with a hydrophobic cluster of Tyr-59, Ile-60 and Leu-63 on one surface. It is this hydrophobic cluster that we suggest is being mimicked by the tryptophanyl-tryptophan groups as illustrated in Fig. 1. We also propose that the hydrophobic cluster in C3a represents a secondary binding site on the molecule, and it is a cooperative interaction on the receptor between the hydrophobic and effector sites (LGLAR) that results in the full activity of C3a and the C3a fragment C3a 57-77. If this concept is accurate, then the secondary binding site, defined properly only when the C-terminal region of C3a is in a helix, can be reproduced in a linear array by selecting hydrophobic groups of the appropriate shape and size. This model explains why the length of the supra-agonist peptides can be optimized since the distance between the acceptor sites on the receptor are relatively fixed.

The importance of this model may be far reaching for complex protein molecules that depend on a folded conformation for activity. Instead of attempting to mimic the secondary or tertiary conformational motif, it seems entirely possible to organize the essential features of the binding site in a linear array. Furthermore, in understanding the nature or structural requirements of the binding sites in biological factors, one may actually improve on the interactions permitting designs of linear supra-agonists without need or benefit of the secondary or tertiary folding constraints.


Received 23 July 1990

S. K. ALEX LAW and ALISTER W. DODDS
M.R.C. Immunochemistry Unit, Department of Biochemistry,
University of Oxford, South Parks Road, Oxford OX1 3QU,
U.K.

The three complement proteins C3, C4 and C5 share the common ability to interact non-specifically with target cell surface structures. Upon activation, C3 and C4 bind covalently to cell surface macromolecules via hydroxyl and amino groups. Activated C5 does not show covalent binding properties, but initiates and focuses the assembly of the terminal components in the membrane. C3, C4 and C5 are related. Even before the availability of information on their primary structure, they were usually discussed as a group by the following criteria [1, 2]. (1) They are of similar sizes. (2) Each is synthesized as a single polypeptide which is subsequently processed into a disulfide-linked multichain structure: C3 and C5 into two chains with M, values of approximately 110 000 for and 75 000 for β, and C4 into three chains with M, values of approximately 95 000 for α, 75 000 for β and 33 000 for γ. The order of the chains in the pro-molecules is β-α for C3 and C5 and β-α-γ for C4. (3) They are activated similarly by enzymic removal of peptides approximately 77 residue long from the N-termini of the α-chains to yield the C3a, C4a and C5a anaphylatoxins. (4) C4 and C5 play analogous roles in the C3- and C5-convertingases of the classical and alternative pathways of complement activation. (5) The activities of C3 and C4 are regulated by the enzyme factor I and proteins of the RCA (regulation of complement activation) superfamily. (6) When activated on the cell surface, only about 5-10% of C3 or C4 bind covalently to the cell surface, and the non-covalent binding of C5 may be as low as 1%, since it is dependent upon the presence of C6 and C7.

When data on the primary structures of these proteins began to accumulate by amino acid sequencing, the assumption that they may have a common ancestry was justified. The complete structures, however, were derived from the sequences of their respective cDNAs [3-7]. Pairwise comparisons between the three translated sequences show post-alignment identity of about 25%. It is interesting to note that
though similar to each other, no segment in any of the three proteins was found to be related to any other known structures. In addition, the similarities between these proteins are observed throughout the length of the proteins (with the exception that C4 has a γ-chain at the C-terminal end). In other words, if these proteins are organized into domains, the component domains have not been described in any other proteins. Yet, the undefined domains must be organized similarly in these proteins, suggesting that the overall organization is preserved for structural and functional reasons. The argument is further substantiated by data on the exon organization of their genes, the coding region of each being split similarly into about 40 exons [8–12].

Another protein which should be included in this discussion is the serum protease inhibitor α1-macroglobulin (α1-M). Its primary structure shows a similar level of identity to those of C3, C4 and C5, and like C3 and C4 it shows covalent binding properties [13].

The subject to be discussed in more detail is the internal thioester and the covalent binding reaction it mediates in C3 and C4, and α1-M. The reaction, using C3 as a model, is schematically illustrated in Fig. 1. The internal thioester in C3 is formed between the thiol group of the cysteine residue and the carbonyl group of the glutamine residue in the tetrapeptide sequence -Cys-Gly-Glu-Gln-. Thus, a 15 member thiolactone ring (Fig. 2) is found in the molecule. In the native protein, the thioester is not accessible to water and other nucleophiles, and is rather stable. Upon activation by the enzymic removal of the anaphylatoxin C3a, a conformational change takes place and the thioester becomes accessible to nucleophiles including water in the fluid phase. Biological macromolecules with amino or hydroxyl groups can also react with the thioester to form amide- or ester-linked covalent complexes. If C3 is activated on the cell surface of a pathogen, some of the molecules become bound to the cell surface. Most of the activated molecules, however, will become inactivated by hydrolysis of the exposed thioester in the fluid phase. The activation scheme is also applicable to C4 and α1-M.

This formulation is satisfactory as to how the complement proteins can function effectively in the immune system. The activated thioester in C3 and C4 can react with all nucleophiles; hence, they have the potential to bind to the surface of all cells, including those of the host, via the hydroxyl and amino groups of biological macromolecules on the cell surface. On the other hand, the very reactive nature of the exposed thioester puts a limit on the effective range of the activated C3 and C4 molecules. Water, at a concentration of 55 m in the aqueous medium, virtually restricts the activated C3 and C4 from binding to any cell other than the one on which they are activated by hydrolysis of the thioester. Cells of the host, which do not activate either the classical or the alternative pathway under normal circumstances, are therefore free from the deposition of large amounts of C3 and C4 and the damaging consequences.

Let us consider the reaction in more detail. It has long been known that C3 and C4 could be inactivated by ammonia and other small nucleophiles such as hydrazine and hydroxylamine [14]. In the later 1970s, it was established that the interaction between C3b and cell surface acceptor molecules (AM) was by an ester bond in the form of C3b-AM-O-CO-O-AM [15, 16]. Since formation of such a bond from a carboxyl and a hydroxyl group would require energy, it was postulated that some kind of chemical bond is already present in the C3 molecule and that the binding reaction is by a transacylation mechanism [16]. The sensitivity of C3 to small nucleophiles could be interpreted as nucleophilic attack on the internal bond. At about the same time, the property of an unusual thiol group in C3 was described [17, 18]. It is not detectable in native C3. Its appearance, however, can be quantitatively titrated with the inactivation of C3 by small nucleophiles. It can also be quantitatively detected in C3b. The conclusion that there is an internal thioester in C3 was inevitable. The experiment that ultimately demonstrated the presence of the thioester was described by Tack et al. [19], where 14C-methylamine was used to inactivate C3, which was subsequently treated with [3H]iodoacetate. The radioactive tryptic peptide was isolated and sequenced and the two radioactivities were recovered three residues apart. The 3H radioactivity was recovered as S-carboxymethylcysteine, whereas the 14C radioactivity was recovered as γ-glutamylmethylamide. The two residues in between the glycine and glutamic acid. It was thus established that the thioester was between the thiol group of the cysteine residue and the carbonyl group of a glutamate or glutamine residue. (It was finally determined by cDNA sequencing that the glutamyl residue was coded for as glutamine [3].)

The thioester in the native C3 and C4 is very stable. A free thioester in aqueous solution would have a half-life in the order of minutes [20], whereas the thioester in the native C3 and C4 is stable for days. In one case, C4 was incubated at 37°C and over 80% of the haemolytic activity was recovered after 77 h [21]. Furthermore, amino-group-bearing compounds of larger size, such as the amino acids, have no effect on the activity of C3 and C4. A more coherent study could be found in the incubation of C3 with a series of alkylamines. The rate of inactivation was found to be related to the size of the alkyl group, with the most effective one being methylene and the potency diminishing progressively with ethylamine and isopropylamine at 0.024 and

![Fig. 2. Internal thioester of C3, C4 and α1-M](image-url)
COMPLEMENT SYSTEM

0.0056 the rate of methylamine, respectively. Inactivation of C3 was not detected when incubated with t-butylamine [22].

If the thioester is internal in C3 and becomes accessible to acceptor molecules upon activation, a conformational change in C3 must take place in the process. The change in conformation of C3 during its activation by trypsin is monitored by its c.d. spectrum at near U.V. [23]. The change in c.d. spectrum has similar kinetics as the rate of conversion of C3 to C3b by trypsin. For the individual C3 molecule, the change in conformation is almost instantaneous upon conversion. This conclusion is in agreement with the observation that activated C3b (C3b* in Fig. 1) has a very short life. The next question was whether enzymic cleavage is absolutely required for the activation of C3. It has long been known that treatment of serum with chaotropes such as KBr and KSCN would destroy C3, C4 and C5 activities, and this method had been employed to prepare reagents lacking the function of these components. In a series of experiments where C3 was incubated with radioactive glycerol in the presence of various chaotropes under conditions that would lead to C3 activation/inactivation, covalent incorporation of glycerol into the C3 chain was detected [24]. Furthermore, when the c.d. spectrum of C3 was monitored in the presence of KBr, it was found that the spectral change followed the same kinetics as the appearance of the free thiol group in C3. Thus conformational change in C3, as induced by chaotropes, is associated with activation. Enzymic cleavage can be regarded as the biological means to promote the necessary conformational change. Since it is impossible to compare the kinetics of conformational change and the binding of activated C3 to acceptor molecules, the cause and effect of these two phenomena could not be properly assessed. However, when C3 is inactivated by methylamine, the change in c.d. spectrum is much slower and is about 10% of the rate of inactivation. Prolonged incubation is required to attain a C3b-like spectrum [23]. Thus, breaking the thioester bond in C3 is not sufficient to induce the rapid conformational change associated with activation. Conversely, appropriate conformational change is sufficient to account for the exposure of the thioester and the activation of C3.

The conformational change is not simply to expose the thioester. In the presence of low concentrations (0.5-1.5 m) of guanidine hydrochloride, covalent binding of glycerol to C3 could be induced. At progressively higher concentrations (> 2 m) of guanidine hydrochloride, the efficiency of binding decreases. Instead, the autolytic cleavage of the α-chain becomes more evident [24]. Khan & Erickson [25, 26] synthesized the thiolactone ring that resembled those of C3, C4 and α.M within the hexapeptide Gly-Cys-Gly-Glu-Glx-Asn (see Fig. 2). In aqueous solution, the thiolactone ring undergoes a reaction similar to the autolytic cleavage reaction. Hydrolysis of the thioester, though detectable, proceeds at a much slower rate. These results, when taken together, suggested that the autolytic cleavage reaction is a property of the thiolactone ring and it is observed when the protein is denatured; that activation requires some form of higher order structure in the protein to be preserved to promote the reactivity of the thioester to nucleophiles. It was thus postulated that some catalytic residues, without defining their roles in detail, must exist [27]. (The autolytic cleavage reaction is observed when native C3, C4 and α.M are treated with denaturants such as guanidine hydrochloride and SDS. Presumably, the peptide nitrogen attacks the thioester to form an internal lactam ring followed by the hydrolysis of the peptide bond. As a result, the α-chain of C3 and C4, and the single chain of α.M, is split into two fragments. Though probably not of physiological importance, since it only takes place when the protein has lost its higher order structure, it is nonetheless an extremely useful way to determine whether the thioester is intact before denaturation. See [28].)

Although the discussion has been carried out in the context of C3, the conclusions drawn are also applicable to C4 and α.M, since similar experiments on C4 and α.M have yielded similar results [29-36].

C5 does not have an internal thioester and it does not bind covalently to cell surface molecules [29, 30]. The four residues at equivalent positions to those that form the internal thioester in C3 and C4 are -Ser-Ala-Glu-Ala- with the cysteine replaced by a serine and the glutamine by an alanine [7]. Since it has been argued that, in the case of C3 and C4, the thioester is probably not crucial to the stability of the conformation of the native molecules, the activation of C5 could be described similarly without the inclusion of the thioester. The conjecture is consistent with the two other properties of C5: that it is also sensitive to chaotropes and that activated C5 also has a relatively short life [37].

Before describing how the catalytic groups were determined, it is necessary to describe the experimental system which was employed to provide a more reliable measurement on the chemical parameter of the covalent binding reaction. It is a fluid phase system, where C3 is incubated with radioactive small molecules at known concentration and specific radioactivity, in the presence of trypsin. Since C3b is rather resistant to further proteolysis by trypsin, the reaction could be terminated at a point when all of the C3 is converted to C3b. The protein is denatured and the chains separated by SDS/PAGE. The α-chain of C3b is identified by staining and the band is cut out and the amount of radioactivity determined. The reaction rate of the small molecules with C3 can be calculated. Binding efficiency (BE), defined as the fraction of C3b labelled with the radioactive small molecules, can be expressed as a function of the concentration of the small molecules (S), the first-order hydrolysis rate of the thioester (k1), and the second-order rate of reaction between the small molecule and the thioester (k2). Thus:

\[
BE = \frac{k_2 S}{k_1 + k_2 S}
\]

which can be rearranged to the form:

\[
k_2 = \frac{1}{BE} k_1 - \frac{1}{BE} S
\]

Since \(S\) is predetermined, and BE a measurable quantity, \(k_2/k_1\) for the small molecules of interest can be calculated [38].

This system was also used to study the binding reaction of C4 and α.M. In the case of C4, the complement subcomponent C1s is used as the convertase. In most of the experiments, glycerol and glycine were generally used as model small molecules bearing hydroxyl and amino groups, respectively.

Identification of the catalytic groups would have been impossible were it not that man has two different forms of C4. The two C4s, known as C4A and C4B, are coded for by tandem genes in the major histocompatibility complex (MHC) region [2]. The two proteins are almost identical in structure, with only six different residues in a total of over 1700 [39]. (This over-simplification of C4 genetics is adequate for the purpose of discussing the thioester and its covalent binding reaction.) Functionally, it was known that the specific haemolytic activity of C4A is about 30% that of C4B [21, 40]. As it turns out, it is their difference in binding efficiency to the erythrocyte surface that caused the observed difference in their haemolytic activities [21, 41]. In the fluid
phase binding system, a more pronounced difference could be measured. C4A has a very fast reaction rate with glycine, but is virtually unreactive with glycerol at neutral pH. C4B reacts with glycine at a rate about 1% that of C4A at neutral pH, but its reaction rate with glycerol, in comparison with that of C4A, is significantly higher. Thus, the specificity of binding to hydroxyl and amino groups must be determined by two or three different residues, and it follows that one or up to all of the six residues are catalytic to the binding reaction [21]. By making similar measurements on some rarer allotypes of C4, the residues responsible for the difference in specificities were narrowed down to four. They are the residues at positions 1101, 2, 5 and 6 with Pro, Cys, Leu, Asp respectively [39, 42]. Similar measurements were also made on C3 and a,M, Table 1. The data on mouse C4 are also included. Structurally it has an overall 75% post-alignment identity with the human C4s. Significantly, it has two each of the C4A and C4B specific residues [44, 45]. Since it functionally resembles C4B, it is tentatively concluded that the Leu-Asp and Ile-His difference is important in the determination of the binding specificity.

Of the two residues, it is likely that the charged ones at position 1106 would play a more important role in the binding reaction. It is possible that the His in C4B and C3 is responsible for enhancing the nucleophilicity of hydroxyl groups and that the Asp in C4A can deprotonate amino groups at neutral pH, αM. While there is no Asp in this position, reacts with amino groups similarly to C4B, and it does not react with hydroxyl groups. C3 has a His and reacts with hydroxyl groups, but its reaction with glycine is not detectable at neutral pH. Some other residue(s), not necessarily in positions 1101-1106, must be involved in this reaction. Since C3 and α,M are only 25% identical to C4, it is presently too early to make a contribution of any particular residue in detail. Work is under way to replace the six residues in C4 by those of C3 and α,M, and thus more meaningful measurements will be made on a C4 background.

Recently, Carroll and Isenman and their colleagues [46] have used site-directed mutagenesis techniques to interchange the residues in the region of 1101-1106 between C4A and C4B and examine the contribution of each to the binding reaction. The assay system used was the comparison of the binding of the expressed molecules to sensitized sheep erythrocytes bearing C1 (hydroxyl-group rich) and IgG immune aggregate (amino-group rich). The main finding was that the Asp/His interchange at position 1106 could largely account for the difference in C4A/C4B activity. However, a Cys/Pro interchange at position 1102 also had an effect on the binding reaction.

Finally, we would like to mention briefly our current work on the C4 of various mammals [47]. The basic puzzle is that man has two forms of C4 with distinct binding specificities and they are different in 4 out of 1700 residues. Mouse only has one C4, which is C4B-like in its binding specificity, but it has two each of the A and B specific residues. In all other mammals studied, a C4B-like activity could be found. However, in addition to the C4B-like activity, C4A-like activity was found, not surprisingly, in primates, but rather unexpectedly also in the sheep and cattle. The evolution of C4 genes is certainly complicated. Gene duplications have probably occurred independently in different mammalian branches. The acquisition of a C4A-like activity in the primates and ungulates, perhaps by gene conversion, could also be unrelated events. Before further speculation, we are currently engaged in obtaining sequence data on the catalytic region of C4 in various animals.

Table 1. Reaction rates of thioester containing proteins with glycine and glycerol and the catalytic residues

<table>
<thead>
<tr>
<th>Residues</th>
<th>Glycine</th>
<th>Glycerol</th>
<th>k_5/k_5(M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human C4A</td>
<td>13400</td>
<td>1.3</td>
<td>P C P V L D</td>
</tr>
<tr>
<td>Human C4B</td>
<td>119</td>
<td>15.5</td>
<td>L S P V I H</td>
</tr>
<tr>
<td>Mouse C4</td>
<td>136</td>
<td>26.0</td>
<td>P C P V I H</td>
</tr>
<tr>
<td>Human C3</td>
<td>0</td>
<td>23.0</td>
<td>D A P V I H</td>
</tr>
<tr>
<td>Human α,M</td>
<td>206</td>
<td>1.2</td>
<td>S G S L L N</td>
</tr>
</tbody>
</table>

Protection against complement lysis

PETER J. LACHMANN
Molecular Immunopathology Unit, Medical Research Council Centre, Hills Road, Cambridge CB2 2QH, U.K.

Introduction
The recognition that mammalian cells are protected from complement attack — and especially from lytic attack by their own species of complement — and that this protection is brought about by the presence of specific proteins on cell membranes is of fairly recent origin. It is now clear that this phenomenon, like so much of the complement system, is not due to a single protein or a single mechanism but occurs as a result of interference with complement action at at least two steps and in both cases probably by more than one protein.

Protection against complement activation at the C3-convertase stage
Two activities of membrane protein limit C3 convertase activity. The first mechanism is used by the decay accelerating factor (DAF) which acts by interfering with the association between the two components of the C3 convertases; namely C4b and C2 in the case of the classical pathway; C3b and factor B in the case of the alternative pathway. There is argument whether DAF dissociates the complex or prevents it from forming or both, but for practical purposes this makes little difference. DAF is a glycolipid anchored membrane protein with a wide distribution. It is coded in the regulator of complement activation (RCA) cluster on chromosome 1q.

DAF is the Cromer blood group antigen and it is now known that isolated DAF deficiency (the INAB phenotype of the anaemia in man. The second mechanism is the accelerated breakdown of C3b and iC3b produced by membrane bound factor I cofactors. DAF itself does not have cofactor activity but two other RCA molecules — CR1 and membrane cofactor protein — do. It can be shown in vitro that this activity limits the amount of C3 and C5 convertase activity generated in cells and it presumably has an analogous effect in vivo. However, for the initial clip of C3 to iC3b, it is not immediately apparent why this would augment the effect of the very large amount of circulating factor H in vivo. On the other hand CR1 is also a cofactor for C4b breakdown where the analogous serum cofactor is C4 binding protein. C4 binding protein has a high molecular mass and may therefore be at relatively low concentration at the extra-vascular sites. This aspect of cofactor activity may therefore be more important in limiting the classical pathway.

Protection from lysis acting at the membrane attack complex (MAC) stage
It is experimentally quite easy to distinguish protection from lysis at these late acting stages from those acting on C3 convertases because only the former will inhibit reactive lysis, the procedure where MAC attack is initiated using activated C56. Here again there are at least two mechanisms though only one applies to erythrocytes.

On nucleated cells that have active membrane metabolism, cells can protect themselves from complement attack by a process which has been called vesiculation [1], where the complement channels are aggregated at one portion of the cell membrane which is then budded off so that it no longer functions as the effective plasma membrane. This allows sublytic membrane attack to provoke cell activation rather than lysis.

The major mechanism however that I wish to discuss involves protection particularly from the species’ own complement.

Homologous restriction
This phenomenon of homologous restriction is dependent upon the presence of specific glycolipid anchored proteins in the cell membrane. The original protein with such an activity was described by Schonermark et al. as C8 binding protein [2]. It has a molecular mass of about 68 kDa, is absent (as are all glycolipid anchored proteins) from paroxysmal nocturnal haemoglobinuria cells and antibodies to it confer on normal human erythrocytes an increased susceptibility to complement lysis, Zalman et al. [3] have studied what appears to be the same protein which they call ‘homologous restriction factor’. They claim further that it is also present in the granules of cytotoxic lymphocytes (NK cells and T cells) and that it restricts cell mediated lysis. This is a surprising finding since NK lysis does not show the phenomenon of homologous restriction and NK cells in general lyse only targets from the same species; and work on a mutant line of cells carrying no glycolipid proteins has failed to show that these