

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Umakhanth Venkatraman Girija, Christopher Furze, Julia Toth, Wilhelm J. Schwaeble, Daniel A. Mitchell, Anthony H. Keeble and Russell Wallis

Article Title: Engineering Novel Complement Activity into a Pulmonary Surfactant Protein

Year of publication: 2010

Link to published version:

<http://dx.doi.org/10.1074/jbc.M109.097493>

Publisher statement: This research was originally published in The Journal of Biological Chemistry. Author(s) Umakhanth Venkatraman Girija, Christopher Furze, Julia Toth, Wilhelm J. Schwaeble, Daniel A. Mitchell, Anthony H. Keeble and Russell Wallis. Engineering Novel Complement Activity into a Pulmonary Surfactant Protein. (2010). Vol. 285, pp. 10546-10552. © the American Society for Biochemistry and Molecular Biology"

# Engineering Complement Activity into a Pulmonary-surfactant Protein<sup>1</sup>

Umakhanth Venkatraman Giriya<sup>\*§</sup>, Christopher Furze<sup>\*§</sup>, Julia Toth<sup>\*</sup>, Wilhelm J. Schwaeble<sup>\*</sup>, Daniel A. Mitchell, Anthony H. Keeble<sup>\*</sup> and Russell Wallis<sup>\*\*‡</sup>

<sup>§</sup>These authors contributed equally to the work

Running title: Engineering complement activity into SP-A

From the Departments of Infection, Immunity and Inflammation<sup>\*</sup> and Biochemistry<sup>‡</sup>,  
University of Leicester, Leicester, United Kingdom

**Address correspondence to:** Dr Russell Wallis RCUK Academic Fellow, Department of Infection, Immunity and Inflammation, Maurice Shock Building, University of Leicester PO Box 138, Leicester, LE1 9HN UK; Tel.: +44 (0)116 252 5089; Fax: +44 (0)116 252 5030; E-mail: [rw73@le.ac.uk](mailto:rw73@le.ac.uk)

Complement neutralises invading pathogens, stimulates inflammatory and adaptive immune responses and targets non- or altered-self structures for clearance. In the classical and lectin activation pathways it is initiated when complexes composed of separate recognition and activation subcomponents bind to a pathogen surface. Despite its apparent complexity, recognition-mediated activation has evolved independently in three separate protein families, C1q, mannose-binding lectins (MBLs) and serum ficolins. Although unrelated, all have bouquet-like architectures and associate with complement-specific serine proteases: MBLs and ficolins with MBL-associated serine protease-2 (MASP-2) and C1q with C1r and C1s. To examine the structural requirements for complement activation, we have created a number of novel MBLs in which the position and orientation of the MASP-binding sites have been changed. We have also engineered MASP-binding into a pulmonary-surfactant protein (SP-A), which has the same domain structure and architecture as MBL, but lacks any intrinsic complement activity. The data reveal that complement activity is remarkably tolerant to changes in the size and orientation of the collagenous stalks of MBL, implying considerable rotational and conformational flexibility in unbound MBL. Furthermore, novel complement activity is introduced concurrently with MASP binding in SP-A, but is uncontrolled and occurs even in the absence of a carbohydrate target. Thus, the active state is default and must be inhibited in circulating MBLs through additional regions

(not present in SP-A) until facilitated by pathogen recognition.

Complement is a central part of the immune system that neutralises pathogens via antibody-dependent and -independent mechanisms and stimulates a variety of protective responses including phagocytosis, inflammation and adaptive immunity (1). It is triggered via three routes called the classical, lectin and alternative pathways. Although the alternative pathway is initially non-specific and depends on host regulatory proteins to avert activation on self structures, the classical and lectin pathways both selectively target pathogen-associated molecular patterns via circulating complexes composed of recognition and zymogen-protease subcomponents (2). Upon binding to a target, conformational changes in the recognition subcomponent trigger activation of the protease subcomponent, which in turn activates the downstream complement cascade. In the classical pathway, C1q binds to microorganisms, immune complexes, apoptotic and necrotic cells as well as amyloids to initiate the stepwise activation of C1r and C1s. In the lectin pathway, mannose-binding lectin (MBL) and serum ficolins bind to terminal mannose-like epitopes or *N*-acetyl groups on pathogens to activate MBL-associated serine protease-2 (MASP-2).

Initiating complexes of the classical and lectin pathways have a number of structural and functional properties in common (2). Although unrelated, recognition subcomponents have similar domain organizations consisting of a short cysteine-containing domain followed by a collagen-like domain, which is linked to a C-terminal recognition domain (Fig. 1; (3)).

During biosynthesis, polypeptides associate into trimeric subunits (4), which assemble to form larger oligomers resembling bouquets (5, 6). In C1q, the N-terminal portions of the collagenous stalks associate with each other and splay apart at a short interruption within the repeating Gly-Xaa-Yaa sequence, called the kink. Although MBL and some ficolins also possess a kink-like region, recent measurements have revealed that the stalks probably diverge nearer the N-terminus, at the junction with the cysteine-containing domain, and thus form spider-like structures rather than the classical bouquets of C1q (7). Recognition subcomponents possess different numbers of subunits. For example, human C1q is a hexamer assembled from three different polypeptide chains (8) and rat MBLs are homooligomers comprising dimers, trimers and tetramers of subunits (9).

The protease subcomponents, C1r and C1s of the classical pathway and MASP-2 of the lectin pathway are homologous, and bind to the collagenous domains of their respective recognition subcomponent through  $\text{Ca}^{2+}$ -dependent interactions (10-13). They each comprise two  $\text{Ca}^{2+}$ -binding CUB domains (domain found in complement component C1r/C1s, Uegf, and bone morphogenic protein 1), separated by a  $\text{Ca}^{2+}$ -binding epidermal growth factor-like domain (EGF), followed by two complement-control protein (CCP) modules and a serine protease (SP) domain (14, 15). C1r and C1s form heterotetramers (C1s-C1r-C1r-C1s) and MASPs are homodimers. Although the stoichiometries of the resulting C1 and MBL-MASP complexes differ, interactions between subcomponents are analogous, with equivalent contacts between the CUB domains of the proteases and the collagenous domains of the recognition subcomponents (2). Overall, each C1r tetramer presents a total of six binding sites, one for each of the collagenous domains of C1q (one site on each CUB1 of C1r and C1s and each CUB2 of C1r). Each MASP-2 dimer presents four binding sites to MBL and ficolins (one site on each CUB domain).

A number of other proteins have bouquet-like architectures, such as adiponectin (16), emilins (17) and some collectins (18), including pulmonary surfactant protein-A (SP-A). Nevertheless, only MBLs, ficolins and C1qs are able to activate complement. As yet, the only established difference between these proteins and their structural analogues is their ability to bind

to MASPs or C1rs tetramers. The MASP-binding site in MBLs and ficolins is characterised by a distinct motif within the collagenous domain: Hyp-Gly-Lys-Xaa-Gly-Pro, where Hyp is hydroxyproline and Xaa is generally an aliphatic residue. Point mutations to this sequence disrupt MASP binding (10, 13). C1q also possesses a similar motif and we have proposed that this region forms the binding sites for C1r and C1s (2).

In this manuscript, we have further probed the functions of the recognition subcomponents by examining the structural requirements for complement activation. Initially, we modified the collagenous domain of MBL to change the position and relative orientations of the MASP-binding site on the collagenous stalks. The data reveal that MASP binding and activation are surprisingly tolerant to such changes, implying considerable rotational and conformational flexibility. To take these studies further, we have engineered the MASP-binding motif into SP-A, which also has a C-type lectin domain and a similar architecture to MBL, but cannot bind MASPs or activate complement. We find that MASP binding comparable to MBL can be engineered through just three amino acid substitutions to the collagenous domain of SP-A. The resulting proteins also activate the MASP, but lack the control mechanisms necessary to target carbohydrate surfaces selectively.

## EXPERIMENTAL PROCEDURES

*Protein Components*—Recombinant rat MBL and modified forms of rat MASP-2 were produced in a Chinese hamster ovary cell-expression system and purified as described previously by affinity chromatography on mannose-Sepharose and nickel-Sepharose columns, respectively (9, 19). MASP-2A is a catalytically inactive form of rat MASP-2 in which the active site serine at position 613 has been replaced by an alanine. MASP-2K is a catalytically active form, in which the arginine residue at the cleavage site for zymogen activation ( $\text{Arg}^{424}$ ) has been changed to a lysine residue to slow down the rate of spontaneous autoactivation during biosynthesis. Mutant forms of MBL were created by PCR and expressed in the same way as the wild-type protein.

*Cloning and production of recombinant wild type and mutant SP-As*—The cDNA of SP-

A isoform A1 was cloned into the polylinker of the eukaryotic expression vector pED-4 (20), which contains the dihydrofolate reductase cDNA as a selectable marker. Proteins were produced by expression in the Chinese hamster ovary cell line DXB11 following amplification using the dihydrofolate reductase inhibitor, methotrexate, to a final concentration of 25  $\mu$ M. Culture media was harvested using a protocol that has been described previously for production of recombinant MBL (9, 21). Protein was purified by affinity chromatography on mannose-Sepharose columns (1 ml column for 250 mls of media). Following washes with high salt (50 mM Tris pH 7.5, containing 150 mM NaCl and 2.5 mM CaCl<sub>2</sub>) and low salt buffers (10 mM Tris pH 7.5, containing 10 mM NaCl and 2.5 mM CaCl<sub>2</sub>) protein was eluted in 0.5 ml fractions of 10 mM Tris pH 7.5, containing 10 mM NaCl and 2.5 mM EDTA. Fractions containing recombinant protein were identified by SDS-polyacrylamide gel electrophoresis. Mutations were introduced into the cDNA of SP-A by PCR and the encoded proteins were produced in the same way as the wild type protein. Although SP-A tolerated relatively high salt concentrations when bound to the Sepharose matrix, concentrations >10 mM caused significant aggregation in solution.

*Surface plasmon resonance*—Measurements were performed using a BIAcore 2000 instrument (GE Healthcare) or a BioRad ProteOn XPR36 biosensor. Protein ligands were diluted into 10 mM sodium acetate (pH 4.5 for MBL or pH 5.0 for SP-A), and immobilized on to the carboxymethylated dextran surface of a CM5 sensor chip (GE Healthcare) or a GLM chip (BioRad), using amine coupling chemistry. Binding was measured in 10 mM Tris-HCl (pH 7.4) containing 140 mM NaCl, 2 mM CaCl<sub>2</sub> and 0.005% surfactant SP40, at a flow rate of 25  $\mu$ l/min and at 25°C. After injection of ligand, the protein surface was regenerated by injection of 10  $\mu$ l of 10 mM Tris-HCl buffer (pH 7.4) containing 1 M NaCl and 5 mM EDTA. Data were analyzed by fitting association and dissociation curves to Langmuir binding models for several protein concentrations simultaneously, using BIAevaluation 4.1 software (GE Healthcare). Increasingly complex models were tested until a satisfactory fit to the data was achieved. Apparent equilibrium

dissociation constants ( $K_D$ ) were calculated from the ratio of the dissociation and association rate constants ( $k_{off}/k_{on}$ ). MBL was immobilized on the chip surface rather than used as a soluble ligand, because it bound to the chip thereby masking analysis of the protein-protein interactions. SP-A was also immobilized because it tended to precipitate on the chip surface.

*MASP-2 activation assays*—Activation was measured by following MASP autolysis using a modified version of the protocol described previously (22). Briefly, MASP-2K was mixed with wild-type or mutant MBL or SP-A and added to a suspension of fucose-Sepharose (5  $\mu$ l of a 1:1 v/v suspension in a total volume of 30  $\mu$ l), in 50 mM Tris-HCl (pH 7.5) containing 150 mM NaCl and 5 mM CaCl<sub>2</sub>, at 37°C with mixing. A 1.2-fold molar excess of MBL or SP-A was used to ensure that all the MASP-2K was bound to the recognition molecule. The mixture was incubated at 37 °C with shaking and aliquots of the suspension were removed from the reaction mix at various times and immediately frozen on dry ice to quench the reaction. Proteins were separated by SDS-polyacrylamide gel electrophoresis under reducing conditions and the amount of MASP cleaved was quantified by densitometry. Data are the mean  $\pm$  S.E. from at least two separate experiments using different protein preparations, unless otherwise stated.

## RESULTS

*MASP-binding and activation are tolerant to changes in the position and orientation of binding sites on MBL*—MASPs bind inside the cone created by the collagenous stalks of MBL and bridge up to four stalks simultaneously (Fig. 1; (3)). Consequently, during complex formation, the binding sites on the MBL stalks must either face inwards or re-orientate so that the correct collagen-CUB domain interactions are formed. If the stalks are relatively rigid or are unable to re-orientate independently, any change in their position or alignment would lead to loss of function. In order to examine the tolerance of MBL to such changes, we made a series of modified MBLs in which the collagenous domain was modified. Because the pitch of collagen is typically  $\sim$ 3.5 residues (23), deletion or insertion of a Gly-Xaa-

Yaa triplet would change both the position of each binding site and its relative orientation. For example, deletion of a triplet would effectively rotate each collagenous stalk by  $\sim 180^\circ$  as well as shift the position of the binding site further towards the N-terminus (Fig. 1). Deletion of a second triplet would restore the original alignment but move the binding sites even further towards the N-terminus. To accommodate all of these changes, the collagenous stalks would not only have to splay further apart (for a deletion) or move closer together (for an insertion), but also rotate independently of each other. Thus, analysis of MASP binding and activation by the modified MBLs provides a useful probe of the conformational and rotational flexibility of MBL.

Two deletion mutants were created in which either one or two Gly-Xaa-Yaa triplets were removed from the collagenous domain, thus moving the MASP-binding motif closer to the N-terminus of MBL (Fig. 1). In addition, a single Gly-Pro-Hyp was inserted near the kink-like region to move the MASP binding motif further towards the C-terminus. We also made an MBL in which the kink-like region itself was removed by insertion of a proline residue to restore the Gly-Xaa-Yaa tandem repeats. A previous study has demonstrated that the kink-like region is not essential for complement activation by MBL (24). However, the kinetics of activation were not investigated in this work, so more subtle changes might have been missed. Analysis of the purified, recombinant MBLs by SDS-polyacrylamide gel electrophoresis under non-reducing conditions, shown previously to be a sensitive indicator of MBL assembly (25), demonstrated that all assembled correctly during biosynthesis (Fig. 2).

Initially MBLs were tested for their abilities to bind to MASP-2A, using surface plasmon resonance. As expected, wild-type bound with complex kinetics compatible with the formation of 1:1 complexes, together with binding of a second MASP at relatively high protein concentrations (19). The  $K_D$ s were 3.7 and 164.5 nM consistent with previous measurements (2). Surprisingly, the entire set of mutant MBLs bound to MASP-2A with similar affinities and kinetics (Table I), indicating that none of the changes perturbed MASP binding (Fig. 2). We next examined their complement activities by monitoring MBL-dependent MASP-2K autoactivation with mannose-Sepharose as an

activating target (Fig. 3). In the presence of the target, all mutant MBLs activated the MASP with comparable rates as wild-type ( $t^{1/2} \sim 20$  mins). The greatest difference observed was for the double-deletion mutant, which activated the MASP just two-fold more slowly ( $t^{1/2} \sim 40$  mins), despite removal of six amino acid residues from the collagenous domain. In the absence of the target, activation was relatively slow and similar to autoactivation by MASP-2 alone ( $>1000$  min), indicating that all MBLs only activate on a carbohydrate target. Thus, neither insertions, deletions nor removal of the kink-like region disrupt MASP-2 binding or activation, despite large changes to the position and relative orientations of the MASP-binding sites. The tolerance of MBL to these changes must reflect significant adjustment and realignment of the collagenous stalks. Given that collagenous domains themselves are relatively rigid and inflexible, changes must occur at the N-terminus where the stalks converge, probably at the junction between the cysteine-rich domain and the collagenous domain.

*Engineering Complement Activity into a Pulmonary-surfactant Protein*—Given its surprising tolerance to even large changes in MBL structure, we sought to further characterize MASP activation by engineering novel activity into a bouquet-like template. We chose SP-A for these studies because of its similar structure, but lack of innate complement activity. SP-A, like MBL, is a member of the collectin family of animal lectins with an N-terminal collagen-like domain and a C-terminal CRD (26). It has similar sugar specificity to MBL (but with a preference for L-fucose over mannose (27)), allowing purification by affinity chromatography and assessment of complement activity using similar activation assays. Although structurally analogous to MBL, the collagenous domains of SP-A and MBL are of different lengths and share no apparent sequence identity (other than the repeating Gly-Xaa-Yaa motif characteristic of all collagens; Fig. 4). SP-A possesses a natural kink near the middle of the collagenous domain where the collagenous stalks splay apart, similar to C1q. Previous studies have shown that upon removal of the kink, the stalks diverge nearer the N-terminus of polypeptides, at the junction between the N-terminal domain and the

collagenous domain, thus becoming more similar to MBL and ficolins (28).

Three new recombinant constructs were created in which all or part of the MASP-binding motif was introduced into the collagenous domain of SP-A (Fig. 4). Changes were made five Gly-Xaa-Yaa triplets towards the C-terminal side of the kink, to ensure that there would be enough space for the MASP to bind within the SP-A bouquet. In SPA-KL, Glu63 and Cys64 were replaced by lysine and leucine residues, respectively. Lysine in the Xaa position is known to be essential for MASP binding and complement activation in both MBLs and ficolins. The adjacent cysteine residue was also replaced to avoid any potential disulfide bond formation which might interfere with MASP binding. In SPA-KLP a proline residue was introduced in place of the glutamate at position 66 to further optimize the MASP-binding motif: Hyp-Gly-Lys-Leu-Gly-Pro. One further change was made in SPA-KLPO, replacing Lys67 by a hydroxyproline residue. Lysine residues in the Yaa position of collagen are often hydroxylated and glycosylated during biosynthesis, so this change was designed to prevent any potential steric inhibition of MASP binding by the sugar residues (23). We created two further variants of SPA-KLPO, in which the kink was removed or modified to change the point at which the stalks splay apart (28), and thus more resemble MBL and ficolins. In one mutant, SPA-KLPOdel the kink sequence (PCPP) was removed completely, and in the other, called SPA-KLPOS, the cysteine residue was replaced by a serine residue to remove the potential for disulfide bond formation, which is believed to tether the separate stalks together in native SP-A (28).

Following production in chinese hamster ovary cells, all proteins bound to mannose- or fucose-Sepharose columns, demonstrating that they were correctly folded. As expected, wild-type SP-A did not bind to MASP-2A at any of the concentrations tested (up to 1  $\mu$ M). Surprisingly however, SPA-KL bound with appreciable affinity (Fig. 5). The data best fitted a two-complex, parallel-reaction model, with apparent dissociation constants  $K_{D1}$  and  $K_{D2}$  of 600 and 2120 nM respectively, compared to 3.8 and 166 nM for MBL-MASP (Table 2). SPA-KLP bound with even higher affinity, with  $K_{Ds}$  of 36 and 104 nM, only slightly weaker than MBL. Additional changes had little effect on

MASP binding, so the naturally occurring lysine residue in SP-A (or its potential glycosylated derivative) does not prevent MASP access. Removal of the kink region also had little effect, indicating that the point at which the stalks splay apart and thus the angle between adjacent stalks is not limiting for MASP binding. Thus, just three amino acid changes to the collagenous domain are sufficient to introduce MASP binding to SP-A almost comparable to MBL itself.

We next examined if the modified SP-As could activate MASP-2K. The half-time for activation by MBL was ~50 mins, using fucose-Sepharose as a target. As expected, wild type SP-A, which does not bind to the MASP at all and SPA-KL, which binds only weakly, did not activate the MASP (Fig. 6). In each case the measured rate was similar to the intrinsic autoactivation rate of MASP-2K alone ( $t^{1/2} >1000$  min). Surprisingly however, SPA-KLP and SPA-KLPO activated the MASP significantly faster than the basal rate ( $t^{1/2} \sim 260$  min). Similar rates were also measured for SPA-KLPOdel and SPA-KLPOS, demonstrating that the kink is neither necessary for MASP activation, nor modulates the rate of activation significantly. To determine whether activation was target dependent, assays were repeated in the absence of fucose-Sepharose. Surprisingly, comparable rates were observed (Fig. 7). Activation was still much faster than the basal rate, so must be driven by SP-A binding, but was constitutive occurring even in the absence of a carbohydrate target. Thus, modified SP-As lack the control mechanisms required to activate MASP selectively.

## DISCUSSION

The data presented here reveal that novel MASP-binding and complement activation can be introduced into a bouquet-like template through relatively minor changes to the chemical make-up of the collagenous stalks. Just three substitutions: Glu63 to Lys, Cys64 to Leu and Glu66 to Pro, establish MASP binding in SP-A, which is almost equivalent to that of MBL. Of these residues, only the lysine is essential for binding (10, 13). Generally, the adjacent residue in the Yaa position (leucine in rat MBL and the SP-A mutants) has an aliphatic side chain, although a small polar residue such as serine can also be accommodated, but glutamate cannot (10). The proline residue in the Xaa position of

the next Gly-Xaa-Yaa triplet is also not essential for MASP binding and can be replaced by an alanine with only a small decrease in affinity. The main function of the proline is probably to help stabilize the binding region (10). It is notable that a significant component of the difference in MASP binding by SPA-KLP relative to SPA-KL is due to faster association, so probably reflects removal of the unfavourable electrostatic interactions mediated by the glutamate, rather than the introduction of the proline residue itself. The presence of any additional binding sites for MASP on MBL can be completely ruled out from the data presented here. Overall therefore, specificity of the MBL-MASP interaction appears to be driven mainly by a single lysine residue in the collagen-like domain.

Several other binding motifs have been identified in collagens and typically extend across two or more Gly-Xaa-Yaa triplets. For example, the I domain of  $\alpha 2$  integrin, recognizes the sequence GFOGER (29) and the extracellular matrix protein SPARC binds to the sequence GVMGFO (30). Thus, the less stringent binding requirements of MASP-2 are relatively unusual. Nevertheless, it is notable that almost all of the other lysine residues in the collagenous domains of MBL and ficolins are in the Yaa position of the Gly-Xaa-Yaa repeat, so are likely to be post-translationally modified by hydroxylation and glycosylation, which probably blocks any potentially incorrect MASP interactions. Perhaps even more importantly, MASP binds to MBL and ficolins through multiple relatively weak interactions involving up to four separate CUB-collagen contacts. Thus, much of the binding specificity is probably mediated through the architectures of the subcomponents and the multivalent nature of the interactions.

Complement activation is thought to be triggered by changes in the angle between collagenous stalks when complexes bind to an activating target (3, 31). The data presented here indicate that significant conformational and rotational flexibility of the stalks must be permitted, implying that unbound MBL is probably quite flexible. Both relaxed-to-strained and strained-to-relaxed mechanisms could explain the activation process. The traditional view is that complexes circulate in an inactive zymogen form and distort upon binding to a target, and the changes to the structure of MBL are transmitted to the MASP to drive

autoactivation. However, in an alternative but equally valid possibility, binding of the zymogen MASP to MBL (or ficolin) induces strain into circulating complexes, which is released on pathogen binding to trigger activation. A major difference between these mechanisms is that the relaxed or default states of the complexes would differ: inactive in a relaxed-to-strained mechanism, but active in a strained-to-relaxed mechanism. In the work described here, MASP-binding leads to constitutive activation by SP-A, even in the absence of a target. Thus, activation is the default state, as would be expected in a strained-to-relaxed mechanism. The possibility that modified SP-As are locked into a high energy conformation that activates MASP constitutively cannot be completely excluded. However, SP-A must also be flexible to bind to the MASP multivalently, which would seem to be at odds with this possibility. Furthermore, large structural changes (through removal of the kink) do not affect MASP activation by SP-A, which also seems incompatible with such a mechanism.

Based on these data, we propose that MBL-MASPs also activate via a strained-to-relaxed mechanism. We have recently suggested that activation of the C1 complex also operates in this way (2) and have proposed that strain is induced into the zymogen complex through the interactions between the protease domains of C1r. In this respect, the mechanism must differ in MBL-MASP complexes because the SP domains of a MASP dimer do not interact significantly until the moment of autocatalysis. Interestingly however, recent analysis suggests that the MBL stalks are not evenly distributed, but rather are highly asymmetrical and separated by  $\sim 35\text{--}40^\circ$  between nearest neighbours (32). In this case, significant strain would be induced when MBL binds to a MASP. Release of this strain upon target recognition might drive the changes that initiate complement activation.

From an evolutionary perspective, these data suggest that recognition-mediated complement activation could have arisen relatively rapidly in a collectin-like molecule such as SP-A, as result of a small number of changes to the collagen-like domain. It would also explain how similar activation mechanisms could have arisen multiple times in structurally analogous, but unrelated recognition components. Clearly however, regulation of activation necessitates additional sequences, and

we are currently undertaking studies to identify these regions in MBL and ficolins.

## REFERENCES

1. Porter, R. R., and Reid, K. B. M. (1978) The biochemistry of complement, *Nature* 275, 699-704.
2. Phillips, A. E., Toth, J., Dodds, A. W., Girija, U. V., Furze, C. M., Pala, E., Sim, R. B., Reid, K. B., Schwaeble, W. J., Schmid, R., Keeble, A. H., and Wallis, R. (2009) Analogous interactions in initiating complexes of the classical and lectin pathways of complement, *J Immunol* 182, 7708-7717.
3. Wallis, R., Mitchell, D. A., Schmid, R., Schwaeble, W. J., and Keeble, A. H. Paths reunited: Initiation of the classical and lectin pathways of complement activation, *Immunobiology* 215, 1-11.
4. Heise, C. T., Nicholls, J. R., Leamy, C. E., and Wallis, R. (2000) Impaired secretion of rat mannose-binding protein resulting from mutations in the collagen-like domain, *J Immunol* 165, 1403-1409.
5. Brodsky-Doyle, B., Leonard, K. R., and Reid, K. B. M. (1976) Circular-dichroism and electron-microscopy studies of human subcomponent C1q before and after limited proteolysis by pepsin, *Biochemical Journal* 159, 279-286.
6. Strang, C. J., Siegel, R. C., Phillips, M. L., Poon, P. H., and Schumaker, V. N. (1982) Ultrastructure of the first component of human complement: electron microscopy of the crosslinked complex, *Proc Natl Acad Sci U S A* 79, 586-590.
7. Jensenius, H., Klein, D. C., van Hecke, M., Oosterkamp, T. H., Schmidt, T., and Jensenius, J. C. (2009) Mannan-binding lectin: structure, oligomerization, and flexibility studied by atomic force microscopy, *J Mol Biol* 391, 246-259.
8. Reid, K. B., and Porter, R. R. (1976) Subunit composition and structure of subcomponent C1q of the first component of human complement, *Biochem J* 155, 19-23.
9. Wallis, R., and Drickamer, K. (1999) Molecular determinants of oligomer formation and complement fixation in mannose-binding proteins, *J Biol Chem* 274, 3580-3589.
10. Girija, U. V., Dodds, A. W., Roscher, S., Reid, K. B., and Wallis, R. (2007) Localization and Characterization of the Mannose-Binding Lectin (MBL)-Associated-Serine Protease-2 Binding Site in Rat Ficolin-A: Equivalent Binding Sites within the Collagenous Domains of MBLs and Ficolins, *J Immunol* 179, 455-462.
11. Reid, K. B., Sim, R. B., and Faiers, A. P. (1977) Inhibition of the reconstitution of the haemolytic activity of the first component of human complement by a pepsin-derived fragment of subcomponent C1q, *Biochem J* 161, 239-245.
12. Teillet, F., Lacroix, M., Thiel, S., Weilguny, D., Agger, T., Arlaud, G. J., and Thielens, N. M. (2007) Identification of the site of human mannan-binding lectin involved in the interaction with its partner serine proteases: the essential role of Lys55, *J Immunol* 178, 5710-5716.
13. Wallis, R., Shaw, J. M., Uitdehaag, J., Chen, C. B., Torgersen, D., and Drickamer, K. (2004) Localization of the serine protease-binding sites in the collagen-like domain of mannose-binding protein: indirect effects of naturally occurring mutations on protease binding and activation, *J Biol Chem* 279, 14065-14073.
14. Feinberg, H., Uitdehaag, J. C., Davies, J. M., Wallis, R., Drickamer, K., and Weis, W. I. (2003) Crystal structure of the CUB1-EGF-CUB2 region of mannose-binding protein associated serine protease-2, *Embo J* 22, 2348-2359.
15. Sim, R. B., and Tsiftoglou, S. A. (2004) Proteases of the complement system, *Biochem Soc Trans* 32, 21-27.
16. Scherer, P. E., Williams, S., Fogliano, M., Baldini, G., and Lodish, H. F. (1995) A novel serum protein similar to C1q, produced exclusively in adipocytes, *J Biol Chem* 270, 26746-26749.
17. Doliana, R., Mongiat, M., Bucciotti, F., Giacomello, E., Deutzmann, R., Volpin, D., Bressan, G. M., and Colombatti, A. (1999) EMILIN, a component of the elastic fiber and a new



- member of the C1q/tumor necrosis factor superfamily of proteins, *J Biol Chem* 274, 16773-16781.
18. Reid, K. B. M., Colomb, M. G., and Loos, M. (1998) Complement component C1 and the collectins: parallels between routes of acquired and innate immunity, *Immunology Today* 12, 56-59.
  19. Chen, C. B., and Wallis, R. (2001) Stoichiometry of complexes between mannose-binding protein and its associated serine proteases. Defining functional units for complement activation, *J Biol Chem* 276, 25894-25902.
  20. Kaufman, R. J., Davies, M. V., Wasley, L. C., and Michnick, D. (1991) Improved vectors for stable expression of foreign genes in mammalian cells by use of the untranslated leader sequence from EMC virus, *Nucleic Acids Research* 19, 4485-4490.
  21. Wallis, R., and Drickamer, K. (1997) Asymmetry adjacent to the collagen-like domain in rat liver mannose-binding protein, *Biochem J* 325 ( Pt 2), 391-400.
  22. Chen, C. B., and Wallis, R. (2004) Two mechanisms for mannose-binding protein modulation of the activity of its associated serine proteases, *J Biol Chem* 279, 26058-26065.
  23. Shoulders, M. D., and Raines, R. T. (2009) Collagen structure and stability, *Annu Rev Biochem* 78, 929-958.
  24. Kurata, H., Cheng, H. M., Kozutsumi, Y., Yokota, Y., and Kawasaki, T. (1993) Role of the collagen-like domain of the human serum mannan-binding protein in the activation of complement and the secretion of this lectin, *Biochem Biophys Res Commun* 191, 1204-1210.
  25. Wallis, R., and Cheng, J. Y. (1999) Molecular defects in variant forms of mannose-binding protein associated with immunodeficiency, *J Immunol* 163, 4953-4959.
  26. Hoppe, H.-J., and Reid, K. B. M. (1994) Collectins - soluble proteins containing collagenous regions and lectin domains - and their roles in innate immunity, *Protein Science* 3, 1143-1158.
  27. Haurum, J. S., Thiel, S., Haagsman, H. P., Laursen, S. B., Larsen, B., and Jensenius, J. C. (1993) Studies on the carbohydrate-binding characteristics of human pulmonary surfactant-associated protein A and comparison with two other collectins: mannan-binding protein and conglutinin, *Biochem J* 293 ( Pt 3), 873-878.
  28. Uemura, T., Sano, H., Katoh, T., Nishitani, C., Mitsuzawa, H., Shimizu, T., and Kuroki, Y. (2006) Surfactant protein A without the interruption of Gly-X-Y repeats loses a kink of oligomeric structure and exhibits impaired phospholipid liposome aggregation ability, *Biochemistry* 45, 14543-14551.
  29. Emsley, J., Knight, C. G., Farndale, R. W., Barnes, M. J., and Liddington, R. C. (2000) Structural basis of collagen recognition by integrin alpha2beta1, *Cell* 101, 47-56.
  30. Hohenester, E., Sasaki, T., Giudici, C., Farndale, R. W., and Bachinger, H. P. (2008) Structural basis of sequence-specific collagen recognition by SPARC, *Proc Natl Acad Sci U S A* 105, 18273-18277.
  31. Wallis, R. (2007) Interactions between mannose-binding lectin and MASPs during complement activation by the lectin pathway, *Immunobiology* 212, 289-299.
  32. Jensenius, H., Klein, D. C., van Hecke, M., Oosterkamp, T. H., Schmidt, T., and Jensenius, J. C. (2009) Mannan-Binding Lectin: Structure, Oligomerization, and Flexibility Studied by Atomic Force Microscopy, *J Mol Biol.*

### Acknowledgements

We thank Robert Freedman and Katrine Wallis, of the Department of Biological Sciences, University of Warwick for use of the Departmental Biacore facility.

### Footnotes

<sup>1</sup>Funding for this work was supplied by Grant G0501425 from the Medical Research Council and Grant 077400 from The Wellcome Trust. R.W. and DM are Research Council U.K. Academic Fellows. AK is a Wellcome Trust Value in People Award Fellow.

<sup>2</sup>Abbreviations used in this paper: MBL, mannan-binding lectin; CRD, carbohydrate-binding domain; MASP, MBL-associated serine protease; CUB, domain found in complement component C1r/C1s, Uegf, and bone morphogenic protein 1; EGF, epidermal growth factor

**Table I Kinetic properties of modified MBLs**

	$k_{on}$ ( $\times 10^4$ ) ( $M^{-1}s^{-1}$ )	$k_{off}$ ( $\times 10^{-3}$ ) ( $s^{-1}$ )	$K_D^a$ (nM)	Half-time for MASP activation (min)
MBL	$40 \pm 1$	$1.5 \pm 0.1$	$3.7 \pm 0.1$	$20 \pm 1$
	$2.9 \pm 0.1$	$4.75 \pm 0.1$	$165 \pm 2$	
Deletion	$38 \pm 4$	$3.8 \pm 1.7$	$11 \pm 6$	$22 \pm 1$
	$3.1 \pm 0.7$	$5.4 \pm 1.3$	$175 \pm 2$	
Double deletion	$39 \pm 1$	$3.1 \pm 0.1$	$8.2 \pm 0.3$	$39 \pm 3$
	$3.3 \pm 0$	$6.4 \pm 0.1$	$195 \pm 3$	
Insertion	$44 \pm 3$	$2.2 \pm 0.6$	$4.9 \pm 1.2$	$21 \pm 1$
	$4.6 \pm 0$	$9.1 \pm 0.2$	$199 \pm 5$	
Hinge	$57 \pm 1$	$6.2 \pm 2.7$	$10 \pm 2$	$21 \pm 1$
	$5.2 \pm 0.2$	$6.3 \pm 0.2$	$136 \pm 9$	

<sup>a</sup> $K_D$  values were calculated from  $k_{off}/k_{on}$  for each experiment and were averaged from two separate experiments

**Table II Kinetic properties of engineered SP-As**

	$k_{on}$ ( $\times 10^4$ ) ( $M^{-1}s^{-1}$ )	$k_{off}$ ( $\times 10^{-3}$ ) ( $s^{-1}$ )	$K_D^a$ (nM)	Half-time for MASP activation <sup>c</sup> (min)
SP-A	N.B <sup>b</sup>	N.B <sup>b</sup>	N.B <sup>b</sup>	$>1000^d$
SPA-KL	$1.2 \pm 0.2$	$7.0 \pm 0.1$	$600 \pm 110$	$>1000^d$
	$4.3 \pm 0.1$	$91 \pm 2$	$2120 \pm 100$	
SPA-KLP	$5.4 \pm 2.9$	$1.4 \pm 0.1$	$36 \pm 20$	$270 \pm 56$
	$36 \pm 24$	$17.6 \pm 1.0$	$104 \pm 77$	
SPA- KLPO	$2.6 \pm 0.8$	$2.5 \pm 0.6$	$97 \pm 7$	$256 \pm 96$
	$9.6 \pm 1.0$	$22 \pm 3.1$	$232 \pm 9$	
SPA- KLPOdel	$5.3 \pm 2.3$	$1.7 \pm 0.1$	$29 \pm 17$	$316 \pm 171$
	$18 \pm 0.0$	$21 \pm 1$	$114 \pm 2$	
SPA- KLPOS	$4.5 \pm 2.6$	$4.6 \pm 1.2$	$99 \pm 18$	$166 \pm 43$
	$21 \pm 2$	$37 \pm 15$	$203 \pm 31$	

<sup>a</sup> $K_D$  values were calculated from  $k_{off}/k_{on}$  for each experiment and were averaged from two separate experiments

<sup>b</sup>No binding detected

<sup>c</sup>data from three separate experiments

<sup>d</sup>activation occurred a rate similar to autoactivation of MASP-2 alone

## Figure Legends

**FIGURE 1. Changes to the collagenous domains of MBLs.** A, Domain organisations of MBL, ficolins and C1q. The position of the kink or kink-like region is marked by an asterisk. The location of the binding sites for MASPs or C1r/C1s on the collagenous domains are shown in *light grey*. The recognition domains are: carbohydrate-recognition domain, fibrinogen-like domain and C1q-globular domain, respectively. B, Aligned sequences of the collagenous domains of wild-type and modified MBLs. Residues of the protease-binding motif are shown in *light blue* and the motif is shaded in *light grey*. The GQ residues that form the kink-like region are *boxed* and shown in *blue*. Residues inserted into the collagenous domain are shown in *red*. Glycosylated hydroxylysine residues are *underlined*. C, Predicted changes to the alignment of the MASP-binding sites on MBL

subunits as a result of a single insertion or deletion of a Gly-Xaa-Yaa into the collagenous domain based upon a pitch of 3.5 for a collagen helix. The figure shows a cross-section through the four collagenous stalks, which are represented by *circles*. Each binding site is shaded in *light grey*. The structure of the CUB1-EGF-CUB2 domains of MASP-2 is from (14). *D*, Schematic representation of the effects of changes to the position of the MASP-binding sites on MBL. Only two MBL stalks are shown for clarity. The MASP is shown as a box and is shaded in *light grey*. CRDs are represented as *grey circles*. The *arrow* shows the position of the insertion/deletion. To accommodate the MASP, the collagenous stalks would have to move closer together for an insertion or splay further apart for a deletion.

**FIGURE 2. MASP-binding by modified MBLs.** *A*, SDS polyacrylamide gel electrophoresis of wild-type and mutant MBLs under reducing (*left*) and non-reducing conditions (*right*). Mutant MBLs are assembled from single polypeptide chains, which migrate as a ladder of covalently-linked polypeptides under non-reducing conditions, characteristic of the heterogeneous nature of oligomers of wild-type MBL (9). *B*, Binding of MASP-2 to immobilised MBLs by surface plasmon resonance. MASP-2 was injected at 333, 167, 83, 42 nM over immobilized MBL (~7000 response units). All data were fitted to a two-complex parallel binding model and the fits are shown by the dotted lines.

**FIGURE 3. MASP activation by modified MBLs.** Kinetics of MBL-MASP-2 activation analyzed by SDS polyacrylamide gel electrophoresis. *A* and *B*, wild-type MBL and the deletion mutant were incubated with MASP-2K together with mannose-Sepharose as an activating target. Proteins were separated on a 4–12% linear gradient gel under reducing conditions and were stained with Coomassie blue. The N-terminal fragment of MASP-2K runs as a double band due to differential glycosylation. MASP activation was measured by quantifying cleavage of the MASP polypeptide. *B*, Comparison of MASP-2K activation by wild-type and mutant MBLs.

**FIGURE 4. Design of modified SP-As.** Sequences of the collagenous domains of MBL (top) and SP-A (bottom) Sequences were crudely aligned based on the positions of the kink of SP-A and the kink-like region of MBL (*boxed*). The MASP-binding motif is shaded *light grey*. Changes to the sequence of SP-A are indicated.

**FIGURE 5. MASP-binding by modified SP-As.** Binding of MASP-2 to immobilised SP-As by surface plasmon resonance. MASP-2 was injected at 764, 447, 261, 152 and 89 nM over each of the immobilized SP-As (~6000 response units). All data were fitted to a two-complex parallel binding model and the fits are shown by the dotted lines.

**FIGURE 6. MASP-2 activation by modified SP-As.** Assays fucose-Sepharose was used as the activating ligand. Proteins were separated on a 15% SDS polyacrylamide gel and were stained with Coomassie blue. *B*, MASP-2K activation by wild-type and mutant SP-As. Averaged data from three separate experiments are shown.