Protein Science (1993), *2*, 706–716. Cambridge University Press. Printed in the USA.

Covalent binding properties of the human complement protein C4 and hydrolysis rate of the internal thioester upon activation

ARMIN SEPP, ALISTER W. DODDS, MICHAEL J. ANDERSON, R. DUNCAN CAMPBELL, ANTONY C. WILLIS, AND S.K. ALEX LAW

The MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, Oxford OX1 3QU, United Kingdom

(RECEIVED August 6, 1992; REVISED MANUSCRIPT RECEIVED January 18, 1993)

Abstract

The complement proteins C3 and C4 have an internal thioester. Upon activation on the surface of a target cell, the thioester becomes exposed and reactive to surface-bound amino and hydroxyl groups, thus allowing covalent deposition of C3 and C4 on these targets. The two human C4 isotypes, C4A and C4B, which differ by only four amino acids, have different binding specificities. C4A binds more efficiently than C4B to amino groups, and C4B is more effective than C4A in binding to hydroxyl groups. By site-directed mutagenesis, the four residues in a cDNA clone of C4B were modified. The variants were expressed and their binding properties studied. Variants with a histidine residue at position 1106 showed C4B-like binding properties, and those with aspartic acid, alanine, or asparagine at the same position were C4A-like. These results suggest that the histidine is important in catalyzing the reaction of the thioester with water and other hydroxyl group-containing compounds. When substituted with other amino acids, this reaction is not catalyzed and the thioester becomes apparently more reactive with amino groups. This interpretation also predicts that the stability of the thioester in C4A and C4B, upon activation, will be different. We measured the time course of activation and binding of glycine to C4A and C4B. The lag in the binding curve behind the activation curve for C4A is significantly greater than that for C4B. The hydrolysis rates (k_0) of the thioester in the activated proteins were estimated to be 0.068 s^{-1} ($t_{1/2}$ of 10.3 s) for C4A and 1.08 s^{-1} ($t_{1/2}$ of 0.64 s) for C4B. These results indicate that the difference in hydrolysis rate of the thioester accounts, at least in part, for the difference in the binding properties of C4A and C4B.

Keywords: C3; C4; complement proteins; covalent binding reaction; internal thioester

The complement proteins participate in a number of activities in host defense against infection (Law & Reid, 1988). Normally present in inert forms, they are activated via the classical and/or the alternative pathways when challenged by the presence of material foreign to the host. Activation of the two complement pathways converges on the activation of the component C3 and the deposition of C3b on the targets. The surface-bound C3b thus serves as a marker for recognition by phagocytes that bear receptors specific for various C3 fragments and as part of the enzyme that activates the lytic components of complement.

Reprint requests to: S.K. Alex Law, The MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, United Kingdom.

C3b binds covalently to target surfaces (Law & Levine, 1977). A thioester bond is present in native C3 as a 15-member thiolactone ring between the thiol group of the cysteine residue and the glutamyl group of the glutamine residue in the tetrapeptide -Cys-Gly-Glu-Gln- and is normally inaccessible to external molecules (Tack et al., 1980; de Bruijn & Fey, 1985). Activation of C3 to C3b involved the removal of C3a from C3 and a concomitant conformational change in C3b (Isenman et al., 1981), resulting in the exposure of the thioester, which is free to react with water as well as other compounds with hydroxyl or amino groups. When C3 is activated on a target cell, the thioester of some of C3b reacts with hydroxyl or amino groups on the cell surface resulting in the covalent binding of C3b to the cell by way of ester or amide bonds (for review see Law [1983a]). This unusual internal thioester is also found

in the complement protein C4 (Harrison et al., 1981) and in a group of protease inhibitors represented by the α_2-macroglobulin (α_2M) (Howard, 1983). These proteins are structurally similar to C3. The covalent binding of C4 to target cells and immune complexes is again the key of its functional activities. The protease inhibitory activities of α_2M and related molecules may also be attributed, at least in part, to their covalent binding via the thioester to the proteases (see Sottrup-Jensen, 1987).

Although the activation and covalent binding of C3, C4, and α_2M to immune targets and proteases may be similarly described, the three proteins have different binding properties to molecules with hydroxyl and amino groups (Isenman & Young, 1984; Law et al., 1984a; Dodds & Law, 1988). At neutral pH, C3 predominantly binds to hydroxyl groups and α_2M to amino groups. Of the two human C4 isotypes, C4A binds to amino groups predominantly and with a much higher reactivity than α_2M, whereas C4B binds to amino groups with efficiency similar to that of α_2M and to hydroxyl groups with efficiency similar to that of C3 (see Table 1). C4A and C4B are virtually identical in sequence except for four residues in a stretch of six from positions 1101 to 1106: the six residues in C4A are PCPVLD and those in C4B are LSPVIH (Belt et al., 1984, 1985; Yu et al., 1986). It was therefore concluded that these residues must play a role in determining the difference in the binding of C4A and C4B to amino and/or hydroxyl groups (Dodds et al., 1986). Mouse C4, which has the hybrid sequence of PCPVIH (Nonaka et al., 1985; Sepich et al., 1985), is C4B-like in its binding activities (Dodds & Law, 1988). Thus, it is likely that the LD in C4A and IH in C4B are the specificity-determining residues, and it was further argued that the key difference is in the D/H (Asp/His) substitution (Dodds & Law, 1988). Both C4A- and C4B-like proteins are found in sheep and cattle (Dodds & Law, 1990), and two different sequences coding for the residues PFPVMD and PCPVIH are found in the C4 genes in both animals (Ren et al., 1993). Although direct correlation between the sequences and binding properties of the proteins had not been made, it is likely that the PFPVMD sequence is found in the C4A-like proteins and the PCPVIH sequence in the C4B-like proteins in both sheep and cattle.

By expressing C4 and a number of variants modified in this region and studying their binding to erythrocyte membranes, Carroll et al. (1990) showed that the Asp/His difference is indeed of major importance in determining the binding specificities of C4A and C4B although other residues in the hexapeptide assume minor roles. In addition, they showed that the replacement of either Asp or His with Ala resulted in a C4 molecule with C4A-like binding properties. In this paper, we report experiments on a number of modified C4 molecules. The binding properties of the expressed proteins were determined in a fluid phase system in which labeled glycine (amino group) and glycerol (hydroxyl group) were used as binding substrates (Law et al., 1984a,b). From these results, we were able to predict, and demonstrate, that the difference in C4A and C4B binding to amino and hydroxyl

Table 1. *The k_2/k_0 values for glycine and glycerol of human C4 and its variants*

	Residues 1101–1106	k_2/k_0: Glycine (M^{-1})					k_2/k_0: Glycerol (M^{-1})				
		PB[a]	1	2	3	4	PB	1	2	3	4
Plasma C4											
C4A	PCPVLD	13,400	–[b]	30,900	NM[c]	–	1.3	–	4.5	NM	–
C4B	LSPVIH	119	176	117	110	–	15.5	14.2	10.3	11.5	–
C4–mouse	PCPVIH	136	–	–	–	–	26.0	–	–	–	–
C3	DAPVIH	0	–	–	–	–	23.0	–	–	–	–
α_2M	SGSLLN	206	–	–	–	–	1.2	–	–	–	–
Expressed C4											
C4A	PCPVLD		61,300	NM	NM	NM		3.8	NM	NM	NM
C4B	LSPVIH		–	NM	147	–		–	–	NM	7.5
C4(mouse)	PCPVIH		245	144	259	–		20.6	12.6	18.3	–
C4(C3)	DAPCIH		–	138	112	127		–	14.3	15.1	13.0
C4(α_2M)	SGSLLN		–	14,100	NM	16,600		–	1.9	NM	0[d]
C4A(D → A)	PCPVIA		–	–	20,400	20,600		–	–	6.6	0.2
C4B(H → A)	LSPVIA		–	–	23,400	21,300		–	–	6.6	0[d]

[a] From previous publications (Dodds & Law, 1988).
[b] Not done.
[c] Not measurable: the experiments were done but the C4 was found to have been inactivated upon storage.
[d] The background on binding is taken to be the amount bound to the β-chain of C4. The background binding of glycerol is rather variable, depending on the batch of [^{3}H]glycerol used, thus accounting for the variable binding results especially for C4A-like molecules.

groups was due, at least in part, to the difference in the hydrolysis rate of the activated C4.

Kinetics of the binding reaction of C4 with a small molecule G

The binding reaction of C4 to a small molecule G (glycine or glycerol in the experiments described in later sections) can be described by the following reaction scheme:

$$\text{C4} \xrightarrow[\text{activation}]{k} \text{C4b}^* \begin{cases} \xrightarrow[\text{hydrolysis}]{k_0} \text{C4b} \\ \xrightarrow[\text{binding}]{k_2} \text{C4b-G} \end{cases} \quad +\text{G}$$

where C4 is activated to C4b* at a first-order rate k. C4b* either hydrolyzes in the medium at a first-order rate of k_0 or reacts with a molecule with an appropriate nucleophilic group, G, at a second-order rate of k_2. At time t, the reaction above can be described by the following equations:

$$\frac{d}{dt}[\text{C4}] = -k[\text{C4}]; \tag{1}$$

$$\frac{d}{dt}[\text{C4b}^*] = k[\text{C4}] - (k_0 + k_2[\text{G}])[\text{C4b}^*]; \tag{2}$$

$$\frac{d}{dt}[\text{C4b}] = k_0[\text{C4b}^*]; \tag{3}$$

$$\frac{d}{dt}[\text{C4b-G}] = k_2[\text{G}][\text{C4b}^*]. \tag{4}$$

Setting initial conditions, and under the pseudo first-order condition where $[\text{G}] \gg [\text{C4}]$ such that the amount of glycine bound to C4b is insignificant in comparison with [G], the concentrations of various species of C4, with respect to total C4, $[\text{C4}_0]$, are given by

$$\frac{[\text{C4}(t)]}{[\text{C4}_0]} = e^{-kt}; \tag{5}$$

$$\frac{[\text{C4b}^*]}{[\text{C4}_0]} = \frac{k}{k_0 + k_2[\text{G}] - k}\{e^{-kt} - e^{-(k_0+k_2[\text{G}])t}\}; \tag{6}$$

$$\frac{[\text{C4b-G}(t)]}{[\text{C4}_0]} = \frac{k_2[\text{G}]}{(k_0 + k_2[\text{G}] - k)} \times \left\{(1 - e^{-kt}) - \frac{k}{(k_0 + k_2[\text{G}])} \times (1 - e^{-(k_0+k_2[\text{G}])t})\right\}; \tag{7}$$

$$\frac{[\text{C4b}(t)]}{[\text{C4}_0]} = \frac{k_0}{(k_0 + k_2[\text{G}] - k)} \times \left\{(1 - e^{-kt}) - \frac{k}{(k_0 + k_2[\text{G}])} \times (1 - e^{-(k_0+k_2[\text{G}])t})\right\}. \tag{8}$$

If the reaction is allowed to proceed to completion, i.e., for time $\to \infty$, then Equation 7 becomes

$$\frac{[\text{C4b-G}(t \to \infty)]}{[\text{C4}_0]} = \frac{k_2[\text{G}]}{k_0 + k_2[\text{G}]} = \text{BE}, \tag{9}$$

which is what has been referred to as the binding efficiency (BE) in previous publications (Law et al., 1981, 1984b). It should be noted that Equation 9 can be rearranged to give

$$\frac{k_2}{k_0} = \frac{1}{[\text{G}]} \cdot \frac{\text{BE}}{(1 - \text{BE})}. \tag{10}$$

Thus k_2/k_0 can be determined by measuring BE at a known concentration of G. BE, however, is determined as a ratio of $[\text{C4b-G}(t \to \infty)]$ and $[\text{C4}_0]$. An independent measurement of $[\text{C4}_0]$ is required since C4 preparations contain varying amounts of inactive material. This can be achieved either by the incorporation of radioactively labeled methylamine, which reacts quantitatively with the internal thioester (Law et al., 1980), or by the measurement of $[\text{C4b-G}(t \to \infty)]$ at different concentrations of G. Equation 9 can also be rearranged to give

$$\frac{1}{[\text{C4b-G}(t \to \infty)]} = \frac{1}{[\text{C4}_0]} + \frac{k_0}{k_2} \cdot \frac{1}{[\text{C4}_0]} \cdot \frac{1}{[\text{G}]}. \tag{11}$$

A double reciprocal plot of $[\text{C4b-G}(t \to \infty)]$ (in arbitrary units) vs. [G] should yield a straight line with the y-intercept giving the value of $1/[\text{C4}_0]$ in the same arbitrary units as those for $[\text{C4b-G}(t \to \infty)]$ (Law et al., 1981). BE can therefore be calculated accordingly.

The fraction of C4 activated at time t can be derived from Equation 5 and is given by

$$\frac{[\text{C4}_0] - [\text{C4}(t)]}{[\text{C4}_0]} = 1 - e^{-kt}. \tag{12}$$

The proportion of C4b-G at time t in comparison with C4b-G at the completion of reaction is given by the ratio:

$$\frac{[\text{C4b-G}(t)]}{[\text{C4b-G}(t \to \infty)]} = 1 - \frac{k_0 + k_2[\text{G}]}{k_0 + k_2[\text{G}] - k} e^{-kt} + \frac{k}{k_0 + k_2[\text{G}] - k} e^{-(k_0+k_2[\text{G}])t}. \tag{13}$$

It should be noted that for $k_0 + k_2[G] \gg k$, Equation 13 becomes $(1 - e^{-kt})$, which is identical to Equation 12. That is, if C4b* is very reactive, then the curves for the time course of activation and binding are superimposable. However, for $k_0 + k_2[G]$ of intermediate values, i.e., similar to k, the binding curve lags behind the activation curve, and the difference, which may be regarded as the proportion of C4 activated but not hydrolyzed nor bound with glycine, is $[C4b^*]/[C4_0]$. Thus

$$\frac{[C4b^*]}{[C4_0]} = \frac{[C4_0] - [C4(t)]}{[C4_0]} - \frac{[C4b\text{-}G(t)]}{[C4b\text{-}G(t \to \infty)]}. \quad (14)$$

Substituting the two terms in Equation 14 with expressions in Equations 12 and 13, $[C4b^*]/[C4_0]$ can be expressed as a function of k, k_0, t, and BE to yield

$$\frac{[C4b^*]}{[C4_0]} = \frac{k(1 - BE)}{k_0 - k(1 - BE)} \{e^{-kt} - e^{-k_0 t/(1-BE)}\}. \quad (15)$$

In the experiments described later, C4 is activated by C1s cleavage. Thus the kinetics of C4 activation is more complicated than stated in Equation 1, because it is dependent on the interaction between C4 and C1s. However, in these experiments, our interest is focused on the generation of C4b*, which is quantitated by the decrease of C4 (or the increase in C4b). This determination is valid on the assumption that C4b* → C4b and C4b-G are irreversible. The activation data for both C4A and C4B displayed a first-order dependence on time (Fig. 1) and gave R^2 (square of regression coefficient) values of 0.99, in both cases, when fitted to Equation 12. Activation of C4 may therefore be approximated by Equation 1 and the rate of activation by k in the equation. The amount of C1s used was carefully chosen in these experiments and was predetermined in trials to give the desired rate of activation. Higher amounts of C1s resulted in an activation rate that was too fast to be determined, and lower amounts of C1s resulted in too slow an activation for the difference in binding and activation to be observed.

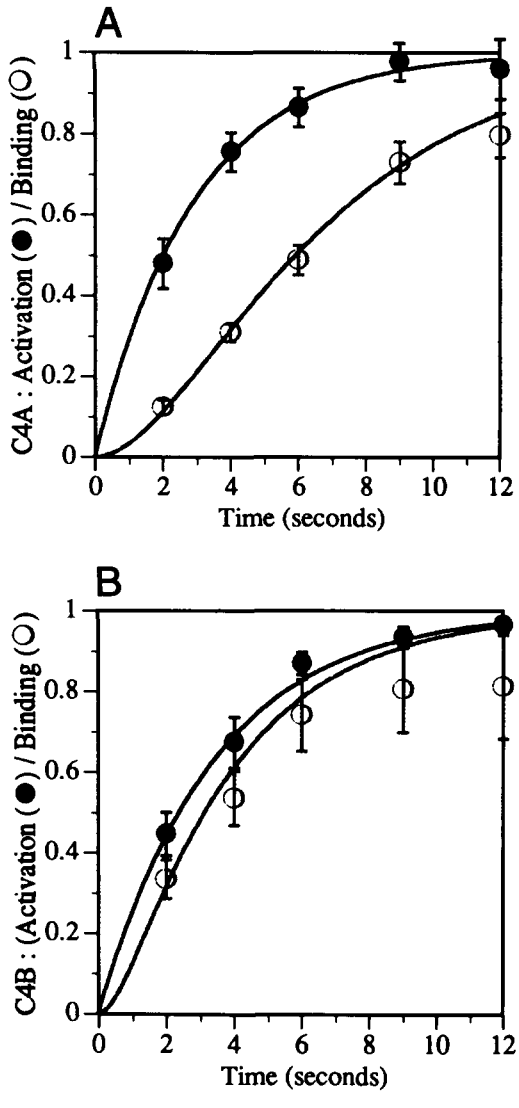

Fig. 1. Time course of activation (●) and binding of glycine (○) to C4A (**A**) and C4B (**B**). Activation was determined by the decrease in stain intensity of the α-chain of C4. The binding at time t was normalized to the binding at completion of the reaction, i.e., $[C4b\text{-}G(t)]/[C4b\text{-}G(t \to \infty)]$. By curve fitting the data on activation (●) to Equation 12, values of k for C4A and C4B were found to be 0.35 and 0.29 s^{-1}, respectively. The binding data (○) were fitted to Equation 13, where the term $k_2[G]$ was substituted with $k_0 BE/(1 - BE)$ (from Equation 10). The values for k were from the activation curves, and the values for BE (0.71 and 0.25 for C4A and C4B, respectively) were obtained from Table 2. The k_0 values of 0.068 and 1.08 s^{-1} were obtained for C4A and C4B, respectively.

Results

Expression of C4 and its variants

Seven C4-producing cell lines were obtained. The C4 produced from each cell line is on a C4B background but with different residues at positions 1101-1106. Depending on the nature of these six residues, the expressed proteins, as well as their corresponding cell lines, are called C4A (PCPVLD), C4B (LSPVIH), C4(mouse) (PCPVIH), C4(C3) (DAPVIH), C4(α_2M) (SGSLLN), C4A(D → A) (PCPVLA), and C4B(H → A) (LSPVIA). Confluent cells were grown in medium supplemented with 2 mM sodium butyrate for 2 days, which was determined to give maximum yield of C4. C4 was purified from 250 mL of tissue culture supernatant. The yield of C4 from different cell lines ranges from 100 to 500 μg.

C4 is synthesized as a single chain molecule (pro-C4), which is subsequently processed by proteolysis into a

three-chain molecule (Hall & Colten, 1977). The order of the three chains in pro-C4 is β–α–γ (Belt et al., 1984). The expressed proteins were analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions. In addition to the α, β, and γ chains, bands of higher molecular weight were also observed (not shown). These bands were blotted onto Immobilon, and N-terminal sequences were obtained. These were identical to those of either the β- or the α-chains of C4 and were tentatively identified as pro-C4, and the β–α- and α–γ-chains in partially processed molecules. If the proteins were treated with labeled methylamine, which specifically binds to the internal thioester in the α-chain of C4, before analysis, some of the higher molecular weight bands were found to be labeled. These findings suggest that not all the expressed C4 molecules were fully processed into the three-chain structure, and the internal thioester is present in some of the partially processed molecules. In our binding experiments, which involved the analysis of the fully processed α- and α'-chains after SDS-PAGE, the presence of partially processed C4 would not affect the results obtained.

Binding studies

The results of the binding assays for the seven expressed C4 constructs, as well as for C4A and C4B purified from plasma, obtained in four experiments, are shown in Table 1. C4A and C4B have k_2/k_0 rates for glycine and glycerol similar to the plasma proteins. C4(mouse) and C4(C3) are similar to C4B, whereas the remainder are C4A-like. It should be noted that the k_2/k_0 values for glycerol show large variation due to the fact that the background binding, taken as the binding to the β-chain, is dependent on the batch of ^{3}H-glycerol used. However, the identification of A-ness and B-ness for the proteins in each experiment is consistent with the distinct patterns obtained with the glycine-binding results. Thus, C4(mouse) is similar to mouse C4 purified from plasma and has C4B-like activity. However, replacing the six residues by the corresponding ones of C3 or α_2M was not sufficient to convert C4 into molecules with C3 or α_2M binding properties, respectively. Hence, other residues in C3 and α_2M must also play a part in determining their respective reaction with amino and hydroxyl groups. The replacement of Asp in C4A and His in C4B with Ala results in molecules with C4A-like activities, thus confirming the results reported by Carroll et al. (1990).

It appears that whereas the His residue at position 1106 is required to confer C4B-like binding activities, the requirement for the Asp residue for C4A-like activities is not absolute. We therefore argue that the His may play a catalytic role in the reaction with a hydroxyl group including the hydrolysis of the thioester. In its absence, the hydrolysis of the thioester is not catalyzed, and the reaction of the thioester with an amino group becomes apparently more efficient. If this argument is correct, the hydrolysis rate of the thioester in C4A and C4B should be different and the time course of binding, with respect to activation, should also be different.

Time course of activation and glycine binding of C4A and C4B

C4A and C4B were activated by C1s in the presence of [^{3}H]glycine. The experimental conditions were chosen such that both activation and binding could be determined reasonably accurately. C1s was chosen to give about 50% activation between 2 and 5 s. [G] was chosen to be 0.05 mM for C4A and 2.5 mM for C4B, giving a BE of about 0.5. The experiment was performed at room temperature (22 °C) for 2–12 s. At the set times, half of the reaction mixture was removed and added to trichloroacetic acid (TCA) to a final concentration of 30%. The remaining half of the reaction mixture was allowed to proceed to completion (about 15 min) and was treated identically. Thus the degree of activation and the level of binding with respect to completion were determined by comparing the two halves of the reaction. The results are presented in Figure 1. Whereas the binding curve of C4A shows significant lag behind the activation curve (Fig. 1A), that of C4B follows the activation curve more closely (Fig. 1B). It should be pointed out, however, that the lag of the C4B binding curve, though small, is consistently observed in similar experiments (not shown).

The absolute values for BE were required for the estimation of the reaction rates, thus the values for $[C4_0]$ had to be determined. Additional binding data of [C4b–G($t \rightarrow \infty$)] were obtained at 2.5 mM glycine for C4A and 10 mM glycine for C4B. Two versions of Equation 9 were set up for each of C4A and C4B. By eliminating k_0 and k_2, $[C4_0]$ can be obtained by the expression $[C4_0] = B_1B_2(G_2 - G_1)/(G_2B_1 - G_1B_2)$, where G_1 and G_2 are the two concentrations of glycine and B_1 and B_2 the respective [C4b–G($t \rightarrow \infty$)] values determined at the two glycine concentrations. The BE values for C4A at 0.05 mM glycine is 0.71 ± 0.06 and that for C4B at 2.5 mM glycine 0.25 ± 0.07 (Table 2).

By fitting the activation data to Equation 12, values of k, 0.35 and 0.29 s^{-1}, are obtained for the experiments on C4A and C4B, respectively ($R^2 = 0.99$ in both cases). Using these values and the mean BE values obtained from Table 2, the binding data are fitted to Equation 13 for a single variable of k_0 ($R^2 = 0.99$ for C4A and 0.85 for C4B). k_0(A) was estimated to be 0.068 s^{-1} and k_0(B) 1.08 s^{-1}. It should be noted that the k_0 values are not very sensitive to the the input values of BE in the curve fitting. For example, k_0(A) values of 0.083 and 0.055 s^{-1} were obtained if the BE values of 0.65 and 0.77 (instead of 0.71) were used for C4A, and similarly, k_0(B) values of 1.18 and 0.98 s^{-1} were obtained for BE values of 0.18 and 0.32 (instead of 0.25) for C4B.

Table 2. *Determination of binding efficiency for C4A and C4B*

C4	[Glycine] (mM)	[C4b-G($t \to \infty$)][a]	[$C4_0$][b]	BE[c]
C4A	0.05	2.31 ± 0.15	3.23 ± 0.34	0.71 ± 0.06
	2.50	3.20 ± 0.04		0.99 ± 0.09
C4B	2.50	0.59 ± 0.07	2.34 ± 0.71	0.25 ± 0.07
	10.00	1.35 ± 0.05		0.58 ± 0.17

[a] Binding of glycine after completion of activation and binding reactions; in arbitrary units.
[b] Calculated according to the formula $[C4_0] = B_1B_2(G_2 - G_1)/(G_2B_1 - G_1B_2)$, where G_1 and G_2 are the two concentrations of glycine and B_1 and B_2 the respective [C4b-G($t \to \infty$)] values at G_1 and G_2; in same arbitrary unit as [C4b-G($t \to \infty$)].
[c] Binding efficiency (BE) in absolute units.

The complementary binding data were also obtained in a similar experiment. C4 and C1s were mixed together at time 0. At time t, an aliquot was added to a solution of labeled glycine and both the activation and binding reactions were allowed to proceed to completion. The rates of activation, however, could not be determined in this experiment and were obtained in a parallel experiment as described in Figure 1. Whereas in the experiment shown in Figure 1 we measured the binding that took place before a specific time, in this experiment we measured the binding that took place after the set times. A significant time lag is again demonstrated for the binding curve behind the activation curve for C4A but not for C4B, indicating that the difference observed for the two isotypes is not affected by the presence of glycine (results not shown).

Activation and binding at different [G]

The k_0 value, which is a reciprocal measure of the difference between the activation and binding curves, is very sensitive to small differences. Because the data for the two curves were obtained by different methods, it is possible that there are systemic biases intrinsic to the methods that may contribute to the differences observed. We therefore performed another set of experiments to allow us to examine the data differently.

From Equations 14 and 6, the difference between activation and binding at time t is a measure of [C4b*(t)] and is dependent on [G]. Thus, we measured activation and binding at a fixed time of 3 s, and at [G] = 0.04 mM, 0.2 mM, and 1 mM for C4A and 2.5 mM, 5 mM, and 10 mM for C4B. The amount of C1s added was such that about 50% of C4 would be activated in 3 s, and the experiment was carried out at 22 °C. For each data point, half of the reaction mixture was stopped by addition to TCA and the remaining half allowed to proceed to completion. The results are shown in Table 3. Because the experiment was performed for three different [G], BE can be determined by double reciprocal plot analysis of [C4b-G($t \to \infty$)] vs. [G] (Fig. 2A). The [C4b*($t = 3$)] values for C4A are different for different [G], but those for C4B do not show significant variation. These values were plotted against BE (Fig. 2B). The data for C4A were fitted to a curve according to Equation 15, yielding a k_0 value of

Table 3. *Activation and binding of C4A and C4B at different concentrations of glycine*

[G] (mM)	[C4b-G($t \to \infty$)][a] (arbitrary units)	[C4b-G($t = 3$)] (arbitrary units)	BE[b]	$\frac{[C4b\text{-}G(t=3)]}{[C4b\text{-}G(t\to\infty)]}$	$\frac{[C4b]}{[C4_0]}$[c]	$\frac{[C4b^*]}{[C4_0]}$[d]
C4A						
0.04	0.745	0.107	0.296	0.144	0.574	0.430
	0.814	0.104	0.323	0.128	0.551	0.423
0.20	1.790	0.360	0.711	0.201	0.488	0.287
	1.741	0.487	0.679	0.280	0.631	0.351
1.00	2.291	1.126	0.909	0.491	0.681	0.190
	2.326	0.966	0.921	0.415	0.601	0.186
C4B						
2.50	0.256	0.101	0.158	0.394	0.586	0.192
[e]	0.242	0.114	0.150	0.472	0.627	0.155
5.00	0.410	0.138	0.254	0.338	0.508	0.170
	0.421	0.149	0.261	0.353	0.514	0.161
10.00	0.706	0.264	0.437	0.374	0.583	0.209
	0.698	0.280	0.432	0.401	0.578	0.177

[a] Binding of glycine after completion of activation and binding reaction (15 min).
[b] Binding efficiency obtained from the double reciprocal plot of [C4b-G($t \to \infty$)] vs. [G] in Figure 2A.
[c] Measured by the decrease in stain intensity of the α-chains of C4 on SDS-polyacrylamide gels (see Materials and methods).
[d] Difference between [C4b]/[$C4_0$] and [C4b-G($t = 3$)]/[C4b-G($t \to \infty$)].
[e] Mistake in experiment: time = 4 s for this particular data point; not used for further analysis.

$0.10\ s^{-1}$ ($R^2 = 0.96$). The curve for C4B is not shown because the best fit gave an R^2 value of 0.12.

Discussion

Using site-directed mutagenesis, we have obtained a number of modified C4 molecules and studied their binding to glycine and glycerol. They were found to be either C4A-like with high glycine-binding activity and low-glycerol binding activity, or C4B-like with intermediate binding activities to both substrates (see Table 1). In addition, this difference in binding may be correlated to a single structural difference among these proteins: the C4B-like proteins have a His residue at position 1106, and the C4A-like proteins have a residue other than His at the same position. Because C4B binds glycerol more efficiently than C4A, we postulated that the His may play a catalytic role in the reaction of the thioester with hydroxyl groups. If this postulate is extended to include the reaction of the thioester with water, the apparently more efficient binding of C4A than C4B with amino groups can also be explained. Furthermore, the difference in hydrolysis rate of the thioester suggests that the binding reactions of C4A and C4B should follow different kinetics.

A time course experiment was performed to measure both C1s cleavage of C4A and C4B and their binding to glycine. The binding curve for C4B is very similar to the activation curve, whereas that of C4A lags behind the activation curve by a significant margin (Fig. 1). The hydrolysis rates (k_0) for the thioester were estimated to be $0.068\ s^{-1}$ for C4A ($t_{1/2} = 10.4$ s) and $1.08\ s^{-1}$ for C4B ($t_{1/2} = 0.64$ s).

The activation and binding curves of C4B (Fig. 1B) are very close and the estimated value of k_0(B), which is a reciprocal measure of the difference between the two curves, is therefore very approximate. We therefore determined the difference between activation and binding by varying the concentration of glycine instead of time. Whereas the $[C4b^*]/[C4_0]$ values for C4A were shown to vary with $[G]$, variation of $[C4b^*]/[C4_0]$ vs. [G] cannot be demonstrated for C4B (Table 3; Fig. 2). This result suggests that the difference between activation and binding in C4B, though consistently observed in repeated experiments, may be due to some systematic biases intrinsic to the experimental methods. If this were the case, the k_0(B) value obtained from the data in Figure 1B may be an underestimate. These biases, however, must also be present in the data for C4A, but they have a lesser effect on the estimate of k_0(A) since the difference between activation and binding is significantly larger. The data in Table 3 were analyzed according to Equation 15 and shown in Figure 2B. Curve fitting to the C4A data gave a k_0(A) value of 0.10 s^{-1}, which is similar to the value of $0.068\ s^{-1}$ from Figure 1A. The data for C4B cannot be fitted to Equation 15 (R^2 of 0.12 for the best fit).

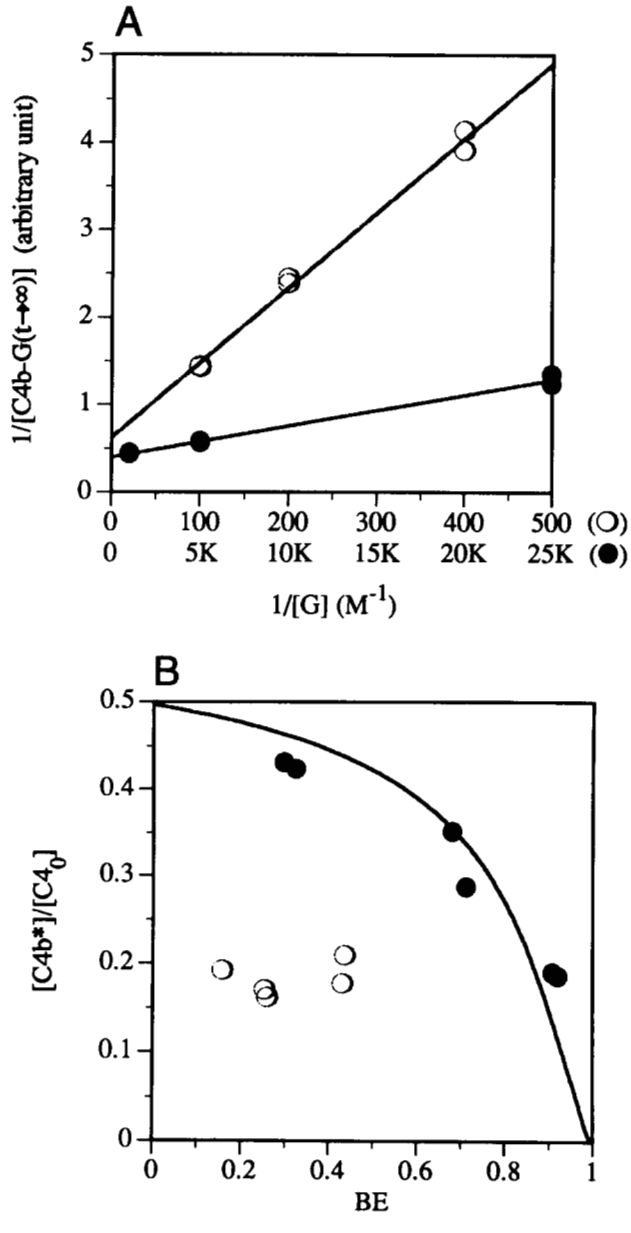

Fig. 2. Activation and glycine binding at different concentrations of glycine. **A:** The double reciprocal plot of $[C4b\text{-}G(t \rightarrow \infty)]$ vs. [G] for C4A (●) and C4B (○). Since $1/[C4b\text{-}G(t \rightarrow \infty)]$ was in arbitrary units, the straight lines drawn through the data points, according to Equation 11, intersect the y-axis to give the values for $1/[C4_0]$ in the same unit. BE values for C4A and C4B were calculated accordingly. ($K = 1{,}000$ in the x-axis label for C4A.) **B:** $[C4b^*]/[C4_0]$ values for C4A (●) and C4B (○) obtained at $t = 3$ s at different concentrations of glycine are plotted vs. BE (data from Table 3). Data for C4A were fitted to Equation 15, giving a k_0(A) value of $0.10\ s^{-1}$. The curve fitted to the data for C4B is not shown because the best fit gave an R^2 value of 0.12.

Although we were unable to determine the hydrolysis rates for C4B with reasonable accuracy, we can conclude that the hydrolysis rate of C4A is significantly slower than that of C4B and that this difference is accountable, at least in part, for the different binding activities of C4A

and C4B. If the difference in hydrolysis rate is the only difference between C4A and C4B, the k_0(B)/k_0(A) value should be on the order of 100, similar to the ratio of the k_2/k_0 values for glycine binding for C4A to C4B (see Table 1). Our best estimates, based on the results presented in Figure 1, give the values of 0.068 s^{-1} for k_0(A) and 1.08 s^{-1} for k_0(B), and a k_0(B)/k_0(A) ratio of about 16. Given the inaccuracy of the k_0(B) value and therefore of this ratio, it remains uncertain whether the difference in hydrolysis rates can solely account for their different binding activities of the two C4 isotypes.

The thioester is extremely stable in native C4A and C4B. Over 80% of either protein remained active, i.e., with the thioester remaining intact, after incubation at 37 °C for 77 h (Law et al., 1984). This has been attributed to the internal location of the thioester and its inaccessibility to water and very limited accessibility to small nucleophiles (Pangburn & Müller-Eberhard, 1980). Activation was shown to be accompanied by a conformational change (Isenman et al., 1981) and exposure of the thioester. Indeed, binding can be activated by reagents, such as KBr and low concentrations (~1 M) of guanidine hydrochloride, known to perturb the higher order structure of proteins (Law, 1983a,b). However, exposure of the thioester is not sufficient to account for its reactivity. The thiolactone ring containing the four amino acids -Cys-Gly-Glu-Gln- was synthesized chemically, and its properties under physiological conditions were studied (Khan & Erickson, 1982). The hydrolysis rate of the thioester in the synthetic peptide was found to be 1×10^{-5} s^{-1} (Erickson & Khan, 1983), which is about 600 times slower than the hydrolysis rate of the thioester in the C4b* form of C4A and about 10,000 times slower than that of C4B. Thus, other residues in the protein must also affect the hydrolysis of the thioester and other binding activities it mediates. Additional support for this hypothesis is found in the study of the hybrid proteins C4(C3) and C4(α_2M). They have binding properties like C4B and C4A, respectively, but not C3 and α_2M (see Table 1). Hybrids containing more extensive regions of C3 on a C4B background are being constructed. Studies of the binding properties of these hybrids may lead to the identification of additional residues involved in the covalent binding reaction unique to the thioester proteins.

Materials and methods

Reagents

Restriction enzymes and other nucleic acid modification enzymes were obtained from Amersham, Boehringer Mannheim, Bethesda Research Laboratories, New England BioLabs, or Pharmacia. Radioactive compounds, including ^{32}P- and ^{35}S-labeled nucleotides, ^{3}H-labeled glycine, glycerol, and methylamine, were purchased from Amersham. The pBluescript plasmid (KS−) was purchased from Stratagene. The pEE6.HCMV.GS expression plasmid was a gift from Celltech. Standard techniques for molecular biology can be found in Sambrook et al. (1989).

Construction of modified C4 cDNA

The C4 cDNA used for expression was a composite from three different clones. The "full-length" clone, pAT-A, contains a rearranged 5′ end and two deletions, of 7 and 9 bp, respectively, in regions 3′ to an internal *Bam*H-I site (Belt et al., 1984, 1985). Hence, the region 5′ to a unique *Kpn*-I site was replaced by one from the partial cDNA clone pAT-F (Belt et al., 1984). Because all the allele-specific nucleotides are located 3′ to the *Bam*H-I site (Yu et al., 1986), the resultant composite cDNA codes for the allotype of pAT-F, i.e., B2 (Belt et al., 1984). In addition, *Xba*-I sites were introduced at both ends to facilitate subsequent cloning of the cDNA into the expression vector. Details of the construction of the cDNA can be found in Anderson (1990).

C4 cDNA clones for expression were constructed by the following scheme. The cDNA of C4B was cloned into pBluescript at the *Xba*-I site. The plasmid was cut with the restriction enzymes *Spe*-I and *Cla*-I, end-filled, and religated. This manipulation resulted in the removal of a segment in the polylinker region inclusive of a *Bam*H-I site. The resultant plasmid was digested with *Bam*H-I and *Sph*-I resulting in the release of a 220-bp fragment from the coding region of C4 containing the isotype-determining region. This fragment was cloned into the corresponding sites of M13mp18. Modification was carried out using the site-directed mutagenesis kit (Amersham) with oligonucleotides prepared on an ABI DNA synthesizer (model 381A). Modified clones were confirmed by sequence analysis using the reagents from the Sequenase sequencing kit, version 2.0 (Cambridge BioScience), before the *Bam*H-I/*Sph*-I fragment was excised, purified, and reintroduced into the major pBluescript/C4 fragment. The C4 cDNA was separated from pBluescript by digestion with *Xba*-I and cloned into the *Xba*-I site of the pEE6.HCMV.GS plasmid (Cockett et al., 1990; Davis et al., 1990).

The conversion of the cDNA of C4B to other C4 variants were carried out in different stages. C4(C3), C4(mouse), and C4B(H → A) were constructed from the C4B clone using the oligonucleotides J89/17, J89/18, and J89/20, respectively. C4A was modified from the C4(mouse) clone using the oligonucleotide J89/22, and C4A(D → A) and C4(α_2M) from C4A using the oligonucleotides J89/19 and K89/07, respectively. The sequences of the oligonucleotides used are as follows:

J89/17	5′-GTATCACTGGAGCGTCGTCCTGGAACGA-3′
J89/18	5′-TGTATCACTGGACAGGGGTCCTGGAACG-3′
J89/19	5′-GCATGCTCCTAGCTAACACTGGAC-3′
J89/20	5′-GCATGCTCCTAGCTATCACTGGAG-3′
J89/22	5′-GCATGCTCCTATCTAACACTGGACAGGG-3′
K89/07	5′-GCATGCTCCTATTTAACAATGAACCGGAGTCCTGGAACGAG-3′

Transfection and expression

C4-pEE6.HCMV.GS plasmids were prepared in CsCl gradients and stored at 4 °C. Transfection and selection of C4 expressing cell lines were performed as described (Cockett et al., 1990; Davis et al., 1990). Briefly, 15 μg of expression constructs were transfected into 1×10^6 Chinese hamster ovary cells (line K1) as precipitates in calcium phosphate. After 24 h, the cells were trypsinized and split into 96 wells in a microtiter plate and selected for growth over the glutamine synthetase inhibitor methionine sulfoximine (Bebbington & Hentschel, 1987). After 10–14 days the supernatants were assayed for C4 hemolytic activity (Law et al., 1984a), and cell lines with high level C4 expression were obtained by standard dilution techniques. C4 producing cell lines were cultured in three 800-mL tissue culture flasks each containing ~80 mL of medium. When the cells reached confluence, the medium was replaced with fresh medium containing 2 mM sodium butyrate (Davis et al., 1990). After 2 days, the supernatants were harvested for the purification of C4.

Purification of C4

Tissue culture supernatant (250 mL) was centrifuged at $10{,}000 \times g$ for 10 min to remove debris. The supernatant was loaded onto a column (16 mm diameter × 25 mm) of Q-Sepharose Fast Flow (Pharmacia) equilibrated with 20 mM Tris, 50 mM ϵ-aminocaproic acid, 5 mM EDTA, 0.02% sodium azide, 100 mM NaCl, pH 7.4, at a flow rate of 4 mL/min. The column was washed with the equilibrating buffer until the absorbance at 280 nm approached zero and the bound protein was eluted with a 30-mL linear gradient to the same buffer with 500 mM NaCl. The major peak eluting between 200 mM and 500 mM NaCl was pooled and loaded onto a column (10 mm diameter × 20 mm) of CNBr-Sepharose (Pharmacia) bearing approximately 2 mg/mL of monoclonal anti-C4 LOO3 (Dodds et al., 1985) at 0.25 mL/min. The column was washed with 50 mM Tris, 50 mM sodium phosphate, 12.5 mM sodium tetraborate, 2.5 mM NaCl, 0.02% sodium azide, pH 7.0, until the absorbance at 280 nm approached zero. C4 was eluted with a 10-mL gradient to the same buffer adjusted to pH 11.5 with NaOH. A single peak was eluted, which was pooled and dialyzed into 10 mM sodium phosphate, 125 mM NaCl, 1 mM EDTA, pH 7.5, for use in binding experiments. Plasma C4A and C4B were prepared from donors having one or the other C4 isotypes as previously described (Dodds et al., 1985).

Binding of glycine and glycerol to C4

The binding of [^{3}H]glycine, [^{3}H]glycerol, and [^{3}H]methylamine to plasma and recombinant C4 and calculation of the BE were performed as described previously (Dodds & Law, 1990).

For the time course experiment, 45 μL of plasma C4 (containing 18 μg of C4) and 10 μL of [^{3}H]glycine (200 mCi/mmol, 25 mM for C4B and 0.5 mM for C4A) in 10 mM sodium phosphate, 125 mM NaCl, 1 mM EDTA, pH 7.5, were mixed and equilibrated at room temperature (~22 °C). Forty-five microliters of activated human C1s (containing 1 μg of C1s), prepared according to Gigli et al. (1976), in the same buffer and equilibrated to room temperature, were added. The reactants were mixed by immediately redrawing and expelling the mixture from the pipettor. At time t (2–12 s from the moment of addition of C1s), 45 μL of the mixture were removed using the same pipette tip, added to 20 μL of 100% (w/v) TCA, and placed on ice. The remaining reactants were allowed to continue to completion (15 min) when a further 45 μL was withdrawn and added to TCA. The TCA-precipitated protein was washed twice with acetone, solubilized in SDS–urea, reduced, run on an SDS–polyacrylamide gel, and stained with Coomassie blue. The intensity of stain in the α-, α′-, and β-chains was determined using a Molecular Dynamics densitometer. The α′-chains were cut from the gel, and the amount of [^{3}H]glycine covalently bound was determined after treatment of the gel slice with NCS tissue solubilizer (Amersham) following the manufacturer's instructions. The β-chains were also excised and counted as a measure of the nonspecific binding of the [^{3}H]glycine to the protein.

C4 preparations usually contain small amounts (<5%) of inactivated C4 (i.e., C4 with a hydrolyzed thioester) and C4b, the degree of activation is calculated as follows. Total amount of activated C4, in arbitrary units, was determined by the difference in stain intensity of the α chain ($I\alpha$), normalized with the stain intensity of the β-chain ($I\beta$), between C4 samples not treated with C1s and C4 treated with C1s to completion, i.e.,

$$\text{total C4} = I\alpha/I\beta \text{ (no C1s)} - I\alpha/I\beta \text{ (+C1s to completion)}.$$

The subtraction of the $I\alpha/I\beta$ value from the completed activated sample is necessary since C4 with hydrolyzed thioester is not cleaved by C1s. The amount of activated C4 at time t is therefore

$$\text{activated C4} = I\alpha/I\beta \text{ (no C1s)} - I\alpha/I\beta \text{ (+C1s at time } t).$$

The ratio of C4 activated at time t is a simple ratio of these two values. Radioactivity associated with the α′-chains are also normalized to the β-chain intensity in gel scanning. Normalization in evaluation of both activation and binding is necessary because recovery of proteins from precipitation may not be 100%. It should be noted that there are alternative ways of calculating C4 activa-

tion, such as by measuring the increase in the intensity of the α'-chain normalized to the β-chain, and that these manipulations do yield slightly, but not significantly, different results. We chose to use the intensity of the α-chains, and not both α- and α'-chains, because the two chains stain slightly differently. In addition, we chose the α-chain in preference to the α'-chain in these experiments since the α-chain can be presumed to be homogeneous with a glutamate residue at the thioester site, whereas the α'-chain may have a glycine covalently bound at the same position.

For the complementary experiment, 45 μL of C1s was added to 45 μL of C4 and mixed as described above. At time t, 45 μL of the mixture were withdrawn and added to 5 μL of [^{3}H]glycine, 0.5 mM for C4A and 25 mM for C4B, giving a final glycine concentration of 0.05 mM and 2.5 mM, respectively, and incubated for 15 min. TCA was added and the precipitated protein treated and analyzed as described above. Note that it was not possible to determine the degree of activation at the time of addition of the [^{3}H]glycine in these experiments.

Data processing

Curve fitting was done using the recursive iteration program in DeltaGraph Professional (DeltaPoint, Inc.). The activation data were fitted to Equation 12 for the single parameter k. The binding data were fitted to Equation 13 for the single parameter k_0. The term k_2[G] was converted to an expression containing k_0 and BE according to Equation 10. The values for k were obtained from the activation data and those for BE from Figure 2. In all cases, the data points were weighted by their respective σ^2. Data points in Figure 2 were entered individually and no weight was used in curve fitting.

Acknowledgment

A.S. was a recipient of a Soros Scholarship.

References

Anderson, M.J. (1990). Molecular genetics of the human complement protein C4. D.Phil. Thesis, University of Oxford.

Bebbington, C.R. & Hentschel, C.C.G. (1987). The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells. In *DNA Cloning: A Practical Approach*, Vol. III (Glover, D.M., Ed.), pp. 163–188. IRL Press, Oxford, UK.

Belt, K.T., Carroll, M.C., & Porter, R.R. (1984). The structural basis of the multiple forms of human complement component C4. *Cell 36*, 907–914.

Belt, K.T., Yu, C.Y., Carroll, M.C., & Porter, R.R. (1985). Polymorphism of human complement component C4. *Immunogenetics 21*, 173–180.

Carroll, M.C., Fathallah, D.M., Bergamaschini, L., Alicot, E.M., & Isenman, D.E. (1990). Substitution of a single amino acid (aspartic acid for histidine) converts the functional activity of human complement C4B to C4A. *Proc. Natl. Acad. Sci. USA 87*, 6868–6872.

Cockett, M.I., Bebbington, C.R., & Yarranton, G.T. (1990). High level expression of tissue inhibitor of metalloproteinases in Chinese hamster ovary cells using glutamine synthetase gene amplification. *Bio/Technology 8*, 662–667.

Davis, S.J., Ward, H.A., Puklavec, M.J., Willis, A.C., Williams, A.F., & Barclay, A.N. (1990). High level expression in Chinese hamster ovary cells of soluble forms of CD4 T lymphocyte glycoprotein including glycosylation variants. *J. Biol. Chem. 265*, 10410–10418.

de Bruijn, M.H.L. & Fey, G.H. (1985). Human complement component C3: cDNA coding sequence and derived primary structure. *Proc. Natl. Acad. Sci. USA 82*, 708–712.

Dodds, A.W. & Law, S.K.A. (1988). Structural basis of the binding specificity of the thioester-containing proteins, C4, C3 and alpha-2-macroglobulin. *Complement 5*, 89–97.

Dodds, A.W. & Law, S.K.A. (1990). The complement component C4 of mammals. *Biochem. J. 265*, 495–502.

Dodds, A.W., Law, S.K.A., & Porter, R.R. (1985). The origin of the very variable haemolytic activities of the common human complement component C4 allotypes including C4-A6. *EMBO J. 9*, 2239–2244.

Dodds, A.W., Law, S.K.A., & Porter, R.R. (1986). The purification and properties of some less common allotypes of the fourth component of human complement. *Immunogenetics 24*, 279–285.

Erickson, B.W. & Khan, S.A. (1983). Synthetic lactam and thiolactone models of protein metastable binding sites. *Ann. N.Y. Acad. Sci. 421*, 167–177.

Gigli, I., Porter, R.R., & Sim, R.B. (1976). The unactivated form of the first component of human complement, C1. *Biochem. J. 159*, 541–548.

Hall, R.E. & Colten, H.R. (1977). Cell-free synthesis of the fourth component of guinea pig complement (C4): Identification of a precursor of serum C4 (pro-C4). *Proc. Natl. Acad. Sci. USA 74*, 1707–1710.

Harrison, R.A., Thomas, M.L., & Tack, B.F. (1981). Sequence determination of the thiolester site of the fourth component of human complement. *Proc. Natl. Acad. Sci. USA 78*, 7388–7392.

Howard, J.B. (1983). Reactive centres in α_2-macroglobulin. *Ann. N.Y. Acad. Sci. 421*, 69–80.

Isenman, D.E., Kells, D.I.C., Cooper, N.R., Müller-Eberhard, H.J., & Pangburn, M.K. (1981). Nucleophilic modification of human complement protein C3: Correlation of conformational changes with acquisition of C3b-like functional properties. *Biochemistry 20*, 4458–4467.

Isenman, D.E. & Young, J.R. (1984). The molecular basis for the difference in immune hemolysis activity of Chido and Rodgers isotypes of human complement component C4. *J. Immunol. 132*, 3019–3027.

Khan, S.A. & Erickson, B.W. (1982). An equilibrium model of the metastable binding sites of α_2-macroglobulin and complement proteins C3 and C4. *J. Biol. Chem. 257*, 11864–11867.

Law, S.K.A. (1983a). The covalent binding reaction of C3 and C4. *Ann. N.Y. Acad. Sci. 421*, 246–258.

Law, S.K.A. (1983b). Non-enzymic activation of the covalent binding reaction of the complement protein C3. *Biochem. J. 211*, 381–389.

Law, S.K.A., Dodds, A.W., & Porter, R.R. (1984a). A comparison of the properties of two classes, C4A and C4B, of the human complement component C4. *EMBO J. 3*, 1819–1823.

Law, S.K.A. & Levine, R.P. (1977). Interaction between the third complement protein and cell surface macromolecules. *Proc. Natl. Acad. Sci. USA 74*, 2701–2705.

Law, S.K.A., Lichtenberg, N.A., & Levine, R.P. (1980). Covalent binding and hemolytic activity of complement proteins. *Proc. Natl. Acad. Sci. USA 77*, 7194–7198.

Law, S.K.A., Minich, T.M., & Levine, R.P. (1981). Binding reaction between the third human complement protein and small molecules. *Biochemistry 20*, 7457–7463.

Law, S.K.A., Minich, T.M., & Levine, R.P. (1984b). Covalent binding efficiency of the third and fourth complement proteins in relation to pH, nucleophilicity and availability of hydroxyl groups. *Biochemistry 23*, 3267–3272.

Law, S.K.A. & Reid, K.B.M. (1988). Complement. (*In Focus* Series [Male, D., Ed.].) IRL Press, Oxford, UK.

Nonaka, M., Nakayama, K., Yuel, Y.D., & Takahashi, M. (1985). Complete nucleotide and derived amino acid sequences of the fourth component of mouse complement (C4): Evolutionary aspects. *J. Biol. Chem. 260*, 10936–10943.

Pangburn, M.K. & Müller-Eberhard, H.J. (1980). Relation of a puta-

tive thioester bond in C3 to activation of the alternative pathway and the binding of C3b to biological targets of complement. *J. Exp. Med. 152*, 1102–1114.
Ren, X.D., Dodds, A.W., & Law, S.K.A. (1993). The thioester and isotypic sites of complement component C4 in sheep and cattle. *Immunogenetics 37*, 120–128.
Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd Ed. Cold Spring Harbor Press, Cold Spring Harbor, New York.
Sepich, S.S., Noonan, D.J., & Ogata, R.T. (1985). Complete cDNA sequence of the fourth component of murine complement. *Proc. Natl. Acad. Sci. USA 82*, 5895–5899.
Sottrup-Jensen, L. (1987). α2-Macroglobulin and related thiol ester plasma proteins. In *The Plasma Proteins*, Vol. 5 (Putnam, F.W., Ed.), pp. 191–291. Academic Press, Orlando, Florida.
Tack, B.F., Harrison, R.A., Janatova, J., Thomas, M.L., & Prahl, J.W. (1980). Evidence for presence of an internal thiolester bond in third component of human complement. *Proc. Natl. Acad. Sci. USA 77*, 5764–5768.
Yu, C.Y., Belt, K.T., Giles, C.M., Campbell, R.D., & Porter, R.R. (1986). Structural basis of the polymorphism of human complement components C4A and C4B: Gene size, reactivity and antigenicity. *EMBO J. 5*, 2873–2881.