

Primary structure of bovine complement activation fragment C4a, the third anaphylatoxin

Purification and complete amino acid sequence

Marjorie A. SMITH,* Linda M. GERRIE, Bryan DUNBAR and John E. FOTHERGILL
Department of Biochemistry, University of Aberdeen, Marischal College, Aberdeen AB9 1AS, Scotland, U.K.

(Received 15 April 1982/Accepted 11 August 1982)

Purification of C4a from heat-activated bovine plasma by elution from CM-Sephadex C-50 at pH 7.4 and gel filtration on Sephadex G-50 gives a 20% yield of pure C4a. The complete amino acid sequence of bovine C4a has been determined by automatic sequencer degradation of CNBr and enzymic fragments, and by carboxypeptidase digestion. The 77-residue bovine sequence shows 12 differences from the human sequence with five of these differences occurring in the C-terminal 11 residues. The sequence of C4a confirms earlier suggestions of homology with C3a and C5a; the three sequences show an almost equal number of identities with each other. The six cysteine residues of the 'disulphide knot' are conserved as well as seven other residues including the C-terminal arginine.

Activation of the complement system generates active proteolytic enzymes that subsequently liberate a number of relatively small protein fragments from C3, C4 and C5 (reviewed by Fothergill & Anderson, 1978; Reid & Porter, 1981). Component C4 is cleaved by C1s to give C4a from the N-terminus of the α -chain of C4 (Booth *et al.*, 1979*b*), and C3 and C5 are cleaved by the classical or alternative pathway convertases to yield C3a and C5a from the α -chains of C3 and C5. C3a and C5a are known as anaphylatoxins because of their substantial biological activities such as releasing histamine from mast cells, causing contraction of smooth muscle, and increasing vascular permeability (reviewed by Hugli & Müller-Eberhard, 1978); they show substantial homology of amino acid sequence.

The first preliminary sequence evidence for C4a (Booth *et al.*, 1979*b*) suggested that bovine C4a might be homologous with C3a and C5a. Similar sequence evidence obtained for human C4a (Gorski *et al.*, 1981) has recently been extended to the whole

Abbreviations used: the nomenclature of complement components and subcomponents is that recommended by the World Health Organisation (1968). Activated components are indicated by a bar, e.g. C1s. SDS, sodium dodecyl sulphate; h.p.l.c., high-pressure liquid chromatography.

* Present address: Department of Pathology, University of Dundee, Ninewells Hospital, Dundee, Scotland, U.K.

molecule (Moon *et al.*, 1981). Comparisons can now be made between bovine and human C4a, and between C4a, C3a and C5a, showing all of these molecules to be homologous. Human C4a has been shown to possess some anaphylatoxin activity (Gorski *et al.*, 1979). Recent studies of the thiol ester regions of C4b (Campbell *et al.*, 1981; Harrison *et al.*, 1981) and C3b (Tack *et al.*, 1980) have emphasized the homology of C4 to C3 as well as to α_2 -macroglobulin (Swenson & Howard, 1980).

Materials and methods

Materials

Fresh bovine blood (9 litres) was collected into 0.136 M-trisodium citrate anticoagulant (1 litre) at the slaughterhouse. Centrifugation at 1800 g for 50 min at 4°C gave clear plasma.

Chemicals were obtained as follows: CM-Sephadex C-50, Pharmacia; Quadrol/trifluoroacetic acid buffer in propanol/water (3:4, v/v), pH 9.0, protected by aminocellulose, Fluka; heptafluorobutyric acid, phenylisothiocyanate, sequencer grade solvents, Rathburn Chemicals, Walkerburn, Peeblesshire, Scotland, U.K. The sources of other chemicals have been described previously (Booth *et al.*, 1979*a,b*; Campbell *et al.*, 1979*a,b*; Nisbet *et al.*, 1981).

Cleavage of C4

Whole plasma was incubated at 37°C at pH 7.4

for 2 h. Conditions similar to these are known to cause activation of bovine C4 by C1r (Booth *et al.*, 1979a) and to cause activation of human (Ziccardi & Cooper, 1976) and bovine C1r (Campbell *et al.*, 1979b). Activated plasma was then frozen by storage at -20°C for 4 h, and allowed to thaw overnight at 4°C . This freezing and thawing caused the aggregation of a sticky precipitate formed during activation, which could be removed conveniently by centrifugation after this process.

Purification of C4a

The purification schedule is outlined in Scheme 1. Thawed, activated plasma (2 litres) was centrifuged at 1800g for 50 min at 4°C , and the supernatant applied to a column (5 cm \times 35 cm) of CM-Sephadex C-50 equilibrated in 28 mM-phosphate buffer, pH 7.4, containing 2 mM-EDTA and 50 mM-NaCl. C4a was eluted with a linear gradient (2 \times 1 litre) of 50–300 mM-NaCl followed by a wash of 1 litre of 300 mM-NaCl (Fig. 1). C4a, detected by electrophoresis in polyacrylamide-gel (10%) containing SDS was pooled and applied to a smaller column (5 cm \times 4 cm) of CM-Sephadex C-50 equilibrated in the same starting buffer. C4a was recovered by stepwise elution with buffer containing 300 mM-NaCl. C4a-containing fractions were dialysed against 1% (w/v) NH_4HCO_3 and freeze dried. A final stage of purification was done by molecular sieving on Sephadex G-50. Some C4a was reduced with a 60-fold molar excess of dithiothreitol in

0.4 M-Tris/HCl, pH 8.6, containing 6 M-guanidine hydrochloride and 2 mM-EDTA, and carboxymethylated with iodo[2- ^{14}C]acetate (Crestfield *et al.*, 1963). Before sequencing this material was further purified by gel permeation h.p.l.c. on a tandem combination of Toyo Soda TSK SW 3000 and SW 2000 columns (0.45 cm \times 60 cm) in 0.1% (v/v) trifluoroacetic acid.

CNBr cleavage of C4a and purification of fragments

Reduced carboxymethylated C4a was treated with a 100-fold molar excess of CNBr in 70% (v/v) formic acid at room temperature for 24 h in the dark. After dilution and freeze-drying, the mixture of peptides was dissolved in 0.1% (v/v) trifluoroacetic acid and applied to the tandem arrangement of TSK SW 3000 and SW 2000 columns (0.45 cm \times 60 cm) and eluted at 1 ml/min by 0.1% trifluoroacetic acid. Peptides were detected by monitoring at 214 nm.

Digestion of C4a by staphylococcal proteinase and purification of peptides

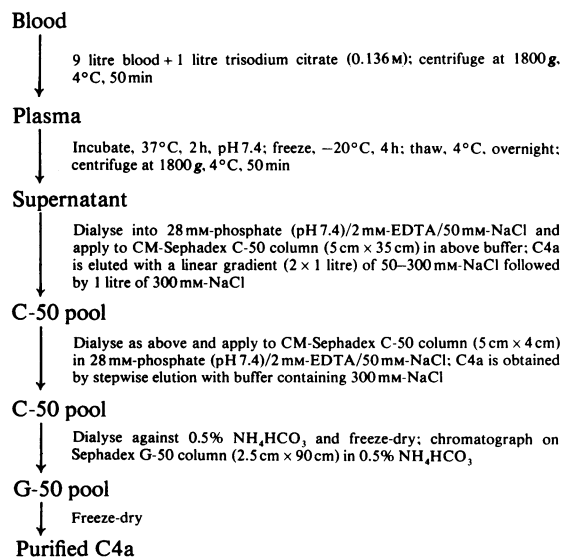
Non-reduced C4a was digested in 1% (w/v) NH_4HCO_3 by staphylococcal proteinase (EC 3.4.21.19) at 1:20 (w/w) enzyme:C4a ratio. The digest was freeze-dried, dissolved in 0.1% trifluoroacetic acid and applied to the TSK h.p.l.c. columns as above.

Amino acid analyses

These were carried out as described by Campbell *et al.* (1979a).

Automatic N-terminal sequence determination

A Beckman 890C sequencer equipped with the Beckman cold-trap accessory was operated with a double-cleavage, double-extraction programme developed by Dr. Jean Gagnon from the method of Hunkapiller & Hood (1978) using 0.25 M-Quadrol. Polybrene (3 mg) was added to the cup before each sample (Klapper *et al.*, 1978). The thiazolinones were extracted from the spinning cup with butyl chloride containing dithioerythritol (10 mg/l) and tributylphosphine (100 $\mu\text{l/l}$) (Frank, 1979). Samples were collected into tubes containing 0.2 ml of freshly prepared 0.1 M-HCl containing 0.1% (v/v) ethanethiol. Solvent was blown off by a stream of nitrogen at room temperature until only the aqueous acid layer remained. Samples were then heated at 80°C for 10 min, frozen and dried in a vacuum desiccator at room temperature. The phenylthiohydantoin derivatives were identified by h.p.l.c. on Waters C_{18} - μ Bondapak columns (0.39 cm \times 30 cm) using an acetate buffer pH 3.4/methanol gradient system (Bridgen *et al.*, 1976) and a 254 nm detector. Yields of serine and threonine were normally about 30%.



Scheme 1. Schedule for the purification of C4a

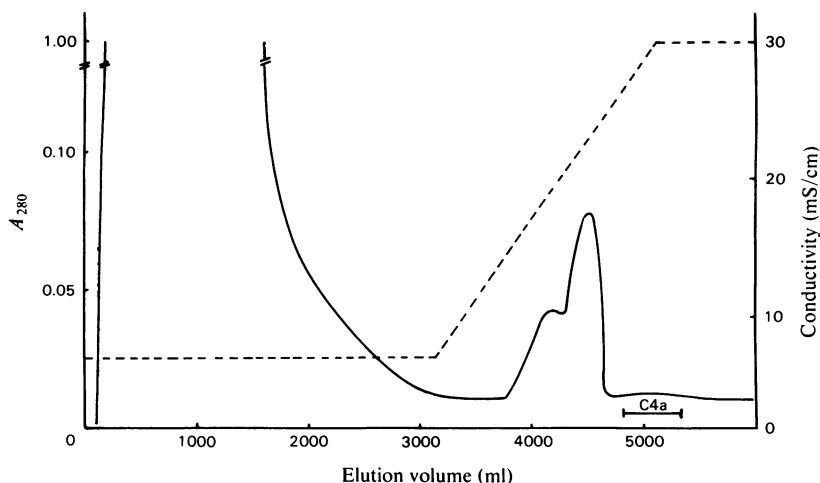


Fig. 1. Chromatography of C4a on CM-Sephadex C-50

Activated plasma from 9 litres of blood was applied to a column (5 cm × 35 cm) of CM-Sephadex C-50 in 28 mM-phosphate buffer (pH 7.4)/2 mM-EDTA/50 mM-NaCl. After the first protein peak had eluted a linear gradient (2 × 1 litre) of 50–300 mM-NaCl was applied, followed by buffer containing 300 mM-NaCl. C4a was located by SDS/polyacrylamide-gel electrophoresis and pooled as shown. —, A₂₈₀; ---, conductivity (mS/cm).

Yields of *S*-carboxymethylcysteine were almost theoretical, and were confirmed by radioactive counting. Amino acid analysis after back hydrolysis with HI was used to confirm any doubtful residues (Smithies *et al.*, 1971).

Sequence determination by carboxypeptidase Y

This was done by the method of Kuhn *et al.* (1974).

Results

Purification of C4a

The final yield of C4a was 9 mg from 2 litres of plasma, representing a yield of approx. 20%. Separations were carried out entirely at pH 7.4 suggesting that acidic conditions are not necessary to release C4a from C4b (Booth *et al.* 1979b; Budzko & Müller-Eberhard, 1970).

The product of this method was C4a without the C-terminal arginine residue. The addition of 6-amino-hexanoic acid (1 M), a plasma carboxypeptidase B inhibitor (Vallota & Müller-Eberhard, 1973) at the activation and following steps was not sufficient to prepare material containing the terminal arginine. This was achieved by purifying whole C4 (Booth *et al.*, 1979a) followed by activation using CIs (Campbell *et al.*, 1979b) in the presence of 1 M-6-amino-hexanoic acid.

Purification of CNBr fragments

Gel permeation h.p.l.c. in 0.1% trifluoroacetic

acid separated undigested material and two fragments, CN-1 and CN-2 (Fig. 2).

Purification of staphylococcal proteinase peptides

Gel permeation h.p.l.c. of non-reduced C4a digested by staphylococcal proteinase showed three main peaks (Fig. 3).

Amino acid analysis

Analyses of reduced carboxymethylated C4a, CNBr fragments and staphylococcal proteinase peptides are shown in Table 1. The single methionine residue of C4a is consistent with the generation of the two CNBr fragments, CN-1 and CN-2. The analyses of the staphylococcal proteinase peptides SP-1, SP-2 and SP-3 suggest that SP-3 is the *N*-terminal decapeptide, SP-2 is the *C*-terminal hexadecapeptide and that SP-1 is mainly the 'disulphide knot' with some undigested material. The 'disulphide knot' probably contains several peptides held together by the disulphide bonds. This digest was intentionally carried out on non-reduced material in order to reduce the complexity of the mixture of peptides to be purified. The amino sugar content was found to be insignificant.

Amino acid sequence determination

The results of automatic sequencer runs on reduced carboxymethylated C4a, CNBr fragment 1 and staphylococcal proteinase peptide 2 are shown in Fig. 4 together with the results of carboxypeptidase Y digestion on the arginine-containing and

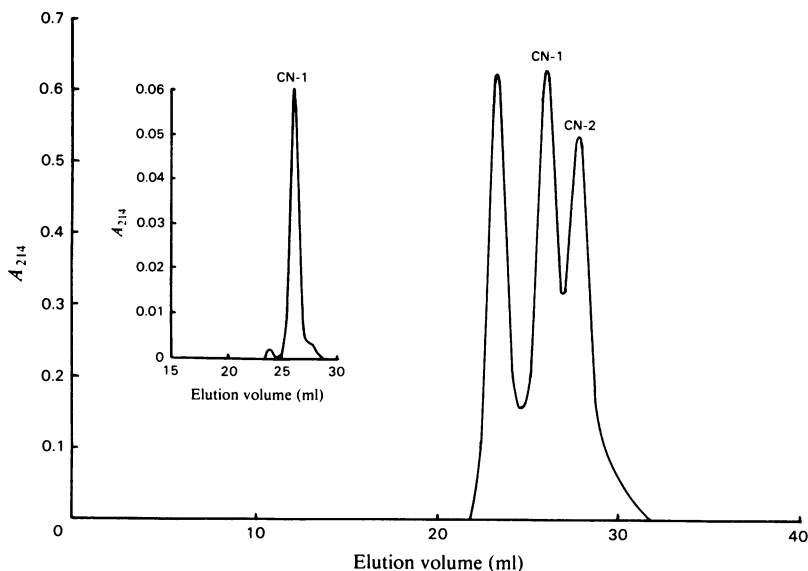


Fig. 2. Separation of reduced carboxymethylated C4a CNBr fragments (50 nmol in 200 μ l) by gel permeation h.p.l.c. in 0.1% (v/v) trifluoroacetic acid

Columns (0.75 cm \times 60 cm) of TSK SW-3000 and SW-2000 were used together in order of decreasing pore size. Flow rate was 1.0 ml/min provided by a Waters 6000A pump. Detection was at 214 nm using a Waters 441 monitor. Inset shows a repeat analytical scale run of peak CN-1 (2 nmol) after pooling several preparative scale runs.

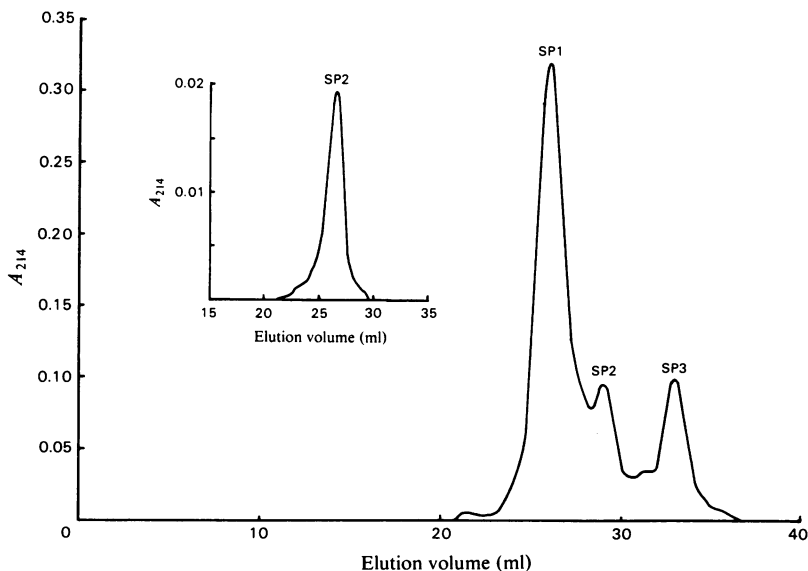


Fig. 3. Separation of staphylococcal proteinase peptides (25 nmol in 100 μ l) of C4a

Columns and conditions as in Fig. 1. Inset shows a repeat analytical scale run of peak SP-2 (1 nmol) after pooling several preparative scale runs.

arginine-lacking forms of C4a. There is good agreement between the amino acid analyses and the sequences, and the overlaps are thoroughly estab-

lished. The molecular weight of bovine C4a calculated from the sequence is 8692. There appears to be no glycosylation.

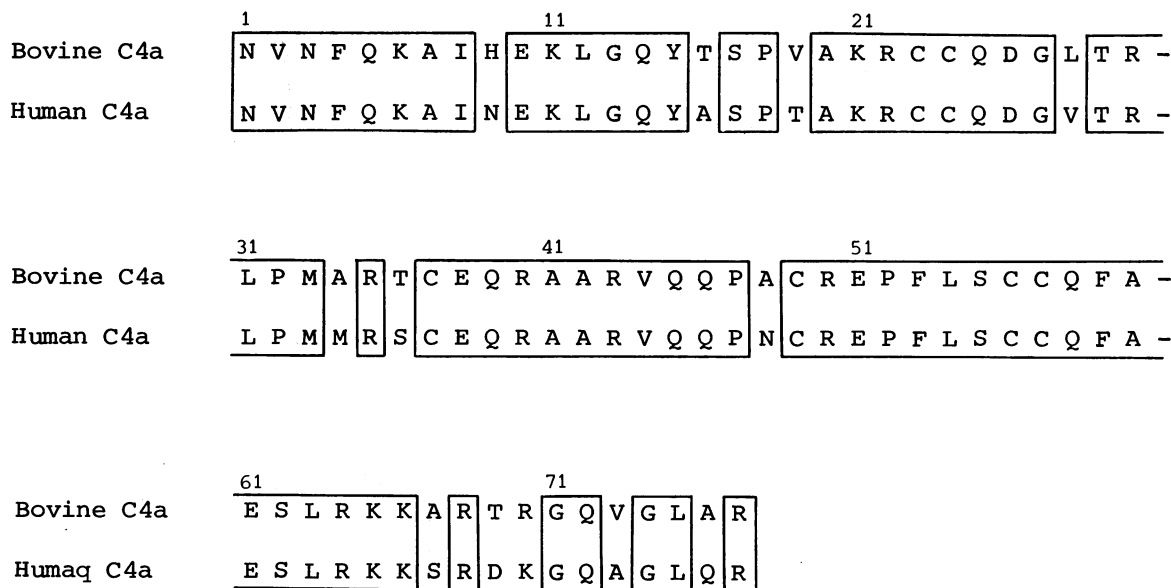


Fig. 5. Comparison of the amino acid sequence of bovine C4a with that of human C4a (Moon *et al.*, 1981). Identical residues are boxed. Continuity between residues 33 and 35 and the position of residue 34 in the human sequence have not been positively established.

are obvious (Fig. 6). Over all six sequences there is complete retention of the six cysteine residues, four as double cysteine sequences and two as single cysteine residues, as well as tyrosine at 15, glycine at 27, arginine at 40, phenylalanine at 54, arginine at 65, leucine at 78 and the C-terminal arginine. Comparing the residues conserved in both species of C4a with those conserved in both species of C3a and C5a, there are 25 residues conserved between C4a and C5a, and 22 residues conserved between C4a and C3a. A similar comparison of C3a and C5a shows 19 conserved residues. These figures are sufficiently similar to suggest that the gene duplication events that gave rise to separate genes for C3, C4 and C5 occurred at approximately the same stage in evolution.

These comparisons are particularly significant when they are correlated with the X-ray diffraction structure established for C3a (Hüber *et al.*, 1980). The crystal structure allows the assignment of the disulphide bridge arrangement of C3a, and also of C4a and C5a if one assumes that the tertiary structures are homologous. As one might expect, the variability between sequences is mainly located in regions of the chain that forms loops away from the 'disulphide knot'. This is presumably the structural basis for the well defined immunochemical distinction among the anaphylatoxins (Moon *et al.*, 1981). The most surprising finding is that so much

variability occurs in the last 15 residues which span between the highly conserved 'disulphide knot' and the highly conserved C-terminal arginine residue that is essential for biological activity in all three anaphylatoxins. Moreover, this variability extends to a deletion that must result in the C-terminal arginine varying in distance from the 'disulphide knot'. From the conserved arginine at position 65, C3a and C4a have 13 residues to the C-terminus, whereas C5a has 12 residues. This has implications for the stereochemistry of binding to the cell surface receptor. In this context the 'disulphide-knot' structure of the anaphylatoxins can be seen as a basic structure that occurs in other biologically active small proteins such as the α -subunits of the glycoprotein hormones (Giudice & Pierce, 1978), the neurophysins (Breslow, 1979) and some snake toxins (Kim & Tamiya, 1981; Walkinshaw *et al.*, 1980).

We thank Dr. Jean Gagnon for much helpful advice with the sequencer and for providing his sequencer program, Dr. Linda Fothergill for amino acid analyses, Mrs. Jean Bathgate, Mr. Ian Davidson and Mr. Charles Dawson for technical services, W. Donald and Son, Portlethen, Kincardineshire, for their generous co-operation in the collection of blood, the S.E.R.C. for providing the sequencer facility and the M.R.C. for financial support.

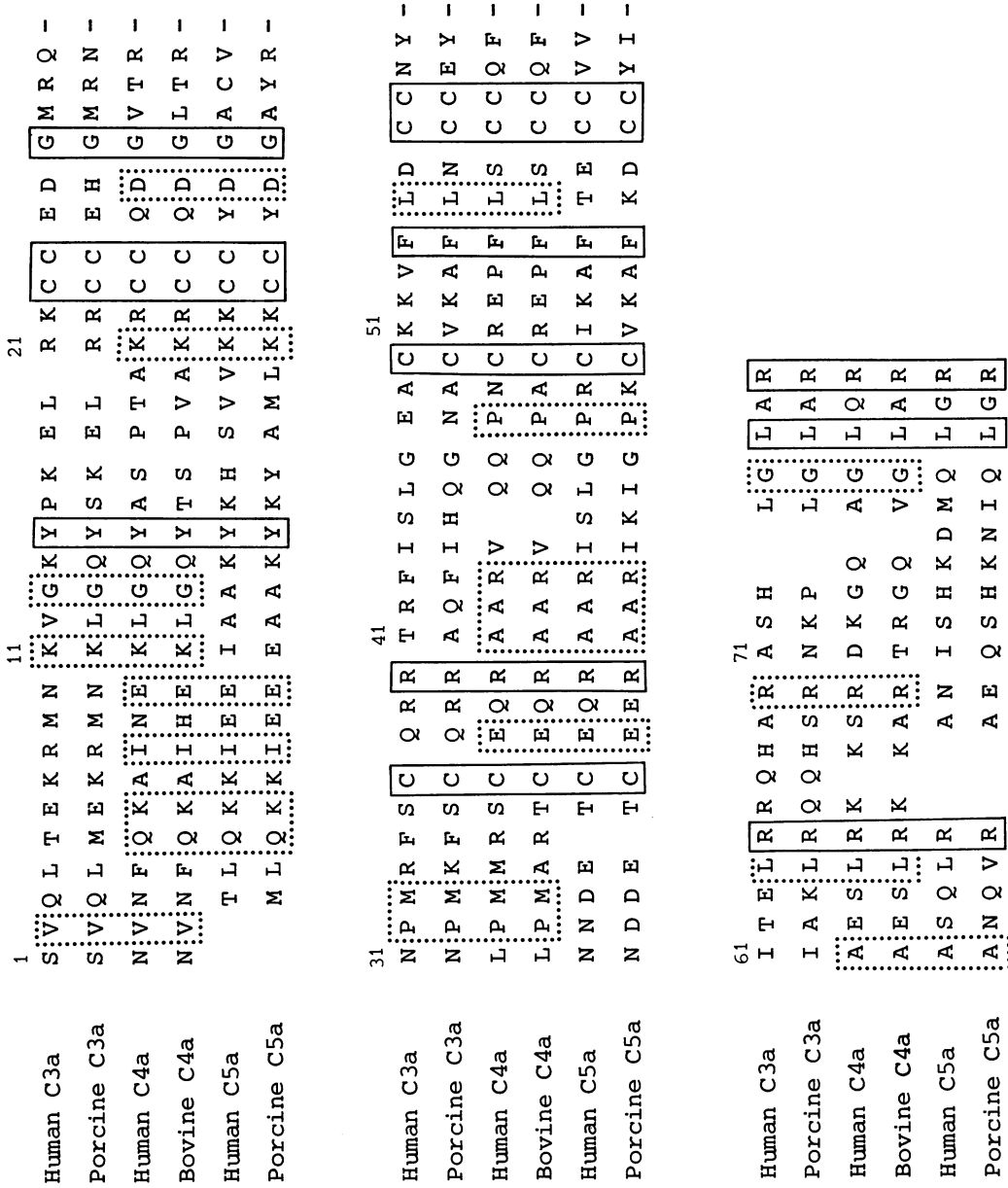


Fig. 6. Comparison of the amino acid sequences of C3a, C4a and C5a. The bovine C4a sequence is shown in alignment with the human C4a sequence (Moon *et al.*, 1981), with human and porcine C3a sequences (Hugli & Müller-Eberhard, 1978), and with human (Hugli & Müller-Eberhard, 1978) and porcine (Gerard & Hugli, 1980) C5a sequences. Sequences have been aligned to maximize homology. Residues identical in all six sequences are boxed with solid lines. Residues identical in both C3a and C4a sequences and residues identical in both C5a and both C4a sequences are boxed with broken lines.

References

- Booth, N. A., Campbell, R. D. & Fothergill, J. E. (1979a) *Biochem. J.* **177**, 959–965
- Booth, N. A., Campbell, R. D., Smith, M. A. & Fothergill, J. E. (1979b) *Biochem. J.* **183**, 573–578
- Breslow, E. (1979) *Annu. Rev. Biochem.* **48**, 251–274
- Bridgen, P. J., Cross, G. A. M. & Bridgen, J. (1976) *Nature (London)* **263**, 613–614
- Budzko, D. B. & Müller-Eberhard, H. J. (1970) *Immunochemistry* **7**, 227–234
- Campbell, R. D., Booth, N. A. & Fothergill, J. E. (1979a) *Biochem. J.* **177**, 531–540
- Campbell, R. D., Booth, N. A. & Fothergill, J. E. (1979b) *Biochem. J.* **183**, 579–588
- Campbell, R. D., Gagnon, J. & Porter, R. R. (1981) *Biochem. J.* **199**, 359–370
- Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622–627
- Dayhoff, M. O. (ed.) (1972) *Atlas of Protein Sequence and Structure*, vol. 5, National Biomedical Research Foundation, Silver Spring, MD
- Fothergill, J. E. & Anderson, W. H. K. (1978) *Curr. Top. Cell. Regul.* **13**, 259–311
- Frank, G. (1979) *Hoppe-Seyler's Z. Physiol. Chem.* **360**, 997–999
- Gerard, G. & Hugli, T. E. (1980) *J. Biol. Chem.* **255**, 4710–4715
- Giudice, L. C. & Pierce, J. G. (1978) in *Structure and Function of the Gonadotropins* (McKerns, K. W., ed.), pp. 81–110, Plenum, New York and London
- Gorski, J. P., Hugli, T. E. & Müller-Eberhard, H. J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5299–5302
- Gorski, J. P., Hugli, T. E. & Müller-Eberhard, H. J. (1981) *J. Biol. Chem.* **256**, 2707–2711
- Harrison, R. A., Thomas, M. L. & Tack, B. F. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7388–7392
- Huber, R., Scholze, H., Paques, E. P. & Deisenhofer, J. (1980) *Hoppe-Seyler's Z. Physiol. Chem.* **361**, 1389–1399
- Hugli, T. E. & Müller-Eberhard, H. J. (1978) *Adv. Immunol.* **26**, 1–53
- Hunkapiller, M. W. & Hood, L. E. (1978) *Biochemistry* **17**, 2124–2133
- Kim, H. S. & Tamiya, N. (1981) *Biochem. J.* **199**, 211–218
- Klapper, D. G., Wilde, C. E. & Capra, J. D. (1978) *Anal. Biochem.* **85**, 126–131
- Kuhn, R. W., Walsh, K. A. & Neurath, H. (1974) *Biochemistry* **13**, 3871–3877
- Moon, K. E., Gorski, J. P. & Hugli, T. E. (1981) *J. Biol. Chem.* **256**, 8685–8692
- Nisbet, A. D., Saundry, R. H., Moir, A. J. G., Fothergill, L. A. & Fothergill, J. E. (1981) *Eur. J. Biochem.* **115**, 335–345
- Reid, K. B. M. & Porter, R. R. (1981) *Annu. Rev. Biochem.* **50**, 433–464
- Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G. & Ballantyne, D. L. (1971) *Biochemistry* **10**, 4912–4921
- Swenson, R. & Howard, J. B. (1980) *J. Biol. Chem.* **255**, 8087–8091
- Tack, B. F., Harrison, R. A., Janatova, J., Thomas, M. L. & Prahl, J. W. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5764–5768
- Vallota, E. H. & Müller-Eberhard, H. J. (1973) *J. Exp. Med.* **137**, 1109–1123
- Walkinshaw, M. D., Saenger, W. & Maelicke, A. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2400–2404
- World Health Organization (1968) *WHO Bull.* **39**, 935–938 [or (1970) *Immunochemistry* **7**, 137–142]
- Ziccardi, R. J. & Cooper, N. R. (1976) *J. Immunol.* **116**, 504–509