

Partial sequence of human complement component Factor B: Novel type of serine protease

(alternative pathway/complement component 3 convertase)

D. L. CHRISTIE, J. GAGNON, AND R. R. PORTER

Medical Research Council Immunochemistry Unit, Biochemistry Department, Oxford University, South Parks Road, Oxford OX1 3QU, England

Contributed by Rodney R. Porter, May 5, 1980

ABSTRACT Factor B (a component of the alternative pathway of complement) is believed to contain the proteolytic site of the complex enzymes C3 convertase (C3bB) and C5 convertase (C3b_nB). Conflicting results have been obtained in regard to the inactivation of these enzymes by diisopropyl phosphorofluoridate but it has been suggested that activated Factor B (Factor B) is a serine protease with the active site in Bb, a COOH-terminal fragment of approximately 60,000 molecular weight. Partial amino acid sequence studies of Bb derived from human Factor B have shown that the NH₂-terminal 40 residues have no homology with NH₂-terminal sequences of other serine proteases. However, positioning of a further 170 residues out of approximately 290 residues in two continuous CNBr fragments from the COOH terminus has shown that there is a strong homology of sequence in this section. The active site residues histidine, aspartic acid, and serine all are present in positions corresponding with those of typical serine proteases. It is suggested that Factor B is a novel type of serine protease with a catalytic chain of molecular weight twice that of proteases previously studied and probably with a different activation mechanism.

In the alternative pathway of activation of complement, activated Factor B (Factor B) in association with C3b forms the complex proteases C3 convertase C3bB and C5 convertase (C3b_nB) (1, 2). On activation by Factor D, Factor B, a glycoprotein of about 90,000 molecular weight, is split into an NH₂-terminal fragment Ba and a COOH-terminal fragment Bb (3-5). Ba is about 30,000 molecular weight and Bb is about 60,000 molecular weight. Medicus *et al.* (6) reported that C3 and C5 convertases of both pathways of activation are inhibited by diisopropyl phosphorofluoridate (iPr₂P-F) and that, if a radioactive reagent is used, it is found in the Bb fragment, suggesting that Factor B is a serine protease with the active center in Bb. An earlier report, however, showed the alternative pathway C3 convertase to be resistant to inhibition by iPr₂P-F (7). Because the suggested catalytic peptide Bb is also about twice the length of the catalytic peptides of all other serine proteases studied so far, it appeared that Factor B could not be a typical serine protease. The amino acid sequences of many of these proteases are known and they show strong homology with each other with many highly conserved features, notably in the NH₂ terminus of the catalytic peptide and in the active site residues, suggesting that all have similar catalytic and activation mechanisms (8).

Sequence studies have therefore been made of Bb and have shown that, although the essential features of the active site residues of serine proteases are present and are in similar positions relative to the COOH terminus, the NH₂-terminal section, which is ≈300 residues longer, does not contain the characteristic NH₂-terminal sequence of other serine proteases.

METHODS

Purification of Factor B and Fragments Ba and Bb. Minor modifications were made to the procedure of Kerr (4) for the isolation of Factor B from outdated human plasma. The column of CM-Sephadex C-50 was equilibrated with 0.05 M (in place of 0.1 M) sodium phosphate buffer at pH 6.0. Chromatography on both "aged" CNBr-activated Sepharose 4B and DEAE-Sepharose utilized a linear gradient of NaCl in 5 mM sodium barbital, pH 8.5/0.5 mM CaCl₂/2.0 mM MgCl₂/40 mM NaCl (VBS⁺⁺).

To prepare fragments Ba and Bb, Factor B (100 mg) was incubated with C3 (10 mg) and Factor D (1 mg) in 50 ml of VBS⁺⁺ for 4 hr at 37°C. The fragments were isolated by chromatography of the digest on a column (2.5 × 20 cm) of DEAE-Sepharose CL-6B in VBS⁺⁺ with a 500-ml linear gradient of NaCl from 40 to 200 mM. Factor D was purified to the stage following gel filtration on Sephadex G-75 as described by Johnson *et al.* (9). Component C3 was kindly provided by R. B. and E. Sim and had been purified as described by Tack and Prah (10). Prior to sequence studies, Bb was reduced and S-[¹⁴C]carboxymethylated as described (9).

Chemical Cleavages. CNBr fragments of Bb were prepared by the addition of a 10-fold weight excess of CNBr to the protein (10 mg/ml) in 70% HCOOH, and the mixture was kept in the dark for 20 hr at 4°C. The mixture was then diluted 1:10 with water and freeze-dried. The digest was chromatographed on a column (3 × 90 cm) of Sephadex G-100 in 10% acetic acid, and 3.6-ml fractions were collected. Peptides were identified by their absorbance at 280 nm and by measurement of radioactivity. Pool CB-VI was purified on a column (1 × 30 cm) of DEAE-Sephadex with a 250-ml linear gradient of NH₄HCO₃ from 0.05 to 0.5 M.

The cleavage of Asp-Pro bonds of Bb was performed by dissolving Bb (10 mg/ml) in 10% acetic acid containing 6 M guanidine hydrochloride adjusted to pH 2.5 with pyridine (11). The vial was flushed with nitrogen and sealed, and the mixture was incubated at 40°C for 96 hr with constant stirring. Products were separated by gel filtration on Sephadex G-100 as described above.

Enzymatic Cleavages. Limited cleavage of Bb was carried out by using clostripain. The enzyme (Institut Pasteur, 500 units/mg) was activated by incubation at a concentration of 0.5 mg/ml in 20 mM NH₄HCO₃, pH 8.0/10 mM dithiothreitol for 1 hr at 37°C. Unmodified Bb (1 mg/ml) in 20 mM NH₄HCO₃

Abbreviations: C2, C3, and C5, second, third, and fifth components of serum complement; iPr₂P-F, diisopropyl phosphorofluoridate; HPLC, high-pressure liquid chromatography; VBS⁺⁺, 5 mM sodium barbital, pH 8.5/0.5 mM CaCl₂/2.0 mM MgCl₂/40 mM NaCl.

at pH 8.0 was incubated with clostripain (enzyme/substrate weight ratio, 1:200) for 1 hr at 37°C. The digest was then freeze-dried, reduced and alkylated as described above, and subjected to gel filtration on Sephadex G-100 equilibrated with 10% acetic acid.

Cleavage at the arginyl residues of CB-II was carried out by using trypsin after succinylation. Succinylation was carried out according to the procedure of Koide *et al.* (12). At the end of the reaction the protein was dialyzed against 0.15 M NH₄HCO₃ at pH 8.5. This resulted in the appearance of a precipitate which partially redissolved on dilution. Succinylated CB-II (0.75 mg/ml) was incubated with TPCK-trypsin (Worthington, 224 units/mg) in 0.15 M NH₄HCO₃ (pH 8.5) for 2 hr at 37°C (enzyme/substrate weight ratio, 1:75). The incubation was repeated with an additional amount of enzyme. The reaction was terminated by freeze-drying. The digest was redissolved in 2.5 ml of 0.1 M NH₄HCO₃ and the soluble material was applied to a column (2 × 96 cm) of Sephadex G-50 (superfine) equilibrated in the same buffer. Fractions (1.5 ml) were collected and their absorbance at 215 nm was determined. Smaller peptides were further purified by high-pressure liquid chromatography (HPLC) using a μ -Bondapak C₁₈ column (Waters) and a 10 mM NH₄HCO₃/CH₃CN elution gradient as described by Waterfield and Scrace (13).

Polyacrylamide Gel Electrophoresis. The purity and apparent molecular weight of Bb and its peptides were assessed by electrophoresis as described by Laemmli (14). Protein bands were visualized after staining with Coomassie brilliant blue.

Amino Acid Analysis. Reduced and alkylated samples were hydrolyzed under reduced pressure at 110°C in constant boiling HCl containing 0.05% 2-mercaptoethanol. Cysteine and cystine were determined as S-carboxymethylcysteine. Tryptophan was not measured. Analyses were carried out by using a Durrum D500 analyzer.

Sequence Determination. Automated amino acid sequence determination was carried out as described by Johnson *et al.* (9). The assignment of S-carboxymethylcysteine residues was based on measurement of radioactivity (LKB-Wallac 1210 Ultrabeta counter) in 200 μ l of the butyl chloride phase in 3 ml of 1,4-dioxane containing 2% (wt/vol) naphthalene and 0.5% 2,5-diphenyloxazole.

COOH-terminal determinations were carried out on Bb, succinylated CB-II, and AC-4 by using carboxypeptidase Y (Boehringer Mannheim, 20 units/mg). Samples were incubated in 50 mM pyridine acetate at pH 5.5 for up to 2 hr at 37°C (enzyme:substrate molar ratio, 1:10). At 30, 60, and 120 min the reaction was stopped by the addition of ice-cold acetone (15). The deproteinized samples were analyzed on a Durrum D500 instrument. No amino acids were detected in control experiments in which the enzyme was incubated alone. Results have been corrected for the recovery of norleucine.

RESULTS

Initial Strategy of Sequence Determination. Fragmentation of a protein the size of Factor B, even by using site-specific chemical or enzymic procedures, would result in a complex mixture of peptides. To simplify matters, Factor B was first cleaved by Factor D in the presence of C3. The larger basic Bb fragment was easily separated from the more acidic Ba fragment and the other components of the reaction mixture by ion-exchange chromatography on DEAE-Sephadex. In this way, each fragment could be treated separately. Our initial efforts have been confined to the Bb fragment because this is considered to be the catalytically active subunit of the C3 convertase.

Table 1. Amino acid compositions of Bb and peptides CB-IV, CB-VIa, CB-II, T-1, and AC-4

Amino acid	Amino acid composition, residues/100 residues					
	Bb	CB-IV	CB-VIa	CB-II	T-1	AC-4
Asp	11.51	13.54	10.62	9.24	10.90	12.25
Thr	4.51	5.81	4.78	4.48	5.93	2.38
Ser	5.59	4.52	7.02	7.39	1.49	4.10
Hse	—	1.28	2.93	—	—	—
Glu	10.76	7.09	10.38	12.15	11.36	10.49
Pro	5.09	3.00	8.85	5.54	6.60	5.58
Gly	7.32	8.48	7.99	7.98	7.65	7.77
Ala	4.98	8.50	2.93	5.53	4.08	3.38
Val	9.53	10.44	7.72	8.67	5.56	10.38
Cys	2.75	1.18	1.93	3.22	3.97	2.12
Met	1.52	—	—	—	—	0.83
Ile	5.75	5.43	6.19	5.19	8.81	5.65
Leu	8.08	8.05	3.13	8.06	9.10	9.79
Tyr	3.92	6.66	2.50	3.50	5.84	1.62
Phe	2.73	1.28	—	3.83	3.10	5.70
His	2.63	1.55	7.52	2.20	2.13	3.69
Lys	9.25	11.20	8.15	8.71	10.00	8.67
Arg	4.08	1.99	7.39	4.31	3.46	5.61

Compositions are based on values from 24-hr hydrolyses except for Bb. In this case they represent the mean of 24-, 48-, and 72-hr hydrolysis values, except for Thr and Ser (extrapolated to zero time) and Val and Ile (72-hr value). Cysteine was determined as S-carboxymethylcysteine. Tryptophan was not determined.

The Bb fragment prepared by this way gave a single band on NaDodSO₄/polyacrylamide gel electrophoresis with an apparent molecular weight of 60,000. Amino acid analysis of Bb showed the fragment to contain approximately 8 methionine residues per mol (Table 1). Thus, cleavage with CNBr would be expected to yield nine peptides. A good separation of the CNBr digest of Bb was achieved by gel filtration, and seven pools of material were collected (Fig. 1). Only three pools are of relevance to the present work, and the full characterization of all of the CNBr peptides of Bb will be presented elsewhere. Pools CB-IV and CB-II each contained a single peptide with molecular weights of approximately 10,000 and 30,000, respectively. Pool CB-VI contained two peptides, of which one, CB-VIa, was isolated by ion-exchange chromatography on DEAE-Sephadex. The amino acid compositions of the peptides sequenced are presented in Table 1. Peptide CB-II represents approximately half of the Bb molecule and contains a high proportion of S-carboxymethylcysteine residues, suggesting a rigid structure. This is the only peptide that was not found to contain homoserine, even when the hydrolysate from an ex-

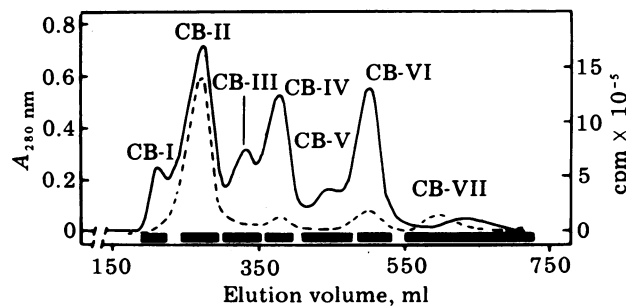


FIG. 1. Elution profile of the CNBr digest of Bb on a column (3 × 90 cm) of Sephadex G-100 in 10% acetic acid. The absorbance at 280 nm (—) and the radioactivity (---) of column fractions was determined. Bars indicated the seven pools collected; these were denoted CB-I to CB-VII. Pool CB-VI was further purified by using DEAE-Sephadex.

cessively large amount of peptide was analyzed. On this basis CB-II was assigned to the COOH-terminal portion of the molecule.

NH₂-Terminal Sequence of Bb. The Bb fragment was subjected to automated Edman degradation and a single sequence extending for 25 residues was obtained. Sequence analysis of the CNBr peptide CB-IV provided an overlap with the NH₂-terminal sequence of Bb, starting at the methionine residue in position 10, and extended the NH₂-terminal sequence to 40 amino acid residues.

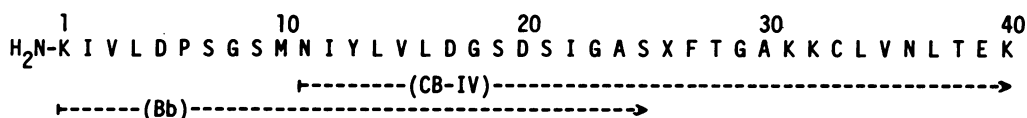
Partial Sequence of COOH-Terminal Half of Bb. The NH₂-terminal sequences of the CNBr peptides CB-VIa and CB-II (the COOH-terminal peptide) were determined for 30 and 42 cycles, respectively. Sequence analysis of a fraction (Cl-3) obtained by gel filtration of a restricted clostripain digest of Bb aligned these two peptides. Analysis of Cl-3 revealed two sequences, one of which was identical to the NH₂-terminal sequence of Bb. We also recognized the other sequence as resulting from cleavage of the Arg-Lys bond at position 5 of CB-VIa. The sequence determination was continued for 35 cycles, extending to residue 8 of CB-II (Fig. 2).

To extend the sequence of CB-II, this peptide was subfragmented by trypsin after succinylation. The digest was fractionated by gel filtration. The largest peptide (T-1), approximately 10,000 molecular weight, was a product of incomplete cleavage. This peptide contained approximately 3 arginine residues per mol (Table 1). The limited digestion probably resulted from the poor solubility of CB-II, even after succinyla-

tion. Peptide T-1 showed a single sequence overlapping the NH₂-terminal sequence of CB-II by eight amino acid residues. This provided a continuous sequence of 110 residues (Fig. 2).

The remaining tryptic peptides were relatively small and some of these were further purified by HPLC. Four peptides were subjected to sequence analysis (Fig. 2). An alternative cleavage procedure was required for their alignment. Fragment Bb was cleaved by dilute acid hydrolysis using conditions selective for Asp-Pro bonds. A peptide (AC-4), estimated 7000 molecular weight, was isolated from the digest by gel filtration. It was evident from carboxypeptidase Y digestion of Bb, CB-II, and AC-4 that the fragment was derived from the COOH-terminal portion of Bb. In all three cases the first amino acids released were leucine followed by phenylalanine, suggesting a COOH-terminal sequence of -Phe-Leu-COOH (Table 2) as reported by Lesavre *et al.* (5). The NH₂-terminal sequence of AC-4 was determined for 48 cycles and this was found to overlap with the last five residues determined for T-5a (Fig. 2). It also contains the sequences of T-4, T-5b, and T-6c, enabling the alignment of these peptides derived from succinylated CB-II. Thus, the sequence must be part of the CB-II peptide. These analyses defined the positions of 60 amino acids proximal to the COOH terminus of Bb (Fig. 2). Sequence analysis of AC-4 showed it to be contaminated with two other peptides, representing approximately 25% of the total material. We recognized the sequence of one of these as resulting from cleavage of the Asp-Pro bond at positions 5 and 6 of Bb and this

NH₂-terminal of Bb:



Partial COOH-terminal half of Bb:

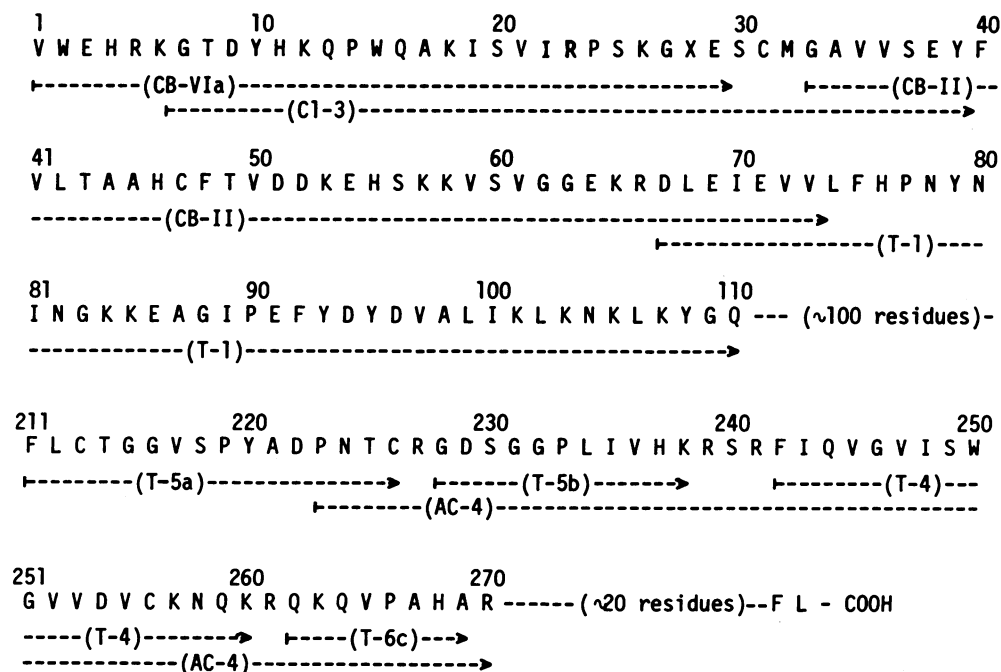


FIG. 2. Partial amino acid sequence of Bb. CB, CNBr peptides; T, tryptic peptides; Cl, clostripain peptide; AC, peptide generated by dilute acid cleavage. Positions sequenced for each peptide are underlined. Residues are in the single-letter code: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; X, unknown; Y, Tyr.

Table 2. Carboxypeptidase Y digests of Bb, CB-II, and AC-4

Time, min	Amino acid, mol/mol peptide					
	Bb		CB-II		AC-4	
	Leu	Phe	Leu	Phe	Leu	Phe
30	0.93	0.43	1.88	1.02	0.88	0.36
60	1.51	0.49	2.36	1.16	1.01	0.42
120	2.49	0.86	2.71	1.39	1.28	0.53

explains the small amount of methionine present in AC-4 (Table 1).

From the size of AC-4 and its amino acid composition, we estimate this peptide to be approximately 70 residues long. This leaves a gap of some 20 residues between the last amino acid of AC-4 positioned (Arg-270) and the COOH terminus of the molecule. With similar reasoning, approximately 100 residues was estimated as the distance separating the 110- and 60-residue segments. The partial sequence of the COOH-terminal portion of the Bb molecule presented in Fig. 2 has been numbered in accordance with these estimates.

DISCUSSION

Sequence analysis of the NH₂ termini of Bb and of CNBr fragment CB-IV provided the assignment of the first 40 amino acid residues with the exception of position 26. This site may well represent an asparagine-linked carbohydrate moiety because the location of a threonine at position 28 is consistent with an Asn-X-Thr sequence requirement for carbohydrate attachment (16). No homology exists between the NH₂-terminal sequence of the catalytic fragment of Factor B and that of other serine proteases.

We have established a partial sequence for the COOH-terminal half of the Bb fragment. This was achieved by automated sequence analysis of fragments obtained by CNBr cleavage, dilute acid cleavage, and enzymatic digestion with trypsin and clostripain. These sequences are contained in two segments of 110 and 60 residues and represent one-third of the amino acids in Bb. Their relative alignment is shown in Fig. 2.

Bb contains most of the functional residues of serine proteases. When the partial sequence of the COOH-terminal portion of Bb was compared to that of the catalytic chains of other serine proteases according to the alignment of Dayhoff (17), extensive homology was observed (Fig. 3). The three active site residues—His-57, Asp-102, and Ser-195 (chymotrypsinogen numbering)—are found at positions 46, 96, and 230, respectively. The linear distribution of these active site residues is in accordance with that found in serine proteases. The spacing between His and Asp compares favorably with that in chymotrypsin (44 residues).

The sequence at residues 228–234 (-Gly-Asp-Ser-Gly-Gly-Pro-Leu-) and that at residues 246–251 (-Gly-Val-Ile-Ser-Trp-Gly-) are homologous to the primary and secondary binding sites, respectively. In serine proteases these two sites are linked by a disulfide bond which is always conserved. In Bb this is probably the cystine residues at positions 226 and 256. Two other disulfide bonds are conserved in all serine proteases. One of these, the "His-loop," is formed by the cysteine residues equivalent to those located at positions 31 and 47 of the Bb sequence. It is likely that residue 213 represents half of the "Met-loop," the other invariant disulfide bond.

Other invariant residues underlined in Fig. 3 are present in the partial sequence with two exceptions. The Met-32 substitutes for Gly. A more important substitution is the replacement of Gly by Arg at position 5. This Gly is one of four conserved residues involved in the activation mechanism of serine proteases, as has been shown for chymotrypsin and trypsin (18, 19). This involves cleavage between a basic residue and this short hydrophobic segment. The newly formed NH₂ terminus then ion-pairs with an Asp (194 in chymotrypsinogen numbering) located in the primary binding pocket. In Factor B, activation is associated with the splitting of a peptide bond some 300 residues nearer the NH₂ terminus. The sequence equivalent to Arg-Ile-Val-Gly-Gly has not been found in the expected position (Fig. 3). This suggests that, whereas the active site of Bb is likely to be similar to that of other serine proteases, the activation mechanism is different. A different type of confor-

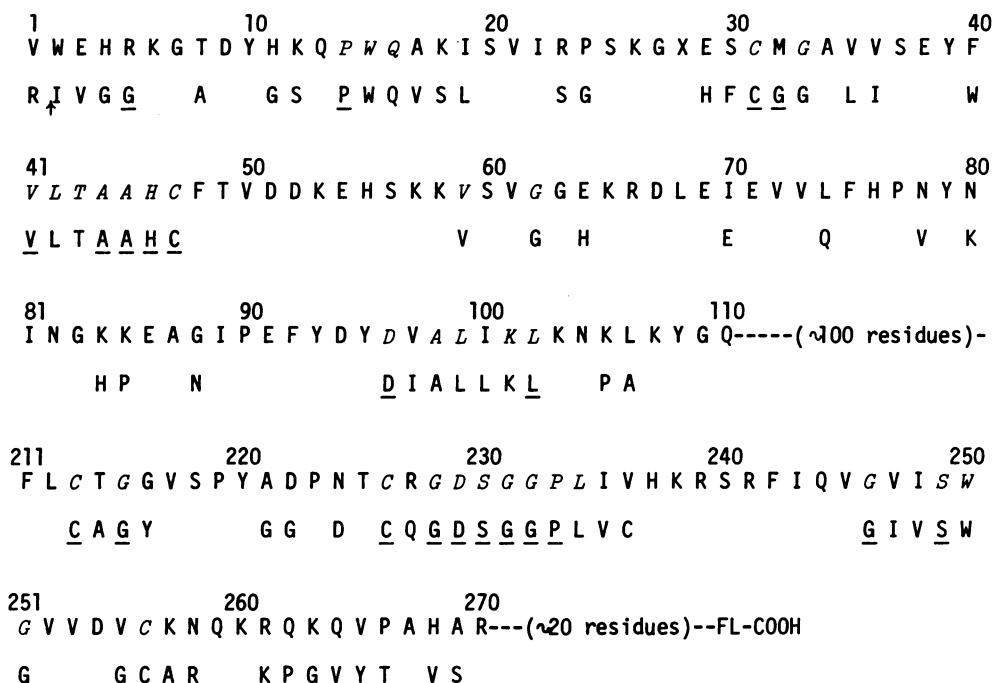


FIG. 3. Comparison of the partial amino acid sequence of the COOH-terminal portion of Factor B with corresponding conserved segments of serine proteases, according to Dayhoff (17). The residues underlined are invariant in the known sequences (17). † denotes the peptide bond cleaved during activation of the zymogens. Conserved residues in Bb are shown in italics.

mational change may occur to make the catalytic site accessible to the substrate. Also surprising is the presence of Asn in position 224 because C3 convertase has a tryptic-like specificity splitting an Arg-Ser bond (20), and Asp would be expected as in trypsin. In most serine proteases the position equivalent to 226 prior to the active Ser is neutral but, in Factor B, Arg is found, whereas in bovine prothombin and human Factor D, Glu and Lys, respectively, are in this position (9, 21).

C2 is equivalent to Factor B in the classical pathway of complement activation and is similar in gross structure and in activation to Factor B, and both are encoded by closely linked genes in the major histocompatibility locus (22). It is probable therefore that C2 and Factor B are both novel types of serine proteases.

We thank Messrs. A. Willis, P. Fisher, and A. Gascoyne for their technical assistance. D.L.C. is a Nuffield Dominions Demonstrator.

1. Porter, R. R. (1979) in *Defense and Recognition IIB*, ed. Lennox, E. S. (University Park Press, Baltimore), Vol. 23, pp. 177-212.
2. Götze, O. & Müller-Eberhard, H. J. (1976) *Adv. Immunol.* **24**, 1-35.
3. Götze, O. (1975) in *Proteases and Biological Control*, Cold Spring Harbor Conferences on Cell Proliferation, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 2, pp. 255-272.
4. Kerr, M. A. (1979) *Biochem. J.* **183**, 615-622.
5. Lesavre, P. H., Hugli, T. E., Esser, A. F. & Müller-Eberhard, H. J. (1979) *J. Immunol.* **123**, 529-534.
6. Medicus, R. G., Götze, O. & Müller-Eberhard, H. J. (1976) *Scand. J. Immunol.* **5**, 1049-1055.
7. Vogt, W., Dames, W., Schmidt, G. & Dieminger, L. (1977) *Immunochimistry* **14**, 201-205.
8. Neurath, H. & Walsh, K. A. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3825-3832.
9. Johnson, D. M. A., Gagnon, J. & Reid, K. B. M. (1980) *Biochem. J.* **187**, 863-874.
10. Tack, B. F. & Prahl, J. W. (1976) *Biochemistry* **15**, 4513-4521.
11. Fraser, K. J., Poulsen, K. & Huber, E. (1972) *Biochemistry* **11**, 4974-4977.
12. Koide, A., Titani, K., Ericsson, L. H., Kumar, S., Neurath, H. & Walsh, K. A. (1978) *Biochemistry* **17**, 5657-5672.
13. Waterfield, M. & Scrace, G. (1980) in *Biological and Biomedical Applications of Liquid Chromatography III*, ed. Hawk, G. L. (Dekker, New York), in press.
14. Laemmli, J. K. (1970) *Nature (London)* **227**, 680-685.
15. Castel, A., De Graaf, J. M. & Kraal, B. (1979) *J. Chromatogr.* **174**, 474-477.
16. Kinkvist, K. E. & Lennarz, W. J. (1978) *J. Supramol. Struct.* **8**, 51-65.
17. Young, C. L., Barker, W. C., Tomaselli, C. M. & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 3, pp. 73-93.
18. Freer, S. T., Kraut, J., Robertus, J. D., Wright, H. T. & Xuong, N. H. (1970) *Biochemistry* **9**, 1997-2009.
19. Huber, R. & Bode, W. (1978) *Acc. Chem. Res.* **11**, 114-122.
20. Tack, B. F., Morris, S. C. & Prahl, J. W. (1979) *Biochemistry* **18**, 1497-1503.
21. Magnusson, S., Petersen, T. E., Sottrup-Jensen, L. & Claeys, H. (1975) in *Proteases and Biological Control*, Cold Spring Harbor Conferences on Cell Proliferation, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 2, pp. 123-149.
22. Barnstaple, C. J., Jones, E. A. & Bodmer, W. F. (1979) in *Defense and Recognition IIA*, ed. Lennox, E. S. (University Park Press, Baltimore), Vol. 22, pp. 151-225.