The EMBO Journal vol.4 no.2 pp.375-382, 1985

The sequence and topology of human complement component C9

Keith K.Stanley, hans-Peter Kocher[1], J.Paul Luzio[2], Peter Jackson[2] and Jürg Tschopp[3] with technical assistance by John Dickson

EMBL, Postfach 10.2209, Meyerhofstrasse 1, 6900 Heidelberg, FRG, [1]Institute of Medical Biochemistry, University of Geneva, Switzerland, [2]Department of Clinical Biochemistry, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge, UK and [3]Institut de Biochimie, Université de Lausanne, CH-1066 Epalinges, Switzerland

Communicated by K.Simons

A partial nucleotide sequence of human complement component C9 cDNA representing 94% of the coding region of the mature protein is presented. The amino acid sequence predicted from the open reading frame of this cDNA concurs with the amino acid sequence at the amino-terminal end of three proteolytic fragments of purified C9 protein. No long stretches of hydrophobic residues are present, even in the carboxy-terminal half of the molecule which reacts with lipid-soluble photoaffinity probes. Monoclonal antibody epitopes have been mapped by comparing overlapping fragments of the C9 molecule to which the antibodies bind on Western blots. Several of these epitopes map to small regions containing other surface features (e.g., proteolytic cleavage sites and N-linked oligosaccharide). The amino-terminal half of C9 is rich in cysteine residues and contains a region with a high level of homology to the LDL receptor cysteine-rich domains. A model for C9 topology based on these findings is proposed.

Key words: complement/epitope analysis/hybrid protein/monoclonal antibody/protein topology

Introduction

C9 is the terminal component of the complement cascade. It is secreted as a soluble monomeric protein which circulates in the blood, but in the presence of membranes containing bound components C5b-8, it undergoes a gross conformational change and becomes inserted into the lipid bilayer. This process is of interest, not only because of the importance of C9 in controlling infection and in autoimmune disease, but also because the molecular mechanism for C9 insertion may provide some insights into other processes of membrane assembly. Up to 18 molecules of C9 can bind to each C5b-8 complex (Tschopp *et al.*, 1984) and their addition correlates with the formation of a stable ion pore and the appearance of characteristic 'hole-like' lesions in the membrane (Hadding and Müller-Eberhard, 1969; Bhakdi *et al.*, 1976; Tschopp, 1984a). Polymerisation of C9 can also be induced *in vitro* by incubation of C9 monomer in the presence of low concentrations of zinc ions (Tschopp, 1984b). This poly-C9 has a similar morphology to the lesions seen on target membranes *in vivo*. While monomeric C9 may be cleaved by a number of proteases (Biesecker *et al.*, 1982), polymerised C9 is peculiarly resistant to proteolytic attack and even to solubilisation by SDS (Yamamoto and Migita, 1981; Podack *et al.*, 1982; Podack and Tschopp, 1982a). This change of conformation from monomeric to poly-C9 has been visualised in the electron microscope as an unfolding of the protein (Podack and Tschopp, 1982b) and is accompanied by changes in the content of α and β secondary structure (Tschopp *et al.*, 1982) and exposure of new antigenic determinants (Podack *et al.*, 1982).

The structure of C9 in either its monomeric globular form or the tubular polymerised form is not known in any detail. We describe the binding of specific monoclonal antibodies to fragments of the protein generated either by proteolysis, chemical cleavage or recombinant DNA techniques. From these data a topological map of C9 is presented.

Results

Sequence and secondary structure of human complement C9

C9 cDNA molecules were identified by cloning human cDNA into the bacterial expression vector pEX and screening recombinant clones with polyclonal and monoclonal antibodies (Stanley and Luzio, 1984). Preliminary sequence data of one clone, C9-7, confirmed that this was a genuine C9 cDNA by comparison with the published sequence at the α-thrombin cleavage site (Biesecker *et al.*, 1982). Using this cDNA as a probe further cDNAs were isolated from the original cDNA library (Woods *et al.*, 1982) by *in situ* colony hybridisation. In total 25 clones were obtained in three screening experiments, each experiment employing a probe isolated from the 5′ end of the previous group of positive clones. In this manner 94% of the open reading frame of mature C9 was obtained as judged from the published amino acid composition (Biesecker *et al.*, 1982). No further attempts were made to obtain a complete cDNA from this library since no clones in the third screening experiment extended the 5′ end of the consensus sequence. Figure 1 shows the relative position of five of these clones and the overlapping M13 subclones that were used to sequence the cDNA. This sequence was confirmed in all areas by the

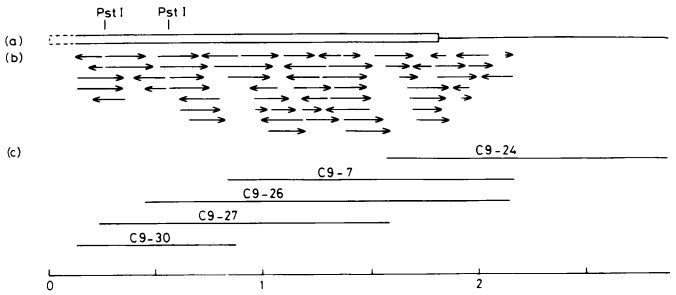

Fig. 1. Cloning of C9 cDNA. The cDNA clones C9-7 and C9-24 were obtained by screening a cDNA library subcloned in the expression vector pEX. C9-24 spanned to the 3′ end of the cDNA. C9-26, 27 and 30 were obtained from a cDNA library cloned in pBR322 by colony hybridization. **(a)** shows the C9 cDNA; open box, coding reading; dashed line, missing portion; **(b)** shows the random M13 clones used for sequencing the cDNA; **(c)** shows some of the clones obtained. The scale is calibrated in kb.

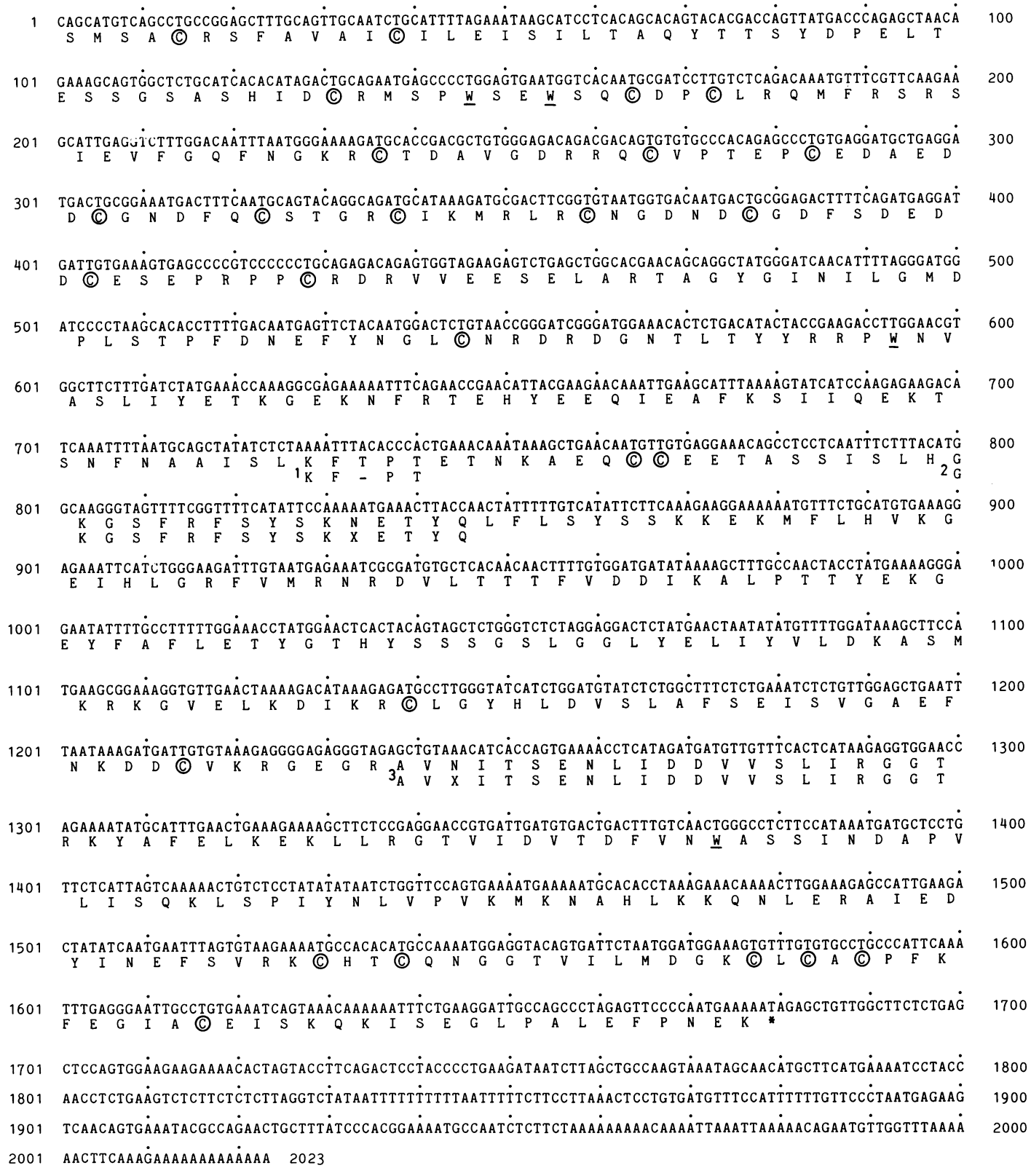

Fig. 2. Partial nucleotide and amino acid sequence of C9. Nucleotide sequences of random fragments from C9 clones were assembled by computer. The first 1681 bases of this sequence contained an open reading frame which contained the amino-terminal sequences of three proteolytic fragments of C9; **1**, chymotrypsin 38-kd fragment; **2**, α-thrombin 37-kd fragment; **3**, trypsin 20-kd fragment. Tryptophan residues are underlined and cysteine residues are encircled.

overlap of more than one different M13 clone and in most areas by fragments cloned in opposite orientations. Figure 2 shows the consensus sequence and the amino acid translation of the single large open reading frame. In addition to the published five amino acids at the amino terminus of the α-thrombin generated C9b fragment the sequence was confirmed over a total of 39 residues from three different proteolytic fragments. These results also served to position accurately

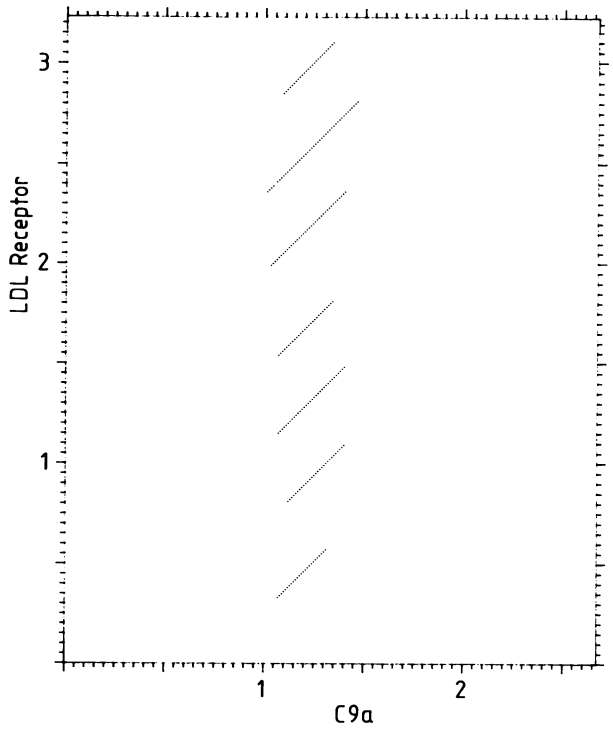

Fig. 3. Homology between C9 and the LDL receptor. The peptide sequence of C9a was compared with the high cysteine region at the amino terminus of the LDL receptor using a dot matrix analysis. The window was set at 40 residues with a stringency of 10 matches. Seven of the eight cysteine repeats of the LDL receptor are homologous to a 39 residue sequence of C9a. Scale calibrations are in 100 amino acid residues.

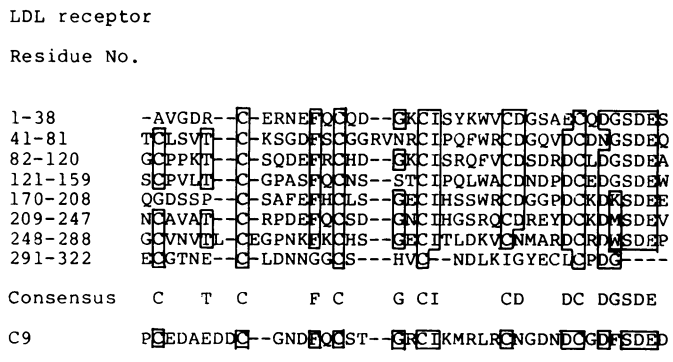

Fig. 4. Alignment of C9a sequence and LDL receptor repeats. Data for the LDL receptor is taken from Yamamoto *et al.* (1984). The portion of C9a sequence, aligned to show maximum homology, is residue 93 to residue 133 with respect to the sequence shown in Figure 2.

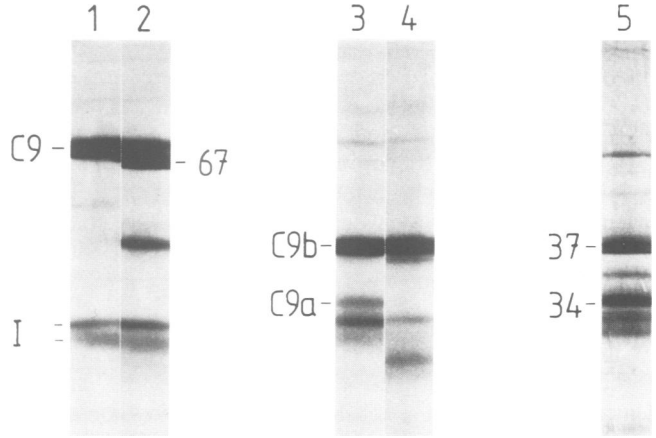

Fig. 5. Proteolytic fragments of C9 monomer. Monomeric C9 was digested with various proteolytic enzymes and analysed on 11.5% SDS-polyacrylamide gels under non-reducing conditions. **Lane 1**, undigested C9; **lane 2**, C9 + V8 *Staphylococcus* protease; **lane 3**, C9 + α-thrombin; **lane 4**, C9 + α-thrombin + V8 *Staphylococcus* protease; **lane 5**, C9 + chymotrypsin. Numbers indicate apparent mol. wts. in kd. I indicates impurities in this preparation of C9.

the proteolytic cleavage fragments on the primary sequence of the protein (see Figure 2). Only 130 bp of the 3′ non-coding region were sequenced corresponding to the 3′ end of clone C9-7. Clone C9-24 however contained 1360 bp of non-coding region ending in a stretch of ~60 As.

The inability to recover the asparagine residues marked X in the amino acid sequence of the thrombin and trypsin fragments is almost certainly due to the presence of N-linked oligosaccharide at these positions. This inference is supported by the high proportion of carbohydrate in the C9b fragment (Biesecker *et al.*, 1982), the presence of the requisite threonine residue in the position next but one to the asparagine, and the large decrease in apparent mobility on polyacrylamide gels of fragments containing this region.

Since C9 can insert into the membrane of target cells it was of immediate interest to search the sequence for domains of high hydrophobicity which might interact with a lipid bilayer. Studies using membrane restricted photoaffinity probes have shown that parts of the C9b fragment on both sides of the trypsin cleavage site are in contact with lipid (Ishida *et al.*, 1982; P.Amiguet, S. Brunner and J.Tschopp, in preparation). Furthermore, Biesecker *et al.* (1982) have shown that the C9b fragment has a higher content of valine and leucine but less charged residues than the C9a fragment and have suggested that C9b inserts into the bilayer. It was therefore interesting to note from the C9 sequence that the valine and leucine residues in C9b are not clustered together in the manner of a transmembrane sequence of an integral membrane protein, but are distributed fairly evenly through the polypeptide chain. The longest stretch without a charged residue is only of 16 amino acids. It is, therefore, most likely that the lipid interacting domain of C9 is contained within an element of its secondary structure. This would also enable nascent C9 to be translocated across the endoplasmic reticulum membrane without integration into the host cell membrane.

The secondary structure of C9 was predicted using the methods of Chou and Fasman (1978) and Garnier *et al.* (1978). Of seven α-helices in the C9b fragment only one, situated just to the amino-terminal side of the trypsin cleavage point, has a high hydrophobic moment. None of the enzyme cleavage sites or probable N-linked oligosaccharide attachment sites falls within regions given a high probability of α or β structure.

C9a contains 18 cysteine residues including one stretch of 100 amino acid residues with 13 cysteine residues at a mean spacing of eight amino acids. Since C9 contains no free SH groups, and C9a and C9b separate on SDS gels after thrombin digestion in non-reducing conditions, these cysteine residues must all be involved in disulphide bond formation within the C9a region. This clustering of cysteine residues is similar to that found in the epidermal growth factor (EGF) and low density lipoprotein (LDL) receptors (Ullrich *et al.*, 1984; Yamamoto *et al.*, 1984). Indeed one region within C9a bears a strong homology to the LDL receptor 'cysteine-

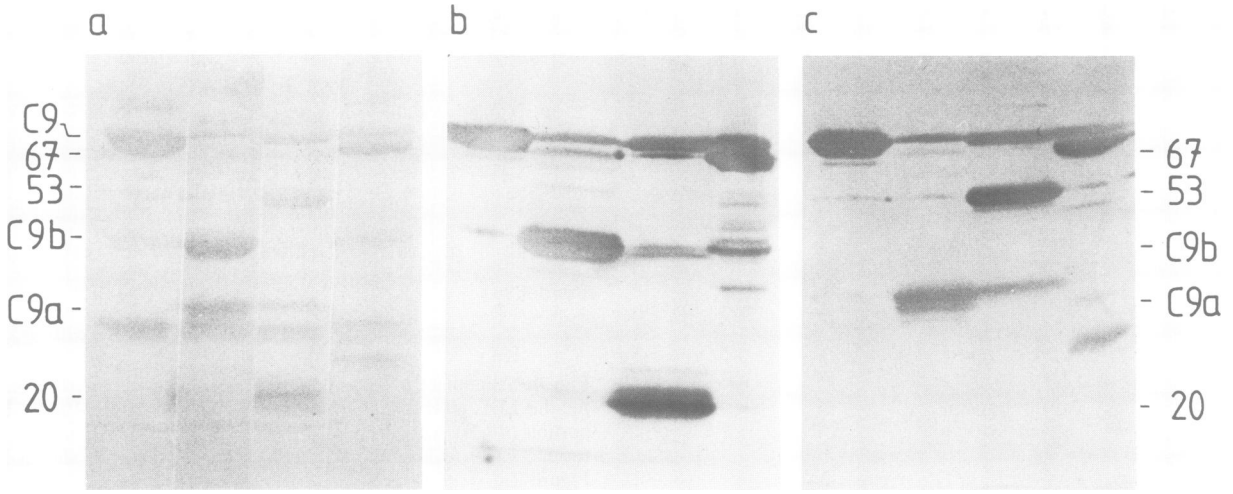

Fig. 6. Monoclonal antibody binding to proteolytic fragments of C9. Proteolytic fragments of C9 on 11.5% SDS-polyacrylamide gels were transferred to nitrocellulose filters and stained with monoclonal antibodies to C9. **Panel a**, Ponceau S stain of fragments; **b**, Western blot using antibody M42; **c**, Western blot using antibody M47. Each panel contains four lanes which show (**left to right**): uncut C9, C9+ α-thrombin, C9+trypsin, C9+V8 *Staphylococcus* protease. Numbers indicate fragment sizes in kd.

domain' consensus sequence (Figure 3). No significant homology was found by dot matrix analysis between the EGF receptor and C9 or within the C9 sequence itself. Figure 4 shows the homology between C9 and the LDL consensus repeat sequence. Fourteen out of 17 conserved residues are found in this C9 sequence.

Proteolytic fragments and hybrid C9 proteins

To map the toplogy of C9 the molecule was divided into small regions by digestion with proteolytic enzymes, chemical cleavage at tryptophan residues, and by the expression of *lacZ*-C9 gene fusions containing various fragments of the C9 cDNA. These polypeptides were then probed with a panel of monoclonal antibodies on Western blots.

Cleavage of C9 with α-thrombin results in two fragments with apparent mobilities after reduction of 34 and 37 kd, (C9a and C9b). This cleavage occurs in the centre of the protein, 294 amino acid residues from the carboxy-terminal end (Figure 2). Trypsin cleaves C9 in the C9b region to give bands on SDS-polyacrylamide gels of 20 and 53 kd (Figure 6). The exact position of this cleavage has also been determined by amino acid sequencing of the amino terminus of the 20-kd fragment (Figure 2). Chymotrypsin cleaves C9 to give two major bands on SDS gels with apparent mobilities similar to the thrombin C9a and C9b fragments (Figure 5). The amino-terminal sequence of this fragment places the chymotrypsin cleavage site 24 amino acid residues away from the thrombin site in the C9a region (Figure 2). C9 can also be digested by V8 *Staphylococcus* protease to give a major band at 67 kd (Figure 5, lane 2). This V8 protease cleavage must occur close to the amino terminus of the protein since double digestion with thrombin and V8 protease changes the size of the C9a but not that of the C9b fragment (Figure 5, lanes 3 and 4). The exact sequence of the cleavage site is not yet known. Two examples of Western blots using these proteolytic fragments are shown in Figure 6. From this it can be seen that monoclonal antibodies M42 and M47 bind to different thrombin and trypsin fragments, but to the same V8 protease fragment.

Figure 7 shows the cleavage fragments obtained using BNPS-skatole [2-(2-nitrophenylsulphenyl)-3-methyl-3′-bromoindolenine]. In this case the cleavage is performed

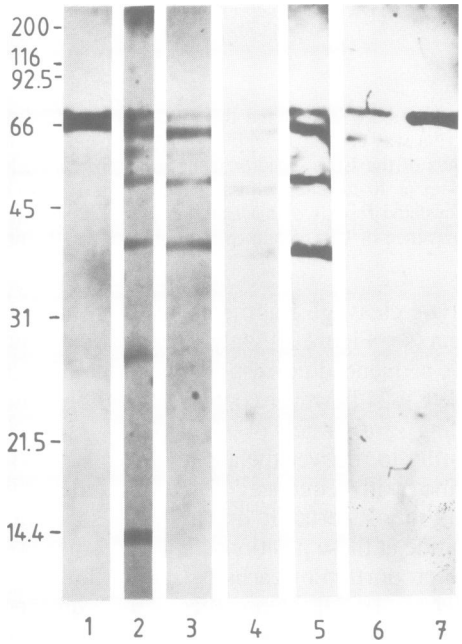

Fig. 7. Monoclonal antibody binding to BNPS-skatole cleaved C9. C9 or BNPS-skatole treated C9 was run in 15% SDS-polyacrylamide gels, Coomassie stained and transferred to nitrocellulose. Immunostaining was performed according to Jackson and Thompson (1984). **Lane 1**, C9 probed with M47; **lane 2**, Coomassie blue stained BNPS-skatole fragments of C9; **lanes 3–6** all show immunoblots of BNPS-skatole fragments using monoclonal antibody: **lane 3**, M47; **lanes 4** and **5**, M42; **lane 6**, M43; **lane 7**, uncut C9 probed with M34. The gel samples were reduced in **lanes 1–4**.

chemically on denatured C9, hence the protein is cleaved at all tryptophan residues irrespective of the domain structure of the protein. This also allows precise identification of the fragments on the C9 sequence since all four tryptophan residues are within the sequenced area of C9 (Figure 2). Five fragments would be produced in a complete digestion with sizes (predicted from the amino acid sequence) of 12.2, 0.3, 13.8, 30.0 and 11.7 kd). The size of the 12.2-kd amino-terminal fragment is estimated from the published composi-

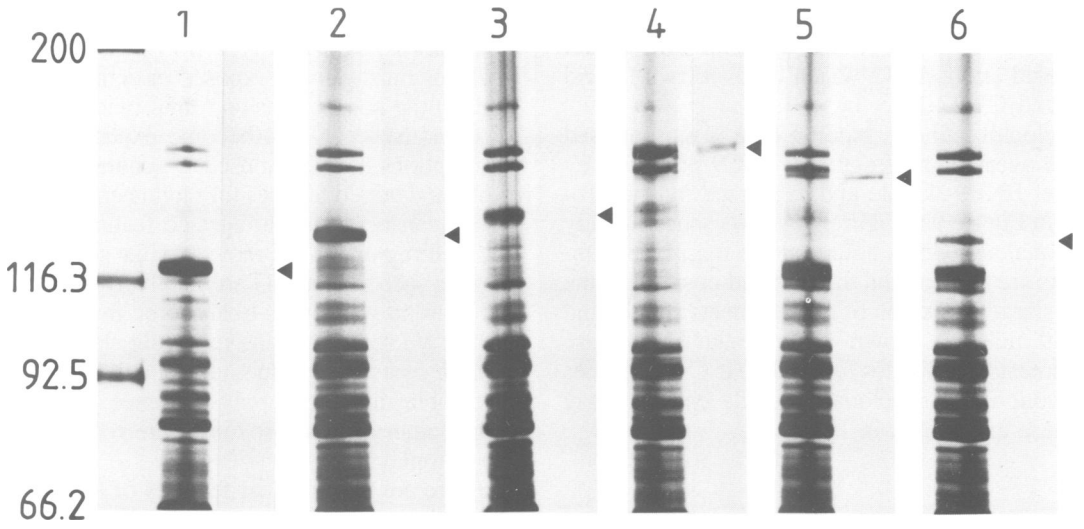

Fig. 8. Monoclonal antibody binding to C9 hybrid proteins. Cultures of *Escherichia coli* transformed with pEX containing different cloned fragments of C9 (see Figure 9) were grown at 30°C to log phase and expressed for 2 h at 42°C. Soluble proteins were extracted from the harvested cells and analysed on a 7.5% SDS-polyacrylamide gel. Each pair of tracks shows a silver stained section of the gel with the corresponding Western blot using antibody M42. Arrows show the positions of hybrid proteins and the numbers give apparent mol. wts. in kd. Only the high mol. wt. region of the gel is shown. **Lane 1**, pEX control (no C9 insert); **lane 2**, C9-30-1; **lane 3**, C9-30-2; **lane 4**, C9-27-2; **lane 5**, C9-7; **lane 6**, C9-24.

tion of C9 together with the amino acid sequence. The other fragments are given in order from the N to the C terminus. Partial digestion products are also obtained with BNPS-skatole so a large number of fragment sizes is possible; however, some of these can be unambiguously resolved on SDS-polyacrylamide gels. The 0.3-kd fragment is lost from 15% polyacrylamide gels and the three small fragments (11.7, 12.2 and 13.8 kd) co-migrate in one band with an apparent mol. wt. of 14-kd (Figure 7). The next three possible fragments have sizes predicted from the amino acid sequence of 26.3 kd (a partial fragment of 12.2 + 0.3 + 13.8 kd), 30.0 kd and 41.7 kd (a partial fragment of 30.0 + 11.7 kd). The latter two fragments, however, contain both predicted attachment sites of N-linked oligosaccharides and run on SDS-polyacrylamide gels with apparent mol. wts. of 35 and 47 kd (Figure 7). Both antibodies M42 and M47 bind to these two fragments, but not to the 26-kd or small fragments. Antibody M34, however, only binds to a much higher mol. wt. partial fragment and is also sensitive to reduction of the C9.

*Pst*I fragments from several clones were subcloned into one of the three pEX vectors so as to obtain fusions with open reading frames between the *lac*Z and C9 cDNA fragments. Inserts with the correct orientation were screened by the size of the hybrid protein bands on SDS gels. The positions of the expressed regions of C9 are shown in Figure 9 and the corresponding β-galactosidase-C9 hybrid proteins in Figure 8. In each case a hybrid protein was expressed with the predicted size except for C9-30-1. This hybrid, however, contains the region of high cysteine content from the C9a region and has a large net negative charge. This might account for its high mobility on SDS gels. C9-27-2 and C9-7 contained the largest cDNA inserts and although some hybrid protein of the predicted size was made (arrows, Figure 8) there was also a noticeable amount of degradation. Only the higher mol. wt. band contains an epitope of C9 suggesting that the cleavage occurs close to the β-galactosidase in these hybrid proteins, possibly at the thrombin cleavage site. In all cases there was

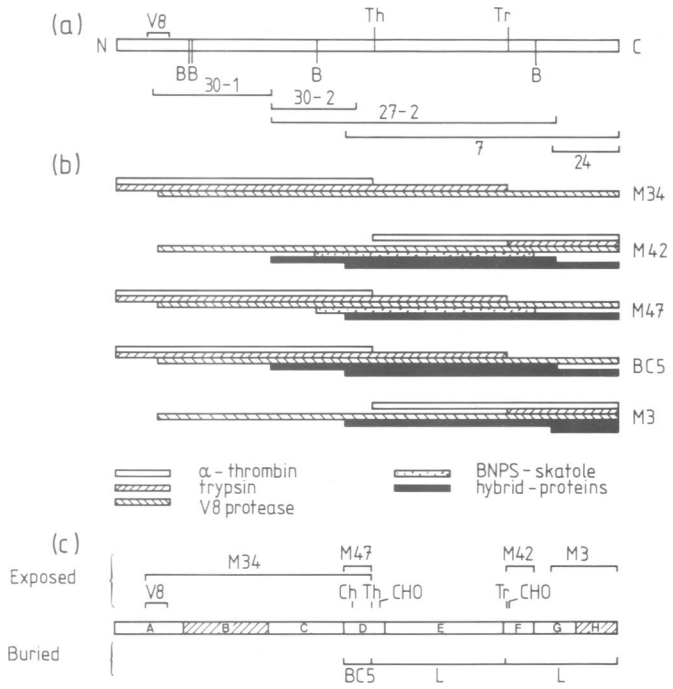

Fig. 9. Topology of complement component C9. The proteolytic, BNPS-skatole and hybrid protein fragments used in this study are shown in **(a)**. V8, V8 *Staphylococcus* protease; Th, α-thrombin; Tr, trypsin; B, BNPS-skatole cleavage sites. *Pst*I fragments from C9-30 were cloned into pEX to give C9-30-1 and C9-30-2. The larger *Pst*I fragment from C9-27 was cloned into pEX to give C9-27-2. C9-7 and C9-24 were isolated from a cDNA library in pEX. The collated results of Western blots are shown in **(b)**. Each bar represents the size and position of a C9 fragment which binds to the monoclonal antibodies shown. Using these data the epitopes for five monoclonal antibodies are shown in **(c)** together with the enzyme cleavage sites, domains containing lipid interacting sites (L), and probable sites of N-linked oligosaccharide attachment (CHO). Brackets denote areas of the molecule which have been mapped. Exposed and buried refer to the accessibility of features in monomeric C9. Domains A–H of C9 are discussed in the text.

sufficient protein for efficient staining with monoclonal antibodies in a Western blot. An example is shown in Figure 7 using monoclonal antibody 42 which binds to two hybrid proteins, C9-7 and C9-27-2.

All the monoclonal antibody binding data is summarised in Figure 9. The overlapping fragments of C9 produced by enzyme cleavage, BNPS-skatole and expressed hybrid proteins are shown in Figure 9a. Figure 9b shows schematically how the monoclonal antibodies bind to these fragments. Only positive results are shown, and the size and position of the bars reflects the size and position of the fragments which bind relative to the C9 molecule shown in Figure 9a. From these overlapping fragments boundaries for the linear C9 sequences necessary to encode the monoclonal antibody epitopes may be deduced. These are shown in Figure 9c.

Discussion

Monoclonal antibodies raised against human complement component C9 have been used to identify a number of distinct structural and functional epitopes in the molecule. Experiments studying the efflux of [^{14}C]sucrose from erythrocyte ghosts through the pore of the membrane attack complex (MAC) have shown that two of the antibodies used in this study, M42 and M47, bind to the external aspect of C9 in the MAC and inhibit sucrose efflux. The same antibodies included in the ghosts with the [^{14}C]sucrose have no effect. In contrast, antibody M34 only inhibits efflux if it is included with the label and therefore is presented to the internal face of C9 (Morgan *et al.*, 1984). These experiments have been taken to indicate the transmembrane integration of the C9 molecule. M42 and M47 have also been shown to bind to different epitopes on the C9 molecule (Morgan *et al.*, 1983; Luzio *et al.*, 1984). The antibody BC5 has been shown to bind to poly-C9 but not to monomeric C9 (T.E.Mollnes, T. Lea, M.Harboe and J.Tschopp, in preparation), and thus recognises a 'neoantigenic' region of C9 which is exposed during the polymerisation of C9.

The epitopes of C9 against which these antibodies have been raised have been mapped onto the primary sequence by determining the minimum region necessary for binding using overlapping fragments of the C9 molecules. In this approach we have chosen to consider only positive results since many different factors could contribute to a negative result (for a review, see Atassi, 1977). For example, a proteolytic cleavage might occur within an epitope, or a hybrid protein containing the right sequence might not fold into the correct conformation.

By using a series of overlapping fragments a rather precise mapping was obtained using relatively large protein fragments. Three of the antibodies were mapped to a region of <46 amino acid residues. M34 was the most poorly mapped antibody since it did not bind to any small BNPS-skatole fragments or any hybrid proteins. It is possible that the binding of this antibody is particularly sensitive to the conformation of C9 which is destroyed during fragmentation. This could also explain its higher sensitivity to reduction of the C9 than the other antibodies.

The position of enzymatic cleavages, carbohydrate attachment and antibody epitopes define some topological constraints for any model of C9 folding. This is based on the assumption that (i) all monoclonal antibodies initially screened by measuring binding to native C9 monomer must have epitopes which are present on the protein surface, (ii) all enzymatic cleavages of monomeric C9 under mild conditions must attack exposed features, (iii) these features become inaccessible after polymerisation of C9 since they are no longer accessible to the proteolytic enzymes, (iv) N-linked oligosaccharides must remain exposed in both conformations of C9, and (v) the neoantigens and lipid binding regions are buried in monomeric C9 but become exposed in poly-C9. These assumptions have been used to generate a simple map of exposed and buried features for monomeric C9 (Figure 9c). From this it can be seen that surface features are often clustered in small regions of the molecule like sections D and F. Other regions such as E and G are relatively free of surface features but overlap with areas likely to be involved in lipid interaction. Two regions of the molecule, B and H, are especially rich in cysteine residues and are likely to be situated extracytoplasmically in poly-C9 since their disulphide bonds might otherwise be broken in the reducing environment of the cytoplasm.

There are two unusual features of this map. Firstly it was surprising to find the epitope for the neoantigen in a surface exposed region of the C9 molecule. Presumably therefore this antibody recognises a surface feature in one conformation which is only present in poly-C9. Secondly it was found that antibody M34, which binds to C9 from the inside of the MAC, has an epitope in the C9a domain despite the fact that the lipid binding domains are found in C9b. This suggests that the molecule might traverse the membrane several times. Since neither thrombin nor trypsin cleavage lead to the separation under non-denaturing conditions of the C9 fragments which are generated, it is clear that a number of contacts must hold monomeric C9 together even after the polypeptide chain has been broken. Regions D and F do not necessarily lie on joining regions between structural domains therefore, but must be located in exposed areas.

Poly-C9 has the lipid binding regions exposed and the epitopes of M42 and M47 close to the central pore. Moving the two surface regions D and F to the inside of the pore could account both for the ability of these antibodies to inhibit sucrose efflux from erythrocyte ghosts (Morgan *et al.*, 1984) and also for the reduced sensitivity of poly-C9 to proteases since these domains also contain the sensitive proteolytic sites of the molecule. During translocation into the lumen of the poly-C9 pore it is possible that antibodies, but not proteases, could maintain access to these regions.

The high cysteine content of the amino-terminal half of the molecule and lack of predicted α- and β-secondary structure (data not shown) suggest that the structure of the C9a region is dominated by the disulphide bond interactions. Disulphide bonds are very hydrophobic, and proteins rich in disulphide bonds can form high cysteine domains which have a core of disulphide bonds joined by loops of the polypeptide chains. The 3-dimensional structure of such a 'cysteine fold' has been solved for the wheat germ agglutinin molecule which has four domains each of which contains eight cysteine residues covalently linked by S-S bridges (Wright, 1977; Drenth *et al.*, 1980; Wright *et al.*, 1984). High cysteine domains are also to be found in the membrane receptors for EGF and LDL (Ullrich *et al.*, 1984; Yamamoto *et al.*, 1984). Dot matrix analysis indeed shows that one region of C9a bears a close resemblance to the 'cysteine-domain' repeat of the LDL receptor. Furthermore, this region contains the highly conserved sequence Asp-Cys-X-Asp-Gly-Ser-Asp-Glu with only one change (Phe for Gly) which Yamamoto *et al.* (1984) ascribe to the function of apoprotein B and E binding. Since C9 polymerisation can be inhibited by HDL, LDL and some

apoproteins (Rosenfeld *et al.*, 1983), it is possible that this conserved sequence could have functional significance for MAC formation or regulation. It cannot be ruled out, however, that this cysteine repeat sequence codes for a structural motif common to a number of extracytoplasmic protein domains. In general, all the high cysteine proteins display a remarkable stability towards proteases and detergents and it seems most likely that this property of C9 is a function of its high cysteine domain which must therefore remain exposed after formation of the MAC.

Evidently further studies will be required to map the exact folding of the C9 polypeptide chain. Studies at present in progress attempt to do this by raising polyclonal antibodies of predetermined monospecificity to small regions of the C9 protein cloned and expressed in pEX.

Materials and methods

Proteolysis of C9

Stock solutions of trypsin (Serva), V8 *Staphylococcus* protease (Miles), chymotrypsin (Serva) and thrombin (prepared according to Lundblad *et al.* (1976) were prepared at 1 mg/ml in 150 mM NaCl, 10 mM Tris-HCl, pH 7.4. C9 (1 mg/ml) was digested at room temperature with 2.5% (w/w) trypsin for 10 min, with 2.5% chymotrypsin for 7 min and with 5% V8 *Staphylococcus* protease for 10 min. C9 digestion by thrombin (2%) was carried out at 37°C for 2 h. Reactions were stopped by the addition of 2% boiling SDS, or, for sequence analysis, by freezing and lyophilising the samples. Since the amino terminus of C9 is blocked no further purification of the fragments was necessary.

Amino acid sequence analysis

50 nmol of enzyme-cleaved C9 was subjected to automated degradation in a Beckman sequencer Model 890 C, using the original Beckman program No. 122974 with 0.25 M Quadrol buffer. Sequencer reagents and solvents were purchased from Fluka, Buchs (Switzerland) with the exception of heptafluorobutyric acid which was from Beckmann. The initial yield of the first degradation step was 60% and the repetitive yield 92%. The conversion of the cleaved thiazolium residues into phenylthiohydantoin (PTH) derivatives was performed by incubation in 1 M HCl at 80°C for 10 min. Analysis of the PTH-derivatives was carried out by h.p.l.c. on a reverse phase C18 column as previously described (Brandt and Jaton, 1978), and by t.l.c. according to Rangarajan and Dabre (1975) for the distinction of PTH-phenylalanine/isoleucine and PTH-phenylvaline/methionine.

BNPS-skatole digestion

C9 was cleaved at tryptophan residues by digestion with BNPS-skatole (Pierce and Warriner, Chester, UK) essentially as described by Fontana (1972). To 250 μg of C9 in 104 μl of water was added 208 μl of glacial acetic acid containing a 200-fold molar excess of BNPS-skatole over the total tryptophan content of the protein and 21 μl of 90% (v/v) aqueous acetic acid solution containing 1.24 mg/ml tyrosine. Digestion was for 18 h at 22°C in the dark, when a second portion of BNPS-skatole in glacial acetic acid, equal to the first was added and digestion continued for a further 6 h. The digest was diluted with 20 ml water and extracted four times with 30 ml diethyl ether. The aqueous layer containing the peptides was stored frozen at −20°C.

Hybrid protein expression

C9-7 and C9-24 were C9 clones obtained by screening a human liver cDNA library subcloned into pEX (Stanley and Luzio, 1984). C9-30-1, C9-30-2 and C9-27-2 were constructed by subcloning *Pst*I fragments from pBR322 clones C9-30 and C9-27 into the *Pst*I site of pEX. C9-30 had a single base change in one *Pst*I site enabling it to be subcloned in two rather than three fragments. Recombinant clones were expressed as previously described (Stanley and Luzio, 1984).

Western blots

Western blots were carried out as described by Burnette (1981) and Jackson and Thompson (1984) using horseradish peroxidase-coupled second antibody and either diaminobenzidine or 4-chloro-1-naphthol as substrate.

DNA sequencing

C9 DNA fragments cloned into M13mp8 were sequenced by the method of Sanger *et al.* (1977) using ^{35}S label and buffer gradient gels according to Biggin *et al.* (1983).

Monoclonal antibodies

Mouse monoclonal antibodies M34, 42 and 47 were raised against C9 by standard methods (Galfre and Milstein, 1981), as described by Morgan *et al.* (1983). In this paper M34 etc. designates the antibody C9-34 etc. in order to minimise confusion with hybrid clone nomenclature. Antibody BC5 was screened by its ability to bind to the MAC or SC5b-9 complex rather than monomeric C9 (T.E.Mollnes, T.Lea, M.Harboe and J.Tschopp, in preparation). M3 has been described previously (J.Tschopp, S.Schäfer, P.Amiguet and J. Tamerius, in preparation).

Acknowledgements

We thank Dr D.Banner for advice with structural interpretation, and Dr T.E.Mollnes, Dr T.Lea, Dr K.Siddle, Dr P.B.Morgan and Dr A.K.Campbell for generously providing their monoclonal antibodies. We also thank the Arthritis and Rheumatism Council for a grant to J.P.L for preparation of monoclonal antibodies. This work has been supported by a Swiss National Science Foundation grant (No. 3.267.82) to J.T.

References

Atassi,M.Z., ed. (1977) *Immunochemistry of Proteins*, Plenum, NY, pp.1-161.

Bhakdi,S., Ey,P. and Bhakdi-Lehnen,B. (1976) *Biochim. Biophys. Acta*, **419**, 445-457.

Biesecker,G., Gerard,C. and Hugli,T.E. (1982) *J. Biol. Chem.*, **257**, 2584-2590.

Biggin,M.D., Gibson,T.J. and Hong,G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963-3965.

Burnette,W.N. (1981) *Anal. Biochem.*, **112**, 195-203.

Brandt,D.C. and Jaton,J.-C. (1978) *J. Immunol.*, **121**, 1194-1198.

Chou,P.Y. and Fasman,G.D. (1978) *Adv. Enzymol.*, **47**, 45-148.

Drenth,J., Low,B.W., Richardson,J.S. and Wright,C.S. (1980) *J. Biol. Chem.*, **255**, 2652-2655.

Fontana,A. (1972) *Methods Enzymol.*, **25**, 419-423.

Galfre,G. and Milstein,C. (1981) *Methods Enzymol.*, **73**, 3-46.

Garnier,J., Osguthorpe,D.J. and Robson,B. (1978) *J. Mol. Biol.*, **120**, 97-120.

Hadding,U. and Müller-Eberhard,H.J. (1969) *Immunology*, **16**, 719-735.

Ishida,B., Wisnieski,B.J., Lavine,C.H. and Esser,A.F. (1982) *J. Biol. Chem.*, **257**, 10551-10553.

Jackson,P. and Thompson,R.J. (1984) *Electrophoresis*, **5**, 35-42.

Lundblad,R.L., Kingdan,H.S. and Mann,K.G. (1976) *Methods Enzymol.*, **45**, 156-176.

Luzio,J.P., Jackson,P., Campbell,A.K., Morgan,P.B. and Stanley,K.K. (1984) *Biochem. Soc. Trans.*, **13**, 105-106.

Morgan,P.B., Daw,R.A., Siddle,K., Luzio,J.P. and Campbell,A.K. (1983) *J. Immunol. Methods.*, **64**, 269-281.

Morgan,B.P., Luzio,J.P. and Campbell,A.K. (1984) *Biochim. Biophys. Res. Commun.*, **118**, 616-622.

Podack,E.R., Tschopp,J. and Müller-Eberhard,H.J. (1982) *J. Exp. Med.*, **156**, 268-282.

Podack,E.R. and Tschopp,J. (1982a) *J. Biol. Chem.*, **257**, 15204-15212.

Podack,E.R. and Tschopp,J. (1982b) *Proc. Natl. Acad. Sci. USA*, **79**, 574-578.

Rangarajan,M. and Darbre,A. (1975) *Biochem. J.*, **147**, 435-438.

Rosenfeld,S.I., Packman,C.H. and Leddy,J.P. (1983) *J. Clin. Invest.*, **71**, 795-808.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

Stanley,K.K. and Luzio,J.P. (1984) *EMBO J.*, **3**, 1429-1434.

Tschopp,J. (1984a) *J. Biol. Chem.*, **259**, 7857-7863.

Tschopp,J. (1984b) *J. Biol. Chem.*, **259**, 10569-10573.

Tschopp,J., Engel,A. and Podack,E.R. (1984) *J. Biol. Chem.*, **259**, 1922-1928.

Tschopp,J., Müller-Eberhard,H.J. and Podack,E.R. (1982) *Nature*, **298**, 534-538.

Ullrich,A., Coussens,L., Hayflick,J.S., Dull,T.J., Gray,A., Tam,A.W., Lee,J., Yarden,Y., Libermann,T.A., Schlessinger,J., Downward,J., Mayes,E.L.V., Whittle,N., Waterfield,M.D. and Seeburg,P.H. (1984) *Nature*, **309**, 418-425.

Woods,D.E., Markham,A.F., Ricker,A.T., Goldberger,G. and Colten,H.R. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5661-5665.

Wright,C.S. (1977) *J. Mol. Biol.*, **111**, 439-457.

Wright,C.S., Gavilanes,F. and Peterson,D.L. (1984) *Biochemistry (Wash.)*, **23**, 280-287.

Yamamoto,K.I. and Migita,S. (1981) *J. Immunol.*, **127**, 423-426.
Yamamoto,T., Davis,C.G., Brown,M.S., Schneider,W.J., Casey,M.L., Goldstein,J.L. and Russell,D.W. (1984) *Cell*, **39**, 27-38.

Received on 28 November 1984

Note added in proof

While this manuscript was in press a similar sequence for human complement C9 cDNA has been published (DiScipio,R.G., Gehring,M.R., Podack,E.R., Kan,C.C., Hugli,T.E. and Fey,G.H. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7298-7302). These authors claim that the mature protein sequence starts at residue 23 of the sequence shown in Figure 2 in contradiction to the previously published amino acid composition (Biesecker *et al.*, 1982). In this case the methionine residue at position 2 is a candidate for the initiating methionine of pre-C9 and residues 2-22 could represent a leader peptide.