J. Immunol. Methods,* **141**, 209–218.

50. Kerr, M. A., Loomes, L. M., Bonner, B. C., Hutchings, A. B. & Senior, B. W. (1997). Purification and characterization of human serum and secretory IgA1 and IgA2 using jacalin. *Methods Mol. Med.* **9**, 265–278.

51. Towns-Andrews, E., Berry, A., Bordas, J., Mant, G. R., Murray, P. K., Roberts, K. *et al.* (1989). Time-resolved X-ray diffraction station: X-ray optics, detectors and data acquisition. *Rev. Sci. Instrum.* **60**, 2346–2349.

52. Narayanan, T., Diat, O. & Bosecke, P. (2001). SAXS and USAXS on the high brilliance beamline at the ESRF. *Nucl. Instrum. Methods Phys. Res. A,* **467–468**, 1005–1009.

53. Heenan, R. K. & King, S. M. (1993). Development of the small-angle diffractometer LOQ at the ISIS pulsed neutron source. In *Proceedings of an International Seminar on Structural Investigations at Pulsed Neutron Sources, Dubna, 1st–4th September 1992. Report E3-93-65,* Joint Institute for Nuclear Research, Dubna.

54. Glatter, O.; Kratky, O. (1982). Editors of *Small-angle X-ray Scattering,* Academic Press, New York.

55. Pilz, I., Kratky, O., Licht, A. & Sela, M. (1973). Shape and volume of anti-poly(D-alanyl) antibodies in the presence and absence of tetra-D-alanine as followed by small-angle X-ray scattering. *Biochemistry,* **12**, 4998–5005.

56. Semenyuk, A. V. & Svergun, D. I. (1991). GNOM–a

program package for small-angle scattering data-processing. *J. Appl. Crystallog.* **24**, 537–540.

57. Laue, T. M., Shah, B. D., Ridgeway, T. M. & Pelletier, S. L. (1992). Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J. & Horton, J. C., eds), pp. 90–125, The Royal Society of Chemistry, Cambridge, UK.

58. Philo, J. (2000). A method for directly fitting the time derivative of sedimentation velocity data and an alternative algorithm for calculating sedimentation coefficient distribution functions. *Anal. Biochem.* **279**, 151–163.

59. Schuck, P. (1998). Sedimentation analysis of non-interacting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys. J.* **75**, 1503–1512.

60. Mizutani, R., Miura, K., Nakayama, T., Shimada, I., Arata, Y. & Satow, Y. (1995). Three-dimensional structures of the Fab fragment of murine N1G9 antibody from the primary immune response and of its complex with (4-hydroxy-3-nitrophenyl)acetate. *J. Mol. Biol.* **254**, 208–222.

61. Chacko, S., Padlan, E. A., Portolano, S., McLachlan, S. M. & Rapoport, B. (1996). Structural studies of human autoantibodies-crystal-structure of a thyroid peroxidase autoantibody Fab. *J. Biol. Chem.* **271**, 12191–12198.

62. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

63. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.

64. Ashton, A. W., Boehm, M. K., Gallimore, J. R., Pepys, M. B. & Perkins, S. J. (1997). Pentameric and decameric structures in solution of the serum amyloid P component by X-ray and neutron scattering and molecular modelling analyses. *J. Mol. Biol.* **272**, 408–422.

65. Perkins, S. J. & Weiss, H. (1983). Low resolution structural studies of mitochondrial ubiquinol-cytochrome c reductase in detergent solutions by neutron scattering. *J. Mol. Biol.* **168**, 847–866.

66. Garcia de la Torre, J. & Bloomfield, V. A. (1977). Hydrodynamic properties of macromolecular complexes. I. Translation. *Biopolymers*, **16**, 1747–1761.

67. Garcia de la Torre, J. & Bloomfield, V. A. (1977). Hydrodynamics of macromolecular complexes. III. Bacterial viruses. *Biopolymers*, **16**, 1779–1793.

68. Garcia de la Torre, J., Navarro, S., Martinez, M. C. L., Diaz, F. G. & Cascales, J. L. (1994). HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules. *Biophys. J.* **67**, 530–531.

Edited by M. Moody