

Reduced order modeling and analysis of the human complement system

Adithya Sagar, Wei Dai[#], Mason Minot[#], and Jeffrey D. Varner^{*}

School of Chemical and Biomolecular Engineering

Cornell University, Ithaca NY 14853

Running Title: A reduced order model of complement

To be submitted: *Scientific Reports*

[#] Denotes equal contribution

^{*}Corresponding author:

Jeffrey D. Varner,

Professor, School of Chemical and Biomolecular Engineering,

244 Olin Hall, Cornell University, Ithaca NY, 14853

Email: jdv27@cornell.edu

Phone: (607) 255 - 4258

Fax: (607) 255 - 9166

Abstract

Complement is an important pathway in innate immunity, inflammation, and many disease processes. However, despite its importance, there have been few validated mathematical models of complement activation. In this study, we developed an ensemble of experimentally validated reduced order complement models. We combined ordinary differential equations with logical rules to produce a compact yet predictive complement model. The model, which described the lectin and alternative pathways, was an order of magnitude smaller than comparable models in the literature. We estimated an ensemble of model parameters from *in vitro* dynamic measurements of the C3a and C5a complement proteins. Subsequently, we validated the model on unseen C3a and C5a measurements not used for model training. Despite its small size, the model was surprisingly predictive. Global sensitivity and robustness analysis suggested complement was robust to any single therapeutic intervention. Only dual knockdown of both C3 and C5 consistently reduced C5a formation when all pathways were activated. Taken together, we developed a reduced order complement model that was computationally inexpensive, and could easily be incorporated into pre-existing or new pharmacokinetic models of immune system function. The model described experimental data, and predicted the need for multiple points of therapeutic intervention to disrupt complement activation.

Keywords: Complement system, systems biology, reduced order models, biochemical engineering

1 Introduction

2 Complement is an important pathway in innate immunity. It plays a significant role in
3 inflammation, host defense as well as many disease processes. Complement was dis-
4 covered in the late 1880s where it was found to 'complement' the bactericidal activity of
5 natural antibodies (1). However, research over the past decade has shown that the im-
6 portance of complement extends well beyond innate immunity. For example, complement
7 contributes to tissue homeostasis by inducing tissue repair (2). Complement has also
8 been linked with several diseases including Alzheimers, Parkinson's disease, multiple
9 sclerosis, schizophrenia, rheumatoid arthritis and sepsis (3, 4). Complement plays both
10 positive and negative roles in cancer; attacking tumor cells with altered surface proteins
11 in some cases, while potentially contributing to tumor growth in others (5, 6). Lastly, sev-
12 eral other important biochemical subsystems are integrated with complement including
13 the coagulation cascade, the autonomous nervous system and inflammation (6). Thus,
14 complement is important in a variety of beneficial and potentially harmful functions in the
15 body.

16 The complement cascade involves over 30 soluble and cell surface proteins, receptors
17 and regulators. The molecular connectivity of complement involves both fluid phase and
18 cell surface events, see the review of Walport (7, 8). The central outputs of complement
19 are the Membrane Attack Complex (MAC), and the inflammatory mediator proteins C3a
20 and C5a. The membrane attack complex, generated during the terminal phase of the re-
21 sponse, forms transmembrane channels which disrupt the membrane integrity of targeted
22 cells, leading to cell lysis and death. On the other hand, the C3a and C5a proteins act as
23 a bridge between innate and adaptive immunity, and play an important role in regulating
24 inflammation (5). Complement activation takes places through three pathways: the clas-
25 sical, the lectin binding and the alternate pathways. The classical pathway is triggered by
26 antibody recognition of foreign antigens or other pathogens. A multimeric protein com-

plex C1 binds to antibody-antigen complexes and undergoes a conformational change, leading to an activated form with proteolytic activity. This activated complex then cleaves soluble complement proteins C4 and C2 into C4a, C4b, C2a and C2b, respectively. The C4a and C2b fragments bind to form the C4bC2a protease, also known as the classical C3 convertase. The lectin pathway is initiated through the binding of L-ficolin or Mannose Binding Lectin (MBL) to carbohydrates on the surfaces of bacterial pathogens. These complexes, in combination with the associated mannose-associated serine proteases 1 and 2 (MASP-1/2), also cleave C4 and C2, leading to additional classical C3 convertase. Thus, the classical and lectin pathways, initiated by the recognition of different cues on foreign surfaces, converge at the classical C3 convertase. However, the alternate pathway works differently. It is activated by a 'tickover' mechanism in which complement protein C3 is spontaneously hydrolyzed to form an activated intermediate C3w; C3w recruits factor B and factor D, leading to the formation of C3wBb. C3wBb cleaves C3 into C3a and C3b, where the C3b fragment further recruits additional factor B and factor D to form C3bBbC3b, the alternate C3 convertase (9). The role of classical and alternate C3 convertases is varied. First, C3 convertases mediate signal amplification. C3 convertases cleave C3 into C3a and C3b; the C3b fragment is then free to form additional alternate C3 convertases, thereby forming a positive feedback loop. Next, C3 convertases link complement initiation with the terminal phase of the cascade through the formation of C5 convertases. Both classical and alternate C3 convertases can recruit C3b subunits to form the classical C5 convertase (C4bC2aC3b), and the alternate C5 convertase (C3bBbC3b), respectively. C5 convertases cleave C5 into the C5a and C5b fragments. The C5b fragment, along with the C6, C7, C8 and multiple C9 complement proteins, form the membrane attack complex. On the other hand, both C3a and C5a are important inflammatory signals involved in several responses (7, 8). Thus, the complement cascade directly attacks the invading pathogen, and signals for help from the immune system.

The complement cascade is regulated by plasma and host cell proteins which balance host cell safety with effectiveness. The initiation of the classical pathway via complement protein C1 is controlled by the C1 Inhibitor (C1-Inh). C1-Inh irreversibly binds to and deactivates the active subunits of C1, preventing chronic complement activation (10). Regulation of upstream processes in the lectin and alternate pathways also occurs through the interaction of the C4 binding protein (C4BP) with C4b, and factor H with C3b (11). Interestingly, both factor H and C4BP are capable of binding their respective targets while in convertase complexes as well. At the host cell surface, membrane cofactor protein (MCP or CD46) can interact with C4b and C3b, which protects the host cell from complement self-activation (12). Delay accelerating factor (DAF or CD55) also recognizes and dissociates both C3 and C5 convertases on host cell surfaces (13). More generally, Carboxypeptidase-N, a well known inflammation regulator, has broad activity against the complement proteins C3a, C4a, and C5a, rendering them inactive by cleavage of carboxyl-terminal arginine and lysine residues (14). Although Carboxypeptidase-N does not directly influence complement activation, it silences the important inflammatory signals produced by complement. Lastly, the assembly of the MAC complex itself can be inhibited by vitronectin and clusterin in the plasma, and CD59 at the host surface (15, 16). Thus, there are many points of control which influence complement across the three activation pathways.

Developing quantitative mathematical models of complement will be crucial to fully understanding its role in the body. Traditionally, complement models have been formulated as systems of linear or non-linear ordinary differential equations (ODEs). For example, Hirayama et al. modeled the classical complement pathway as a system of linear ODEs (17), while Korotaevskiy and co-workers modeled the classical, lectin and alternate pathways (up to C3a) as a system of non-linear ODEs (18). More recently, large mechanistic models of sections of complement have also been proposed. For example, Liu et al. an-

79 analyzed the formation of the classical and lectin C3 convertases, and the regulatory role
80 of C4BP using a system of 45 non-linear ODEs with 85 parameters (19). Zewde and
81 co-workers constructed a detailed mechanistic model of the alternative pathway which
82 consisted of 107 ODEs and 74 kinetic parameters and delineated between the host cell
83 and pathogen (16). However, these previous modeling studies involved little experimen-
84 tal validation. Thus, while these models are undoubtedly important theoretical tools, it
85 is unclear if they can describe or quantitatively predict complement measurements. The
86 central challenge of complement model identification is the estimation of model parame-
87 ters from experimental measurements. Unlike other important cascades, such as coagu-
88 lation where there are well developed experimental tools and publicly available data sets,
89 the data for complement is relatively sparse. Missing or incomplete data sets, and limited
90 quantitative data make the identification of large mechanistic complement models difficult.
91 Thus, reduced order modeling approaches, which describe the key biology with a limited
92 number of parameters, will be important for developing validated models of complement,
93 and to fully understandings its role in the body.

Results

In this study, we developed an ensemble of experimentally validated reduced order complement models. The modeling approach combined ordinary differential equations with logical rules to produce a complement model with a limited number of equations and parameters. The reduced order model, which described the lectin and alternative pathways, consisted of 18 differential equations with 28 parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable models in the literature. We estimated an ensemble of model parameters from *in vitro* time series measurements of the C3a and C5a complement proteins. Subsequently, we validated the model on unseen C3a and C5a measurements that were not used for model training. Despite its small size, the model was surprisingly predictive. After validation, we performed global sensitivity and robustness analysis to estimate which parameters and species controlled model performance. These analyses suggested complement was robust to any single therapeutic intervention; only the knockdown of both C3 and C5 consistently reduced C3a and C5a formation for all cases. Taken together, we developed a reduced order complement model that was computationally inexpensive, and could easily be incorporated into pre-existing or new pharmacokinetic models of immune system function. The model described experimental data, and predicted the need for multiple points of intervention to disrupt complement activation.

Reduced order complement network. The complement model described the alternate and lectin pathways (Fig. 1). A trigger event initiated the lectin pathway, which activated the cleavage of C2 and C4 into C2a, C2b, C4a and C4b respectively. Classical Pathway (CP) C3 convertase (C4aC2b) then catalyzed the cleavage of C3 into C3a and C3b. The alternate pathway was initiated through the spontaneous hydrolysis of C3 into C3a and C3b (not C3w). The C3b fragment generated by hydrolysis (or by CP C3 convertase) could then form the alternate pathway (AP) C3 convertase (C3bBb). We did not consider

C3w, nor the formation of the initial alternate C3 convertase (C3wBb). Rather, we assumed C3w was equivalent to C3b and only modeled the formation of the main AP C3 convertase. Both the CP and AP C3 convertases catalyzed the cleavage of C3 into C3a and C3b. A second C3b fragment could then bind with either the CP or AP C3 convertase to form the CP or AP C5 convertase (C4bC2aC3b or C3bBbC3b). Both C5 convertases catalyzed the cleavage of C5 into the C5a and C5b fragments. In this initial study, we simplified the model by assuming both Factor B and Factor D were in excess. However, we did explicitly account for two control proteins, Factor H and C4BP. Lastly, we did not consider MAC formation, instead we stopped at C5a and C5b. Lectin pathway activation, and C3/C5 convertase activity was modeled using a combination of saturation kinetics and non-linear transfer functions, which facilitated a significant reduction in the size of the model while maintaining performance. Binding interactions were modeled using mass-action kinetics, where we assumed all binding was irreversible. Thus, while the reduced order complement model encoded significant biology, it was highly compact consisting of only 18 differential equations and 28 model parameters. Next, we estimated an ensemble of model parameters from time series measurements of the C3a and C5a complement proteins.

Estimating an ensemble of reduced order complement models. A critical challenge for the development of any dynamic model is the estimation of model parameters. We estimated an ensemble of complement model parameters in a hierarchical fashion using *in vitro* time-series data sets generated with and without zymosan, a lectin pathway activator (20). The residual between model simulations and experimental measurements was minimized using the dynamic optimization with particle swarms (DOPS) routine, starting from an initial random parameter guess. Unless otherwise specified, all initial conditions were assumed to be their mean physiological values. A hierarchical approach was taken in which the alternate pathway parameters were estimated first and then fixed during the

estimation of the lectin pathway parameters. While we had significant training data, the parameter estimation problem was underdetermined (we were not able to uniquely determine model parameters). Thus, instead of using a best-fit yet uncertain parameter set, we estimated an ensemble of probable parameter sets ($N = 50$, see materials and methods). The reduced order complement model ensemble captured the behavior of both the alternative and lectin pathways (Fig. 2). For the alternative pathway, we used C3a and C5a measurements in the absence of zymosan, and only allowed the alternative parameters to vary (Fig. 2A and B). On the other hand, lectin pathway parameters were estimated from C3a and C5a measurements in the presence of 1g zymosan with alternate pathway parameters fixed (Fig. 2C and D). The reduced order model reproduced a panel of alternate and lectin pathway data sets in the neighborhood of physiological factor and inhibitor concentrations. However, it was unclear whether the reduced order model could predict new data, without updating the model parameters. To address this question, we fixed the model parameters and simulated data sets not used for model training.

We tested the predictive power of the reduced order complement model with data not used during model training (Fig. 3). Six validation cases were considered, three for C3a and C5a, respectively. All model parameters and initial conditions were fixed for the validation simulations (with the exception of zymosan, and other experimentally mandated changes). The ensemble of reduced order models predicted the qualitative dynamics of C3a formation (Fig. 3, top), and C5a formation (Fig. 3, bottom) at three inducer concentrations. The rate of C3a formation and C3a peak time were directly proportional to initiator dose. Similarly, the C5a plateau and rate of formation were also directly proportional to initiator dose, with the lag time being indirectly proportional to initiator exposure. However, there were shortcomings with model performance. First, while the overall C3a trend was captured (within the 99% confidence interval), the C3a dynamics were too fast with the exception of the low dose case. We believe the C3a timescale was related to

our choice of training data, how we modeled the tickover mechanism, and factor B and D limitation. We trained the model using either no or 1 mg/ml zymosan, but predicted cases in a different initiator range; comparing the training to prediction, the model performance e.g., the shape of the C3a trajectory was biased towards either high or very low initiator doses. Next, tickover was modeled as a first-order generation processes where C3wBb formation and activity was lumped into the AP C3 convertase. Thus, we skipped an important step which could strongly influence AP C3 convertase formation by slowing down the rate C3 cleavage into C3a and C3b. We also assumed both factor B and factor D were not limiting, thereby artificially accelerating the rate of AP C3 convertase formation. The C5a predictions followed a similar trend as C3a; we captured the long-time C5a behavior but over predicted the time scale of C5 cleavage. However, because the C5a time scale depends strongly upon C3 convertase formation, we can likely correct the C5 issues by fixing the rate of C3 cleavage. Despite these shortcomings, we qualitatively predicted unseen experimental measurements within the 99% confidence of the ensemble, for three inducer levels. Next, we used global sensitivity and robustness analysis to determine which parameters and species controlled the performance of the complement model.

Global analysis of the reduced order complement model. We conducted sensitivity analysis to estimate which parameters controlled the performance of the reduced order complement model. We calculated the total sensitivity of the C3a and C5a residual to changes in model parameters with and without zymosan (Fig. 4). In the absence of zymosan (where only the alternative pathway is active), the most sensitive parameter was the rate constant governing the assembly of the AP C3 convertase, as well as the rate constant controlling basal C3b formation. The C5a trajectory was sensitive to the AP C5 convertase kinetic parameters (Fig. 4A). Interestingly, neither the rate nor the saturation constants governing AP C3-convertase activity were sensitive in the absence of zymosan. Thus, C3a formation in the alternative pathway was more heavily influenced by

the spontaneous hydrolysis of C3, rather than AP C3-convertase activity, in the absence of zymosan. In the presence of zymosan, the C3a residual was controlled by the formation and activity of the CP C3 convertase, as well as tickover and degradation parameters. On the other hand, the C5a residual was controlled by the formation and activity of CP C5 convertase, and tickover C3b formation in the presence of zymosan (Fig. 4B). The lectin initiation parameters were sensitive, but to a lesser extent than CP convertase kinetic parameters and tickover C3b formation. Thus, sensitivity analysis suggested that CP C3/C5 convertase formation and activity dominated in the presence of zymosan, but tickover parameters and AP C5 convertase were more important without initiator. Next, we compared the sensitivity results to current therapeutic approaches; pathways involving sensitive parameters have been targeted for clinical intervention (Fig. 4C). In particular, the sensitivity analysis suggested AP/CP C5 convertase inhibitors, or interventions aimed at attenuating C3 or C5 would most strongly influence complement performance. Thus, there was at least a qualitative overlap between sensitivity and the potential of biochemical efficacy. However, sensitivity coefficients are only a local measure of how small changes in parameters affect model performance. To more closely simulate a clinical intervention e.g., administration of an anti-complement inhibitor, we performed robustness analysis.

Robustness coefficients quantify the response of a marker to a macroscopic structural or operational perturbation to a biochemical network. Here, we computed how the C3a and C5a trajectories responded to a decrease in the initial abundance of C3 and/or C5. Robustness analysis suggested there was no single intervention that inhibited complement activation in the presence of both initiation pathways (Fig. 5). We calculated robustness indices for C3a and C5a for the parameter ensemble ($N = 50$) with and without the lectin pathway initiator. We simulated the addition of different doses of anti-complement inhibitor cocktails by decreasing the initial concentration of C3 or C5 or the combination of C3 and C5 by 50% and 90%. This would be conceptually analogous to the administration

of a C3 inhibitor e.g., compstatin alone or combination with eculizumab (Fig. 4E). A \log_{10} transformed robustness index of zero indicated no effect due to the perturbation, whereas an index of less than zero indicated decreased C3a or C5a. The response of the complement model to different knock-down magnitudes was non-linear; a 90% knock-down had an order of magnitude more impact than a 50% knock-down. As expected, a C5 knock-down had no effect on C3a formation for either the alternate (Fig. 5A, lanes 1 or 3) or lectin pathways (Fig. 5B, lanes 1 or 3). However, C3a abundance and to a lesser extent C5a abundance decreased with decreasing C3 concentration in the alternate pathway (Fig. 5A or B, lanes 1 or 2). This agreed with the sensitivity results; changes in AP C3-convertase formation affected the downstream dynamics of C5a formation. Thus, if we only considered the alternate pathway, C3 alone could be a reasonable target, especially given that C5a formation was surprisingly robust to C5 levels in the alternate pathway (Fig. 5A or B, lane 2). Yet, when both pathways were activated, C5a levels were robust to the initial C3 concentration (Fig. 5A or B, lane 4); C5a formation was catalyzed by CP C3 and C5 convertases. Thus, the only reliable intervention that consistently reduced both C3a and C5a formation for all cases was a dual-knockdown of C3 and C5. For example, a 90% decrease of both C3 and C5 reduced the formation of C5a by over an order of magnitude (Fig. 5B, lane 4), while C3a was reduced to a lesser extent (Fig. 5B, lane 3).

Discussion

In this study, we developed an ensemble of experimentally validated reduced order complement models. The modeling approach combined ordinary differential equations with logical rules to produce a complement model with a limited number of equations and parameters. The reduced order model, which described the lectin and alternative pathways, consisted of 18 differential equations with 28 parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable mathematical models in the literature. We estimated an ensemble of model parameters from *in vitro* time series measurements of the C3a and C5a complement proteins. Subsequently, we validated the model on unseen C3a and C5a measurements that were not used for model training. Despite its small size, the model was surprisingly predictive. After validation, we performed global sensitivity and robustness analysis to estimate which parameters and species controlled model performance. These analyses suggested complement was robust to any single therapeutic intervention. The only intervention that consistently reduced C5a formation for all cases was a dual-knockdown of both C3 and C5. Taken together, we developed a reduced order complement model that was computationally inexpensive, and could easily be incorporated into pre-existing or new pharmacokinetic models of immune system function. The model described experimental data, and predicted the need for multiple points of intervention to disrupt complement activation.

Despite its importance, there has been a paucity of validated mathematical models of complement pathway activation. To our knowledge, this study is one of the first complement models that combined multiple initiation pathways with experimental validation of important complement products like C5a. However, there have been several theoretical models of components of the cascade in the literature. Liu and co-workers modeled the formation of C3a through the classical pathway using 45 non-linear ODEs (19). In contrast, in this study we modeled lectin mediated C3a formation using only five ODEs.

Though we did not model all the initiation interactions in detail, especially the cross-talk between the lectin and classical pathways, we successfully captured C3a dynamics with respect to different concentrations of lectin initiators. The model also captured the dynamics of C3a and C5a formed from the alternate pathway using only seven ODEs. The reduced order model predictions of C5a were qualitatively similar to the theoretical complement model of Zewde et al which involved over 100 ODEs (16). However, we found that the quantity of C3a produced in the alternate pathway was nearly 1000 times the quantity of C5a produced. Though this was in agreement with the experimental data (20), it differed from the theoretical predictions made by Zewde et al. who showed C3a was 10^8 times the C5a concentration (16). In our model, the time profile of C5a generation from the lectin pathway changed with respect to the quantity of zymosan (the lectin pathway initiator). The lag phase for generation was inversely proportional to the initiator concentration. Korotaevskiy et al. showed a similar trend using a theoretical model of complement, albeit for much shorter time scales (18). Thus, the reduced order complement model performed similarly to existing large mechanistic models, despite being significantly smaller.

Global analysis of the complement model estimated potential important therapeutic targets. Complement malfunctions are implicated in a number of diseases, however the development of complement specific therapeutics has been challenging (3, 21). Previously, we have shown that mathematical modeling and sensitivity analysis can be useful tools to estimate therapeutically important mechanisms in biochemical networks (22–25). In this study, we analyzed a validated ensemble of reduced order complement models to estimate therapeutically important mechanisms. In presence of an initiator, C5a formation was primarily sensitive to the lectin initiation parameters, and parameters governing the conversion of C5 to C5a and C5b. This result agrees well with the current protease inhibitors targeting initiating complexes, including mannose-associated serine proteases 1 and 2 (MASP-1,2) (26). The most commonly used anti-complement drug eculizumab

(21), targets the C5 protein which is cleaved to form C5a. Our sensitivity analysis showed that kinetic parameters governing C5 conversion were sensitive in both lectin initiated and alternate pathways, thus agreeing with targeting C5 protein. The formation of basal C3b was also a sensitive parameter in the formation of C3a through the alternate pathway. Thus, this mechanism can act as a target for both C3a and C5a inhibitors. Lectin initiated C3a formation showed a number of sensitive parameters. This included the lectin initiation parameters that controlled C5a formation, C3 convertase inhibition by C4BP, and parameters governing C3 convertase activity. All these mechanisms are potential drug targets.

To further validate these results from sensitivity analysis about potential drug targets we did a robustness analysis. We knocked down C3 and C5 levels and studied their impact on the generation of C3a and C5a. The C3a and C5a levels in the lectin pathway were strongly influenced by initial levels of C3 and C5. Thus direct inhibition of C3 and C5, or targeting complexes (MASP complex, C3 and C5 convertases) that act on C3 and C5 have a direct impact on production of C3a and C5a. This is also in agreement with sensitivity analysis that C5 is a good drug target. A number of drugs targeting C5 are being developed. For example LFG316 by Novartis is being used to target C5 in cases of Age-Related Macular Degeneration (27), Mubodina is an antibody that targets C5 in the treatment of Atypical Hemolytic-Uremic Syndrome (aHUS) (28), Coversin is a small molecule targeting C5 (29), Zimura is an aptamer targeting C5 (30), small peptides and RNAi are also being used to inhibit C5 (31). Another important conclusion that can be drawn together from sensitivity and robustness analysis is that C3 and C5 convertases can be important therapeutic targets. Though knockdown of C3 and C5 affects C3a and C5a levels downstream, the abundance and turnover rate (32, 33) of these proteins make them difficult targets. Thus targeting C3 and C5 directly will require high dosage of drugs. It is also well known that eculizumab dosage needs to be adjusted while treating for Atyp-

320 ical Hemolytic-Uremic Syndrome (aHUS), a disease that is caused due to uncontrolled
321 complement activation (34). The issue of high dosage can potentially be circumvented
322 by targeting convertases or fragile mechanisms that involve C3, C5 or their activated
323 components. Our analysis shows that formation and assembly of these convertases are
324 sensitive mechanisms that strongly impact downstream proteins like C5a. Formation of
325 convertases is inhibited by targeting upstream protease complexes like MASP-1,2 from
326 lectin pathway (or C1r, C1s from classical pathway). For example, Omeros is a protease
327 inhibitor that targets MASP-2 complex and thereby inhibits formation of downstream con-
328 vertases (35). Lampalizumab (an immunoglobulin) and Bikacimab (an antibody frag-
329 ment) target Factor B and Factor D respectively. Factor B and Factor D are crucial to
330 formation alternate pathway convertases (36, 37). Novemed Therapeutics recently de-
331 veloped antibody, NM9401 against propeptin, a small protein that stabilizes alternate C3
332 convertase (38). Cobra Venom Factor (CVF), an analogue of C3b has been used to bind
333 to Factor B to regulate alternate convertases (39). Thus, analysis of the ensemble of com-
334 plement models identified potentially important therapeutic targets that are consistent with
335 therapeutic strategies that are under development.

336 The performance of the complement model was impressive given its limited size. How-
337 ever, there are several questions that should be explored further. A logical progression for
338 this work would be to expand the network to include the classical pathway and the forma-
339 tion of the membrane attack complex (MAC). However, it is unclear whether the addition
340 of the classical pathway will decrease the predictive quality of our existing model. Liu
341 et al have shown cross-talk between the activation of the classical and lectin pathways
342 that could influence model performance (19). One potential approach to address such
343 difficulties would be to incorporate C reactive proteins (CRP) and L-ficolin (LF) into the
344 model, both of which are involved with the initiation of classical and lectin pathways. Liu
345 et al. showed that under inflammation conditions interactions between lectin and classical

pathways was mediated through CRP and LF (19). Thus incorporating these two proteins would help us in modeling cross talk. Time course measurements of MAC abundance (and MAC formation dynamics) are also scarce, making the inclusion of MAC challenging. Next, we should address the under-prediction of C3a. We believe the C3a under-prediction can be attributed to how we modeled C4BP interactions. C4BP interactions were modeled as irreversible binding steps resulting in completely inactive complexes; however, the binding of C4BP with complement proteins is likely reversible and C4BP-bound convertases may have residual activity. We also did not capture the maximum concentration of C3a at low initiator levels. One possible reason for this could be the C2-by-pass pathway, which was not included in the model. This pathway further accelerates C3a production without the involvement of a C3 convertase. Currently the C3a in the model is generated only through the activity of a C3 convertase. Incorporating this additional step within the reduced order modeling framework would be a future direction that we need to consider. We should test alternative model structures which include reversible C4BP binding, and partially active convertases. Alternatively, we could also perform sensitivity analysis on the C3a prediction residual to determine which parameters controlled the C3a prediction.

Materials and Methods

Formulation and solution of the complement model equations. We used ordinary differential equations (ODEs) to model the time evolution of complement proteins (x_i) in the reduced order model:

$$\frac{1}{\tau_i} \frac{dx_i}{dt} = \sum_{j=1}^{\mathcal{R}} \sigma_{ij} r_j(\mathbf{x}, \epsilon, \mathbf{k}) \quad i = 1, 2, \dots, \mathcal{M} \quad (1)$$

where \mathcal{R} denotes the number of reactions and \mathcal{M} denotes the number of protein species in the model. The quantity τ_i denotes a timescale parameter for species i which captures unmodeled effects; by default $\tau_i = 1$, for all species unless otherwise specified. The quantity $r_j(\mathbf{x}, \epsilon, \mathbf{k})$ denotes the rate of reaction j . Typically, reaction j is a non-linear function of biochemical and enzyme species abundance, as well as unknown model parameters \mathbf{k} ($\mathcal{K} \times 1$). The quantity σ_{ij} denotes the stoichiometric coefficient for species i in reaction j . If $\sigma_{ij} > 0$, species i is produced by reaction j . Conversely, if $\sigma_{ij} < 0$, species i is consumed by reaction j , while $\sigma_{ij} = 0$ indicates species i is not connected with reaction j . Species balances were subject to the initial conditions $\mathbf{x}(t_o) = \mathbf{x}_o$.

Rate processes were written as the product of a kinetic term (\bar{r}_j) and a control term (v_j) in the complement model. The kinetic term for the formation of C4a, C4b, C2a and C2b, lectin pathway activation, and C3 and C5 convertase activity was given by:

$$\bar{r}_j = k_j^{max} \epsilon_i \left(\frac{x_s^\eta}{K_{js}^\eta + x_s^\eta} \right) \quad (2)$$

where k_j^{max} denotes the maximum rate for reaction j , ϵ_i denotes the abundance of the enzyme catalyzing reaction j , η denotes a cooperativity parameter, and K_{js} denotes the saturation constant for species s in reaction j . We used mass action kinetics to model

382 protein-protein binding interactions within the network:

$$\bar{r}_j = k_j^{max} \prod_{s \in m_j^-} x_s^{-\sigma_{sj}} \quad (3)$$

383 where k_j^{max} denotes the maximum rate for reaction j , σ_{sj} denotes the stoichiometric coef-
 384 ficient for species s in reaction j , and $s \in m_j^-$ denotes the set of *reactants* for reaction j .
 385 We assumed all binding interactions were irreversible.

386 The control terms $0 \leq v_j \leq 1$ depended upon the combination of factors which in-
 387 fluenced rate process j . For each rate, we used a rule-based approach to select from
 388 competing control factors. If rate j was influenced by $1, \dots, m$ factors, we modeled this re-
 389 lationship as $v_j = \mathcal{I}_j(f_{1j}(\cdot), \dots, f_{mj}(\cdot))$ where $0 \leq f_{ij}(\cdot) \leq 1$ denotes a regulatory transfer
 390 function quantifying the influence of factor i on rate j . The function $\mathcal{I}_j(\cdot)$ is an integration
 391 rule which maps the output of regulatory transfer functions into a control variable. Each
 392 regulatory transfer function took the form:

$$f_{ij}(Z_i, k_{ij}, \eta_{ij}) = k_{ij}^{\eta_{ij}} Z_i^{\eta_{ij}} / (1 + k_{ij}^{\eta_{ij}} Z_i^{\eta_{ij}}) \quad (4)$$

393 where Z_i denotes the abundance of factor i , k_{ij} denotes a gain parameter, and η_{ij} denotes
 394 a cooperativity parameter. In this study, we used $\mathcal{I}_j \in \{min, max\}$ (40). If a process has
 395 no modifying factors, $v_j = 1$. The model equations were implemented in MATLAB and
 396 solved using the ODE23s routine (The Mathworks, Natick MA). The complement model
 397 code and parameter ensemble is freely available under an MIT software license and can
 398 be downloaded from <http://www.varnerlab.org>.

399 **Estimation of an ensemble of complement model parameters.** We minimized the
 400 residual between simulations and experimental C3a and C5a measurements using Dy-

401 namic Optimization with Particle Swarms (DOPS). DOPS minimized the objective:

$$\min_{\mathbf{k}} \sum_{\tau=1}^{\tau} \sum_{j=1}^S \left(\frac{\hat{x}_j(\tau) - x_j(\tau, \mathbf{k})}{\omega_j(\tau)} \right)^2 \quad (5)$$

402 where $\hat{x}_j(\tau)$ denotes the measured value of species j at time τ , $x_j(\tau, \mathbf{k})$ denotes the sim-
 403 ulated value for species j at time τ , and $\omega_j(\tau)$ denotes the experimental measurement
 404 variance for species j at time τ . The outer summation is with respect to time, while the
 405 inner summation is with respect to state. DOPS is a novel metaheuristic that combines
 406 multi swarm particle swarm optimization (PSO) with a greedy global optimization algo-
 407 rithm called dynamically dimensioned search (DDS). DOPS is faster than conventional
 408 global optimizers and has the ability to find near optimal solutions for high dimensional
 409 systems within a relatively few function evaluations. It uses an adaptive switching strat-
 410 egy based on error convergence rates to switch from the particle swarm to DDS search
 411 phases. This enables DOPS to quickly estimate globally optimal or near optimal solutions
 412 even in the presence of many local minima. In the swarm search, for each iteration the
 413 particles compute error within each sub-swarm by evaluating the model equations using
 414 their specific parameter vector realization. From each of these points within a sub-swarm
 415 a local best is identified. This along with the particle best within the sub-swarm S_k is used
 416 to update the parameter estimate for each particle using the following rules:

$$z_{i,j} = \theta_{1,j-1} z_{i,j-1} + \theta_2 r_1 (\mathcal{L}_i - z_{i,j-1}) + \theta_3 r_2 (\mathcal{G}_k - z_{i,j-1}) \quad (6)$$

417 where $z_{i,j}$ is the parameter vector, $(\theta_1, \theta_2, \theta_3)$ were adjustable parameters, \mathcal{L}_i denotes the
 418 best solution found by particle i within sub-swarm S_k for function evaluations $1 \rightarrow j-1$, and
 419 \mathcal{G}_k denotes the best solution found over all particles within sub-swarm S_k . The quantities
 420 r_1 and r_2 denote uniform random vectors with the same dimension as the number of

unknown model parameters ($\mathcal{K} \times 1$). At the conclusion of the swarm phase, the overall best particle, \mathcal{G}_k , over the k sub-swarms was used to initialize the DDS phase. For the DDS phase, the best parameter estimate was updated using the rule:

$$\mathcal{G}_{new}(J) = \begin{cases} \mathcal{G}(\mathbf{J}) + \mathbf{r}_{normal}(\mathbf{J})\sigma(\mathbf{J}), & \text{if } \mathcal{G}_{new}(\mathbf{J}) < \mathcal{G}(\mathbf{J}). \\ \mathcal{G}(\mathbf{J}), & \text{otherwise.} \end{cases} \quad (7)$$

where \mathbf{J} is a vector representing the subset of dimensions that are being perturbed, \mathbf{r}_{normal} denotes a normal random vector of the same dimensions as \mathcal{G} , and σ denotes the perturbation amplitude:

$$\sigma = R(\mathbf{p}^U - \mathbf{p}^L) \quad (8)$$

where R is the scalar perturbation size parameter, \mathbf{p}^U and \mathbf{p}^L are $(\mathcal{K} \times 1)$ vectors that represent the maximum and minimum bounds on each dimension. The set \mathbf{J} was constructed using a monotonically decreasing probability function \mathcal{P}_i that represents a threshold for determining whether a specific dimension j was perturbed or not. DDS updates are greedy; \mathcal{G}_{new} becomes the new solution vector only if it is better than \mathcal{G} . At the end of DDS phase we obtain the optimal vector \mathcal{G} which we use for plotting best fits against the experimental data, and for generating a parameter ensemble.

An ensemble of parameters was obtained by randomly perturbing the optimal parameter set within bounds established by analyzing repeated DOPS runs. Parameters were selected using goodness of fit (model residual) and numerical stability criteria. We selected an ensemble of $N = 50$ parameter sets for this study using this sampling procedure. The best fit parameter set, and parameter ensemble is available in a tab delimited plain text format from the Complement model GitHub repository. The DOPS routine was implemented in MATLAB (The Mathworks, Natick MA) and can be downloaded from <http://www.varnerlab.org> under an MIT software license.

Sensitivity and robustness analysis of complement model performance. We conducted global sensitivity and robustness analysis to estimate which parameters and species controlled the performance of the reduced order model. We computed the total variance-based sensitivity index of each parameter relative to the training residual for the C3a alternate, C5a alternate, C3a lectin, and C5a lectin cases using the Sobol method (41). The sampling bounds for each parameter were established from the minimum and maximum value for that parameter in the parameter ensemble. We used the sampling method of Saltelli *et al.* to compute a family of $N(2d + 2)$ parameter sets which obeyed our parameter ranges, where N was the number of trials per parameters, and d was the number of parameters in the model (42). In our case, $N = 200$ and $d = 28$, so the total sensitivity indices were computed using 11,600 model evaluations. The variance-based sensitivity analysis was conducted using the SALib module encoded in the Python programming language (43).

Robustness coefficients quantify the response of a marker to a structural or operational perturbation to the network architecture. Robustness coefficients were calculated as shown previously (44). Log-transformed robustness coefficients denoted by $\hat{\alpha}(i, j, t_o, t_f)$ are defined as:

$$\hat{\alpha}(i, j, t_o, t_f) = \log_{10} \left[\left(\int_{t_o}^{t_f} x_i(t) dt \right)^{-1} \left(\int_{t_o}^{t_f} x_i^{(j)}(t) dt \right) \right] \quad (9)$$

Here t_o and t_f denote the initial and final simulation time, while i and j denote the indices for the marker and the perturbation, respectively. A value of $\hat{\alpha}(i, j, t_o, t_f) > 0$, indicates increased marker abundance, while $\hat{\alpha}(i, j, t_o, t_f) < 0$ indicates decreased marker abundance following perturbation j . If $\hat{\alpha}(i, j, t_o, t_f) \sim 0$, perturbation j did not influence the abundance of marker i . In this study, we perturbed the initial condition of C3 or C5 or a combination of C3 and C5 by 50% or 90% and measured the area under the curve

465 (AUC) of C3a or C5a with and without lectin initiator. Log-transformed robustness coeffi-
466 cients were calculated for every member of the ensemble, where the mean $\pm 1 \times$ standard-
467 deviation are reported.

Competing interests

The authors declare that they have no competing interests.

Author's contributions

J.V directed the study. A.S developed the reduced order complement model and the parameter ensemble. A.S, W.D and M.M analyzed the model ensemble, and generated figures for the manuscript. The manuscript was prepared and edited for publication by A.S, W.D, M.M and J.V.

Acknowledgements

We gratefully acknowledge the suggestions from the anonymous reviewers to improve this manuscript.

Funding

This study was supported by an award from the US Army and Systems Biology of Trauma Induced Coagulopathy (W911NF-10-1-0376) to J.V. for the support of A.S.

References

1. Nuttall G (1888) Experimente über die bacterienfeindlichen Einflüsse des thierischen Körpers. Z Hyg Infektionskr 4: 353-394.
2. Ricklin D, Hajishengallis G, Yang K, Lambris JD (2010) Complement: a key system for immune surveillance and homeostasis. NAT IMMUNOL 11: 785–797.
3. Ricklin D, Lambris JD (2007) Complement-targeted therapeutics. NAT BIOTECHNOL 25: 1265-1275.
4. Rittirsch D, Flierl MA, Ward PA (2008) Harmful molecular mechanisms in sepsis. NAT REV IMMUNOL 8: 776–787.
5. Sarma JV, Ward PA (2011) The complement system. Cell Tissue Res 343: 227–235.
6. Ricklin D, Lambris JD (2013) Complement in immune and inflammatory disorders: pathophysiological mechanisms. J Immunol 190: 3831–3838.
7. Walport MJ (2001) Complement. first of two parts. N Engl J Med 344: 1058-66.
8. Walport MJ (2001) Complement. second of two parts. N Engl J Med 344: 1140-4.
9. Pangburn MK, Müller-Eberhard HJ (1984) The alternative pathway of complement. Semin Immunopathol .
10. Walker D, Yasuhara O, Patston P, McGeer E, McGeer P (1995) Complement c1 inhibitor is produced by brain tissue and is cleaved in alzheimer disease. Brain Res 675: 75–82.
11. Blom AM, Kask L, Dahlbäck B (2001) Structural requirements for the complement regulatory activities of c4bp. J Biol Chem 276: 27136–27144.
12. Riley-Vargas RC, Gill DB, Kemper C, Liszewski MK, Atkinson JP (2004) Cd46: expanding beyond complement regulation. Trends Immunol 25: 496–503.
13. Lukacik P, Roversi P, White J, Esser D, Smith G, et al. (2004) Complement regulation at the molecular level: the structure of decay-accelerating factor. Proc Natl Acad Sci USA 101: 1279–1284.

14. Liszewski MK, Farries TC, Lublin DM, Rooney IA, Atkinson JP (1995) Control of the complement system. *Adv Immunol* 61: 201–283.
15. Chauhan A, Moore T (2006) Presence of plasma complement regulatory proteins clusterin (apo j) and vitronectin (s40) on circulating immune complexes (cic). *Clin Exp Immunol* 145: 398–406.
16. Zewde N, Gorham Jr RD, Dorado A, Morikis D (2016) Quantitative modeling of the alternative pathway of the complement system. *PloS one* 11: e0152337.
17. Hirayama H, Yoshii K, Ojima H, Kawai N, Gotoh S, et al. (1996) Linear systems analysis of activating processes of complement system as a defense mechanism. *Biosystems* 39: 173–185.
18. Korotaevskiy AA, Hanin LG, Khanin MA (2009) Non-linear dynamics of the complement system activation. *Math Biosci* 222: 127–143.
19. Liu B, Zhang J, Tan PY, Hsu D, Blom AM, et al. (2011) A computational and experimental study of the regulatory mechanisms of the complement system. *PLoS Comput Biol* 7: e1001059.
20. Morad HO, Belete SC, Read T, Shaw AM (2015) Time-course analysis of c3a and c5a quantifies the coupling between the upper and terminal complement pathways in vitro. *J Immunol Methods* 427: 13–18.
21. Morgan BP, Harris CL (2015) Complement, a target for therapy in inflammatory and degenerative diseases. *NAT REV DRUG DISCOV* 14: 857-877.
22. Luan D, Zai M, Varner JD (2007) Computationally derived points of fragility of a human cascade are consistent with current therapeutic strategies. *PLoS Comput Biol* 3: e142.
23. Nayak S, Salim S, Luan D, Zai M, Varner JD (2008) A test of highly optimized tolerance reveals fragile cell-cycle mechanisms are molecular targets in clinical cancer trials. *PLoS One* 3: e2016.

24. Tasseff R, Nayak S, Salim S, Kaushik P, Rizvi N, et al. (2010) Analysis of the molecular networks in androgen dependent and independent prostate cancer revealed fragile and robust subsystems. PLoS One 5: e8864.
25. Rice NT, Szlam F, Varner JD, Bernstein PS, Szlam AD, et al. (2016) Differential contributions of intrinsic and extrinsic pathways to thrombin generation in adult, maternal and cord plasma samples. PLoS One 11: e0154127.
26. Héja D, Harmat V, Fodor K, Wilmanns M, Dobó J, et al. (2012) Monospecific inhibitors show that both mannan-binding lectin-associated serine protease-1 (masp-1) and-2 are essential for lectin pathway activation and reveal structural plasticity of masp-2. J Biol Chem : 20290–20300.
27. Roguska M, Splawski I, Diefenbach-Streiber B, Dolan E, Etemad-Gilbertson B, et al. (2014) Generation and characterization of lfg316, a fully-human anti-c5 antibody for the treatment of age-related macular degeneration. Investigative Ophthalmology & Visual Science : 3433–3433.
28. Melis JP, Strumane K, Ruuls SR, Beurskens FJ, Schuurman J, et al. (2015) Complement in therapy and disease: Regulating the complement system with antibody-based therapeutics. Molecular immunology : 117–130.
29. Weston-Davies WH, Nunn MA, Pinto FO, Mackie IJ, Richards SJ, et al. (2014) Clinical and immunological characterisation of coversin, a novel small protein inhibitor of complement c5 with potential as a therapeutic agent in pnh and other complement mediated disorders. Blood : 4280–4280.
30. Epstein D, Kurz JC (2007). Complement binding aptamers and anti-c5 agents useful in the treatment of ocular disorders. US Patent App. 12/224,708.
31. Borodovsky A, Yucius K, Sprague A, Banda NK, Holers VM, et al. (2014) Aln-cc5, an investigational rna therapeutic targeting c5 for complement inhibition. Complement 20: 40.

32. Sissons J, Liebowitch J, Amos N, Peters D (1977) Metabolism of the fifth component of complement, and its relation to metabolism of the third component, in patients with complement activation. *J Clin Invest* 59: 704.
33. Swaak A, Hannema A, Vogelaar C, Boom F, van Es L, et al. (1982) Determination of the half-life of c3 in patients and its relation to the presence of c3-breakdown products and/or circulating immune complexes. *Rheumatol Int* : 161–166.
34. Noris M, Galbusera M, Gastoldi S, Macor P, Banterla F, et al. (2014) Dynamics of complement activation in ahus and how to monitor eculizumab therapy. *Blood* : 1715–1726.
35. Schwaeble HW, Stover CM, Tedford CE, Parent JB, Fujita T (2011). Methods for treating conditions associated with masp-2 dependent complement activation. US Patent 7,919,094.
36. Katschke KJ, Wu P, Ganesan R, Kelley RF, Mathieu MA, et al. (2012) Inhibiting alternative pathway complement activation by targeting the factor d exosite. *Journal of Biological Chemistry* : 12886–12892.
37. Hu X, Holers VM, Thurman JM, Schoeb TR, Ramos TN, et al. (2013) Therapeutic inhibition of the alternative complement pathway attenuates chronic eae. *Mol Immunol* : 302–308.
38. Bansal R (2014). Humanized and chimeric anti-properdin antibodies. US Patent 8,664,362.
39. Vogel CW, Fritzinger DC, Hew BE, Thorne M, Bammert H (2004) Recombinant cobra venom factor. *Molecular immunology* : 191–199.
40. Sagar A, Varner JD (2015) Dynamic modeling of the human coagulation cascade using reduced order effective kinetic models. *Processes* 3: 178.
41. Sobol I (2001) Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *MATH COMPUT SIMULAT* 55: 271 - 280.

- 585 42. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, et al. (2010) Variance based
586 sensitivity analysis of model output. design and estimator for the total sensitivity index.
587 Comput Phys Commun 181: 259–270.
- 588 43. Herman J. Salib: Sensitivity analysis library in python (numpy). con-
589 tains sobol, morris, fractional factorial and fast methods. available online:
590 <https://github.com/jdherman/salib>.
- 591 44. Tasseff R, Nayak S, Song SO, Yen A, Varner JD (2011) Modeling and analysis
592 of retinoic acid induced differentiation of uncommitted precursor cells. Integr Biol
593 (Camb) 3: 578-91.

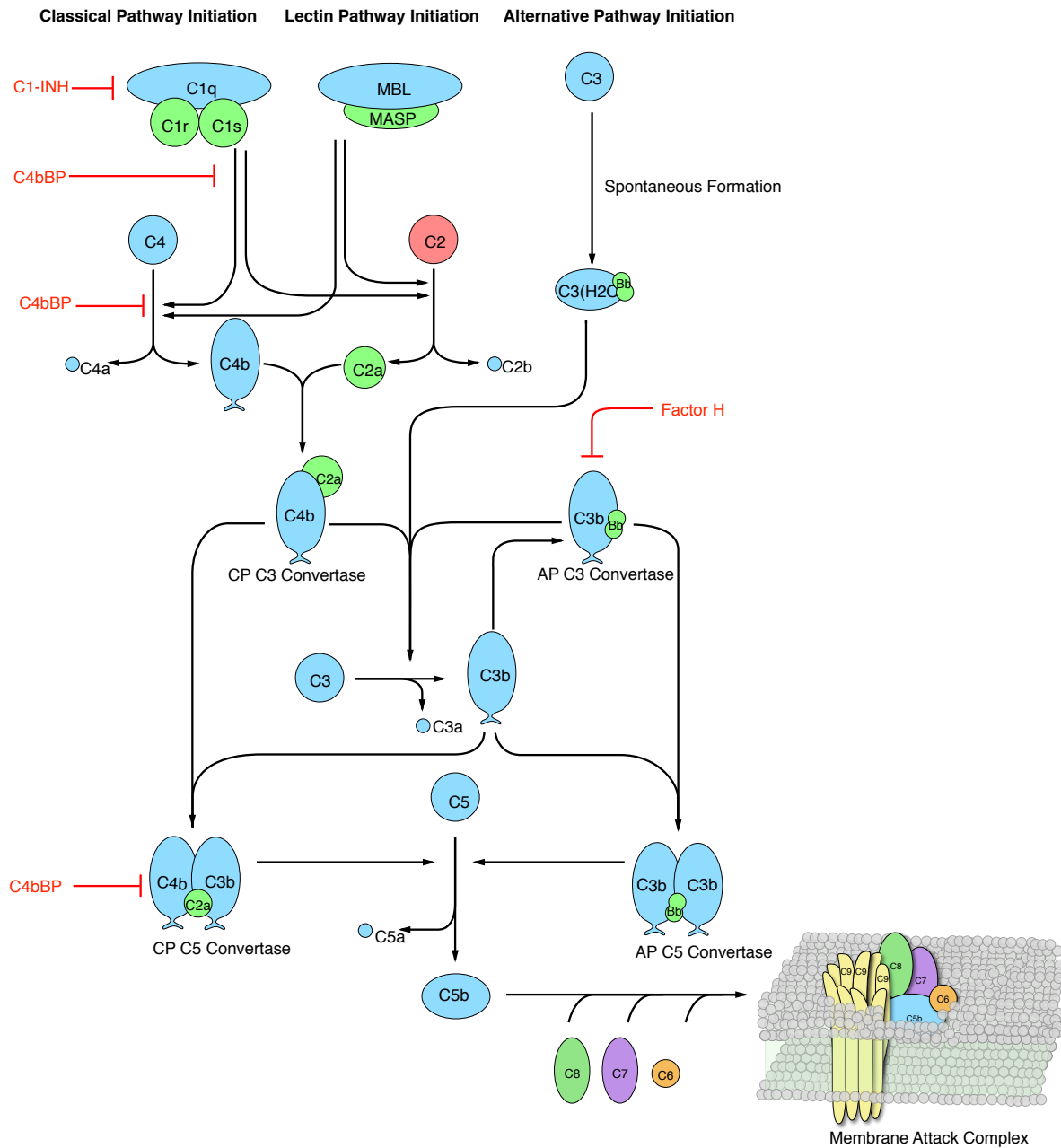


Fig. 1: Simplified schematic of the human complement system. The complement cascade is activated through three pathways: the classical, the lectin, and the alternate pathways. Complement initiation results in the formation of classical or alternative C3 convertases, which amplify the initial complement response and signal to the adaptive immune system by cleaving C3 into C3a and C3b. C3 convertases further react to form C5 convertases which catalyze the cleavage of the C5 complement protein to C5a and C5b. C5b is critical to the formation of the membrane attack complex (MAC), while C5a recruits an adaptive immune response.

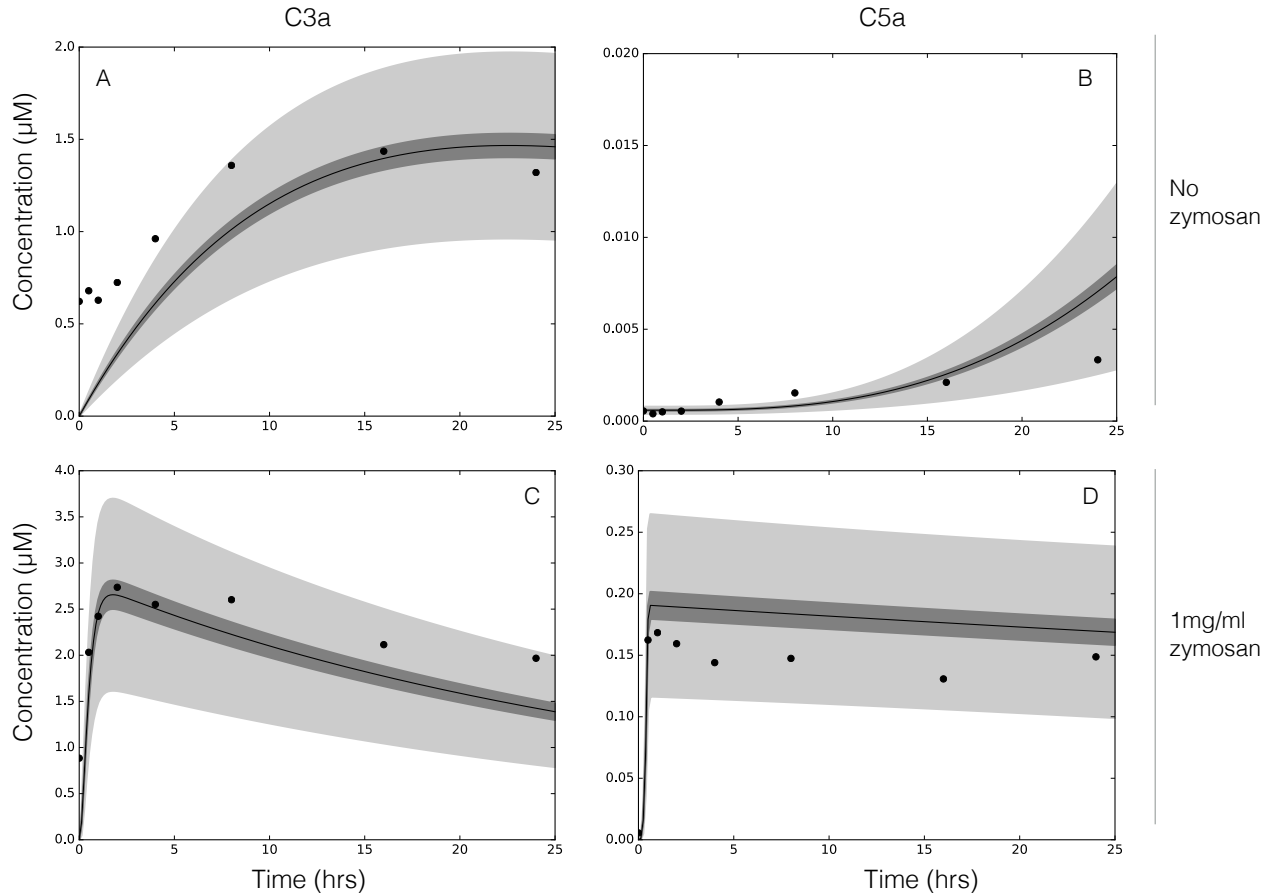


Fig. 2: Reduced order complement model training. Model parameters were estimated using Dynamic Optimization with Particle Swarms (DOPS) from C3a and C5a measurements with and without zymosan (20). The model was trained using C3a and C5a data generated from the alternative pathway (**A–B**) and lectin pathway initiated with 1g zymosan (**C–D**). The solid red line shows the simulation with the best-fit parameter set, the solid black lines show the simulated mean value of C3a or C5a for the ensemble ($N = 50$). The dark shaded region denotes the 99% confidence interval of the simulated mean concentrations, while the light shaded region denotes the 99% confidence interval of the best-fit simulation for C3a and C5a. All initial conditions were assumed to be at their physiological serum levels unless otherwise noted.

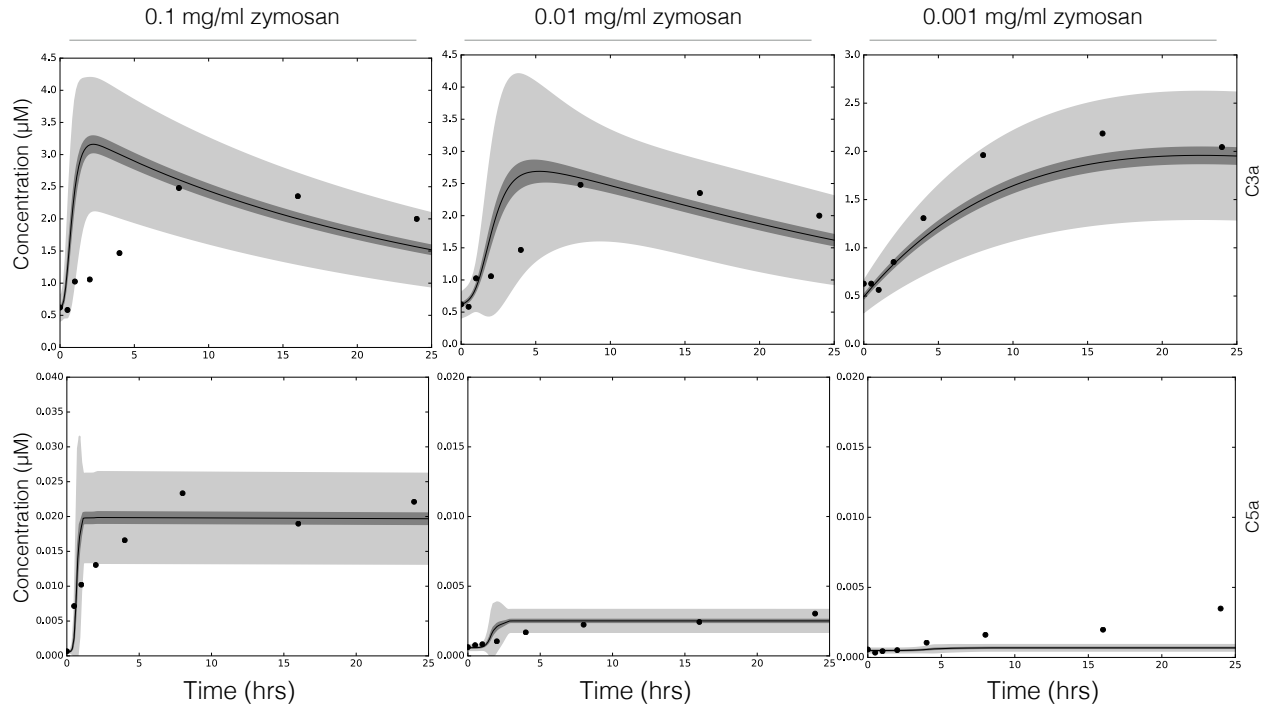


Fig. 3: Reduced order complement model predictions. The reduced order coagulation model parameter estimates were tested against data not used during model training (20). Simulations of C3a and C5a generated in the lectin pathway using 0.1g, 0.01g, and 0.001g zymosan were compared with the corresponding experimental measurement. The solid red line shows the simulation with the best-fit parameter set, the solid black lines show the simulated mean value of C3a or C5a for the ensemble ($N = 50$). The dark shaded region denotes the 99% confidence interval of the simulated mean concentrations, while the light shaded region denotes the 99% confidence interval of the best-fit simulation for C3a and C5a. All initial conditions were assumed to be at their physiological serum levels unless otherwise noted.

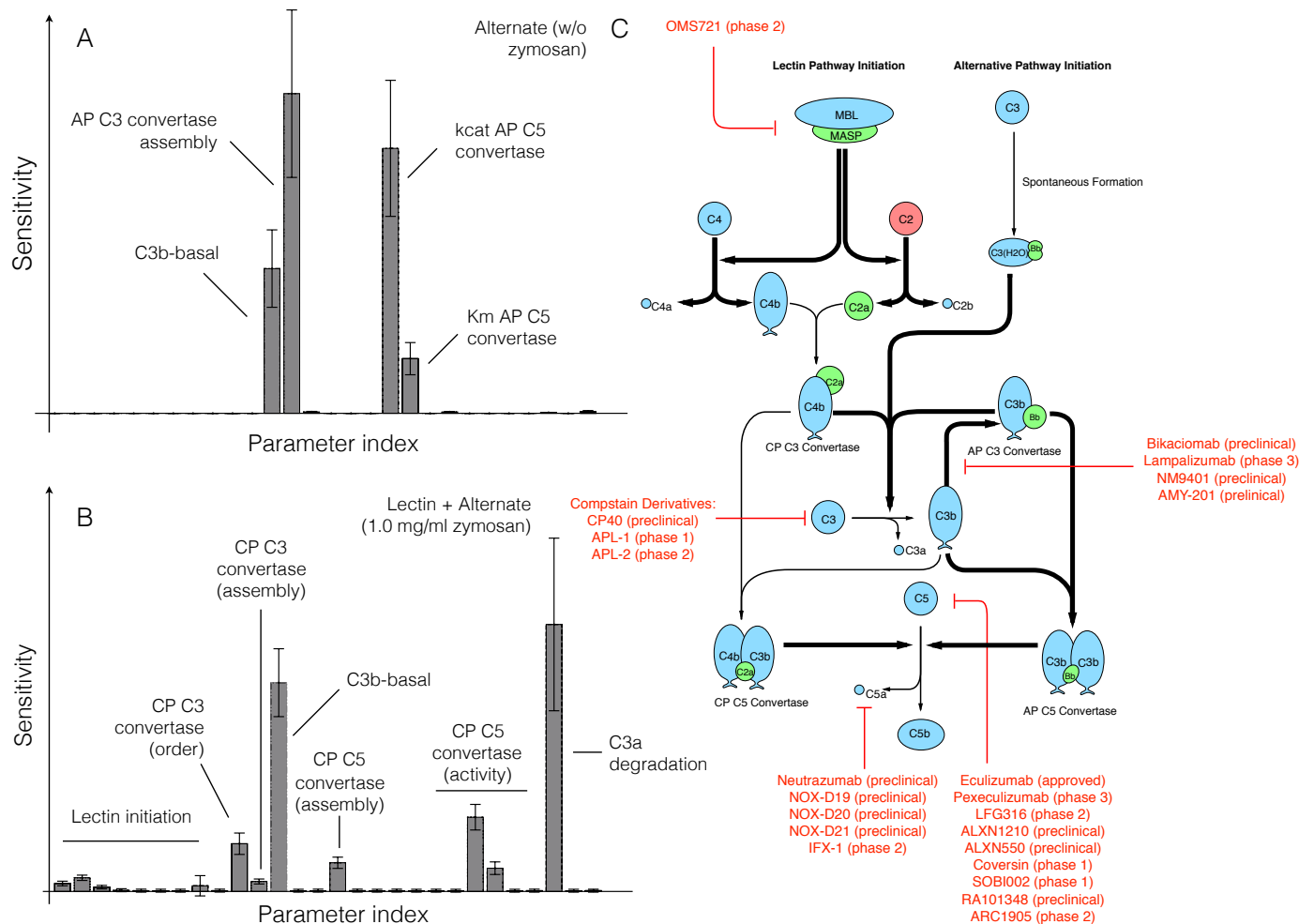
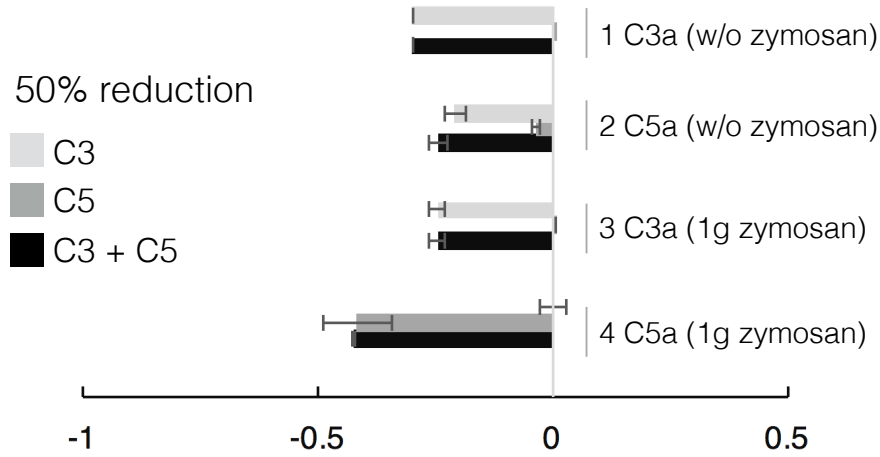


Fig. 4: Global sensitivity analysis of the reduced order complement model. Sensitivity analysis was conducted on the two objectives used for model training. **A:** Sensitivity of the C3a and C5a residual w/o zymosan, **B:** Sensitivity of the C3a and C5a residual with 1mg/ml zymosan. The bars denote the mean total sensitivity index for each parameter, while the error bars denote the 95% confidence interval. **C:** Pathways controlled by the sensitivity parameters. Bold black lines indicate the pathway involves one or more sensitive parameters, while the red lines show current therapeutics targets. Current complement therapeutics were taken from the review of Morgan and Harris (21).

A



B

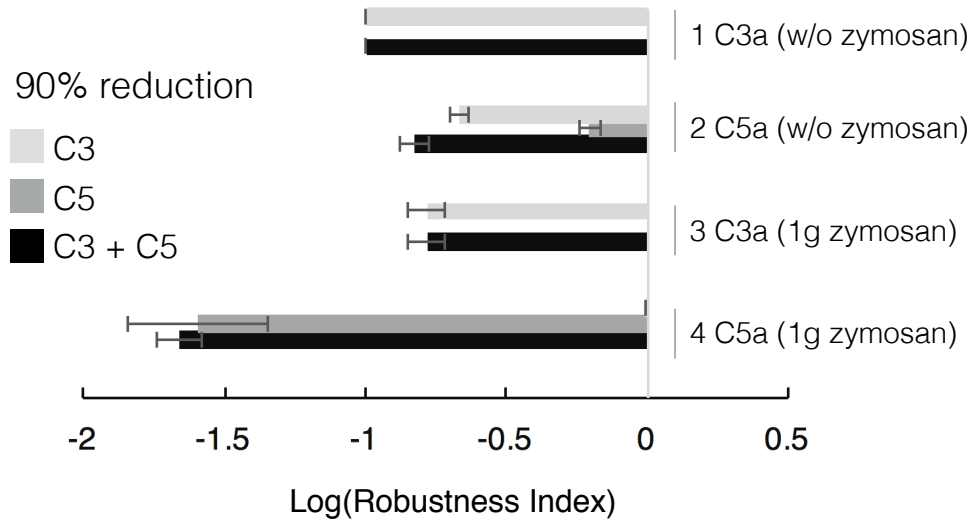


Fig. 5: Robustness analysis of the complement model with respect to C3 and C5 initial conditions. Robustness analysis was conducted on the four cases used for model training: C3a alternate (w/o zymosan), C5a alternate (w/o zymosan), C3a lectin (1g zymosan), and C5a lectin (1g zymosan), by reducing the initial concentration of C3 and/or C5 by 50% or 90 %. **A:** Robustness results for a 50% decrease in the C3, C5, or C3 and C5 initial condition. **B:** Robustness results for a 90% decrease in the C3, C5, or C3 and C5 initial condition. The bars denote the log-transformed robustness index while error bars denote one standard deviation. At zero, the perturbed initial concentration has no impact on the measured output. A log-transformed robustness index less than zero indicates a negative relation between the perturbed initial concentration and the measured output.