Modelling of the serine-proteinase fold by X-ray and neutron scattering and sedimentation analyses: occurrence of the fold in factor D of the complement system

Stephen J. PERKINS,*§ Kathryn F. SMITH,* John M. KILPATRICK,† John E. VOLANAKIS† and Robert B. SIM‡

*Department of Biochemistry and Chemistry, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, U.K.,

†Department of Medicine, Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, UAB Station, Birmingham, AL 35294, U.S.A., and ‡MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

Solution scattering is a powerful means of determining the overall arrangement of domains in the multidomain proteins of complement. The serine-proteinase domain is central to all proteolytic events during complement activation. As models of this domain, bovine β -trypsin, trypsinogen, α -chymotrypsin and chymotrypsinogen A were studied by neutron and X-ray synchrotron solution scattering. At pH 7, all the X-ray and neutron M_r values corresponded to monomeric proteins. The X-ray radii of gyration, R_g , of β -trypsin, trypsinogen, α -chymotrypsin and chymotrypsinogen A (measured in positive solute-solvent contrasts) were 1.59 nm, 1.78 nm, 1.71 nm and 1.76 nm (± 0.05 – 0.11 nm) in that order. Neutron contrast variation showed that the R_g at infinite contrast, R_c , for these four proteins were 1.57 nm, 1.70 nm, 1.67 nm and 1.78 nm (± 0.03 nm) in that same order. The radial inhomogeneity of neutron-scattering density, α , was positive at $(5\text{--}13) \times 10^{-5}$, and corresponds to the preponderance of hydrophilic residues near the protein surface. On trypsinogen activation, a small reduction in the R_g value of 0.13 ± 0.07 nm was just detectable, while the R_g of chymotrypsinogen A was unchanged after activation. The R_c and α values

of the four proteins can be calculated by using crystallographic co-ordinates. The reduced R_c of β -trypsin relative to trypsinogen was explained in terms of the removal of the extended N-terminal hexapeptide of trypsinogen. The full X-ray and neutron-scattering curves in positive and negative contrasts agreed well with scattering curves calculated from crystallographic co-ordinates to a nominal structural resolution of 4.5 nm, provided that the internal structure was considered in neutron modelling, and that the hydration was considered in X-ray modelling. Sedimentation-coefficient data also provide information on the disposition of domains in multidomain proteins. It was found that the hydrated X-ray sphere model could be directly utilized to calculate sedimentation coefficients. X-ray scattering on factor D showed from its R_g of 1.78 nm that this is monomeric and very similar in structure to β -trypsin. The X-ray-scattering curve of factor D was readily modelled using the β -trypsin crystal structure after allowance for sequence changes. The success of these modellings provides a basis for the constrained modelling of solution scattering data for the multidomain proteins of complement.

INTRODUCTION

Almost all the complement components of immune defence are constructed from the assembly of independently folded structural motifs, modules or domains (Reid et al., 1986; Sim and Perkins, 1989). The three-dimensional arrangement of domains in multidomain proteins is key to the understanding of their physiological roles. Solution scattering is a powerful means of deducing the overall structure of multidomain proteins under near-physiological conditions (Perkins, 1988a,b). Atomic structures are now rapidly becoming available for single domains within multidomain proteins. In order to build up a realistic model of how the domains are located within the whole protein (and thus to determine its high-resolution structure), shape data from solution scattering are essential. Atomic structures for individual domains can be used to refine or constrain solution scattering data (Perkins et al., 1991). The success of quantitative comparisons of neutron and X-ray scattering curves with the corresponding crystal or two-dimensional (2D) n.m.r. structures is central to this strategy. The calculability of scattering curves has been demonstrated for proteins such as myoglobin, lysozyme and hexokinase by starting from their atomic co-ordinates (Stuhr-

mann, 1973; Stuhmann and Fuess, 1976; McDonald et al., 1979). However, proteins (and glycoproteins in particular) are significantly inhomogeneous in their internal scattering density, and this is detectable by scattering measurements in different solvents and buffers. The requirement of the modelling calculations is that, starting from one or more known crystal or 2D n.m.r. structures, the radius of gyration, R_g , any available cross-sectional radii of gyration, and the full scattering curve $[I(Q)]$ can be routinely obtained as a function of the solvent contrast. The calculations can be performed more rapidly in a more straightforward manner if the atomic co-ordinates are represented by an array of small spheres. The sphere model can be readily adapted to other homologous structures by inspection of sequence data without the need for detailed molecular modelling. Internal density inhomogeneities can be incorporated into these models. To date, this method has been critically evaluated for neutron or synchrotron X-ray scattering applied to two glycoproteins, the single-domain α_1 -antitrypsin structure (M_r 51 500) and the multidomain IgM structure and its fragments (M_r between 49 000 and 980 000) (Smith et al., 1990; Perkins et al., 1991).

Among the complement proteins, the serine proteinase (SP) domain is found in single copies in C1r, C1s and C2, and in

Abbreviations used: DFP, di-isopropyl fluorophosphate; NPGB, *p*-nitrophenyl *p*'-guanidinobenzoate; TAME, *p*-tosyl-L-arginine methyl ester; SP, serine proteinase; 2D, two-dimensional.

§ To whom correspondence and requests for reprints should be addressed.

factors B, D and I. In the complement cascade, the first five of these multidomain components are responsible for the key proteolytic events that lead to the activation of C1r, C1s, C2, factor B, C3, C4 and C5, whereas factor I is responsible for the proteolytic inactivation of C3b and C4b (Reid, 1986). β -Trypsin and α -chymotrypsin are archetypal SP structures that can be used to represent these SP domains [see the accompanying paper (Perkins and Smith, 1993)]. The further testing of this modelling method with SPs (M_r 23 500) will facilitate the routine analysis of multidomain protein structures by scattering. It is also necessary to assess whether the difference in structure upon SP activation is low enough to be unimportant in the modelling of multidomain structures containing the SP domain. Calculation of sedimentation coefficients, $s_{20,w}^0$, offers a further source of information on the overall arrangement of domains in multidomain proteins. The availability of experimental $s_{20,w}^0$ data and crystallographic co-ordinates for the SP proteins permits this second method to be evaluated as well.

For these reasons, we describe comparative scattering and sedimentation studies on β -trypsin versus trypsinogen, and α -chymotrypsin versus chymotrypsinogen A, and the comparison of these data with all four sets of crystal structures.

(i) The transition from inactive trypsinogen to active β -trypsin by the proteolytic removal of its N-terminal hexapeptide is typical of SP activation. X-ray crystallography on trypsinogen showed that 41 (out of 229) residues were not observed in the electron-density map, whereas these were visible in β -trypsin (Huber and Bode, 1978). Of these, 35 residues are immediately adjacent to the active site of β -trypsin, so their structural disorder accounts for the much reduced proteolytic activity of trypsinogen compared with β -trypsin. Further crystallographic work at low temperature, together with ^1H -n.m.r. studies at 37 °C, showed that the six N-terminal residues of trypsinogen that are cleaved on activation were predominantly flexible (Perkins and Wüthrich, 1980; Singh et al., 1980; Walter et al., 1982). It was inferred that the remaining 35 residues in trypsinogen adopt a range of conformational substates.

(ii) Chymotrypsinogen A is activated to α -chymotrypsin by the removal of two dipeptides, Ser¹⁴-Arg¹⁵ and Thr¹⁴⁷-Asn¹⁴⁸. Comparison of chymotrypsinogen A and α -chymotrypsin by protein crystallography (Wang et al., 1985) showed that full proteolytic activity is caused by small polypeptide chain rearrangements of about 0.05 nm in magnitude near to the active site that result in the complete formation of the side-chain specificity pocket and the oxyanion hole for catalytic activity. In chymotrypsinogen A, these nearby residues adopt a range of conformational substates in a manner related to that found for trypsinogen.

The X-ray scattering analysis of factor D of complement is then described, and the scattering curve modelling method is adapted to this homologous structure. Factor D is a single-domain complement protein with one known substrate, factor B. It is one of the least abundant complement proteins, has no known inactive precursor form, and is the same protein as mouse and human adipsin (Cook et al., 1985; Min and Spiegelman, 1985; Phillips et al., 1986; Rosen et al., 1989; White et al., 1992).

MATERIALS AND METHODS

Preparations of β -trypsin, α -chymotrypsin and their zymogens

β -Trypsin, the uncleaved main-chain form of bovine trypsin, was separated from α -trypsin, ψ -trypsin and minor polypeptides using a scaled-up version of the Schroeder and Shaw (1968) method, as adapted from Liu et al. (1971) and Fehllhammer and Bode (1975). A sulphopropyl-Sephadex C-50 column (5 cm ×

90 cm) was employed at 4 °C with a buffer comprising 0.1 M Tris, 20 mM CaCl₂, 1 mM benzamidinium and 3 mM NaN₃, pH 7.4–7.5. After centrifugation, 2 g of commercial bovine trypsin (Merck) was applied in 100 ml of buffer at 75–85 ml/h, and washing was continued at the same flow rate. α -Trypsin was eluted after 3.0–3.3 days and β -trypsin after 3.8–4.5 days. The pooled eluates were acidified to pH 3 with HCl, excess benzamidinium was added, and ultrafiltration was used to reduce the volume to 8–9 ml (Amicon 402 and 8AMC concentrators with UM10 membranes). After pre-²H labelling of the buffer salts by evaporation under vacuum from ²H₂O solution, β -trypsin was dialysed into 33 mM CaCl₂/5 mM benzamidinium, pH 3, in 0–100 % ²H₂O with five buffer changes in 24 h. Freshly purchased commercial trypsin was found by this method to contain 60–62 % β -trypsin, 16–17 % α -trypsin, 3–4 % ψ -trypsin, and 17–20 % inert products. SDS/PAGE showed one sharp band (70–72 % β -trypsin) and two additional bands (28–30 % α -trypsin) from commercial trypsin (Weber et al., 1972; Price, 1976). *p*-Nitrophenyl *p*'-guanidinobenzoate (NPGB) assays (Chase and Shaw, 1967) showed 82 % active protein (β -, α - and ψ -trypsin). The neutron scattering of the purified samples was measured at 37 °C in 0, 20, 60, 80 and 100 % ²H₂O buffer contrasts at pH 7; samples were 100 % β -trypsin as shown by SDS/PAGE and NPGB assays. Commercial cationic bovine trypsinogen (Merck) was found to be homogeneous by SDS/PAGE (Weber et al., 1972; Price, 1976). *p*-Tosyl-L-arginine methyl ester (TAME) assays in freshly-prepared solutions showed zero activity (Walsh, 1970). Since these observations parallel those of Bode et al. (1978) and Bode (1979), these trypsinogen samples were used for neutron data collection in these five buffer contrasts, which was accomplished within a few hours after re-adjustment of the pH to 7.

The success of solution scattering experiments on trypsinogen depends on the completion of the measurements before auto-activation to trypsin has taken place. This was important for neutron data collection in ¹H₂O buffers, where counting times are lengthy as the result of the high incoherent background of ¹H₂O and for dilution series which can take several hours to complete. Trypsinogen is stable at pH 3 (Walsh, 1970), and, as the pH is increased, it undergoes conformational transitions until at pH 7–8 it is able to autoactivate (Lazdunski and Delaage, 1967; Perkins and Wüthrich, 1980). Kay and Kassell (1971) showed that autoactivation rates increase with concentration, but decrease in the presence of inhibitors. Ca²⁺ is important to maintain both the protein structure (Sipos and Merkel, 1970; Bode and Schwager, 1975; Gomez et al., 1977; Perkins and Wüthrich, 1980), and the N-terminus hexapeptide (Abita et al., 1969; Radhakrishnan et al., 1969; Cliffe and Grant, 1981). Trypsinogen samples were assayed during and after data collection. Table 1 shows that, in the presence of 5 mM benzamidinium at pH 7, *I*_{0.1}, activation was successfully delayed by several hours. Table 1 also shows that increase of the ionic strength reduces the lag time, as also reported by Lonsdale-Eccles et al. (1979). Trypsinogen solutions at pH 3 in 5 mM benzamidinium/33 mM CaCl₂ were found to be stable, as also were β -trypsin solutions at pH 7. β -Trypsin at pH 3 was occasionally found to lose activity during storage.

Further bovine SP samples (trypsin, T8003; trypsinogen, T1143; α -chymotrypsin, C7762; α -chymotrypsinogen A, C4879; all from Sigma) were treated repetitively with 2 mM di-isopropyl fluorophosphate (DFP; D0879; Sigma) to stabilize the zymogen forms or to inactivate active enzymes. These were purified by ion-exchange chromatography on a Mono S f.p.l.c. column (Pharmacia), equilibrated in 10 mM KH₂PO₄, pH 4.7. Bound proteins were eluted on a linear salt gradient of up to 1 M KCl

Table 1 Autoactivation of trypsinogen solutions

Conditions	[Trypsinogen] (mg/ml)	Mean distance between molecules (nm)	[Ca ²⁺] (mM)	Activation lag-time (h)
In 5 mM benzamidine, 99.8% ² H ₂ O, pH 7.0, 37 °C	18	13	33	> 22–29.5
	77–88	8	33	3–9
	84	8	5	9
	84	8	33	5–7
	84	8	95	3
In 0.1 M Tris (¹ H ₂ O), pH 8.1 (Kay and Kassell, 1971)	0.1	730	50	60
	7.2	18	50	0.5

in the starting buffer. Minor contaminants and minor forms of the proteinases were eliminated by this procedure, and the major peak from each separation was shown to migrate as a single band on SDS/PAGE under reducing and non-reducing conditions. Subsequent samples of the same proteinases were stabilized by repetitive use of the irreversible proteinase inhibitor Pefabloc SC [4-(2-aminoethyl)benzenesulphonyl fluoride; Pentapharm, Basle, Switzerland] at a final concentration of 2 mM. Further samples of materials separated by mono-S chromatography were examined by Superose 12 f.p.l.c. in 10 mM KH₂PO₄ to determine whether aggregates were present, but no aggregation was detected by this method.

Sample preparation for solution scattering

For neutron data collection at 20 °C, the DFP-treated β -trypsin, trypsinogen and chymotrypsinogen A samples were dialysed into a buffer with 12.5 mM sodium/potassium phosphate, 200 mM NaCl, pH 7.0, and 0, 80 or 100% ²H₂O over at least 36 h with four buffer changes. Further neutron data sets were also obtained for untreated commercial trypsinogen, chymotrypsinogen A and α -chymotrypsin (Sigma) at 20 °C in 33 mM CaCl₂/5 mM benzamidine, pH 7.0, in 0, 80 or 100% ²H₂O. For synchrotron X-ray data collection at 20 °C, protein samples were dialysed into buffers containing either 12.5 mM sodium/potassium phosphate, 145 mM NaCl/0.5 mM EDTA, pH 7.0, or 10 mM Pipes/140 mM NaCl, pH 7.1, or if untreated with DFP or Pefabloc, into 33 mM CaCl₂/5 mM benzamidine. Full X-ray dilution series were obtained for DFP-treated β -trypsin, trypsinogen and chymotrypsinogen A, Pefabloc-treated α -chymotrypsin and trypsinogen, and benzamidine-inhibited trypsinogen. For X-ray scattering, human factor D of complement was purified from the urine of a patient with Fanconi's syndrome (Volanakis and Macon, 1987), and concentrated by forced evaporation to 5.7 mg/ml. This was dialysed into 12.5 mM sodium/potassium phosphate/145 mM NaCl/0.5 mM EDTA, pH 7.0, for data collection. The sequence of factor D is taken from White et al. (1992). No oligosaccharide sites are seen in the factor D sequence, and none was detected on assay (Tomana et al., 1985).

Protein concentrations *c* were calculated from the sample absorption at 280 nm using sequence data and the modified Wetlaufer procedure (Perkins, 1986). This gave final $A_{280}^{1\% \cdot 1\text{cm}}$ values of 16.1 (trypsinogen), 16.6 (β -trypsin), 20.6 (chymotrypsinogen A), 21.0 (α -chymotrypsin), and 12.2 (factor D). That for β -trypsin agrees well with a mean experimental value of 15.5,

which is taken from published values of 15.3 (Liu et al., 1971), 15.4 (Keil, 1974), 15.4 (Gomez et al., 1974) and 15.9 (Yon, 1959). That for trypsinogen is also similar to the reported values of 15.0 (Walsh, 1970), 15.3 (Walter et al., 1982), 15.4 (Keil, 1974) and 15.4 (Kay and Kassell, 1971), whereas lower values of 13.7 (Gomez et al., 1974) and 13.9–14.4 (Abita et al., 1969) have, however, been reported. Those for α -chymotrypsin and chymotrypsinogen A are reported to be 20.0–20.7 (Schwert and Kaufmann, 1951; Laskowski, 1955), in good agreement with the calculated values.

Neutron and X-ray solution scattering measurements and n.m.r. spectroscopy

All neutron-scattering data were obtained on Instrument D17 at the Institut-Laue-Langevin, Grenoble, France. This is analogous in construction to Instrument D11 (Ibel, 1976). For the benzamidine-inhibited material, use of a sample-detector distance (*SD*) of 1.42 m and a wavelength (λ) of 1.009 nm resulted in a *Q* range of 0.26–1.58 nm⁻¹ ($Q = 4\pi \sin \theta / \lambda$; 2θ = scattering angle). Four further independent sessions on Instrument D17 were used to complete data collection on the DFP-treated samples, using an *SD* of 1.400 m, a λ of 1.106 nm and detector-main-beam angles of 0° or 20°. This resulted in a usable *Q* range of 0.16–1.41 nm⁻¹ for a detector-main-beam angle of 0°, which is extended to 0.16–3.25 nm⁻¹ for a detector-main-beam angle of 20°. Samples were measured in rectangular quartz Hellma cuvettes of pathlength 2 mm for samples in 80 and 100% ²H₂O buffers and 1 mm for samples in ¹H₂O buffers; these were also used for A_{280} absorbance measurements. Sample and buffer transmissions were measured relative to an empty cell transmission for use in data reduction and analyses. Individual sample counting times typically ranged between 3 min in 100% ²H₂O buffers to 60 min in 0% ²H₂O buffers. Data reduction was performed using standard Grenoble software (RNILS, SPOLLY, RGUIM and RPLOT; Ghosh, 1989). A cadmium background was first subtracted from each scattering curve. The buffer background run was subtracted from that of the sample run, and the result was normalized for the detector response by using a water run from which an empty cell background had been subtracted.

X-ray scattering data were obtained in four independent sessions using the low-angle solution scattering camera at Station 8.2 (Towns-Andrews et al., 1989) at the Synchrotron Radiation Source, Daresbury, Warrington, U.K. Experiments were performed with beam currents in a range of 92–230 mA and a ring energy of 2.00 GeV. Samples were measured for 10 min for protein concentrations that ranged between 2 and 12 mg/ml. The use of a 500-channel quadrant detector with *SD* values of 3.45 m, 3.54 m, 4.16 m and 4.44 m resulted in usable *Q* ranges between 0.17 and 2.19 nm⁻¹ to between 0.12 and 1.56 nm⁻¹. The *Q* range was calibrated using fresh, wet, slightly stretched rat tail collagen, based on a diffraction spacing of 67.0 nm. Samples were held in Perspex (Lucite) cells of sample volume 20 μ l, contained within mica windows of thickness between 10 and 15 μ m. The cells were positioned in a brass temperature-controlled sample holder connected to a water-bath at 20 °C. Buffers and samples were measured in alternation for equal times to minimize background subtraction errors. Data were only accepted if the subsequent Guinier plots were linear and led to reproducible data in repeated measurements. Curves were recorded in ten time frames, which were individually checked as a control for time-dependent radiation damage effects, in particular for the curves used for model-fitting. Data reduction was performed using the standard Daresbury software package

OTOKO [P. Bendall, J. Bordas, M. H. C. Koch and G. R. Mant (European Molecular Biology Laboratory, Hamburg, Germany, and the SERC Daresbury Laboratory, Warrington, U.K.), unpublished software]. The raw scattering curves were normalized on the basis of an ion-chamber monitor positioned after the sample and a detector response measured for 8 h using a uniform ^{55}Fe radioactive source. Reduced curves were calculated by subtraction of the buffer runs from those of the samples.

^1H -n.m.r. spectra of β -trypsin and trypsinogen were recorded at 37 °C at 250 MHz on a Cameca 250 n.m.r. spectrometer in its Fourier-transform mode. Samples were extensively dialysed at pH 7 and 6 °C into 99.98 % $^2\text{H}_2\text{O}$ buffers containing 33 mM CaCl_2 and 5 mM benzamidine, and stored under N_2 . Pulses were separated by 5.9 s to allow full n.m.r. signal-intensity recovery. Protein concentrations were 53–62 mg/ml (2.3–2.6 mM). Activity tests were performed as described above, and also using the self-monitoring of trypsinogen based on the high-field shifted methyl signals (Perkins and Wüthrich, 1980).

Analysis of neutron and X-ray scattering data

At small Q values, analyses of the scattering curves by the Guinier equation gave the radius of gyration R_g (which is a measure of macromolecular elongation) and the forward scattered intensity at zero angle $I(0)$ (which is a measure of composition) (Guinier and Fournet, 1955; Glatter and Kratky, 1982):

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

This expression is valid for $I(Q)$ values in a Q range corresponding to $Q \cdot R_g$ up to about 1.5–2.0 for an ellipsoidal shape with axial ratios close to 1:1:1.8. Absolute relative molecular masses (M_r) were obtained from the neutron $I(0)/c$ values measured in $^1\text{H}_2\text{O}$ buffers with the use of the incoherent scattering of $^1\text{H}_2\text{O}$ as an intensity standard, in which systematic errors of measurements were minimized (Kratky, 1963; Jacrot and Zaccari, 1981). From the X-ray $I(0)/c$ values, relative M_r values can be calculated using the $I(0)/c$ values for other proteins of known concentrations and M_r and measured in the same buffer during the same data collection session. At the Royal Free Hospital School of Medicine, London, neutron and X-ray Guinier analyses were performed jointly using a common interactive analysis and plotting program SCTPL3 (Perkins and Sim, 1986; A. S. Nealis and S. J. Perkins, unpublished software). Results were processed using Microsoft EXCEL spreadsheets. Statistical analyses were performed using MINITAB (Ryan et al., 1985).

In neutron-contrast-variation studies, the protein matchpoint (at which the protein is invisible to neutrons, i.e. when $I(0)$ is zero) was obtained experimentally from a plot of $\sqrt{\{I(0)/cT_s t\}}$ against volume % $^2\text{H}_2\text{O}$ (T_s , sample transmission; t , pathlength). This was also calculated from the amino acid and carbohydrate composition on the basis that the protein volume was unhydrated and there was 10% of non-exchange of the main-chain NH protons (Perkins, 1986). This percentage of NH non-exchange was measured from the ^1H -n.m.r. spectrum in $^2\text{H}_2\text{O}$ of β -trypsin and trypsinogen by integration of the envelope of signals corresponding to the amide NH protons (Bolton et al., 1976) using the aromatic signals of ten Tyr, three Phe, four Trp and three His residues as a standard, but excluding the spectral region of overlap of the ^1H -n.m.r. signals from the His C(2)H, Trp C(4)H and C(7)H resonances with those from benzamidine at 7.61–7.77 p.p.m. downfield from the 2,2-dimethyl-2-silapentane-5-sulphonate reference signal at 0 p.p.m.

Since hydrophilic and hydrophobic residues in the SPs have different scattering densities, the R_g depended on the solvent

scattering density. This contrast-dependence is given by the Stuhrmann equation (Ibel and Stuhrmann, 1975; Perkins, 1988a,b):

$$R_g^2 = R_c^2 + \alpha/\Delta\rho - \beta/\Delta\rho^2$$

where R_c is the R_g at infinite contrast, $\Delta\rho$ is the difference in scattering density (ρ) between the solute and solvent, α is the radial inhomogeneity of scattering density fluctuations within the protein, and β corresponds to the distance between the centres of two components of very different scattering densities. α is positive if the outermost scattering densities are higher than those nearer the centre. β is not usually measurable for glycoproteins.

Indirect transformation of the scattering data in reciprocal space $[I(Q)]$ into that in real space $[P(r)]$ was carried out using the ITP-81 and ITP-91 program of Glatter (1982). This offered an alternative calculation of the R_g that is based on the full scattering curve measured, and gives the maximum dimension of the macromolecule (L). For β -trypsin, the neutron $I(Q)$ in 100 % $^2\text{H}_2\text{O}$ contained 67 data points which extended to $Q = 2.84 \text{ nm}^{-1}$, and which were fitted with five splines assuming a maximum dimension (D_{max}) of 8 nm (in which D_{max} should ideally be 50 % greater than the anticipated value of L), whereas the X-ray $I(Q)$ contained 121 data points extending out to 2.10 nm^{-1} and was fitted with six splines with D_{max} set as 8 nm. For trypsinogen, the neutron $I(Q)$ contained 77 data points extending out to 3.21 nm^{-1} and was fitted with five splines with D_{max} set as 8 nm, whereas the X-ray $I(Q)$ contained 126 data points extending out to 2.22 nm^{-1} and was fitted with five splines with D_{max} set as 8.5 nm. Comparable ITP parameters were used with α -chymotrypsin, chymotrypsinogen A and factor D, which are similar in size to β -trypsin and trypsinogen. In all cases the number of points used to define $P(r)$ in real space was 101. Criteria for an acceptable calculation of $P(r)$ were: (i) $P(r)$ should exhibit positive values; (ii) the R_g from ITP should agree with the R_g from Guinier analyses; (iii) $P(r)$ should be zero when r is zero; (iv) $P(r)$ should be stable and reproducible for different experimental $I(Q)$ curves when the number of splines and D_{max} is varied over a reasonable range. L was determined from $P(r)$ when this became zero at large r ; errors in L can be significant as a result of the low intensity of $P(r)$ in this region.

Modelling of neutron and X-ray scattering data

The crystallographic co-ordinates used to calculate scattering curves for β -trypsin, trypsinogen, α -chymotrypsin and chymotrypsinogen A were taken from the Brookhaven database (Bernstein et al., 1977). Their reference numbers were as follows: β -trypsin (monomeric), 2PTN, 1TPO, 1TLD, 3PTN, 3PTB, 1TPP, 4PTP; trypsinogen (monomeric), 1TGN, 2TGA, 1TGC, 1TGT, 2TGT, 1TGB, 2TGD; α -chymotrypsin (dimeric), 4CHA, 5CHA, 6CHA; chymotrypsinogen A (dimeric), 2CGA. When required, molecular-graphics displays were analysed using Silicon Graphics Personal Iris 4D25TG and Indigo XS24 Workstations with Insight II software (Biosym Technologies Inc.).

Hydrogen co-ordinates were generated at all exchangeable and non-exchangeable positions (Perkins, 1982). The neutron R_g of the full co-ordinate set was then calculated from:

$$R_g^2 = \sum_i b_i r_i^2 / \sum_i b_i$$

where b_i is the neutron scattering length for an atom i (Perkins, 1988b) and r_i is the distance of the i th atom from the centroid of the macromolecule. The Stuhrmann R_c (the R_g at infinite contrast) was obtained by setting all b_i as unity (including the

protons at exchangeable sites), since the contrast difference between different atoms becomes negligible. In the case of X-ray scattering, b_i was replaced by the atomic scattering factor, f_i . The Stuhrmann α was calculated from:

$$\alpha = j^{-1} \sum_j (\rho_j - \rho_m) r_j^2$$

where r_j is the distance between the centroid of each amino acid residue j and the centroid of the protein, ρ_j is the mean scattering density of the residue, and ρ_m is the mean scattering density of the whole protein. The ρ_j and ρ_m values were calculated from the dry volume (Chothia, 1975; Perkins, 1988b).

Calculations of the full scattering curves for the SPs utilized small Debye spheres to follow standard procedures (Glatter and Kratky, 1982) in application to neutron and synchrotron X-ray scattering data (Perkins and Weiss, 1983; Perkins and Sim, 1986). To constrain the simulations, the unhydrated volume (23.2 nm³ and 23.9 nm³ for β -trypsin and trypsinogen respectively) was calculated from the composition using sequence data (Chothia, 1975). For β -trypsin, the crystal co-ordinates were subdivided into 188 cubes of side 0.54 nm. Trypsinogen was represented by 193 cubes (see the Results and discussion section). This cut-off scheme meant that the total volume of the cubes equalled that of the dry protein volume. Overlapping spheres of the same volume as each cube were placed at the centre of each cube. Neutron data correspond largely to this unhydrated volume, since the hydration shell that surrounds the macromolecule is invisible by neutron contrast variation (Perkins, 1986). [This means that partial specific volumes ($\bar{v}$) are larger if calculated from neutron matchpoints instead of by densitometric means.] The volumes of the hydrophilic and hydrophobic residues determined the totals of 88 outermost and 106 innermost spheres in the model, which were assigned manually by inspection of the model. To represent the experimental solute-solvent contrasts, the scattering densities of the more hydrophilic outermost and more hydrophobic innermost spheres were respectively set to 55% ²H₂O and 35% ²H₂O. Similar procedures were followed for α -chymotrypsin and chymotrypsinogen A, in which 204 and 208 spheres respectively were used. The R_c , R_g and α values of the sphere models were calculated as controls of the co-ordinate conversion. The full scattering curve [$I(Q)$] was calculated using the program SCT (S. J. Perkins, unpublished software). Standard wavelength spread and beam divergence corrections were applied to the simulated neutron curves before these were compared with the experimental curves (Cusack, 1981; Perkins and Weiss, 1983). A goodness-of-fit coefficient (R) was calculated to follow the procedure of Smith et al. (1990). Note that a good curve fit does not lead to the determination of a unique structure as the result of spherical averaging which is implicit in solution scattering.

X-ray curve-fitting for the SPs corresponded to a high positive solute-solvent contrast. Models were based on the hydrated volume, which was calculated using a hydration of 0.3 g of water/g of glycoprotein and an electrostricted volume of 0.0245 nm³ per bound water molecule (Perkins, 1986). This choice of hydration is within error of the value of 0.53 ± 0.26 g of water/g of protein determined by Squire and Himmel (1979) for 21 proteins in application to hydrodynamic data. The hydration of 0.3 g of water/g of protein corresponds approximately to a monolayer of water molecules on the protein surface, which is consistent with the current molecular view of protein hydration from high-resolution crystallographic studies (Saenger, 1987; Thanki et al., 1988). The above 188- or 193-sphere Debye models for β -trypsin or trypsinogen were modified with each overlapping

sphere volume reset to be that of the cube of side 0.589 nm. In two-density models, the positive solute-solvent contrast was represented by setting the scattering densities of the more hydrophilic outermost and more hydrophobic innermost spheres to be 405 e/nm³ and 428–430 e/nm³ respectively relative to a solvent density of 334 e/nm³. Similar procedures were also followed for α -chymotrypsin and chymotrypsinogen A. No corrections were made for instrumental errors, since these were negligible with synchrotron radiation.

Calculations of sedimentation coefficients

Modelling of the sedimentation coefficient $s_{20,w}^0$ was based on the hydrated volume as in the X-ray curve-fitting. The frictional coefficient (f) was calculated from:

$$f = M_r (1 - \bar{v} \cdot \rho_{20,w}) / (N \cdot s_{20,w}^0)$$

where $\rho_{20,w}$ is the density of water at 20 °C, and N is Avogadro's constant. The hydrated partial specific volumes ($\bar{v}$) of β -trypsin, trypsinogen, α -chymotrypsin and chymotrypsinogen A were calculated from the sequence to be 0.728 ml/g, 0.726 ml/g, 0.735 ml/g and 0.735 ml/g in that order, again using a hydration of 0.3 g of water/g of glycoprotein and an electrostricted volume of 0.0245 nm³ per bound water molecule (Perkins, 1986). The $\bar{v}$ value for β -trypsin agrees within 0.001 ml/g with the value cited by Squire and Himmel (1979), and those for α -chymotrypsin and chymotrypsinogen A agree within 0.002 ml/g and 0.003 ml/g of those experimentally determined by Lee and Timasheff (1974). For reason of computer-hardware limitations on the CDC Cyber computer previously used for simulations, a relatively small number of hydrodynamic spheres had been used to calculate f (Perkins, 1989). For these, the hydration had been reset to a notional value of 0.39 g of water/g of protein in order to compensate for the volume lost in the void spaces between the non-overlapping spheres required to calculate f (Perkins, 1989). An improved approach is based on the use of the hydrated X-ray scattering model (but now with 188–208 non-overlapping spheres of diameter 0.589 nm) directly in the simulations. This is a closer representation of the protein structure, and is strongly advantageous in that both the scattering and hydrodynamic data can be represented by the same model. This calculation is possible with the Silicon Graphics 4D35S Workstation running under IRIX 4.0 used in the present study. Calculations were performed by the modified Oseen tensor procedure using the program GENDIA (Garcia de la Torre and Bloomfield, 1977a,b). Use of the updated hydrodynamic modelling program HYDRO and TRV (Garcia de la Torre, 1989) was not possible in the present study because the upper limit on the number of spheres that can be handled is just over 100 on presently available hardware.

RESULTS AND DISCUSSION

Neutron and X-ray scattering of β -trypsin and trypsinogen

For β -trypsin at pH 7 in 100% ²H₂O and 0% ²H₂O buffers, representative neutron and X-ray Guinier analyses are shown in Figures 1(b) and 1(f). These exhibited excellent fitted linear plots out to a $Q \cdot R_g$ of 1.4 (Q of 0.93 nm⁻¹). This linearity extended to a $Q \cdot R_g$ of at least 2. For benzamidine-inhibited or DFP-inhibited β -trypsin in the presence or absence of EDTA, no dependence of the Guinier R_g and $I(0)/c$ values was found in the concentration range of 1.3–13.8 mg/ml in 0, 80 or 100% ²H₂O buffers by neutrons or X-rays. Further experiments at pH 3 also showed no difference in the R_g and $I(0)/c$ values of β -trypsin, where n.m.r. showed that the non-activation region of the protein structure is

Figure 1 Representative neutron and X-ray Guinier plots for the four SPs

The $\ln[I(Q)]$ data points used for Guinier fits are denoted by filled symbols. The $Q \cdot R_G$ values at the start and end of the fits are arrowed. The maximum Q employed in fits was 0.93 nm^{-1} . The numbering of the vertical scale is arbitrary, but the relative displacement of each pair of Guinier plots corresponds to the differences in sample concentration. (a–d) The neutron sample concentrations (100% $^2\text{H}_2\text{O}$ buffers) were as follows: benzamidine-inhibited trypsinogen, 11.6 mg/ml and 7.9 mg/ml; DFP-treated β -trypsin, 7.0 and 3.6 mg/ml; DFP-treated chymotrypsinogen A, 10.2 mg/ml and 5.1 mg/ml; α -chymotrypsin, 17.0 mg/ml and 6.9 mg/ml. Note that the α -chymotrypsin neutron samples were affected by small amounts of association. (e–h) The X-ray sample concentrations were as follows: Pefabloc-treated trypsinogen, 11.8 mg/ml and 5.9 mg/ml; DFP-treated β -trypsin, 11.2 and 5.6 mg/ml; DFP-treated chymotrypsinogen A, 5.7 mg/ml and 1.9 mg/ml; Pefabloc-treated α -chymotrypsin, 12.0 mg/ml and 6.0 mg/ml.

affected by the loss of calcium (Perkins and Wüthrich, 1980). In negative solute–solvent contrast in 100% $^2\text{H}_2\text{O}$, the mean neutron R_G of β -trypsin was 1.41 nm, whereas in positive contrast in 0% $^2\text{H}_2\text{O}$, that for neutrons was 1.78 nm and that by X-rays was 1.59 nm (Table 2). The smaller R_G value in 100% $^2\text{H}_2\text{O}$ is attributed to the dependence of the R_G on the contrast (see below). The X-ray R_G of 1.51 nm reported for β -trypsin by Damaschun et al. (1968) is in good agreement with the present determination.

For trypsinogen in Figures 1(a) and 1(e), similar satisfactory neutron and X-ray Guinier analyses for the benzamidine-, DFP- and Pefabloc-inhibited samples are shown. Guinier fits were linear out to a $Q \cdot R_G$ value of 1.6. Again, for these samples, no significant differences in the R_G or $I(0)/c$ values were observed in the concentration range of 1.8–15.6 mg/ml, either at pH 3 or pH 7, or in 0, 80 or 100% $^2\text{H}_2\text{O}$, or in the presence or absence of EDTA. Compared with β -trypsin, the neutron R_G values of

trypsinogen were slightly increased to 1.57 nm in 100% $^2\text{H}_2\text{O}$ and 1.83 nm in 0% $^2\text{H}_2\text{O}$ (Table 2). The mean X-ray R_G value was increased to 1.78 nm. Although the increase in R_G was small at 0.05–0.19 nm, this change was nonetheless reproducibly observed. This may be attributed to the presence of the extended N-terminal activation hexapeptide Val-(Asp)₄-Lys in trypsinogen which is removed on activation to β -trypsin (see the Introduction).

M_r values were calculated from the mean neutron $I(0)/c$ value in $^1\text{H}_2\text{O}$ buffers where systematic errors due to $\bar{v}$ and ^1H – ^2H exchange are minimized. For β -trypsin, an experimental M_r of 21900 was in good agreement with the sequence value of 23200 (Table 2), and confirmed the 280 nm absorption coefficient of $16.6 \text{ litre} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ (see the Materials and methods section). For trypsinogen, an experimental M_r of 24100 was in good agreement with the sequence value of 23900 if the absorption coefficient is taken as $16.1 \text{ litre} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$. Graphs of $\sqrt{[I(0)]/c \cdot t \cdot T_s}$ (see the Materials and methods section) plotted against volume % $^2\text{H}_2\text{O}$ result in the matchpoint determinations of β -trypsin and trypsinogen at which $I(0)$ is zero. These plots exhibited excellent linearity (results not shown). Linear regressions of these data gave matchpoints of 43.0% $^2\text{H}_2\text{O}$ and 43.1% $^2\text{H}_2\text{O}$ for β -trypsin and trypsinogen respectively. Predictions of the matchpoint from the sequence requires knowledge of the non-exchanged exchangeable proton content in the protein. The non-exchanged proton contents of β -trypsin and trypsinogen were determined to be $10 \pm 1\%$ and $7 \pm 1\%$ respectively by ^1H -n.m.r. spectroscopy (see the Materials and methods section). Lazdunski and Delaage (1967) report an estimate of 14% for both β -trypsin and trypsinogen. Interestingly, these values were lower than the 24% reported from neutron diffraction studies of β -trypsin crystals soaked in $^2\text{H}_2\text{O}$ buffer (Kossiakoff, 1982), and this suggests that ^1H – ^2H exchange is more limited in the crystal state. The neutron matchpoint was predicted as 41.3–41.5% $^2\text{H}_2\text{O}$ from the dry volume (Chothia, 1975). The experimental matchpoints were thus close to the predicted values, and the $I(0)/c$ data were seen to be compatible with known M_r values. Together with the linearity of the Guinier plots (Figure 1), these agreements showed that structural data can be derived from the R_G analyses.

The Stuhmann plot of Figure 2(a) summarizes the contrast-dependence of the full neutron R_G data. A linear dependence of R_G^2 on $\Delta\rho^{-1}$ was observed for both proteins. Linear least-squares fitting gave the R_G from the intercept at zero $\Delta\rho^{-1}$ as 1.57 nm for β -trypsin and 1.70 nm for trypsinogen (Table 2). The small decrease of 0.13 nm in R_G on going from trypsinogen to β -trypsin is verified in most of the contrasts used (Figure 2a). The radial inhomogeneity of scattering density (α) was found to be 11×10^{-5} and 13×10^{-5} respectively from the slopes. The positive values of α correspond to the location of residues with higher scattering densities in the surface regions of both proteins, i.e. the hydrophilic amino acid residues, as expected for soluble globular proteins.

The full scattering curves $[I(Q)]$ were analysed using the Indirect Transformation Procedure of Glatter (1982) in order to calculate the distance distribution function $[P(r)]$ (see the Materials and methods section). The small size of the proteins in the observed Q range caused $I(Q)$ to decline relatively slowly at the largest Q used, and the $P(r)$ calculations were more satisfactory if the neutron data extended to the maximum Q of 3.25 nm^{-1} . For β -trypsin and trypsinogen, the R_G and $I(0)$ values agreed to within 0.02–0.03 nm of the Guinier R_G and within 1–2% of the $I(0)$ values. The maximum, M , in the neutron $P(r)$ curves was located at 1.84 nm in both β -trypsin and trypsinogen in 100% $^2\text{H}_2\text{O}$ buffers. The point, L , at which $P(r)$ becomes zero

Table 2 Scattering and sedimentation data for β -trypsin, α -chymotrypsin and their zymogens

	β -Trypsin	Trypsinogen	α -Chymotrypsin	Chymotrypsinogen A
Guinier analyses				
R_G (nm) (0% $^2\text{H}_2\text{O}$)	1.78 ± 0.17 (4)*	1.83 ± 0.07 (5)	1.71 ± 0.16 (3)	1.90 ± 0.12 (6)
R_G (nm) (100% $^2\text{H}_2\text{O}$)	1.41 ± 0.07 (11)	1.57 ± 0.05 (18)	1.69 ± 0.06 (3)	1.69 ± 0.05 (6)
R_G (nm) (X-rays)	1.59 ± 0.08 (6)	1.78 ± 0.05 (18)	1.91 ± 0.03 (6)	1.76 ± 0.11 (7)
M_r (neutrons)	$21\,900 \pm 3400$ (4)	$24\,100 \pm 4000$ (5)	$24\,000 \pm 3000$ (3)	$25\,300 \pm 4200$ (6)
M_r (sequence)	23 200	23 900	25 200	25 600
Matchpoint (% $^2\text{H}_2\text{O}$)	43.0 ± 0.9 (20)	43.1 ± 0.7 (32)	40.3 ± 1.1 (9)	40.1 ± 0.6 (18)
Matchpoint (seq.)	41.5	41.3	40.7†	40.9†
Stuhrmann R_G (nm)	1.57 ± 0.03 (20)	1.70 ± 0.02 (32)	1.67 ± 0.05 (9)	1.78 ± 0.02 (18)
$10^{-5} \times$ Stuhrmann α	11 ± 3	13 ± 2	5 ± 5	13 ± 2
Modelling calculations				
Stuhrmann R_G (nm) (co-ords.‡)	1.616 ± 0.003 (7)	1.62 (7)	1.643 ± 0.002 (6)	1.671 ± 0.002 (2)
$10^{-5} \times$ Stuhrmann α (co-ords.‡)	5	5	5	4
Stuhrmann R_G (nm) (spheres)	1.64	1.72	1.66	1.70
$10^{-5} \times$ Stuhrmann α (spheres)	7	9	6	7
X-ray R_G (nm) (spheres)	1.80	1.86	1.82	1.86
Sedimentation analyses				
Experimental $s_{20,w}^0$ (S)	2.50	2.48–2.7	2.4–2.7	2.4–2.75
Calculated $s_{20,w}^0$ (S)	2.63	2.47	2.56	2.49

* Values in parentheses give the number of determinations used to calculate the mean.

† Calculated assuming 10% non-exchange of main-chain NH protons.

‡ Based on the co-ordinate sets specified in the Materials and methods section.

Figure 2 Neutron-contrast-variation analysis of trypsinogen, β -trypsin, chymotrypsinogen A and α -chymotrypsin

In the Stuhrmann plots of R_G^2 against $\Delta\rho^{-1}$ for β -trypsin (○), trypsinogen (●) and chymotrypsinogen A (●), the mean R_G^2 values at each contrast are shown. For α -chymotrypsin (○), the R_G^2 values obtained by linear extrapolation of the R_G^2 values to zero concentration were used. Error bars denote the full range over which the S.D.s of the R_G^2 data vary. The statistical analyses are summarized in Table 2. (a) Regressions for β -trypsin (—) and trypsinogen (—); (b) Regressions for α -chymotrypsin (—) and chymotrypsinogen A (—).

Figure 3 Distance distribution functions $P(r)$ for trypsinogen, based on the indirect transformation of the scattering curves $I(Q)$

The experimental curve is denoted by the continuous line, the maximum in $P(r)$ is denoted by M , and the maximum dimension is denoted by L . The $P(r)$ were calculated using neutron data in 100% $^2\text{H}_2\text{O}$ at 15.4 mg/ml. The calculated $P(r)$ curve (—) from the neutron $I(Q)$ curve fit based on 100% $^2\text{H}_2\text{O}$ (see Figure 5 below) is shown as well. The corresponding $P(r)$ curves for β -trypsin, α -chymotrypsin and chymotrypsinogen A are similar, and are not shown. Statistical error bars are not shown, but are low except when r is less than 0.8 nm. The $P(r)$ curve was rescaled to an arbitrary M value of 100.

at large r defines the maximum dimension within the protein, although this point may not be well defined for reason of the low intensities in $P(r)$ and larger errors. Figure 3 shows that L is close to 5 nm for trypsinogen. Irrespective of whether β -trypsin or trypsinogen was examined, both $P(r)$ curves exhibited closely similar profiles.

Neutron and X-ray scattering of α -chymotrypsin and chymotrypsinogen A

Less extensive neutron and X-ray solution scattering experiments were carried out for α -chymotrypsin and chymotrypsinogen A in similar buffers to see whether the R_G changed upon zymogen activation. The possibility of α -chymotrypsin dimerization was

Figure 4 Structures of trypsinogen and β -trypsin to show the α -carbon co-ordinates and the representation of the proteins as small spheres

The sphere diameter is 0.54 nm. The N-terminal hexapeptide of trypsinogen is represented by five spheres at the lower left. For clarity, the upper Figures are viewed perpendicular to the lines of spheres, while the lower Figures are rotated slightly to offer a perspective view. N and C denote the N-terminus and C-terminus of the α -carbon co-ordinates.

minimized by working at pH 7 in the relatively low concentration range between 6.0 and 17.0 mg/ml, since this favours the presence of monomers (Schwert, 1951; Neet and Brydon, 1970; Timasheff, 1969; Aune and Timasheff, 1971). The neutron and X-ray Guinier plots of both proteins (Figures 1c, 1d, 1g and 1h) exhibited excellent linearity out to a $Q \cdot R_g$ of 1.8, except for α -chymotrypsin in the presence of $^2\text{H}_2\text{O}$. There, a slight tendency for the neutron $I(0)$ data to increase at low Q values in 80% $^2\text{H}_2\text{O}$, and more so in 100% $^2\text{H}_2\text{O}$, suggested that some partial association of α -chymotrypsin had occurred. This association was successfully circumvented by performing Guinier fits in a slightly larger Q range up to 1.18 nm^{-1} instead of 0.93 nm^{-1} with the 100% $^2\text{H}_2\text{O}$ data, i.e. to a $Q \cdot R_g$ of 2.0 (Figure 1d). The R_g and $I(0)/c$ data for α -chymotrypsin were linearly extrapolated to zero c over the concentration range of 6–17 mg/ml in all three neutron contrasts as well as the X-ray data. For chymotrypsinogen A, no difference in the R_g by neutron scattering was observed before or after DFP treatment. No concentration-dependence of the R_g or $I(0)$ parameters in the range of 1.9–17.6 mg/ml for chymotrypsinogen A was observed in the X-ray or neutron data, and their mean values were used.

The extrapolated neutron R_g of α -chymotrypsin in 0% $^2\text{H}_2\text{O}$ was 1.71 nm, which was within error of the R_g from X-ray scattering of 1.91 nm (Table 2). The corresponding mean neutron and X-ray values for chymotrypsinogen A were 1.90 nm and 1.76 nm respectively (Table 2). These results were supported by the neutron M_r determinations in 0% $^2\text{H}_2\text{O}$ buffers of 24 000 and 25 300 for α -chymotrypsin and chymotrypsinogen A, which were in good agreement with the sequence values of 25 200 and 25 600 respectively. For α -chymotrypsin, these M_r values show that possible effects of dimerization had been circumvented by extrapolation to zero c . X-ray R_g values of 1.80 nm for α -chymotrypsin and 1.81 nm for chymotrypsinogen A at pH 7.0 have been reported by Krigbaum and Godwin (1968), and they are in good agreement with the present data.

The Stuhmann plots of Figure 2(b) summarize the contrast variation data for α -chymotrypsin and chymotrypsinogen A (Table 2). The matchpoint graphs (not shown) exhibited excellent

linearity. For α -chymotrypsin, an experimental matchpoint of 40.3% $^2\text{H}_2\text{O}$ was determined, which agrees well with the predicted value of 40.7% $^2\text{H}_2\text{O}$. For chymotrypsinogen A, the experimental value was 40.1% $^2\text{H}_2\text{O}$, which also agrees well with the predicted value of 40.9% $^2\text{H}_2\text{O}$. The dependence of the R_g values on the contrast (Figure 2) gave R_g values of 1.67 nm and 1.78 nm for α -chymotrypsin and chymotrypsinogen A respectively. The slopes (α) for α -chymotrypsin and chymotrypsinogen A were positive at 5×10^{-5} and 13×10^{-5} respectively, and corresponded to the location of the hydrophilic residues on the protein surface. The R_g s of α -chymotrypsin and chymotrypsinogen A (Figure 2) were thus similar to within 0.11 nm, although less extensive data had been collected than for β -trypsin and trypsinogen above. The lack of detectable change is compatible with the surface location of the proteolytic cleavages in chymotrypsinogen A to excise the Ser¹⁴-Arg¹⁵ and Thr¹⁴⁷-Asn¹⁴⁸ dipeptides that lead to the formation of α -chymotrypsin.

Molecular modelling of the SP scattering curves

Both the Stuhmann neutron R_g and α parameters, the X-ray R_g data and the full $I(Q)$ and $P(r)$ curves were calculated from crystallographic co-ordinates for the four proteins (see the Materials and methods section). This permitted the direct comparison of crystal structures with scattering data (Table 2).

When all the scattering lengths, b_i , were set at unity to meet the condition of infinite contrast, the R_g for β -trypsin was calculated from the crystal co-ordinates to be 1.616 nm (Table 2). This value was increased by 0.02–0.03 nm in comparison with the situation when b_i was set to their individual atomic values to correspond to a finite contrast *in vacuo*. If the exchangeable protons were completely removed from this calculation (since they make part of the macromolecule look like the solvent; Worcester, 1988), the R_g calculation with b_i values of unity showed that the R_g increased by 0.01–0.02 nm. The experimental neutron R_g value of 1.57 ± 0.03 nm was within 0.05 nm of the calculated R_g . The calculation of the Stuhmann α for β -trypsin gave 5×10^{-5} , which was less than, but comparable with, the

Figure 5 Scattering curve-fits for the trypsinogen models

The upper three curves show the neutron $I(Q)$ data and the modelled curve fits for trypsinogen in 0%, 80% and 100% $^2\text{H}_2\text{O}$ buffers. The calculated curves are shown by the continuous line, and the experimental $I(Q)$ data are shown by $\bullet$. To indicate the contrast-dependence of the curve-fits, the modelled curve-fits in 0% and 100% $^2\text{H}_2\text{O}$ are compared with those in 100% and 0% $^2\text{H}_2\text{O}$ respectively (denoted by the broken line). The maximum Q used for the R_G fits of Figure 1 is 0.93 nm^{-1} . The lowest curve corresponds to the curve-fit for the synchrotron X-ray data, using a single-density model with a hydration of $0.3\text{ g of water/g of protein}$. The goodness-of-fit coefficients (R) were determined to be 0.011 , 0.011 and 0.008 for the 0, 80 and 100% $^2\text{H}_2\text{O}$ contrasts in the Q range out to 1.7 nm^{-1} . Similar good curve-fits were also obtained for β -trypsin with R values of 0.017 , 0.019 and 0.006 in 0, 80 and 100% $^2\text{H}_2\text{O}$ buffers (the increased R values were the result of weaker signal-to-noise ratios) and chymotrypsinogen A (R of 0.009 , 0.006 and 0.012 , in that order).

observed neutron value of $(11 \pm 3) \times 10^{-5}$. The R_G or α for trypsinogen could not be correctly calculated because in one crystal structure the four polypeptide segments corresponding to the activation region in trypsinogen possessed no significant electron density in the final Fourier map (Fehlhammer et al., 1977), whereas in the other the N-terminal peptide is not visible (Kossiakoff et al., 1977). The best calculation for trypsinogen gave an apparent R_G of 1.62 nm (Table 2), which was slightly less than the observed neutron R_G value of $1.70 \pm 0.02\text{ nm}$, as expected given the absence of the extended N-terminal hexapeptide on the protein surface in the crystal co-ordinates.

The calculation of the R_G for α -chymotrypsin and chymotrypsinogen A from the crystal co-ordinates gave values of 1.643 nm and 1.671 nm respectively (Table 2). These were comparable with the experimental neutron R_G values of $1.67 \pm 0.05\text{ nm}$ and $1.78 \pm 0.02\text{ nm}$ respectively. This shows that agreement within 0.11 nm of the experimental R_G data for these two proteins can be obtained.

The crystallographic co-ordinates for β -trypsin and trypsinogen were converted into 188 small Debye spheres in order to calculate the full neutron scattering curves $[I(Q)]$ (Figure 4). In order to represent the N-terminal hexapeptide of the zymogen, an additional five spheres were added in an extended line to the trypsinogen model at Ile¹⁶ in protein co-ordinate sets in which Ile-16 was exposed on the protein surface. This turned out to be the case in the Brookhaven co-ordinate sets 1TGT and 2TGT.

The R_G values of the β -trypsin and trypsinogen models were 1.64 nm and 1.72 nm , in good agreement with the experimental neutron R_G values of 1.57 nm and 1.70 nm respectively. This calculation now accounted for the observed slightly increased R_G of trypsinogen relative to β -trypsin. Representation of the sphere model as a two-density object with a hydrophilic surface and hydrophobic core (see the Materials and methods section) resulted in calculated Stuhmann α values of 7×10^{-5} and 9×10^{-5} respectively. These were in reasonable agreement with the observed neutron values of $(11 \pm 3) \times 10^{-5}$ and $(13 \pm 2) \times 10^{-5}$ (Figure 2). That the agreement in α is not closer is attributed to other factors that have not been considered, such as the effect of a monolayer of hydrated water on the protein surface, the choice of $\bar{v}$ values for individual amino acid residues, and the possibility of non-uniform ^1H - ^2H exchange within the protein (Zaccai and Jacrot, 1983; Perkins, 1986; Worcester, 1988), as well as experimental error. These other factors do not, however, appear to be significant in calculations of the R_G values. These results imply that scattering curve-fits should be made under conditions of high positive and negative solute-solvent contrasts in order to minimize the effects of these errors.

Neutron scattering curve-fits for β -trypsin and trypsinogen (Figure 5) were based on $I(Q)$ data obtained using 80% and 100% $^2\text{H}_2\text{O}$ buffers to correspond to high negative solute-solvent contrasts, and 0% $^2\text{H}_2\text{O}$ buffer for the high positive contrast. For trypsinogen, good scattering curve-fits out to a Q value of 1.7 nm^{-1} were observed for all three contrasts in Figure 5. The contrast-dependence of the scattering curves is clearly visible in Figure 5 (----). The goodness-of-fit coefficients (R), which should be less than 0.02 for a good fit (Smith et al., 1990), were determined to be 0.008 – 0.011 . Similar good curve-fits (results not shown) were also obtained for β -trypsin and chymotrypsinogen A (Table 2 and Figure 5). For α -chymotrypsin, the presence of slight association affected the $I(Q)$ data points below Q of 0.6 nm^{-1} ; otherwise good curve-fits were obtained. In addition, comparison of the experimental and modelled $P(r)$ curves for trypsinogen in Figure 3 showed excellent correspondence. In conclusion, for all four SPs, the neutron data in all three contrasts could be readily fitted using unhydrated two-density models derived directly from the crystal co-ordinates.

Synchrotron X-ray curve-fits based on the crystal structures were reasonably successful (Figure 5). Starting from the dry neutron model, a hydration of $0.3\text{ g of water/g of protein}$ caused the R_G to increase by 0.14 – 0.15 nm . Use of a two-density model instead of a single-density model also caused the R_G to increase by 0.04 – 0.05 nm . The R_G values for β -trypsin, trypsinogen, α -chymotrypsin and chymotrypsinogen A were calculated to be 1.80 , 1.86 , 1.82 and 1.86 nm on the basis of a hydrated, single-density, model. These were within 0.08 – 0.21 nm of the observed X-ray R_G values of 1.59 , 1.78 , 1.91 and 1.76 nm (errors ± 0.05 – 0.11 nm) in that order. These agreements concur with the observation that the incorporation of hydration and two-density effects in the modelling have similar effects on the R_G calculation (Perkins et al., 1991).

As exemplified in Figure 5 for the hydrated, single-density, model of trypsinogen, good agreements were obtained between the calculated and experimental X-ray $I(Q)$ curves up to a Q value of 1.6 nm^{-1} , and the R coefficient was 0.014 . Similar fits were readily obtained for β -trypsin, α -chymotrypsin and chymotrypsinogen A in the Q range up to 2 nm^{-1} . The quality of the curve-fit was, however, variable, where either the agreement of fit was very close, as in Figure 5, or this deviated at Q values below 0.4 nm^{-1} or above 1.4 nm^{-1} . Investigation of the most likely cause of this showed that the X-ray buffer background level was much increased below Q of 0.4 nm^{-1} (since this angular range in

Figure 6 Radiation damage effects on factor D

The time-course of the dependence of the R_G and $I(0)$ parameters on X-ray beam exposure for 10 min is shown for four different samples at concentrations of 7.0 mg/ml (● and ▼) and 5.3 mg/ml (○ and ▽).

Q is close to the position of the main beam), and that the difference between the sample and buffer X-ray curves was low when Q was above 1.4 nm^{-1} (when the sample scattered less strongly). Small errors in the buffer subtraction can lead to disproportionate effects on the measured scattering curves on Station 8.2. It may be relevant that the low M_r of these proteins means that they are less strong scatterers. In summary, although the modelling based on crystal structures gave good curve-fits to the X-ray-scattering data, neutron scattering gave more reproducible fits over a larger Q range.

Hydrodynamic analyses of the SPs

The advantage of hydrodynamic simulations is that these provide an independent check on the modelling of the scattering data. The experimentally observed values of the sedimentation coefficient ($s_{20,w}^0$) for trypsinogen was 2.5 S between pH 1.9 to pH 6.9 (Lazdunski and Delaage, 1967), and 2.48 S and 2.7 S by another worker (Walsh, 1970). This agreed well with a calculated $s_{20,w}^0$ value from the hydrated X-ray scattering model for trypsinogen of 2.47 S. An experimental value of 2.50 S for β -trypsin (Walsh, 1970) agreed well with the calculated $s_{20,w}^0$ value of 2.63 S from the scattering model. The $s_{20,w}^0$ for α -chymotrypsin is dependent on its degree of dimerization, which depends on the pH (Schwert, 1949; Aune and Timasheff, 1971). That of 2.4–2.5 S reported for monomeric α -chymotrypsin in phosphate buffer (Schwert and Kaufmann, 1951) or 2.7 S at low ionic strength ($I(0.03)$) (Tellam and Winzor, 1977) was in good agreement with the calculated value of 2.56 S for the X-ray-scattering model of α -chymotrypsin. No dimerization is observed for chymotrypsinogen A. Values of 2.5 S at pH 4–6 (Schwert, 1951), 2.7 S at

pH 7.92 (Massey et al., 1955), 2.58 S at pH 7.5 and pH 9.30 (Dreyer et al., 1955; Nichol, 1968), and 2.4–2.75 S at pH 4.0–6.2 (Neet and Brydon, 1970) for chymotrypsinogen A were all in good agreement with the calculated value of 2.49 S from the X-ray-scattering model. In summary, the mean difference between experiment and calculation was ± 0.12 S (12 values), and the largest difference seen in any of these comparisons was 0.26 S (Table 2). This agreement is satisfactory when compared with other hydrodynamic simulations (Perkins, 1989). This success shows that the hydrated X-ray-scattering sphere models (based on a hydration of 0.3 g of water/g of protein) can be employed directly in the calculation of $s_{20,w}^0$ data. For comparison, recalculation on the basis of hydrations of 0 g and 0.6 g of water/g of protein shows that the $s_{20,w}^0$ values are increased by 9.1 % (0.23 S) and decreased by 7.2 % (−0.18 S) respectively. The choice of hydration level thus has a moderate effect on the outcome of the $s_{20,w}^0$ calculations.

Solution scattering analyses of factor D of complement

In order to compare the scattering curves of the SPs with that of the single-domain SP factor D of complement, synchrotron X-ray scattering was used to study factor D in the concentration range 1.8–7.0 mg/ml. Unlike β -trypsin, α -chymotrypsin and their zymogens, initial data analyses showed that factor D was significantly aggregated. Guinier analyses of the ten individual 1 min time frames obtained during data collection (Figure 6) showed that this was caused by radiation damage, and that to a good approximation both the R_G and $I(0)$ parameters increased linearly with time. In one analysis, the corrected R_G and $I(0)$ values for each run could be determined by linear extrapolation to zero time, based on Guinier fits in the Q range $0.40\text{--}0.93 \text{ nm}^{-1}$ for each time frame, as used in Figure 1. The mean of the four extrapolated R_G determinations at 5.3 and 7.0 mg/ml gave $1.82 \pm 0.03 \text{ nm}$. Alternatively, the ten time-frame curves were extrapolated to zero time on the assumption that each time frame represented different points in a concentration series which could be analysed by the Indirect Transformation Program ITP-91. Guinier analyses of the zero-time curves in the Q range $0.030\text{--}0.093 \text{ nm}^{-1}$ obtained from each of eight full curves showed excellent linearity (Figure 7a), and resulted in a mean radiation-damage-free R_G determination of $1.78 \pm 0.04 \text{ nm}$, in good agreement with the preceding analysis. No concentration-dependence of R_G or $I(0)/c$ was observed. Both R_G determinations were close to the values of 1.78 nm (0% $^2\text{H}_2\text{O}$, neutrons) and 1.59 nm (X-rays) obtained for β -trypsin above, and also to the R_G values of 1.83 nm (0% $^2\text{H}_2\text{O}$, neutrons) and 1.78 nm (X-rays) for trypsinogen. Since the M_r of factor D is 24400 (228 residues) and is similar to those of 23200 and 23900 respectively for β -trypsin and trypsinogen (223 and 229 residues), this close agreement in the R_G of factor D with that of β -trypsin is as expected.

Indirect Transformation of the $I(Q)$ curves to give $P(r)$ was based on the ITP-91 version of Glatter's Transformation, in which the scattering curve at 7.0 mg/ml of factor D was extrapolated to zero time as described above to remove radiation damage, and transformed into the $P(r)$ curve (Figure 7b). The zero-time $P(r)$ curve of Figure 7(b) was significantly different from the non-corrected $P(r)$ curve. Comparison of the zero-time $P(r)$ curve with those calculated for β -trypsin and trypsinogen (Figure 3) showed that the maximum (M) in $P(r)$ was at 2.24 nm, which was close to values of M of 1.68–1.96 nm for β -trypsin and trypsinogen. The maximum dimension (L) of factor D was close to 5.5 nm and similar to that reported in Figure 3 for trypsinogen.

Molecular modelling of the factor D scattering curve was based on the adaptation of the β -trypsin model of Figure 4 to

Figure 7 Synchrotron X-ray scattering analyses for factor D of complement

(a) The Guinier plots correspond to sample concentrations of 7.0, 5.3 and 3.5 mg/ml. The notation is as for Figure 1. (b) The $P(r)$ distance distribution function was based on data at a concentration of 7.0 mg/ml. —, Curve extrapolated to zero time; ---, modelled curve; ·····, radiation-damaged original curve. (c) The scattering curve-fit for factor D starting from the β -trypsin sphere model of Figure 4 and the sequence of factor D. The region of the curve used for Guinier R_G fits is arrowed.

allow for sequence insertions and deletions as shown in Figure 1 of the accompanying paper (Perkins and Smith, 1993). The final dry model with 198 spheres and an R_G of 1.66 nm was derived by adding 11 spheres at the positions of eight small sequence insertions in factor D, and deleting one sphere just after β -strand L. On hydration of this model by 0.3 g of water/g of protein, the R_G increased to 1.82 nm, which was in good agreement with the observed value of 1.78 ± 0.04 nm. The use of a two-density model increased the R_G further to 1.87 nm. Comparison of the Debye curve calculated from the model with the R_G of 1.82 nm with the

experimental X-ray curve extrapolated to zero time in Figure 7(c) gave excellent agreement up to a Q of 1.2 nm^{-1} , and satisfactory agreement to Q of 2.2 nm^{-1} . The R coefficient of this fit was 0.014 in the full Q range, which was considered to be satisfactory. Good agreement was also obtained between the experimental and modelled $P(r)$ curves in Figure 7(b). It is concluded that the overall solution structure of factor D is well modelled as a monomeric protein.

CONCLUSIONS

Structure of factor D and other SPs by scattering

Control of the reactivity of SPs is crucial to their physiological role. For both β -trypsin and α -chymotrypsin this is achieved by the use of a deactivated proenzymic form which is activated by cleavage so that the key salt bridge with Asp¹⁹⁴ is formed in order to yield a well-organized catalytic site. This is a localized structural rearrangement. Both neutron and X-ray scattering and the molecular modelling of the four proteins show that no gross structural rearrangement of the SP fold is detectable during activation. This is strongly indicative that the basic SP structure can be employed in scattering curve modelling of the complement proteins containing SP domains, irrespective of the activation state of the SP domain.

Scattering curve modelling of proteins homologous in structure to a known crystal structure can be carried out. In application to factor D, scattering showed first that, under conditions close to physiological, this was a fully monomeric protein at pH 7 between 1.8–7.0 mg/ml, with an R_G value close to that of β -trypsin. The sequence analyses of Perkins and Smith (1993) then showed that the β -trypsin crystal structure could be readily adapted to model successfully the scattering data, as shown in Figure 7(c).

The importance of the modelling of scattering curves for multidomain proteins, in which a domain has a recognized structural homology with a known atomic structure, is that the scattering analyses are constrained by the knowledge of the crystal or 2D n.m.r. structure, and thus become more informative. Two types of analyses are possible. For multidomain proteins such as C1 inhibitor and factor I, one domain is homologous with a known structure. Here solution scattering can be used to derive overall dimensions for both the macromolecule and the other non-homologous domains to see where the latter are positioned relative to the known structure. For other proteins, such as factor H, the atomic structure of all the domains are known. Here, it is then possible to model their disposition within the whole structure by a fully constrained procedure which allows for the correct connectivity between the domains.

Molecular-modelling strategies for multidomain proteins

The same rapid molecular-modelling strategy using small spheres for scattering curves of homologous proteins starting from known atomic structures has been applied to proteins that range in M_r values from 24000 to 980000. In the three studies to date (the present study; Smith et al., 1990; Perkins et al., 1991), this procedure is summarized as follows.

(i) A sequence alignment with the known crystal structure is required that can be strongly justified on the grounds of conserved catalytic residues, buried disulphide bridges, solvent-exposed glycosylation sites, and the surface location of sequence insertions and deletions.

(ii) The crystal or 2D n.m.r. co-ordinates are converted into cubes of total volume set to that predicted from its known sequence (Chothia, 1975; Perkins, 1986). Overlapping spheres of

the same volume as the cubes are placed at the centre of each cube (Figure 4); their total is edited in the light of the sequence alignment above and comparison with the co-ordinates.

(iii) This dry model is used directly to calculate neutron-scattering curve-fits. Internal-scattering-density inhomogeneities are allowed by the assignment of the outermost and innermost spheres in the ratio of hydrophilic and hydrophobic residues to densities equivalent to about 50–55% $^2\text{H}_2\text{O}$ and 30–35% $^3\text{H}_2\text{O}$ (Figure 5).

(iv) For X-ray analysis, the dry model is converted on the basis of adding a volume of hydration of 0.3 g of water/g of glycoprotein into the total dry volume, which entails an appropriate increase in the sphere diameter. Hydration effects on a protein structure studied by X-ray scattering are equivalent in their effect to the addition of surface carbohydrate residues or to dividing the sphere model into hydrophilic and hydrophobic regions of differing electron densities. Minor variation in the hydration level within a range of 0.2–0.3 g of water/g of protein can be sufficient to allow for these three factors for modelling purposes (Perkins et al., 1991).

(v) Curve fits should be made under conditions of both high positive and high negative solute-solvent contrasts. The internal inhomogeneity of the protein or glycoprotein can have a large effect on its measured scattering curve, and the use of both contrasts ensures that no large systematic errors will occur in scattering curve-modelling.

Guidelines to the extent of agreement expected between the scattering data and the crystal structure modelling can be obtained from the observations that the R_G values are generally in agreement to within ± 0.1 nm as an upper limit. Since the Stuhmann α is less directly measured than the R_G data, only semi-quantitative agreement with the internal inhomogeneity of the protein can be obtained (Figure 2). In curve modelling of $I(Q)$, good fits in a Q range extending to approx. 2 nm^{-1} are possible with a goodness-of-fit coefficient (R) of less than 0.02. The hydrodynamic analyses based on the final scattering model should result in experimental and calculated $s_{20,w}^0$ values that are within ± 0.3 S of each other (Perkins, 1989). It should, however, be remembered that a good curve-fit does not correspond to a unique structure determination, since the observed structure is spherically averaged over all macromolecular orientations by solution scattering.

We thank the Wellcome Trust and the Royal Society for support, and Mr. A. S. Nealis for computational assistance. Initial solution scattering was carried out at the European Molecular Biology Laboratory (EMBL) in Grenoble during an EMBL Fellowship held by S.J.P., and support from Dr. F. Borras and Dr. B. Schoot is gratefully acknowledged. Beamtime at the Institut-Laue-Langevin (Grenoble, France) and the Synchrotron Radiation Source (Daresbury, Warrington, U.K.) was provided with support from the Science and Engineering Research Council. We thank Dr. J. Torbet, Dr. A. de Geyer and Dr. W. Bras for their support during data collection.

REFERENCES

- Abita, J. P., Delaage, M., Lazdunski, M. and Savrda, J. (1969) *Eur. J. Biochem.* **8**, 314–324.
- Aune, K. C. and Timasheff, S. N. (1971) *Biochemistry* **10**, 1609–1617.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) *J. Mol. Biol.* **112**, 535–542.
- Bode, W. (1979) *J. Mol. Biol.* **127**, 357–374.
- Bode, W. and Schwager, P. (1975) *J. Mol. Biol.* **98**, 693–717.
- Bode, W., Schwager, P. and Huber, R. (1978) *J. Mol. Biol.* **118**, 99–112.
- Bolton, P. H., Jones, C. R., Bastedo-Lerner, D., Wong, K. L. and Kearns, D. R. (1976) *Biochemistry* **15**, 4370–4377.
- Chase, T. and Shaw, E. (1967) *Biochem. Biophys. Res. Commun.* **29**, 508–514.
- Chothia, C. (1975) *Nature (London)* **254**, 304–308.
- Cliffe, S. G. R. and Grant, D. A. W. (1981) *Biochem. J.* **193**, 655–658.
- Cook, K. S., Groves, D. L., Min, H. Y. and Spiegelman, B. M. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6480–6484.
- Cusack, S. (1981) *J. Mol. Biol.* **145**, 539–541.
- Damaschun, G., Fichtner, P., Pürschel, H.-V. and Reich, J. G. (1968) *Acta Biol. Med. Germ.* **21**, 309–316.
- Dreyer, W. J., Wade, R. D. and Neurath, H. (1955) *Arch. Biochem. Biophys.* **59**, 145–156.
- Fehlhammer, H. and Bode, W. (1975) *J. Mol. Biol.* **98**, 683–692.
- Fehlhammer, H., Bode, W. and Huber, R. (1977) *J. Mol. Biol.* **111**, 415–438.
- Garcia de la Torre, J. (1989) in *Dynamic Properties of Biomolecular Assemblies* (Harding, S. E. and Rowe, A. J., eds.), pp. 3–31, Royal Society of Chemistry, Cambridge.
- Garcia de la Torre, J. and Bloomfield, V. A. (1977a) *Biopolymers* **16**, 1747–1761.
- Garcia de la Torre, J. and Bloomfield, V. A. (1977b) *Biopolymers* **16**, 1779–1793.
- Ghosh, R. E. (1989) Internal Publication 89GH02T, Institut-Laue-Langevin, Grenoble.
- Glatter, O. (1982) in *Small Angle X-ray Scattering* (Glatter, O. and Kratky, O., eds.), pp. 119–196, Academic Press, London.
- Glatter, O. and Kratky, O. (eds.) (1982) *Small-Angle X-ray Scattering*, Academic Press, New York.
- Gomez, J. E., Birnbaum, E. R. and Darnell, D. W. (1974) *Biochemistry* **13**, 3745–3750.
- Gomez, J. E., Birnbaum, E. R., Royer, G. P. and Darnell, D. W. (1977) *Biochim. Biophys. Acta* **495**, 177–182.
- Guinier, A. and Fournet, G. (1955) *Small-Angle Scattering of X-rays*, Wiley, New York.
- Huber, R. and Bode, W. (1978) *Acc. Chem. Res.* **11**, 114–122.
- Ibel, K. (1976) *J. Appl. Crystallogr.* **9**, 269–309.
- Ibel, K. and Stuhmann, H. B. (1975) *J. Mol. Biol.* **93**, 255–266.
- Jacrot, B. and Zaccari, G. (1981) *Biopolymers* **20**, 2413–2426.
- Kay, J. and Kassell, B. (1971) *J. Biol. Chem.* **246**, 6661–6665.
- Keil, B. (1974) in *The Enzymes* (Boyer, P. D., ed.), 3rd edn., vol. 3, pp. 249–275, Academic Press, New York.
- Kossiakoff, A. A. (1982) *Nature (London)* **296**, 713–721.
- Kossiakoff, A. A., Chambers, J. L., Kay, L. M. and Stroud, R. M. (1977) *Biochemistry* **16**, 654–664.
- Kratky, O. (1963) *Prog. Biophys. Chem.* **13**, 105–173.
- Krigbaum, W. R. and Godwin, R. W. (1968) *Biochemistry* **7**, 3126–3131.
- Laskowski, M. (1955) *Methods Enzymol.* **2**, 8–26.
- Lazdunski, M. and Delaage, M. (1967) *Biochim. Biophys. Acta* **140**, 417–434.
- Lee, J. C. and Timasheff, S. N. (1974) *Biochemistry* **13**, 257–265.
- Liu, W. K., Trzeciak, H., Schüssler, H. and Meienhofer, J. (1971) *Biochemistry* **10**, 2849–2855.
- Londsdale-Eccles, J. D., Kerr, M. A., Walsh, K. A. and Neurath, H. (1979) *FEBS Lett.* **100**, 157–160.
- Massey, V., Harrington, W. F. and Hartley, B. S. (1955) *Discuss. Faraday Soc.* **20**, 24–32.
- McDonald, R. C., Steiz, T. A. and Engelman, D. M. (1979) *Biochemistry* **18**, 338–342.
- Min, H. Y. and Spiegelman, B. M. (1985) *Nucleic Acids Res.* **14**, 8879–8892.
- Neet, K. E. and Brydon, S. E. (1970) *Arch. Biochem. Biophys.* **136**, 223–227.
- Nichol, J. C. (1968) *J. Biol. Chem.* **243**, 4065–4069.
- Perkins, S. J. (1982) *Biol. Magn. Reson.* **4**, 193–336.
- Perkins, S. J. (1986) *Eur. J. Biochem.* **157**, 169–180.
- Perkins, S. J. (1988a) *Biochem. J.* **254**, 313–327.
- Perkins, S. J. (1988b) *New Comp. Biochem.* **18B**, 143–264.
- Perkins, S. J. (1989) in *Dynamic Properties of Biomolecular Assemblies* (Harding, S. E. and Rowe, A. J., eds.), pp. 226–245, Royal Society of Chemistry, London.
- Perkins, S. J. and Sim, R. B. (1986) *Eur. J. Biochem.* **157**, 155–168.
- Perkins, S. J. and Smith, K. F. (1993) *Biochem. J.* **295**, 109–114.
- Perkins, S. J. and Weiss, H. (1983) *J. Mol. Biol.* **168**, 847–866.
- Perkins, S. J. and Wüthrich, K. (1980) *J. Mol. Biol.* **138**, 43–64.
- Perkins, S. J., Nealis, A. S., Sutton, B. J. and Feinstein, A. (1991) *J. Mol. Biol.* **221**, 1345–1366.
- Phillips, M., Djan, P. and Green, H. (1986) *J. Biol. Chem.* **261**, 10821–10827.
- Price, N. C. (1976) *Anal. Biochem.* **73**, 447–457.
- Radhakrishnan, T. M., Walsh, K. A. and Neurath, H. (1969) *Biochemistry* **8**, 4020–4027.
- Reid, K. B. M. (1986) *Essays Biochem.* **27**, 27–68.
- Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T. and Tack, B. F. (1986) *Immunol. Today* **7**, 230–234.
- Rosen, B. S., Cook, K. S., Yaglom, J., Groves, D. L., Volanakis, J. E., Damm, D., White, T. and Spiegelman, B. M. (1989) *Science* **244**, 1483–1487.
- Ryan, B. F., Joiner, B. L. and Ryan, T. A. (1985) *Minitab Handbook*, 2nd edn. PWS-Kent Publishing Company, Boston.
- Saenger, W. (1987) *Annu. Rev. Biophys. Chem.* **16**, 93–114.
- Schroeder, D. D. and Shaw, E. (1968) *J. Biol. Chem.* **243**, 2943–2949.
- Schwert, G. W. (1949) *J. Biol. Chem.* **179**, 655–664.
- Schwert, G. W. (1951) *J. Biol. Chem.* **190**, 799–806.
- Schwert, G. W. and Kaufmann, S. (1951) *J. Biol. Chem.* **190**, 807–816.
- Sim, R. B. and Perkins, S. J. (1989) *Curr. Top. Microbiol. Immunol.* **153**, 209–222.
- Singh, T. P., Bode, W. and Huber, R. (1980) *Acta Crystallogr.* **B36**, 621–627.
- Sipos, T. and Merkel, J. R. (1970) *Biochemistry* **9**, 2766–2775.

- Smith, K. F., Harrison, R. A. and Perkins, S. J. (1990) *Biochem. J.* **267**, 203–212
- Squire, P. G. and Himmel, M. E. (1979) *Arch. Biochem. Biophys.* **196**, 165–177
- Stuhrmann, H. B. (1973) *J. Mol. Biol.* **77**, 363–369
- Stuhrmann, H. B. and Fuess, H. (1976) *Acta Crystallogr.* **A32**, 67–74
- Tellam, R. and Winzor, D. J. (1977) *Biochem. J.* **161**, 687–694
- Timasheff, S. N. (1969) *Arch. Biochem. Biophys.* **132**, 165–169
- Thanki, N., Thornton, J. M. and Goodfellow, J. M. (1988) *J. Mol. Biol.* **202**, 637–657
- Tomana, M., Niemann, M., Garner, C. and Volanakis, J. E. (1985) *Mol. Immunol.* **22**, 107–111
- Towns-Andrews, E., Berry, A., Bordas, J., Mant, G. R., Murray, P. K., Roberts, K., Summer, I., Worgan, J. S., Lewis, R. and Gabriel, A. (1989) *Rev. Sci. Instrum.* **60**, 2346–2349
- Volanakis, J. E. and Macon, K. J. (1987) *Anal. Biochem.* **163**, 242–246
- Walsh, K. A. (1970) *Methods Enzymol.* **19**, 41–63
- Walter, J., Steigemann, W., Singh, T. P., Bartunik, H., Bode, W. and Huber, R. (1982) *Acta Crystallogr.* **B38**, 1462–1472
- Wang, D., Bode, W. and Huber, R. (1985) *J. Mol. Biol.* **185**, 595–624
- Weber, K., Pringle, J. R. and Osborn, M. (1972) *Methods Enzymol.* **26**, 3–27
- White, R. T., Damm, D., Hancock, N., Rosen, B. S., Lowell, B. B., Usher, P., Flier, J. S. and Spiegelman, B. M. (1992) *J. Biol. Chem.* **267**, 9210–9213
- Worcester, D. L. (1988) *J. Appl. Crystallogr.* **21**, 669–674
- Yon, J. (1959) *Biochim. Biophys. Acta* **31**, 75–85
- Zaccai, G. and Jacrot, B. (1983) *Ann. Rev. Biophys. Bioeng.* **12**, 139–157