Human C4b-binding Protein, Structural Basis for Interaction with Streptococcal M Protein, a Major Bacterial Virulence Factor

Received for publication, October 25, 2005, and in revised form, November 30, 2005. Published, JBC Papers in Press, December 5, 2005. DOI 10.1074/jbc.M511563200

Huw T. Jenkins, Linda Mark, Graeme Ball, Jenny Persson, Gunnar Lindahl, Dusan Uhrin, Anna M. Blom, and Paul N. Barlow

From the Edinburgh Biological NMR Unit, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, Scotland, United Kingdom, the Department of Laboratory Medicine, Section of Clinical Chemistry, University Hospital Malmö, Lund University, S-20502 Malmö, Sweden, and the Department of Laboratory Medicine, Division of Medical Microbiology, Lund University, Sölvegatan 23, SE-22362 Lund, Sweden

Human C4b-binding protein (C4BP) protects host tissue, and those pathogens able to hijack this plasma glycoprotein, from complement-mediated destruction. We now show that the first two complement control protein (CCP) modules of the C4BP α-chain, plus the four residues connecting them, are necessary and sufficient for binding a bacterial virulence factor, the Streptococcus pyogenes M4 (Arp4) protein. Structure determination by NMR reveals two tightly coupled CCP modules in an elongated arrangement within this region of C4BP. Chemical shift perturbation studies demonstrate that the N-terminal, hypervariable region of M4 binds to a site including strand 1 of CCP module 2. This interaction is accompanied by an intermodular reorientation within C4BP. We thus provide a detailed picture of an interaction whereby a pathogen evades complement.

Bacteria that enter human blood or tissues will be opsonized by phagocytes unless they have a means of overcoming attack by the complement system. An extensively studied example of such a virulence mechanism is the ability of many Streptococcus pyogenes M proteins to bind the key human complement regulator C4b-binding protein (C4BP), an interaction that endows the bacteria with resistance to phagocytosis (1). C4BP is a polymeric, soluble, glycoprotein (~200 mg/liter in plasma) (2). The M proteins, classical bacterial virulence factors, first recognized over 75 years ago (3), are fibrillar surface structures exhibiting antigenic variation within a 50–100-amino acid residue hypervariable N terminus (4, 5). Expression of M protein is important for the ability of S. pyogenes to cause the diseases associated with this pathogen: acute pharyngitis and impetigo (6). Many other pathogens share the ability to sequester C4BP (7–12). Thus escape from complement attack by hijacking of C4BP is emerging as a widespread contributor to immune evasion strategies and is a therapeutic target.

The complement system, a vital molecular component of innate immunity, consists of plasma and cell-surface proteins that conspire to rid the body of infectious particles (13, 14). Because complement is potentially harmful to host tissue it is tightly regulated. Like other members of the regulators of complement activation (RCA) protein family, the plasma protein C4BP acts upon the bimolecular C3 convertases that are the main drivers of the complement cascade. The C3 convertases cleave and thereby activate the third component of complement, C3, leading to deposition of C3b on the surface of an unprotected particle such as an invading microorganism. This process, known as opsonization, marks the particle as a target for phagocytosis. C3b also nucleates assembly of further convertase complexes and progression of the cascade toward formation of the membrane attack complex. C4BP regulates the C3 convertase of the classical pathway, a complex of C4b and C2a, by acting as a cofactor for the proteolysis of C4b (15, 16) and accelerating the decay of C4b-C2a complexes deposited on host cells (17). It binds self-surface glycosaminoglycans where it is maximally effective in preventing complement activation. C4BP additionally acts as a regulator of the alternative pathway (18), participates in apoptosis (19), and binds CD40 (20) and DNA (21).

The predominant isoform of C4BP has seven α-chains each containing eight complement control protein (CCP) modules (or short consensus repeats) and a single β-chain with three CCP modules (22, 23). Transmission electron microscopy indicated that up to six molecules of C4b bind to one molecule of C4BP at the distal (N-terminal) ends of the seven α-chains (2). Mutants lacking CCP1, CCP2 (in particular), or CCP3 in their α-chains had severely impaired ability to bind to immobilized C4b (24), a result confirmed by Fukui et al. (25). Four residues, Arg99, Lys103, Arg104, and His105, predicted to reside within a positively charged cluster at the CCP1/2 interface, are crucial for C4b binding (26).

The C4BP-binding M proteins of S. pyogenes interact with a C4BP region close to the C4b-binding site, but sites for binding natural and pathogen-borne proteins appear non-identical (27). The bacterial protein adheres more tightly than C4b and is able to displace C4b from its binding site. When bound by a streptococcal M protein, C4BP is still capable of regulating complement and thus protecting the organism upon which it is sequestered from complement-mediated destruction.

Unlike in the case of all other RCA proteins there has, until now, been no experimental three-dimensional structural information available for C4BP. Hence the atomic resolution basis of its interactions with patho-
gen-borne and natural ligands has remained unknown. Opportunities for rational design of therapeutic interventions, e.g. to counter streptococcal infections or to inhibit complement-mediated inflammation, have consequently been limited. Here we demonstrate which domains of C4BP are involved in binding to an M protein of S. pyogenes and present the solution structure of the relevant fragment of C4BP. This has allowed us to delineate in detail the binding sites on the surface of C4BP for the M protein.

EXPERIMENTAL PROCEDURES

Expression, Purification, and Sample Preparation—Human C4BP CCP1–2 (C4BP12) cDNA was amplified by PCR yielding the protein sequence: MNGPPPTLFSAAPMDITLTETRFTGTKTLKYTCLPGY-VR5HSTQLTTCNSDEGWVYNTTFCYRCRHPRNLGQRVQIEKT-DLSFGQSIEFSCGFLIYSTT5SRCVEQDRGYW5HPLPQCELIE-HH1111111. This was cloned in the Bluescript vector using the PCR-Script cloning kit (Stratagene, La Jolla, CA) and transferred into pET16 vector (Novagen, Merck Biosciences, Nottingham, UK) using Ndel and Xhol. The DNA was transfected into BL21 (DE3) CodonPlus-RP Escherichia coli, which were cultured in Luria-Bertani broth containing kanamycin and chloramphenicol (37 °C). After cooling (30 °C), expression was induced with isopropyl-β-D-thiogalactopyranoside and bacteria grown for 5 h. The bacteria were then centrifuged and resuspended in cold phosphate-buffered saline, lysed, and sonicated. The sonicate was centrifuged and the pellet resuspended in 6 M guanidine HCl, 20 mM Tris-HCl, pH 8.0, 10 mM reduced glutathione. The crude material was derived from recombinant C4BP molecules in which one of the eight CCP modules of the α-chain was deleted (A) or in which two Ala residues were inserted into the linker between CCP1 and CCP2 from human, bovine, murine, and rat C4BP α-chain with human DAF and CR1. The hypervariable loops of C4BP modules 1 and 2 are underlined. D, isothermal titration calorimetry profile for the interaction of M4-N with C4BP12. Top, heat differences obtained for 20 injections. Bottom, integrated curve with experimental points (■) and the best fit (—). Data were fitted using a one-site model resulting in the following: stoichiometry, n = 0.787 ± 0.015; Kd = 0.5 ± 0.07 μM; ΔH = −9450 ± 243 cal mol⁻¹; ΔS = −2.96 cal mol⁻¹ K⁻¹.

FIGURE 1. M4 binding to C4BP. A and B, binding of CCP module deletion mutants (A) and double-alanine insertion mutants (B) of C4BP to M4. Microtiter plates coated with M4 were exposed to recombinant C4BP molecules in which one of the eight CCP modules of the α-chain was deleted (A) or in which two Ala residues were inserted into the linker between CCPs 1 and 2, 2 and 3, or 3 and 4 (B). Binding was expressed as a percentage of the maximum binding of wild type observed in each experiment. C, multiple sequence alignment of CCP1 and CCP2 from human, bovine, murine, and rat C4BP α-chain with human DAF and CR1. The hypervariable loops of C4BP modules 1 and 2 are underlined. D, isothermal titration calorimetry profile for the interaction of M4-N with C4BP12. Top, heat differences obtained for 20 injections. Bottom, integrated curve with experimental points (■) and the best fit (—).
phage (Profos, Regensburg, Germany) with a 2H splitting of 0.9 Hz. Sample conditions were: 0.70 mM C4BP12, 7.3 mg/ml Pf1 phage, 20 mM NaOAc, pH 5.5, 90 mM NaCl. 1D NH, 1D CH, and 1D CN couplings were measured for aligned and non-aligned samples.

Structure Calculation—NOE intensities were converted into distance categories of 0–2.7 Å, 0–3.3 Å, 0–5.0 Å, and 0–6.0 Å. The distance restraints, along with four disulfide bonds (defined by homology) and torsion angle restraints calculated from chemical shift index data, were used as input for the structure calculation performed with CNS-solve (30). As calculations progressed, iterative “filtering” of ambiguously assigned NOEs removed duplicates and assignments contributing <1% to the total NOE. The RDC restraints were used only in the final refinement by including the TENSOR energy term within CNS-solve using a harmonic potential. Residual dipolar couplings for some residues were excluded on the basis of heteronuclear NOE data. A total of 100 structures were calculated.

Titration with M4-N—The peptide dimer, M4-N, was expressed in E. coli, as described. Briefly, the part of the M4 gene corresponding to the hypervariable region (residues 1–45) was amplified from plasmid pARP401 (31), cloned into expression vector pET3a, and expressed in E. coli. The codon for a cysteine was added to the 3′ end of the construct to facilitate dimerization of the expressed protein via a disulfide bridge. Dimerization was achieved by oxidation of bacterial lysates with 20 mM sodium carbonate, pH 9.6. Wells were washed with 50 mM Tris-HCl, NaOAc, pH 5.5, 90 mM NaCl. 1DNH, 1DC, and 1DCO were subsequently recorded at M4-N concentrations of 100, 50, 25, and 10 μM (dimer concentrations) and compared with a reference spectrum for a sample containing no M4-N.

M4 Binding Assay—Microwell plates were incubated (overnight, 4°C) with 50 μl of solution containing 10 μg/ml full-length M4 protein, expressed and purified as described in Stenberg et al. (33) in 75 mM sodium carbonate, pH 9.6. Wells were washed with 50 mM Tris-HCl, 0.15 M NaCl, 0.1% Tween 20, pH 7.5 (washing buffer) and then incubated (room temperature) with 200 μl of quench solution (washing buffer plus 3% fish gelatin). After re-washing, increasing concentrations of recombinant C4BP mutants (see Ref. 24) diluted in 50 mM Tris-HCl, 150 mM NaCl, pH 8.0, supplemented with 0.1% bovine serum albumin and 0.1% Tween 20 were added and the plates incubated (4 h, room temperature). Plates were then washed and incubated with biotinylated monoclonal antibody 67 (against CCP4) or monoclonal antibody 104 (against CCP1) diluted in quench solution. After 1 h of incubation, plates were washed and incubated (1 h) with streptavidin-conjugated horseradish peroxidase (Dakopatts, Glostrup, Denmark), washed again, and developed. The results are shown as a percentage of the maximal binding of recombinant wild type C4BP obtained in each set of triplicate experiments.

Isothermal Titration Calorimetry—Experiments were performed (25°C) on a VP-isothermal titration calorimeter (ITC) microcalorimeter (MicroCal Inc., Northampton, MA). The cell contained 0.01 mM C4BP12 (1.4 mM) and the syringe contained 0.15 mM M4-N. Both solutions were in 10 mM sodium phosphate buffer, pH 6.0. For titrations, one preliminary injection of 1 μl was made, followed by 19 injections of 10 μl with an injection speed of 0.5 μl/s. The stirring speed was 310 rpm; delay time between injections = 3 min. A blank titration, i.e. injecting M4-N into buffer, was used to correct the M4-N:C4BP12 titration.

### RESULTS

CCPs 1 and 2 of C4BP α-Chain Are Necessary and Sufficient for Binding the Hypervariable Domain of M4—To define the C4BP region involved in M protein binding, deletion mutants of the α-chain were tested for binding to the M protein, M4 (Fig. 1A). None of CCP modules 3–8 were required for M4 binding, but deletion of CCP2 led to significantly reduced affinity. Deletion of CCP1 destroyed binding almost completely. In further investigations, the linking sequences between CCPs 1, 2, 3, and 4 of the α-chain were engineered by insertions of Ala residues (Fig. 1B). Addition of residues between CCP2 and CCP3, or between CCP3 and CCP4, had no effect on binding, whereas insertions between CCP1 and CCP2 caused a loss of M4 binding. A recombinant protein consisting of residues 1–124 of C4BP α-chain showed binding. His-tagged at the C terminus, was expressed in E. coli and refolded. ITC was used (Fig. 1D) to measure the binding affinity between C4BP12 and a dimerised peptide corresponding to the hypervariable region of M4 (residues 1–45 of M4 protein with a C-terminal Cys added. A dimer (M4-N) was created because previous studies had shown dimerization of M4 strongly enhanced C4BP binding (32). The results are consistent with a 1:1 stoichiometry, and $K_d = 0.5 \mu M$. Taken together, the data summarized in Fig. 1 prove that the N-terminal two CCP modules of C4BP α-chain along with the four residues that link

| Table 1 Structural statistics for the 40 lowest energy structures |
|---------------------------------------------------------------|
| **Unambiguous NOEs:**                                        |
| Intrasidue                                                   | 1639 |
| Sequential                                                  | 899  |
| Short range (2 ≤ i − j ≤ 4)                                  | 240  |
| Long range (i − j > 4)                                      | 763  |
| Intermodular                                                | 7    |
| Intralinker                                                 | 111  |
| From module 1 to linker                                      | 35   |
| From module 2 to linker                                      | 21   |
| **Total unambiguous NOEs**                                   | 3541 |
| **Total ambiguous NOEs**                                    | 895  |
| **Residual dipolar couplings**                               |      |
| $D_{NH}$                                                    | 62   |
| $D_{NC}$                                                    | 65   |
| $D_{CN}$                                                    | 41   |
| $D_{CHi}$                                                   | 24   |
| **r.m.s.d. for ensemble of 40 structures ± S.D.**            |      |
| NOE (Å)                                                     | 0.0423 ± 0.0028 |
| Bond lengths (Å)                                            | 0.0025 ± 0.0001 |
| Bond Angles (°)                                             | 0.5213 ± 0.0177 |
| **Backbone atoms (C*, N, CO): excluding loops (all residues)** |
| Module 1                                                      | 0.528 (0.895) |
| Module 2                                                      | 0.488 (0.597) |
| Both modules                                                 | 0.856 (1.042) |
| **All heavy atoms: excluding loops (all residues)**          |      |
| Module 1                                                      | 0.851 (1.487) |
| Module 2                                                      | 0.922 (1.108) |
| Both modules                                                 | 1.147 (1.507) |
| **Ramachandran assessment (%)**                             |      |
| Most favored                                                 | 52.4 |
| Additionally allowed                                         | 37.5 |
| Generously allowed                                           | 7.1  |
| Disallowed                                                  | 3.1  |
| **Skew, twist, and tilt angles (°)**                         |      |
| Skew                                                         | 204.14 | −6.25 | 31.75 |
| Twist                                                       | 288.89 | 21.70 | 48.74 |
| Tilt                                                        | 249.85 | 3.66  | 38.42 |
| **Standard deviation**                                       | 17.93 | 5.68  | 4.39  |

*Excluding residues 15–21, 35–45 (module 1) and 78–82, 107–111 (module 2)." Residues from CysI to CysIV of each module.
Structure-Function of C4b-binding Protein

The Structure of C4BP12 Has Been Determined—To further characterize the M4-binding site of C4BP, the structure of C4BP12 was solved using NMR. There were no double peaks in the $^{15}$N-$^1$H HSQC spectrum of C4BP12, i.e., there was no indication of multiple conformations. Each of the eight Pro residues is trans as indicated by the difference in chemical shifts, $\Delta\delta^{\beta-\alpha}$ (34), and appropriate, strong NOE cross-peaks. A total of 1250 interresidue restraints were employed in the structure calculation, along with 895 ambiguous restraints. These were supplemented by 192 residual dipolar couplings (RDCs) collected on an aligned sample of C4BP12 prepared in the presence of filamentous phage. An ensemble of 40 structures was selected from 100 calculated, on the basis of lowest energy arising from experimental restraints (Table 1 and Fig. 2, A–C). The quality of the structures, as judged by the Ramachandran plot (Table 1), is acceptable for a small protein domain with disordered loops (see below); 90% of residues within the ensemble occur in the most favored and additionally allowed regions of the plot.

The structures of the individual modules converge well (Table 1, Fig. 2, A and B). Each has a similar, elongated shape (Fig. 2D) in which short $\beta$-strands and other extended segments (that do not satisfy the criteria used to define $\beta$-strands (35)) are aligned with the long axis. Five strands/extended segments wrap around a hydrophobic core that is bounded by the two invariant disulfide bridges. The strands/extended segments run up-down-up-down such that the N and C termini are at opposite ends of the module. The connecting turns and loops also generally lie toward the ends of the module, but the "hypervariable loop," a site of high sequence variation and of insertions or deletions (see Fig. 1C), projects laterally (Fig. 2, A–C), and it lies close to the intermodular interface. In general, $T_1/T_2$ ratios and heteronuclear NOEs are more variable in CCP1 than in CCP2 (Fig. 3, A and B), reflecting a higher level of flexibility in the first module.

The Two N-terminal CCP Modules of C4BP $\alpha$-Chain Form a Semi-rigid Unit—The ensemble of structures converge well when superimposed over both modules (Fig. 2C). The intermodular interface is mainly hydrophobic with contributions from all four linker residues and from Tyr$^{37}$ and Val$^{108}$ of CCP1, along with Phe$^{84}$, Val$^{108}$, and Val$^{113}$ of CCP2. The Tyr$^{37}$ and Arg$^{44}$ side chains line up on one side, while on the other, Val$^{108}$ and Val$^{108}$ bracket Phe$^{84}$ and all three contact Lys$^{63}$, which in turn lies alongside Ile$^{61}$. There is no evidence from the relaxation measurements for fast or slow timescale motion in the backbone of the linker (Fig. 3, A and B). Nonetheless, a superposition based on individual modules is better than a superposition of the module pair (Fig. 2, A–C), reflecting a small range of intermodular angles among the ensemble (values calculated (Table 1) using for reference a vector connecting the principal inertia tensor of each module with the C$^\text{a}$ of its consensus Trp). This could be interpreted in terms of a restricted degree of intermodular movement but could also reflect a limitation of the experimental data. Both the mean tilt and twist angles are relatively small so that the two modules form an elongated structure in which equivalent features, such as the hypervariable loops, lie on the same face of the molecule (Fig. 2D).

Residues Involved in Binding to M4 Mapped by Chemical Shift Perturbation—Significant changes in the $^{15}$N-$^1$H HSQC spectrum of C4BP12 occurred upon addition of M4-N (Fig. 4). At a low ratio (0.5:1) of M4-N to C4BP12, some cross-peaks disappear, while others move a small distance within the spectrum. As this ratio is increased, those peaks that had previously vanished reappear at a new frequency, while those that had moved shift further from their original positions. The
disappearance and subsequent reappearance of some peaks reflects exchange between free and bound forms of C4BP12 with a frequency comparable with the difference between the frequencies of the original peak (free) and the new one (bound). On the other hand, when the difference between the frequencies of the original peak and the new one is smaller, then the exchange rate is faster relative to the frequency difference, giving rise to the peaks that move incrementally during the course of the titration. The line widths of the cross-peaks for the complex are not significantly broader than those of free C4BP12, despite the increased molecular mass of the complex (15 kDa for C4BP12 + 11.3 kDa for M4-N). One explanation is that the complex is less anisotropic than free C4BP12 and tumbles accordingly; this could happen if the bound form of M4-N is relatively compact and binds toward the center of C4BP12, as opposed to at either end.

Most of the significantly perturbed residues (Fig. 5) are in the intermodular linking sequence or in loops or turns near the intermodular interface. From the surface representation (Fig. 5A) it is obvious that these residues do not all lie on one face of the molecule. It is therefore improbable that they are all simultaneously involved in contacting the peptide. More likely the chemical shifts of some of these residues are
perturbed by re-orientation of the two modules upon M4-N binding. Within CCP2, however, two stretches of residues distant from the interface have significantly perturbed chemical shifts. These do form a contiguous surface and together comprise a feasible binding patch. The protein engineering experiments described above establish that residues from the interface-proximal region of module 1, and/or in the linker, additionally contribute directly to binding.

**DISCUSSION**

C4b-binding protein is sequestered by several S. pyogenes strains (36) whereupon the complement regulatory activity of this polymeric protein protects the invading bacteria (1). C4b-binding protein adheres to the streptococcal M proteins, but the structural details of the protein-protein interactions involved have remained obscure. Our module deletion and ITC-based approach showed that the two N-terminal CCP modules of the C4BP α-chain, and a four-residue linker between them, are necessary and sufficient for binding of the S. pyogenes M protein, M4. This observation is in good agreement with previous module deletion experiments (37), using whole bacteria, that localized the S. pyogenes binding site to CCPs 1–3, with CCP2 absolutely required, while deletion of CCP3 had only a marginal effect. The ITC-based measurements proved that a recombinantly expressed fragment of the C4BP α-chain (C4BP12) composed of residues 1–124, and therefore encompassing CCP1 and CCP2, binds tightly to a dimerized form of a peptide derived from the N-terminal region of M4 (M4-N) and thus that the C4BP-M protein interaction is amenable to structural analysis. To circumvent the possibility that crystal-packing forces, or the presence of high concentrations of precipitants, might influence inter-modular angles (38) the structure of C4BP12 was determined in solution.

**Comparison of the Structure of C4BP12 with Other CCP Structures**

While each module has the structural features expected of a CCP module, a pair-wise comparison of CCP1 and CCP2 (using combinatorial extension (39)) yielded a Cα r.m.s.d. = 3.39 Å (over 57 residues). The second module of the C4BP α-chain is highly similar in structure to CCP16 of CR1 (CD35, Protein Data Bank code = 1GKN) and CCP3 of DAF (CD55, Protein Data Bank code = 1H03) (Cα r.m.s.d. of ~1.6 Å in both cases). It is also similar to CCP2 of MCP (CD46, Protein Data Bank code = 1CKL). All of these modules belong to the same sequence cluster (cluster C, Soares et al., (40)) and occupy the second module positions within the C3b/C4b-binding sites of their respective parent proteins (41). CCP1 of C4BP is a more structurally divergent module. Its closest
known structural relatives include CCP15 of CR1 and CCP2 of DAF, both of which are first modules within C3b/C4b-binding sites. Thus both at the level of individual module structures and from a functional perspective there are parallels between module pairs C4BP12, DAF23, and CR1–1516. However, the linker sequence of C4BP12 (IYKR) is replaced by KRKR in CR1–1516 and KKKK in DAF23. Furthermore, while Val100 in CCP1 of C4BP is semiconservatively replaced by a Tyr in CR1, this position is occupied by an Arg in DAF. It is not surprising therefore that the intermodular angles of C4BP12 are at variance with those calculated for the solution structure of CR1–1516 (42) and the crystal and solution structures of DAF23 (Fig. 6). Based on these results, the ability of a protein to bind C3b/C4b does not appear to require a specific intermodular arrangement in the free protein.

A prominent structural feature of C4BP12 is the loop in CCP2 formed by the residues between Val108 and Val113 (sequence: QDRG). The presence of such a loop, four mostly polar residues bounded by two hydrophobic residues that can contribute to the interface with the preceding module, is unique to cluster C CCP modules (40). Since cluster C members form the second modules of C3b/C4b recognition sites in C4BP, MCP, DAF, and both sites 1 and 2 of CR1, some conservation among these modules in the use of specific features for C3b/C4b binding might be expected. Indeed this loop is critical to the C3b/C4b binding activities of CR1 sites 1 and 2 (43, 44). It is therefore an obvious candidate for future mutagenesis.

**Interaction with M Proteins**—Perturbations of C4BP12 chemical shifts that accompany association with M4-N implicate in binding residues toward the middle of the module pair, while the N terminals of CCP1 and the C terminals of CCP2 are not involved. This is consistent with a complex that is globular, rather than linear, as also inferred from the relatively narrow line widths of its NMR signals. The extent and distribution of affected residues are most convincingly explained by a ligand-induced intermodular conformational adjustment. Three side chains exert ring current shifts in the vicinity (Fig. 5C), one from each of CCPs 1 and 2 and one in the linker (Tyr37, Tyr62, and Phe84, respectively). Thus movements between modules and associated conformational adjustments of loops would result in large chemical shift changes.

In earlier studies, K63Q, a substitution that could not be accommodated without perturbing the intermodular interface observed in the structure, had little effect on M4 binding (27). Considered in conjunction with the Ala insertion experiments, this implies that while the four-residue length of the linker is critical for M4 binding, the interface composition, and hence the relative arrangement and flexibility of the two modules in the free protein, is not. This conclusion is consistent with a change of modular orientation upon binding, i.e. the bound conformation being different from the free one.

The mutants R64Q and H67Q each displayed reduced M4 binding; on the other hand R66Q and K79Q increased affinity. In the current study, all four of these mutated residues experience significant perturbations of chemical shift upon addition of M4-N. From the structure it is apparent that none of these mutations involve a residue that participates in the interface, rather Arg64 and His67 form a potential M4-binding patch on the surface of CCP2. Note that bovine, murine, and rat C4BP lack one or both of Arg64 and His67 and are indeed unable to bind M protein (37). Arg66 and Lys70 flank His67, and the gain of affinity resulting from the R66Q and K79Q substitutions could be explained if the native residues impede access to the M4 binding site either because of their positive charge (see below) or by virtue of their steric bulk. The lack of dependence on salt and pH suggest that binding of M4 requires other forces apart from electrostatics (27). In general, the pattern of effects on M4 binding that result from mutagenesis of charged residues is consistent with an interaction in which electrostatics steer the two components toward a productive interaction, but other forces are critical to stabilize the complex eventually formed. Thus while some mutations inhibit binding, others increase it, and some mutations compensate for one another. It is unsurprising that selected substitutions in C4BP improve binding; the interaction is unlikely to be optimal in terms of affinity, since the two partners are under opposing evolutionary influences in this respect. Furthermore, the interaction is not specific to M4; other M proteins with different hypervariable domains also bind to C4BP.

**Interpretation of Mutagenesis Data for C4b and Glycosaminoglycan Binding**—The current structural and chemical shift perturbation data demonstrate that M4 binding involves an intermodular reorientation of C4BP12. This is the first experimental evidence in support of intermodular flexibility being important for the ligand-binding properties of a mammalian RCA protein. In the light of this, it seems likely that such conformational flexibility could also be important for binding to the natural ligands, C4b and glycosaminoglycans. Indeed, critical residues, Arg59, Lys63, Arg64, and His67, implicated by mutagenesis are not clustered together in the structure of the free protein (Fig. 5B) but could be brought into juxtaposition by a twist between the modules creating an interaction that is dominated by electrostatic forces. In the case of DAF23 a similar hypothesis was suggested as an explanation for the pattern of functionally critical mutants within the solution structure (45). Moreover, the aforementioned lack of consistency among intermodular angles within C3b/C4b-binding regions of the free RCAs could
be explained if each undergoes a conformational change upon binding, perhaps converging on a common C3b/C4b-bound orientation that is necessary for cofactor and or/decay-accelerating activity.

The avoidance of phagocytosis necessitates sequestration at the bacterial surface of sufficient, functionally active, C4BP. For its own protection, the host cell may also sequester C4BP, a process that relies on the presence of glycosaminoglycan-binding sites; each α-chain carries a glycosaminoglycan recognition site within its N-terminal pair of CCP modules that overlaps with the C4-binding site. The multivalency of this polymeric protein presumably permits occupancy of some sites by the M protein or glycosaminoglycan, leaving others to mediate decay acceleration and cofactor activity. The extra affinity provided by the hydrophobic component of the binding site enables M proteins to compete for binding with glycosaminoglycans and with the convertase even when the latter is in excess following amplification. Because M proteins exploit some positively charged residues that are important for both glycosaminoglycan binding and complement regulation, it is less likely that C4BP variants with low affinity for M proteins will provide an advantage to the host, as proposed for CR1 and resistance to Plasmodium falciparum-induced rosetting in malaria-exposed populations (46). The conformational change that accompanies binding was unexpected, given that the free form of the protein has a pre-existing defined orientation. On the other hand the viral RCA, VCP, also undergoes a conformational transition upon heparin binding (47). An interesting possibility is that the conformational change brought about by M protein mimics a putative one induced by glycosaminoglycans, which in turn might exert a positive effect on the regulatory activity of C4BP.

Acknowledgements—We thank Prof. Björn Dahlbäck who developed monoclonal antibodies used in this study, Dr. Ingemar André and Prof. Sara Linse for help with cloning and purification of M protein constructs, Dr. Lena Kask for help in experimental work, Dinesh Soares for help with Combinatorial Extension, and Dr. Brian Smith for advice on structure calculations.

REFERENCES

1. Carlsson, F., Berggård, K., Stahlhammar-Carlsmalm, M., and Lindahl, G. (2003) J. Exp. Med. 198, 1057–1068
2. Dahlbäck, B., Smith, C. A., and Müller-Eberhard, H. J. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 3461–3465
3. Lancefield, R. (1928) J. Exp. Med. 47, 91–103
4. Fischetti, V. A. (1989) Clin. Microbiol. Rev. 2, 285–314
5. Navarre, W. W., and Schneewind, O. (1999) Microbiol. Mol. Biol. Rev. 63, 174–229
6. Stevens, D. L., and Kaplan, E. L. (2000) Strepトtosic Infections. Clinical Aspects, Microbiology, and Molecular Pathogenesis, Oxford University Press, New York
7. Berggård, K., Lindahl, G., Dahlbäck, B., and Blom, A. M. (2001) Eur. J. Immunol. 31, 2771–2780
8. Jarva, H., Ram, S., Vogel, U., Blom, A. M., and Meri, S. (2003) J. Immunol. 170, 5328–5336
9. Frithz, E., Heden, L.-O., and Lindahl, G. (1989) Mol. Microbiol. 3, 1111–1119
10. Morfeldt, E., Berggård, K., Persson, J., Drakenberg, T., Johnsson, E., Lindahl, E., Linse, S., and Lindahl, G. (2001) J. Immunol. 167, 3870–3877
11. Stark, G., and Argos, P. (1995) Protein Eng. 8, 566–579
12. Simonson, T., and Warren, G. L. (1998) J. Biol. Chem. 273, 8623–8629
13. Uhrin, D., Uhrin, M., and Uhlíř, K. (2000) Protein Eng. Des. Sel. 13, 127–133
14. Soares, D. C., Gerloff, D. L., Syme, N. R., Coulson, A. F., Parkinson, J., and Barlow, P. N. (2000) Protein Eng. Des. Sel. 13, 1327–13278
15. Simonson, T., and Warren, G. L. (2000) J. Biol. Chem. 275, 1058–1066
16. Riepenhofer, P., and Barlow, P. N. (1994) J. Biol. Chem. 269, 547–556
17. Krych, M., Clemenza, L., Hovdenesh, D., Hauhart, R., Hourcade, D., and Atkinson, J. P. (1994) J. Biol. Chem. 269, 13273–13278
18. Koradi, R., Billeter, M., and Wuthrich, K. (1996) J. Mol. Graph. 14, 29–32, 51–55
19. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 10037–10041