The Structure and Enzymic Activities of the C1r and C1s Subcomponents of C1, the First Component of Human Serum Complement

By ROBERT B. SIM,* RODNEY R. PORTER, KENNETH B. M. REID and IRMA GIGLI†
MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford,
South Parks Road, Oxford OX1 3QU, U.K.

(Received 20 September 1976)

The subcomponents C1r and C1s and their activated forms C1r and C1s were each found to have mol.wts. in dissociating solvents of about 83 000. The amino acid compositions of each were similar, but there were significant differences in the monosaccharide analyses of subcomponents C1r and C1s, whether activated or not. Subcomponents C1r and C1s have only one polypeptide chain, but subcomponents C1r and C1s each contain two peptide chains of approx. mol.wts. 56 000 ('a' chain) and 27 000 ('b' chain). The amino acid analyses of the 'a' chains from each activated subcomponent are similar, as are those of the 'b' chains. The *N*-terminal amino acid sequence of 29 residues of the C1s 'a' chain was determined, but the C1r 'a' chain has a blocked *N*-terminal amino acid. The 20 *N*-terminal residues of both 'b' chains are similar, but not identical, and both show obvious homology with other serine proteinases. The difference in polysaccharide content of the subcomponents C1r and C1s is most marked in the 'b' chains. When tested on synthetic amino acid esters, subcomponent C1s hydrolysed both lysine and tyrosine ester bonds, but subcomponent C1r did not hydrolyse any amino acid esters tested nor any protein substrate except subcomponent C1s. The lysine esterase activity of subcomponent C1s provides a rapid and sensitive assay of the subcomponent.

The first component of complement, C1,‡ consists of three subcomponents C1q, C1r and C1s, which form a complex in the presence of Ca²⁺. Component C1 binds to aggregated antibody and is converted into an active proteinase, which splits the second and fourth components, C2 and C4, and which is inhibited by iPr₂P-F (see review by Reid & Porter, 1975). It was recognized early that subcomponent C1s is responsible for the activation of components C2 and C4 and that it exists in serum in a proenzymic form (Lepow *et al.*, 1963).

Subcomponent C1r is also a proteinase and is concerned in the activation of C1s (Naff & Ratnoff, 1968); these authors also reported that subcomponent C1r would hydrolyse some synthetic lysine and arginine esters. Subcomponent C1r has been isolated in proenzymic form (Ziccardi & Cooper, 1976; Sim

& Porter, 1976; Gigli *et al.*, 1976), as well as in the activated form C1r (de Bracco & Stroud, 1971; Takahashi *et al.*, 1975b; Gigli *et al.*, 1976). It was suggested that subcomponent C1r was also inhibited by iPr₂P-F (Sakai & Stroud, 1974), and this has been confirmed (Takahashi *et al.*, 1975b; Sim & Porter, 1976).

The isolation of subcomponents C1r and C1s in both proenzymic and activated forms in good yield (Gigli *et al.*, 1976) has made possible a detailed comparison of the chemical and enzymic properties of these two components. The results confirm some, but not all, of the previous reports, and show that the marked similarities in chemical structure between subcomponents C1r and C1s are in contrast with the different specificities of the active enzymes.

Materials and Methods

Materials

The sources of chemicals used were as follows: iPr₂P-F, Sigma (London) Chemical Co., Kingston-on-Thames, U.K.; *p*-nitrophenyl esters of *N*-benzyloxycarbonyl derivatives of L-leucine, L-phenylalanine, L-tyrosine and glycine, Sigma; α -*N*-Boc-Lys-ONp, Calbiochem Ltd., Hereford, U.K.; toluene-*p*-sulphonic acid, 3-(2-aminoethyl)indole

* Present address: D.R.F./Biochimie, C.E.A.-C.E.N.G., 85X, 38041 Grenoble Cedex, France.

† Present address: New York University Medical Center, Skin and Cancer Unit, 562 First Avenue, New York, NY 10016, U.S.A.

‡ Abbreviations: the nomenclature of complement components and subcomponents is that recommended by the World Health Organization (1968). Activated components are indicated by a bar, e.g. C1s. iPr₂P-F, diisopropyl phosphorofluoridate; Boc-, benzyloxycarbonyl-; -ONp, *p*-nitrophenyl ester.

hydrochloride and silylating reagents, Pierce and Warriner Ltd., Chester, U.K.; mannitol (organic analytical standard) and hexosamine hydrochlorides, BDH Chemicals Ltd., Poole, Dorset, U.K.; neutral sugars and sialic acid standards, Sigma; pre-packed columns for g.l.c., Phase-Sep Ltd., Queensferry, Clwyd, Wales, U.K.

Proteins used as molecular-weight standards were purchased from Sigma or from Merck, Darmstadt, West Germany. Trypsin, treated with 1-chloro-4-phenyl-3-L-tosylamidobutan-2-one (TPCK), was from Worthington Biochemical Corp., Freehold, NJ, U.S.A. α -Chymotrypsin from Sigma was treated with 7-amino-1-chloro-3-L-tosylamidoheptan-2-one (TLCK) (Shaw *et al.*, 1965). Plasmin was purchased from Kabi, Stockholm, Sweden. The unactivated and activated forms of the subcomponents C1r and C1s were isolated as described previously (Gigli *et al.*, 1976), except that in preparation of the activated subcomponents iPr₂P-F was added only to the serum (final concentration 5 mM) before fractionation and the C1 fraction from the Sepharose 6B column was incubated at 37°C and pH 7.4 in the presence of 5 mM-CaCl₂ for 1 h to ensure full activation. Subcomponent C1q was isolated as described by Reid (1974).

Haemolytic and enzymic activities of subcomponents C1r, C1r, C1s and C1s

Haemolytic assays were carried out as described previously (Gigli *et al.*, 1976).

Enzymic assay of subcomponent C1s was by hydrolysis of Boc-Tyr-ONp at 25°C, pH 8.1 (Bing, 1969) or by hydrolysis of Boc-Lys-ONp at 25°C, pH 6.0 (Silverstein, 1975). A unit of esterase activity is defined as the amount of enzyme required to hydrolyse 1 nmol of substrate/min at 25°C.

Subcomponent C1r was assayed by incubating subcomponent C1s (450 µg/ml) with 2–10% (w/w) of C1r in 57 mM-sodium phosphate/140 mM-NaCl/5 mM-EDTA, pH 7.4, for 10–15 min at 37°C. Subcomponent C1s formed was assayed as above.

The presence of activated enzymes in the pro-enzymic subcomponent preparations was estimated by their sensitivity to iPr₂P-F treatment in the haemolytic assay, activity in enzymic assay and the behaviour of the reduced protein on sodium dodecyl sulphate/polyacrylamide-gel electrophoresis (Gigli *et al.*, 1976).

Rates of hydrolysis of *N*-benzyloxycarbonyl-L-amino acid *p*-nitrophenyl esters by subcomponents C1r and C1s were measured at 25°C and compared with the rates of hydrolysis obtained with trypsin, chymotrypsin and plasmin. Assays were done in 15 mM-Tris/HCl/135 mM-NaCl, pH 8.1, except for the hydrolysis of α -*N*-Boc-Lys-ONp, when 100 mM-sodium phosphate/100 mM-NaCl, pH 6.0, was used. Final substrate concentration for all assays was

30 µM. The K_m and K_{cat} , defined as mol of substrate hydrolysed/s per mol of enzyme at infinite substrate concentration, for Boc-Lys-ONp and Boc-Tyr-ONp was determined from Lineweaver–Burk plots.

Molecular-weight estimation by sodium dodecyl sulphate/polyacrylamide-gel electrophoresis

Electrophoresis in 5.6% (w/v) polyacrylamide gels in buffers containing sodium dodecyl sulphate was carried out as described by Fairbanks *et al.* (1971). Portions (30–100 µg) of protein in 0.1 M-Tris/HCl/4 M-urea/1% (w/v) sodium dodecyl sulphate, pH 8.0, were reduced by incubation at 100°C for 4 min in the presence of 20 mM-dithiothreitol and then alkylated in the presence of 45 mM-iodoacetamide for 2 min at 100°C. Unreduced samples were incubated with 20 mM-iodoacetamide for 4 min at 100°C. After electrophoresis, gels were stained with Coomassie Blue. Apparent molecular weights were calculated by comparison of mobilities with those of reduced or non-reduced bovine serum albumin (mol.wt. 67 000), human transferrin (83 000), bovine liver catalase (60 000), rabbit immunoglobulin G (150 000), hen ovalbumin (42 000) and bovine chymotrypsinogen (25 000).

Separation of the polypeptide chains of subcomponents C1r and C1s, and estimation of their molecular weights

The polypeptide chains of subcomponents C1r and C1s were separated, after complete reduction, by gel filtration on a column (80 cm × 1.6 cm diam.) of Sephadex G-200 in 6 M-guanidine hydrochloride. Subcomponents C1r and C1s were reduced by incubation in 6 M-guanidinium chloride/50 mM-dithiothreitol for 1 h at 37°C, then alkylated for 20 min at 37°C, in the presence of 110 mM-iodoacetamide. Molecular weights of the 'a' and 'b' chains and of the reduced single polypeptide chains of subcomponents C1r and C1s were estimated by comparing their elution volumes with those of reduced marker proteins mentioned above.

N-Terminal determination and automatic sequencing

N-Terminal amino acids were identified by the dansyl procedure (Gray, 1972). Dansyl-amino acids were resolved by t.l.c. on polyamide sheets (Woods & Wang, 1967). Quantitative *N*-terminal analysis of the isolated 'a' chain of subcomponent C1r was performed by the thioacetylthioglycolic acid method (Mross & Doolittle, 1971) as described by Reid *et al.* (1972). Four cycles of thioacetylthioglycolic acid coupling and cleavage were done. *N*-Terminal amino acid sequences of polypeptide chains, carboxyamido-methylated in the presence of iodo[¹⁴C₂]acetamide, were determined by automated Edman degradation in a Beckman 890C sequencer by using a 0.1 M-Quadrol [0.1 M-NNN'-N'-tetrakis-(2-hydroxypropyl)-ethylenediamine trifluoroacetate, pH 9.5] programme

as described by Brauer *et al.* (1975). Thiazolines released were converted into phenylthiohydantoin derivatives, which were identified by g.l.c. and t.l.c. and measured by amino acid analysis after HI hydrolysis, as summarized by Reid (1976). Release of radioactivity into the butyl chloride phase collected by the sequencer was determined by liquid-scintillation counting. A 0.3 ml portion was added to 15 ml of 1,4-dioxan containing 0.5% (w/v) 2,5-diphenyloxazole and 2% (w/v) naphthalene.

Amino acid and carbohydrate analysis

Amino acid analysis was as described by Reid (1974), except that tryptophan was determined after hydrolysis of samples in 3M-toluene-*p*-sulphonic acid (Liu & Chang, 1971). Glucosamine and galactosamine were measured in the same hydrolysate (Allen & Neuberger, 1975). γ -Carboxyglutamic acid was measured after alkaline hydrolysis (Fernlund *et al.*, 1975).

Neutral sugars and sialic acid were determined by g.l.c. of *O*-trimethylsilyl ethers of the methyl glycosides (Clamp *et al.*, 1971). Protein samples were freed of contaminating sugars by precipitation with 10% (w/v) trichloroacetic acid at 4°C, or by gel filtration on Bio-Gel P150 columns in 0.2M-NH₄HCO₃, pH 7.8. Protein (10–25 nmol) containing 40 nmol of mannitol as internal standard was methanolysed at 90°C for 6 h in 1.5M-HCl in anhydrous methanol. Sugar derivatives were examined on a Hewlett-Packard 5830A gas chromatograph with a single column (200 cm $\times$ 0.32 cm) of 3% SE-30 Ultraphase on Chromosorb AW-DCMS (85–100 mesh). Quadruplicate analyses were made of single samples, before and after removal of contaminating sugars.

Antisera to subcomponents C1r and C1s

Antisera to subcomponents C1r and C1s were raised in rabbits by two injections of 100 μ g of the two proteins in 200 μ l of Freund's complete adjuvant into the popliteal node region at 21-day intervals. The rabbits were bled 7 and 15 days after the last injection.

Absorption coefficients of subcomponents C1r and C1s

The absorption coefficients at 280 nm of subcomponents C1r and C1s were measured after drying samples to constant weight over P₂O₅ and redissolving in 0.2M-Tris/HCl/8M-urea, pH 7.4. They were found to be $A_{1\text{cm}}^{1\%} = 11.7$ for C1r and $A_{1\text{cm}}^{1\%} = 9.4$ for C1s. The values for the unactivated subcomponents were assumed to be the same.

Determination of Stokes radii and sedimentation coefficients

Sedimentation coefficients relative to those of bovine liver catalase, bovine thyroglobulin and rabbit immunoglobulin G were calculated by sucrose-

density-gradient centrifugation (Martin & Ames, 1961). Protein samples were layered on linear gradients of 5–25% (w/v) sucrose in 5 mM-triethanolamine/HCl/145 mM-NaCl, pH 7.5, containing 5 mM-EDTA or 5 mM-CaCl₂. Gradients were loaded in a Beckman SW 40 Ti head and centrifuged for 12 h, 4°C at 195 000g ($r_{\text{av.}} = 10.93$ cm) in a Beckman L5 65 ultracentrifuge.

Stokes radii were calculated from relative diffusion rates on agarose plates in 5 mM-triethanolamine/HCl/145 mM-NaCl, containing 5 mM-CaCl₂ or 5 mM-EDTA (Allison & Humphrey, 1959), or were determined by gel filtration (Ackers, 1964) on Sephadex G-200 columns (80 cm $\times$ 1.6 cm diam.) in the same buffers.

The partial specific volumes of subcomponents C1r and C1s, calculated from amino acid and carbohydrate compositions (Schachman, 1957) are 0.714 cm³/g and 0.717 cm³/g respectively. Molecular weights and frictional ratios were calculated as described by Siegel & Monty (1966).

Results

Stability of subcomponents

The isolated subcomponents appeared to be homogeneous, as reported previously (Gigli *et al.*, 1976), and were fully active as shown by haemolytic assay. The haemolytic activity of subcomponents C1r and C1r was the same within the accuracy of the method, but that of subcomponent C1s was consistently somewhat higher than that of subcomponent C1s. The subcomponents were stable over weeks at 2°C in solutions of pH 5.3–7.3 in the absence of Ca²⁺.

Molecular weights of subcomponents

The molecular weights of subcomponents C1r, C1s and their constituent peptide chains, as measured in dissociating conditions by comparison with standard proteins, are given in Table 1. Results obtained by using non-reduced proteins compared with non-reduced standards are shown for comparison with other published estimates. The agreement, for reduced proteins in polyacrylamide-gel electrophoresis in sodium dodecyl sulphate buffers and on Sephadex G-200 columns in 6 M-guanidinium chloride, is good for subcomponents C1r and C1s and for the larger 'a' chains of C1r and C1s. The value for the 'b' chains was higher from the electrophoresis measurements and probably reflected their carbohydrate content. Previous experience (Reid & Porter, 1976) suggested that values found by gel filtration in guanidinium chloride were most likely to be correct.

In neutral buffers containing 5 mM-CaCl₂ or 5 mM-EDTA, subcomponents C1r and C1r behaved as dimers, of approx. mol.wt. 166 000 (Table 2). Under the same conditions, subcomponents C1s and C1s

behaved as monomers of about 81 000 mol.wt., in the presence or absence of Ca^{2+} .

Amino acid and carbohydrate analysis

The amino acid and carbohydrate analyses of the activated subcomponents and their polypeptide chains are summarized in Table 3. No significant differences were found between analyses of activated and unactivated subcomponents (Sim & Porter, 1976). The molecular weights of the peptide chains of subcomponents C1r and C1s are too high to make any precise comparison of total amino acids with those of the whole molecules, but it is apparent from the data and from the molecular-weight estimations that no large peptides are lost during activation. Though the amino acid content of the two subcomponents is very similar, there are small but significant differences in the quantities of some amino acids such as valine, leucine and arginine. The higher carbohydrate content of subcomponent C1r than of C1s appears to be largely confined to the 'b' chain and probably explains the lower mobility of the 'b' chain

of subcomponent C1r relative to that of C1s on polyacrylamide-gel electrophoresis in buffers containing sodium dodecyl sulphate.

Less than 1 mol of γ -carboxyglutamic acid/mol of protein was present in alkaline hydrolysates of subcomponents C1r, C1s and C1q; this amino acid has been associated with Ca^{2+} binding in various proteins (Magnusson *et al.*, 1974; Hauschka *et al.*, 1975). The results indicate, however, that the binding of Ca^{2+} ions in the C1 complex does not involve γ -carboxyglutamic acid residues.

N-Terminal amino acid sequences

N-Terminal amino acid identification by the dansyl technique gave isoleucine as N-terminal for the 'b' chains of both subcomponents C1r and C1s. Glutamic acid was N-terminal on both the 'a' chain of subcomponent C1s and on the single polypeptide chain of subcomponent C1s. No N-terminal acid could be found for the 'a' chain of C1r, nor for the single polypeptide chain of subcomponent C1r. This negative result was confirmed for the subcomponent-C1r 'a' chain by the thioacetylthioglycolic acid method and also by automated Edman degradation of whole subcomponent C1r, in which only residues from the 'b' chain were cleaved. This result was in contrast with the N-terminal isoleucine reported for subcomponent-C1r 'a' chain by Takahashi *et al.* (1975b).

The N-terminal sequences found for the three chains with free amino groups are given in Tables 4, 5 and 6 and the homology between the sequences of the 'b' chains of both subcomponents was apparent.

Enzymic activity of subcomponents C1r and C1s

None of the five synthetic amino acid esters tested was hydrolysed by subcomponent C1r (Table 7). Subcomponent C1r was also inactive when tested with component C4 and with C2, nor would it hydrolyse the B chain of insulin (E. M. Press, unpublished work). The only enzymic activity of subcomponent C1r detected so far was the hydrolysis of subcomponent C1s, suggesting that subcomponent C1r was a remarkably specific proteinase. Subcomponent C1s hydrolysed the lysine ester and, at a slower rate, the tyrosine ester. It had little or no effect

Table 1. *Estimates of molecular weight in dissociating conditions*

Apparent molecular weights were calculated from mobilities on polyacrylamide-gel electrophoresis in buffers containing sodium dodecyl sulphate, and by gel filtration on Sephadex G-200 in 6M-guanidinium chloride. Full details are given in the text. Subcomponents C1r and C1s give two chains on reduction.

Sub-component	Apparent molecular weight		
	By electrophoresis		By gel filtration
	Unreduced proteins, compared with unreduced standards	Reduced proteins, compared with reduced standards	Reduced proteins, compared with reduced standards
C1r	98 000	83 000	83 000
C1r	107 000	58 000 + 36 000	57 000 + 26 000
C1s	92 000	83 000	83 000
C1s	92 000	58 000 + 29 000	57 000 + 26 000

Table 2. *Estimates of molecular weights in non-dissociating conditions*

Experimental details are given in the text. Results are given ± 1 s.d., for the numbers of determinations shown in parentheses. No differences were observed between proenzymic and activated subcomponents or between measurements in the presence of Ca^{2+} or EDTA.

Subcomponent	Stokes radius (nm)	Sedimentation coefficient (S)	Mol.wt.	Frictional ratio
C1r or C1r	5.0 ± 0.2 (8)	7.9 ± 0.1 (6)	$166\,000 \pm 4000$	1.41 ± 0.05
C1s or C1s	4.1 ± 0.3 (10)	4.7 ± 0.2 (4)	$81\,000 \pm 3500$	1.46 ± 0.07

Table 3. Amino acid (a) and carbohydrate (b) compositions of subcomponents C1r and C1s and their constituent polypeptide chains

Amino acid compositions of subcomponents C1r and C1s were calculated from duplicate 24h, 48h and 72h HCl hydrolysates. Corrections were made for destruction of serine and threonine. Amino acid compositions of 'a' and 'b' chains were calculated from duplicate 48h HCl hydrolysates, and serine and threonine yields were corrected assuming the same rate of destruction as observed for subcomponents C1r and C1s. Other details are given in the text.

	Content (residues/100 amino acid residues)					
	C1r 'a' chain	C1r 'b' chain	Subcomponent C1r	C1s 'a' chain	C1s 'b' chain	Subcomponent C1s
(a) Amino acid						
Asp	10.3	9.6	10.2	11.0	10.7	10.9
Thr	4.9	5.3	4.9	4.9	5.7	5.1
Ser	6.0	7.1	6.3	6.6	5.4	6.3
Glu	12.3	11.7	12.1	11.6	10.1	10.9
Pro	6.5	4.5	5.9	7.0	6.9	7.1
Gly	9.8	12.0	10.4	10.0	10.7	10.2
Ala	5.0	7.2	5.3	5.5	7.2	5.9
$\frac{1}{2}$ Cys	3.0	1.5	2.5	3.1	1.2	2.5
Val	4.9	6.6	5.3	6.7	8.5	7.2
Met	2.3	1.4	2.0	1.9	1.3	1.8
Ile	4.2	3.6	4.0	4.2	3.5	3.9
Leu	7.8	8.0	7.9	6.0	7.0	6.4
Tyr	3.9	4.2	4.2	4.0	4.1	4.1
Phe	4.9	4.0	4.7	5.0	4.1	4.7
Trp	1.1	2.0	1.6	1.0	1.5	1.4
Lys	4.7	4.6	4.7	4.4	5.4	4.6
His	3.3	2.7	3.1	2.9	2.5	2.7
Arg	5.1	4.1	4.9	4.2	4.5	4.0
(b) Monosaccharide						
Fucose	0.2	0.0	0.0	0.0	0.0	0.0
Mannose	1.0	1.2	1.1	0.8	0.9	0.9
Galactose	1.5	1.4	1.4	0.8	0.8	0.8
N-Acetylglucosamine	1.2	4.0	2.4	1.1	1.2	1.2
N-Acetylgalactosamine	0.3	0.4	0.3	0.1	0.2	0.1
Sialic acid	0.6	0.8	0.8	1.0	1.1	1.0
Carbohydrate (% of total weight)*	7.4	12.4	9.4	6.5	7.5	7.1

* Hexosamines were calculated as N-acetylhexosamines, and sialic acid as N-acetylneuraminic acid.

on the glycine, phenylalanine and leucine esters. This was clearly distinct from the specificities of trypsin and chymotrypsin. The specificity of subcomponent C1s resembled most closely, on these esters, that of plasmin, but it was distinct in that, in contrast with plasmin (Groskopf *et al.*, 1968), it did not hydrolyse the B chain of insulin (E. M. Press, unpublished work).

Subcomponent C1s had a specific activity of 160–190 esterase units/mg when assayed with Boc-Tyr-ONp, at a final substrate concentration of 30 μ M at pH 8.1. The specific activity was similar to that of the subcomponent-C1s preparations described by Bing (1969) and Assimeh *et al.* (1974). The K_m of subcomponent C1s for this substrate was 7×10^{-5} M, and the K_{cat} , taking a mol.wt. of 83000 with a single catalytic centre, was 0.73 s^{-1} .

Hydrolysis of α -N-Boc-Lys-ONp was found to

provide a more sensitive assay for subcomponent C1s than did hydrolysis of the tyrosine ester. Subcomponent C1s had a specific α -N-Boc-Lys-ONp esterase activity of 2700–3100 esterase units/mg at a substrate concentration of 0.14 mM and pH 6.0. The K_m for this substrate was 8×10^{-4} M and $K_{cat} = 20.8 \text{ s}^{-1}$.

Subcomponent C1r activated subcomponent C1s at an initial rate corresponding to the development of 24–39 units of subcomponent-C1s Boc-Tyr-ONp esterase activity formed/min per mg of C1r at pH 7.4, 37°C.

Discussion

As reported previously (Gigli *et al.*, 1976), it was possible to isolate the subcomponents C1r, C1s and their activated forms C1r and C1s in good yield, and

Table 4. Automated sequence analysis of subcomponent C1s 'a' chain

Subcomponent C1s 'a' chain (150nmol) was subjected to 40 cycles of Edman degradation. Full details are given in the text. Yields of amino acids ranged from 16 to 19nmol at steps 1-3, down to 4-6nmol at steps 21-23. Unambiguous identification of amino acids released was possible up to step 18. Residues 20-23, 25, 27 and 29 were also clearly identified. No radioactivity was released in the first 25 cycles. Abbreviations: ND, not done; Abu, 2-aminobutyric acid. —, nothing observed.

Residue no.	Method of identification of amino acid phenylthiohydantoin derivative		Amino acid analysis	Proposed sequence
	G.l.c.	T.l.c.		
1	—	Glu	Glu	Glu
2	Thr/Pro	Pro	Pro	Pro
3	Thr/Pro	Thr/Ser	Abu*	Thr
4	Met	Met	—*	Met
5	—	Tyr	Tyr	Tyr
6	—	Gly	Gly	Gly
7	ND	Glu	Glu	Glu
8	Ile/Leu	Ile/Leu	Ile	Ile
9	Ile/Leu	Ile/Leu	Leu	Leu
10	—	Thr/Ser	Ala*	Ser
11	Thr/Pro	Pro	Pro	Pro
12	ND	Asn	Asp*	Asn
13	ND	Tyr	Tyr	Tyr
14	Thr/Pro	Pro	Pro	Pro
15	ND	Gln	Glu*	Gln
16	Ala	Ala	Ala	Ala
17	ND	Tyr	Tyr	Tyr
18	Thr/Pro	Pro	Pro	Pro
19	—	—	—	—
20	ND	Glu	Glu	Glu
21	Val	Val	Val	Val
22	ND	Glu	Glu	Glu
23	ND	Lys	Lys	Lys
24	—	—	—	—
25	ND	Trp	Gly + Ala*	Trp
26	—	—	—	—
27	ND	Ile/Leu	Ile	Ile
28	—	—	—	—
29	Val	Val	ND	Val

* After hydrolysis the phenylthiohydantoin derivatives of Asn and Gln are detected as Asp and Glu respectively, that of Thr as Abu, that of Ser as Ala, and that of Trp as Gly + Ala; that of Met is destroyed.

they gave stable products. This was possible for the proenzyme form only by repeated addition of iPr_2P-F to inactivate contaminating proenzymes, which appear to be slowly converted into active proteinases during handling. Repeated addition of iPr_2P-F during the isolation does not decrease the haemo-

Table 5. Automated sequence analysis of subcomponent C1s 'b' chain

Subcomponent C1s 'b' chain (153nmol) was subjected to 20 cycles of Edman degradation. Full details are given in the text. Yields of amino acids at each cycle ranged from 10 to 12nmol at cycles 1-3 down to 5-7nmol at cycles 14-16. Unambiguous identification of amino acid released was possible at all steps. No radioactivity was released in the 20 cycles. ND, not done.

Residue no.	Method of identification of amino acid phenylthiohydantoin derivative		Amino acid analysis	Proposed sequence
	G.l.c.	T.l.c.		
1	Ile/Leu	Ile/Leu	Ile	Ile
2	Ile/Leu	Ile/Leu	Ile	Ile
3	ND	Gly	Gly	Gly
4	Gly	Gly	Gly	Gly
5	—	Thr/Ser	Ala*	Ser
6	—	Asp	Asp	Asp
7	Ala	Ala	Ala	Ala
8	ND	Asp	Asp	Asp
9	Ile/Leu	Ile/Leu	Ile	Ile
10	ND	Lys	Lys	Lys
11	ND	Asn	Asp*	Asn
12	Phe	Phe	Phe	Phe
13	Thr/Pro	Pro	Pro	Pro
14	—	Trp	Gly + Ala*	Trp
15	ND	Gln	Ala*	Gln
16	Val	Val	Val	Val
17	Phe	Phe	Phe	Phe
18	Phe	Phe	Phe	Phe
19	ND	Asp	Asp	Asp
20	ND	Asn	ND	Asn

* See footnote to Table 4.

lytic activity of the proenzymic forms of the two subcomponents.

The molecular weights of subcomponents C1r and C1s have been estimated by two methods, namely polyacrylamide-gel electrophoresis in the presence of sodium dodecyl sulphate, and gel filtration in 6M-guanidinium chloride. The results agree for the whole molecules, 83000mol.wt., and for the heavier 'a' chains of the activated forms, 56000mol.wt. For the 'b' chains, gel filtration, the more reliable method, gave a mol.wt. of 27000 for both. Since the sum of molecular weights of the 'a' and 'b' chains equals, within the limitations of the method, the molecular weight of the single proenzymic chain, it appears that only limited hydrolysis occurs during activation, and no large peptides are lost. Lepow *et al.* (1958) have reported no release of trichloroacetic acid-soluble N during activation of a partially purified C1 preparation. The same close similarity between the two

components was shown by their amino acid analysis, and the major significant difference found is in the carbohydrate content of the 'b' chains, which is higher in subcomponent C1r.

Table 6. Automated sequence analysis of subcomponent C1r 'b' chain

Subcomponent-C1r 'b' chain (90nmol) was subjected to 38 cycles of Edman degradation. Full details are given in the text. Yields of amino acids at each step ranged from 10 to 12nmol at cycles 1-3, down to 4-5nmol at cycles 17-19. The first 20 steps are shown. Release of radioactivity was not monitored. ND, not done. Abu, 2-aminobutyric acid.

Residue no.	Method of identification of amino acid phenylthiohydantoin derivative		Amino acid analysis	Proposed sequence
	G.l.c.	T.l.c.		
1	Ile/Leu	Ile/Leu	Ile	Ile
2	Ile/Leu	Ile/Leu	Ile	Ile
3	ND	Gly	Gly	Gly
4	ND	Gly	Gly	Gly
5	ND	Gln	Glu*	Gln
6	ND	Lys	Lys	Lys
7	Ala	Ala	Ala	Ala
8	ND	Lys	Lys	Lys
9	Met	Met	—*	Met
10	ND	Gly	Gly	Gly
11	ND	Asn	Asp*	Asn
12	Phe	Phe	Phe	Phe
13	Thr/Pro	Pro	Pro	Pro
14	ND	Trp	Gly + Ala*	Trp
15	ND	Gln	Glu*	Gln
16	ND	Val	Val	Val
17	ND	Phe	Phe	Phe
18	ND	Thr/Ser	Abu*	Thr
19	ND	Asn	Asp*	Asn
20	ND	—	Glu	Glx

* See footnote to Table 4.

The results in Table 2 show subcomponent C1r, like C1r, to be a dimer in non-dissociating conditions, as found by Ziccardi & Cooper (1976). Subcomponent C1s behaves as a monomer, in agreement with Haines & Lepow (1964) and Loos *et al.* (1976), though it was found to be a dimer in the absence of Ca²⁺ by Okamura *et al.* (1973) and in the presence of Ca²⁺ by Valet & Cooper (1974).

Both components C1r and C1s are inhibited to similar extents by iPr₂P-F (Gigli *et al.*, 1976), and inactivation is associated with the binding of 1mol of di-isopropyl phosphate/mol of enzyme (Sim & Porter, 1976). Both subcomponents appear to be typical serine proteinases, the catalytic site being in the smaller 'b' chain. For subcomponent C1s, it has been shown that there is homology of sequence around the reactive serine residue and the equivalent sequences of trypsin, chymotrypsin, plasmin and thrombin (Barkas *et al.*, 1973). It is likely that a similar sequence is present around the reactive serine residue of subcomponent C1r. Close homology of the N-terminal tetrapeptide sequence of subcomponent-C1s 'b' chain with the N-terminal sequences of the di-isopropyl phosphate-binding chains of the same enzymes has also been reported (Takahashi *et al.*, 1975a). The extended N-terminal sequences reported here for 'b' chains of both subcomponents C1r and C1s have been compared with the corresponding regions of trypsin, chymotrypsin, plasmin and thrombin (Fig. 1). All follow the same characteristic pattern of valine or isoleucine at positions 1 and 2, usually glycine at positions 3 and 4 and alanine at position 7, followed by random variation, and then proline, tryptophan or tyrosine, glutamine and valine at positions 13, 14, 15 and 16 respectively. The most marked differences between 'b' chains of subcomponents C1r and C1s is in the lysine/aspartic acid changes at positions 6 and 8 and glycine/lysine change at position 10. If, as is highly probable, these residues are in the same position in the steric structure as the equivalent residues in chymotrypsin, they would not contribute to the

Table 7. Hydrolysis of N-benzoyloxycarbonyl L-amino acid p-nitrophenyl esters by subcomponents C1r, C1s, trypsin, chymotrypsin and plasmin

Assays were done at 25°C, at a final substrate concentration of 3μM. The pH for all assays was 8.1, except for the lysine ester, which was assayed at pH 6.0. The absolute purity of the plasmin was not established, but specific activities for plasmin are calculated assuming 100% purity. Other details are given in the text.

Enzyme	Substrate hydrolysed (mol/min per mol of enzyme)				
	Boc-Tyr-ONp	Boc-Gly-ONp	Boc-Phe-ONp	Boc-Leu-ONp	α-N-Boc-Lys-ONp
C1r	0.0	0.0	0.0	0.0	0.0
C1s	14.0	0.0	1.74	0.0	74.7
Trypsin	409	3.50	116	45.5	298
Chymotrypsin	1790	16.7	407	72.2	6.58
Plasmin	31.5	1.83	10.1	0.52	110

Fig. 1. Comparison of the *N*-terminal sequences of 'b' chains of subcomponents C1r and C1s with the corresponding polypeptide chains of other serine proteinases

Sequence data and residue numbering for trypsin, chymotrypsin, plasmin and thrombin are from de Haën *et al.* (1975). Residues shown in **bold** typeface indicate an identity with both subcomponents C1r and C1s; residues shown in *italic* typeface denote identity with subcomponents C1r or C1s.

Residue no.	(5)	(10)	(15)	(20)
Human C1r 'b' chain	Ile -Ile -Gly-Gly-Gln-Lys-Ala-Lys-Met-Gly-Asn-Phe-Pro-Trp-Gln-Val-Phe-Thr-Asn-Glx			
Human C1s 'b' chain	Ile -Ile -Gly-Gly-Ser -Asp-Ala-Asp-Ile -Lys-Asn-Phe-Pro-Trp-Gln-Val-Phe-Phe-Asp-Asn			
Bovine chymotrypsin A	Ile -Val-Asn-Gly-Glu-Glu-Ala-Val-Pro-Gly-Ser-Trp-Pro-Trp-Gln-Val-Ser -Leu-Gln-Asp			
Bovine trypsin	Ile -Val-Gly-Gly-Tyr-Thr-Cys-Gly-Ala-Asn-Thr-Val-Pro-Tyr-Gln-Val-Ser -Leu-Asn-Ser			
Bovine thrombin B chain	Ile -Val-Glu-Gly-Gln-Asp-Ala-Glu-Val-Gly-Leu-Ser-Pro-Trp-Gln-Val-Met-Leu-Phe-Arg			
Human plasmin B chain	Val-Val-Gly-Gly-Cys-Val-Ala-His-Pro-His-Ser-Trp-Pro-Trp-Gln-Val-Val -Leu-Leu-Arg			

specificity of the active enzymes, as they would be exposed to the solvent on an external loop distant from the active site (Birktoft *et al.*, 1970). In solution, subcomponent C1r will hydrolyse one or more bonds in subcomponent C1s but not in C1r (R. B. Sim, R. R. Porter & A. W. Dodds, unpublished work). These differences in sequence are near to the bond hydrolysed in subcomponent C1s, and it is possible that they contribute to the difference in susceptibility to hydrolysis of subcomponents C1s and C1r by subcomponent C1r.

There is a 60% identity of sequence in the *N*-terminal 20-residue section of the 'b' chains of subcomponents C1r and C1s, emphasizing the similarity of the two subcomponents which is apparent from their size and amino acid composition, and reinforcing the suggestion that they are likely to have arisen by gene duplication. As the *N*-terminal sequence of the 'a' chain of subcomponent C1r was not obtained, no comparison with the 'a' chain of subcomponent C1s is possible. The latter shows no homology with any region of the published sequences of human plasminogen (Claeys *et al.*, 1976; Wiman, 1973) or bovine prothrombin (Claeys *et al.*, 1976). Takahashi *et al.* (1975a) reported the *N*-terminal dipeptide of the subcomponent-C1s 'a' chain to be Glx-Ile, contrary to our finding of Glu-Pro.

On the present results, subcomponent C1r appears to show exceptional specificity. In contrast with the results of Assimeh *et al.* (1974), subcomponent C1r does not hydrolyse Boc-Tyr-ONp. Naff & Ratnoff (1968) reported that partially purified subcomponent C1r would hydrolyse a limited number of arginine and lysine esters. The results of Arlaud *et al.* (1976) suggest that hydrolysis of a similar arginine ester, benzoylarginine ethyl ester, by subcomponent C1r is negligible. Lowe (1973) showed that the haemolytic and esterase activities associated with partially purified rabbit subcomponent C1r could be partially separated during further purification procedures, suggesting that the use of basic amino acid esters to measure subcomponent-C1r activity may produce

misleading results. The difference in these findings on subcomponent-C1r specificity may arise from the presence of contaminating proteinases in the partially purified preparations discussed in other reports. Even with the highly purified C1r-subcomponent preparations described here, contamination with small amounts of other proteases occasionally occurred. Four of 11 subcomponent-C1r preparations tested contained very low Boc-Lys-ONp esterase activity, but this activity did not correlate with the ability of the C1r preparations to activate subcomponent C1s nor with its haemolytic activity. The esterase activity is therefore assumed to be that of a contaminant.

The results reported here confirm that subcomponents C1r and C1s appear to be closely related proteinases, and that both show sequence homology with other serine proteinases of the trypsin family. The gross structures of subcomponents C1r and C1s before and after activation have been compared, and assays for the activated forms have been developed which are sufficiently rapid and sensitive to permit kinetic studies of the activation process itself.

We are grateful to Mr. T. Gascoyne for amino acid analyses, and to Ms. J. Addis for excellent technical assistance. We thank Miss E. M. Press and Dr. M. A. Kerr for useful discussions. R. B. S. was in receipt of a Medical Research Council Research Studentship.

References

- Ackers, G. K. (1964) *Biochemistry* **3**, 723-730
- Allen, A. K. & Neuberger, A. (1975) *FEBS Lett.* **60**, 76-80
- Allison, A. C. & Humphrey, J. H. (1959) *Nature (London)* **183**, 1590-1592
- Arlaud, G. J., Meyer, C. M. & Colomb, M. G. (1976) *FEBS Lett.* **66**, 132-136
- Assimeh, S. N., Bing, D. H. & Painter, R. H. (1974) *J. Immunol.* **113**, 225-234
- Barkas, T., Scott, G. K. & Fothergill, J. E. (1973) *Biochem. Soc. Trans.* **1**, 1219-1220

- Bing, D. H. (1969) *Biochemistry* **8**, 4503–4510
- Birktoft, J. J., Blow, D. M., Henderson, R. & Steitz, T. A. (1970) *Philos. Trans. R. Soc. London Ser. B* **257**, 67–76
- Brauer, A. W., Margolies, M. N. & Haber, E. (1975) *Biochemistry* **14**, 3029–3035
- Claeys, H., Sottrup-Jensen, L., Zajdel, M., Petersen, T. E. & Magnusson, S. (1976) *FEBS Lett.* **61**, 20–24
- Clamp, J. R., Bhatti, T. & Chambers, R. E. (1971) *Methods Biochem. Anal.* **19**, 229–344
- de Bracco, M. M. E. & Stroud, R. M. (1971) *J. Clin. Invest.* **50**, 838–848
- de Haën, C., Neurath, H. & Teller, D. C. (1975) *J. Mol. Biol.* **92**, 225–259
- Fairbanks, G., Steck, T. L. & Wallach, D. F. H. (1971) *Biochemistry* **10**, 2606–2617
- Fernlund, P., Stenflo, J., Roepstorff, P. & Thomsen, J. (1975) *J. Biol. Chem.* **250**, 6125–6133
- Gigli, I., Porter, R. R. & Sim, R. B. (1976) *Biochem. J.* **157**, 541–548
- Gray, W. R. (1972) *Methods Enzymol.* **25**, 121–138
- Groskopf, W. R., Hsieh, B., Summari, L. & Robbins, K. C. (1968) *Biochim. Biophys. Acta* **168**, 376–379
- Haines, A. L. & Lepow, I. H. (1964) *J. Immunol.* **92**, 456–467
- Hauschka, P. V., Lian, J. B. & Gallop, P. M. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 3925–3929
- Lepow, I. H., Ratnoff, O. D. & Levy, R. (1958) *J. Exp. Med.* **107**, 451–474
- Lepow, I. H., Naff, G. B., Todd, E. W., Pensky, T. & Hinz, C. F., Jr. (1963) *J. Exp. Med.* **117**, 983–1008
- Liu, T.-Y. & Chang, Y. H. (1971) *J. Biol. Chem.* **246**, 2842–2848
- Loos, M., Hill, H.-U., Wellek, B. & Heinz, H.-P. (1976) *FEBS Lett.* **64**, 341–345
- Lowe, D. M. (1973) D.Phil. Thesis, University of Oxford
- Magnusson, S., Sottrup-Jensen, L., Petersen, T. E., Morris, H. R. & Dell, A. (1974) *FEBS Lett.* **44**, 189–193
- Martin, R. G. & Ames, B. N. (1961) *J. Biol. Chem.* **236**, 1372–1379
- Mross, G. A. & Doolittle, R. F. (1971) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **30**, 1241
- Naff, G. B. & Ratnoff, O. D. (1968) *J. Exp. Med.* **128**, 571–593
- Okamura, K., Muramatsu, M. & Fujii, S. (1973) *Biochim. Biophys. Acta* **295**, 252–257
- Reid, K. B. M. (1974) *Biochem. J.* **141**, 189–203
- Reid, K. B. M. (1976) *Biochem. J.* **155**, 5–17
- Reid, K. B. M. & Porter, R. R. (1975) *Contemp. Top. Mol. Immunol.* **4**, 1–22
- Reid, K. B. M. & Porter, R. R. (1976) *Biochem. J.* **155**, 19–23
- Reid, K. B. M., Lowe, D. M. & Porter, R. R. (1972) *Biochem. J.* **130**, 749–763
- Sakai, K. & Stroud, R. M. (1974) *Immunochemistry* **11**, 192–196
- Schachman, H. K. (1957) *Methods Enzymol.* **4**, 32–104
- Shaw, E., Mares-Guia, M. & Cohen, W. (1965) *Biochemistry* **4**, 2219–2224
- Siegel, L. M. & Monty, K. J. (1966) *Biochim. Biophys. Acta* **112**, 346–362
- Silverstein, R. M. (1975) *Anal. Biochem.* **65**, 500–506
- Sim, R. B. & Porter, R. R. (1976) *Biochem. Soc. Trans.* **4**, 127–129
- Takahashi, K., Nagasawa, S. & Koyama, J. (1975a) *FEBS Lett.* **50**, 330–333
- Takahashi, K., Nagasawa, S. & Koyama, J. (1975b) *FEBS Lett.* **55**, 156–160
- Valet, G. & Cooper, N. R. (1974) *J. Immunol.* **112**, 339–350
- Wiman, B. (1973) *Eur. J. Biochem.* **39**, 1–9
- Woods, K. R. & Wang, K.-T. (1967) *Biochim. Biophys. Acta* **133**, 369–370
- World Health Organization (1968) *Bull. W.H.O.* **39**, 935–938
- Ziccardi, R. J. & Cooper, N. R. (1976) *J. Immunol.* **116**, 496–503