*Biochem. J.* (1981) **199**, 359–370
*Printed in Great Britain*



# Amino acid sequence around the thiol and reactive acyl groups of human complement component C4

R. Duncan CAMPBELL, Jean GAGNON and Rodney R. PORTER
*M.R.C. Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU*

(*Received 13 April 1981/Accepted 7 July 1981*)

Activation of the fourth component of complement (C4) by $C\bar{1}s$ results in the generation of a reactive acyl group, able to react with putrescine, and in the release of a free thiol group that cannot be detected in the native haemolytically active molecule. Both the reactive acyl group and the free thiol group have been shown to reside in C4d, a fragment of the α′-chain of C4b derived from digestion of the molecule with the control proteins C3b inactivator and C4-binding protein. Peptides derived from CNBr digestion of [1,4-$^{14}$C]putrescine-labelled and iodo[2-$^{14}$C]acetic acid-labelled C4d have been obtained and used to establish a continuous sequence of 88 residues from the *N*-terminus of the molecule. The thiol and reactive acyl groups are contained in an octapeptide that shows near identity with the equivalent sequences reported for $\alpha_2$-macroglobulin and C3. Other adjacent short sections also show homology of sequence between the three proteins, and it is highly likely that they contribute to the overall structure that gives a unique reactivity to the thiol ester bond postulated to exist in the native forms of the three proteins.

Complement is a complex mixture of plasma proteins that play an important role in the immune defence against infection. Activation by immunoglobulins and fungal and bacterial products results in the sequential interaction of the component proteins and in the destruction and elimination of foreign material from the body. The classical pathway of complement is initiated by the binding of the first component, C1, to the Fc region of the antibody molecule in antibody–antigen aggregates and results in the formation of an active proteinase, subcomponent $C\bar{1}s$. This activates both C4 and C2, which form the complex proteinase $C\bar{4}\bar{2}$, the classical-pathway C3 convertase, capable of continuing the activation process by cleaving C3 (Lachmann, 1979; Porter & Reid, 1979; Reid & Porter, 1981).

The fourth component of complement, C4, a glycoprotein of approx. mol.wt. 200000, is composed of three disulphide-linked polypeptide chains, an α-chain of mol.wt. 93000, a β-chain of mol.wt. 75000 and a γ-chain of mol.wt. 33000 (Schreiber & Müller-Eberhard, 1974; Nagasawa & Stroud, 1976). Activation of the molecule by $C\bar{1}$ results in the cleavage of an 8000-mol.wt. peptide, C4a, from the *N*-terminus of the α-chain and the generation elsewhere in the α-chain of a reactive group that allows nascent C4b to bind covalently to the Fd region of the heavy chain of IgG (Campbell *et al.*, 1980) or to surface structures on erythrocyte membranes (Law *et al.*, 1980). In this respect, C4 appears to be similar to C3, which also generates a reactive group (Law & Levine, 1977), allowing the molecule to bind covalently to either IgG (Gadd & Reid, 1981) or to erythrocyte membranes (Law & Levine, 1977; Law *et al.*, 1979). The reactive group in C4 exposed after activation reacts rapidly with water. Thus, if the C4 does not bind immediately after activation, it remains in the fluid phase as inactive C4b. C4b, whether bound or in the fluid phase, can be further degraded by specific control proteins C3b inactivator (C3bINA) and C4-binding protein to give rise to the fragments C4c (mol.wt. 150000) and C4d (mol.wt. 44500) (Shiraishi & Stroud, 1975; Nagasawa & Stroud, 1977; Fujita

Abbreviations used: The nomenclature of complement components is that recommended by the World Health Organisation (1968); $iPr_2P$-F, di-isopropyl phosphorofluoridate; IgG, immunoglobulin G; Fd, *N*-terminal half of the heavy chain of IgG; SDS, sodium dodecyl sulphate; $Nbs_2$, 5,5′-dithiobis-(2-nitrobenzoic acid); Tos-Phe-$CH_2Cl$, 1-chloro-4-phenyl-3-L-tosylamidobutan-2-one ('TPCK'); h.p.l.c., high-pressure liquid chromatography.



0306-3275/81/110359-12$01.50/1 

*et al.*, 1978; Nagasawa *et al.*, 1980; Press & Gagnon, 1981).

During activation of C4 by $C\bar{1}s$, a free thiol group is generated that cannot be detected in the native haemolytically active molecule (Janatova *et al.*, 1979; Reboul *et al.*, 1980; Janatova & Tack, 1981). Treatment of native haemolytically active C4 with small amine nucleophiles such as $NH_3$, methylamine and hydrazine causes inactivation of the molecule (Gordon *et al.*, 1926; Müller-Eberhard & Biro, 1963; Gorski & Howard, 1980; Janatova & Tack, 1981) and generation of the free thiol group (Janatova & Tack, 1981). The free thiol group has been localized to the C4d fragment of the α′-chain [the preceding paper (Press & Gagnon, 1981)]. A similar finding has also been demonstrated for C3 (Janatova *et al.*, 1980*a*,*b*; Pangburn & Müller-Eberhard, 1980), and it has been proposed that in the native haemolytically active molecule there exists a thiol-ester bond that is broken during activation, resulting in the generation of a reactive acyl group, involved in the covalent binding reaction, and the free thiol group (Tack *et al.*, 1980; Sim *et al.*, 1981). Studies have shown that the amino acid labelled with methylamine is a glutamic acid residue (Howard, 1980; Tack *et al.*, 1980) and that this residue is only three residues away from the free thiol group in the primary sequence (Tack *et al.*, 1980).

Another protein that has many similarities to C4 and C3 is the plasma proteinase inhibitor $\alpha_2$-macroglobulin. This protein also appears to be able to form covalent bonds with proteins (Harpel & Hayes, 1979; Salvesen & Barrett, 1980) and to incorporate methylamine covalently by amide linkage to the side-chain carboxyl group of a glutamic acid residue (Swenson & Howard, 1979) released probably from a thiol ester (Sottrup-Jensen *et al.*, 1980, 1981; Salvesen *et al.*, 1981).

The present paper describes further studies in an attempt to elucidate the nature of the reactive group in C4. The amino acid sequence around the putrescine-binding site and the free thiol group has been determined and a comparison is made with the corresponding sequences from C3 and $\alpha_2$-macroglobulin.

## Materials and methods

### *Materials*

*Plasma and chemicals.* Fresh frozen human plasma was obtained from the Churchill Hospital, Oxford. It was made 20mM with $CaCl_2$ and left to clot overnight at 4°C. The clot was removed by centrifugation and filtration through muslin, and the serum was stored at −20°C.

Chemicals were obtained as follows: cyanogen bromide, tributylphosphine, Aldrich Chemical Co., Gillingham, Dorset, U.K.; putrescine dihydrochloride, $Nbs_2$, Sigma Chemical Co., Poole, Dorset, U.K.; [1,4-$^{14}$C]putrescine dihydrochloride (116Ci/mol), iodo[2-$^{14}$C]acetic acid (54Ci/mol) and iodo[1-$^{14}$C]-acetamide (53Ci/mol), The Radiochemical Centre, Amersham, Bucks, U.K.; trypsin [Tos-Phe-$CH_2Cl$ ('TPCK')-treated], Worthington Biochemical Corp., Freehold, NJ, U.S.A. Sources of other reagents were as described previously [Campbell *et al.*, 1980; the preceding paper (Press & Gagnon, 1981)].

*Complement components.* Human C4 was prepared by the sequential chromatography of the pseudoglobulin fraction of human serum through DEAE-Sephadex A-50, Sepharose 6B and DEAE-Sepharose CL-6B as described previously (Gigli *et al.*, 1977; Campbell *et al.*, 1980). The C4 pool from the DEAE-Sephadex A-50 column of the C3 purification procedure described in Sim *et al.* (1981) was also used as a source of this component and was kindly donated by Dr. R. B. Sim and Miss C. Parkes. The C4, which had usually been converted into C4b, was purified from this pool by passage through columns of DEAE-Sepharose CL-6B and Sephacryl S-200 (superfine grade) and was used to prepare C4d. C4d and C4-binding protein were prepared as described in the preceding paper by Press & Gagnon (1981). C3bINA was prepared from human plasma as described by Crossley & Porter (1980) and was provided by Miss E. M. Press. $C\bar{1}s$ was prepared from $C\bar{1}$ as described by Gigli *et al.* (1976).

### *Methods*

*Labelling of the reactive acyl group in C4 with [1,4-$^{14}$C]putrescine.* Before use, the [1,4-$^{14}$C]-putrescine [250μCi in 5ml of 2% (v/v) ethanol] was freeze dried and dissolved in 250μl of 0.1M-NaCl/0.02%-$NaN_3$/0.002M-$CaCl_2$/0.05M-Tris/HCl, pH7.4.

The reactive acyl group in C4 was labelled with [1,4-$^{14}$C]putrescine by incubating C4 with $C\bar{1}s$ in the presence of 10mM-[1,4-$^{14}$C]putrescine (25Ci/mol) as described by Campbell *et al.* (1980). The position of the label in C4b was determined by fluorographic analysis of samples on SDS/polyacrylamide slab gels.

For maximum radioactive incorporation of [1,4-$^{14}$C]putrescine into C4b used in sequence studies, C4 was incubated with $C\bar{1}s$ in the presence of undiluted [1,4-$^{14}$C]putrescine (116Ci/mol) as follows: 150μCi of [1,4-$^{14}$C]putrescine was added to 45mg of C4 in 26ml of 100mM-NaCl/0.02%-$NaN_3$/2mM-$CaCl_2$/0.01M-Tris/HCl, pH7.4, such that the final concentration of [1,4-$^{14}$C]putrescine was 0.049mM. $C\bar{1}s$ (1:40 weight ratio of $C\bar{1}s$ over C4) was added and the sample was incubated at 37°C for 30min. $iPr_2P$-F was then added at 4°C to a concentration of 2.5mM to inactivate $C\bar{1}s$. The free thiol group generated during cleavage of C4 by $C\bar{1}s$ was

alkylated by incubating the C4b with a 1000-fold molar excess of iodoacetic acid at 37°C for 30min. The [1,4-$^{14}$C]putrescine-labelled C4b was dialysed against four 1 litre changes of 0.4M-NaCl/0.02M-putrescine/0.02%-$NaN_3$/0.005M-EDTA/0.05M-Tris/HCl, pH7.4, at 4°C for 48h, followed by two 1 litre changes of 0.15M-NaCl/0.02%-$NaN_3$/0.02M-Tris/HCl, pH7.0, at 4°C for 24h. This change in buffer was necessary to dialyse the C4b into the buffer used in the C3bINA cleavage.

*Labelling of the free thiol group in C4b with iodo[2-$^{14}$C]acetic acid.* The free thiol group in samples of C4b subsequently used to prepare C4d for sequence studies was labelled with iodo[2-$^{14}$C]acetic acid as described by Press & Gagnon (1981) in the preceding paper.

*Preparation of C4d from C4b.* [1,4-$^{14}$C]-Putrescine-labelled or iodo[2-$^{14}$C]acetic acid-labelled C4b was digested with C3bINA in the presence of C4-binding protein, and the C4d thus produced was purified on a column of Sephacryl S-200 (superfine grade) as described by Press & Gagnon (1981) in the preceding paper.

*Reduction and alkylation of C4d.* Freeze-dried C4d (90–180nmol) was dissolved in 6ml of 6M-guanidine hydrochloride/5mM-EDTA/0.4M-Tris/HCl, pH8.6. The flask was flushed with $N_2$, dithiothreitol (Calbiochem) was added to a concentration of 20mM, the flask was again flushed with $N_2$ and the sample was incubated at 37°C for 3h with frequent mixing. Iodoacetic acid was added to a final concentration of 44mM, the flask was flushed with $N_2$, and the sample was incubated on ice for 45min. The reduced and alkylated sample was then dialysed at 4°C against five 1 litre changes of 50mM-$NH_3$ for 36h and freeze-dried.

*CNBr cleavage of C4d and purification of the peptides.* Reduced and alkylated C4d (90–170nmol) was dissolved in 2ml of 70% (v/v) formic acid containing 100mg of CNBr and incubated at 4°C in the dark for 24h. The mixture was then diluted 1:10 with ice-cold water and freeze-dried. The digest was redissolved in 1ml of 1% (w/v) $NH_4HCO_3$, pH8.1, and applied to a column (1.6cm × 93cm) of Sephadex G-50 (superfine grade) equilibrated in 1% (w/v) $NH_4HCO_3$, pH8.1. Peptides were identified by their absorbance at 280nm and 220nm and by measurement of radioactivity. The peak of radioactivity, pool C (containing the CNBr peptide designated 'CN-4'), was freeze-dried, redissolved in 0.1% (w/v) $NH_4HCO_3$ and purified by h.p.l.c. using a μBondapak C-18 column equilibrated at room temperature in 0.1% $NH_4HCO_3$/acetonitrile in the ratio 17:3 (v/v). The column was pumped at 1ml/min and eluted with a linear 1h gradient to give a final ratio of 0.1% $NH_4HCO_3$/acetonitrile of 9:11 (v/v). The CNBr peptide CN-4 in pool B was also purified by h.p.l.c. under identical conditions. CNBr peptides CN-5 and CN-7 were purified from pools D and E respectively in a similar manner, except that the column was equilibrated in 0.1% (w/v) $NH_4HCO_3$/acetonitrile in the ratio 19:1 (v/v) and eluted with a linear 1h gradient to give a final ratio of 0.1% (w/v) $NH_4HCO_3$/acetonitrile of 9:11 (v/v).

*Purification of the peptides derived from tryptic digestion of CNBr peptide CN-4.* Peptide CN-4 (57nmol) was dissolved in 100μl of 0.5% (w/v) $NH_4HCO_3$, pH8.3, 5μl of a 0.6mg/ml solution of Tos-Phe-$CH_2$Cl-treated trypsin (Worthington Biochemicals) was added such that the enzyme/substrate ratio on a molar basis was 1:375, and the sample was incubated at 37°C for 2h. The incubation was repeated for a further 2h after the addition of the same amount of enzyme; then 300μl of water was added and the peptides were purified by h.p.l.c. as described for CN-5 and CN-7.

*Acid cleavage of CN-4 and automated sequence analysis of the product CN-4AC.* CN-4 (15nmol) was resuspended in 200μl of 50% (v/v) formic acid, incubated at 37°C for 72h, then freeze-dried. The sample was dissolved in 200μl of *N*-methylmorpholine/water (1:1) containing 0.2% β-mercaptoethanol and applied to the sequencer cup. *O*-Phthalaldehyde (200μl of a 0.3mg/ml solution in water) was added to block any primary amine (Mendez & Marco, 1980) leaving only the proline imino group available for sequencing. The mixture was incubated for 5min under $N_2$ at 57°C. The sequencer was then started after a wash with 3ml of butyl chloride.

*Trypsin digestion of C4d and purification of the [1,4-$^{14}$C]putrescine-labelled peptide T-5P14.* Reduced and alkylated [1,4-$^{14}$C]putrescine-labelled C4d (60nmol) was resuspended in 1ml of 1% (w/v) $NH_4HCO_3$, pH8.3, and incubated with Tos-Phe-$CH_2$Cl-treated trypsin (enzyme/substrate ratio 1:200, w/w) at 37°C for 2h. After the addition of the same amount of enzyme and incubation for a further 2h, the trypsin was inactivated with 1mM-$iPr_2P$-F and the digest was fractionated on a column (1.6cm × 91cm) of Sephadex G-50 (superfine grade) equilibrated in 1% (w/v) $NH_4HCO_3$. The fractions containing the labelled peptide (T-5P14) were pooled, freeze-dried, redissolved in 400μl of 0.1% (w/v) $NH_4HCO_3$ and the peptide was purified by h.p.l.c. under the same conditions described for the tryptic peptides of CN-4.

*Amino acid analysis.* Samples were hydrolysed under vacuum at 110°C for 24h in constant-boiling HCl containing 4mM-phenol. No corrections were made for destruction of threonine and serine, and tryptophan was not determined. Cysteine and cystine were determined as *S*-carboxymethylcysteine after acid hydrolysis under vacuum of alkylated or reduced and alkylated samples in the presence of 0.05% 2-mercaptoethanol.

*Automated amino-acid-sequence determination.* Automated amino-acid-sequence determination was performed as described by Johnson *et al.* (1980) in a Beckman 890c sequencer. The thiazolinones released were extracted with butyl chloride containing 10mg of dithioerythritol and 100$\mu$l of tributylphosphine per litre (Frank, 1979). At each cycle the extract was collected under $N_2$ in a fraction-collector tube containing 200$\mu$l of 0.1M-HCl/0.1% ethanethiol (prepared fresh daily). The thiazolinones were then converted into phenylthiohydantoins as soon as possible. By using these precautions we were able to obtain serine and threonine residues in 20–30% yield.

The phenylthiohydantoin derivatives were identified by h.p.l.c. as previously described (Johnson *et al.*, 1980). With high-quality reagents and solvents we were able to detect as little as 0.1nmol of the derivatives.

*SDS/polyacrylamide-gel electrophoresis*

Protein samples to be examined by fluorography were run on 12% (w/v) polyacrylamide slab gels containing 0.06% bisacrylamide and 0.1% SDS, by using the discontinuous buffer system described by Laemmli (1970). Reduction and alkylation of protein samples was performed as described by Gigli *et al.* (1976). The gels were stained for protein with Coomassie Blue, then treated with dimethyl sulphoxide and PPO as described by Bonner & Laskey (1974). The gels were dried on Whatman 3MM chromatography paper then laid on a sheet of X-ray film (Fuji Rx Medical X-ray film) that had been pre-exposed to a brief flash of light (Laskey & Mills, 1975), and left in the dark at −70°C. Fluorographs were developed with Kodak DX-80 developer and Kodak FX-40 fixer.

*Determination of the thiol content of C4 and C4b samples.* Determinations with $Nbs_2$ were performed as described by Janatova *et al.* (1980*b*). To 1ml of a C4 solution (0.8–1.6mg of protein) in 5mM-EDTA/50mM-sodium phosphate, pH7, was added 100$\mu$l of 10mM-$Nbs_2$ and the reaction was monitored by measuring $A_{412}$. The thiol content was calculated by using an absorption coefficient of $13\,600\,M^{-1}\cdot cm^{-1}$ for $Nbs^{2-}$ (Ellman, 1959).

To determine the thiol content with iodo[1-$^{14}$C]acetamide, samples of C4 and C4b in 5mM-EDTA/200mM-Tris/HCl, pH8.0, were incubated with a 20-fold molar excess of iodo[1-$^{14}$C]acet-

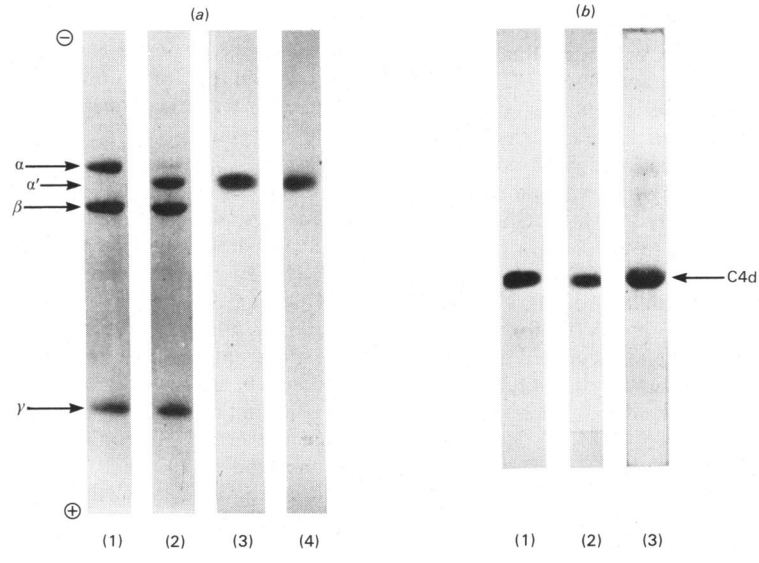


Fig. 1. *SDS/polyacrylamide-gel electrophoresis of samples of C4b showing the location of the [1,4-$^{14}$C]putrescine-reactive site and the free thiol group generated during cleavage of C4 by C1s*

(*a*) Samples of [1,4-$^{14}$C]putrescine-labelled C4b or iodo[1-$^{14}$C]acetamide-labelled C4b were prepared as described in the text and run under reducing conditions on SDS/12% (w/v)-polyacrylamide slab gels by the method of Laemmli (1970). Protein was stained with Coomassie Blue and fluorographs were prepared as described in the Materials and methods section. Samples shown are: (1) C4 (5$\mu$g) stained for protein; (2) C4b (5$\mu$g) stained for protein; (3) fluorograph of 20$\mu$g of [1,4-$^{14}$C]putrescine-labelled C4b; (4) fluorograph of 20$\mu$g of iodo[1-$^{14}$C]acetamide-labelled C4b. (*b*) Samples (15$\mu$g) of [1,4-$^{14}$C]putrescine-labelled C4b or iodo[1-$^{14}$C]acetamide-labelled C4b were incubated with 15$\mu$g of C4-binding protein and 0.3$\mu$g of C3bINA at 37°C for 3h. The digest was then run under reducing conditions on SDS/12% (w/v)-polyacrylamide slab gels by the method of Laemmli (1970). Samples shown are: (1) fluorograph of a C3bINA digest of [1,4-$^{14}$C]putrescine-labelled C4b; (2) fluorograph of a C3bINA digest of iodo[1-$^{14}$C]acetamide-labelled C4b; (3) purified C4d (15$\mu$g) run under the same conditions and stained for protein.

amide (53 Ci/mol) at 37°C for 1h. The radiolabelled sample was exhaustively dialysed against 0.15 M-NaCl/0.1 M-sodium phosphate buffer, pH 7.4, and the degree of incorporation was determined by liquid-scintillation counting as described previously (Campbell *et al.*, 1980).

## Results

### *Location of the putrescine-reactive site and the free thiol group generated during cleavage of C4 by $\bar{C}1s$*

In a previous publication (Campbell *et al.*, 1980) it was shown that activation of C4 by C1s in the presence of 10 mM-[1,4-$^{14}$C]putrescine results in the covalent incorporation of approx. 0.6 mol of putrescine/mol of C4. The position of the putrescine label in C4b was determined by fluorographic analysis of SDS/polyacrylamide gels. All of the label was found to be associated with the α′-chain, with none in the β- or γ-chains (Fig. 1). If the labelled C4b was further digested with C3bINA in the presence of C4-binding protein, all of the label associated with the α′-chain was now located in the C4d fragment (Fig. 1). No label was found to be associated with the other α′-fragments.

During cleavage of C4 by $\bar{C}1s$, a free thiol group is generated that cannot be detected in the native haemolytically active molecule. By using $Nbs_2$ or iodo[1-$^{14}$C]acetamide, 0.91 or 0.81 mol of thiol can be detected respectively per mol of C4b. Only 0.06–0.11 mol of thiol could be detected per mol of native C4. Fluorographic analysis of iodo[1-$^{14}$C]-acetamide-labelled C4b shows that the label is covalently bound to the α′-chain (Fig. 1), though if putrescine is not present during cleavage by $\bar{C}1s$, some label can also be detected in higher-molecular-weight bands, which may be dimers and higher aggregates of the α′-chain (Campbell *et al.*, 1980). Cleavage of C4b with C3bINA in the presence of C4-binding protein shows that the label associated with the α′-chain is in the C4d fragment (Fig. 1) [the preceding paper (Press & Gagnon, 1981)].

### *Purification of C4d, CNBr digestion and purification of the labelled peptide*

For maximum radioactive incorporation of [1,4-$^{14}$C]putrescine into C4b, C4 was incubated with $\bar{C}1s$ in the presence of undiluted [1,4-$^{14}$C]putrescine (final concn. 0.049 mM, 116 Ci/mol) at 37°C for 30 min, then dialysed as described in the Materials and methods section to remove unbound radioactivity. At this low concentration, but high specific radioactivity, of [1,4-$^{14}$C]putrescine, approx. 0.04–0.06 mol of the diamine were bound per mol of C4b, and the radioactivity in the sample ranged from 40000 to 68000 c.p.m. per mg of protein.

C4d, prepared from radioactively labelled C4b as described in the preceding paper (Press & Gagnon, 1981), was digested with CNBr and the digest was fractionated on a column of Sephadex G-50 (superfine grade) equilibrated in 1% (w/v) $NH_4HCO_3$. An elution profile is shown in Fig. 2. About 45 and 36% of the radioactivity was found to elute with peptides of apparent mol.wts. 8000 and 13000 respectively, with the remainder of the radioactivity (19%) eluting in the void volume. However, if the CNBr digest was run on a Sephadex G-75 column equilibrated in 10% (v/v) formic acid, 80% of the radioactivity was found to elute with a peptide of approx. mol.wt. 4700, suggesting that gel filtration on Sephadex G-50 equilibrated in 1% (w/v) $NH_4HCO_3$ leads to aggregation of the labelled peptide. Fractionation of the CNBr digest of C4d on the Sephadex G-50 column equilibrated in 1% (w/v) $NH_4HCO_3$ was preferred, since acid degradation of the peptide in 10% (v/v) formic acid was frequently observed, suggesting the presence of an acid-labile Asp–Pro

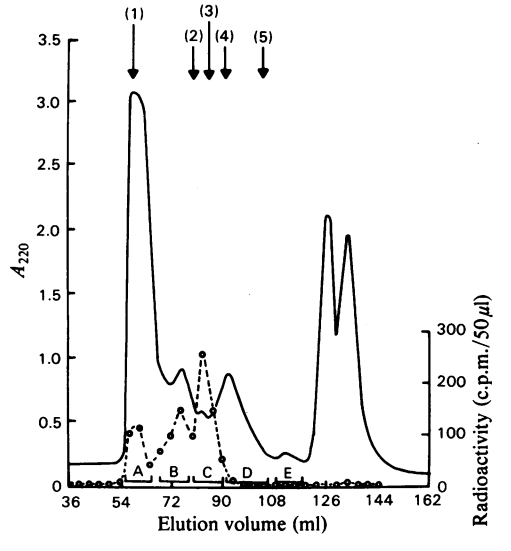


Fig. 2. *Purification on Sephadex G-50 of peptides obtained by CNBr digestion of C4d*

The freeze-dried CNBr digest of C4d was redissolved in 1 ml of 1% (w/v) $NH_4HCO_3$ and applied to a column (1.6 cm × 93 cm) of Sephadex G-50 (superfine grade) equilibrated in 1% (w/v) $NH_4HCO_3$ and run at a flow rate of 3 ml/h. ——, $A_{220}$; ----, radioactivity of *S*-[$^{14}$C]carboxymethylcysteine. The arrows indicate the elution positions of the following marker proteins: (1) reduced and alkylated bovine serum albumin (mol.wt. 67000); (2) reduced and alkylated cytochrome *c* (mol.wt. 12400); (3) and (5) peptides CN-2 and CN-3 respectively from a CNBr digest of reduced and alkylated horse myoglobin (mol.wts. 6200 and 2500); (4) unreduced insulin (mol.wt. 5800).

bond, and this has been confirmed by sequence studies (see Fig. 5).

The labelled CNBr peptide, designated 'CN-4', from pool C of the Sephadex G-50 column was further purified by h.p.l.c. using a μBondapak C-18 column. Multiple peaks were obtained that were eluted 20–26 min after the start of a linear gradient, but with different proportions of label associated with each peak (Fig. 3). Amino acid composition and *N*-terminal-sequence analysis have shown that these peptides are identical. Why multiple peaks are obtained is not clear, but it may be due to a homoserine↔lactone interconversion or to varying extents of deamination of glutamine and asparagine residues. Peptide CN-4 from pool B of the Sephadex G-50 column was also purified by h.p.l.c. and elutes in a similar position to that from pool C.

The lower-molecular-weight CNBr peptides of C4d, peptide CN-5 (mol.wt. 3900) and peptide CN-7 (mol.wt. 1400) were also purified by h.p.l.c. from pools D and E of the Sephadex G-50 column respectively.

*Amino acid compositions, and N-terminal amino acid sequences of the CNBr peptides CN-4, CN-5 and CN-7*

The amino acid compositions of peptides CN-4, CN-5, and CN-7 are shown in Table 1.

The *N*-terminal amino acid sequence of peptide CN-4 was found to overlap with the *N*-terminal amino acid sequence of C4d established by Press & Gagnon (1981) in the preceding paper, and this positions peptide CN-4 with respect to the *N*-terminus of C4d (Fig. 5). The amino acid composition of peptide CN-7 suggests that this peptide corresponds to the *N*-terminus of C4d, and this has been confirmed by sequence studies. In all, 26 residues were identified unambiguously in peptide CN-4 (Fig. 5). If the peptide was labelled with [1,4-$^{14}$C]putrescine, the radioactivity was found to elute from the sequencer at cycle 45, whereas if iodo[2-$^{14}$C]acetic acid was the label, the radioactivity was found to elute at cycle 42.

The *N*-terminal sequence of peptide CN-5 was also determined (Fig. 5), as evidence, presented below, suggested that this peptide was *C*-terminal to peptide CN-4 in C4d.

*Enzymic and chemical cleavage of peptide CN-4 and amino acid sequences of the purified fragments*

As the amino acid composition of peptide CN-4 (Table 1) suggested the presence of three arginine residues and no lysine residues, tryptic digestion should produce four peptides. A portion (57 nmol) of peptide CN-4 was digested with trypsin as des-

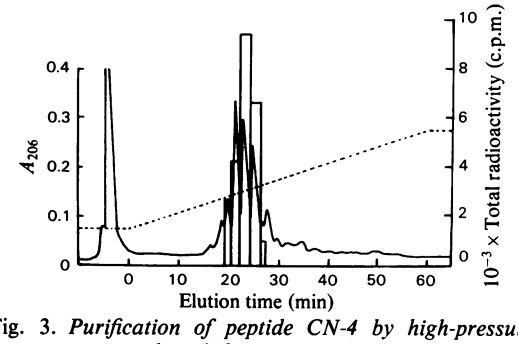


Fig. 3. *Purification of peptide CN-4 by high-pressure liquid chromatography*

Pool C, containing peptide CN-4, from the Sephadex G-50 column was freeze-dried, redissolved in 400 μl of 0.1% (w/v) $NH_4HCO_3$ and loaded on to a μBondapak C-18 column pumped at 1 ml/min and equilibrated at room temperature in 0.1% (w/v) $NH_4HCO_3$/acetonitrile (17:3, v/v). After the breakthrough peak had eluted, a linear 1 h 15–55% acetonitrile gradient (----) was applied and individual peaks were collected. The total radioactivity (open bars) in each peak was determined by counting 25 μl portions. ——, $A_{206}$.

Table 1. *Amino acid compositions of CNBr peptides CN-4, CN-5, CN-4/5 and CN-7, and the tryptic peptide T-5P14*

The amino acid compositions are given as mol of residue/mol of peptide. The samples were hydrolysed at 110°C for 24 h. No corrections were made for serine and threonine destruction, and tryptophan was not determined. Cysteine was determined as *S*-carboxymethylcysteine. Numbers in parentheses correspond to the integral number of residues determined from the composition.

| Amino acid | Amino acid composition (mol/mol) of: CN-4 | CN-5 | CN-7 | T-5P14 | CN-4/5 |
|---|---|---|---|---|---|
| Asp | 5.26 (5) | 2.92 (3) | 3.30 (3) | 0.76 (1) | 8.21 (8) |
| Thr | 2.80 (3) | 3.24 (3) | 1.05 (1) | 1.93 (2) | 5.50 (6) |
| Ser | 3.94 (4) | 1.74 (2) | 0.90 (1) | 1.30 (1) | 5.83 (6) |
| Hse | 0.85 (1) | 0.68 (1) | 1.15 (1) | | 1.07 (1) |
| Glu | 3.43 (3) | 3.58 (4) | 1.15 (1) | 2.40 (2) | 7.83 (8) |
| Pro | 2.85 (3) | 2.14 (2) | 2.19 (2) | 0.94 (1) | 4.87 (5) |
| Gly | 7.22 (7) | 2.38 (2) | 1.21 (1) | 1.99 (2) | 9.28 (9) |
| Ala | 3.28 (3) | 4.17 (4) | | 2.77 (3) | 6.64 (7) |
| Val | 3.11 (3) | 1.64 (2) | | 0.42 (0) | 5.22 (5) |
| Cys | 0.82 (1) | | | 0.60 (1) | 0.87 (1) |
| Met | | | | 0.75 (1) | 0.30 (1) |
| Ile | 1.07 (1) | 1.60 (2) | 1.15 (1) | 1.02 (1) | 2.65 (3) |
| Leu | 5.99 (6) | 4.37 (4) | 1.16 (1) | 2.39 (2) | 9.00 (9) |
| Tyr | 0.98 (1) | 2.60 (3) | | 0.88 (1) | 3.17 (3) |
| Phe | 0.79 (1) | 0.19 (0) | | 0.23 (0) | 2.25 (2) |
| His | | 0.90 (1) | | 0.18 (0) | 1.05 (1) |
| Lys | 0.36 (0) | 2.61 (3) | 0.15 (0) | 0.37 (0) | 2.79 (3) |
| Arg | 2.94 (3) | 1.07 (1) | | 1.00 (1) | 3.67 (4) |

cribed in the Materials and methods section and the peptides were purified by h.p.l.c. with a μBondapak C-18 column (Fig. 4). Most of the radioactivity was found to be associated with the first peak to be eluted (peptide CN-4T1) and the remainder with the peak eluting immediately afterwards. Amino acid analyses of the tryptic peptides are shown in Table 2. Sequence analysis of the labelled peptide CN-4T1 revealed the sequence shown in Fig. 5. Since this is the only tryptic peptide that contains homoserine, this peptide must represent the *C*-terminal tryptic peptide of peptide CN-4. All of the residues shown by analysis to be present in the peptide have been sequenced, though it was not possible to detect the homoserine residue. If the peptide was labelled with iodo[2-$^{14}$C]acetic acid, the radioactivity was found at cycle 2, whereas if the peptide was labelled with [1,4-$^{14}$C]putrescine, the radioactivity was found at cycle 5. Further evidence that the glutamic acid at position 5 of CN-4T1 is the reactive acyl group generated during cleavage of C4 by C1s is derived from sequence analysis of material labelled with 6-aminohexanoic acid. An unknown amino acid phenylthiohydantoin derivative was found by h.p.l.c. at this position, which on back hydrolysis gave equimolar amounts of 6-aminohexanoic acid and glutamic acid.

Peak 4 (peptide CN-4T3) has the expected composition of the *N*-terminal tryptic peptide of peptide CN-4 from isoleucine at position 1 to arginine at position 11. Peak 5 (peptide CN-4T4) has the composition of the tryptic peptide that would be expected to start at residue 12 (valine) of peptide CN-4, and this has been confirmed by sequence analysis (Fig. 5). It was not possible to sequence this peptide to the *C*-terminal arginine residue, since no sequence information was obtained after the proline residue at cycle 18. Comparison of the amino acids identified in the sequence with those present in the composition of the peptide reveals that Ser, $Gly_2$, Ala, Val, $Leu_2$ and Arg residues have not been accounted for. The presence of Val and Leu residues suggests that there may be a hydrophobic sequence after the proline residue at position 18, which may have resulted in the peptide being washed from the sequencer cup.

The peptide in peak 2 has the composition of peptide CN-4T1 plus peptide CN-4T2, suggesting that this peptide is the result of a partial tryptic digestion of peptide CN-4. *N*-Terminal sequence

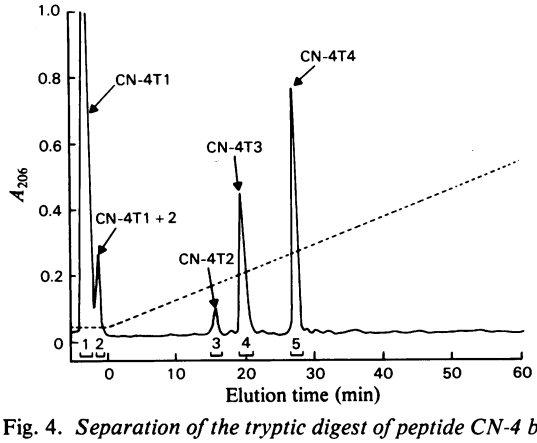


Fig. 4. *Separation of the tryptic digest of peptide CN-4 by high-pressure liquid chromatography*

Peptide CN-4 (57nmol) was dissolved in 100μl of 0.5% (w/v) $NH_4HCO_3$, pH8.3, and incubated at 37°C for 2h with trypsin (enzyme/substrate molar ratio 1:375). The incubation was repeated for a further 2h with an additional amount of enzyme, then 300μl of water was added and the digest was fractionated on a μBondapak C-18 column pumped at 1ml/min and equilibrated at room temperature in 0.1% (w/v) $NH_4HCO_3$/acetonitrile (19:1, v/v). After the breakthrough peaks had eluted, a linear 1h gradient of 5–55% acetonitrile (----) was applied and individual peaks were collected. ——, $A_{206}$.

Table 2. *Amino acid compositions of the tryptic peptides from CNBr peptide CN-4*

The amino acid compositions are given as mol of residue/mol of peptide. The samples were hydrolysed at 110°C for 24h. No corrections were made for serine and threonine destruction, and tryptophan was not determined. Cysteine was determined as *S*-carboxymethylcysteine. Numbers in parentheses correspond to the integral number of residues determined from the composition.

| Amino acid | Amino acid composition (mol/mol) of: CN-4T1 | CN-4T2 | CN-4T3 | CN-4T4 |
|---|---|---|---|---|
| Asp | 0.13 (0) | 0.33 (0) | 3.34 (3) | 2.21 (2) |
| Thr | 1.04 (1) | 0.20 (0) | 0.28 (0) | 1.82 (2) |
| Ser | 0.11 (0) | 0.19 (0) | 1.03 (1) | 3.31 (3) |
| Hse | 1.00 (1) | | | |
| Glu | 2.26 (2) | 0.08 (0) | 0.33 (0) | 1.12 (1) |
| Pro | | 0.81 (1) | 1.11 (1) | 1.75 (2) |
| Gly | 2.26 (2) | 0.14 (0) | 1.34 (1) | 4.28 (4) |
| Ala | | 0.12 (0) | | 3.31 (3) |
| Val | 0.16 (0) | 0.32 (0) | 1.13 (1) | 2.35 (2) |
| Cys | 0.72 (1) | | 0.06 (0) | |
| Met | | | | |
| Ile | | 0.16 (0) | 1.15 (1) | |
| Leu | | 0.99 (1) | 0.27 (0) | 5.30 (5) |
| Tyr | | 0.14 (0) | 1.00 (1) | |
| Phe | | 0.11 (0) | 0.89 (1) | |
| His | | | | |
| Lys | 0.13 (0) | 0.16 (0) | | |
| Arg | | 1.00 (1) | 1.00 (1) | 1.00 (1) |

| Peptide | Amount used in sequencer (nmol) | Recovery at first step (nmol) | Stepwise yield (%) |
|---|---|---|---|
| CN-4 | 10.8 | 3.5 | 90 |
| CN-5 | 42 | 25 | 95 |
| CN-4T1 | 34 | 24 | 90 |
| CN-4T4 | 37 | 14 | 90 |
| CN-4AC | 15 | 1.7 | 92 |
| T-5P14 | 13 | 2.5 | 90 |

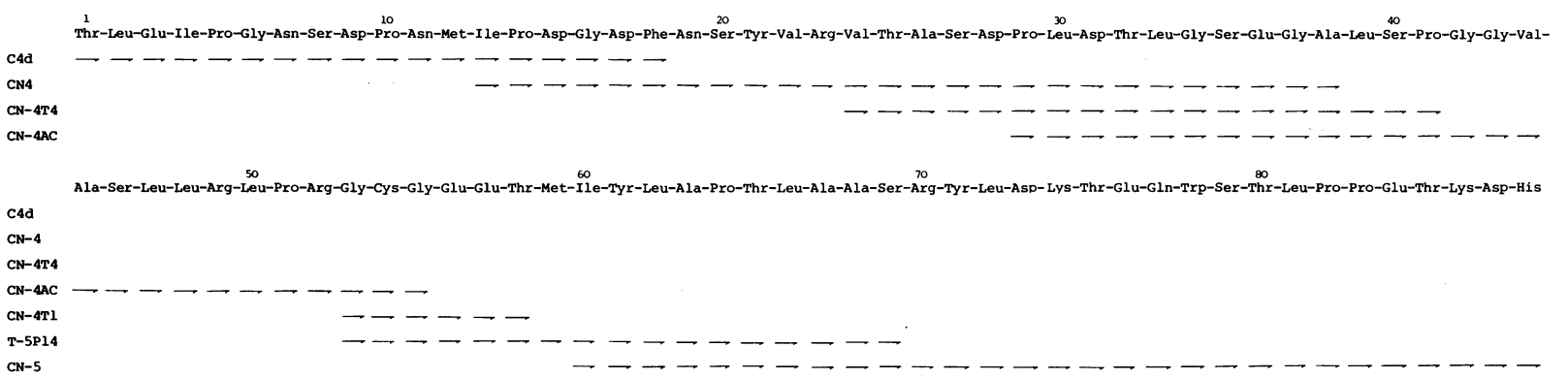


Fig. 5. *Automated sequence analysis and alignment of CNBr peptides CN-4 and CN-5, the tryptic and acid-cleavage peptides of peptide CN-4, and the tryptic peptide T-5P14*

Details of the automated Edman degradation are given in the text. Amino acid phenylthiohydantoin derivatives were identified by h.p.l.c. The detection of the *S*-carboxymethylcysteine phenylthiohydantoin derivatives were also confirmed by detection of radioactive label. — Denotes that positive identification of a particular residue was made. The *N*-terminal amino acid sequence of C4d is taken from the preceding paper (Press & Gagnon, 1981).

analysis of 2.8nmol of this peptide gave the sequence Leu-Pro-Xaa-Gly-Xaa-Gly, thus suggesting the alignment of peptides CN-4T1 and CN-4T2 in the order:

CN-4T2 — CN-4T1

In order to obtain the *C*-terminal sequence of the tryptic peptide CN-4T4 and to confirm the alignment of peptide CN-4T4, peptide CN-4T2 and peptide CN-4T1, the CNBr peptide CN-4 was incubated with 50% (v/v) formic acid at 37°C for 72h to acid-cleave the Asp–Pro bond between residues 16 and 17 and to generate a new peptide with *N*-terminal proline. In order to avoid purification of the cleavage products, the sample was freeze-dried, then incubated with *o*-phthalaldehyde as described in the Materials and methods section to block the *N*-terminal isoleucine residue of peptide CN-4 and any other primary amine (Mendez & Marco, 1980). *o*-Phthalaldehyde was very effective at blocking any residual *N*-terminal amino acid, with the exception of proline, and thus it was possible to sequence the acid-cleavage product, peptide CN-4AC. Although a low yield was obtained, owing to the presence of residual reagents in the first three cycles of the sequencer run, only one sequence was obtained, and with a very clean background after the third cycle. The sequence obtained is shown in Fig. 5 and confirms the alignment of the tryptic peptides of peptide CN-4 in the order:

CN-4T4 — CN-4T2 — CN-4T1

The sequence of the acid-cleavage peptide also completes the sequence of the CNBr peptide CN-4.

All the amino acids in the composition of CN-4 are accounted for in the sequence. Only the values for serine and proline are higher, by one amino acid, than those expected from the composition.

*Evidence for the overlap of peptides CN-4 and CN-5*

On the occasions when the CNBr digest of [1,4-$^{14}$C]putrescine-labelled or iodo[2-$^{14}$C]acetic acid-labelled C4d was fractionated on Sephadex G-75 equilibrated in 10% (v/v) formic acid, a peak of radioactivity, corresponding to approx. 15% of that applied, was found to elute with a peptide of mol.wt. 8500, suggesting that this peptide was a partial-digest product containing peptide CN-4 and another of the CNBr peptides of C4d. On a molecular-weight basis it is probable that the partial-digest product is composed of peptide CN-4 and peptide CN-5 (mol.wts. 4700 + 3900 = 8600), and the amino acid composition of the purified peptide CN-4/5 (Table 1) is very similar to the composition of peptides CN-4 + CN-5.

In order to confirm the alignment of peptide CN-4 and peptide CN-5, 60nmol of reduced and alkylated[1,4-$^{14}$C]putrescine-labelled C4d was incubated with trypsin and the digest was fractionated on a Sephadex G-50 (superfine grade) column equilibrated in 1% (w/v) $NH_4HCO_3$. The radioactive peptide, which elutes with an approx. mol.wt. of 2500, was purified by h.p.l.c. on a μBondapak C-18 column. The peptide (T-5P14) was found to elute approx. 29min after the start of a linear 5–55% (v/v) acetonitrile gradient. The amino acid composition of T-5P14 is shown in Table 1. A 13nmol portion was subjected to automated sequence analysis and the sequence obtained, shown in Fig. 5, overlaps the *C*-terminal sequence of peptide CN-4 and the *N*-terminal sequence of peptide CN-5 and thus aligns these peptides in the order:

CN-4 — CN-5.

## Discussion

When C4 is activated by $C\bar{1}$, a reactive acyl group is generated that allows the molecule to covalently bind to the Fd region of the heavy chain of IgG (Campbell *et al.*, 1980) or to surface structures on erythrocyte membranes (Law *et al.*, 1980). Incubation of C4 with $C\bar{1}$s in the presence of [1,4-$^{14}$C]putrescine, which can inhibit the binding of C4 to antibody–antigen aggregates (Campbell *et al.*, 1980), leads to the covalent incorporation of the diamine. Fluorographic analysis of [1,4-$^{14}$C]-putrescine-labelled C4b (Fig. 1) shows that all the radioactivity is associated with the α′-chain. Concomitant with the production of the reactive acyl group, a free thiol group is generated that cannot be detected in the native molecule, and again this group has been localized to the α′-chain (Fig. 1). The expression of a free thiol group on activation of C4 with $C\bar{1}$s has also been documented by Reboul *et al.* (1980) and Janatova & Tack (1981). Digestion of C4b with C3bINA in the presence of C4-binding protein shows that both the reactive acyl group and the free thiol group are located in C4d, a 44500-mol.wt. fragment shown by Press & Gagnon (1981), in the preceding paper, to occupy a central position in the α′-chain. Thus the structure of C4d is of considerable interest as it is that part of the C4 molecule that contains the potentially reactive acyl group involved in the covalent-binding reaction.

The amino acid sequence for the first 88 residues of C4d has been established by automated sequence analysis of fragments obtained by CNBr cleavage, dilute acid cleavage and enzymic digestion with trypsin. The sequence of the CNBr peptide that contains both the [1,4-$^{14}$C]putrescine and iodo-[2-$^{14}$C]acetic acid labels has been found to overlap with the *N*-terminal sequence of C4d obtained by Press & Gagnon (1981) in the preceding paper, and thus it has been possible to align this peptide with respect to the *N*-terminus of C4d (Fig. 5). The amino acid residue labelled with putrescine during acti-

vation of C4 by $C\bar{1}s$ is the glutamic acid residue at position 57 of C4d, and this residue is only three residues away from the free thiol group at position 54 also generated during activation. Since the order of the α′-chain fragments of C4b produced by cleavage with C3bINA is known [the preceding paper (Press & Gagnon, 1981)], it is possible to calculate that the free thiol group and the reactive acyl group are approx. 261 and 264 amino acids respectively from the *N*-terminus of the α′-chain. Thus the cleavage of a single peptide bond, which occurs during C4 activation, at approx. 80 residues from the *N*-terminus of the α-chain and 264 residues before the reactive acyl group, has a profound effect on this region of the molecule, resulting in the capacity to covalently bind to acceptor molecules. However, although the $C\bar{1}s$ cleavage site is distant from the reactive acyl group in the primary sequence, they may be much closer together in the tertiary structure.

C3 and $\alpha_2$-macroglobulin show many similarities to C4. All three proteins can form covalent bonds after a single peptide-bond cleavage, all three are inactivated by small amine nucleophiles, and, concomitant with either activation by proteolytic cleavage or inactivation by amines, a free thiol group is generated that cannot be detected in the native molecules. The present results show that C4, C3 and $\alpha_2$-macroglobulin have very similar amino acid sequences (Fig. 6), especially in the section containing the reactive acyl group and the free thiol group. The octapeptide sequence is identical except for the conservative replacement of valine in $\alpha_2$-macroglobulin for isoleucine at position 60 (C4d numbering) and threonine for asparagine in C4d at position 58. Thus all three proteins appear to share a common functional site that is involved in the covalent-binding reaction.

The simultaneous appearance in C3, on activation by proteolytic cleavage, of the thiol group and the reactive acyl group has given rise to the suggestion that in the native molecule these residues form a thiol-ester bond (Tack *et al.*, 1980; Sim *et al.*, 1981). A similar proposal has also been put forward for these residues in $\alpha_2$-macroglobulin (Sottrup-Jensen *et al.*, 1980; Salvesen *et al.*, 1981), and it is highly likely that the cysteine residue at position 54 and the glutamic acid residue at position 57 of C4d also exist as a thiol ester in the native molecule. A thiol-ester bond will not, in itself, form a new covalent bond by transesterification spontaneously at neutral pH unless activated by a unique mechanism such as that proposed by Davies & Sim (1981). These authors have suggested that the reactivity of the proteins is the result of a hydrogen-bonding intramolecular general-acid-catalysis mechanism (Davies & Sim, 1981). This will be dependent on associated structures in other sections of the

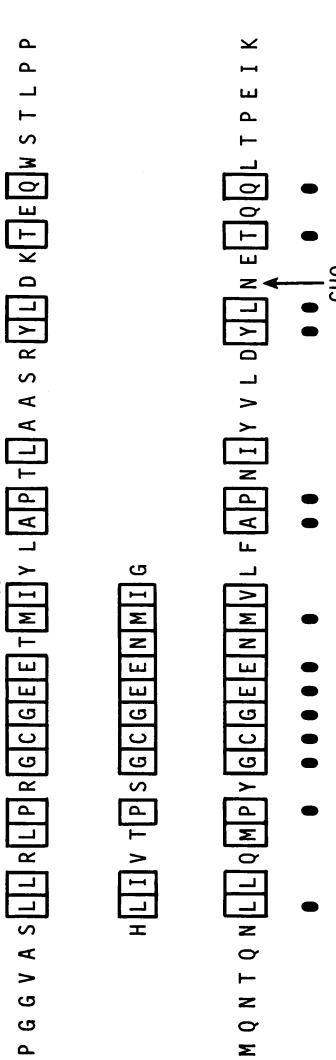


Fig. 6. *Comparison of the sequences of C4, C3 and $\alpha_2$-macroglobulin around the proposed thiolester bond*

Sequence data for C3 and $\alpha_2$-macroglobulin are from Tack *et al.* (1980) and Swenson & Howard (1980). The residue numbering shown is that of C4d. The cysteine and glutamic acid residues involved in the thiolester bond are marked with an asterisk (*). □ represents functionally conserved amino acids, designated by Dayhoff *et al.* (1972), from chemical similarity and accepted point-mutation data and consist of (A, P, G), (N, Q), (D, E), (S, T), (C), (V, I, M, L), (K, R, H) and (F, Y, W). Identical residues are marked thus: ●. The single-letter code for amino-acid residues is used: 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; Y, Tyr. CHO signifies a carbohydrate attachment site.

polypeptide chain for shielding of the H-bond activated thiol ester in the native proteins. It is evident from Fig. 6 that there are homologous sequences near the octapeptide containing the proposed intrachain thiol ester, and it is highly likely that they contribute to the protected environment of this bond. The two hydrophobic residues at positions 47–48 and the proline residue at position 51 are notable features. Although no sequence information on C3 *C*-terminal to the thiol-ester bond is yet available, comparison of the sequences of C4 and $\alpha_2$-macroglobulin in this region of the molecule shows many homologies. The alanine and proline residues at positions 63 and 64 are notable, as is the section between residues 71 and 77. This region shows four identities, though in $\alpha_2$-macroglobulin there is a polysaccharide attachment site to the asparagine residue, whereas in C4 the corresponding residue is aspartic acid and the adjacent residue is lysine, not glutamic acid. These may be important differences that might be expected to have a substantial effect on the behaviour of the thiol ester and might be responsible for the observed differences in the half-life of the reactive acyl group (Davies & Sim, 1981; Salvesen *et al.*, 1981; Sim *et al.*, 1981). However, it is unlikely that a definitive description of this novel structure can be made until the conformation of the adjacent peptide chains of the native or activated forms of one of these proteins has been obtained from X-ray crystallography.

We thank Miss E. M. Press and Dr. K. B. M. Reid, Dr. D. N. Chakravarti and Dr. R. B. Sim for helpful discussions, Mr. T. Gascoyne for performing the amino acid analyses, and Miss A. Moseby and Mr. A. C. Willis for excellent technical assistance. R. D. C. holds a Wolfson College Junior Research Fellowship.

## References

Bonner, W. M. & Laskey, R. A. (1974) *Eur. J. Biochem.* **46**, 83–88

Campbell, R. D., Dodds, A. W. & Porter, R. R. (1980) *Biochem. J.* **189**, 67–80

Crossley, L. G. & Porter, R. R. (1980) *Biochem. J.* **191**, 173–182

Davies, S. G. & Sim, R. B. (1981) *Biosci. Rep.* **1**, 461–468

Dayhoff, M. O., Eck, R. V. & Park, C. M. (1972) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., ed.), pp. 89–99, National Biomedical Research Foundation

Ellman, G. (1959) *Arch Biochem. Biophys.* **82**, 70–78

Frank, G. (1979) *Hoppe-Seyler's Z. Physiol. Chem.* **360**, 997–999

Fujita, T., Gigli, I. & Nussenzweig, V. (1978) *J. Exp. Med.* **148**, 1044–1051

Gadd, K. J. & Reid, K. B. M. (1981) *Biochem. J.* **195**, 471–480

Gigli, I., Porter, R. R. & Sim, R. B. (1976) *Biochem. J.* **157**, 541–548

Gigli, I., von Zabern, I. & Porter, R. R. (1977) *Biochem. J.* **165**, 439–446

Gordon, J., Whitehead, H. R. & Wormall, A. (1926) *Biochem. J.* **20**, 1028–1034

Gorski, J. P. & Howard, J. B. (1980) *J. Biol. Chem.* **255**, 10025–10028

Harpel, P. C. & Hayes, M. B. (1979) in *The Physiological Inhibitors of Coagulation and Fibrinolysis* (Collen, D., Wiman, B. & Verstraete, M., eds.), pp. 231–238, Elsevier/North-Holland Biomedical Press, Amsterdam

Howard, J. B. (1980) *J. Biol. Chem.* **255**, 7082–7084

Janatova, J. & Tack, B. F. (1981) *Biochemistry* **20**, 2394–2402

Janatova, J., Prahl, J. W. & Tack, B. F. (1979) *Congr. Int. Union Biochem. 11th, Toronto, July, 1979*, p. 524 (Abstr. 08)

Janatova, J., Tack, B. F. & Prahl, J. W. (1980*a*) *Biochemistry* **19**, 4479–4485

Janatova, J., Lorenz, P. E., Schechter, A. N., Prahl, J. W. & Tack, B. F. (1980*b*) *Biochemistry* **19**, 4471–4478

Johnson, D. M. A., Gagnon, J. & Reid, K. B. M. (1980) *Biochem. J.* **187**, 863–874

Lachmann, P. J. (1979) in *The Antigens*, (Sela, M., ed.), vol. 5, pp. 284–354, Academic Press, New York and London

Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685

Laskey, R. A. & Mills, A. D. (1975) *Eur. J. Biochem.* **56**, 335–341

Law, S. K. & Levine, R. P. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2701–2705

Law, S. K., Lichtenberg, N. A. & Levine, R. P. (1979) *J. Immunol.* **123**, 1388–1394

Law, S. K., Lichtenberg, N. A., Holcombe, F. H. & Levine, R. P. (1980) *J. Immunol.* **125**, 634–639

Mendez, E. & Marco, R. (1980) in *Methods in Peptide and Protein Sequence Analysis* (Birr, C. H. R., ed.), pp. 295–308, Elsevier/North-Holland, Amsterdam

Müller-Eberhard, H. J. & Biro, C. E. (1963) *J. Exp. Med.* **118**, 447–466

Nagasawa, S. & Stroud, R. M. (1976) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **35**, 654

Nagasawa, S. & Stroud, R. M. (1977) *Immunochemistry* **14**, 749–756

Nagasawa, S., Ichihara, C. & Stroud, R. M. (1980) *J. Immunol.* **125**, 578–582

Pangburn, M. K. & Müller-Eberhard, H. J. (1980) *J. Exp. Med.* **152**, 1102–1114

Porter, R. R. & Reid, K. B. M. (1979) *Adv. Protein Chem.* **33**, 1–71

Press, E. M. & Gagnon, J. (1981) *Biochem. J.* **199**, 351–357

Reboul, A., Thielens, N., Villiers, M. B. & Colomb, M. G. (1980) *FEBS Lett.* **115**, 118–122

Reid, K. B. M. & Porter, R. R. (1981) *Annu. Rev. Biochem.* **50**, 433–464

Salvesen, G. S. & Barrett, A. J. (1980) *Biochem. J.* **187**, 695–701

Salvesen, G. S., Sayers, C. A. & Barrett, A. J. (1981) *Biochem. J.* **195**, 453–461

Schreiber, R. D. & Müller-Eberhard, H. J. (1974) *J. Exp. Med.* **140**, 1324–1335

Shiraishi, S. & Stroud, R. M. (1975) *Immunochemistry* **12**, 935–939

Sim, R. B., Twose, T. M., Paterson, D. S. & Sim, E. (1981) *Biochem. J.* **193**, 115–127

Sottrup-Jensen, L., Petersen, T. E. & Magnusson, S. (1980) *FEBS Lett.* **121**, 275–279

Sottrup-Jensen, L., Hansen, H. F., Mortensen, S. B., Petersen, T. E. & Magnusson, S. (1981) *FEBS Lett.* **123**, 145–148

Swenson, R. & Howard, J. B. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4313–4316

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

World Health Organisation (1968) *Bull. W.H.O.* **39**, 935–938