Springer Semin Immunopathol (1984) 7: 143–162

Springer Seminars in Immunopathology

Initiation of Complement Activation

P. J. Lachmann and N. C. Hughes-Jones

MRC Unit on Mechanisms in Tumour Immunity, The MRC Centre, Hills Road, Cambridge CB2 2QH, England

Introduction

Although the complement system represents a formidably complex array of interacting proteins, at its centre there lies a single enzymatic reaction, namely the cleavage of C3 into its two primary fragments C3a and C3b. This reaction can be brought about by a number of enzymes but physiologically it is principally caused by the two homologous enzymes which are generated as a consequence of complement activation. These are the two C3 convertases of the classical and alternative pathways of complement activation, and it is the generation of these enzymes which is the culmination of the two principal pathways of complement activation. Both the classical and the alternative pathway are examples of what Macfarlane [29] has called the "triggered enzyme systems" of plasma, and the two pathways use in large part homologous proteins that are clearly gene duplicates of each other [24]. In spite of this similarity, however, the two pathways show fascinating differences in the strategies that are employed in the generation of their products, and in this chapter the initiation mechanisms of the two pathways are to be compared and contrasted.

Strategies of Triggered Enzymes Cascades

In its simplest form the triggered enzyme cascade can be pictured as the situation where the product of one reaction is the enzyme of the next. This may be described as the "enzyme generated" strategy. This strategy is used in a number of triggered enzyme systems of blood plasma and is found in the complement system, particularly in the classical pathway where the generation of $C\bar{1}$ – the enzymically active form of C1 – is the essential product of the initiation reaction and where its substrates, C4 and C2, are normally present in serum.

This is, however, not the only strategy employed in these triggered enzyme cascades. The second strategy is where the enzyme is present in fully active form in the circulation and it is the substrate that has to be generated – "substrate generated" strategy. This strategy requires the evolution of enzymes which resist the protease inhibitors that form so large a part of normal plasma protein and which must also have a highly restricted substrate specificity. This strategy is also used in the complement system, particularly in the reactions which break down C3b, the essential control mechanism of the alternative pathway. Here the enzyme concerned, Factor I, is present normally in plasma and its substrate C3b is generated during the complement activation.

There is yet a further strategy employed in the complement system where both the enzyme and its substrate are normally present in plasma, but what requires to be generated is a modifier which is required to react with the substrate before the enzyme can cleave it ("modifier generated" strategy). This again is seen in the alternative complement pathway where Factor D, the enzyme that cleaves Factor B from its zymogen form into its enzymatically active fragment Bb, is normally present in plasma, as is Factor B. However, the cleavage of Factor B by Factor D in physiological circumstances is wholly dependent on the prior combination of Factor B with C3b and it is only in this association that the reaction can take place. The C3b itself is not altered in the reaction and can be repeatedly reused. As will be discussed below this particular strategy, where substrate-modifying proteins are necessary before an enzyme can cleave its substrate, is used frequently in the complement system and is highly characteristic of it. It is quite possible that it evolved only once and that the various examples again represent gene duplicates. Table 1 shows examples of the various strategies and Table 2 gives a list of the complement reactions which employ the modifier strategy.

Initiation Mechanisms

In all these strategies there remains the question of how the initial molecules required are to be generated. If, for example, one discusses the generation of an enzyme by a previous enzymatic reaction there comes at the initial point the question of where do the original enzyme molecules come from to generate the first product. There are again three separate mechanisms for so doing which are used in the complement system. These are, first, the production of an "active zymogen". This describes the situation where the inactive precursor (zymogen) of an enzyme which normally requires to be activated by proteolytic cleavage can nevertheless acquire some enzymatic activity in the precursor state by virtue of a conformational

Table 1. Triggering strategies used in the complement system

Enzyme	Substrate	Modifier
Enzymes generated (with or without generated modifier)		
C1	[C4] (→C4b)	–
C1	[C2] (→C2b)	*C4b*
C4b2b	[C3] (→C3b)	–
C4b2b	[C5] (→C5b)	*C3b*
Substrate generated		
[I]	*C3b* (→iC3b)	[H]
[I]	*iC3b* (→C3c)	[CR1]
Modifier ge.erated		
[D]	[B] (→Bb)	*C3b*

Boxed factors are present in normal blood
Underlined factors require to be generated by previous triggered steps

Table 2. Complement reactions where substrate modification is needed for an enzymatic reaction to proceed

Enzyme	Substrate	Modifier	Requirement for modifier
C1	C2	C4b	Necessary when surface bound Enhances in free solution
FD	FB	C3b	Necessary
C4b2b or C3bBb	C5	C3b	? Necessary
FI	C3b	FH	Necessary in free solution Enhances when surface bound
FI	iC3b	CR1	Necessary
FI	C4b	C4 binding protein	Necessary

change. This system is well known in the contact-activatable protease systems where there is mutual zymogen activation of kallikrein and Hageman factor (Factor XII of the clotting cascade) when both are complexed on a negatively charged surface with high molecular weight kininogen. [50]. This appears to be the method employed in the classical pathway in the activation of C1r subsequent to the interaction of C1 with immune complexes, as will be discussed below.

The second initiation strategy is that described as the "tickover". The analogy here is with the motor car. If the engine is idling or "ticking-over" any movement of the throttle will accelerate the engine and cause the car to move. However, if the

engine is not ticking-over no amount of throttle movement will produce any effect. Similarly, in such mechanisms there is a continuously proceeding slow reaction sequence which is not sufficient to produce any observable effect but which can be accelerated ("fired") to produce measurable effects. This is the mechanism by which the alternative complement pathway works and will be described in detail below.

The third mechanism of initiation is to use as initiator some enzyme which is not intrinsic to the pathway under consideration. This is used in a number of circumstances and is seen in both complement pathways. For example, proteolytic enzymes derived from polymorphonuclear leukocytes or from bacteria, or occasionally from other enzyme systems such as the fibrinolytic system or the coagulation system, may all serve to activate the complement pathway. Examples will again be given below.

Activation of the Classical Complement Pathway

The initiation reaction of the classical pathway lies solely in the activation of C1 which is therefore described in detail.

Structure of C1

C1. The C1 molecule is a complex composed of three subcomponents: C1q, C1r, and C1s. Two molecules each of C1r and C1s combine firmly to form a C1r-C1s tetramer and each tetramer combines relatively loosely with one C1q molecule to form the C1 complex. The function of C1q is to bind the whole C1 complex to immunoglobulin or other activating surfaces; both C1r and C1s are proenzymes which, on activation, become converted to serine proteases and thus initiate the classical complement pathway.

C1q. The structure of C1q (Fig. 1) has been elucidated from electron microscopy studies [23, 46] and from amino acid sequencing [39]. The molecule (mol. wt. 410 000) is composed of six peripheral globular heads, each of which is joined by a connecting strand to a fibril-like central portion. The fibrils and central portion have a collagen-like structure. Electron microscopy studies [44] have shown that the most favoured angle in the splaying of the collagen strands is 100°, with the range extending from 40° to 160°.

C1r and C1s. Both C1r and C1s have similar molecular weights (about 83 000 daltons) and amino acid content; it is probable that they have arisen by gene duplication. Both molecules are in the proenzymic form and consist of a single polypeptide chain. Formation of the active serine esterase site results from a splitting of the molecules into heavy (56 000 daltons) and light chains (27 000 daltons), held together by a disulphide bond, the active site being present on the light chain. Electron microscopy of the purified C1r-C1s tetramer has shown it to be a linear structure of eight globular domains and thus it is presumed that each of the C1r and C1s molecules has two separate domains (Fig. 1). On the basis that C1r readily forms dimers whereas C1s does not, it is suggested that the two C1r molecules are on the inside of the chain [49]. Electron microscopy of the native C1 molecule indicates that the C1r-C1s subunit is bound to the collagen portion of C1q.

Binding and Activation of C1

In the early studies it was thought that C1 could be activated only following binding to immunoglobulin (Ig). Recently it has become clear that the classical pathway can be activated by certain viruses, bacteria, polyions, and other particles such as mitochondria through direct binding of C1 in the absence of antibody [8]. Whether the mode of activation is the same or different in the two circumstances is not known, but it is unlikely on general grounds that there are two separate mechanisms. Only binding and activation to antigen-immunoglobulin complexes will be discussed here.

It is well recognized that C1 is activated by immunoglobulin only when the latter is bound to antigen or aggregated in other ways. The fundamental questions are thus why monomer Ig in the plasma does not bind and activate C1, and what alterations in the immunoglobulin molecules follow attachment to antigens and result in C1 binding and activation. It is only with the formation of a firm bond between C1 and a potential activating particle that conditions are optimal for the subsequent activation process. The binding of C1q must take place through two of its globular heads in order to result in a firm bond and this divalent binding only takes place when immunoglobulin is bount to antigen. In plasma the reaction between C1q and Ig is monovalent and relatively weak and does not lead to activation. The mode of binding of C1q by IgM and IgG will be discussed separately.

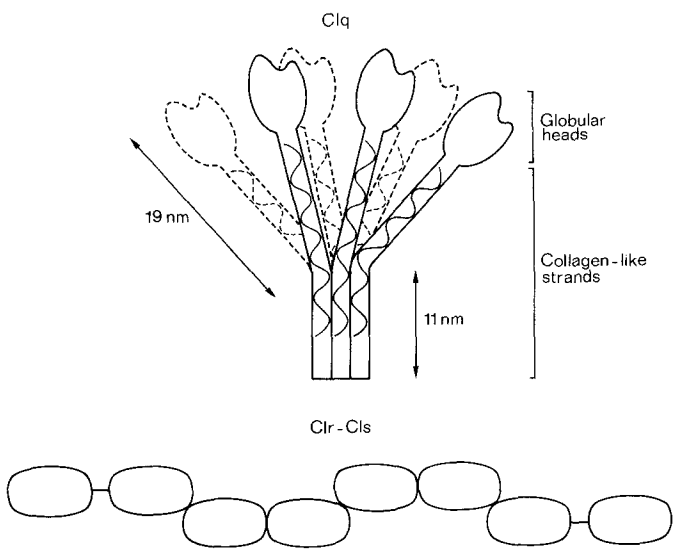

Fig. 1. Diagram of the structure of C1q (redrawn from Reid and Porter [38]) and the C1r-C1s tetramer (redrawn from Tschopp et al. [49]). Each C1r and C1s molecule has two domains and the outer two domains at each end are thought to be comprised of C1s molecules. The C1q and C1r-C1s components are drawn approximately to scale

Binding of C1q by IgG Antibodies

Two theories have been proposed to answer the question why IgG complexes bind C1q tightly but IgG free in solution does not do so. First, it is postulated that an allosteric conformational change takes place within the IgG molecule following the antibody-antigen reaction and that this results in the appearance of a binding site on the Fc piece. Much effort has been expended by the proponents of this theory and although conformational changes have been found, they are for the most part confined to the F(ab) fragments and no convincing change has been shown to take place in the Fc piece. It appears to be unlikely that the flexible hinge joining the F(ab) arms to Fc could relay an allosteric change from one domain to the other. Moreover, there is evidence that neither the strength of binding of C1q to IgG nor the rate of C1 activation is affected by the binding of the IgG molecule to antigen. Thus, it has been found that the binding of C1q to an IgG antibody whose binding site was completely occupied by hapten was exactly the same as when a related hapten occupying only 47% of the antibody-combining site was used [26]. Similarly, Tschopp et al. [48] have found that the rate of activation of C1 by IgG polymers was the same in the presence or absence of antigen.

On the other hand there is a considerable amount of evidence which is compatible with the view that monomer IgG has two pre-existing binding sites for C1q. The two binding sites are on opposite sides of the Fc piece and the C1q molecule is insufficiently flexible to allow binding by two of its globular heads to a single IgG molecule. The essential feature leading to C1q binding is the aggregation of IgG molecules so that they are sufficiently close together in the complex to allow two of the C1q heads to bind to separate IgG molecules.

The best evidence that monomer IgG free in solution has a binding site for C1q comes from the demonstration that IgG monomers can be shown to combine weakly with C1q. Estimates of the value of K have fallen in the range $1–5 \times 10^4 M^{-1}$ [14, 43], a binding so weak that the IgG concentration has to be about 5 mg/ml in order to demonstrate it with certainty. Not all IgG subclasses bind equally well, the order being IgG3, IgG1, IgG2, and IgG4, the latter showing a very low and not always detectable bond. The binding site for C1q is on the $C\gamma 2$ domain of the Fc piece, since antibody with the terminal $C\gamma 3$ domains removed is still capable of binding C1q whereas the $F(ab)_2$ fragments with no Fc piece lose their ability to bind [32]. Since IgG molecules are symmetrical, there must be two sites on each of the $C\gamma 2$ domains in the IgG molecule [35].

The nature and situation of the C1q binding site on the $C\gamma 2$ domain are still uncertain. It is clear that ionic groups are involved and probably contribute most of the binding energy. This follows from observations on the effect of ionic strength of the medium on the binding of C1q, a lowering of the ionic strength greatly increasing the strength of the bond. This accounts for the fact that complement assays have been found to be much more sensitive when carried out at low ionic strength. Estimations of the number of oppositely charged ionic pairs involved vary from 4–6 [7, 14]. There are two differing views on the position of the site on the Fc piece. There is evidence that the charged amino acids involved are two glutamic acids at positions 318 and 333, and two lysines at positions 320 and 322 [7]. On the other hand, evidence based on the reaction with small peptides derived from the C2

portion of IgG indicates that the positively charged residues histidine-285, lysine-288 and -290, and arginine-292 are involved [28]. Both of these sites on the Fc piece are situated near the hinge region (Fig. 2), which is consistent with the finding that a myeloma IgG1, which lacks the hinge region, does not bind C1, presumably due to steric hindrance by the F(ab) arms [22].

There is very good evidence from a variety of sources that each C1q molecule must bind to two IgG molecules in order to bind C1 firmly enough for C1 activation to take place. The first piece of evidence for this came from an electron microscopy study of complement-lysed sheep red cells by Humphrey and Dourmashkin [18]. They studied the relationship between the number of "holes" in the red cell membrane and the amount of antibody bound to the cell. Using an IgM antibody, it was found that 2–3 molecules could produce one "hole" and that this ratio was constant as the amount of antibody bound to the cells was increased. In contrast, about 1000 IgG molecules had to be bound in order to produce a single "hole"; moreover, as the amount of antibody was increased above this value, the number of holes increased exponentially. A similar relationship between C1 uptake and the amount of antibody bound was found by Borsos and Rapp [3]. The interpretation of these data is that a single C1q molecule has to bind to a pair of IgG molecules which are sufficiently close together to allow the C1q to span between them, the greater the density of antibody molecules, the greater the chance of finding suitable pairs of binding sites.

The other evidence for bivalent binding of C1q comes from studies on the strength of binding of C1q to antigen-antibody complexes. The binding of C1q to

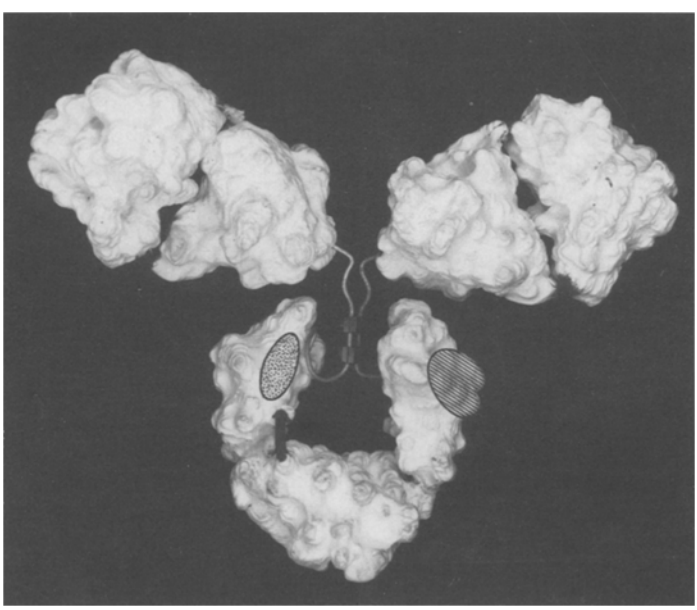

Fig. 2. A model of an IgG molecule showing the suggested sites for the C1q-binding site; on the left, that of Emanuel et al. [7] and on the right that of Lukas et al. [28]

monomer IgG is a weak one, being of the order $5 \times 10^4 M^{-1}$. Binding to complexes on the other hand is much firmer and two sets of values can be obtained, one of about $5 \times 10^7 M^{-1}$ and the other at about $1 \times 10^{10} M$. The most reasonable explanation of the results is that C1q binds to monomer IgG through a single binding site on one of its heads but that binding to IgG complexes takes place either by two heads ($5 \times 10^7 M^{-1}$) or by three heads ($1 \times 10^{10} M^{-1}$) [14]. Support is given to this interpretation by the finding that isolated globular C1q "heads" contain the binding site and that they bind to IgG complexes with the same weak binding constant ($5 \times 10^4 M^{-1}$) as that seen between C1q and monomer IgG [15]. The single-headed binding is so weak that the average duration of the bond is considerably less than one second, too short a time for C1 activation to be initiated. Hence C1 activation does not occur to any extent in plasma as a result of the reaction between C1q and IgG monomers, although the concentrations of C1q and IgG are such that over half the C1q is bound to IgG at any given time. On the other hand, divalent binding by two heads has an average duration of several minutes and is sufficient for activation of the associated C1r-C1s tetramer. The requirement for firm two-headed binding over an extended period is thus a device to prevent activation within the plasma by the pre-existing C1q-binding site present on monomer IgG.

The requirement for two-headed binding of C1q could partly explain the failure of certain antibodies to activate complement. The outstanding example of this is the human blood group antibody anti-D. The number of D-antigen sites on red cells is known to vary between 10 000 and 30 000 [40]. It is probable that the D-antigen is distributed at random on the red cell surface and that a few pairs of molecules will be present by chance within 30 nm of each other, the distance that can be spanned by a C1q molecule. This has proved to be the case, as it has been shown that there are as many as 500–1000 sites available for C1q binding when about 10 000 anti-D molecules are present on a red cell surface [16]. The low density of potential C1 binding sites could be partly responsible for the failure to be lysed by complement, but as there are some sites present there must also be either a failure of bound C1 to activate, or an insufficient number of activated C1 molecules present on the cells to result in an adequate production of the C5-9 membrane attack complex.

Another phenomenon which is dependent on divalent binding of C1q is the synergistic cooperation which is seen between antibody molecules in their ability to bind and activate C1. This has been demonstrated using monoclonal antibodies specific for MHC antigens present on rat red cells [17]. These monoclonal antibodies when used singly are poor at bringing about complement-mediated lysis; usually not more than about 20% lysis occurs even though almost all the antigen sites are saturated by using an excess of antibody. Complete lysis can be brought about, however, by the addition of a second monoclonal antibody which is specific for a different epitope on the MHC antigen. The presence of two antibodies at the same time probably results in the formation of tetramers composed of two antigen and two antibody molecules combined as a closed circle, and also in the formation of short chains or catenars composed of alternating antibody and antigen molecules. When a single antibody is used the molecules are probably distributed randomly on the red cell surface, being attached either monovalently to one or bivalently to two antigen molecules. This random distribution results in only a few

IgG molecules being sufficiently close to allow C1q binding, but the circle and catenary formations occurring when two monoclonals are combined bring many antibodies close together. This "aggregation" of IgG has three effects: (1) there is an increase in the number of C1q binding sites on each cell, (2) there is an increase in the binding constant for C1q as it allows a third head to bind, and (3) there is an increase in the rate of activation of C1 by a mechanism as yet unknown. The net result of these three factors is an increase in the number of activated C1 molecules on the red cell surface, leading to an increase in the extent of the lysis. This synergistic mechanism almost certainly has an important physiological role in improving the efficiency of complement activation as it is a reasonable assumption that polyclonal antisera will contain antibodies against different epitopes on the same surface antigen, whether the cells involved are bacteria, viruses, parasites, or red cells.

Binding of C1q by IgM Antibodies

As with IgG, IgM cannot bind and activate C1 when both are present free in solution. The present evidence suggests that monomer IgM has only one binding site available for C1q binding. In contrast to IgG, further C1q binding sites appear on the IgM molecule following binding of IgM to antigen.

When in solution, electron microscopy has shown that the IgM molecule is present in a star-shaped and planar form [10]. Studies with ^{125}I-labelled C1q indicate that IgM in this form binds weakly with C1q, the binding constant being of the order of $5 \times 10^5 M^{-1}$, a bond so weak that any C1q-IgM complex formed will have a half-life of only a few seconds. This indicates that only one binding site is involved, since association by two C1q heads would be much stronger. If there is only one site available on the IgM molecule, this implies that the IgM molecule, which is composed of five 7S subunits, is asymmetrical. The binding of C1 to IgM in plasma is therefore very weak and only transient so no activation of C1 can take place [11].

There are two ways in which an IgM molecule can combine with antigen. It can remain in the star-like or planar form and cross-link antigenic sites on separate particles. This conformation is seen especially in antigen excess. The alternative conformation is seen when IgM binds to several antigen molecules on the same particle. In this form of binding, the $F(ab)_2$ units are dislocated and the molecule has the appearance of a "staple" [10]. This form is most frequently seen in antibody excess.

Of these two forms, it is the dislocated "staple" form which is the more potent activator of C1. The evidence for this comes from several sources. First, there is the evidence from electron microscopy (EM) of IgM-dextran complexes. Only complexes made in antibody excess activate C1 and EM shows that most of the IgM is in the staple form. In antigen excess, the planar form predominates and C1 activation is negligible [11]. Second, there is the evidence from multivalency. C1 activation is greater when IgM is bound to liposomes containing a high density of hapten than when a low density is present [4, 19]. The supposition is that with a high hapten density more of the bound IgM can assume the staple form by binding through at least two of its $F(ab)_2$ arms. It has also been found that IgM bound bivalently to a ligand will activate C1, whereas monovalent binding does not [20].

What then are the changes that occur in the IgM molecule on assuming the staple form that allow the binding and activation of C1? First, there is the appearance of at least one further site for the binding of C1q heads. This is indicated by increase in the binding constant for C1q to a value of about $5 \times 10^7 M^{-1}$, which is consistent with binding by two heads. Rosse et al. [42] using an elegant technique involving electron bombardment showed that there are probably at least three potential C1q sites on an IgM molecule and that at least two are required for C1 activation. Based on the evidence from C1q binding to fragments of IgM, the most likely site for binding is on the Cμ3 or Cμ4 domains. The binding sites for C1q do not appear to be present on the Fc_5 core of IgM after removal of the $F(ab)_2$ arms, and thus the most likely explanation for their appearance is that the dislocation of The $F(ab)_2$ arms on assuming the staple form distorts the Fc_5 core and results in the appearance of the binding sites [11].

The changes associated with the staple form that give rise to the activation of C1 which follows binding are not at all clear, mainly due to the fact that the necessary stimulus for activation is still unknown. Binding of C1 alone is an insufficient stimulus as binding to IgM-dextran complexes in antigen excess takes place by two heads but does not lead to activation [11]. Nor does activation take place when C1 is bound to 7S IgM subunits. The findings are compatible with the suggestion that a binding site for C1r-C1s on IgM staples is required for activation (see next section), but other explanations are possible.

Activation of C1

The precise mode of activation of the C1 molecule following its binding to immune complexes is still unknown. There are two characteristics of C1 activation which must be made to fit into any theory. First, the binding of C1 and its subsequent activation are two separate processes. The separation of the two processes was first clearly shown by Colten et al. [6] who found that up to 80% of C1 bound to IgG-coated erythrocytes could remain unactivated. Another instance is the finding that C1 will bind to glutaraldehyde-aggregated IgG just as firmly as to immune complexes, but that activation does not occur [12]. Second, the rate of activation is a relatively slow process and shows a considerable amount of variation in speed depending on circumstances. Activation may be more than 70% complete in 8–12 min [17] or may take 2–3 hours to approach completion [37].

It is known that activation of C1r is the first step in the activation process and that activated C1r then acts on C1s [53]. Two types of mechanism can then be considered in C1 activation. First, there is the possibility that a conformational change takes place in C1q on binding to complexes and as a result of the close contact between C1r and C1q, a conformational change then takes place in C1r which reveals the active enzyme site. Second there is the possibility that C1r can autoactivate. It must then be postulated that when present within the C1 molecule free in solution, autoactivation is very slow or does not occur at all; following binding of C1 to complexes, conditions may be altered allowing autoactivation to take place.

There is some recent evidence that C1q undergoes a conformational change on binding, as it has been possible to produce monoclonal antibodies which only react

with C1q when the latter is bound to complexes and thus has revealed neoantigens [27]. However, there is no direct evidence as yet that this conformational change is responsible for C1r-C1s activation. On the other hand, there is now good evidence that C1r is able to autoactivate without contact with other molecules [5]. When purified, C1r exists as a dimer and autocatalytic activation takes place within the dimer. Many enzymes in their proenzyme form possess some enzymic activity, i. e. they contain some "active zymogen", although the amount is small when compared to their activated state. When intradimer activation takes place, it is presumed that one member of the pair in the proenzymic form is able to split the other member and that the latter, having been activated, then splits the former. This mode of activation implies that there is considerable flexibility at the bond between the members of the dimer and that, with random motion due to thermal agitation, the enzymic site on one molecule can contact the site to be split on the other molecule.

If the C1r-C1s tetramer can autoactivate when C1 is bound to complexes, then there must be some mechanism which accounts for the fact that activation is non-existent or negligible when C1 is free in plasma but rapid following binding. Three possibilities may be considered. First, it is possible that C1 inhibitor prevents activation under these circumstances. C1 inhibitor has been shown to prevent self-activation of purified C1 and even under certain circumstances to inhibit the activation of C1 when the latter is combined with complexes [12, 51]. Combination of the C1 molecule with complexes may thus prevent access by the C1-inhibitor in much the same way as the C3 convertase is activated in the alternative pathway. A second possible mechanism is based on the observation that 5mM Ca^{2+} will inhibit the autoactivation of C1r-C1s tetramers. The binding of C1q could result in the local sequestration of calcium thus allowing activation to take place [33].

A third suggestion for the mechanism of activation is that the C1r-C1s tetramer requires to be bound to another site in addition to the binding to C1q. This suggestion was first made by Goers et al. [13] and was based on observations of C1 activation by a dinitrophenol-lysine-antibody complex. They found that the C1r-C1s tetramer could bind to the DNP-polylysine hapten and also that the strength of binding of C1 to the hapten-antibody complex was ten-fold higher than the binding of C1q, suggesting that the C1r-C1s tetramer is attached both to C1q and to the complex. Additional evidence was given by Bartholomew and Esser [2] who found that mouse leukaemia virus will bind and activate human C1 and that this correlated with the binding of C1 through both the C1q and C1s subunits. In contrast, guinea pig C1s did not bind to the retrovirus, nor was guinea pig C1 activated. These experiments certainly establish that the binding of C1r-C1s to an additional site on a C1-activator does lead to activation. The main question which remains is whether this mode of activation is also the normal one when C1 is bound to immune complexes.

The binding between the C1q and C1r-C1s components of C1 is only moderately firm, the value of the binding constant being of the order of 3×10^7–$6 \times 10^7 M^{-1}$ [52, 45]. An increase in the strength of this bond when C1 is activated by immune complexes also supports the suggestion that the C1r-C1s has to bind to complexes in order to be activated [52]. IgG complexes aggregated by glutaraldehyde bind C1 but do not activate it. Under these conditions, the strength of the bond between C1r-C1s and the C1q was only slightly higher than the bond between C1r-C1s and

C1q when free in solution. In contrast, when C1 was bound to strongly activating ovalbumin-IgG complexes, there was a 10–30 fold increase in the strength of the bond between C1r-C1s and C1q. The findings could be explained if the binding of C1r-C1s was stronger when the tetramer was bound to a site on C1q and at the same time to a site on the complex, than when the tetramer was bound to C1q alone. Furthermore, it was found that in the presence of a considerable molar excess of C1q in relation to C1r-C1s, there was a considerable increase in the rate of C1 activation. It is suggested that under these circumstances, the C1q molecules bound to the complex are in sufficiently high density for the C1r-C1s tetramer to span between two neighbouring molecules and that this configuration of C1r-C1s leads to rapid activation.

A tentative model has been put forward for C1 activation based on these findings [52]. If the C1r-C1s tetramer has the two C1s molecules at either end of the linear chain, then it is probable that there are two binding sites available on the polymer. In solution the tetramer may be bound to C1q by both binding sites and in this configuration the relative spatial arrangement of the two C1r molecules does not allow autocatalytic activation. Following binding of C1 to complexes, one of the C1r-C1s binding sites may detach from C1q and become attached to a site on the complex, or to a neighbouring C1q molecule if it were available. It is suggested that in this new configuration of the tetramer, the C1r molecules can interact and eventually split each other.

This model will explain three of the characteristic features of C1 activation, namely (1) that binding of C1 is not always followed by activation, (2) that activation is a relatively slow process, and (3) that the rate of activation is variable under different conditions. The model indicates that the rate of activation is dependent on the density of binding sites for C1r-C1s. No activation could take place in the absence of these binding sites, and the maximum rate would be achieved when there was a suitable site for C1r-C1s adjacent to each C1q binding site. The maximum rate of activation would then be dependent on the time taken for one of the two C1r-C1s bonds to C1q to dissociate and become attached to the new binding site on the complex or other activating surface. This dissociation and reassociation could be the rate limiting step and is relatively slow. The variability in the rate of activation would be explained by variation in the density of C1r-C1s binding sites. With lower densities of C1r-C1s binding sites, activation would be slower as C1r-C1s molecules at nonactivating sites would have to dissociate and then reassociate at a more favourable site.

Activation of the Alternative Complement Pathway

The discovery that there were methods of activating the complement system distinct from those involving the activation of C1 by antigen-antibody complexes is already inherent in the original discovery of 'C'3'. C'3 was defined as the component of complement that was inactivated by yeast or by cobra venom factor treatment of serum, neither of which removed either mid-piece (C1 in modern parlance) or end-piece (C2). The essential correctness of this observation had to await studies of sera with isolated deficiencies of C1, C4, and C2, all of which again confirm that there is a group of activators that are capable of depleting the late-acting complement

components, C3–C9, without affecting the early components of the classical pathway. The activators that are capable of doing this were originally particulate polysaccharides of which yeast cell walls or zymosan was the canonical example but agar, agarose, particulate inulin, and a variety of similar particulate polysaccharides have the same effect. To this group of compounds, however, have been added many others over the years. These fall into a number of groups. Beside the particulate polysaccharides there are a number of mammalian cells that share the same property. Thus rabbit erythrocytes are one of the most potent activators of the human alternative pathway. Rabbit erythrocytes do not, however, activate the alternative pathway in rabbit serum and this demonstrates that the activation of the alternative pathway is not completely nonspecific but involves a sufficient degree of recognition to distinguish complement of one species from another. However, cells infected with measles or influenza viruses are capable of activating the alternative pathway by their own species of complement as well as others. Cells transformed by the Epstein-Barr virus and a variety of malignant cell lines behave similarly as do almost all parasites. A third important group of alternative pathway activators are antibodies. It is interesting that the capacity of antibodies to activate the alternative pathway is not subclass restricted in the same way as is the case with C1 activation and furthermore that $F(ab')_2$ fragments activate as well as IgG. However, it is again necessary for the aggregates to be particulate and it is preformed antigen-antibody precipitates that activate the alternative pathway rather than soluble complexes. It is a little difficult to see what structural features this heterogeneous group of activators have in common. One factor is certainly their particulate nature. This is particularly clear with a substance like inulin where the soluble material is absolutely noncomplement activating. It has been suggested that the absence of sialic acid is of importance [21] but it is clearly not true to say that all particulate matter free of sialic acid activates the alternative pathway. Bacterial flagella are a good exception. They will fix complement but only in the presence of IgM antibodies. The mechanism by which this group of activators work will be returned to later.

Besides working in the absence of C1, C4, and C2, the alternative pathway has a number of other characteristics. Since it does not utilize C1, it has no requirement for calcium ions but it does need magnesium ions for the combination of C3 with Factor B. For this reason in whole serum and in the absence of the classical pathway, the alternative pathway can be activated in the presence of magnesium EGTA, a chelating agent that leaves free magnesium but no free calcium in solution.

The alternative pathway has several factors peculiar to itself. The alternative pathway analogue of C2 is Factor B. This protein is even more heat labile than C2, at least in human serum, and by very critical heating (50 °C for 20 min) it is possible to make a reagent which no longer supports alternative pathway activation but will still support classical pathway activation, albeit not as well as an unheated serum. The protein which in the alternative pathway serves some (but not all) of the functions of C1 is Factor D. This has the distinction of being far and away the smallest of the complement proteins (mol. wt. 25 000) and this again provides a method for distinguishing alternative and classical pathway activation. Serum that has been excluded from G75 Sephadex has all the complement components except Factor D. It therefore allows activation of complement by the classical pathway but

its alternative pathway activity is abrogated. A third protein that the alternative pathway requires for optimal activity is properdin. This has no obvious homologue in the classical pathway and its role will be considered again below.

The Reaction Sequence of the Alternative Pathway

In its essential outline the reaction sequence for the generation of the alternative pathway C3 convertase is clearly homologous with that of the classical pathway C3 convertase (Fig. 3). The essential difference is that C3b, the product of the reaction catalysed by the C3 convertase, forms a subunit of the C3 splitting enzyme itself. This means that whereas in the classical pathway the generation of C4b is the product of the initiation reaction, in the case of the alternative pathway the analogous product C3b is the product of the C3 convertase itself, and the alternative pathway is therefore a positive feedback amplification loop rather than a linear cascade. Furthermore the enzyme that splits C3bB (the zymogen) to C3bBb (the active enzyme) is Factor D and this is a further contrast with the classical pathway, since Factor D unlike $C\bar{1}$ does not require to be generated and is found in fully active form in plasma. Once $C\overline{3b}$ is generated, therefore, there is no need for the generation of any other active product to allow the alternative pathway to cycle until either all the available C3 or all the available Factor B is depleted. The reason that the amplification loop does not run to exhaustion whenever any C3b is generated is that there also exists in plasma an efficient system for destroying C3b to a form where it can no longer bind Factor B. This process is proteolytic and involves the cleavage of C3b by Factor I. Factor I is another proteolytic enzyme normally present in fully active form in plasma whose sole known substrates are C3b and C4b. Factor I, however, cleaves neither C3b nor C4b unless these proteins are complexed with

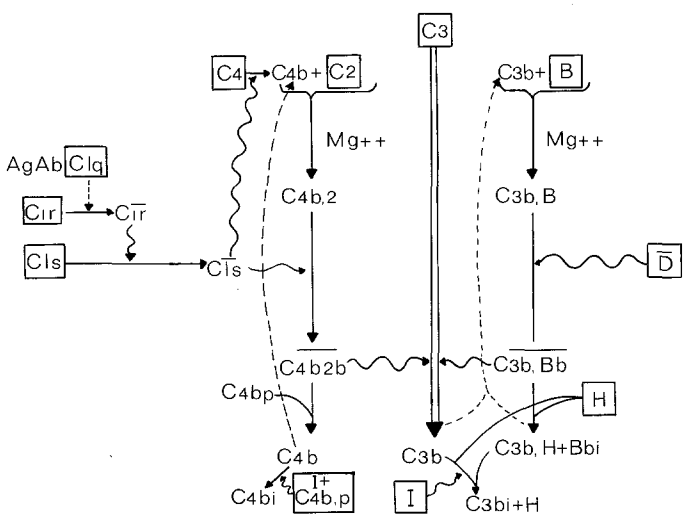

Fig. 3. A comparison of the classical and alternative pathways based on the central role of C3

modifiers. In the case of C4b this modifier is the C4 binding protein which allows cleavage at two sites in the chain to give C4d and C4c. In the case of C3b the situation is more complicated. When C3b is bound to Factor H, a normally occurring serum protein, Factor I, cleaves the alpha chain in two sites in close proximity to each other, eliminating a small peptide, C3f (mol. wt. about 3000). The resulting large molecular weight product is now called iC3b and no longer has the capacity to bind Factor B or C5 or further Factor H. This iC3b fragment, when cell bound however, exposes a carbohydrate which can bind to bovine conglutinin and it was in fact this property of Factor I to render bound C3b liable to conglutination, by which it was first described as the "conglutinating activating factor" [25]. iC3b can, however, be split further by Factor I to C3d, g and C3c, but for this purpose the iC3b requires to be bound to the C3b receptor (CR1) found on primate red cells, platelets of many species of non-primate mammals, and on the polymorphs and monocytes of most mammalian species [30, 31, 41]. However, as far as the regulation of the alternative pathway is concerned, it is the competition between two opposing reactions which controls the rate at which the alternative pathway progresses. On the one hand, there is the combination of C3b with Factor B and the subsequent splitting by Factor D; on the other hand, there is the binding of C3b by Factor H and the subsequent splitting by Factor I to iC3b. This is shown schematically in Fig. 4. That this is indeed the way the system works was first demonstrated by an experiment of nature. A patient was described who had a genetic deficiency of Factor I [1] and his serum contained low levels of C3 all of which was present as C3b. Factor B was present at very low levels and largely as Bb + Ba. When further Factor B was added to his serum it was rapidly cleaved to Bb + Ba. Exactly similar effects could be produced by the immunochemical depletion of Factor I from normal serum [34]. It is of particular interest that when Factor I is removed immunochemically under conditions where no complement activation should occur (using $F(ab')_2$ anti-I at 0 °C, or in the presence of EDTA) that C3 cleavage starts immediately the temperature is raised or magnesium ions are re-added. This led to the idea that there was no initiating event required to fire the pathway in the absence of one of the control proteins and that the system must therefore be "ticking-over" as described earlier.

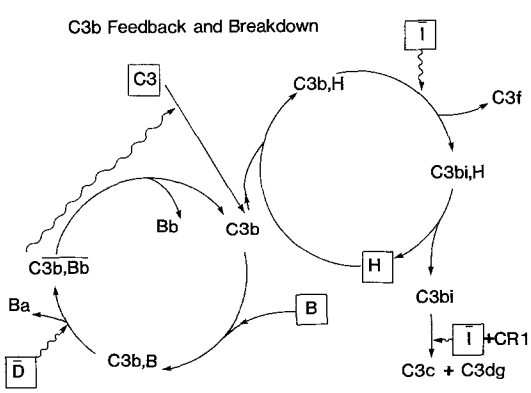

Fig. 4. C3b feedback and breakdown. The progression of the alternative pathway is determined by the rate at which C3b progresses to C3bBb or to C3bi

If the alternative pathway is seen as a positive feedback amplification system controlled by the relative speeds of the feedback cycle on the one hand and the C3 destruction cycle on the other, it can readily be predicted that the system will be activated by those factors that accelerate C3b production on the one hand or those that limit its destruction on the other. These are listed separately in Table 3.

Mechanisms of increased C3 production include enzymes that are exogenous to the alternative pathway itself. Of these clearly the most important is the classical pathway C3 convertase, C4b2b, and indeed the classical pathway can quite fairly be regarded as a gene duplicate of the alternative pathway that provides a powerful stimulus to the system and which also provides the interface between the complement system and antibody-mediated reactions. Other exogenous enzymes may also be important, such as plasmin, polymorph elastase, and perhaps other polymorph enzymes. Various bacterial proteases can all slit C3b (or a fragment sufficiently like C3b to be able to bind Factor B). Mechanisms that stabilise the C3bBb enzyme and prevent its dissociation by Factor H or its natural instability will also enhance C3b formation. This is the physiological function of properdin, and stabilisation of C3bBb by properdin is necessary under physiological conditions in order to get a firing of the alternative pathway. There is also an autoantibody to C3bB which stabilises C3bBb. This autoantibody which is usually called nephritic factor (NeF) is found in association with partial lipodystrophy, the dense deposit form of mesangiocapillary glomerulonephritis, and occasionally in systemic lupus erythematosus and causes activation of the alternative pathway in vivo. Unlike the patients with Factor I deficiency when C3 levels fall to zero, patients with nephritic factor stabilise at low C3 levels. Low C3 levels are due not only to hypercatabolism but also to the depression of C3 synthesis by the late products of C3 metabolism. When the C3 levels are sufficiently low that even the stabilisation by NeF fails to make the feedback cycle fire, the catabolic rate of C3 falls to approximately normal values and Factor B levels also return to normal. An analogous but unphysiological situation can be produced by the administration of cobra venom factor to mammalian plasma.

Table 3. Initiation of alternative complement pathway. Tick-over maintained by spontaneous hydrolysis of C3 internal thioester bond to 'C3b-like' molecule

Tick-over "fired" by
1. Mechanisms that increase C3b production
a) 'Exogenous' C3 splitting enzymes eg $\overline{C4b,2b}$
Plasmin
Leucocyte proteases eg elastase
Bacterial proteases
b) Stabilisation of $\overline{C3,Bb}$
Properdin
Nephritic Factor
$\overline{CVF,Bb}$
2. Mechanisms that reduce C3b destruction
a) Fixation on "protected surface"
b) Deficiency of Factor H and Factor I
c) Local sequestration of Factor H

Cobra venom factor behaves functionally just like C3b except that it is capable of taking part only in the feedback cycle and not in the destructive cycle, i. e. cobra venom factor can bind mammalian Factor B but fails to bind mammalian Factor H and is thus not destroyed. Both antigenically and in its protein structure, cobra venom factor resembles C3b and it seems to represent cobra C3b which for one reason or another is present in cobra venom in the absence of the cobra control proteins.

Mechanisms that reduce the destruction of C3b include the genetic deficiency of Factor I already described. More recently genetic deficiency of Factor H has been described [47], which although it is not complete is sufficient to decompensate the C3b destructive cycle and lead to a physiological state analogous to Factor I deficiency. It has also been found that local sequestration of Factor H on polyanions may cause local activation of the alternative pathway.

None of these mechanisms, however, so far explain the activity of the common activators of the alternative pathway: particulate polysaccharides, parasites, and virus-infected cells. The mechanism by which they destabilise the C3b feedback cycle was first described by Fearon and Austen [9] who pointed out that C3b fixed on certain surfaces is hindered in its reaction with Factor H compared to its capacity to react with Factor B. This capacity of the physical environment of the fixed C3b to affect its ligand reactivity is extremely interesting and the molecular mechanism involved has still not been elucidated. Whether the protected surface has some secondary binding site for C3b that hinders Factor H attachment or whether unprotected surfaces actually enhance Factor H attachment, is not wholly clear.

The remaining question about the nature of the tickover is what maintains it? It was originally suggested that it was maintained by the spontaneous generation of C3b from C3 and this indeed turns out to be essentially correct, although the mechanism involved is not the proteolysis originally believed to be responsible. It has more recently been shown that C3i, the C3 molecule formed by the spontaneous

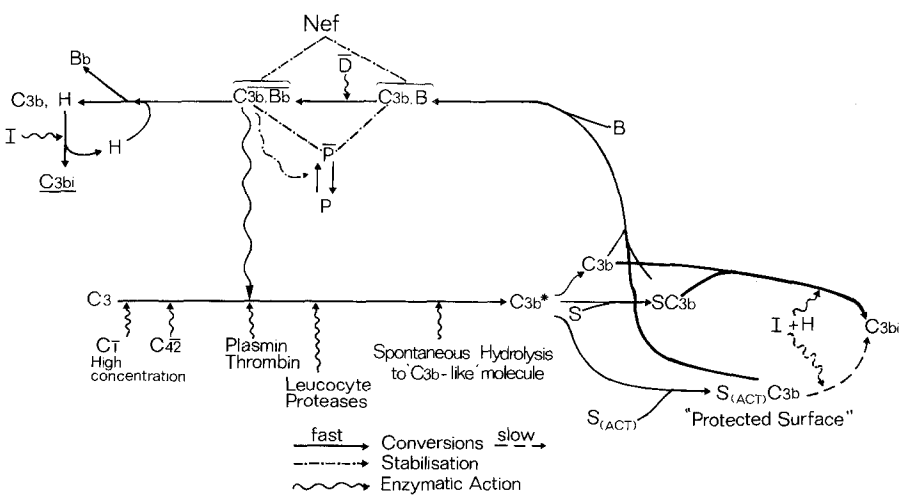

Fig. 5. The C3-tickover cycle

hydrolysis of the thiolester bond, is C3b-like in so far as it can bind Factor B to give rise to a C3-cleaving enzyme and Factor H in its ability to be subsequently cleaved by Factor I [36]. It is this slow spontaneous hydrolysis of C3 to C3i which provides the driving force of the tickover and enables the other factors already listed to activate the alternative pathway. A pictorial representation of the tickover is given in Fig. 5. It is an example of great evolutionary economy that the internal thiolester bond of C3 is not only of great importance in the formation of the metastable binding site and the covalent attachment of C3 to surfaces, but is also utilised as a method of generating the tickover which controls the alternative pathway.

Summary

The complement system is an excellent one to look at as an example of a triggered enzyme system. It is found to use almost all the strategies of enzyme cascades and for initiation which were discussed in the introduction. The classical pathway is a fine example of an active zymogen form of activation and the alternative pathway is the canonical example of a tickover. Both may also be activated by exogenous enzymes from other systems. The strategies of enzymes waiting for substrates, of substrates waiting for enzymes, and of both waiting for modifying proteins are all seen in various stages of the reaction. There is an association of a linear cascade with the positive feedback amplification loop. This degree of evolutionary adaptation is not only aesthetically pleasing but must be taken to mean that the system is of considerable biological importance and has been over a long evolutionary time span.

References

1. Abramson N, Alper CA, Lachmann PJ, Rosen FS, Jandl JH (1971) Deficiency of C3 inactivator in man. J Immunol 155: 96
2. Bartholomew RM, Esser AF (1980) Mechanism of antibody-independent activation of the first component of complement (C1). Biochemistry 19: 2847
3. Borsos T, Rapp HJ (1965) Complement fixation on cell surfaces by 19S and 7S antibodies. Science 150: 505
4. Borsos T, Chapuis RM, Langone JJ (1981) Distinction between fixation of C1 and the activation of complement by natural IgM antihapten antibody: effect of cell surface hapten density. Mol Immunol 18: 863
5. Colomb MG, Bensa JC, Villier CL, Arlaud GJ (1982) Aspects biochimiques de l'activation intrinsique de C1. Ann Immunol (Paris) 133C: 155
6. Colten HR, Borsos T, Rapp JR (1969) Titration of the first component of complement on a molecular basis: suitability of IgM and non-suitability of IgG haemolysins as sensitizers. Immunochemistry 6: 461
7. Emanuel EJ, Brampton AD, Baston DR, Dwek RA (1982) Formation of complement subcomponent C1q-immunoglobulin G complex. Biochem J 205: 361
8. Esser AF (1982) Interactions between complement proteins and biological and model membranes. In: Chapman D (ed) Biological membranes. Academic Press, London, p 277
9. Fearon DT, Austen KF (1977) Activation of the alternative pathway with rabbit erythrocytes by circumvention of the regulatory action of endogenous control proteins. J Exp Med 146: 22

10. Feinstein A, Beale D (1977) Models of immunoglobulins and antigen-antibody complexes. In: Glynn LE, Steward MW (eds) Immunochemistry, and advanced text book. Wiley, Chichester, p 263
11. Feinstein A, Richardson NE, Gorick BD, Hughes Jones NC (1983) Immunoglobulin M conformational change is a signal for complement activation. In: Celada F, Shumaker VN, Sercarz EE (eds) Protein conformation as an immunological signal. Plenum Press, New York, p 47
12. Folkerd EJ, Gardner B, Hughes-Jones NC (1980) The relationship between the binding ability and the rate of activation of the complement component C1. Immunology 41: 179
13. Goers JW, Ziccardi RJ, Schumaker VN, Glovsky MM (1977) The mechanism of activation of the first component of complement by a univalent hapten-IgG antibody complex. J Immunol 118: 2182
14. Hughes-Jones NC (1977) Functional affinity constants of the reaction between ^{125}I-labelled C1q and C1q binders. Immunology 32: 191
15. Hughes-Jones NC, Gardner B (1979) Reaction between the isolated globular sub-units of the complement component C1q and IgG complexes. Mol Immunol 16: 697
16. Hughes-Jones NC, Ghosh S (1981) Anti-D-coated Rh-positive red cells will bind the first component of the complement pathway, C1q. FEBS Lett 128: 318
17. Hughes-Jones NC, Gorick BD, Howard JC (1983) The mechanisms of synergistic complement-mediated lysis of rat red cells by monoclonal IgG antibodies. Eur J Immunol 13: 635
18. Humphrey JH, Dourmashkin RR (1965) Electron microscopy studies of immune cell lysis. In: Wolstenholme GEW, Knight J (eds) Complement. Churchill, London, p 175
19. Humphries GMK, McConnell HM (1977) Membrane-controlled depletion of complement activity by spin-label specific IgM. Proc Natl Acad Sci USA 74: 3537
20. Karush F, Chua M-M, Rodwell JD (1979) Interaction of a bivalent ligand with IgM anti-lacton antibody. Biochemistry 18: 2226
21. Kazatchkine MD, Fearon DT, Austen KF (1979) Human alternative pathway: membrane associated sialic acid regulates the competition between B and B1H for cell bound C3b. J Immunol 122: 75
22. Klein M, Haeffner-Cavaillou N, Isenman DE, Rivat C, Mavia NA, Davies DR, Dorrington KJ (1981) Expression of biological effector functions by immunoglobulin G molecules lacking the hinge region. Proc Natl Acad Sci USA 78: 524
23. Knobel HR, Villiger W, Isliker H (1975) Chemical analysis and electron microscopy studies of human C1q prepared by different methods. Eur J Immunol 5: 78
24. Lachmann PJ (1979) An evolutionary view of the complement system. Behring Inst Mitt 63: 25
25. Lachmann PJ, Müller-Eberhard HJ (1968) The demonstration in human serum of "conglutinogen-activating factor" and its effect on the third component of complement. J Immunol 100: 691
26. Liberti PA, Bansch DM, Schoenberg LM (1982) On the mechanism of C1q binding to antibody. I. Aggregation and distribution of the IgG versus combining site-transmitted effects. Mol Immunol 19: 143
27. Loos M (1982) Antibody-independent activation of C1, the first component of complement. Ann Immunol (Paris) 133: C 165
28. Lukas TJ, Munoz H, Erickson BW (1981) Inhibition of C1-mediated immune haemolysis by monomeric and dimeric peptides from the second constant domain of human immunoglobulin G. J Immunol 127: 2555
29. Macfarlane RG (1969) The blood clotting mechanism. The development of a theory of blood coagulation. Proc R Soc London Ser B 173: 261
30. Medicus RG, Arnaout MA (1982) Release of C3c from bound C3bi by C3b inactivator. Mol Immunol 19: 1386.
31. Medof ME, Iida K, Mold C, Nussenzweig V (1982) Unique role of the complement receptor CR1 in the degradation of C3b associated with immune complexes. J Exp Med 156: 1739
32. Meyer CM, Colomb MG (1976) Use of an IgG fragment prepared with particulate plasma to study C1 binding and activation. FEBS Lett 66: 132
33. Naff GB, Ratnoff OD (1968) The enzymatic nature of C1r. J Exp Med 128: 571
34. Nicol PAE, Lachmann PJ (1973) The alternative pathway of complement activation: the role of C3 and its inactivator (KAF). Immunology 24: 135
35. Painter RH, Foster DB, Gardner B, Hughes-Jones NC (1982) Functional affinity constants of subfragments of immunoglobulin G for C1q. Mol Immunol 19: 127
36. Pangburn MK, Schreiber RD, Muller-Eberhard HJ (1981) Formation of the initial C3 convertase of the alternative pathway. Acquisition of C3b-like activities by spontaneous hydrolysis of the putative thiolester bond in native C3. J Exp Med 154: 856

37. Price JW, McConnell HM, Bartholomew RM, Esser AF (1980) Kinetics of antibody-dependent activation of the first component of complement on lipid bilayer membranes. Biochem Biophys Res Commun 93: 235
38. Reid KBM, Porter RR (1976) Subunit composition and structure of subcomponent C1q of the first component of human complement. Biochem J 155: 19
39. Reid KBM, Gagnon H, Frampton J (1982) Completion of the amino acid sequences of the A and B chains of subcomponent C1q of the first component of human complement. Biochem J 203: 559
40. Rochna E, Hughes-Jones NC (1965) The use of purified ^{125}I-labelled antiglobulin in the determination of the numbers of D antigen sites on red cells of different phenotypes. Vox Sang 10: 675
41. Ross GD, Lambris JD, Cain JA, Newman SL (1982) Generation of three different fragments of bound C3 with purified factor I or serum. I. Requirements for Factor H versus CR1 cofactor activity. J Immunol 129: 2051
42. Rosse WF, Rapp HJ, Borsos T (1967) Structural characteristics of haemolytic antibodies as determined by the effects of ionizing radiation. J Immunol 98: 1190
43. Schumaker VN, Calcott MA, Spiegelberg HL, Muller-Eberhard HJ (1976) Ultracentrifuge studies of the binding of IgG of different subclasses to the C1q subunit of the first component of complement. Biochemistry 15: 5175
44. Schumaker VN, Poon PH, Seegan GW, Smith CA (1981) Semiflexible joint in the C1q subunit of the first component of human complement. J Mol Biol 148: 191
45. Siegel RC, Schumaker VN (1983) Measurements of the association constants of the complexes between intact C1q or pepsin-treated C1q stalks and the unactivated or activated $C1r_2C1s_2$ tetramers. Mol Immunol 20: 53
46. Svehag SE, Bloth B (1970) The ultrastructure of human C1q. Acta Pathol Microbiol Scand B 78: 260
47. Thompson RA, Winterborn MH (1981) Hypocomplementaemia due to a genetic deficiency of B1H globulin. Clin Exp Immunol 46: 110
48. Tschopp J, Shulthers T, Engel J, Jaton J-C (1980) Antigen-independent activation of the first component of complement C1 by chemically cross-linked rabbit IgG oligomers. FEBS Lett 112: 152
49. Tschopp J, Villiger W, Fuchs H, Kilcherr E, Engel J (1980) Assembly of subcomponents C1r and C1s of the first component of complement. Proc Natl Acad Sci USA 77: 7014
50. Wiggins RC, Cochrane CG, Griffin JH (1979) Rabbit blood coagulation Factor XI. Mechanism of activation of rabbit Hageman factor (Factor XII). Thromb Res 15: 475
51. Ziccardi RJ (1982) A new role for $\overline{C1}$-inhibitor in homeostasis: control of activation of the first component of human complement. J Immunol 128: 2505
52. Hughes-Jones NC, Gorick BD (1982) The binding and activation of the C1r-C1s subunit of the first component of human complement. Mol Immunol 19: 1105
53. Ziccardi RJ, Cooper NR (1976) Physicochemical and functional characterization of the C1r subunit of the first complement component. J Immunol 116: 406