Reduced order modeling and analysis of the human complement subsystem

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

School of Chemical and Biomolecular Engineering

Cornell University, Ithaca NY 14853

Running Title: Reduced order model of complement

To be submitted: *PLoS ONE*

[#] 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 a central part of innate immunity and plays a significant role in regulating the inflammatory response. In this study, we build a reduced order model of complement to study the human complement system. The key novelty of our approach is the use of Ordinary Differential Equations (ODEs) along with logical rules to capture the behavior of a complex biochemical network. Using this framework we constructed a model of complement that analyzed the dynamics of C3a and C5a when initiated through the alternate and lectin pathways. The reduced order model consisted of only 18 differential equations with 28 kinetic and control parameters. Thus, the model was an order of magnitude smaller that any existing model of complement that includes alternate and lectin pathways.

Keywords: Biochemical engineering, systems biology, reduced order models, complement system

Introduction

Complement is a central part of innate immunity and plays a very significant role in regulating the inflammatory response. Complement was first discovered in the 1890s where it was found to 'complement' the bactericidal activity of natural antibodies. Complement is mediated through a set of approximately 30-35 soluble and cell surface proteases [1]. The central process in complement activation involves the formation of Membrane Attack Complex (MAC) and a protein called C5a. C5a acts as a bridge between innate and adaptive immunity and plays a very important role in regulating inflammation and coagulation [2]. Complement activation takes places through three different pathways: the alternate, the classical and the lectin. Each of these pathways involves a different initiator signal that leads to a cascade of downstream reactions in the complement system. The classical pathway is triggered when antibodies form complexes with foreign antigens or other pathogens. A multimeric protein complex C1 binds to the antigen-antibody complex and undergoes a conformational change. This activated complex cleaves proteins C4 and C2 to C4a, C4b, C2a and C2b respectively. C4a and C2b combine to form a protease C4bC2a also known as the classical C3 convertase. The lectin pathway is initiated through the binding of L-ficolin or Mannose Binding Lectin (MBL) to the carbohydrates on the surfaces of bacterial pathogens. This bound complex in turn cleaves C4 and C2 and leads to the formation of C4bC2a. The alternate pathway involves a 'tickover' mechanism in which C3 is hydrolyzed to form C3b. In presence of foreign pathogens C3b binds to these surfaces and recruits additional factors called factor B and factor D that lead to the formation of alternate C3 convertase - C3bBb [3]. The formation of classical and alternate C3 convertases on bacterial surfaces is followed by the formation of proteases called C5 convertases. The classical and alternate C3 convertases recruit C3, Factor B and Factor D to form classical C5 convertase (C4bC2aC3b) and alternate C5 convertase (C3bBbc3B) respectively. The C5 convertases then cleave C5 to form C5a and C5b re-

spectively. The cleavage of C5 is followed by a series of sequential cleavages of proteins C6, C7, C8 and C9 that combine with C5b to form the membrane attack complex (MAC) [4].

The activation of complement and formation of C5a and MAC is regulated at different points through a number of plasma and host cell proteins. The initiation of the classical pathway through the attachment of C1 to an antibody is controlled by the C1 Inhibitor (C1-Inh), a protease inhibitor belonging to the serpin superfamily. C1-Inh irreversibly binds to and deactivates the active subunits of component C1 to prevent spontaneous fluid phase and chronic activation of complement [5]. The serum and host-tissue regulation of the upstream elements of the complement system is also achieved through the binding of C4 binding protein (C4BP) to C4b and through the binding of factor H to C3b [6]. These proteins are also capable of binding their respective components in the convertase form. Membrane cofactor protein (MCP or CD46) possesses a cofactor activity for C4b and C3b, which protects the host from self-activation of complement [7]. Delay accelerating factor (DAF or CD55) is able to recognize and dissociate both convertases [8]. MAC is inhibited by vitronectin and clusterin in the plasma and CD59 at the host surface. Carboxypeptidase-N, a regulator of inflammation is a plasma zinc metalloprotease that cleaves carboxyl-terminal arginines and lysines of complement complement components and anaphylatoxins such as C3a, C4a, and C5a. [9].

Research over the past decade has also shown that the role of complement extends beyond as a simple player in the immune system. Complement has been shown to contribute to homeostasis by inducing growth factors involved in tissue repair [4]. The malfunction of complement has been linked with several diseases including Alzheimers, glaucoma, Parkinson's disease, multiple sclerosis, schizophrenia, rheumatoid arthritis and sepsis [10, 11]. Complement has also been shown to play both a positive and negative role in certain cancers; attacking tumor cells with altered surface patterns in some

cases potentially contributing to tumor growth in others [2, 12]. Several networks maintain cross talk with complement including coagulation, autonomous nervous response and the ability to regulate inflammation [12]. The system's involvement in a variety of beneficial regulatory processes coupled with the role it can play in disease makes it important to understand complement in a more holistic perspective.

Thus given the complexity and importance of complement, developing models of complement within an integrative framework (that includes other biochemical networks) are crucial to understanding its role. Traditionally, complement models have been formulated as linear or non-linear Ordinary Differential Equation (ODE) systems. Hirayama et al. used a system of linear ODEs to model the classical pathway of complement [13]. Korotaevskiy and co-workers built a theoretical model of complement using a system of non-linear ODEs that included classical, lectin and alternate pathways [14]. However both these studies involve no validation studies with experimental data. Liu et al analyzed the formation of classical and lectin C3 convertases and the regulatory role of C4BP using a system of 45 non-linear ODEs with 85 parameters [15]. Recently, Zewde and co-workers built a detailed mechanistic model of alternative complement activation was built using 107 ODEs and 74 kinetic parameters [16]. This model delineated the response of complement on a host cell and a foreign antigen. However, these previous models were largely based upon mechanistic knowledge. Given the complexity of complement and its interactions with other networks it is unfeasible and computationally expensive to build such large mechanistic models. In addition is much more difficult to experimentally interrogate the response of various complement proteins under different conditions. This also presents with the problem of estimation of a large number of parameters with little or no experimental data. Thus there exists a need to reduce the mechanistic complexity while capturing dynamics of complement accurately.

In this study, we present a hybrid modeling approach to build a reduced order model of

complement. The key innovation of this approach is the use of simple equations to capture the behavior of a complex biochemical network. The hybrid approach combines ODEs with logical rules to model biochemical processes that are complex or for which a complete mechanistic understanding is missing. We used this framework to capture dynamics of C3a and C5a formation in the lectin and alternative pathways. The reduced order model consisted of only 18 differential equations with 28 kinetic and control parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable ODE models in the literature. We estimated the model parameters from in vitro time series data of C3a and C5a from Morad and coworkers [17]. Subsequently we validated the model on unseen C3a and C5a experimental data that were not used for model training. After validation, we performed a sensitivity analysis on the model to estimate which parameters were critical to model performance under different experimental conditions. Given its small size, the hybrid approach produced a surprisingly predictive human complement model, similar to an earlier study on human coagulation using the same modeling framework [18]. Taken together, the combined analysis of alternate and lectin pathways along with the incorporation of the downstream reactions involving C5 convertase elucidated new insight into the roles of parameters that govern the complement system. A deeper understanding about how these parameters influence complement dynamics will greatly aid in the development of drugs for strategic therapeutic targets. Due to the low computational cost relative to the existing models and accuracy of our predictions, we believe that our reduced order complement network is the first step towards building a computation toolbox for screening drug potential drug targets or therapeutic agents that can be targeted against complement.

Results

Formulation of a reduced order complement model We developed a reduced order human complement network consisting of the most crucial steps of the human complement system (Fig. 1). The core of our model was based upon the experimental measurements of Morad and coworker's earlier work [17], we only consider the activation of complement system through the alternate and the lectin pathways. In doing so we aim to capture a complex biological phenomenon using a few simple ordinary differential equations. A trigger event initiates the lectin pathway in the presence of zymosan, which activates the cleavage of C2 and C4 into C2a and C2b, and C4a and C4b respectively. Classical Pathway (CP) C3 convertase (C4aC2b) is a combination of C4a and C2b, which catalyzes the cleavage of C3 into C3a and C3b. Similarly, the activation of the alternative pathways happens through the spontaneous hydrolysis of C3 which facilitates the cleavage of C3. C3b then could combine with with C3 to form alternate pathway (AP) C3 convertase. Both C3 convertases catalyze the cleavage of C3 into C3a and C3b, and C3b can then combine with either CP or AP C3 convertase to form C5 convertase, CP or AP respectively that is responsible for the cleavage of C5 to C5a and C5b. Lectin pathway activation was approximated using a combination of saturation kinetics and Hill-like function control functions. These control coefficients then modified the rates of model processes at each time step. Hill-like transfer functions $0 \leq \mathbf{f}(\mathbf{Z}) \leq 1$ quantified the contribution of components upon a target process, in this study, $\mathbf{Z}$ represents the abundance of the initiator. Taken together, while the reduced order human complement model encodes significant biological complexity, it is highly compact (consisting of only 18 differential equations). Thus, it will serve as an excellent proof of principle example to study the reduction of a highly complex human subsystem.

An ensemble of complement models was estimated using dynamically dimensioned search. A critical challenge for any dynamic model is the estimation of kinetic param-

eters. We estimated kinetic and control parameters in a hierarchical fashion using two $in\ vitro$ time-series human complement data sets with and without zymosan present. The residual between simulation and experimental measurements were minimized using dynamically dimensioned search (DDS). An initial parameter set was initialized with randomized kinetic and control parameters and allowed to search for parameter vectors that minimized the residual. Knowing that the kinetic and control parameters of the lectin pathway does not affect the dynamics of the alternate pathway, we used a hierarchal approach that estimated the parameters for the alternative pathway and lectin pathway separately. For the alternative pathway, we utilized the time-course experimental measurements of Morad and coworkers [17] of C3a and C5a in the absence of zymosan and only allowed the alternative parameters to vary (Fig. 2 A and B). The estimated alternate parameters was then fixed for the determination of lectin pathway parameters. The training for the lectin parameters, we used the experimental measurements of C3a and C5a in the presence of 1 g of zymosan published by Morad et al [17] (Fig. 2 C and D). The reduced human complement model captured the behavior of the alternative and lectin pathways through the time-course abundance of C3a and C5a (Fig. 2). However we were not able to capture the curvature of the C5a alternative (Fig. 2). The decreasing slope of the experimental measurements may be an indication of the decreasing cofactors that are required for the spontaneous hydrolysis in the alternative pathway, which we neglected. Taken together, the model identification results suggested that our reduced order approach could reproduce a panel of lectin pathway initiation 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.

We tested the predictive power of the reduced order human complement model with validation data sets not used during model training. Six validation data sets were used, three for C3a and C5a respectively at different zymosan concentrations. All kinetic and

control parameters were fixed for the validation simulations. The reduced order model predicted the C3a and C5a time-course profiles at a qualitative level (Fig. 3). For the prediction of the $C3a$ at the three different inducer concentrations, our model was able to accurately capture the dynamic behaviors at a qualitative level, including the time the concentration of $C3a$ plateaus. For the prediction for $C5a$, our model was able to capture the behavior at a quantitative level for all three inducer concentrations. However, the our model was not able to capture the concave down curvature for the cases of $0.001\ g$ and $0.01\ g$ inducer. We believe some of the shortcomings of the prediction may be attributed by our reduced complexity in the modeling of $C4B$ interactions in our system. $C4BP$ interaction was modeled as irreversible mass action kinetics, in reality the binding of $C4BP$ onto other complement components are reversible. Due to this discrepancy in between our model and the experimental results, we believe that the interaction and the reversibility of the $C4BP$ is a key player in creating a threshold where C5a is resistant to minor changes in inducer concentration even though C3a is sensitive.

Global Sensitivity analysis of the reduced order complement model We conducted a Sobol's sensitivity analysis to estimate which parameters controlled the performance of the reduced order model. We calculated the sensitivity of the change in C3a and C5a profiles using the residuals between simulation and experimentally measured data for the cases of 0 and 1g zymosan (Fig. 4). For the cases in absence of zymosan where only the alternative pathway is active, we observed that only a few variables are responsible for the system response. For C3a alternate, the sensitivity analysis found that $k_{c3b\ basal}$ and $k_{degradationC3a}$ are the only sensitive parameters. This gives us new insight in which of the parameters play a role in complement activation. Even though AP C3 convertase is also responsible in the conversion of C3 and the production of C3a, the kinetic parameters that govern the equation was not sensitive at all. This elucidated that the activation of alternative pathway is more heavily governed by the spontaneous hydrolysis of C3 rather than

the activity of AP C3 Convertase. Surprisingly, closely examining the sensitive parameters that control C5a, in addition to the expected kinetic and control parameters related to the formation of AP C5 Convertase, we observed that $k_C3\ Convertase2$, the was previously not sensitive to C3a, to be sensitive in the formation of C5a. The AP C3 Convertase is a substrate required for the formation of AP C5 Convertase and the formation of C3b. The change in activity of AP C3 Convertase will not drastically change the C3a dynamics, but will effect AP C5a Convertase formation and C5a formation. The our reduced order human complement model in combination with Sobol's sensitivity analysis was able to unravel important indirect parameter interaction. Our sensitivity analysis yielded expected results for the lectin pathway analyzes (Fig. 4 (C and D)). One key difference that was observed between the sensitivity of the parameters between C3a an C5a was their respective degradation terms. The degradation constant of C3a was sensitive between the two different cases of zymosan that was tested while the degradation constant of the C5a was not sensitive. We believe this different is attributed to the magnitude of the parameters and their respective concentrations.

Discussion

The discussion has three (sometimes four) paragraphs:

1. **First paragraph**: Present a modified version of the last paragraph of the introduction. In this study, [...]. Taken together, [killer statement]

   In this study, we present a hybrid modeling approach to build a reduced order model of complement. The key innovation of this approach is the use of simple equations to capture the behavior of a complex biochemical network. The hybrid approach combines ODEs with logical rules to model biochemical processes that are complex or for which a complete mechanistic understanding is missing. We used this framework to capture dynamics of C3a and C5a formation in the lectin and alternative pathways. The reduced order model consisted of only 18 differential equations with 28 kinetic and control parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable ODE models in the literature. We estimated the model parameters from in vitro time series data of C3a and C5a from Morad and coworkers [17]. Subsequently we validated the model on unseen C3a and C5a experimental data that were not used for model training. After validation, we performed a sensitivity analysis on the model to estimate which parameters were critical to model performance under different experimental conditions. Given its small size, the hybrid approach produced a surprisingly predictive human complement model, similar to an earlier study on human coagulation using the same modeling framework [18]. Taken together, the combined analysis of alternate and lectin pathways along with the incorporation of the downstream reactions involving C5 convertase elucidated new insight into the roles of parameters that govern the complement system. A deeper understanding about how these parameters influence complement dynamics will greatly aid in the development of drugs for strategic

therapeutic targets. Due to the low computational cost relative to the existing models and accuracy of our predictions, we believe that our reduced order complement network is the first step towards building a computation toolbox for screening drug potential drug targets or therapeutic agents that can be targeted against complement.

2. **Second paragraph**: Contrast the key findings of the study with other computational/experimental studies

   Though the role of complement in immune response has been well known since long, there has been a paucity of mathematical models of complement. To our knowledge this is the first model of complement that combines different pathways of initiation and validates the dynamics of downstream proteins like C5a using experimental data. Liu and co-workers modeled formation of C3a through the classical pathway using 45 non-linear ODEs [15]. The hybrid modeling framework, however, allowed us to model the lectin mediated formation of C3a using only 5 ODEs. Though we do not model all the interactions of initiation in detail, especially the cross-talk between lectin and classical pathways like Liu et al. [15] we successfully captured C3a dynamics with respect to different concentrations of lectin initiators. The model was also surprisingly accurate in capturing the quantitative dynamics of C3a and C5a formed from the alternate pathway with only 7 ODEs. The lag phase in the initiation of C5a is followed by an acceleration in the production of C5a. This is also qualitatively similar to the predicted C5a time profiles using a theoretical model of complement with 107 equations [16]. Similarly, our model was able to capture C3a formation from the alternate pathway that showed the same qualitative trends as Zewde et al [16]. We also observe in our model simulation that the quantity of C3a produced in the alternate pathway is nearly 1000 times the quantity of C5a pro-

duced. Though this is in agreement with the experimental data [17], it differs from the theoretical predictions by Zewde et al.[16] who show that C3a is 10^8 times the C5a concentration. It is possible that the experimental system has trace amounts of C5a which leads to this difference. In our model the time profile of C5a generation from the lectin pathway changes with respect to the quantity of zymosan (the lectin pathway initiator). We see that the lag phase for generation progressively decreases with increasing concentration of the initiator. Korotaevskiy et al. [14] show a similar lag phase followed by an accelerated production of C5a using a theoretical model of complement for much smaller time scales. Zewde et al. [16] show a similar time profile for C5a generated in the alternate pathway as well. We do observe a similar trend in our model. However given the different time scale in our study we do not observe a very prominent lag phase. Our model also captures the maximum concentrations of C5a with different initiators and in the alternate pathway. However we do not capture the maximum concentration of C3a with low initiator levels though we capture the overall trend which includes the initiation and amplification phases. One of the possible reasons for this could be the non-inclusion of C2-by pass pathway which further accelerates production of C3a 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. Taken together we surprisingly do very well in capturing the dynamics of key complement proteins and observe similar trends as in large mechanistic models.

The global sensitivity analysis of our reduced order model provides interesting insights into the regulation of C3a and C5a which are important therapeutic targets. We see that the formation of C5a in the absence of an initiator is the most sensitive to the formation of basal C3b through the tick over mechanism. This mechanism

could be a potential target for C5a inhibitors especially in cases of autoimmune disorders where there is an absence of initiators of the lectin or the classical pathway. In presence of an initiator we see that C5a formation is primarily sensitive to the lectin initiation parameters and the kinetic parameters that govern the conversion of C5 to C5a and C5b. This result agrees well with the current protease inhibitors targeting initiating complexes that includes mannose-associated serine proteases 1 and 2 (MASP-1,2) [19]. The most commonly used anti-complement drug eculizumab [19], targets C5 protein which is cleaved to form C5a. Our sensitivity analysis shows that kinetic parameters governing C5 conversion are sensitive in both lectin initiated and alternate pathways, thus agreeing with targeting C5 protein. The formation of basal C3b is 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. The lectin initiated formation of C3a shows a number of sensitive parameters. This includes the lectin initiation parameters that control the formation of C5a along with C3 convertase inhibition using C4BP and the parameter governing C3 convertase. All these mechanisms involve potential drug targets and literature shows that C4BP, C3 convertase inhibitors are the being developed as the next generation inhibitors of complement [19]. Taken together it is surprising to see that reduced order model has the ability to identify mechanisms for potential drug targets that are in agreement with existing literature.

3. **Third paragraph**: Present future directions. If you had more time, what would like to do? Highlight the key shortcomings of the approach and how will we address them in the future. In this case, we will have a scaling issue if we extend to genome scale. We should extend to dynamic cases, and we need to experimentally validate the findings.

The performance of the reduced order complement model was impressive given its limited size. However, there are several critical questions that should be explored following this study. A logical progression for this work would include expanding the network to include the classical pathway and the formation of the membrane attack complex (MAC). It is unclear whether the addition of the classical pathway will decrease the prediction of our existing model due to the cross-talk between the classical and lectin activation shown by Liu et al [15]. One potential approach in addressing such difficulties would be the incorporation of additional species such as C reactive proteins (CRP) and L-ficolin (LF) that involved in complement initiation of classical and lectin pathways. The influence of CRP, LF and the cross-talk can be captured through additional control functions that act upon the initiation pathways in a logical integration rule developed by Wayman and coworkers [20]. Another issue with our reduced order model involve the omitted species that are implicitly lumped together with our effective kinetics and control parameters. Due to the reduction of parameters, the model cannot determine the dynamics or explicit impact of the omitted species on the system. However, we have created a hierarchy approach for parameter estimation that can be used to uncouple the kinetic parameter and contribution of any additional complement proteins and regulators. Using this simple and versatile modeling approach that we created, we took the first step in the development of a computation toolkit that can be readily used in a clinical setting. Our reduced order complement model is computationally inexpensive, and versatile so it could easily be incorporated into pre-existing or new pharmacokinetic models. Furthermore this approach model has the potential to create individualized treatment plans for patients with complement deficiency.

Materials and Methods

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

$$\frac{dx_i}{dt} = \sum_{j=1}^{\mathcal{R}} \sigma_{ij} r_j \left(\mathbf{x}, \epsilon, \mathbf{k}\right) \qquad i = 1, 2, \ldots, \mathcal{M} \tag{1}$$

where $\mathcal{R}$ denotes the number of reactions, $\mathcal{M}$ denotes the number of protein species in the model. The quantity $r_j \left(\mathbf{x}, \epsilon, \mathbf{k}\right)$ denotes the rate of reaction j. Typically, reaction j is a non-linear function of biochemical species abundance, as well as unknown kinetic 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 $\mathbf{x}\left(t_o\right) = \mathbf{x}_o$.

The reaction rates controlling formation C4a, C4b, C2a and C2b were written as a product of a kinetic term ($\bar{r}_j$) and a control term (v_j) such that $r_j \left(\mathbf{x}, \mathbf{k}\right) = \bar{r}_j v_j$. The kinetic term for these rates was modeled using saturation kinetics. The control term $0 \leq v_j \leq 1$ for these reaction rates was modeled using regulatory transfer functions which took the form:

$$f_{ij} \left(\mathcal{Z}_i, k_{ij}, \eta_{ij}\right) = k_{ij}^{\eta_{ij}} \mathcal{Z}_i^{\eta_{ij}} / \left(1 + k_{ij}^{\eta_{ij}} \mathcal{Z}_i^{\eta_{ij}}\right) \tag{2}$$

where $\mathcal{Z}_i$ denotes the abundance factor i, k_{ij} denotes a gain parameter, and η_{ij} denotes a cooperativity parameter.

We used saturation kinetics to model the lectin pathway activation and C3 and C5 convertase activity $\bar{r}_j$:

$$\bar{r}_j = k_j^{max} \epsilon_i \left(\frac{x_s^{\eta}}{K_{js}^{\eta} + x_s^{\eta}}\right) \tag{3}$$

where k_j^{max} denotes the maximum rate for reaction j, ϵ_i denotes the enzyme abundance

which catalyzes reaction j, η denotes a cooperativity parameter (similar to a Hill coefficient), and K_{js} denotes the saturation constant for species s in reaction j. On the other hand, we used mass action kinetics to model the protein conversion reactions within the network $\bar{r}_j$:

$$\bar{r}_j = k_j^{max} \prod_{s \in m_j^-} x_s \tag{4}$$

where k_j^{max} denotes the maximum rate for reaction j, ϵ_i denotes the enzyme abundance which catalyzes reaction j. The product in Eqn (4) was carried out over the set of *reactants* for reaction j (denoted as m_j^-).

Estimation of an ensemble of model parameters. Model parameters were estimated by minimizing the difference between simulations and experimental C3a and C5a measurements (squared residual):

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

where $\hat{x}_j(\tau)$ denotes the measured value of species j at time τ, $x_j(\tau, \mathbf{k})$ denotes the simulated value for species j at time τ, and $\omega_j(\tau)$ denotes the experimental measurement variance for species j at time τ. The outer summation is with respect to time, while the inner summation is with respect to state.

We minimized the model residual using Dynamic Optimization with Particle Swarms (DOPS). DOPS is a novel metaheuristic that combines multi swarm particle swarm optimization (PSO) with a greedy global optimization algorithm called dynamically dimensioned search (DDS). DOPS is much faster than conventional global optimizers and has the ability to find near optimal solutions for high dimensional systems within a relatively few function evaluations. It uses an adaptive switching strategy based on error convergence rates to switch from swarms search to DDS search. This enables it to find quickly,

globally optimal or close to globally optimal solutions even in the presence of many local minima. In the swarm search, for each iteration the particles compute error within each sub-swarm by evaluating the model equations using their specific parameter vector realization. From each of these points within a sub-swarm a local best is identified. This along with the particle best within the sub-swarm $\mathcal{S}_k$ is used to update the parameter estimate for each particle using the following rules:

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

where $z_{i,j}$ is the parameter vector, $(\theta_1, \theta_2, \theta_3)$ were adjustable parameters, $\mathcal{L}_i$ denotes the best solution found by particle i within sub-swarm $\mathcal{S}_k$ for function evaluations $1 \rightarrow j-1$, and $\mathcal{G}_k$ denotes the best solution found over all particles within sub-swarm $\mathcal{S}_k$. The quantities 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:

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

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:

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

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 $\mathcal{G}$ for our model which we use for plotting best fits against the experimental data. We perturb this parameter vector to generate an ensemble of parameter vectors. The quality of parameter estimates was measured using goodness of fit (model residual). The DOPS routine was implemented in the MATLAB programming language.

Global sensitivity analysis of model performance We conducted a global sensitivity analysis, using the variance-based method of Sobol, to estimate which parameters controlled the performance of the reduced order model [21]. We computed the total sensitivity index of each parameter relative to four performance objectives, each objective was based on the sum of squared errors between model and experimental data for C3a alternate, C5a alternate, C3a lectin, and C5a lectin simulations. We established the sampling bounds for each parameter from the minimum and maximum value of that parameter in the parameter set ensemble. We used the sampling method of Saltelli *et al.* [22] to compute a family of $N\,(2d+2)$ parameter sets which obeyed our parameter ranges, where N was the number of trials, and d was the number of parameters in the model. In our case, N = 200 and d = 28, so the total sensitivity indices were computed from 11,600 model evaluations. The variance-based sensitivity analysis was conducted using the SALib module encoded in the Python programming language [23].

Acknowledgements

This study was supported by an award from [FILL ME IN].

References

1. Walport MJ (2001) Complement. first of two parts. The New England journal of medicine .
2. Sarma JV, Ward PA (2011) The complement system. Cell and tissue research 343: 227–235.
3. Pangburn MK, Müller-Eberhard HJ (1984) The alternative pathway of complement. Springer seminars in immunopathology .
4. Ricklin D, Hajishengallis G, Yang K, Lambris JD (2010) Complement: a key system for immune surveillance and homeostasis. Nature immunology 11: 785–797.
5. 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 research 675: 75–82.
6. Blom AM, Kask L, Dahlbäck B (2001) Structural requirements for the complement regulatory activities of c4bp. Journal of Biological Chemistry 276: 27136–27144.
7. Riley-Vargas RC, Gill DB, Kemper C, Liszewski MK, Atkinson JP (2004) Cd46: expanding beyond complement regulation. Trends in immunology 25: 496–503.
8. 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. Proceedings of the National Academy of Sciences of the United States of America 101: 1279–1284.
9. Liszewski MK, Farries TC, Lublin DM, Rooney IA, Atkinson JP (1995) Control of the complement system. Advances in immunology 61: 201–283.
10. Ricklin D, Lambris JD (2007) Complement-targeted therapeutics. Nature biotechnology .
11. Rittirsch D, Flierl MA, Ward PA (2008) Harmful molecular mechanisms in sepsis. Nature Reviews Immunology 8: 776–787.
12. Ricklin D, Lambris JD (2013) Complement in immune and inflammatory disorders:

pathophysiological mechanisms. The Journal of Immunology 190: 3831–3838.

13. 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.
14. Korotaevskiy AA, Hanin LG, Khanin MA (2009) Non-linear dynamics of the complement system activation. Mathematical biosciences .
15. 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 .
16. Zewde N, Gorham Jr RD, Dorado A, Morikis D (2016) Quantitative modeling of the alternative pathway of the complement system. PloS one .
17. 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. Journal of immunological methods 427: 13–18.
18. Sagar A, Varner JD (2015) Dynamic modeling of the human coagulation cascade using reduced order effective kinetic models. Processes .
19. Morgan BP, Harris CL (2015) Complement, a target for therapy in inflammatory and degenerative diseases. Nature Reviews Drug Discovery .
20. Wayman JA, Sagar A, Varner JD (2015) Dynamic modeling of cell-free biochemical networks using effective kinetic models. Processes .
21. Sobol I (2001) Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. Mathematics and Computers in Simulation 55: 271 - 280.
22. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, et al. (2010) Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. Computer Physics Communications .
23. Herman J. Salib: Sensitivity analysis library in python (numpy). con-

tains sobol, morris, fractional factorial and fast methods. available online: https://github.com/jdherman/salib.

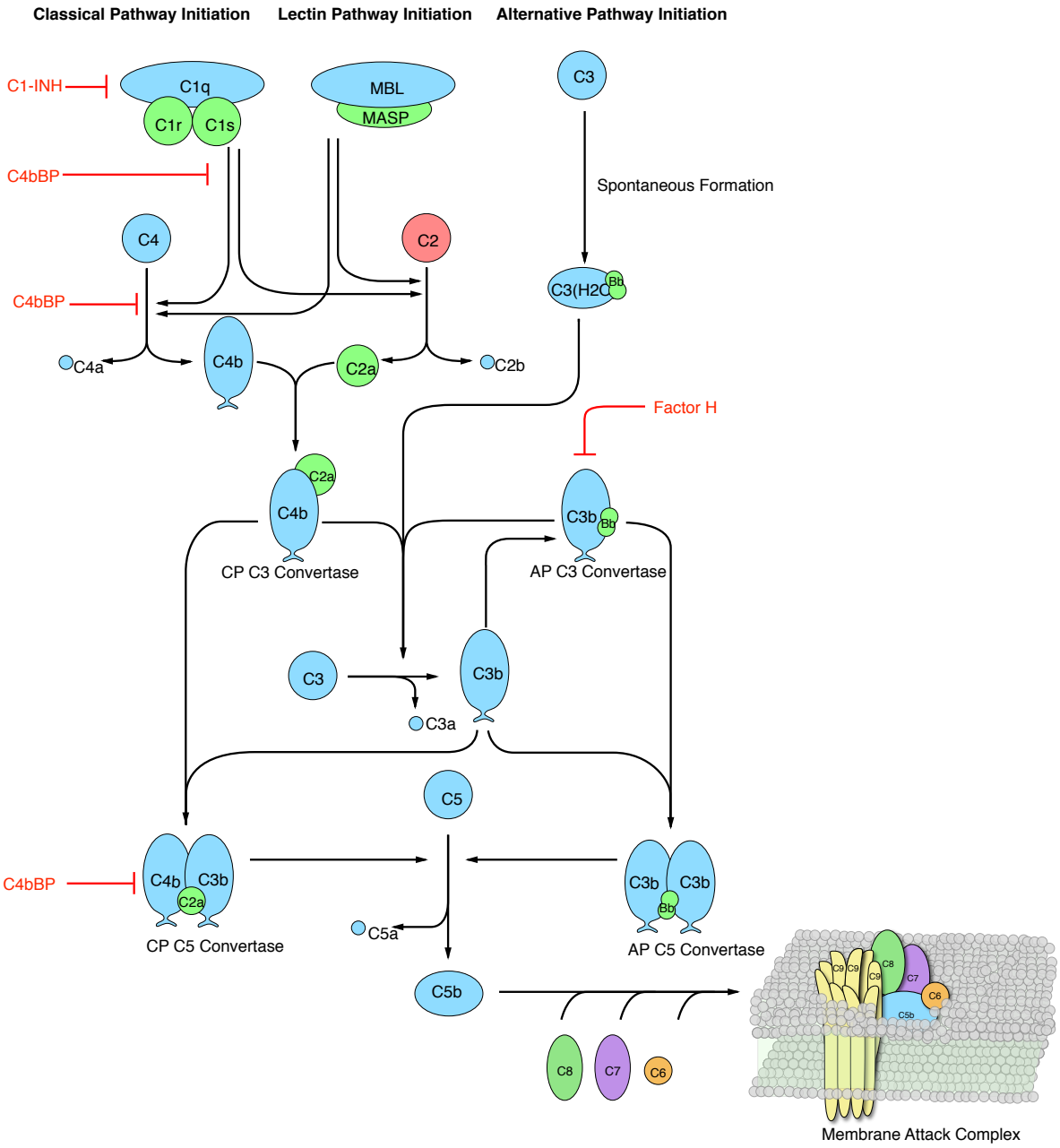

Fig. 1: Simplified schematic representation of the human complement system. The complement cascade is activated through any one, or more, of the three pathways: classical, lectin, and alternate pathway. The classical pathway is activated by the complex formation of $C1q$, $C1r$, and $C1s$ by the recognition of antibody:antigen complexes. Similarly, the lectin pathway is initiated by binding mannan-binding lectin to mannose on pathogen surfaces. The alternative pathway is activated through a spontaneous tick-over mechanism by the hydrolysis of C3 to form fluid phase C3 convertase. The activation from the three pathways creates a cascades of reactions that forms the proteases, $C3$ Convertase that cleaves $C3$ into $C3a$, and $C3b$, the main effector molecule of the complement system. $C3b$ can find to a $C3$ convertase and form a $C5$ convertase that cleaves $C5$ into $C5a$, and $C5b$ that undergoes a series of reactions to form the membrane attach complex (MAC).

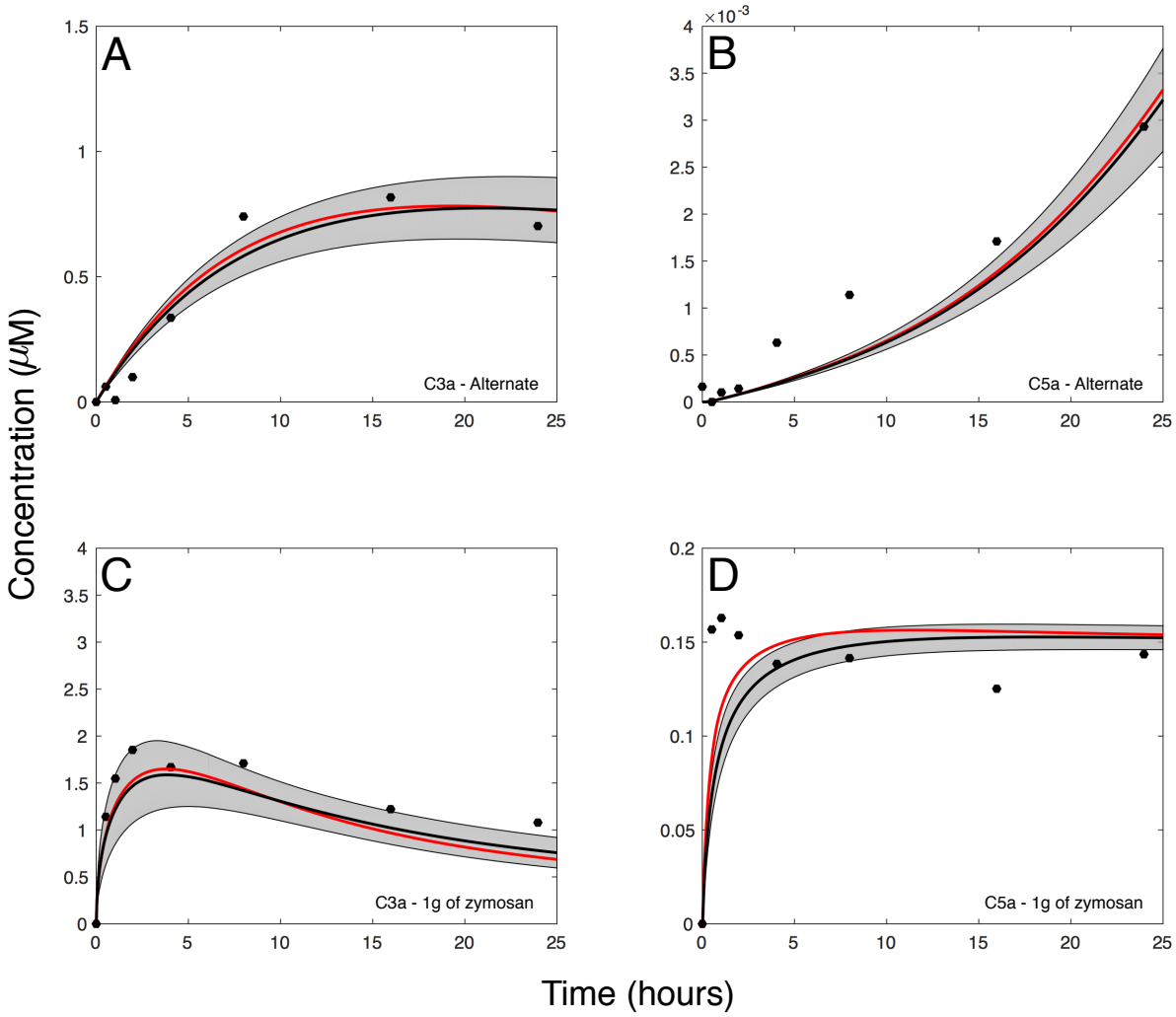

Fig. 2: Reduced order complement model training simulation for lectin and alternative pathway in presence of zymosan. Reduced order complement model parameters were estimated using Dynamic Optimization with Particle Swarms (DOPS). The model was trained against experimental data from Shaw and co-workers [17] in the presence and absence of zymosan. The red line shows the best-fit parameter, the black lines denotes the simulated mean value of $C3a$ or $C5a$ for a 50 parameter set ensemble. The shaded region denotes 99 % confidence interval on the simulated mean concentration.

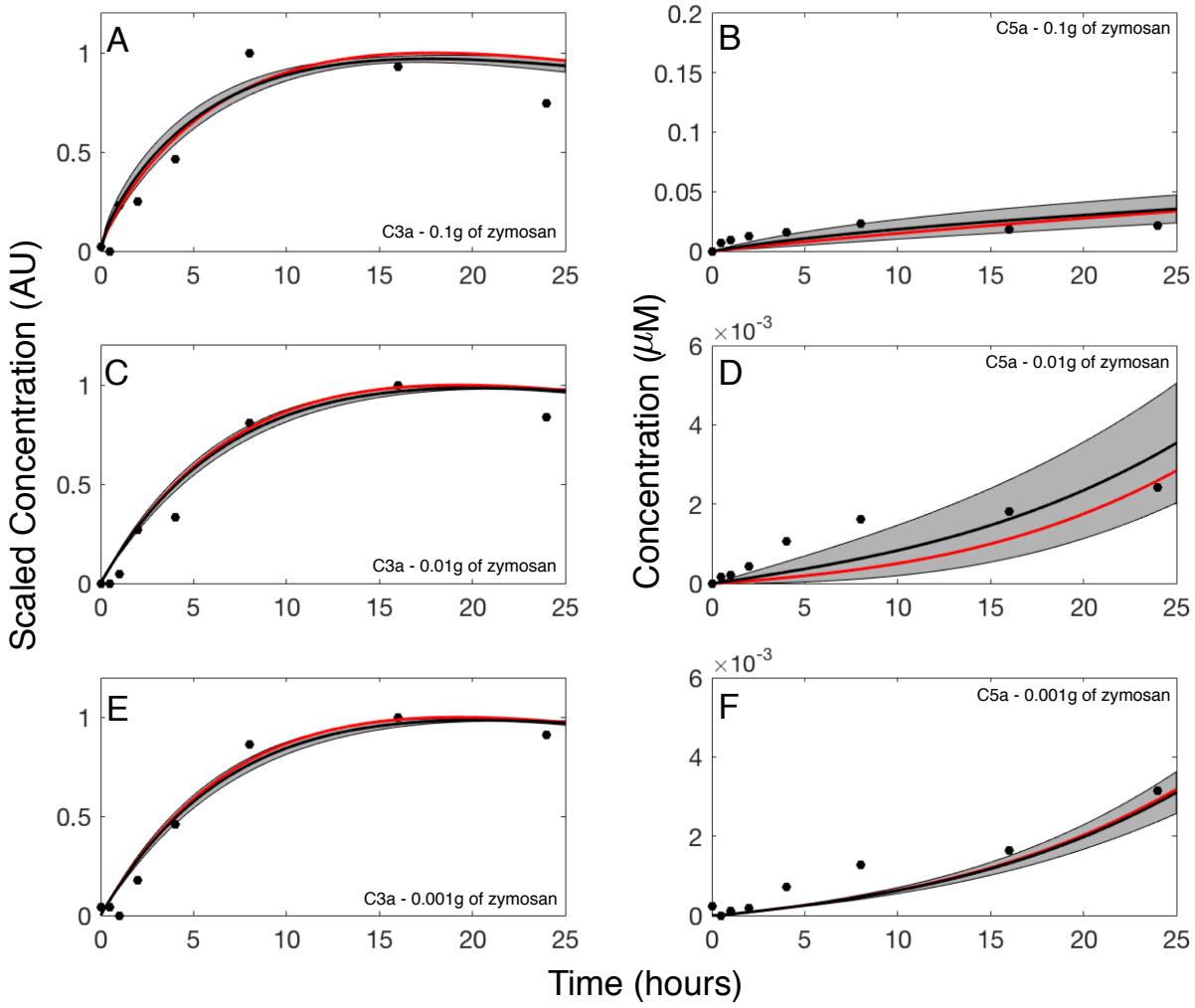

Fig. 3: Reduced order complement model predictions of lectin and alternative pathway in presence of zymosan. (A-F) Simulation of complement dynamics in the presence of zymosan were conducted for a range of trigger values (0.1, 0.01, and 0.001 grams of zymosan). The time-course profiles of $C3a$ and $C5a$ under three different zymosan concentrations were predicted using 50 ensembles of trained parameter sets against experimental data of Shaw and co-workers [17]. The red curve represents the best fit parameter, the black curve is the mean of the ensemble. The shaded region denotes 99 % confidence interval on the simulated mean concentration. All complement protein initial concentrations are at human serum levels unless otherwise noted.

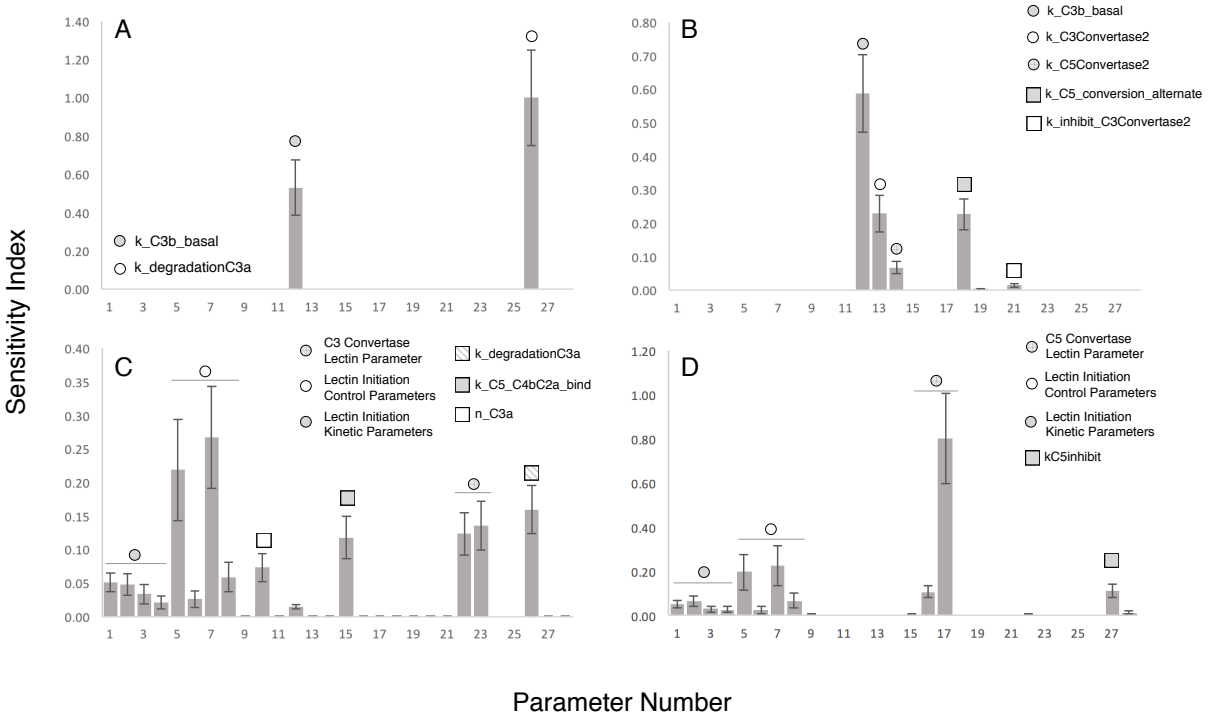

Fig. 4: Sobol's sensitivity analysis of the reduced order complement model with respect to the modeling parameters. Sensitivity analysis was conducted on the four cases we used to train our model: (A) C3a at 0 zymosan, (B) C5a 0 zymosan, (C) C3a 1 g zymosan, and (D) C5a 1 g zymosan. The bars denote total sensitivity index which includes local contribution of each parameter and global sensitivity of significant pairwise interactions. The error bars are the 95 percent confidence interval. k represents association rate, km denote Michaelis-Menten saturation constants, and $alpha$ and n refers to the exponentials of the control functions.

Supplemental materials.

Model equations. The reduced-order complement model consisted of 18 ordinary differential equations, 12 rate equations, and two control equations:

$$\frac{dx_1}{dt} = -r_1 f_1 \quad \text{(S1)}$$
$$\frac{dx_2}{dt} = -r_2 f_2 \quad \text{(S2)}$$
$$\frac{dx_3}{dt} = r_1 f_1 \quad \text{(S3)}$$
$$\frac{dx_4}{dt} = r_1 f_1 - r_6 \quad \text{(S4)}$$
$$\frac{dx_5}{dt} = r_2 f_2 - r_6 \quad \text{(S5)}$$
$$\frac{dx_6}{dt} = r_2 f_2 \quad \text{(S6)}$$
$$\frac{dx_7}{dt} = r_3 - r_4 - r_5 \quad \text{(S7)}$$
$$\frac{dx_8}{dt} = r_3 + r_4 + r_5 - k_{deg,c3a} * C3a \quad \text{(S8)}$$
$$\frac{dx_9}{dt} = r_3 + r_4 + r_5 - r_7 \quad \text{(S9)}$$
$$\frac{dx_{10}}{dt} = r_6 - r_{10} - r_8 \quad \text{(S10)}$$
$$\frac{dx_{11}}{dt} = r_7 - r_{11} - r_9 \quad \text{(S11)}$$
$$\frac{dx_{12}}{dt} = r_{10} - r_{14} \quad \text{(S12)}$$
$$\frac{dx_{13}}{dt} = r_{10} \quad \text{(S13)}$$
$$\frac{dx_{14}}{dt} = -r_{12} - r_{13} \quad \text{(S14)}$$
$$\frac{dx_{15}}{dt} = r_{12} + r_{13} - k_{deg,c5a} \quad \text{(S15)}$$
$$\frac{dx_{16}}{dt} = r_{12} + r_{13} \quad \text{(S16)}$$
$$\frac{dx_{17}}{dt} = -r_8 - r_{14} \quad \text{(S17)}$$
$$\frac{dx_{18}}{dt} = -r_9 \quad \text{(S18)}$$
$$\quad \text{(S19)}$$

where the rate equations are given by:

$$r_1 = \frac{k_{i1}(C4)}{(K_{1s} + C4)} \quad \text{(S20)}$$

$$r_2 = \frac{k_2(C2)}{(K_{2s} + C2)} \quad \text{(S21)}$$

$$f_1 = \frac{Zymo^{\eta_1}}{(Zymo^{\eta_1} + \alpha_1^{\eta_1})} \quad \text{(S22)}$$

$$f_2 = \frac{Zymo^{\eta_2}}{(Zymo^{\eta_2} + \alpha_2^{\eta_2})} \quad \text{(S23)}$$

$$r_3 = k_3(C3) \quad \text{(S24)}$$

$$r_4 = \frac{k_4(C3C_L)(C3^{\eta_3})}{(K_{4s}^{\eta_3} + C3^{\eta_3})} \quad \text{(S25)}$$

$$r_5 = \frac{k_5(C3C_A)(C3)}{(K_{5s} + C3)} \quad \text{(S26)}$$

$$r_6 = k_6(C4b)(C2a) \quad \text{(S27)}$$

$$r_7 = k_7(C4b)(C2a) \quad \text{(S28)}$$

$$r_8 = k_8(C3C_L)(C4b)(C4BP) \quad \text{(S29)}$$

$$r_9 = k_9(C3C_A)(FactorH) \quad \text{(S30)}$$

$$r_{10} = k_{10}(C3C_L)(C3b) \quad \text{(S31)}$$

$$r_{11} = k_{11}(C3C_A)(C3b) \quad \text{(S32)}$$

$$r_{12} = \frac{k_{12}(C5C_L)(C5^{\eta_4})}{(K_{12s}^{\eta_4} + C5^{\eta_4})} \quad \text{(S33)}$$

$$r_{13} = \frac{k_{13}(C5C_A)(C5)}{(K_{13s} + C5)} \quad \text{(S34)}$$

$$r_{14} = k_{14}(C5C_L)(C4BP) \quad \text{(S35)}$$