Self-association and domain rearrangements between complement C3 and C3u provide insight into the activation mechanism of C3

Keying Li, Jayesh Gor, Stephen J Perkins

Self-association and domain rearrangements between complement C3 and C3u provide insight into the activation mechanism of C3

Keying Li[1], Jayesh Gor[1] and Stephen J. Perkins[1*]

[1]Department of Structural and Molecular Biology,
Division of Biosciences
Darwin Building,
University College London,
Gower Street,
London WC1E 6BT, U. K.

Running Title: Solution structure of C3 and C3u
Keywords: Complement; X-ray scattering; modelling; ultracentrifugation
Abbreviations: C3, third component of complement.

* Author to whom correspondence and requests for reprints should be addressed (Tel: 020-7679-7048; Fax: 020-7679-7193; Email: s.perkins@medsch.ucl.ac.uk). The text consists of 31 pages (including 8 Figures and 1 Table, and a two-page Supplemetary Material).

SYNOPSIS

Component C3 is the central protein of the complement system. During complement activation, the thioester group in C3 is slowly hydrolysed to form C3u, then the presence of C3u enables the rapid conversion of C3 to functionally-active C3b. C3u shows functional similarities to C3b. To clarify this mechanism, the self-association properties and solution structures of C3 and C3u were determined using analytical ultracentrifugation and X-ray scattering. Sedimentation coefficients identified two different dimerisation events in both proteins. A fast dimerisation was observed in 50 mM NaCl that was removed in 137 mM NaCl. Low amounts of a slow dimerisation was observed for C3u and C3 in both buffers. The X-ray radius of gyration R_G values were unchanged for both C3 and C3u in 137 mM NaCl, but depend on concentration in 50 mM NaCl. The C3 crystal structure gave good X-ray fits for C3 in 137 mM NaCl. By randomisation of the TED/CUB domains in the C3b crystal structure, X-ray fits showed that the TED/CUB domains in C3u are extended and differ from the more compact arrangement of C3b. This TED/CUB conformation is intermediate between those of C3 and C3b. The greater exposure of the TED domain in C3u (which possesses the hydrolysed reactive thioester) accounts for the greater self-association of C3u in low salt. This conformational variability of the TED/CUB domains would facilitate their interactions with a broad range of antigenic surfaces. The second dimerisation of C3 and C3u may correspond to a dimer observed in one of the crystal structures of C3b.

INTRODUCTION

The complement system of innate immunity plays a major role in the recognition and elimination of microbial intruders and other pathogenic cells [1,2]. Complement is activated by three pathways, the classical, alternative and lectin pathways, each of which leads to the activation of C3, the so-named third component of complement. C3 is the most abundant complement component in plasma at about 1.0 mg/ml, and its level can be elevated considerably during inflammation and infection. The three pathways lead to a C3 convertase enzyme complex that activates C3 by proteolytic cleavage to yield the small anaphylatoxin C3a and active C3b [1,2]. The removal of C3a induces a major conformational change in C3b to expose an internal thioester bond, which is inaccessible in C3 [3-10]. The thioester forms covalent bridges with antigenic surfaces. C3 is also continuously activated at a slow "tick-over" rate in plasma to form C3u, sometimes named $C3_{H2O}$. C3u also participates in convertase enzyme complexes that cleave C3 into C3a and C3b, and triggers the rapid activation of C3 to C3b, but C3u is unable to attach covalently to surfaces because its thioester bond has been hydrolysed.

C3 is synthesised as a single chain pre-pro-molecule in which the α- and β-chains are linked by a tetra-Arg sequence that is enzymatically removed during post-translational processing. The α- and β-chains (molecular masses 115 kDa and 75 kDa respectively) are linked by a single disulphide bond and non-covalent interactions. Major aspects of complement activation have been revealed by three C3 and C5 crystal structures (where C5 is a homologue of C3) and five C3b crystal structures [3-10]. C3 is formed from a compacted arrangement of eight macroglobulin (MG) domains, to which are added the thioester-containing domain (TED) domain containing the C3 active site, a C1r/C1s-Uegf-Bmp1 (CUB) domain that links the TED and MG domains, and three additional C345C, L and C3a domains. In C3, the TED and CUB domains are positioned near the C345C domain and the thioester is buried [3-5]. In C3b, four similar crystal structures show that the TED and CUB domains are extended along the major axis of C3b to make contact with the MG1 domain at the base of the structure [6-9]. A fifth deviant crystal structure suggested that the TED and CUB domains are significantly extended away from the MG domains [10], however the crystallographic evidence for the fifth C3b structure was disputed [11-13]. Electron microscopy studies of C3b *in vacuo* supported the location of the TED and CUB domains seen in the four C3b crystal structures [14]. H-D exchange rates that were measured in mass spectrometry studies of C3, C3u and C3b indicated that large conformational changes occur between the three different protein forms [15,16].

Knowledge of the solution properties of C3 and C3u will clarify the initial stage of the mechanism of the conversion of C3 to C3b. A notable aspect of these crystal structures is that many were crystallised in buffers with low salt, and it is not known if these crystallographically-observed C3 structures or their properties are maintained in near-physiological buffers. A two-segmented structure for C3 and its homologues C4 and C5 had been identified from low resolution solution scattering modelling, in which large conformational differences between C4 and C4b were detectable [17-19]. Because the smaller segment seen earlier by scattering resembles the distinct TED-CUB domains seen in recent crystal structures, this indicates that scattering modelling is able to monitor the positions of the TED-CUB domains in C3 and C3u. In recent years, X-ray scattering has benefitted from improved signal-noise ratios at the ESRF facility, alongside with new constrained modelling methods that result in molecular structural determinations that are deposited in the Protein Data Bank [20,21]. Analytical ultracentrifugation has been much improved by the use of size-distribution *c(s)* analyses to analyse self-association [22]. Here,

we combine both methods to show that both self-association and conformational differences exist between C3 and C3u. We show that C3 and C3u are principally monomeric, however the monomer exists in equilibria with two different types of dimers. Constrained scattering modelling demonstrated that the position of the TED-CUB domains relative to the MG1-MG8 domains is significantly altered between C3 and C3u. We conclude that the conformational properties of the TED-CUB domains are more variable than believed from crystallography. This variability would enable activated C3 and its homologues to bind to a broad range of antigenic targets and provides insight for the mechanism of C3 regulation by complement factor H and other inhibitors.

EXPERIMENTAL

Purification of C3 and C3u

Wild-type C3 was purified from fresh human plasma essentially by Q-Sepharose fast flow anion-exchange column (Amersham Biosciences, GE Healthcare, Uppsala, Sweden) and a MonoQ 5/50 GL column (Pharmacia GE Healthcare, Uppsala, Sweden) as described previously with slight modifications [23]. Reduced SDS-PAGE confirmed the α chain (115 kDa) and β chain (75 kDa) (Figure 1). For ultracentrifugation and scattering experiments, C3 was dialysed into phosphate buffer saline (PBS: 8.1 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 2.7 mM KCl, 50 mM or 137 mM NaCl, pH 7.4) overnight at 4ºC, and concentrated using Amicon® Ultra-15 centrifugal filter devices with a molecular weight cut-off of 50 kDa at a speed of 2,500 g. Here and below, the buffer is denoted as 137 mM NaCl or 50 mM NaCl, even though phosphate is also present (Methods). The C3 concentrations were determined using an absorption coefficient of 9.4 (1%, 280 nm, 1 cm path length) calculated from its composition (SWISSPROT code P01024) assuming the presence of three high mannose type oligosaccharides at Asn63, Asn917 and Asn1597 [17,24,25]. The sequence-predicted molecular mass of C3 and C3u was 189.0 kDa, to be compared with that of 179.3 kDa for C3b. A partial specific volume of 0.739 g/ml was calculated for C3 and C3u from its composition [26]. The buffer density of 137 mM NaCl and 50 mM NaCl were measured at 20°C using an Anton-Paar DMA5000 density meter to be 1.00542 ml/g and 1.00171 ml/g, in good agreement with the predicted values of 1.00543ml/g and 1.00175 ml/g from SEDNTERP [27].

Haemolytically inactive C3 (C3u, also known as C3(H_2O)) was produced by incubating C3 with 100 mM hydrazine at 37ºC for 1 hr [28], and then left at 4 ºC overnight. C3u was separated from unconverted C3 by chromatography using the MonoQ column as above for C3, where C3u eluted earlier than C3. C3u but not C3 was active in functional assays using factor I and factor H [29].

Analytical ultracentrifugation data for C3 and C3u

Sedimentation velocity experiments were performed at 20°C using two Beckman XL-I analytical ultracentrifuges (Beckman-Coulter Inc, Palo Alto, CA) at rotor speeds of 40,000 r.p.m. and 50,000 r.p.m.. An eight-hole AnTi50 rotor was used with standard double-sector cells with column heights of 12 mm at 20ºC, monitoring sedimentation using absorbance optics at 280 nm and interference optics. C3 was studied in a concentration series of 0.26 mg/ml to 2.28 mg/ml (50 mM NaCl) and 0.48 mg/ml to 2.20 mg/ml (137 mM NaCl). C3u was studied between 0.35 mg/ml to 2.13 mg/ml (50 mM NaCl) and 0.27 mg/ml to 2.36 mg/ml (137 mM NaCl). The continuous *c(s)* analysis method was used to determine the $s_{20,w}$ values of C3 and C3u using SEDFIT software (version 11.7) [30,31]. The final *c(s)* fits were determined using a resolution of 200 and by floating the meniscus, frictional ratio f/f_0, and baseline and holding the cell bottom, partial specific volume and solvent density fixed until the overall root-mean-square deviations and visual appearance of the fits were satisfactory (Figure 2). Because f/f_0 is floated in the fits, the final f/f_0 values may be affected by oligomerisation.

X-ray scattering data for C3 and C3u

X-ray solution scattering data for C3 and C3u were collected in five sessions on the camera at the ID02 high brilliance beamline at the European Synchrotron Radiation Facility (ESRF) at Grenoble, France with a ring energy of 6.0 GeV [32]. In the first and second sessions, data were collected for C3 in 16-bunch mode using beam currents of 67-77 mA and 68-86 mA. In the third, fourth and fifth sessions, data were collected for C3 and C3u using

beam currents of 30-37 mA (4-bunch mode), 69-83 mA (16-bunch mode) and 30-43 mA (4-bunch mode). Data were acquired using an improved fibre optically-coupled high sensitivity and dynamic range CCD detector (FReLoN) with a smaller beamstop. The sample-to-detector distance was 2.0 m. Sample flow cells were used to eliminate radiation damage in conjunction with on-line checks to confirm the absence of radiation damage by optimisation of the exposure times. Each sample was measured using four sets of ten time frames, each of length 0.05, 0.08, 0.1, 0.15, 0.2 or 0.5 sec for C3 and 0.05, 0.08, 0.1, 0.15, 0.2 or 0.25s for C3u, then these were averaged. The sample temperature was 20°C. C3 in 50 mM NaCl was studied at seven concentrations between 0.18 mg/ml and 0.98 mg/ml. C3 in 137 mM NaCl was studied at 18 concentrations between 0.21 mg/ml and 1.95 mg/ml. C3u in 50 mM NaCl was studied at five concentrations between 0.32 mg/ml and 1.67 mg/ml. C3u in 137 mM NaCl was studied at seven concentrations between 0.28 mg/ml and 1.12 mg/ml.

Guinier plot at low Q values gives the R_G and scattering at zero angle $I(0)$ [33]:

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

This expression is valid in a $Q.R_G$ range up to 1.5. In a given solute-solvent contrast, the radius of gyration R_G is a measure of structural elongation if the internal inhomogeneity within C3 or C3u has no effect. The $I(0)/c$ value is proportional to the molecular mass. For an elongated structure, at larger Q values than those used for the R_G analysis, the radius of gyration of the cross-sectional structure R_{XS} and the mean cross-sectional intensity at zero angle $[I(Q).Q]_{Q\to 0}$ parameters are obtained from:

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

The R_{XS} value monitors the elongation of its cross-section shape along its longest axis. The R_G and R_{XS} analyses were performed using an interactive PERL script program SCTPL7 (J. T. Eaton and S. J. Perkins, unpublished software) on Silicon Graphics OCTANE Workstations. Indirect Fourier transformation of the full scattering curve $I(Q)$ in reciprocal space gives the distance distribution function $P(r)$ in real space. This yields the maximum dimension of the macromolecule L and its most commonly occurring distance vector M in real space:

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

The transformation was implemented using GNOM software [34]. For C3, up to 345 data points were used for Q values between 0.13 nm^{-1} and 1.73 nm^{-1}, while for C3u, up to 373 data points were used for Q values between 0.16 nm^{-1} and 1.85 nm^{-1}.

Constrained scattering and sedimentation coefficient modelling for C3 and C3u

The scattering modelling of human C3 was based on its unmodified crystal structure (PDB code 2a73) [3]. That for C3u was based on four crystal structures for human C3b (PDB codes 2i07, 2icf, 2win, 2wii), in which the 4.7% volume contribution from C3a (77 residues) was taken to be negligible [6-9]. The three high-mannose oligosaccharide chains at Asn63, Asn917 and Asn1597 contribute 2.3% of the volume. Inspection of these five crystal structures indicated that C3 and C3b could be considered in terms of a C3c region connected to the CUB and TED domains. For the constrained modelling of C3u based on the C3b 2i07 crystal structure, in which four residues of the oligosaccharides at Asn63 and Asn917 were visible, these three regions were joined by two pairs of linker peptides. Linkers 1 and 2 connect the MG7 and MG8 domains respectively in C3c with the CUB domain. Linkers 3 and 4 connect the CUB and TED domains. Their sequences are as follows: linker 1, 908-KVVPEGI-914; linker 2, 1328-AKAKDQLTC-1336; linker 3, 964-PVAQMTEDAV-973; linker 4, 1263-DAPDHQEL-1270 (C3 sequence numbering [3]). Each linker was created as an extended β-strand structure, each of which was then conformationally randomised to

follow Method 2 used for modelling human IgA1 [35]. For linker 1, 6,000 conformers were created in which the distance between the first and last a-carbon atoms ranged between 1.21 nm to 2.27 nm. For linker 2, the 4,650 conformers ranged in length between 0.59 nm to 2.72 nm. For linker 3, the 6,000 conformers ranged in length between 1.35 nm to 3.38 nm. For linker 4, the 4,200 conformers ranged in length between 0.86 nm to 2.89 nm. Five sets of randomised C3u models were formed by combining the linkers and three regions. In the first of these, the crystallographically-observed linkers 2, 3 and 4 were fixed, the connection between linker 2 and the MG8 domain was severed, and 6,000 randomised linker 1 conformations joined the TED and CUB domains to the C3c region. The second set of 4,650 C3u models employed the linker 2 conformations. The third and fourth sets of models varied the position of the TED domain relative to the fixed CUB and C3c regions using the 6,000 linker 3 or 4,200 linker 4 conformations respectively. A fifth set of 8,000 models was created from a randomised selection of linkers 1 and 3 conformations while holding linkers 2 and 4 fixed. In the best-fit models, the distance between the TED and MG1 domains was taken to be the α-carbon separation between GlnA87 and LysB1062.

Each C3 or C3b molecular model was used to calculate its X-ray scattering curve for comparison with the experimental C3 or C3u curve in 137 mM NaCl using Debye sphere models [36]. A cube side length of 0.540 nm in combination with a cutoff of 4 atoms was used to convert the atomic coordinates into Debye sphere models with 1557 spheres (C3) or 1478 spheres (C3b). The spheres corresponded to the unhydrated structure and allow for the amino acid and carbohydrate residues expected to be present, even though residues were not visible in the crystal structure for reason of disorder. The hydration shell corresponding to 0.3 g H_2O/ g protein was created using HYPRO [37], where the optimal total of hydrated spheres in the C3 or C3u models is 2047 (C3) or 1943 (C3b) The X-ray scattering curve *I(Q)* was calculated using the Debye equation adapted to spheres [36]. No instrumental corrections were applied because these are considered to be negligible for the pinhole optics in use. No ad-hoc assumptions were required in relation to hydration effects, and details are given elsewhere [38]. Steric overlap between the TED, CUB and C3c regions in the models was assessed using the number of spheres *N* in the models after grid transformation, where models showing less than 95% of the required total of 2047 or 1943 spheres were discarded. Over 54% of the 10,000 models met this "absence of steric overlap" criterion. Next, the X-ray R_G and R_{XS} values were calculated from the modelled curves in the same *Q* ranges used for the experimental Guinier fits. This allowed for any approximations inherent in the use of the $Q.R_G$ fit range up to 1.5. Models that passed R_G and R_{XS} filters of ± 5% 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 1.73 nm^{-1} (C3) or 1.85 nm^{-1} (C3u).

Sedimentation coefficients $s^0_{20,w}$ were calculated directly from the hydrated Debye sphere models using the program HYDRO as a confirmation of the GENDIA analysis [39]. They were also calculated from the atomic coordinates in the HYDROPRO shell modelling program using the default value of 0.31 nm for the atomic element radius for all atoms to represent the hydration shell [40]. Previous applications of these calculations to proteins are reviewed elsewhere [21].

Protein Data Bank accession number

The five best-fit C3u models for 137 mM NaCl were deposited in the Protein Data Bank with the accession code 3mmq.

RESULTS

Sedimentation velocity data analysis for C3 and C3u

Immediately prior to ultracentrifugation and scattering experiments, purified C3 and C3u (Materials and Methods) were subjected to gel filtration on the same Superose 6 column to remove potential pre-existing aggregates or degradation products. Both proteins eluted as single symmetric peaks in 137 mM NaCl and near-symmetric peaks in 50 mM NaCl (Figures 1(a,b)). C3u (and less so with C3) eluted earlier in 50 mM NaCl than in 137 mM NaCl, suggesting that there were hydration changes, conformational changes or altered self-association in reduced salt. C3 and C3u before and after experiments revealed a single band in non-reduced SDS-PAGE and two bands corresponding to the α- and β-chains in reduced SDS-PAGE (Figure 1(c)).

Analytical ultracentrifugation studies macromolecular structures in solution by monitoring the time-course of sedimentation under high centrifugal force [22]. The peaks in size distribution analyses *c(s)* plots monitor the number of species present, and macromolecular elongation is monitored through the sedimentation coefficient $s_{20,w}$ values. Using SEDFIT, good boundary fits were obtained for each of C3 and C3u in 50 mM and 137 mM NaCl buffer (Figure 2(a-d)). A major *c(s)* peak was visible in all runs, and differences were seen between the two buffers:

(i) In 137 mM NaCl, both the interference and absorbance data sets for C3 gave an $s^0_{20,w}$ value of 7.85 ± 0.05 S, and both data sets for C3u gave an $s^0_{20,w}$ value of 7.44 ± 0.07 S (Figure 2(g,h); Table 1(a)). No concentration dependence was seen, showing that both the shape and size of C3 and C3u were each unchanged with concentration. The molecular masses calculated from the *c(s)* peaks corresponded to 192 ± 8 kDa and 172 ± 16 kDa for C3 and C3u respectively, in good agreement with their sequence-predicted molecular weight of 189.0 kDa. Accordingly the $s_{20,w}$ changes showed that C3u is more elongated in solution than C3. The fitted frictional ratios f/f_o (where f_O is the frictional coefficient of the sphere with the same volume as the hydrated glycoprotein) for C3 and C3u were 1.37 ± 0.10 and 1.40 ± 0.10 respectively.

(ii) In 50 mM NaCl, the major *c(s)* peak for C3 and C3u both showed a clear concentration dependence (Figure 2(e,f)). The molecular weight and f/f_o value for C3 were 190 ± 18 kDa and 1.38 ± 0.09 respectively; those for C3u were 175 ± 20 and 1.37 ± 0.12 respectively. The $s_{20,w}$ values for C3 ranged from 8.04 S to 8.25 S, while those for C3u increased by a larger proportion from 7.59 S to 8.02 S for C3u. By extrapolation, the $s^0_{20,w}$ values for C3 and C3u were 8.02 ± 0.12 S and 7.66 ± 0.19 S respectively. While these values are slightly increased in 50 mM compared to 137 mM, and may imply that both proteins became more compact in 50 mM NaCl, the $s_{20,w}$ change is not greater than their error.

For the C3 or C3u monomer, the large concentration dependence in 50 mM NaCl indicated a fast equilibrium with a higher oligomer, presumed to be a C3 or C3u dimer. Reaction boundaries are peaks that correspond to the co-sedimentation of two different species, and exhibit sedimentation coefficients between those for the two different species [41]. Interestingly, the C3d fragment of C3 showed related observations, where a single unchanged $s_{20,w}$ value was seen at 3.0 S in 137 mM NaCl, while three separate *c(s)* peaks at 3.0 S, 4.3 S and 6.0 S were seen in 50 mM NaCl that indicated monomer, dimer and trimer forms in a slow equilibrium [42,43]. Because the C3 and C3u concentration dependence resembles that for C3d (which is equivalent to its TED domain), the self-association of C3 and C3u may be attributable to the TED domain. The observation of a slow equilibrium for C3d and a fast one for C3 and C3u indicated that TED self-association within C3 and C3u is

weaker. The greater salt dependence seen for C3u in 50 mM NaCl compared to 137 mM NaCl indicated that the TED domain in C3u is more exposed than in C3.

The *c(s)* analyses for C3 and C3u consistently revealed a second minor $s_{20,w}$ peak at approximately 11.2 ± 1.0 S and 11.5 ± 1.3 S respectively (Figure 3(a)). No concentration dependence of this peak was observed. The corresponding molecular masses were 330 ± 40 kDa and 310 ± 40 kDa for C3 and C3u respectively, which are comparable to 378 kDa for a predicted dimer. The stability of this peak showed that this is not a reaction boundary, but is a second different dimer. Integrations of the *c(s)* peaks showed that the estimated proportion of dimer is low at 5 ± 2% for C3 in 50 mM NaCl and 5 ± 3% for C3 in 137 mM NaCl. For C3u, this is low at 7 ± 2% in 137 mM NaCl, but increases up to 19% in 50 mM NaCl (Figure 3(b,c)). The difference between the C3 and C3u dimer in 50 mM NaCl indicated that C3 changes conformation when C3u is formed. This may reflect the large movements of the MG1-MG8 domains between C3 and C3b on activation [3-9], suggesting that the C3u conformation may resemble that of C3b.

X-ray solution scattering data for C3 and C3u

Solution scattering is a diffraction technique that studies the overall structure of macromolecules in solution [21,44]. The domain arrangement and self-association of freshly prepared C3 and C3u in both buffers were assessed by Guinier analyses of the X-ray *I(Q)* data in two different *Q* ranges ($Q = 4\pi \sin\theta/\lambda$; 2θ = scattering angle; λ = wavelength). All four sets of Guinier R_G analyses showed linear fits within satisfactory $Q.R_G$ and $Q.R_{XS}$ limits (Figure 4). The concentration dependence of the R_G values again reflected the buffer in use:

(i) In 137 mM NaCl, the R_G values for C3 and C3u were unchanged at 4.52 ± 0.08 nm and 4.88 ± 0.23 nm with protein concentration (Figure 5(a,c); Table 1(a)). The larger value for C3u reflected its more elongated structure, in agreement with the $s^0_{20,w}$ values. The *I(0)/c* values were similar at 0.0186 ± 0.0009 and 0.0196 ± 0.0013 (Figure 5(b,d)), showing that both C3 and C3u were similar in mass as expected. The R_{XS} values were the same at 2.54 ± 0.06 nm and 2.55 ± 0.16 nm for C3 and C3u (Figure 5(e,f)).

(ii) In 50 mM NaCl, the R_G values for C3 and C3u increased linearly with concentration. At zero concentration, the R_G values were 4.87 ± 0.26 nm and 5.16 ± 0.22 nm respectively. That these values are slightly larger than those at 137 mM is attributed to the increased proportion of dimer in reduced salt, meaning that the R_G value is an average of monomer and a small amount of dimer. The *I(0)/c* values were similar at 0.0199 ± 0.0016 and 0.0178 ± 0.0011, and agree with those at 137 mM NaCl when extrapolated to zero concentration. Thus the *I(0)/c* values correspond to similar masses, indicating that the monomer-dimer equilibrium is reversible. The R_{XS} values are also the same at 2.50 ± 0.17 nm and 2.56 ± 0.03 nm for C3 and C3u when extrapolated to zero concentration.

The distance distribution function *P(r)* leads to the determination of overall lengths *L* following an assumption of the value of the maximum dimension D_{max}. The R_G values calculated from the *P(r)* analyses agree with those from the Guinier analyses (Table 1(a)). In 137 mM NaCl, the length *L* of C3 and C3u in 137 mM NaCl is 16 nm (Figure 6(a,c)). In 50 mM NaCl, the *L* values increase from 15-16 nm at the lowest concentrations of 0.5-0.6 mg/ml up to 20 nm at 1.0-1.7 mg/ml. This concentration dependence is consistent with Figure 5.

Previous C3 X-ray scattering was performed in different buffers (50 mM potassium phosphate, 100 mM KC1, 5 mM EDTA, pH 7.0, or 12.4 mM sodium phosphate, 200 mM NaC1, 0.5 mM EDTA, pH 7.0) [17]. The previous R_G value in 200 mM NaCl [17] was 5.2 ± 0.1 nm, which is larger than the present C3 R_G values (Table 1(a)). The reduced R_G value is

attributable to the avoidance of protein aggregation by gel filtration immediately prior to ESRF beam time, and instrumental improvements at the ESRF such as better detector sensitivities, and a larger Q range. The previous R_{XS} value in 200 mM NaCl of 2.6 ± 0.1 nm [17] agrees with the current C3 R_{XS} value of 2.54 nm (Table 1(a)).

Constrained modelling of the C3 and C3u solution structures

C3 is converted to C3b by the removal of its small anaphylatoxin C3a domain. The TED/CUB domains move from their position next to the C345C domain in C3 to a position adjacent to the MG1 domain in four C3b crystal structures [3-9]. Comparatively large conformational changes in the MG domains also occur during this conversion. C3b has 12 domains, namely the MG1-MG8, LNK, and C345c domains within its C3c region and the CUB and TED domains outside this region (Supplementary Figure 1(a)). The CUB and TED domains are connected by linker peptides of length seven to ten residues to the MG7 and MG8 domains, denoted by 1-4 in Supplementary Figure 1(a). These peptides are significantly altered in conformation between C3 and C3b. Here, the conformational movement of the TED/CUB domains was investigated by constrained scattering modelling of C3 and C3u.

First, six crystal structures were tested for their compatibility with sedimentation coefficients. For the monomer *c(s)* peak, the C3 crystal structure gave a calculated $s^0_{20,w}$ value of 8.11 S, which agreed well with the experimental value of 8.04 S in 137 mM NaCl, and within error of that of 7.85 S in 50 mM NaCl, given that the method is accurate to ± 0.21 S [21] (Table 1(a,b)). The four C3b crystal structures gave calculated $s^0_{20,w}$ values of 7.51-7.69 S, which agreed well with the experimental value of 7.66 S for C3u in 50 mM NaCl. Intriguingly, if the TED/CUB domains are extended away from the MG domains, as seen in the deviant C3b crystal structure [10-13], the calculated $s^0_{20,w}$ value is reduced to 7.37 S, in good agreement with the experimental value of 7.47 S for C3u in 137 mM NaCl. Even though the change is at the limit of experimental error, the transition from an extended TED/CUB arrangement to a more compact one would account for this difference in $s^0_{20,w}$ values seen for C3u between 50 mM and 137 mM NaCl. The calculated f/f_0 ratios from the five C3 and C3b crystal structures gave values close to 1.3. They are slightly less than the observed fitted values of 1.37-1.40 above from SEDFIT, which is as expected because the fitted values will reflect both monomer and oligomer. For the dimer *c(s)* peak, one of the four C3b crystal structures (PDB code 2win) revealed dimer formation at the MG4-MG5 interface. The calculated $s^0_{20,w}$ values calculated from two dimers in this crystal structure were 11.8 S and 12.1 S, in good accord with the observed values of 11.2 ± 1.0 S and 11.5 ± 1.3 S (Figure 3(a)). While this crystallographic dimer may lead to the observed *c(s)* dimer peak, this peak assignment is not unequivocal.

Next, the monomeric crystal structures were tested for their compatibility with the scattering data. The experimental R_G values of 4.52-4.80 nm for C3 in both buffers were similar to the calculated R_G value of 4.35 nm for C3 within error (Table 1(a,b)). The experimental R_G values of 4.88-5.35 nm for C3u in both buffers were also similar to the calculated R_G values of 4.58-4.64 nm for the four C3b crystal structures within error. That the experimental values were slightly larger than the calculated values is attributed to the sensitivity of scattering at low Q to small amounts of dimer. Full curve fits were performed with the experimental scattering data in 137 mM NaCl, where dimerisation was assumed to be sufficiently low not to affect the fits. A good C3 curve fit was obtained for the C3 crystal structure with a low R-factor of 3.1% (Figure 7(a)). This showed that the crystal and solution structures are similar. The four C3b crystal structures gave larger R-factors of 4.5-5.1% with C3u in 137 mM NaCl, while an intermediate R-factor of 3.9% was obtained with the deviant

C3b crystal structure. This showed that, if the TED/CUB domains are extended away from the MG domains, this resulted in improved agreement between the solution and crystal structures.

Third, because the C3u solution structure is unknown, the extended location of the CUB and TED domains in C3u was evaluated by constrained scattering modelling starting from the C3b crystal structure (Materials and Methods). C3b was used for modelling because of the functional similarity between C3u and C3b and the consistency of the sedimentation data with the C3b crystal structure. The four linkers that connect the CUB and TED domains to the MG1-MG8 domains (Supplementary Figure 1(a)) were each conformationally randomised in four separate searches in order to generate up to 8,000 trial TED/CUB structures per search. The predicted X-ray curves from these models were compared with the experimental C3u curve in 137 mM NaCl buffer. The four graphs of *R*-factor (goodness-of-fit) vs. the R_G values showed that the experimental R_G value of 4.97 nm is close to the minimum *R*-factor values in each case (Supplementary Figure 1(b-f)). Filters based on the number of spheres *N* in the best-fit models and the R_G and R_{XS} values (Table 1) identified 5-15 best-fit C3u models from each search (pink circles in Supplementary Figure 1(b-f); the best-fit model is denoted in green). All four searches consistently showed that the TED domain was well separated from the MG1 domain in C3u, unlike in the C3b crystal structures.

The overall best-fit C3u model with an *R*-factor of 3.6% was determined from the Linker 2 search (Table 1(c)). Comparison of the calculated and experimental C3u *I(Q)* curves showed good visual agreement (Figure 7(b)). Likewise the modelled *P*(*r*) curve showed a major peak at r = 5.45 nm and L of 16 nm, indicating good agreement with Figure 6(b). In the best-fit models, the separation of 6.0 ± 0.9 nm between the TED and MG1 domains is significantly wider than that of 3.1 ± 0.1 nm seen in the other four crystal structures for C3b (PDB codes 2i07, 2icf, 2win, 2wii) [6-9].

Control calculations were explored. First, the effectiveness of the modelling in detecting conformational change between C3 and C3u was shown by clear intensity differences observed in the data between Q of 0.1-0.6 nm^{-1} (Figure 7(b)). Second, to test whether the curve fits could identify the C3 crystal structure, models were created by varying both Linker 1 and 3 (Supplementary Figure 1(f)). When these were fitted using the experimental C3 curve in 137 mM NaCl, the *R*-factor *vs.* R_G analysis differed significantly from that for C3u. The minimum was shifted to lower R_G values, and showed even lower *R*-factors than the C3u analysis. Of the 15 best-fit structures, seven resembled the C3 crystal structure itself, while the other eight showed a less specific range of other conformations. Even though this search did not allow for the significant shifts of the MG domains between C3b and C3, the similarity of many best-fit models with the C3 crystal structure validated the modelling procedure used for C3u.

DISCUSSION

C3, C3u and C3b are important for the initiation, amplification and effector functions of complement [1,2]. The "tick-over" conversion of C3 to C3u leads to alternative pathway activation, because C3u exhibits C3b-like function and forms an initial convertase with the Bb fragment of Factor B to enable the rapid generation of active C3b. We have determined solution structures for C3 and C3u by constrained scattering modelling. The high quality of the curve fits are similar to those obtained for other complement proteins and antibodies [20,45]. While the C3 solution structure corresponded well with its crystal structure, the C3u solution structure is distinct to that of four similar C3b crystal structures (Figures 8(a,b)). The TED domain in C3u is significantly extended away from the MG1-MG8 domains in solution, and is not proximate to the MG1 domain. This is different from the four C3b crystal structures that show that the TED and MG1 domains are in contact with each other. A small variability in the four C3b crystal structures suggests that the TED and MG1 domains do not form a precise interaction. A weak interaction between the TED and MG1 domains would explain why the C3b crystal and C3u solution structures differ. Electron microscopy images are viewed in two-dimensions in comparatively harsh conditions. While shape differences between C3 and C3u were detected by electron microscopy [14], their structures could not be defined in more detail. Hydrogen/deuterium exchange studies also revealed differences between C3 and C3u, and showed that C3u possessed a more open structure [15]. Our results indicate that the arrangement of the TED and CUB domains relative to the eight MG domains exhibit previously unrecognised conformational flexibility.

Mechanistic insights are obtained by this study. Our results show that C3u represents a structure that is intermediate between those of C3 and C3b. Functionally, the ability of the TED domain to move independently from the MG domains in the C3 proteins will facilitate the attachment of C3b to an antigenic surface in a broad range of conformations independent of the remaining C3b structure. The four similar C3b crystal structures do not show that the TED domain is optimally positioned for antigen interactions, because the MG domains can potentially block steric access to one half of this. Versatile TED arrangements in the C3 proteins would permit a more efficient innate immune response by C3 and also its complement C4 and C5 homologues. Functionally also, the mobility of the TED domain relative to the fixed MG domains clarifies the need for Factor H to act as a cofactor for the Factor I-mediated cleavage of C3b. The Factor H SCR-4 domain is seen to link together the MG1 and TED domains in the crystal structure [8]. It is deduced that Factor H restricts what would otherwise be a high conformational variability of the TED domain in order to provide an appropriate rigid template at the CUB domain for Factor I-mediated cleavage of C3.

The existence of both C3 and C3u monomer and dimer in solution was previously unrecognised. By size-exclusion chromatography, no dimer peak was visible, however both C3 and C3u eluted earlier in 50 mM NaCl than in 137 mM NaCl buffer (Figure 1). Our *c(s)* size-distribution analyses and X-ray scattering data explain this observation [14] in terms of reversible greater self-association in lower salt. While our analysis of reversible systems has been achieved using sedimention boundary analysis, it is not rigorous. Even though there are good indications that dimerisation is occurring and differs with salt, the possible effects of conformational change, different hydration levels and non-ideality may require consideration. Concentration studies by sedimentation equilibrium and other methods will clarify the association kinetics and dissociation constant K_D values between each of the C3 and C3u dimers. Physiologically, C3 concentrations in serum range between 1 and 1.5 mg/ml (5.3 μM to 7.9 μM) in 137 mM NaCl [2]. Given that C3 is an acute-phase reactant, its concentration is increased during inflammation, and any localised production could also result in higher C3 levels at inflammatory sites. Here, we have observed both C3 and C3u to be largely monomeric up to 2.3 mg/ml in physiological 137 mM NaCl, in agreement with the observation of monomeric C3d in this buffer [43]. In 50 mM NaCl, greater amounts of dimer were seen, especially for C3u, in a reversible self-association with monomer (Figures 2-3). Self-association was also observed with C3d in this buffer, which is equivalent to the TED domain [42]. The interaction between factor H and C3d is seen to be strengthened in 50 mM NaCl compared to 137 mM NaCl buffer, even though this interaction is complicated by its multimeric nature [46]. It is expected that the factor H-C3u interaction will be likewise increased in 50 mM NaCl. A rigorous understanding of these events will require K_D determinations for C3 and C3u dimerisation as well as the factor H-C3u interaction. If substantiated, the larger dimerisation change seen with C3u, but not C3, may be explained by the existence of a more exposed TED domain in the C3u solution structure (Figure 8(a)). In addition to this, relatively low amounts of C3 and C3u dimer were also seen in all buffers (Figure 3). This may be best explained by a second dimerisation site. The crystal structure of the C3b-SCIN complex (where SCIN is a bacterial *Staphylococcus aureus* protein that inhibits C3b) revealed a different non-covalent C3b dimer that is formed through interactions between the MG4-MG5 domains and stabilised by SCIN [9]. In support of this, the modelling of our sedimentation coefficient data for this second dimer based on this dimeric crystal structure was satisfactory.

Many C3 and C3b crystal structures were crystallised from low salt buffers using a diverse range of additives. Here, we have shown that alterations in buffer conditions affect the spatial arrangement of the 13 domains in the structures of C3 and C3u. Interestingly, examination of the crystallisation buffers for all four similar C3b crystal structures showed that comparatively low NaCl concentrations up to 50 mM were used, while the deviant C3b crystal structure was determined using 100 mM NaCl. Even though the deviant crystal structure is disputed [10-13], the higher NaCl concentration used for these crystals apparently resulted in a more extended TED/CUB domain arrangement that gave a better account of the C3u scattering and ultracentrifugation data in 137 mM NaCl. The resolution of separate monomer and dimer peaks by ultracentrifugation showed that the arrangement of 13 domains in both the C3 and C3u monomers is slightly more compact in 50 mM NaCl than in 137 mM NaCl. Functionally-related proteins such as Factor H with 20 SCR domains also demonstrated conformational changes with change of salt concentration [47]. Accordingly functional studies performed in 50 mM NaCl buffer should be interpreted with caution. Buffers containing 137 mM NaCl will maintain C3 and C3u as largely monomeric proteins in physiologically-appropriate concentrations.

Conformational and self-asssociation variability will also apply to C3 homologues, most notably C4 and C5, where this may have significant roles in complement mechanisms during their activation, and the formation of the C5 convertase by the covalent attachment of a C3b molecule to the C3bBb complex [1,2]. Thus the degree of self-association of C3 and C3u and conformational changes in these is important for both physiological function and laboratory investigations. The appropriate choice of buffer will be crucial to study their interactions with physiological ligands or pathogenic proteins.

AUTHOR CONTRIBUTION

Keying Li designed and performed experiments, analysed data and wrote the manuscript; Jayesh Gor designed and performed experiments; and Stephen Perkins designed the study, analysed data and wrote the manuscript.

ACKNOWLEDGEMENTS

K. L. thanks the UCL Graduate School, the Overseas Research Fund and the Mercer Fund of the Fight For Sight Charity for a postgraduate scholarship. We are very grateful to Dr Azubuike I. Okemefuna for useful discussions, Dr A. Shukla and Dr T. Narayanan (ESRF, Grenoble) for excellent instrumental support, Dr A. W. Dodds (MRC Immunochemistry Unit, Oxford) for generous advice in purifying C3, and the Biotechnology and Biological Sciences Research Council, the Henry Smith Charity and the Mercer Fund of the Fight For Sight Charity for equipment grant support.

REFERENCES

1 Walport, M. J. (2001) Complement. N. Engl. J. Med. **344**, 1058-1066 and 1140-1144.

2 Law, S. K. A. and Reid, K. B. M. (1995) Complement (Second edition) IRL Press, Oxford.

3 Janssen, B. J., Huizinga, E. G., Raaijmakers, H. C., Roos, A., Daha, M. R., Nilsson-Ekdahl, K., Nilsson, B. and Gros, P. (2005) Structures of complement component C3 provide insights into the function and evolution of immunity. Nature, **22**, 505-511.

4 Fredslund, F., Jenner, L., Husted, L. B., Nyborg, J., Andersen, G. R. and Sottrup-jensen, L. (2006) The structure of bovine complement component 3 reveals the basis for thioester function. J. Mol. Biol. **361**, 115-127.

5 Fredslund, F., Laursen, N. S., Roversi, P., Jenner, L., Oliveira, C. L., Pedersen, J. S., Nunn, M. A., Lea, S. M., Discipio, R., Sottrup-Jensen, L. and Andersen, G. R. (2008) Structure of and influence of a tick complement inhibitor on human complement component 5. Nat. Immunol. **9**, 753-760.

6 Janssen, B. J., Christodoulidou, A., McCarthy, A., Lambris, J. D. and Gros, P. (2006) Structure of C3b reveals conformational changes that underlie complement activity. Nature, **444**, 213-216.

7 Wiesmann, C., Katschke, K. J., Yin, J., Helmy, K. Y., Steffek, M., Fairbrother, W. J., McCallum, S. A., Embuscado, L., DeForge, L., Hass, P. E. and van Lookeren Campagne, M. (2006) Structure of C3b in complex with CRIg gives insights into regulation of complement activation. Nature, **444**, 217-220.

8 Wu, J., Wu, Y. Q., Ricklin, D., Janssen, B. J., Lambris, J. D.and Gros, P. (2009) Structure of complement fragment C3b-factor H and implications for host protection by complement regulators. Nat. Immunol. **10**, 728-733.

9 Rooijakkers, S, H., Wu, J., Ruyken, M., van Domselaar, R., Planken, K. L., Tzekou, A., Ricklin, D., Lambris, J. D., Janssen, B. J., van Strijp, J. A. and Gros, P. (2009) Structural and functional implications of the alternative complement pathway C3 convertase stabilized by a staphylococcal inhibitor. Nat. Immunol. **10**, 721-727.

10 Ajees, A. A., Gunasekaran, K., Volanakis, J. E., Narayana, S. V. L., Kotwal, G. J.and Murthy, K. H. M. (2006) The structure of complement C3b provides insights into complement activation and regulation. Nature, **444**, 221-225.

11 Janssen, B. J. C., Read, R. J., Brűnger, A. T. and Gros, P. (2007) Crystallographic evidence for deviating C3b structure? Nature, **448**, E1-E2.

12 Ajees, A. A., Gunasekaran, K., Narayana, S. V. L. and Murthy, K. H. M. (2007) Reply. Nature, **448**, E2-E3.

13 Borrell, B. (2009) Fraud rocks protein community. Nature, **462**, 970-970.

14 Nishida, N., Walz, T. and Springer, T. A. (2006) Structural transitions of complement component C3 and its activation products. Proc. Natl. Acad. Sci. U. S. A. **103**, 19737-19742.

15 Winters, M. S., Spellman, D. S. and Lambris, J. D. (2005) Solvent accessibility of native and hydrolyzed human complement protein 3 analyzed by hydrogen/deuterium exchange and mass spectrometry. J. Immunol. **174**, 3469-3474.

16 Schuster, M. C., Ricklin, D., Papp, K., Molnar, K. S., Coales, S. J., Hamuro, Y., Sfyroera, G., Chen, H., Winters, M. S. and Lambris, J. D. (2008) Dynamic structural changes during complement C3 activation analyzed by hydrogen/deuterium exchange mass spectrometry. Mol. Immunol. **45**, 3142-3151.

17 Perkins, S. J. and Sim, R. B. (1986) Molecular modelling of human complement component C3 and its fragments by solution scattering.Eur. J. Biochem. **157**, 155-168.

18 Perkins, S. J., Nealis, A. S. and Sim, R. B. (1990) Molecular modelling for human complement component C4 and its fragments by X-ray and neutron solution scattering. Biochemistry, **29**, 1167-1175.

19 Perkins, S. J., Smith, K. F., Nealis, A. S., Lachmann, P. J. and Harrison, R. A. (1990) Structural homologies of component C5 of human complement with components C3 and C4 by neutron scattering. Biochemistry, **29**, 1175-1180.

20 Bonner, A., Almogren, A., Furtado, P. B., Kerr, M. A. and Perkins, S. J. (2009) Location of secretory component on the Fc edge of dimeric IgA1 reveals insight into the role of secretory IgA1 in mucosal immunity. Mucosal Immunology (Nature Publishing Group), **2**, 74-84.

21 Perkins, S. J., Okemefuna, A. I., Nan, R., Li, K. and Bonner, A.(2009) Constrained solution scattering modelling of human antibodies and complement proteins reveals novel biological insights. J. Roy. Soc. Interface. **6**, S679-S696.

22 Cole, J. L., Lary, J. W., Moody, T. P. and Laue, T. M. (2008) Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. Meth. Cell Biol. **84**, 143-79.

23 Dodds, A. W. (1993) Small-scale preparation of complement components C3 and C4. Methods Enzymol. **223**, 46-61.

24 Hase, S., Kikuchi, N., Ikenaka, T. and Inoue, K. (1985) Structures of sugar chains of the third component of human complement. J. Biochem., **98**, 863-74.

25 Hirani, S., Lambris, J. D. and Müller-Eberhard, H. J. (1986) Structural analysis of the asparagine-linked oligosaccharides of human complement component C3. Biochem. J. **233**, 613-6.

26 Perkins, S. J. (1986) Protein volumes and hydration effects. The calculations 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.

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

28 Sim, E., Wood, A. B., Hsiung, L.M. and Sim, R. B. (1981) Pattern of degradation of human complement fragment C3b. FEBS Lett. **132**, 55-60.

29 Nan, R., Gor, J., Lengyel, I. and Perkins, S. J. (2008) Uncontrolled zinc- and copper-induced oligomerisation of the human complement regulator Factor H and its possible implications for function and disease. J. Mol. Biol. **384**, 1341-1352.

30 Dam, J. and Schuck, P. (2004) Calculating sedimentation coefficient distribution by direct modeling of sedimentation velocity concentration profiles. Methods Enzymol. **384**, 185–212.

31 Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. J. Biophys. **78**, 1606–1619.

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

33 Glatter, O. and Kratky, O., Eds. (1982) Small-angle X-ray Scattering. Academic Press, New York.

34 Semenyuk, A. V. and Svergun, D. I. (1991) GNOM – a program package for small-angle scattering dataprocessing. J. Appl. Crystallog. 24, 537–540.

35 Boehm, M. K., Woof, J. M., Kerr, M. A. and 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.

36 Perkins, S. J. and 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.

37 Ashton, A. W., Boehm, M. K., Gallimore, J. R., Pepys, M. B. and 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.

38 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.

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

40 Garcia de la Torre, J., Huertas, M. L. and Carrasco, B. (2000) Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. Biophys. J. **78**, 719-730.

41 Dam, J. and Schuck, P. (2005) Sedimentation velocity analysis of heterogeneous protein-protein interactions: Sedimentation coefficient distributions c(s) and asymptotic boundary profiles from Gilbert-Jenkins theory. Biophys. J. **89**, 651–666.

42 Gilbert, H. E., Eaton, J. T., Hannan, J. P., Holers, V. M., and Perkins, S. J. (2005) Solution structure of the complex between CR2 SCR 1-2 and C3d of human complement: an X-ray scattering and sedimentation modelling study. J. Mol. Biol. **346**, 859-873.

43 Li, K., Okemefuna, A. I., Gor, J., Hannan, J. P., Asokan, R., Holers, V. M. and Perkins, S. J. (2008) Solution structure of the complex formed between human complement C3d and full length complement receptor Type 2. J. Mol. Biol. **384**, 137-150.

44 Perkins, S. J., Okemefuna, A. I., Fernando, A. N., Bonner, A., Gilbert, H. E. and Furtado, P. B. (2008) X-ray and neutron scattering data and their constrained molecular modelling. Meth. Cell Biol. **84**, 375-423.

45 Furtado, P. B., Huang, C. Y., Ihyembe, D., Hammond, R. A., Marsh, H. C. and Perkins, S. J. (2008) The partly-folded back solution structure arrangement of the 30 SCR domains in human complement receptor type 1 (CR1) permits access to its C3b and C4b ligands. J. Mol. Biol. **375**, 102-118.

46 Okemefuna, A. I., Li, K., Nan, R., Ormsby, R. J., Sadlon, T., Gordon, D. L. and Perkins, S. J. (2009) Multimeric interactions between complement Factor H and its C3d ligand provide new insight on complement regulation. J. Mol. Biol. **391**, 119-135.

47 Okemefuna, A. I., Nan, R., Gor, J. and Perkins, S. J. (2009) Electrostatic interactions contribute to the folded-back conformation of wild-type human Factor H. J. Mol. Biol. **391**, 98-118.

Table 1. X-ray scattering and sedimentation coefficient modelling fits for the solution structures of C3 and C3u

	Filter	Models	Hydrated spheres N	R_G (nm) [a]	R_{XS} (nm)	D_{max} (nm) [b]	R-factor (%)	$s^o_{20,w}$ (S)
(a) Experimental								
C3 (50 mM NaCl)				4.87 ± 0.26 4.80 ± 0.18	2.50 ± 0.17	15-20		8.02 ± 0.12
C3 (137 mM NaCl)				4.52 ± 0.08 4.67 ± 0.19	2.54 ± 0.04	16		7.85 ± 0.05
C3u (50 mM NaCl)				5.16 ± 0.22 5.35 ± 0.16	2.56 ± 0.03	16-20		7.66 ± 0.19
C3u (137 mM NaCl)				4.88 ± 0.23 5.07 ± 0.21	2.55 ± 0.10	16		7.44 ± 0.07
(b) Crystal structure								
C3 (PDB 2a73)		1	1863	4.35	2.45	16.4	3.1	8.11
C3b (PDB 2i07)		1	1845	4.64	2.48	17.7	4.8	7.69
C3b (PDB 2icf)		1	1895	4.58	2.52	17.5	4.5	7.56
C3b (PDB 2win)		4	1804-1856	4.59-4.62	2.46-2.50	17.3-17.6	4.6-5.1	7.51-7.59
C3b (PDB 2wii)		1	1856	4.50	2.56	16.0	4.5	7.62
(c) C3u modelling								
Linker 1 search	None	6000	1553-1895	4.15-5.38	1.67-2.83	n. a.	3.8-14.7	n. a.
	N, R_G, R_{XS} and R factor	13	1853-1877	4.75-4.88	2.56-2.65	17.1-18.2	3.8-4.5	7.27-7.53
	Best fit	1	1867	4.78	2.59	18.6	3.8	7.46
Linker 2 search	None	4650	1526-1902	4.08-5.55	1.66-2.90	n. a.	3.6-15.8	n. a.
	N, R_G, R_{XS} and R factor	5	1870-1886	4.73-4.81	2.55-2.68	18.0-19.0	3.6-3.8	7.33-7.53
	Best fit	1	1877	4.78	2.55	19.0	3.6	7.48
Linker 3 search	None	6000	1622-1894	3.97-5.01	1.97-2.85	n. a.	3.7-14.1	n. a.
	N, R_G, R_{XS} and R factor	6	1864-1883	4.77-4.82	2.54-2.58	17.9-18.2	3.9-4.1	7.33-7.46
	Best fit	1	1871	4.77	2.55	17.9	4.1	7.44
Linker 4 search	None	4200	1623-1881	3.96-4.82	1.97-2.78	n. a.	3.8-14.2	n. a.

	N, R_G, R_{XS} and R factor	15	1851-1873	4.74-4.77	2.55-2.67	17.3-18.4	3.9-4.0	7.33-7.59
	Best fit	1	1860	4.74	2.59	18.0	3.9	7.46
Linker 1 and 3 search	None	8000	1453-1890	3.93-5.45	1.52-2.90	n. a.	3.8-17.6	n. a.
	N, R_G, R_{XS} and R factor	13	1853-1888	4.71-4.91	2.50-2.69	16.1-18.9	3.9-4.7	7.29-7.55
	Best fit	1	1878	4.76	2.50	18.9	4.1	7.48

n. a. Not available.

[a] The experimental values are extrapolated to zero concentration. The first value is from the Guinier R_G analyses (Figure 4 and 5) and the second value is from the GNOM *P(r)* analyses (Figure 6).

[b] The values in (a) are from *P(r)* analyses, while the values in (b) and (c) are calculated using HYDROPRO.

FIGURE LEGENDS

Figure 1. Size-exclusion gel filtration and SDS-PAGE analyses of C3 and C3u.
(a,b) The elution of C3 (dashed line) and C3u (continuous line) are labelled as shown. The elution positions of four standards are shown in kDa. In 137 mM NaCl buffer (black), C3 and C3u eluted as a single peak at approximately the same position. In 50 mM NaCl buffer (grey), C3u eluted slightly earlier than C3, and both eluted earlier compared to 137 mM NaCl buffer.
(c) Non-reducing and reducing SDS-PAGE analyses of C3 and C3u from the 137 mM NaCl buffer. The protein markers are labelled at the right side in kDa.

Figure 2. Sedimentation velocity size distribution analyses *c(s)* of C3 and C3u.
(a-d) Representative interference boundary fits for C3 and C3u at about 2 mg/ml in 137 mM NaCl (black) and 50 mM NaCl (grey) buffers. Only every sixth or tenth scan of the fitted boundaries are shown for clarity.
(e) The *c(s)* distributions for C3 in 137 mM (black) and 50 mM (grey) NaCl buffers. In 137 mM NaCl, C3 concentrations ranged from 0.5 mg/ml to 2.2 mg/ml. In 50 mM NaCl, C3 concentrations ranged from 0.3 mg/ml to 2.3 mg/ml. Here and in (f), the dotted lines indicate the movement of $s_{20,w}$ with change of concentration.
(f) The *c(s)* distributions for C3u in 137 mM (black) and 50 mM (grey) NaCl buffers. In 137 mM NaCl, C3u concentrations range from 0.3 mg/ml to 2.4 mg/ml. In 50 mM NaCl, C3u concentrations range from 0.4 mg/ml to 2.1 mg/ml.
(g) The $s_{20,w}$ values for C3 in 50 mM (grey) and 137 mM (black) were fitted by linear regression to result in $s^o_{20,w}$ values at zero concentration of 7.85 ± 0.05 S in 137 mM NaCl and 8.02 ± 0.12 S in 50 mM NaCl. Statistical error bars are shown where visible.
(h) The corresponding $s_{20,w}$ values for C3u in 50 mM (grey) and 137 mM (black) were fitted by linear regression to result in $s^o_{20,w}$ values at zero concentration of 7.47 ± 0.07 S in 137 mM NaCl and 7.66 ± 0.19 S in 50 mM NaCl.

Figure 3. Dimer formation in C3 and C3u from size distribution analyses *c(s)*.
(a) C3u in 50 mM NaCl buffer is shown at 0.7 mg/ml, 1.4 mg/ml and 2.1 mg/ml. The monomer peak is labelled as 1 (Figure 2(e,f) and the dimer peak is labelled as 2.
(b) The concentration dependence of the percentage C3u dimer formation in 50 mM NaCl (grey) and 137 mM NaCl buffers (black) by linear regression.
(c) The concentration dependence of the percentage C3 dimer formation in 50 mM NaCl and 137 mM NaCl buffers. The mean values are shown by the lines.

Figure 4. Experimental X-ray Guinier R_G and cross-section R_{XS} analyses of C3 and C3u.
The Q ranges used for the R_G fits were 0.13 to 0.30 nm^{-1} for C3, and 0.16 to 0.30 nm^{-1} for C3u. . The Q ranges used for the R_{XS} fits were 0.35 to 0.50 nm^{-1} for C3 and C3u. The data points are denoted by open circles; the filled circles were used to determine the R_G (a-d) and R_{XS} (e-h) values based on the best fit lines as shown.
(a,e) C3 concentrations between 0.5 mg/ml to 1.2 mg/ml in 137 mM NaCl (black) from bottom to top.
(b,f) C3 concentrations between 0.2 mg/ml to 2.0 mg/ml in 50 mM NaCl (grey).
(c,g) C3u concentrations between 0.5 mg/ml to 1.1 mg/ml in 137 mM NaCl.
(d,h) C3u concentrations between 0.5 mg/ml to 1.7 mg/ml in 50 mM NaCl.

Figure 5. Concentration dependence of the Guinier R_G, $I(0)/c$ and R_{XS} values for C3 and C3u. Each value was measured in quadruplicate and averaged for linear regression fits. Statistical error bars are shown where visible.

(a, c) The R_G values are shown for 137 mM NaCl (black) and 50 mM NaCl buffer (grey).
(b, d) The corresponding *I(0)/c* values are shown for 137 mM and 50 mM NaCl buffer. The fitted line shown for 137 mM NaCl buffer is the mean value.
(e, f) The R_{XS} values are shown for 137 mM NaCl and 50 mM NaCl buffer.

Figure 6. X-ray distance distribution function *P(r)* analyses for C3 and C3u. For C3 and C3u in 137 mM NaCl, the peak maximum *M* is observed at 5.0 nm. For C3 and C3u in 50 mM NaCl, M is observed at 5.0 to 5.4 nm for C3 and 5.0 to 5.2 for C3u.
(a) C3 between 0.5 mg/ml to 1.2 mg/ml in 137 mM NaCl buffer showed an unchanged maximum length *L* of 16 nm.
(b) C3 in 50 mM NaCl buffer showed *L* values of 15 nm (0.6 mg/ml), 17 nm (0.7 mg/ml) and 20 nm (0.9 mg/ml).
(c) C3u between 0.5 mg/ml to 1.1 mg/ml in 137 mM NaCl buffer showed an unchanged *L* value of 16 nm.
(d) C3u in 50 mM NaCl buffer showed L values of 16 nm (0.5 mg/ml), 17 nm (1.0 mg/ml) and 20 nm (1.7 mg/ml).

Figure 7. X-ray scattering curve fits for C3 and C3u. The experimental data for C3 and C3u in 137 mM NaCl buffer correspond to 1.2 mg/ml and 1.1 mg/ml respectively.
(a) The experimental *I(Q)* and *P(r)* curves for C3 are shown in green. The modelled curves based on the C3 crystal structure (PDB code 2a73)) are shown in black.
(b) The experimental *I(Q)* and *P(r)* curves for C3u are shown in red. The curves for the best-fit C3u model are shown in black. The experimental I(Q) curves for C3 (green) and C3u (red) are also shown superimposed upon each other, where the main difference between the curves is highlighted by a black bar
(c) A ribbon view of the C3 crystal structure, with the TED and CUB domains highlighted in red. Here and in (d), the interdomain linker peptides are highlighted in blue.
(d) A ribbon view of the C3u best fit model, with the TED and CUB domains highlighted in red.

Figure 8. Comparison of ribbon views of the solution and crystal structures for C3u, C3b and C3. The TED and CUB domains are shown in red and pink respectively, the linkers are shown in blue, and the remaining C3c segment is tangerine.
(a) The best-fit C3u solution structure model; (b) Superimposition of four C3b crystal structures (PDB codes 2i07, 2icf, 2wii and 2win); (c) The C3 crystal structure (PDB code 2a73).

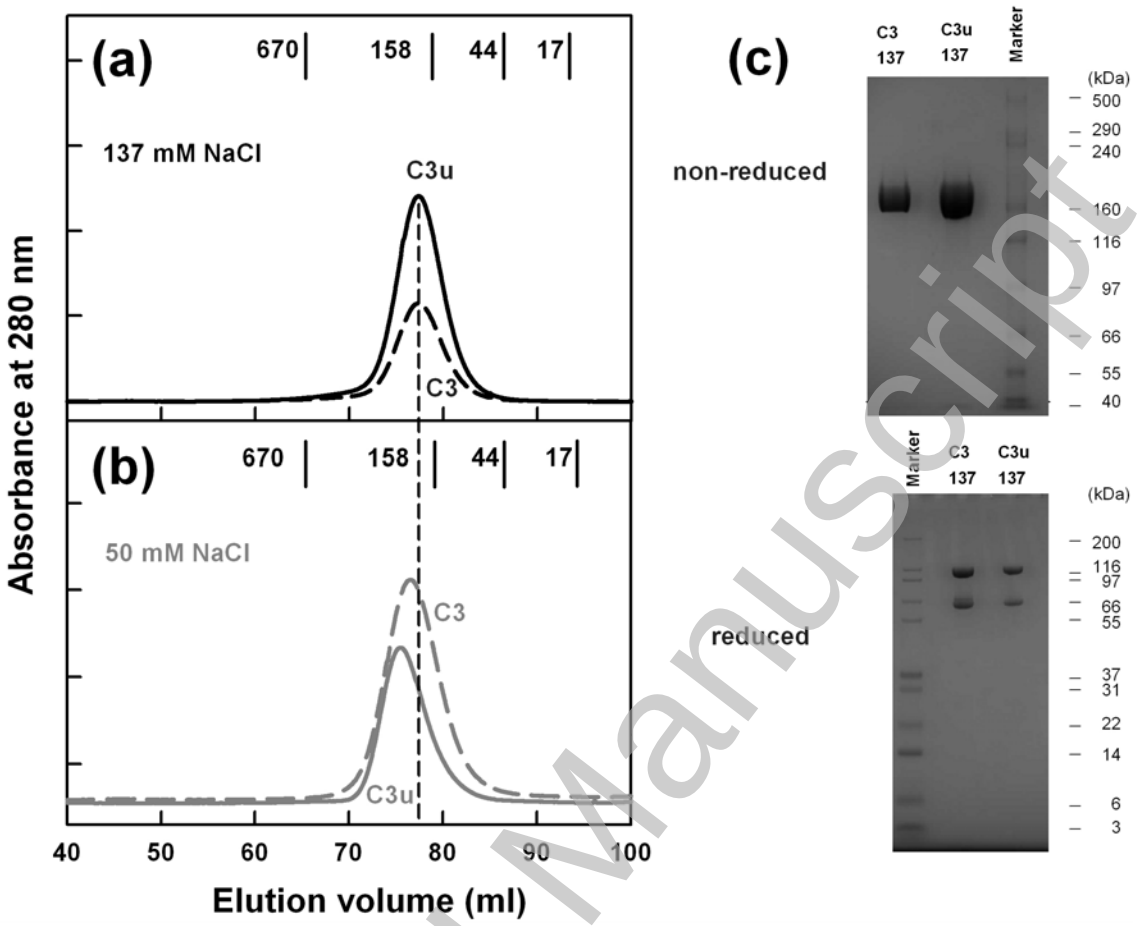

Figure 1

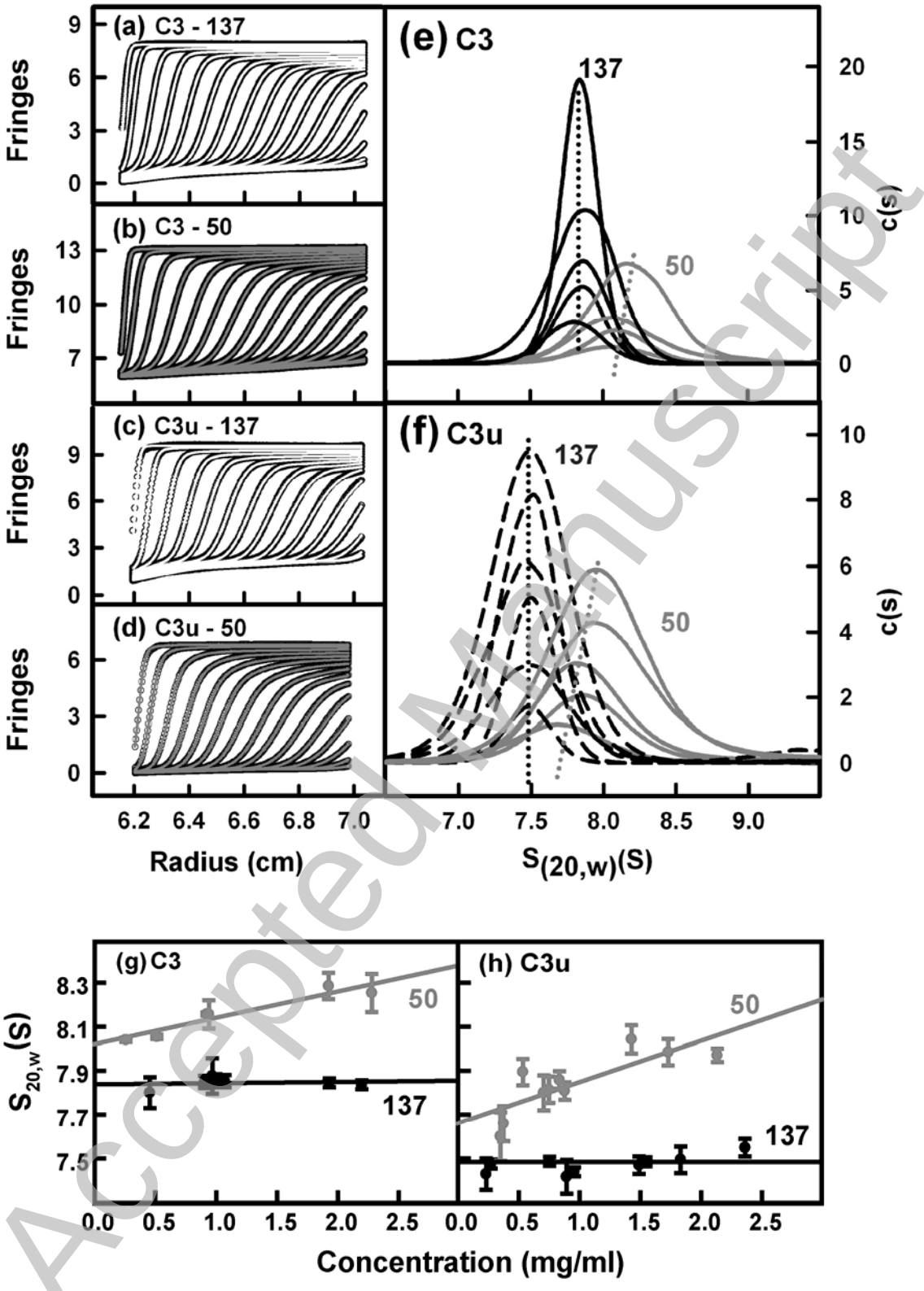

Figure 2

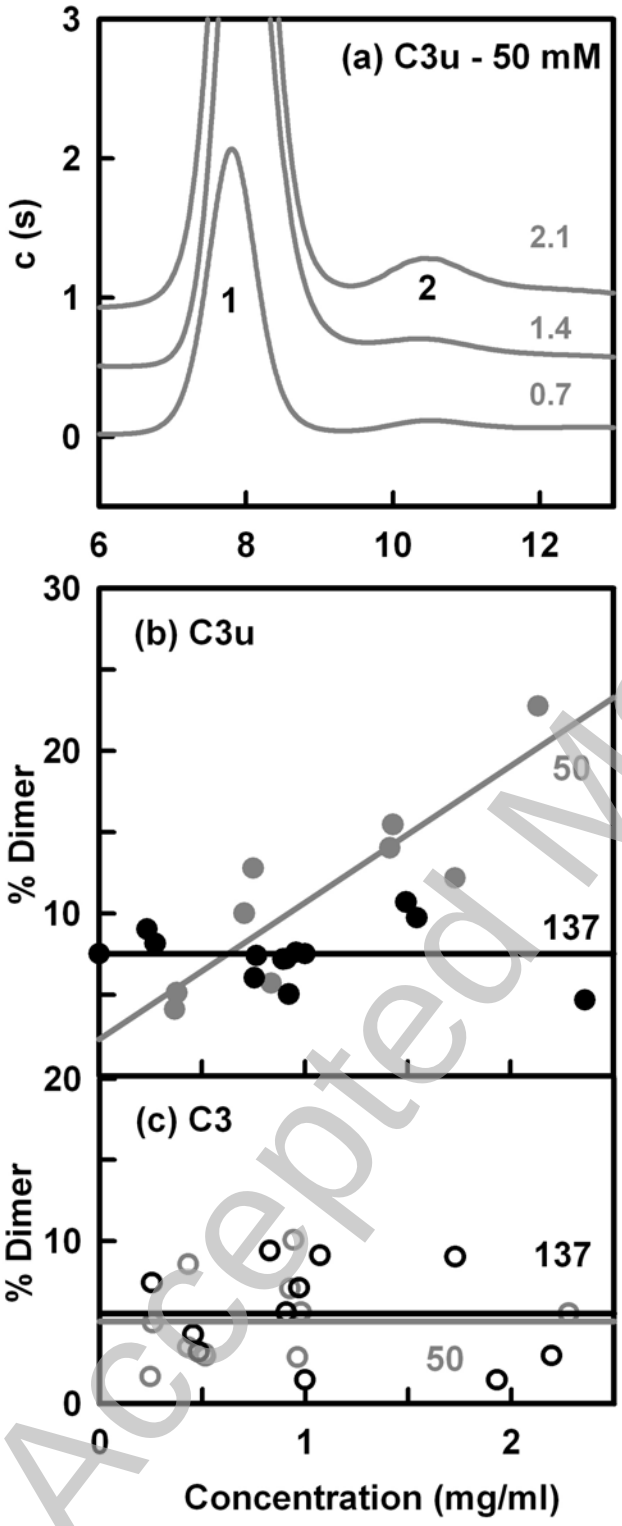

Figure 3

Biochemical Journal Immediate Publication. Published on 29 Jul 2010 as manuscript BJ20100759

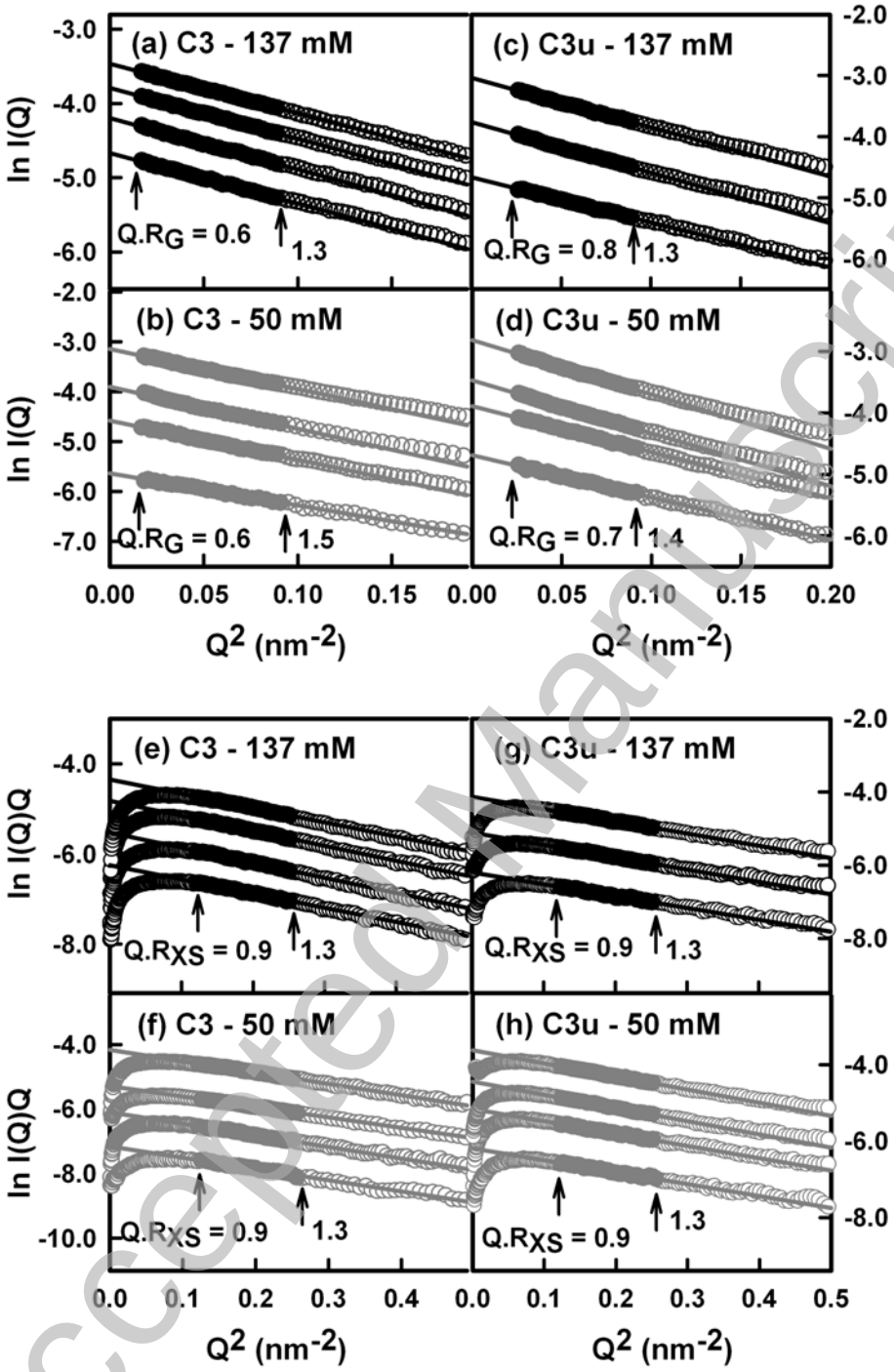

Figure 4

Biochemical Journal Immediate Publication. Published on 29 Jul 2010 as manuscript BJ20100759

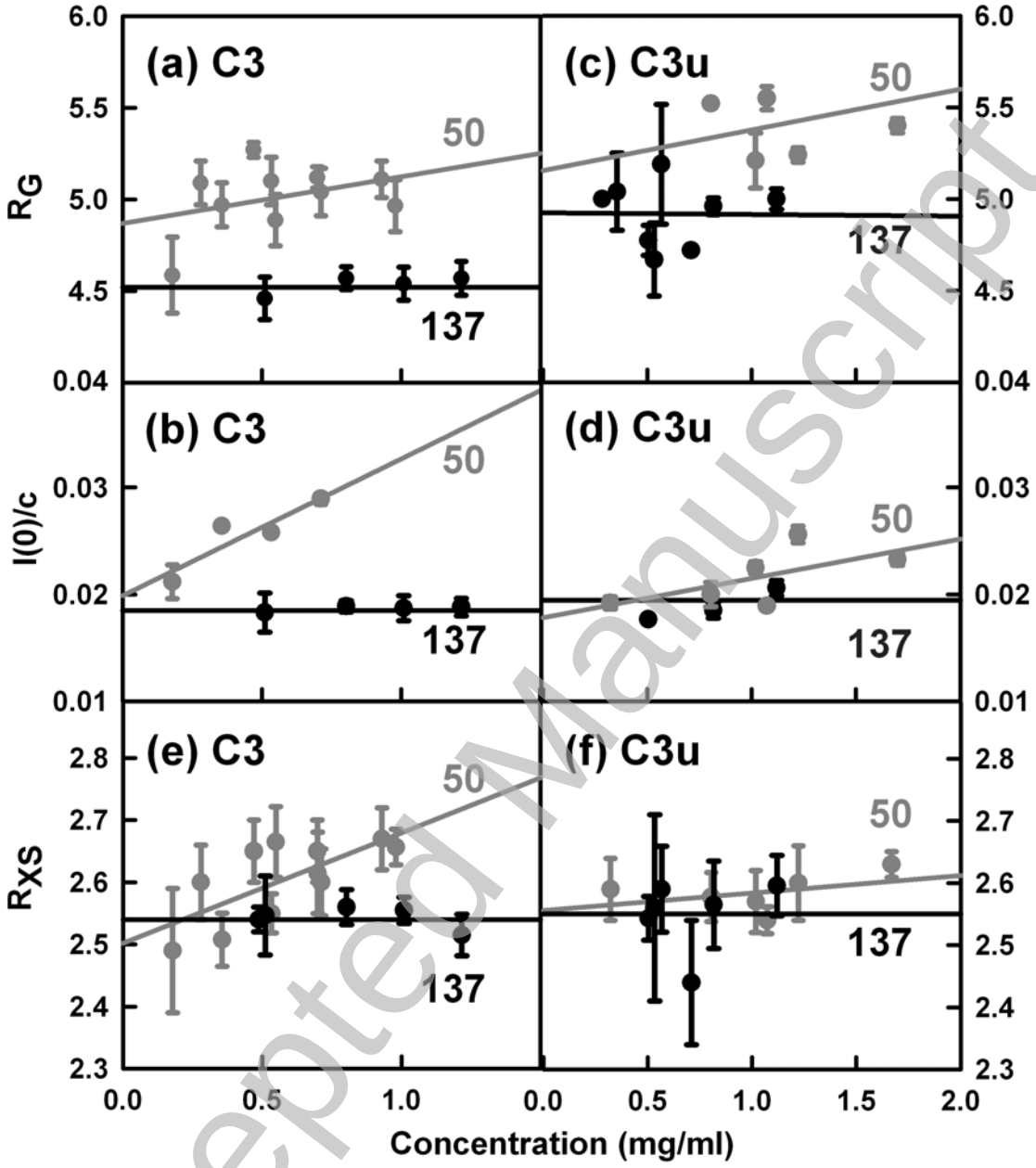

Figure 5

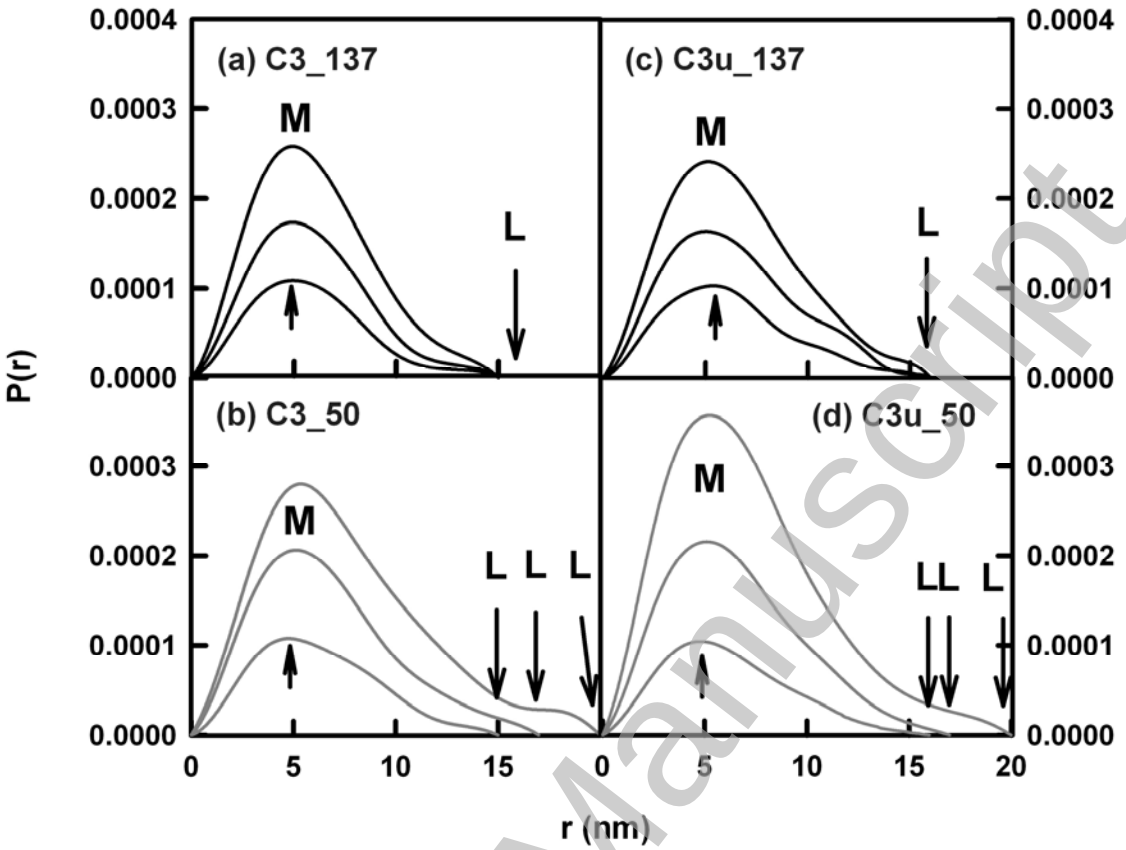

Figure 6

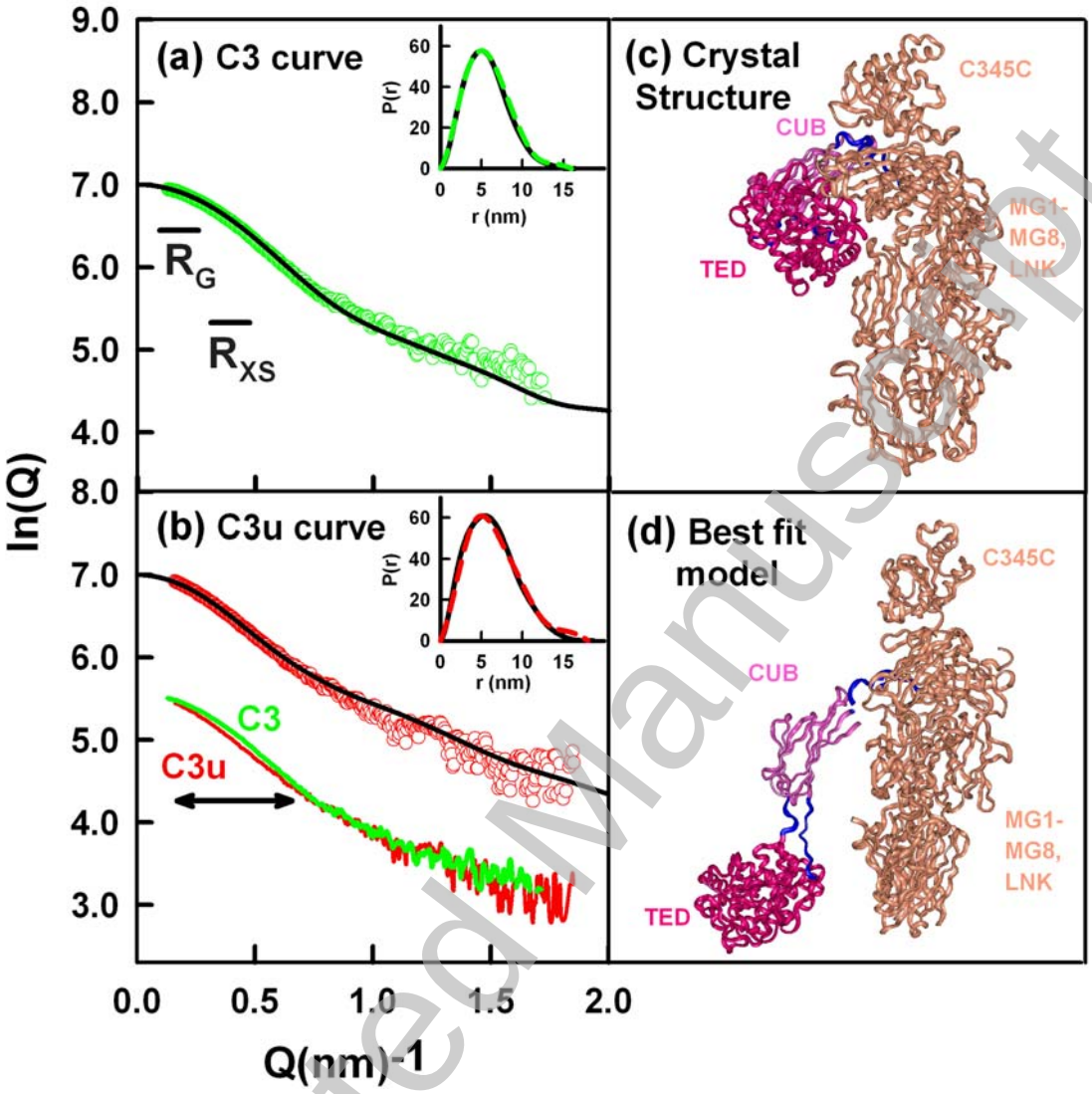

Figure 7

Biochemical Journal Immediate Publication. Published on 29 Jul 2010 as manuscript BJ20100759

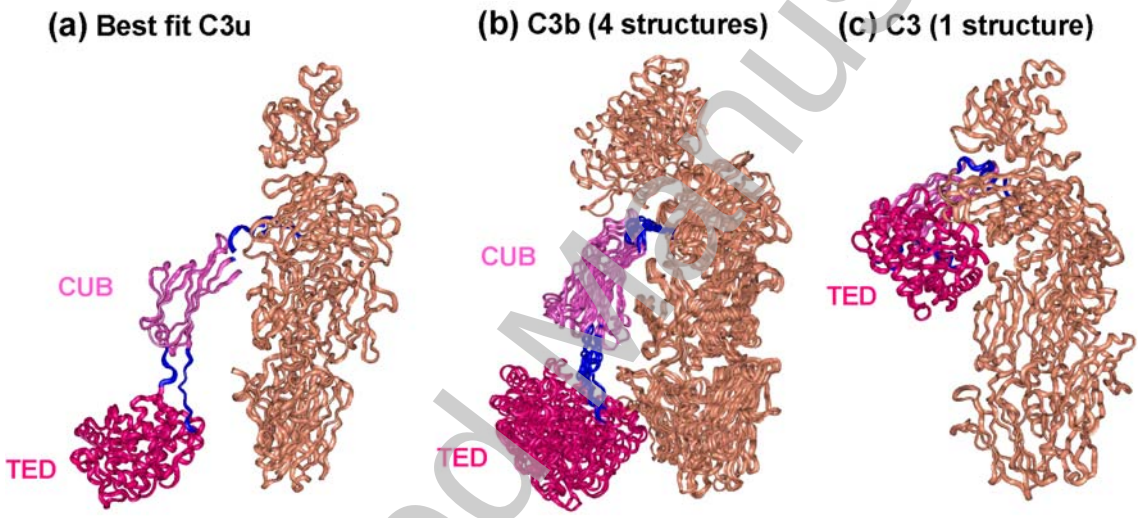

Figure 8