THE COMPLEMENT COMPONENTS OF THE MAJOR HISTOCOMPATIBILITY LOCUS

Author: **R. R. Porter**
Department of Biochemistry
University of Oxford
Oxford, England

Referee: K. Frank Austen
Brigham and Women's Hospital
Boston, Massachusetts

I. INTRODUCTION

The importance of the complement system in the prevention of infectious disease is apparent when a central component such as C3 is missing. Patients with an inherited deficiency of C3 are subject to repeated bacterial infection controlled only by the use of antibiotics. Deficiencies of other components cause related diseases of differing complexity and severity varying considerably between individuals. Inability to remove immune aggregates is perhaps the most common feature, and this leads to a variety of symptoms typical of systemic lupus erythematosus (SLE).[1] A less direct disease association has been found with the products, including complement components of the genes in the major histocompatibility complex. Susceptibility to illness of an autoimmune character such as alkylosing spondylitis, myasthenia gravis, multiple sclerosis, juvenile onset diabetes, and celiac disease correlates with certain haplotypes of the histocompatibility antigens. Reviews of the biology and biochemistry of the major histocompatibility complex (MHC) will be found in References 2, 3, and 4.

The HLA-A-B and C genes (Figure 1) code for antigens which are present on the surface of nucleated cells and which determine histocompatibility. They are referred to as MHC type I genes. Their physiological role appears to be in controlling the specific interaction of cytotoxic T lymphocytes with their target cells. Identity of the one type I antigen between cytotoxic cell and target, in addition to the antigenic specificity, is usually necessary for lysis to occur. HLA-D antigens, which are much more restricted in their cellular distribution, were first recognized as differences between cell surface proteins which caused blast cell formation when lymphocytes from different individuals are cultured together in vitro, the mixed lymphocyte culture (MLC) response. Inheritance of the D antigens was established and the genetic locus was found to be linked closely to the HLA-A-B and C loci. Antisera were then found which would inhibit the MLC response and which were not directed to the HLA-A, B, and C antigens. The antigens identified by these antisera were found to be similar to the Ia antigens of mice coded by the immune response (Ir) genes. The surface antigens responsible for the MLC response and those defined serologically by the inhibiting antisera correspond closely but as their identity is not proven the latter are described as DR antigens. The D and DR genes are named (as the Ir genes) as Class II genes. The physiological role of the Class II antigens is in the presentation of foreign antigen to B lymphocytes by macrophages and helper T cells and in other lymphocyte interactions. Identity is usually essential between one of the Class II antigens of these interacting cells. Both Class I and Class II antigens are exceptionally polymorphic with many loci and many alleles at each locus coding for these cell surface proteins.

Also polymorphic, but less so than the Class I and II antigens, are the three complement components, C2, C4, and factor B, which are coded by genes in the MHC and which are

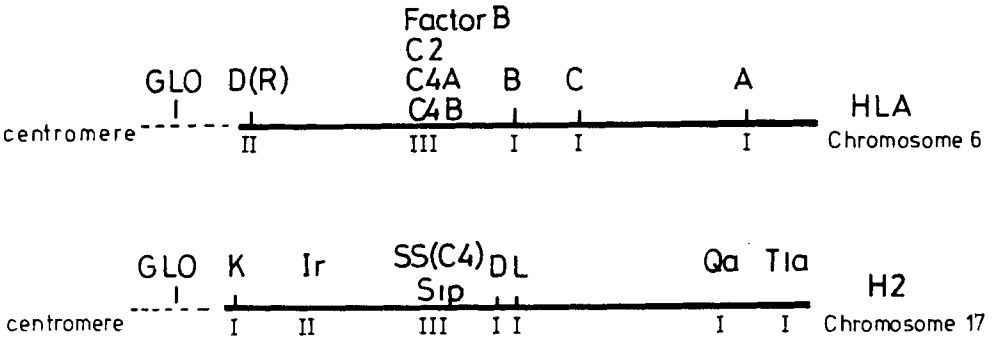

FIGURE 1. The major histocompatibility complex of man (HLA) and mouse (H2) showing the approximate positions of the genes coding for histocompatibility antigens (Class I), the D(R) and Ir genes (Class II), and the complement components C2, C4A, C4B, Factor B (Class III) in HLA and Ss (C4) and Slp in H2. GLO indicates the relative position of the glyoxalase gene.

referred to as Class III antigens. They are found between the HLA-D(DR) and the HLA-B loci, probably closer to the HLA-B. It is to this section of the MHC that susceptibility to autoimmune disease has been mapped with linkage disequilibrium to D(R) and B haplotypes and correlation has also been found between the haplotypes of C2, C4, and factor B and the incidence of some of these diseases.[5,6]

How far, if at all, the diseases associated with the MHC are related to defects in the complement components of the MHC is uncertain, though null alleles at the C2 and C4 loci may determine susceptibility to SLE. Nor is the biological significance of the presence of their structural genes in this complex apparent, though most of the products of the complex identified so far appear to have an essential function in immunity. However, the glyoxalase gene lies close to HLA and H2 in an equivalent position on the centromere side.[7,8] The gene coding for neuraminidase has been mapped into the center of the H2 in the S region[9,10] and the 21-hydroxylase deficiency gene is closely linked to HLA. None of these enzymes have any known role in immune reaction, but the presence of the complement component genes in the MHC does seem to be a sufficiently striking association to justify reviewing together the structure and genetics of the three proteins C2, C4, and factor B as at present understood.

The chemical structure of these proteins, of which much is now known, will be considered first and then such evidence as is available on the organization of their genes, together with a speculation as to the origin of the association of these genes with each other and with susceptibility to disease.

II. THE CHEMICAL STRUCTURE OF C2, C4, AND FACTOR B

The complement system is a cascade of proteolytic activation with two pathways leading to the formation of a complex of five proteins able to lyse animal and bacterial cells (for reviews see Porter[11] and Lachmann[12]). The central feature is the formation of a C3 convertase which activates C3, the major component of the system common to both pathways of activation (Figure 2).

In the classical pathway, C3 convertase is a complex of two activated proteins C4 and C2, the proteolytic active site being in C2, and this is bound noncovalently to C4 which is bound covalently to the antibody-antigen aggregates or antibody-coated cells. The C3 convertase of the alternative pathway is formed from activated C3 and factor B, the proteolytic activity being in factor B. B is bound noncovalently to C3 which is attached covalently to the activating particles, polysaccharide or antibody aggregates. Though formed from different proteins, the two C3 convertases appear to be functionally and catalytically similar with C3

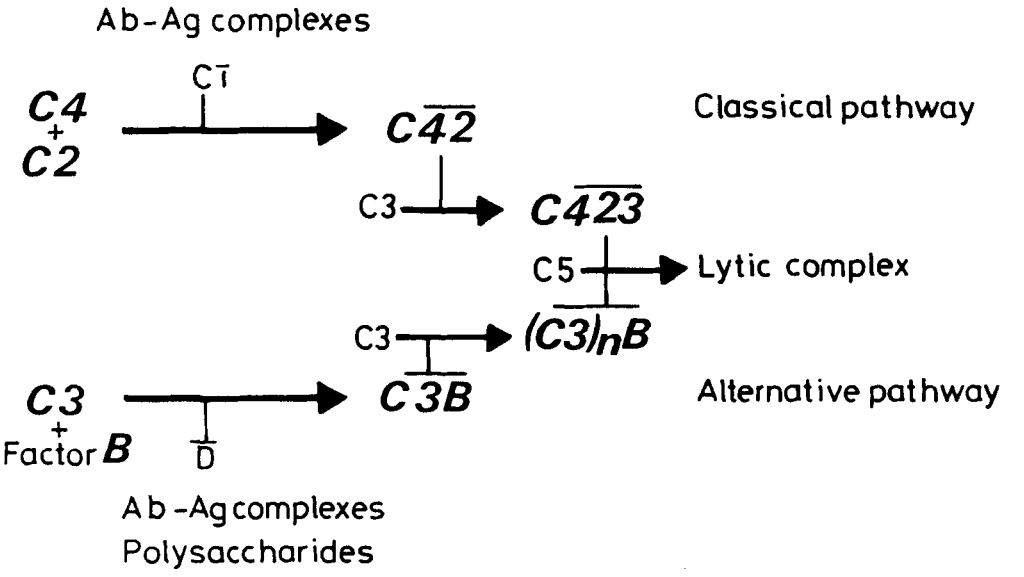

FIGURE 2. A simplified scheme of the activation of complement by the classical and alternative pathways. In the classical pathway C1 binds to antibody-antigen complexes and becomes an active protease $\overline{C1}$. $\overline{C1}$ activates C4 and C2 which become bound to the aggregates and form a complex protease $\overline{C42}$, which activates C3. $\overline{C3}$ associates with $\overline{C42}$ to form $\overline{C423}$, C5 convertase. This protease activates C5 which initiates the formation of the lytic complex of the late components. The alternative pathway is activated by antibody complexes or substances such as the high-molecular-weight polysaccharides found in bacterial and yeast cell walls. Factor B, in association with $\overline{C3}$, is activated by $\overline{D}$ to give the C3 convertase $\overline{C3B}$. More C3 binds to $\overline{C3B}$ to give the C5 convertase $(\overline{C3})_n\overline{B}$. C3 and factor B have equivalent roles in the convertases and related structures to C4 and C2, respectively.

and C4 having equivalent roles as do C2 and factor B. Both require Mg^{++} for formation and both dissociate spontaneously with loss of activity; and in both, association of their reaction product $\overline{C3}$ with the C3 convertase changes its specificity to that of a C5 convertase. C2 and factor B are similar structurally as are C3 and C4. Though in vivo it is likely that the C3 convertases are functionally active, only when bound to antibody aggregates or activating particles they can be formed in solution experimentally. The covalent binding of C4 and particularly C3 with immune aggregates causes their partial dissociation and facilitates the uptake of the aggregates by phagocytic cells through C3 and C4 receptors on the cell surfaces. The complex proteases of the complement system have been reviewed recently.[13,14]

A. C2 and Factor B

Both C2 and Factor B are synthesized by macrophages[15] and are single-chain glycoproteins of molecular weight of approximately 100,000 and 90,000, respectively. The structure of factor B, which is present in blood in much higher concentrations (200 mg/ℓ) than C2 (15 mg/ℓ), has been studied in greater detail and will be described first.

1. Factor B

There were conflicting reports on the inactivation of factor B by diisopropyl fluorophosphate and, hence, as to whether it was a serine protease or not. In the many serine proteases which have been studied, the catalytic chain has always been of a molecular weight of about 25,000 ± 3000 but activation of factor B by $\overline{D}$ results in the splitting of a single peptide bond giving chains of 30,000 (Ba) and 60,000 (Bb) molecular weights which are not disulfide-bonded and which dissociate easily. The catalytic site is in the 60,000 mol wt fragment and its relation to the catalytic chain of other serine proteases has been established by the solution of its complete amino acid sequence[16-20] (Figure 3).

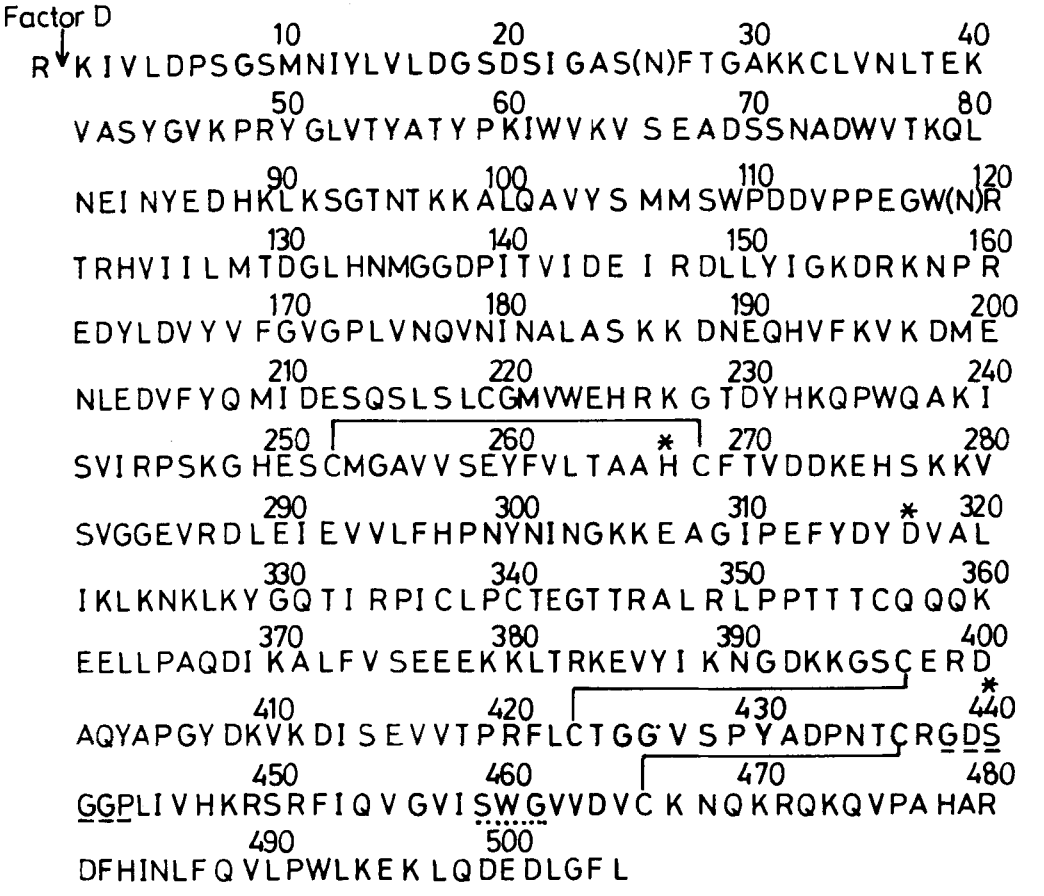

FIGURE 3. Amino acid sequence of the C-terminal Bb fragment of factor B.
*: Active site residues histidine, aspartic acid and serine
–: Underlined residue show primary binding site
. . .: Underdotted lined residues show secondary binding site
The disulfide bonds shown are placed by homolog with those of chymotrypsinogen 158—168 (His loop), 168—182 (Met loop) and 191—220. The half-cystine residue at position 33 carries the only free sulfhydryl group. Asparagine residue (N) in brackets has a polysaccharide group attached. Single letter code for amino acids A, ala; B, Asx; C, cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu,; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr; Z, Glx. (From Gagnon, J. and Christie, D. L., *Biochem. J.*, 209, 51, 1983. With permission.)

The C-terminal section of Bb (residues 230 to 505) shows a strong homology with conserved sequences found in all serine proteases.[21] The three active site residues, His-57, Asp 102, and Ser 195 in chymotrypsinogen numbering, which are found in positions 267, 317, and 440 of the Bb sequence and the primary binding site residues Gly-Asp-Ser-Gly-Gly-Pro (438-443) are entirely conserved (Figure 3). The secondary binding site Ser-Trp-Gly (459-461) is also conserved. The half cysteine residues at 252 and 268, 396 and 423, and 436 and 466 probably correspond to the disulfide bonds between 42 and 58 (His loop), 168 and 182 (Met loop), and 191 and 220 found in chymotrypsinogen.[20] It is clear, therefore, that the catalytic site of factor B has a structure common to all other serine proteases.

Major differences are, however, found between the Bb sequence and that of chymotrypsinogen. For example, while the distance between the active site residues His and Asp is 49 relative to 44 in chymotrypsinogen, that between Asp and Ser is 122 in Bb but only 92 in chymotrypsinogen. More importantly, the N-terminal sequence of the catalytic chains of

all other serine proteases is missing. Activation is usually by the splitting of an Arg-Ile bond (residue 16-17 in chymotrypsinogen) giving N-terminal Ile-Val-Gly-Gly or some very similar sequence. The movement of this section of the chain to form an ion pair between the α-amino group of the N-terminal isoleucine with Asp 194 (chymotrypsinogen numbering) in the primary binding pocket is believed to lead to exposure of the catalytic site which is masked in the zymogen.[22] In factor B, activation is by splitting an Arg-Lys bond some 222 residues N-terminal of the section equivalent to the activation site of chymotrypsinogen. In Bb the unrelated sequence Trp-Glu-His-Arg (residues 223-226 in Bb) replaces Ile-Val-Gly-Gly (Figure 3).

Factor B is clearly a zymogen of a serine protease, but of a novel type with a different activation mechanism from that found in other serine proteases. It is likely that the N-terminal 220 residues of Bb form a domain which interacts with C3 to give the active C3 convertase. If the conformation of this domain in unactivated factor B is such as to enable it to bind to C3, a change of conformation when Ba is split off would lead to dissociation of C3b and Bb with loss of the convertase activity.

That Bb alone has no C3 convertase activity could arise from the suggested conformational change in the N-terminal half if this contained a binding site for C3, or possibly there is a C3 binding site in the associated $C\overline{3}$. It is an unusual situation where the product of the catalysis is also part of the enzyme complex.

2. C2

Much less is known of the structure of C2, but enough to confirm that it is homologous to factor B. It is a single-chain protein split on activation by $C\overline{1}s$ into two chains of approximately 30,000 (C2b) + 70,000 (C2a) molecular weight. (Note that the nomenclature of C2a and C2b is the opposite to that used for Ba and Bb.) Homology of amino acid sequence is apparent at the N-terminus of Bb and C2a[23,24] (Figure 4), but not in the short sequence published for C2b and Ba.[23,25] Peptides derived from the catalytic site section of C2a show that C2, as factor B, is the zymogen of a serine protease with strong similarities in structure to factor B.[145] For example, a segment of 36 residues around the active site serine shows 16 identical amino acids and 7 conservative replacements (Figure 5).

The C3 convertase of the classical pathway, $C\overline{42}$, dissociates and loses activity with a decay rate constant k of 2.0 min^{-1} at 37°C. Removal of C2b from the complex, after activation by $C\overline{1}s$, does not affect the decay rate but this section of C2 is likely to be involved in the initial interaction of C2 with C4b.[26] A remarkable feature of the convertase is its stabilization by oxidation of C2 with weak iodine solution prior to the activation of C2.[27] The $C\overline{42}$oxy convertase is 10- to 20-fold more active than $C\overline{42}$. It was suggested at the time that this stabilization of the complex might be due to oxidation of SH to S–S groups,[28] but it has been found subsequently that both factor B and C2 each have only a single sulfhydryl group,[24] contrary to a previous report.[29] Reaction of the sulfhydryl group with pCMB inactivates C2 but not factor B, and similarly oxidation with I_2 solution does not affect the activity of factor B. Presumably, therefore, the sulfhydryl of factor B is not in a section of the molecule interacting directly with C3b. The sulfhydryl groups of both proteins are near the N-terminus of the catalytic chain, but not in equivalent positions[24] (Figure 4). The cysteine residue of factor B has no role in its biological activity, but that in C2 appears to be involved, at least indirectly, in the interaction of C2a and C4b as oxidation increases the half-life of the complex from a few minutes to several hours. It is suggested that the sulfhydryl group is oxidized to a sulphenyl iodide, which reacts with another amino acid side chain displacing the iodide and forming a labile intramolecular covalent bond. Such a reaction could stabilize the conformation of the part of C2a which interacts with C4b and hence increase the stability of the complex.

By analogy with factor B, it is unlikely that the section of the peptide chain containing

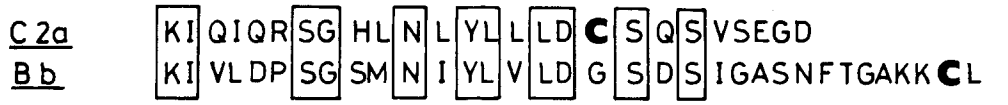

FIGURE 4. The N-terminal sequences of the active site containing fragment of factor B (Bb) and of C2 (C2a). Identical residues are boxed and the single sulfhydryl group of each protein is in the half-cystine residue in heavy type (C). (From Parkes, C., Gagnon, J., and Kerr, M. A., *Biochem J.*,)

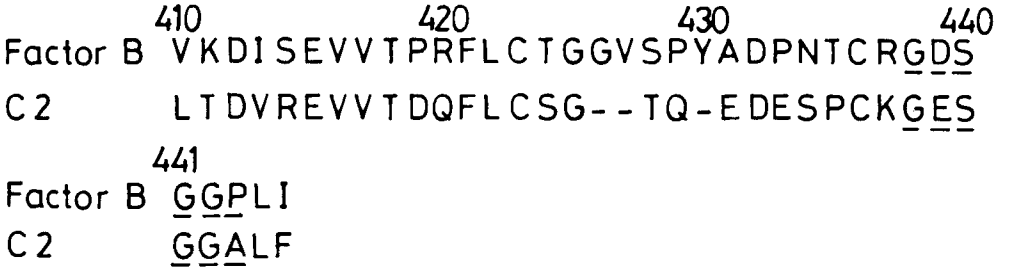

FIGURE 5. The amino acid sequence around the primary binding site residues (underlined) in C2 and factor B. Gaps (−) have been left in the C2 sequence to maximize homology. This gives 16 identical positions and 7 conservative replacements out of a sequence of 36 positions.

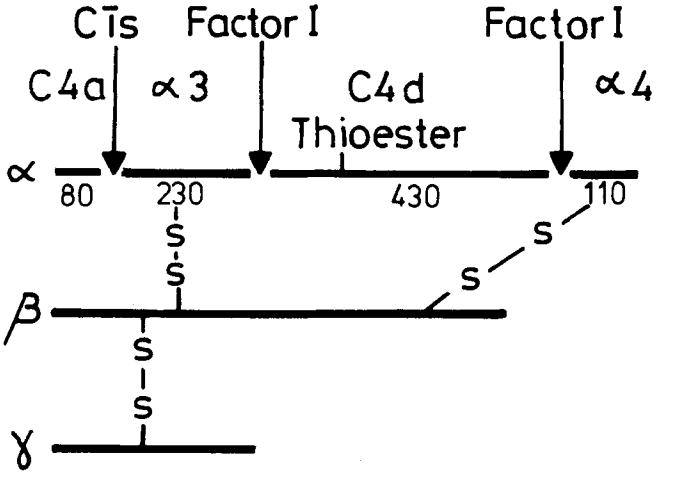

FIGURE 6. The three-peptide chain structure of C4 showing the chain activated by C1s to give C4a and C4b. The latter is inactivated by hydrolysis at two positions by Factor I and the protein cofactor C4b giving the three additional chain fragments α3, C4d, and α4. Numbers under the chain show the approximate number of amino acid residues in each fragment. The position of the intrachain thiolester bond near the midpoint of the chain is also shown. The interchain disulfide bonds are diagrammatic only but show that, after inactivation, C4d is released and C4c contains the β and α chains joined to the α3 and α4 fragments of the chain.

the sulfhydryl group is itself part of the C4b binding site. More probably, the proposed bond is stabilizing the native conformation of the section of C2 containing the C4b binding site. If the nature of the reacting amino acid side chain could be found it should help to identify the C4b binding site, but the lability of this postulated bond has prevented its characterization.[24]

B. C4

C4 (as C3 and C5) in guinea pigs[30] and mice[31] is synthesized in macrophages as a single peptide chain and is subsequently split to a disulfide-bonded multichain protein on secretion[32,31] (Figure 6). Small amounts of unsplit C4 have been found in serum,[33,34] and also a partially

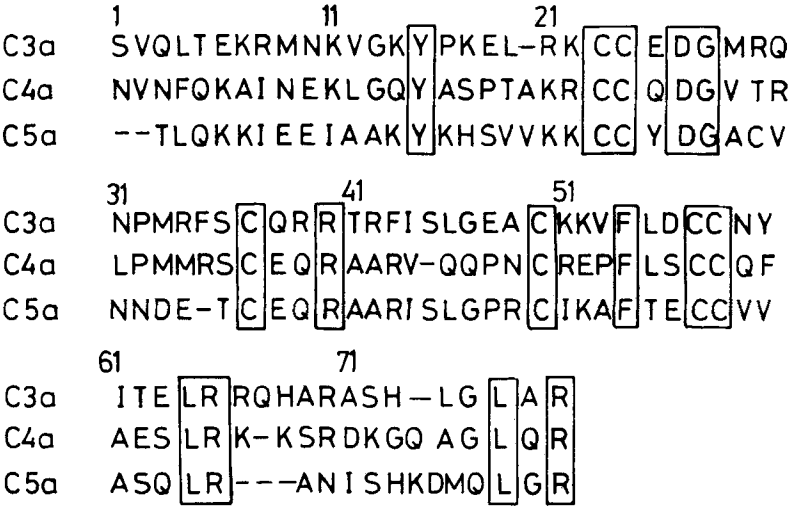

FIGURE 7. Comparative sequences of the human C3a, C4a, and C5a N-terminal peptides formed on activation of C3, C4, and C5. Residues are arranged to give maximum homology and identical residues are boxed. Data from References 41 and 42.

processed form is present in about 8% of the total in which the α chain is extended by about 50 amino acid residues at the C-terminal end.[35] A similar higher-molecular-weight α chain had been reported previously in mice.[31] It is suggested that the 5000 mol wt peptide is split off after secretion. C4 contains three chains: α 95,000, β 75,000, and γ 30,000.[36,37] C3 and C5 contain two chains: α 110,000 and β 70,000. All the proteins are glycosylated but the importance of the polysaccharide units is uncertain. Human C4 protein secreted by macrophages which had been incubated with tunicamycin to inhibit the glycosylation had similar hematological activity to normal C4,[38] suggesting that the polysaccharide had no role in this reaction. However, mouse C4 occurs in several variant structural forms with different hematological activities. Most strains of mice have an α chain of 98,000 mol wt, but in some of the molecular weight of the α chain is 94,000. The difference has been shown by chemical removal of the polysaccharide to be due to the lower glycosylation of the α chain in the smaller-molecular-weight form.[39] In these strains the hemolytic activity of the C4 is some 75% lower than in those with the higher glycosylation, and it was suggested that the polysaccharide might be responsible for the difference in activity. Definitive evidence on this point will probably require much more detailed knowledge of the multiple interactions of C4 with other proteins such as antibody, C2, and $C\bar{1}s$.

The chemical structure of human C4 has been studied most extensively and it has been found that its activation by $C\bar{1}s$ and inactivation by proteolysis by factor I, together with the protein cofactor C4bp, all depend on the splitting of peptide bonds in the α chain (Figure 6). The amino acid sequences of several of these proteolytic fragments have been reported. C4a from both bovine[40] and human C4[41] contains 77 amino acid residues from the N-terminal end of the α chain. They show close homology with each other and also with C3a and C5a (Figure 7). Entirely conserved is the C-terminal arginine which is essential for the anaphylotoxin activity and which is found in all these peptides.[42,43] Also striking is the presence of six half-cystine residues in equivalent positions in C3a, 4a, and C5a. It has been suggested that they form a "disulfide knot" and determine much of the secondary structure of these peptides.[42] The crystal structure of human C3a has since been determined and has shown that it has a drumstick-like appearance with all the disulfide bonds in the globular head. It was possible to establish the positions of the disulfide bonds between residues 22 to 49, 23 to 56, and 36 to 57.[44]

The three fragments of the α′ chain formed by proteolysis with factor I and the C4bp have been aligned by determining the N- and C-terminal sequences of the whole chain and of the fragments. This showed them to be α3-C4d-α4 with approximate molecular weight of 25,000, 44,500, and 12,000, respectively.[45] Together with C4a (molecular weight 9000), these four peptides are believed to account for the whole α chain of C4 with probably no small peptides being lost, though the sum (90,500) is rather less than that given for the α chain (95,000). Again there is homology between C3 and C4 in the N-terminal sequence of the α′ chain.[45]

More information is available on the C4d sequence as considerable interest has been aroused by the finding in this section (and the equivalent section of C3) of evidence of an intrachain thiolester bond which plays an essential role in the activation and assembly of the complement proteins.

1. The Thiolester Bond of C4 and C3

The first suggestion of an unusual structure in C3 came from the observation of Law and Levine[46] that if complement is activated by antibody-coated red cells, C3 forms a stable, probably covalent, bond with components of the red cell surface. From its stability it was suggested that it was probably an ester bond formed between an acyl group on C3 and a hydroxyl group on the cell surface.[47] Further studies from several laboratories (reviewed in Reference 14) showed that both the serum protease inhibitor, α_2 macroglobulin and C4 could also form covalent bonds, on proteolytic activation, with proteins and also with small molecules such as amines and glycerol. It had also been observed that in C3 and C4, when denatured or activated, a sulfhydryl group was exposed which could be labeled by a radioactive iodoacetic acid.[48,49] Methylamine, which inactivates C3 and C4, is incorporated by an amide bond to a glutamic acid residue. Sequence determinants around these reactive groups showed a common sequence in this section and led to the suggestion that an intrachain thiolester bond was present,[48,50] as shown (Figure 8). More extensive sequencing of C4d positioned some 100 residues from the N-terminus, including the postulated thiolester bond at residues 54—57 (Figure 9).[51] Most of the remaining sequence of C4d has been completed and, so far, alternative residues hae been found in four positions. These presumably arise from inherited variations though their relation to the multiple polymorphic forms of C4 has not yet been established.[52]

The thiolester structure has several unusual features. First, its stability in C3 and C4 (half-life of several days) is greater than that of a simple thiolester in aqueous solution (half-life 30 min). On activation by splitting off C3a or C4a, however, the half-life is 0.1 to 1 msec.[53] There must be special features in the adjacent structure responsible for this greatly increased reactivity of the thiolester bond on activation, and the homology of the adjacent amino acid sequence (Figure 9) between C4, C3, and α_2M suggests that this is so. The Arg-Ala bond hydrolyzed by $C\bar{1}s$ in the activation of C4 is some 300 residues N-terminal of the thiolester bond making it likely that subsequent conformational change rather than a direct effect of the peptide bond breakage is responsible for the greatly increased lability of the bond. As discussed later, there is evidence of a marked conformation change from C4 to C4b and Isenman et al.[54] showed that an immediate change in the circular dichroism spectrum occurred when C3 was activated by trypsin. C3 can be activated to bind small molecules by denaturation with a variety of reagents such as KBr, KSCN, guanidine, urea, and SDS, and by comparing the rates of reaction under these conditions with those when C3 is activated by trypsin, Law[55] concluded that a conformational change was the essential feature in the increased reactivity of the thiolester bond.

Further, it was observed first in α_2M,[56,57] and then in C3[48,58,59] and C4,[58] that denaturation of the native protein by heat or other means caused breakage of a peptide bond between the adjacent glutamic acid residues leaving an N-terminal pyroglutamic acid residue on the

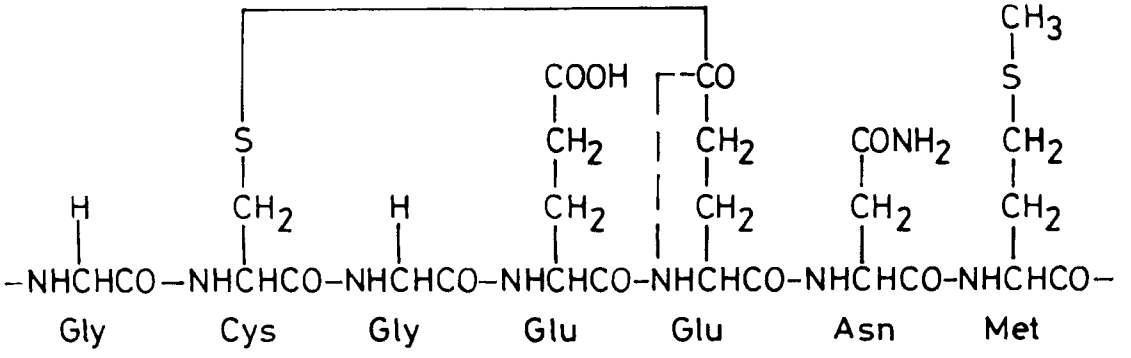

FIGURE 8. The postulated thiolester structure of C3 and α_2M. In C4, the sequence is identical except that asparagine is replaced by threonine. The dashed line shows the probable formation of an N-terminal pyrollidone carboxylic acid residue with the breaking of the Glu-Glu peptide bond on denaturation of the native protein.

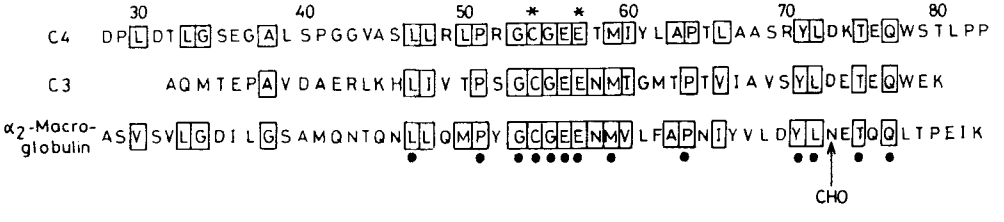

FIGURE 9. The amion acid sequences around the postulated thiolester bond in C4, C3 and α_2M. Boxed residues are homologous and spots show identity. Asterisks denote the cysteine and glutamic residues forming the intrachain thiolester bond.

carboxyl half of the chain (Figure 8). Law[55] compared the rates of the binding of small molecules to C3 with the rates of autolytic cleavage in increasing concentrations of guanidine hydrochloride. He found that C3 was stable up to 1 *M*, the binding reaction predominated between 1 and 2 *M*, and the autolytic split was faster above 2 *M* guanidine HCl when the structure becomes more random. Khan and Erickson[60,61] synthesized a thiolactone ring resembling the thiolester site of C3 and found that it also undergoes autolytic cleavage in aqueous solution. Law[55] suggested that autolytic cleavage is a function of the thiolester structure alone, but that the binding reaction depends on more complex features dependent on adjacent sections of polypeptide chains.

It is likely that these deductions concerning C3 are also applicable to the structure of the thiolester bond in C4.

A hypothesis to explain these reactions has been put forward by Davies and Sim,[62] who suggest that in the Gly-Cys-Gly-Glu-Glu sequence the reactivity of the activated proteins is a result of hydrogen bonding of the first glutamic acid residue and the carboxyl oxygen of the thiolester. Formation of such a bond, which can be accommodated in a molecular model, would increase the electrophilicity of the thiolester and greatly enhance the nucleophilic attack on the ester carboxyls. In the native protein, it is suggested that the activated thiolester is completely protected from solvent by a portion of the polypeptide chain. Activation of the proteins results in a conformational change which exposes the thiolester. The bond breakage on denaturation would involve addition of nucleophilic nitrogen to the hydrogen bond-activated thiolester with the formation of a five-membered imide ring releasing the thiol. Hydrolysis of the imide would cause peptide bond cleavage and formation of a new cyclic N-terminal pyroglutamyl residue. Another suggestion for a mechanism for activation of the thiolester, namely, charge distortion by a neighboring ionized carboxyl group, has been put forward by Howard.[63]

The biological significance of covalent bond formation between C4 (and C3) and antibody or cell surface molecules may lie in the need to anchor the C4 molecule after a conformational change occurs on activation. A weak noncovalent interaction of C4 and antibody may bring the potential acyl group into close proximity with an acceptor residue which is in the N-terminal half of the antibody heavy chain.[64] The very short half-life of the acyl group in aqueous solution requires that the reactive groups be very close. That there is a marked conformational change in C4 when C4a is split off is apparent from the exposure of the acyl group, the susceptibility of C4b but not native C4 to hydrolysis by factor I and C4b binding protein, the affinity of C4b but not C4 for C2, and the change in differential iodination of the three peptide chains from C4 to C4b.[65] The anchoring of C4b by a covalent bond together with the short half-life of the $C\overline{42}$ convertase ensures that complement activation is localized to the immediate vicinity of the antibody antigen interaction. As complement activation could lead to lysis of the animal's own cells this limitation of activation is probably an essential feature of the system.

III. GENETICS OF C2, C4, AND FACTOR B

A. C2

C2 deficiency in man is linked to HLA[66-68] and it has since been shown that the C2 deficiency gene is an allele of the structural locus of C2.[69-71] Using polymorphic forms recognized by isoelectrophoretic focusing,[72-74] it was confirmed that the structural gene for C2 is in the HLA complex. A series of papers summarized by Weitkamp and Lamm[75] place the C2 gene between HLA-B and HLA-DR (Figure 1).

B. Factor B

Simultaneously with the finding that C2 deficiency in man is linked with HLA, factor B

was also reported to be linked to HLA.[76] Microheterogeneity of factor B could be seen when it was electrophoresed at alkaline pH in agarose and identified by specific antisera.[77] The patterns were interpreted as representing 4 allelic forms at a single autosomal locus, but subsequent work has shown the number of alleles to be at least 11,[78] and probably higher.[79] A series of family studies summarized by Weitkamp and Lamm[75] showed factor B, as C2, to be coded by a structural gene between HLA-DR and HLA-B, probably closer to HLA-B.[80,81] There has been no convincing example reported, so far, of a crossover between the C2 and factor B genes.[82]

The factor B gene[147] and probably the C2 gene[148] have been linked to the H2 region in mice.

C. C4 in Man

The genetics of C4 is more complex than that of C2 and factor B in that it has now become clear that there are probably two and possibly more structural gene loci for C4, with alleles at each.

Studies of C4-deficient families[83-85] and of inheritance of electrophoretically polymorphic forms of C4[73,86,87] showed that the gene coding for C4 was linked to HLA. At the same time, newly discovered red cell antigens, Chido and Rodgers, were also shown to be linked to HLA.[88-90] The Chido and Rodgers antigens were known to be in plasma as well as on red cells, and O'Neill and colleagues[91] showed that they were on the C4 molecule as C4 would inhibit anti Chido and anti Rodgers serum. Results from following these antigens were therefore applicable to C4 genetics. The relation of Chido and Rodgers antigens to C4 became more precise when Tilley et al.[92] showed that the C4 fragment bound to red cells was C4d. This was in agreement with the structural studies described above which suggested that C4 becomes bound to the red cell surface through an ester bond[93] in the C4d section of the α' chain and that subsequent hydrolysis by factor I and C4bp would release C4c and leave C4d still bound. This observation has, however, been challenged recently by Chu et al.[94] who could not find C4d on red cells activated in serum at 37° but did detect noncovalently bound C4c. In cells activated at 0° C4d was present but apparently not covalently bound as it was not in a high-molecular-weight form. Anti Chido and anti Rodgers serum was neutralized by whole C4, α chain, C4d and the 25,000 mol wt α' chain fragment of C4c, suggesting that antigens were on both C4d and C4c. It is possible that the specificity of the antisera is complex and that not all are identical. Tryptic digest fragments of C4 α' chain, apparently equivalent to C4d, carried the Chido and Rodgers antigen and were distinguishable by their size.[95]

Multiple forms of C4 were reported first by Rosenfeld et al.[96] using crossed immuno electrophoresis, but greater complexity has been found by prolonged electrophoresis of plasma or serum in agarose gels and identification of C4 by immunofixation.[97] This technique showed four distinct bands on C4F and four in C4S, and as they were not overlapping they have eight bands in the heterozygote C4FS. C4F carried the Rodgers antigenic specificity and C4S the Chido specificity. The resolution depended on the buffer used and the precise experimental conditions. From this study of 248 children in 79 families O'Neill et al.[97] concluded that there were two structural loci for C4, Chido or S and Rodgers or F, contrary to previous suggestions[86,87,98] of one locus with three or four alleles. Whether all individuals have two C4 loci is still not clear.[99] Treatment of serum with neuraminidase before electrophoresis gave sharper bands[100] and increased the number of detectable bands, some overlapping between homozygous S and F. A new nomenclature of A for F and B for S was suggested and has now been generally accepted. The polymorphism of C4 appears to be confined to the α chain as judged by two-dimensional electrophoresis.[101]

Using functional assays as well as immune fixation it appeared that C4A was less hemolytically active than C4B. A hemolytically inactive form of C4 has been reported by Teisberg

et al.,[102] and this has been confirmed.[103] Remarkably, O'Neill et al.[104] found the C4F1 (now A6) allotype to be inactive when associated with HLA-B17 antigen but not with other HLA-B alleles.

The classification of C4A and C4B alleles in any individual depends upon assessment of a complex pattern of bands after electrophoresis and immunofixation or hemolytic overlay. Agreement between laboratories is still being sought, but it seems likely that there are not less than 13 alleles of C4A and 22 alleles of C4B.[105]

D. C4 in Mice

In 1963 Schreffler and Owen[106] reported that there was an inherited variation in the content of a serum protein in mouse sera. It was named Ss and the inheritance was shown to be linked to the H2 locus. It was later shown that the S locus controlled complement levels[107] and that Ss was mouse C4.[108-110] An antigenically related protein Slp was found and its inheritance was also linked to H2. Ss protein has C4 hemolytic activity[11] but Slp has no hemolytic activity by in vitro assay[112] nor is it split by $C\bar{1}$. Partial chemical characterization of biosynthetic intermediary products suggest that the structure of the Ss and Slp proteins are similar.[113] Short N-terminal sequences show homology between Ss and Slp and the apparent difference in molecular weight of the α chain appears to be due to different carbohydrate contents.[114] Substantial differences by peptide mapping between Ss and Slp proteins had been observed earlier in the β chains, though the α and γ chains were similar.[115] A further similarity between Ss and Slp proteins is that both are synthesized as a precursor molecule of molecular weight about 185,000.[31] The genes coding for Ss and Slp lie between the I region and H-2D in the H2 complex[116] (Figure 1).

There are, therefore, strong correlations between the properties of the human and mouse C4 genes and proteins. Both are synthesized as single peptide chains and subsequently split into three, both occur in two structurally distinct forms, one of which in man is less active by hemolytic assay than the other and one is inactive in the mouse. Both are coded by genes in equivalent positions close to the immune response genes in mice and in man.

E. C3b Receptors

Another group of complement-associated proteins has been reported to be coded by genes in the MHC. These are the C3 receptors which are found on macrophages and lymphocytes, and are responsible for the adherence to them of cells coated with antibody and complement components. This interaction facilitates opsonization and possibly other cellular interaction in the immune response. The most common assay is measurement of rosette formation between lymphocytes and red cells coated with antibody and complement.

Such an assay was reported to be blocked by anti H2 or anti HLA antisera, and using mouse human somatic cell hybrids expression of C3b and C3d receptors correlated with expression of HLA antigens.[117] Others reported blocking with some, but not all, anti HLA specificities[118,119] and some failed to find specific blocking of the rosette assay.[120,121] Using mice, Bishop et al.[122] found nonspecific inhibition of rosetting with some anti H2 sera, but also some specific inhibition. It was suggested that secondary disturbance of the cell surface components following specific interaction of anti MHC sera might be involved and there is no clear evidence at present that the genes coding for any of the C3 receptor proteins are in the MHC.

F. C4 Binding Protein

C4 binding protein, the cofactor required for the hydrolysis of C4b by factor I, has been found to show inherited polymorphism in mice and using inbred strains it was mapped close to D and L but on the opposite side to the S region.[123] There is no evidence so far on the position of the gene coding for C4bp in man.

IV. MOLECULAR GENETICS OF C4 AND FACTOR B

Rapid progress is being made in precisely identifying the structures of the genes of the Class I and Class II antigens in the major histocompatibility locus of both mouse and man. In the H2 for example, much of the DNA coding for the Class II antigens in the I region has been included in overlapping cosmid clones[124] and some of the Class I genes have been identified.[125-127] Similarly, the structures of the HLA genes are being solved.[128-130]

Progress on the complement genes is, however, slower. Cloned cDNA coding for most of the C3 components in mouse has been obtained,[131] but the gene for C3 is not in the MHC. In the mouse it is on the same chromosome but shows no linkage to H2,[134,133] and in man it is on a different chromosome — chromosome 19.[132] Of the MHC complement genes, cDNA has been cloned for factor B.[135] It was identified in a human liver cDNA library by hybridization with synthetic oligonucleotide probes prepared on the basis of the known amino acid sequence of factor B. Inserts of up to 2300 base pairs were found and account, therefore, for most of the mRNA coding for factor B. Nucleotide sequencing showed agreement with published amino acid sequences with one discrepancy and no doubt when the sequence is complete the position of allelic variants will be established. In a similar experimental approach Campbell and Porter[136] isolated a shorter cDNA insert and used it to identify the factor B gene in a cosmid library of total human DNA where inserts of genomic DNA of 40,000 to 50,000 kilobase pairs occur.[137] Four cosmid clones containing the factor B gene were found, and from their restriction enzyme map appear to contain overlapping sections of the same DNA sequence of approximately 47 kb, except in one where a deletion containing a Cla digest site occurred. This was well separated from the factor B gene and no evidence was obtained for more than one factor B gene. Nor have restriction enzyme digests shown up any allelic differences so far.

Sequencing of the factor B gene and comparison with the known amino acid sequence has given the structure from the C-terminal end of the Ba fragment to the C-terminal end of Bb, about $^2/_3$ of the whole molecule, and has shown it to be about 4 kb in length. The nucleotide sequence has been completed in 3.3 kb, covering the whole coding sequence for the C-terminal 418 amino acid residues which include all those sections of the peptide chain containing the catalytic site. Figure 10 shows the arrangement of the coding sequences which are contained in eight exons. Exons 2, 3, and 7 contain the three active site residues Asp, His, and Ser, respectively, while exon 8 contains the secondary binding site and exon 6 the binding pocket containing the specificity-determining Asp residue. Exon 4 contains a conserved and structurally important disulfide bond. This conforms in part to the concept of exons corresponding to structural units, though the three active site residues are, of course, parts of the same functional unit and their precise steric relationship is essential for catalytic activity. However, Craik et al.[138] have suggested that exon-intron boundaries occur at protein surfaces. From the data which they quote for trypsinogen and chymotrypsinogen there is exact correspondence of a boundary immediately before the glycine residue of the invariant active site serine sequence, i.e., at Arg-Gly residues 437-438 of factor B (Figure 3). Others show varying similarities, but the full significance will not become apparent until more information on other serine proteases becomes available.

The first report has appeared on the molecular genetics of C4.[139] Cloned cDNA, identified by hybridization with a synthetic nucleotide mixture based on a known amino acid sequence, has been used as a probe in the same cosmid library of human DNA as was used for the factor B gene study. Six clones containing C4 genes were found and restriction enzyme mapping showed that all six probably contained a C4 gene placed centrally in the 40-kb-long insert. Carrol and Belt[146] obtained a cDNA clone of 5.2 kb and, by comparing the nucleotide sequence at the 5′ and 3′ ends with known amino acid sequences, they showed that it coded for the whole length of the C4 including a leader sequence. This was used to

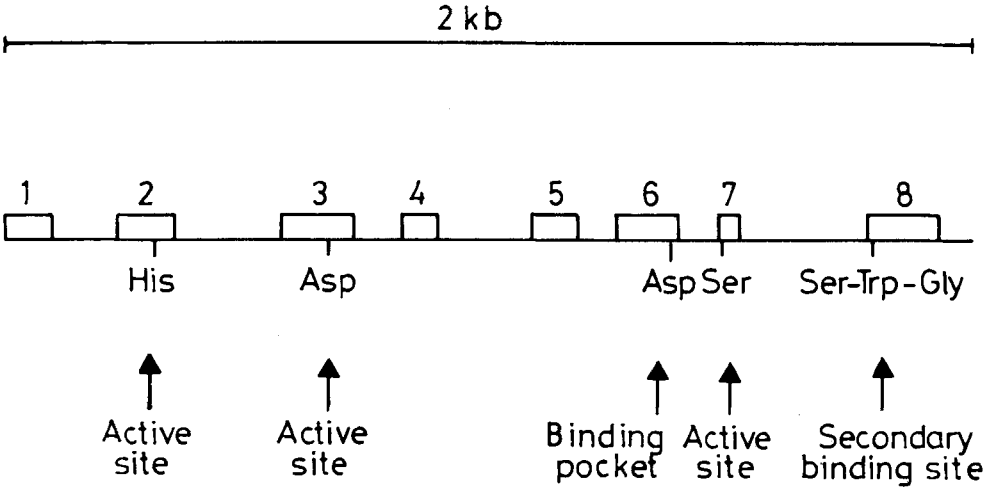

FIGURE 10. Arrangement of the exons coding for the serine protease domain of factor B.[136] Each section of the peptide chain with a defined catalytic function is in a separate exon. Exon 5 codes for a sequence which appears to be an addition relative to trypsin and chymotrypsin.[138]

show that in a cosmid clone, one full-length gene was placed centrally and that a large part of a second C4 gene was also present. The two genes were separated by about 10 kb and, surprisingly, were in opposite sense, i.e., 3′ to 3′. Other cosmid clones containing C4 genes were investigated and restriction enzyme digests suggested that a third gene may be present in this individual, though it is not known whether one or more may be pseudogenes.

Digestion with Kpn restriction enzyme showed that two different-sized fragments hybridizing with C4 cDNA were present. Comparison of the Kpn digests of individuals typed for C4[140] showed that the 7.5 digest product correlated with C4a and the 8.5 product with C4B, heterozygotes carrying C4A and C4B gave both fragments. Presumably, more detailed analysis will show variation of nucleotide sequence corresponding with the different alleles of C4A and C4B. The population studies discussed above suggest that the genetics of C4 are complex, but extension of the present work should give a full picture of the number of C4 genes and of the total number of different allelic forms which are present in different individuals.

As work on the molecular genetics of the complement components of both mice and humans is going on in several laboratories, it should soon be possible to orient the genes of C2, C4, and factor B to each other and also to the Class I and Class II genes which are adjacent to them.

V. HLA, COMPLEMENT POLYMORPHISM, AND RELATED DISEASES

As mentioned in the introduction, there is an association between the relative susceptibility of individuals to certain autoimmune diseases and the different allelic forms which they carry of the HLA-D, HLA-B, C2, C4, and factor B genes.[141,142] Also the absence of complement components C2[1] and C4A or C4B[5] correlates with susceptibility to systemic lupus erythematosis (SLE). In an attempt to relate these findings to each other, it has been suggested that an important factor might be the efficiency of the complement activation system and that this could depend on the association in any individual of the particular polymorphic forms of C2, C4, and factor B which are present.[141] An exceptionally efficient complement system might make an individual more prone to damage of his own tissues by complement and hence exacerbation and possibly initiation of autoimmune damage. An

exceptionally inefficient complement system would lead to failure to dissociate and remove antibody-antigen aggregates and lead to symptoms of SLE.

The argument put forward was based on the following points, all of which have been discussed above:

1. C4, which occupies a central role in the activation of the classical pathway, interacts in the activation and inactivation scheme with many other proteins: $C\bar{1}s$ which activates C4, antibody and cell surfaces to which C4 binds covalently, C2 which associates with C4 to form C3 convertase, C3 which associates with $C\overline{42}$ to form C5 convertase, C4b receptors in phagocytic cells which facilitate removal of immune aggregates, factor I and its cofactor C4 binding protein which together inactivate C4b by proteolysis.
2. All these interactions, as so far understood, involve the α chain which is split in activation and inactivation and which contains the intrachain thiolester group.
3. Of the 40 or more polymorphic forms of C4 so far reported, it is assumed that they correspond to amino acid residue changes and the four replacements so far identified occur in positions on the α chain, C-terminal to the thiolester bond.
4. It is therefore likely that the α chain occupies much of the molecular surface of C4 and is involved in the interactions with the seven other proteins on which the activation and inactivation of the classical pathway depends. If, as appears probable, most of the amino acid replacements also occur in the α chain, it's likely that they will have a marked influence on the strength of these interactions and hence on the efficiency of the complement system. In standard hemolytic assays using heterologous components, C4A is less hemolytically active than C4B, and in the mouse Ss protein but not Slp shows C4 hemolytic activity. In a completely autologous system different results may be obtained, but that variation in activity will occur seems certain.
5. Because many of the other proteins interacting with C4 are also polymorphic, efficient and inefficient associations of different forms of each would be expected, possibly extending to Class I and Class II antigens as major components of cell surfaces and, hence, potential targets for covalent bond formation with C4.

In view of these data, it is suggested that optimum association of Class I, II, and III polymorphic forms may occur and could contribute to the known linkage disequilibrium of certain of their many polymorphic forms.[142] It is a distorted association leading to under- or over-efficiency of the complement system, which predisposes some individuals to SLE or autoimmune disease, respectively. This is only speculation but offers one reason for the association of these genes for these disparate proteins in the MHC. Why the polymorphism of C4, C2, and factor B arose in the first place, and what biological advantage has ensured its survival, is not clear. If, however, lysis of pathogens and dissociation and removal of immune aggregates is the primary function of complement, and if an important factor is the binding of C4 (and subsequently C3) to pathogen surfaces, then differences in the C4 molecular surface particularly adjacent to the thiolester bond might be significant. Close association between C4 and the wide variety of structures present in pathogen surfaces might well be facilitated by the availability of a range of different amino acid residues in the sensitive area of the C4 surface. Much more information about all these structures and reactions will be required to determine if any of these speculations have any basis in fact.

REFERENCES

1. **Rynes, R. I.,** *Clinics in Rheumatic Diseases,* Vol. 8, Hughes, G. R. V., Ed., W. B. Saunders, Philadelphia, 1982, 21.
2. **Barnstable, C. J., Jones, E. A., and Bodmer, W. F.,** in *Defense and Recognition IIA,* Lennox, E. S., Ed., University Park Press, Baltimore, 1979, 151.
3. **Dorf, M. E., Ed.,** *The Role of the Major Histocompatibility Complex in Immunobiology,* John Wiley & Sons, New York, 1981, 405.
4. **Klein, J., Juretic, A., Baxevanis, C. N., and Nagy, Z. A.,** *Nature (London),* 291, 455, 1981.
5. **Fielder, A. H. L., Walport, M. J., Batchelor, J. R., Rynes, R. I., Black, C. M., Dodi, I. A., and Hughes, G. R. V.,** *Br. Med. J.,* 286, 425, 1983.
6. **Rittner, C. and Bertrams, J.,** *Hum. Genet.,* 56, 235, 1981.
7. **Bender, K. and Grzeschik, K. H.,** *Hum. Genet.,* 32, 341, 1976.
8. **Weitkamp, L. R. and Guttormsen, S. A.,** *Cytogenet. Cell Genet.,* 16, 364, 1976.
9. **Figueroa, F., Klein, D., Tewarson, S., and Klein, J.,** *J. Immunol.,* 129, 2089, 1982.
10. **Francke, U. and Weitkamp, L. R.,** *Cytogenet. Cell Genet.,* 25, 32, 1979.
11. **Porter, R. R.,** in *Defense and Recognition IIB Structural Aspects,* Lennox, E. S., Ed., University Park Press, Baltimore, 1979, 177.
12. **Lachmann, P. J.,** in *The Antigens,* Vol. 5, Sela, M., Ed., Academic Press, New York, 1979, 28.
13. **Porter, R. R.,** *Proc. R. Soc. Ser. B.,* 210, 477, 1980.
14. **Reid, K. B. M. and Porter, R. R.,** *Ann. Rev. Biochem.,* 50, 433, 1981.
15. **Colten, H. R., Oiy, M., and Edelson, P. J.,** *Ann. N. Y. Acad. Sci.,* 332, 482, 1979.
16. **Christie, D. L., Gagnon, J., and Porter, R. R.,** *Proc. Natl. Acad. Sci. U.S.A.,* 77, 4923, 1980.
17. **Mole, J. E. and Niemann, M. A.,** *J. Biol. Chem.,* 255, 8472, 1980.
18. **Christie, D. L. and Gagnon, J.,** *Biochem. J.,* 201, 555, 1982.
19. **Christie, D. L. and Gagnon, J.,** *Biochem. J.,* 209, 61, 1983.
20. **Gagnon, J. and Christie, D. L.,** *Biochem. J.,* 209, 51, 1983.
21. **Young, C. L., Barker, W. C., Tomasalli, C. M., and Dayhoff, M. O.,** in *Atlas of Protein Sequence,* Vol. 5 (suppl. 3), Dayhoff, M. E., Ed., National Biomedical Research Foundation, Washington, D.C., 1978, 73.
22. **Huber, R. and Bode, W.,** *Acc. Chem. Res.,* 11, 114, 1978.
23. **Kerr, M. A.,** *Biochem. J.,* 183, 615, 1979.
24. **Parkes, C., Gagnon, J., and Kerr, M. A.,** *Biochem. J.,* 213, 201, 1983.
25. **Kerr, M. A. and Gagnon, J.,** *Biochem. J.,* 205, 59, 1982.
26. **Kerr, M. A.,** *Biochem. J.,* 189, 173, 1980.
27. **Polley, M. J. and Muller-Eberhard, M. J.,** *J. Exp. Med.,* 126, 1013, 1967.
28. **Polley, M. J. and Muller-Eberhard, M. J.,** *J. Immunol.,* 102, 1339, 1969.
29. **Lesavre, P. H., Hugli, T. E., Esser, A. F., and Muller-Eberhard, H. J.,** *J. Immunol.,* 23, 529, 1979.
30. **Hall, R. E. and Colten, H. R.,** *J. Immunol.,* 118, 1903, 1977.
31. **Roos, M. H., Atkinson, J. P., and Schreffler, D. C.,** *J. Immunol.,* 121, 1106, 1978.
32. **Parker, K. L., Roos, M. H., and Schreffler, D. C.,** *Proc. Natl. Acad. Sci. U.S.A.,* 76, 5853, 1979.
33. **Gorski, T. P. and Muller-Eberhard, H. J.,** *J. Immunol.,* 120, 1775, 1978.
34. **Gigli, I.,** *Nature (London),* 272, 836, 1978.
35. **Chan, A. C., Mitchell, K. R., Munns, T. W., Karp, D. R., and Atkinson, J. P.,** *Proc. Natl. Acad. Sci. U.S.A.,* 80, 268, 1978.
36. **Schreiber, R. D. and Muller-Eberhard, H. J.,** *J. Exp. Med.,* 140, 1324, 1974.
37. **Nagasawa, S. and Stroud, R. M.,** *Fed. Proc. Fed. Am. Soc. Exp. Biol.,* 35, 654, 1976.
38. **Matthews, W. T., Goldberger, G., Marino, J. T., Einstein, L. P., Gash, D. J., and Colten, H. R.,** *Biochem. J.,* 204, 839, 1982.
39. **Karp, D. R., Atkinson, J. P. and Schreffler, D. C.,** *J. Biol. Chem.,* 257, 7330, 1982.
40. **Smith, M. A., Gerrie, L. M., Dunbar, B., and Fothergill, J. E.,** *Biochem. J.,* 207, 253, 1982.
41. **Moon, K. E., Gorski, J. P., and Hugli, T. E.,** *J. Biol. Chem.,* 256, 8685, 1981.
42. **Hugli, T. E. and Muller-Eberhard, H. J.,** *Adv. Immunol.,* 26, 1, 1978.
43. **Gorski, T. P., Hugli, T. E., and Muller-Eberhard, H. J.,** *Proc. Natl. Acad. Sci. U.S.A.,* 76, 5299, 1979.
44. **Huber, R., Scholze, H., Paques, E. F., and Diesenhofer, J.,** *Hoppe-Seyler Z. Physiol. Chem.,* 361, 1389, 1980.
45. **Press, E. M. and Gagnon, J.,** *Biochem. J.,* 199, 351, 1981.
46. **Law, S. K. and Levine, R. F.,** *Proc. Natl. Acad. Sci. U.S.A.,* 74, 2701, 1977.
47. **Law, S. K., Lichtenberg, N. A., and Levine, R. P.,** *J. Immunol.,* 123, 1388, 1979.
48. **Janatova, J., Tack, B. F., and Prahl, J. W.,** *Biochemistry,* 19, 4479, 1980.

49. **Janatova, J., Lorenz, P. E., Schecter, A. N., Prahl, J. W., and Tack, B. F.,** *Biochemistry,* 19, 4471, 1980.
50. **Tack, B. F., Harrison, R. A., Janatova, J., Thomas, M. L., and Prahl, J. W.,** *Proc. Natl. Acad. Sci. U.S.A.,* 77, 5764, 1980.
51. **Campbell, R. D., Gagnon, J., and Porter, R. R.,** *Biochem. J.,* 199, 359, 1981.
52. **Chakravarti, D. N., Campbell, R. D., and Gagnon, J.,** *FEBS Lett.,* 154, 387, 1983.
53. **Sim, R. B., Twose, T. M., Patterson, D. S., and Sim, E.,** *Biochem. J.,* 193, 115, 1981.
54. **Isenman, D. E., Kells, D. I. C., Cooper, N. R., Muller-Eberhard, H. J., and Pangburn, M. K.,** *Biochemistry,* 20, 4458, 1981.
55. **Law, S. K.,** *Biochem. J.,* in press.
56. **Harpel, P. C., Hayes, M. B., and Hugli, T. E.,** *J. Biol. Chem.,* 254, 8669, 1979.
57. **Svenson, R. P. and Howard, J. B.,** *J. Biol. Chem.,* 254, 4452, 1979.
58. **Sim, R. B. and Sim, E.,** *Biochem. J.,* 193, 129, 1981.
59. **Janatova, J. and Tack, B. F.,** *Biochemistry,* 20, 2394, 1981.
60. **Khan, S. A. and Erickson, B. W.,** *J. Am. Chem. Soc.,* 103, 7374, 1981.
61. **Khan, S. A. and Erickson, B. W.,** *J. Biol. Chem.,* 257, 11864, 1982.
62. **Davies, S. G. and Sim, R. B.,** *Biosci. Rep.,* 1, 461, 1981.
63. **Howard, J. B.,** *Proc. Natl. Acad. Sci. U.S.A.,* 78, 2235, 1981.
64. **Campbell, R. D., Dodds, A. W., and Porter, R. R.,** *Biochem. J.,* 189, 67, 1980.
65. **Reboul, A., Thielens, N., Villiers, M. B., and Colomb, M. G.,** *FEBS Lett.,* 103, 156, 1979.
66. **Fu, S. M., Kunkel, H. G., Brusman, H. P., Allen, F. H., and Fotino, M.,** *J. Exp. Med.,* 140, 1108, 1974.
67. **Fu, S. M., Stern, R., Kunkel, H. G., Dupont, B., Hansen, J. A., Day, N. K., Good, R. A., Jersild, C., and Fotino, M. J.,** *J. Exp. Med.,* 142, 495, 1975.
68. **Day, N. K., Lesperance, P., Good, R. A., Michael, A. F., Hansen, J. A., Dupont, B., and Jersild, C.,** *J. Exp. Med.,* 141, 1464, 1975.
69. **Pariser, K. M., Raum, D., Berkman, E. M., Alper, C. A, and Agnello, V.,** *J. Immunol.,* 131, 2580, 1978.
70. **Marshall, W. H., Grandy, R., and Schroeder, M. L.,** *Histocompatibility Testing 1980,* Teraski, P. I., Ed., Munksgaard, Copenhagen, 1980, 937.
71. **Mortensen, J. P., Buskjaer, L., and Lamm, L. U.,** *Immunology,* 39, 541, 1980.
72. **Meo, T., Atkinson, J., Bernoco, M., Bernoco, D., and Cepellini, R.,** *Eur. J. Immunol.,* 6, 916, 1976.
73. **Hobart, M. J. and Lachmann, P. J.,** *Transplant. Rev.,* 32, 26, 1976.
74. **Raum, D., Glass, D., Carpenter, C. B., Alper, C. A., and Schur, P. H.,** *J. Clin. Invest.,* 58, 1240, 1976.
75. **Weitkamp, L. R. and Lamm, L. U.,** *Cytogenet. Cell Genet.,* 32, 130, 1982.
76. **Allen, F. H.,** *Vox. Sang.,* 27, 382, 1974.
77. **Alper, C. A., Boenish, T., and Watson, L.,** *J. Exp. Med.,* 135, 68, 1972.
78. **Mauff, G., Hauptmann, G., Hitzerothe, H. W., Ganchell, F., and Scherz, R.,** *Z. Immunitatforschung. Immunobiol.,* 27, 382, 1978.
79. **Dykes, D. D., Polesky, H. F., and Crawford, M. H.,** *Electrophoresis,* 2, 320, 1981.
80. **Arnason, A., Larsen, B., Marshall, W. H., Edwards, J. H., Mackintosh, P., Olaisen, B., and Teisberg, P.,** *Nature (London),* 268, 527, 1977.
81. **Olaisen, B., Teisberg, P., Jonassen, R., Gedde-Dahl, T., Moen, T., and Thorsby, E.,** *Hum. Immunol.,* 2, 247, 1981.
82. **Alper, C. A.,** *J. Exp. Med.,* 144, 1111, 1976.
83. **Hauptmann, G., Grosshans, E., and Heid, E.,** *Ann. Dermatol. Syphiligr.,* 101, 479, 1974.
84. **Rittner, C., Hauptmann, G., Gross-Wilde, H., Grosshaus, E., Tongio, M. M., and Mayer, S.,** *Histocompatibility Testing 1975,* Kissmeyer-Nielsen, F., Ed., Munksgaard, Copenhagen, 1975.
85. **Ochs, H. S., Rosenfeld, S. I., Thomas, E. D., Giblett, E. R., Alper, C. A., Dupont, B., Schaller, J. G., Gilliland, B. C., Hansen, J. A., and Wedgwood, R. J.,** *N. Engl. J. Med.,* 296, 470, 1977.
86. **Teisberg, P., Akesson, I., Olaisen, B., Gedde-Dahl, T., and Thorsby, E.,** *Nature (London),* 264, 253, 1976.
87. **Teisberg, P., Olaisen, B., Jonassen, R., Gedde-Dahl, T., and Thorsby, E.,** *J. Exp. Med.,* 146, 1380, 1977.
88. **Middleton, J., Crookson, M. C., Falk, J. A., Robson, E. B., Cook, P. J., Batchelor, J. R., Bodmer, J., Ferrara, G. B., Festenstein, H., Harris, R., Kissmeyer-Nielsen, F., Lawler, S. D., Sachs, J. A., and Wolfe, E.,** *Tissue Antigens,* 4, 3365, 1974.
89. **Gedde-Dahl, T., Giles, C. M., Thorsby, E., Olaisen, B., and Robson, E. B.,** *Cytogenet. Cell Genet.,* 6, 307, 1976.
90. **Longster, G. and Giles, C. M.,** *Vox Sang.,* 30, 175, 1976.

91. **O'Neill, G. J., Yang, S., Tepoli, J., Berger, R., and Dupont, B.,** *Nature (London)*, 273, 668, 1978.
92. **Tilley, C. A., Romans, D. G., and Crookston, M. C.,** *Nature (London)*, 276, 713, 1978.
93. **Law, S. K., Lichtenberg, N. A., Holcombe, F. H., and Levine, R. P.,** *J. Immunol.*, 125, 634, 1980.
94. **Chu, V. F. H., March, W. L., and Gigli, I.,** *J. Immunol.*, 128, 181, 1982.
95. **Lundwall, A., Hellman, U., Eggersten, G., and Sjoquist, J.,** *Molec. Immunol.*, 19, 1655, 1982.
96. **Rosenfeld, S. I., Ruddy, S., and Austen, K. F.,** *J. Clin. Invest.*, 48, 2283, 1969.
97. **O'Neill, G. J., Yang, S., and Dupont, B.,** *Proc. Natl. Acad. Sci. U.S.A.*, 75, 5165, 1978.
98. **Mauff, G., Bender, K., and Fischer, B.,** *Vox Sang.*, 34, 296, 1978.
99. **Olaisen, B., Teisberg, P., Nordhagen, R., Michaelson, T., and Gedde-Dahl, T.,** *Nature (London)*, 279, 736, 1979.
100. **Awdeh, Z. L. and Alper, C. A.,** *Proc. Natl. Acad. Sci. U.S.A.*, 77, 8576, 1980.
101. **Mevag, B., Olaisen, B., and Teisberg, P.,** *Scand. J. Immunol.*, 14, 303, 1981.
102. **Teisberg, P., Olaisen, B., Nordhagen, R., Thorsby, E., and Gedde-Dahl, T.,** *Immunobiology*, 158, 91, 1980.
103. **Bruun-Peterson, G., Lamm, L. U., Sorenson, I. J., Buskjaer, L., and Morrtensen, J. P.,** *Hum. Genet.*, 58, 260, 1981.
104. **O'Neill, G. J., Miniter, P., Pollack, M. S., and Dupont, B.,** *Hum. Immunol.*, 1, 23, 1980.
105. **Mauff, G., Alper, C. A., Awdeh, Z., Batchelor, J. R., Bertrams, J., Bruun-Peterson, G., Dawkins, R. L., Demant, P., Edwards, J. H., Grosse-Wilde, H., Hauptmann, G., Klouda, P., Lamm, L. U., Molenaur, E., Nwerl, C., Olaisen, B., O'Neill, G., Rittner, C., Roos, M. H., Skanes, V., Teisberg, P., and Wells, L.,** *Immunobiology*, 164, 184, 1983.
106. **Schreffler, D. C. and Owen, R. D.,** *Genetics*, 48, 9, 1963.
107. **Demant, P., Capkova, J., Hinzova, E., and Voracova, B.,** *Proc. Natl. Acad. Sci. U.S.A.*, 70, 863, 1973.
108. **Meo, T., Krasteff, T., and Schreffler, D. C.,** *Proc. Natl. Acad. Sci. U.S.A.*, 72, 4536, 1975.
109. **Lachmann, P. J., Grennan, D., Martin, A., and Demant, P.,** *Nature (London)*, 258, 243, 1975.
110. **Curman, B., Ostberg, L., Sanberg, L., Malmheden-Eriksson, I., Stalenhaim, G., Rask, L., and Peterson, P. A.,** *Nature (London)*, 258, 243, 1975.
111. **Carroll, M. C. and Capra, J. D.,** *Proc. Natl. Acad. Sci. U.S.A.*, 75, 2424, 1978.
112. **Ferreira, A., Nussenzweig, V., and Gigli, I.,** *J. Exp. Med.*, 148, 1186, 1978.
113. **Karp, D. R., Capra, J. D., Atkinson, J. P., and Schreffler, D. C.,** *J. Immunol.*, 128, 2336, 1982.
114. **Karp, D. R., Parker, K. L., Schreffler, D. C., Slaughter, C., and Capra, J. D.,** *Proc. Natl. Acad. Sci. U.S.A.*, 79, 6347, 1982.
115. **Carroll, M. C., Passmore, H. C., and Capra, J. D.,** *J. Immunol.*, 124, 1745, 1980.
116. **McDevitt, H. O., Deak, B. D., Shreffler, D. C., Klein, J., Stimpfling, J. H., and Snell, G. D.,** *J. Exp. Med.*, 135, 1259, 1972.
117. **Curry, R. A., Dierich, M. P., Pellegrino, A., and Hoch, J. A.,** *Immunogenetics*, 3, 465, 1976.
118. **Arnaiz-Villena, A. and Festenstein, H.,** *Nature (London)*, 258, 732, 1975.
119. **Galili, U., Branbar, C., and Schlesinger, M.,** *Tissue Antigens*, 10, 99, 1977.
120. **Few, A. J. and Bright, S.,** *Tissue Antigens*, 10, 315, 1977.
121. **Ferreira, A., Fontina, M., and Nussenzweig, V.,** *Eur. J. Immunol.*, 5, 832, 1976.
122. **Bishop, C., Demant, P., and Capel, P. J. A.,** *Tissue Antigens*, 15, 31, 1980.
123. **Kaido, T., Natsuume-Bakai, S., and Takahashi, M.,** *Proc. Natl. Acad. Sci. U.S.A.*, 78, 3794, 1981.
124. **Steinmetz, M., Minard, K., Horrath, S., McNicholas J., Srelinger, J., Wake, C., Long, E., Mach, B., and Hood, L.,** *Nature (London)*, 330, 35, 1982.
125. **Steinmetz, M., Winoto, A., Minard, K., and Hood, L.,** *Cell*, 28, 489, 1982.
126. **Goodenow, R. S., McMillan, M., Nicolson, M., Sher, B. T., Eakle, K., Davidson, N., and Hood, L.,** *Nature (London)*, 300, 231, 1982.
127. **Pease, L. R., Nathenson, S. G., and Leinwand, L. A.,** *Nature (London)*, 298, 382, 1982.
128. **Wake, C. T., Long, E. D., and Mach, B.,** *Nature (London)*, 300, 372, 1982.
129. **Lee, J. S., Trowsdale, J., and Bodmer, W. F.,** *Proc. Natl. Acad. Sci. U.S.A.*, 79, 545, 1982.
130. **Lee, J. S., Trowsdale, J., Travers, P. J., Carey, J., Grosveld, F., Jenkins, J., and Bodmer, W. F.,** *Nature (London)*, 299, 750, 1982.
131. **Domdey, H., Wiebauer, K., Kazmaier, M., Muller, V., Odink, K., and Fey, G.,** *Proc. Natl. Acad. Sci. U.S.A.*, 79, 7619, 1982.
132. **Whitehead, A. S., Solomon, E., Chambers, S., Bodmer, W. F., Povey, S., and Fey, G.,** *Proc. Natl. Acad. Sci. U.S.A.*, 79, 5021, 1982.
133. **Natsuume-Sakai, S., Hayakawa, J., Amano, S., and Takahashi, M.,** *J. Immunol.*, 123, 947, 1979.
134. **Rubinstein, P., Vienne, K., Hoecker, G. F.,** *J. Immunol.*, 122, 2584, 1979.
135. **Woods, D. E., Markham, A. F., Ricker, A. T., Goldberger, G., and Colten, H. R.,** *Proc. Natl. Acad. Sci. U.S.A.*, 79, 5661, 1982.

136. **Campbell, R. D. and Porter, R. R.,** *Proc. Natl. Acad. Sci. U.S.A.*, 80, 4464, 1983.
137. **Grosveld, F. G., Dahl, H. M., Boer, E. D., and Flavell, R. A.,** *Gene*, 13, 227, 1981.
138. **Craik, C. S., Sprang, S., Fletterick, R., and Rutter, W. J.,** *Nature (London)*, 299, 180, 1982.
139. **Carroll, M. C. and Porter, R. R.,** *Proc. Natl. Acad. Sci. U.S.A.*, 80, 264, 1983.
140. **Palsdottir, A., Arnason, A., Cross, S., and Carroll, M. C.,** Abstr. 10th Int. Complement Workshop, 1983.
141. **Batchelor, J. R. and Welsh, K. I.,** in *Clinical Aspects of Immunology*, Lachmann, P. J., and Peters, D. K., Eds., Blackwell Scientific, Oxford, 1982.
142. **Moller, G., Ed.,** *Immunol. Rev.*, 70, Munksgaard, Copenhagen, 1983, 218.
143. **Porter, R. R.,** *Molec. Biol. Med.*, 1, 161, 1983.
144. **Awdeh, Z. L., Raum, D., Yunis, E. J., and Alper, C. A.,** *Proc. Natl. Acad. Sci. U.S.A.*, 80, 259, 1983.
145. **Gagnon, J.,** unpublished data.
146. **Carrol, M. C. and Belt, ,** private communication.
147. **Roos, M. H. and Demant P.,** *Immunogenetics*, 15, 23, 1982.
148. **Gorman, J. C., Jackson, R., Desantola, J. R., Schreffler, D., and Atkinson, J. P.,** *J. Immunol.*, 125, 344, 1980.