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: *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 which plays a significant role in the inflammatory response, and many other disease processes. In this study, we analyzed an ensemble of experimentally validated reduced order complement models. Our reduced order modeling approach combined ODEs with logical rules to produce a predictive model with a limited number of equations and parameters. We used this framework to capture the dynamics of C3a and C5a formation in the lectin and alternative pathways. The reduced order model consisted of only 18 differential equations with 28 parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable ODE 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. Given its small size, the hybrid approach produced a surprisingly predictive human complement model. After validation, we performed a global sensitivity analysis on the model ensemble to estimate which parameters were critical to model performance under different experimental conditions.

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

Introduction

Complement is a central part of innate immunity which plays a significant role in the inflammatory response. Complement was discovered in the 1890s where it was found to 'complement' the bactericidal activity of natural antibodies [REF]. However, research over the past decade has shown the importance of complement extends well beyond innate immunity. For example, complement contributes to tissue homeostasis by inducing growth factors involved in tissue repair [1]. Complement malfunctions have been linked with several diseases including Alzheimers, glaucoma, Parkinson's disease, multiple sclerosis, schizophrenia, rheumatoid arthritis and sepsis [2, 3]. Complement can also play both a positive and negative role in certain cancers; attacking tumor cells with altered surface proteins in some cases, while potentially contributing to tumor growth in others [4, 5]. Several other important biochemical networks are integrated with complement including the coagulation cascade, the autonomous nervous response and the ability to regulate inflammation [5]. Thus, complement is an important system involved in a variety of both beneficial and potentially harmful functions in the body.

Complement is mediated by over 30 soluble and cell surface proteins that are present as inactive forms in the circulation [6]. The central output of complement activation is the formation of the Membrane Attack Complex (MAC) and a key protein called C5a. The membrane attack complex forms transmembrane channels which disrupt the cell membrane of targeted cells, leading to cell lysis and death [REF]. On the other hand, the C5a protein acts as a bridge between innate and adaptive immunity, and plays an important role in regulating inflammation and coagulation [4]. Complement activation takes places through three pathways: the alternate, the classical and the lectin binding pathway. Each of these pathways involves a different initiator signal which leads to a cascade of downstream events 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 then cleaves complement proteins C4 and C2 into C4a, C4b, C2a and C2b respectively. The C4a and C2b fragments combine to form the C4bC2a protease, also known as the classical C3 convertase. The lectin binding pathway is initiated through the binding of L-ficolin or Mannose Binding Lectin (MBL) to carbohydrates on the surfaces of bacterial pathogens. This bound complex in turn cleaves C4 and C2, leading to the formation of C4bC2a. The alternate pathway involves a 'tickover' mechanism in which complement protein C3 is hydrolyzed to form C3b. In presence of pathogens, the C3b fragment binds foreign surfaces and recruits the additional proteins, factor B and factor D, which lead to the formation of C3bBb, the alternate C3 convertase [7]. 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 the classical C5 convertase (C4bC2aC3b), and alternate C5 convertase (C3bBbc3B) respectively. The C5 convertases then cleave C5 to form the C5a and C5b fragments. The cleavage of C5 is followed by a series of sequential cleavage steps involving the C6, C7, C8 and C9 complement proteins which combine with C5b to form the membrane attack complex [1].

Complement activation is regulated by many plasma and host cell proteins. The initiation of the classical pathway via complement protein C1 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 C1, preventing spontaneous fluid phase and chronic activation of complement [8]. Regulation of the upstream elements of complement is also achieved through the interaction of the C4 binding protein (C4BP) with C4b, as well as through the interaction of factor H with C3b [9]. These regulatory proteins are also capable of binding their respective targets in the convertase form as well. Membrane cofactor protein (MCP or CD46) possesses a cofactor activity for C4b and C3b,

which protects the host from self-activation of complement [10]. Delay accelerating factor (DAF or CD55) is also able to recognize and dissociate both C3 and C5 convertases [11]. Carboxypeptidase-N, a well known inflammation regulator, cleaves carboxyl-terminal arginines and lysines of the complement proteins C3a, C4a, and C5a rendering them inactive [12]. Lastly, the assembly of the MAC complex is inhibited by vitronectin and clusterin in the plasma, and CD59 at the host surface [REF]. Thus, there are many points of control which influence complement activation across the three activation pathways.

Developing quantitative mathematical models of complement could be crucial to 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 [13], while Korotaevskiy and co-workers modeled the classical, lectin and alternate pathways as a system of non-linear ODEs [14]. More recently, large mechanistic models of sections of complement have also been proposed. For example, Liu et al. analyzed the formation of the 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 constructed a detailed mechanistic model of the alternative pathway which consisted of 107 ODEs and 74 kinetic parameters and delineated the complement response of the host and pathogen [16]. However, these previous modeling studies involved little (if any) experimental validation. Thus, while these models are undoubtably important theoretical tools, it is unclear if they can describe or quantitatively predict experimentally validated complement dynamics. The central challenge is the estimation of model parameters from experimental data. Unlike other important cascades, such as coagulation for which there are well developed experimental tools and many publicly available data sets, the data for complement is relatively sparse. Missing or incomplete data sets, and limited quantitative data make the identification of mechanistic complement models difficult.

In this study, we analyzed an ensemble of reduced order complement models. Our reduced order modeling approach combined ODEs with logical rules to produce a predictive model with a limited number of equations and parameters. We used this framework to capture the dynamics of C3a and C5a formation in the lectin and alternative pathways. The reduced order model consisted of only 18 differential equations with 28 parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable ODE models in the literature. We estimated an ensemble of model parameters from *in vitro* time series measurements of C3a and C5a from Morad and coworkers [17]. Subsequently, we validated the model on unseen C3a and C5a measurements that were not used for model training. Given its small size, the hybrid approach produced a surprisingly predictive human complement model. After validation, we performed a global sensitivity analysis on the model ensemble to estimate which parameters were critical to model performance under different experimental conditions.

Results

Reduced order complement network. The reduced order 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. On the other hand, activation of the alternative pathway was initiated through the spontaneous hydrolysis of C3 into C3a and C3b. The C3b fragment then recombined with C3 to form the alternate pathway (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. The C5 convertase catalyzed the cleavage of C5 into the C5a and C5b fragments. Lectin pathway activation was approximated using a combination of saturation kinetics and non-linear transfer functions which modified the rates of model processes at each time step. These transfer functions allowed a significant reduction in the size of the model, while maintaining the biological performance. Thus, while the reduced order complement model encoded significant biological complexity, 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 any dynamic model is the estimation of model parameters. We estimated the complement model parameters in a hierarchical fashion using two *in vitro* time-series data sets generated with and without zymosan, a lectin pathway activator [17]. The residual between model simulations and experimental measurements was minimized using the dynamic optimization with particle swarms (DOPS) approach, starting from an initial random parameter guess. Knowing the parameters of the lectin pathway did not affect the dynam-

ics of the alternate pathway, we used a hierarchal approach that estimated the alternative and lectin pathway parameters separately. The reduced order complement model captured the behavior of the alternative and lectin pathways (Fig. 2). For the alternative pathway, we used the C3a and C5a measurements in the absence of zymosan, and only allowed the alternative parameters to vary (Fig. 2A and B). The putative alternate parameters were then fixed, and the lectin parameters were estimated. Lectin parameters were estimated from C3a and C5a measurmemts in the presence of 1g zymosan (Fig. 2C and D). Taken together, the reduced order model reproduced 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. To address this question, we fixed the model parameters and simulated data 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 at different zymosan concentrations. All model parameters were fixed for the validation simulations. The ensemble of reduced order models captured the qualitative dynamics of C3a formation (Fig. 3, left column), and C5a formation (Fig. 3, right column) at three inducer concentrations. However, there were shortcomings, especially for the C3a prediction. First, while the C3a dynamics were captured, and particularly the peak time, the overall level of C3a was under-predicted in all cases (Fig. 3, inset left column). 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 convertases may have residual activity even in the bound form. Thus, we likely over-predicted the influence of C4BP. We also failed to capture the concave down curvature for the 0.001 g and 0.01 g zymosan cases in the

C5a validation studies. The decreasing slope of the C5a measurements may indicate decreasing cofactors abundance, or missing biology which we have not explicitly accounted for in the reduced order approach. However, despite these shortcomings, we qualitatively predicted unseen experimental data, including correctly capturing the dynamic time scale of C3a formation, and the correct order of magnitude for the concentration of C5a for three inducer levels. Next, we used global sensitivity analysis to determine which parameters controlled the performance of the complement model.

Global sensitivity analysis of the reduced order complement model We conducted global sensitivity analysis to estimate which parameters controlled the performance of the reduced order complement model. We calculated the sensitivity of the C3a and C5a residuals with and without zymosan for the ensemble of parameter sets (Fig. 4). In the absence of zymosan (where only the alternative pathway is active), $k_{f,C3b}$ (formation of C3b) and $k_{d,C3a}$ (degradation rate constant governing C3a) were largely responsible for the system response. Interestingly, $k_{c,C3}$ (the rate constant governing AP C3-convertase activity) was not sensitive in the absence of zymosan. Thus, the behavior of the alternative pathway was more heavily influenced by the spontaneous hydrolysis of C3, rather than AP C3-convertase activity. On the other hand, $k_{c,C3}$ was one of the parameters that controlled C5a formation, in addition to the expected parameters related to AP C5-convertase formation. The AP C3-convertase is required for AP C5-convertase formation, and the formation of the C3b fragment. Thus, changes in the activity of AP C3-convertase will not drastically change the C3a dynamics, but will effect AP C5-convertase activity and C5a formation. The sensitivity analysis yielded the expected results for the lectin pathway (Fig. 4C and D). One key difference observed between the sensitivity of C3a an C5a parameters, was their respective degradation constants. The rate constant governing C3a degradation was sensitive, while the degradation constant for C5a was not. This difference was likely attributable to the magnitude of the degradation parameters and the

respective concentration of C3a and C5a. Taken together, global sensitivity analysis identified important indirect parameter interactions that could have therapeutic significance.

Discussion

In this study, we analyzed an ensemble of reduced order complement models. The reduced order modeling approach combined ODEs with logical rules to produce a predictive model with a limited number of equations and parameters. The reduced order model consisted of only 18 differential equations with 28 parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable ODE models in the literature. We used this framework to simulate the dynamics of C3a and C5a formation in the lectin and alternative pathways. We estimated an ensemble of model parameters from *in vitro* time series measurements of C3a and C5a abundance. Subsequently, we validated the model on unseen C3a and C5a measurements that were not used for model training. Given its small size, the hybrid approach produced a surprisingly predictive complement model. After validation, we performed a global sensitivity analysis on the model ensemble to estimate which parameters were critical to model performance under different experimental conditions. The global sensitivity analysis identified important indirect parameter interactions that could have therapeutic significance. Using a simple and versatile modeling approach, we developed a 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.

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 [15]. 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 [17], 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 [14]. Thus, the reduced order complement model performed similarly to existing large mechanistic models, despite being significantly smaller.

Global sensitivity 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 [2]. Previously, we have shown that mathematical modeling and sensitivity analysis can be useful tools to estimate therapeutically important mechanisms in biochemical networks [18–21]. 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) [22]. The most commonly used anti-complement drug

eculizumab [22], 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. For example, both C4BP and C3 convertase inhibitors are currently being developed as next generation complement inhibitors [REF ORIGINAL STUDIES NOT REVIEW]. Thus, analysis of the ensemble of complement models identified potentially important therapeutic targets.

The performance of the reduced order complement model was impressive given its limited size. However, there are several questions that should be explored further. A logical progression for this work would be to expand the network to include the classical pathway and the formation of the membrane attack complex (MAC). However, it is unclear whether the addition of the classical pathway will decrease the predictive quality of our existing model. Liu et al have shown cross-talk between the activation of the classical and lectin pathways that could influence model performance [15]. One potential approach to address such difficulties would be to incorporate C reactive proteins (CRP) and L-ficolin (LF) into the model, both of which are involved with the initiation of classical and lectin pathways [HOW DOES THIS HELP THE CROSSTALK ISSUE?]. 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 reasons for this could 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

We used ordinary differential equations (ODEs) to model the time evolution of complement proteins (x_i) in the reduced order 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 and $\mathcal{M}$ 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 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 $\mathbf{x}\left(t_o\right) = \mathbf{x}_o$.

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:

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

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 protein-protein binding interactions within the network:

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

where k_j^{max} denotes the maximum rate for reaction j, σ_{sj} denotes the stoichiometric coefficient for species s in reaction j, and $s \in m_j$ denotes the set of *reactants* for reaction j. The control terms $0 \leq v_j \leq 1$ depended upon the combination of factors which influenced rate process j. For each rate, we used a rule-based approach to select from competing control factors. If rate j was influenced by $1, \ldots, m$ factors, we modeled this relationship as $v_j = \mathcal{I}_j\left(f_{1j}\left(\cdot\right), \ldots, f_{mj}\left(\cdot\right)\right)$ where $0 \leq f_{ij}\left(\cdot\right) \leq 1$ denotes a regulatory transfer function quantifying the influence of factor i on rate j. The function $\mathcal{I}_j\left(\cdot\right)$ is an integration rule which maps the output of regulatory transfer functions into a control variable. Each regulatory transfer function took the form:

$$f_{ij}\left(\mathcal{Z}_i, k_{ij}, \eta_{ij}\right) = k_{ij}^{\eta_{ij}} \mathcal{Z}_i^{\eta_{ij}} / \left(1 + k_{ij}^{\eta_{ij}} \mathcal{Z}_i^{\eta_{ij}}\right) \tag{4}$$

where $\mathcal{Z}_i$ denotes the abundance of factor i, k_{ij} denotes a gain parameter, and η_{ij} denotes a cooperativity parameter. In this study, we used $\mathcal{I}_j \in \{mean\}$ [23]. If a process has no modifying factors, $v_j = 1$. The model equations implemented in MATLAB and solved using the `ODE15s` routine (The Mathworks, Natick MA). The complement model code and parameter ensemble can be downloaded from http://www.varnerlab.org.

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):

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

where $\hat{x}_j\left(\tau\right)$ denotes the measured value of species j at time τ, $x_j\left(\tau, \mathbf{k}\right)$ denotes the simulated value for species j at time τ, and $\omega_j\left(\tau\right)$ 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 [24]. 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.* [25] 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 [26].

Acknowledgements

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

References

1. Ricklin D, Hajishengallis G, Yang K, Lambris JD (2010) Complement: a key system for immune surveillance and homeostasis. Nature immunology 11: 785–797.
2. Ricklin D, Lambris JD (2007) Complement-targeted therapeutics. Nature biotechnology .
3. Rittirsch D, Flierl MA, Ward PA (2008) Harmful molecular mechanisms in sepsis. Nature Reviews Immunology 8: 776–787.
4. Sarma JV, Ward PA (2011) The complement system. Cell and tissue research 343: 227–235.
5. Ricklin D, Lambris JD (2013) Complement in immune and inflammatory disorders: pathophysiological mechanisms. The Journal of Immunology 190: 3831–3838.
6. Walport MJ (2001) Complement. first of two parts. The New England journal of medicine .
7. Pangburn MK, Müller-Eberhard HJ (1984) The alternative pathway of complement. Springer seminars in immunopathology .
8. 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.
9. Blom AM, Kask L, Dahlbäck B (2001) Structural requirements for the complement regulatory activities of c4bp. Journal of Biological Chemistry 276: 27136–27144.
10. Riley-Vargas RC, Gill DB, Kemper C, Liszewski MK, Atkinson JP (2004) Cd46: expanding beyond complement regulation. Trends in immunology 25: 496–503.
11. 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.
12. Liszewski MK, Farries TC, Lublin DM, Rooney IA, Atkinson JP (1995) Control of the

complement system. Advances in immunology 61: 201–283.

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. 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.
19. 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.
20. 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.
21. 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.

22. Morgan BP, Harris CL (2015) Complement, a target for therapy in inflammatory and degenerative diseases. Nature Reviews Drug Discovery .
23. Sagar A, Varner JD (2015) Dynamic modeling of the human coagulation cascade using reduced order effective kinetic models. Processes 3: 178.
24. Sobol I (2001) Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. Mathematics and Computers in Simulation 55: 271 - 280.
25. 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 .
26. Herman J. Salib: Sensitivity analysis library in python (numpy). contains sobol, morris, fractional factorial and fast methods. available online: https://github.com/jdherman/salib.

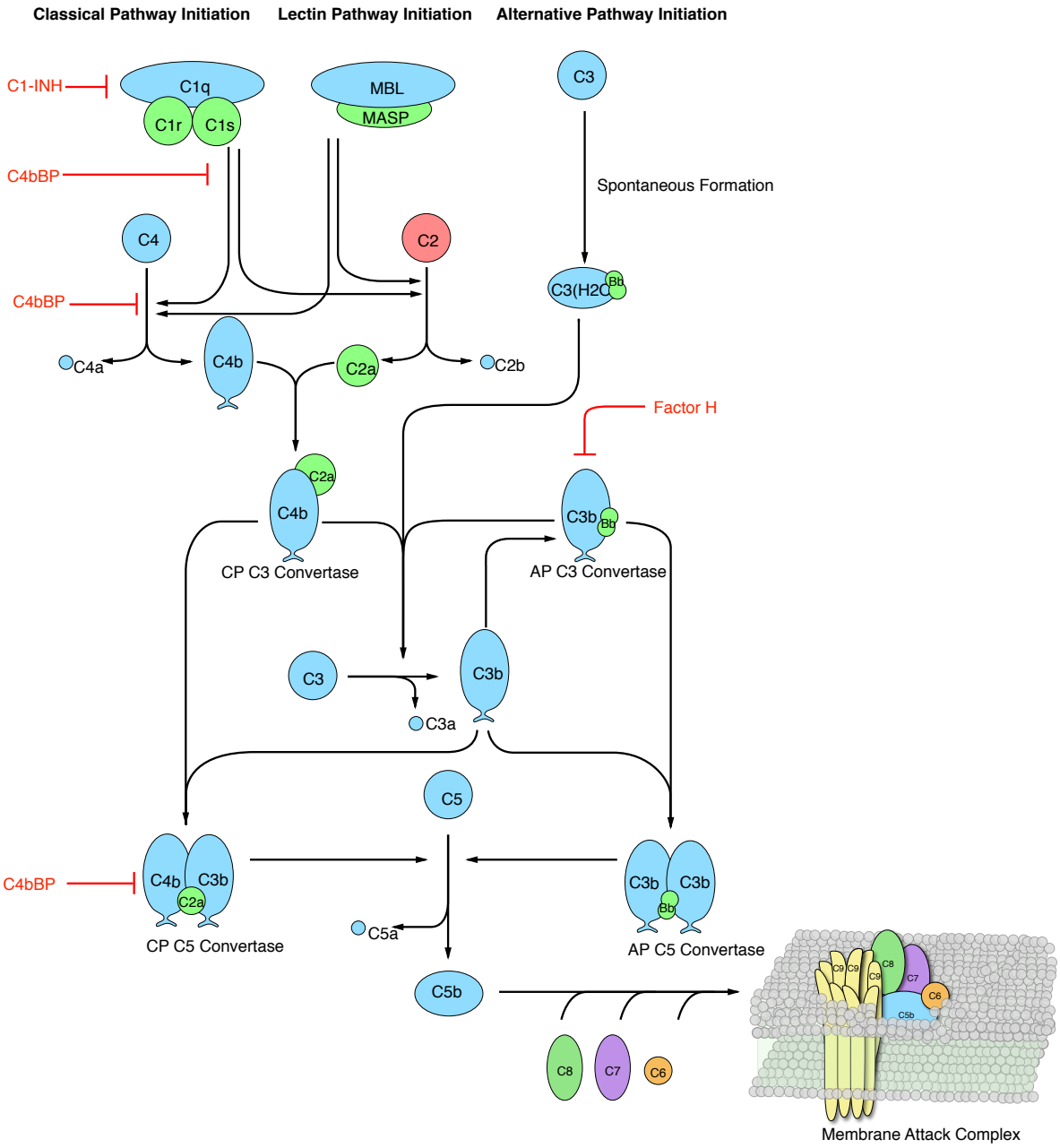

Fig. 1: Simplified schematic of the human complement system. The complement cascade is activated through any one, or more, of the three pathways: the classical, the lectin, and the alternate pathways. The classical pathway is activated by the binding of C1 complex through the C1q subunit to the IgG or IgM immune complex. This binding leads to conformational changes in the C1 complex that leads to the activation of C1r and C1s subunits. Activated C1-antibody complex cleaves C4 and C2 to form the classical C3 convertase. The lectin pathway is initiated by the binding mannose-binding lectins (MBL) and ficolins to carbohydrate moieties on the pathogen surfaces. This results in the formation mannose-binding lectin-associated serine proteases (MASPs). The MBL-MASP complex cleaves C4 and C2 to form the lectin C3 convertase. The alternative pathway is activated through a spontaneous tick-over mechanism by the hydrolysis of C3 to form fluid phase C3 convertase. The C3 convertases cleaves C3 into C3a, and C3b. C3b combines with C4b and C2a to form classical C5 convertase (C4bC3aC3b). The C3b binds with Factor B to form the alternate C5 convertase (C3bBbC3b) . The C5 convertases cleave C5 into C5a, and C5b that undergoes a series of reactions to form the membrane attack complex (MAC).

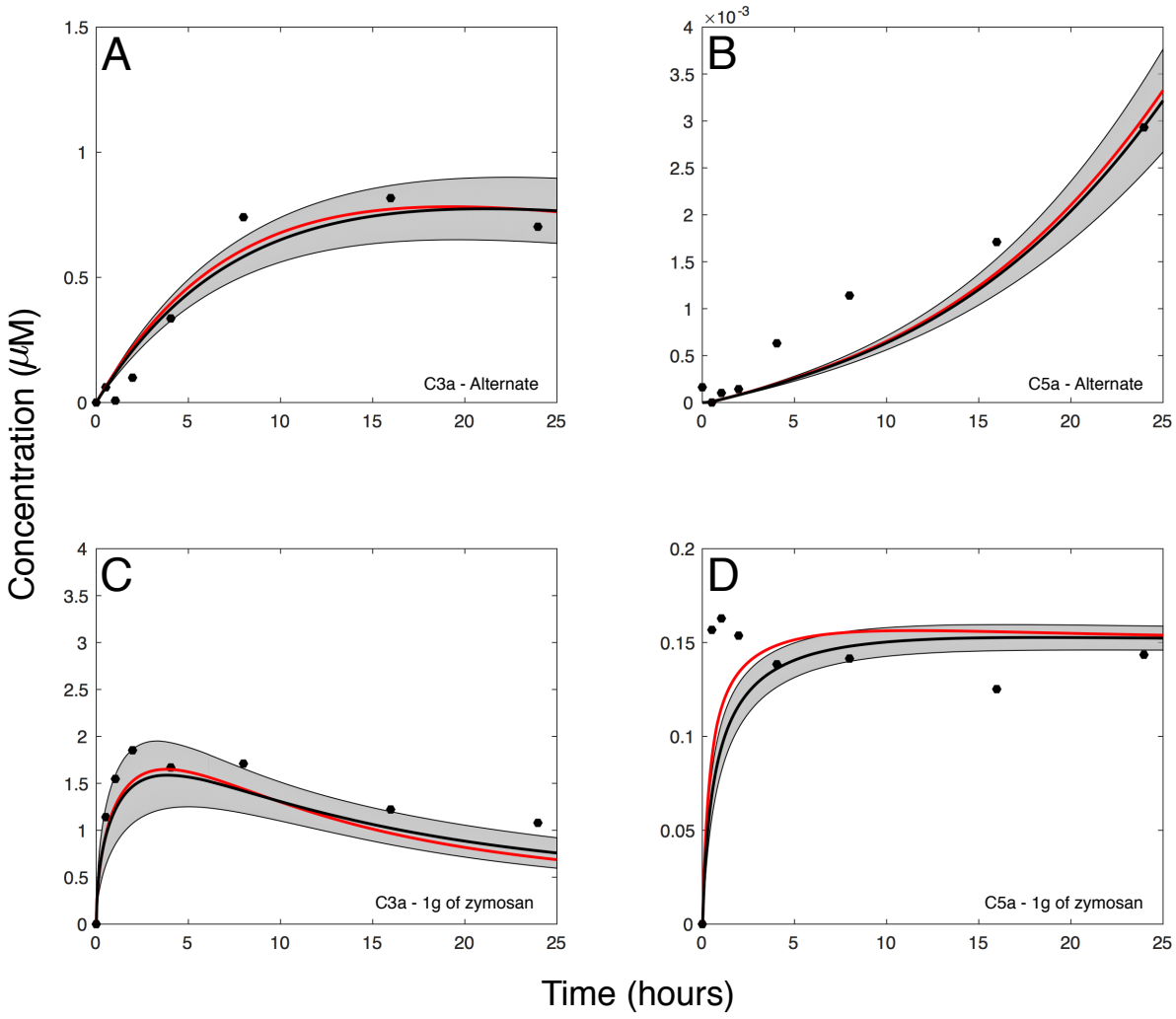

Fig. 2: Reduced order complement model training simulations. Reduced order complement model parameters were estimated using Dynamic Optimization with Particle Swarms (DOPS). The model was trained against experimental data from Shaw and co-workers [17] in the presence and absence of zymosan. The model was trained using C3a and C5a data generated from the alternative pathway (**A**–**B**) and lectin initiated pathway with 1g zymosan (**C**–**D**). The solid red line shows the simulation with the best-fit parameter, the solid black lines show the simulated mean value of C3a or C5a for 50 independent particles. The shaded region denotes 99 % confidence interval on the simulated mean concentration of C3a or C5a (uncertainty in the model simulation). All initial concentrations of complement proteins are at human serum levels unless otherwise noted.

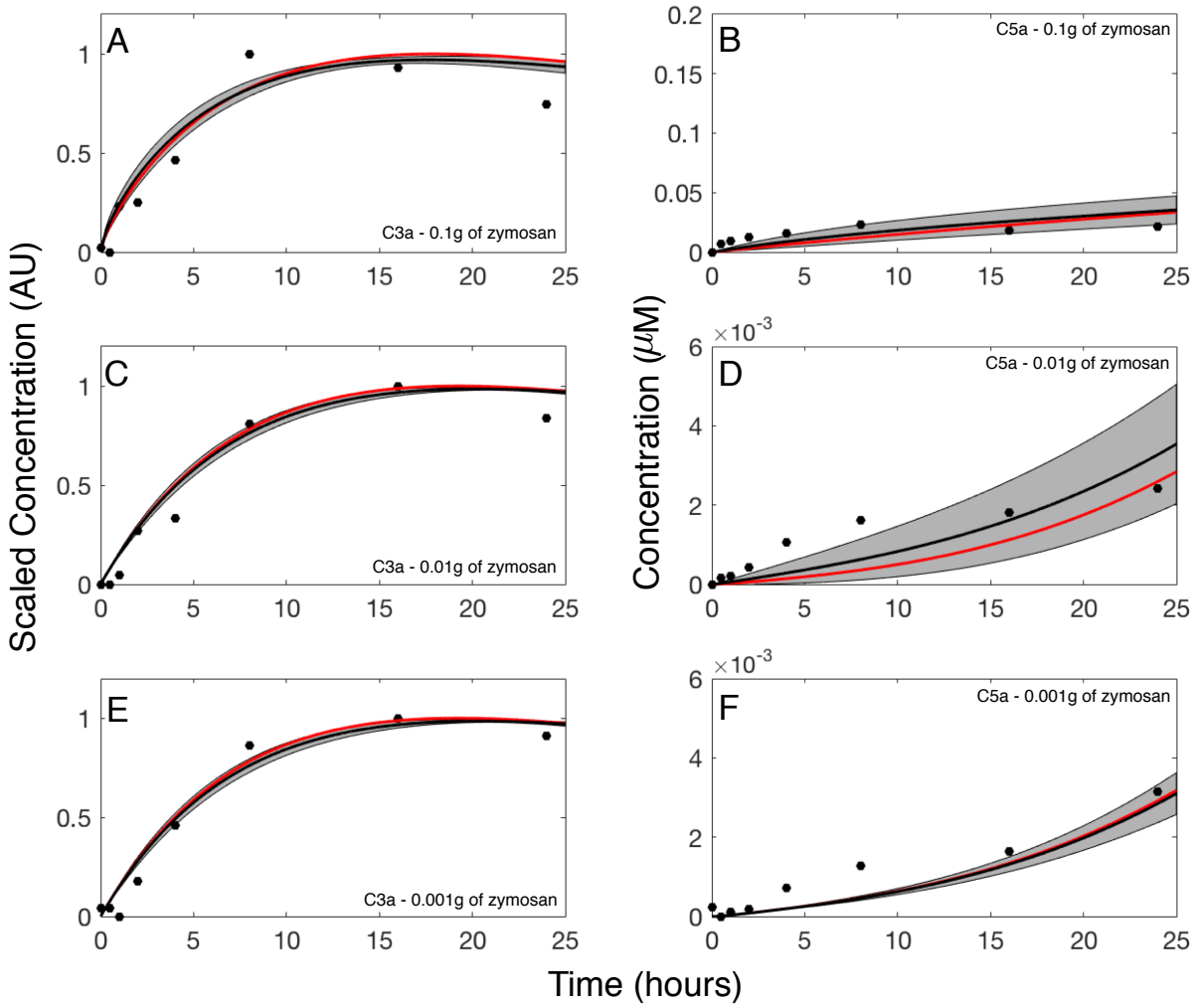

Fig. 3: Reduced order complement model predictions vs experimental data for C3a and C5a generated in the lectin pathway. The reduced order coagulation model parameter estimates were tested against data not used during model training. Simulations of C3a and C5a generated in the lectin pathway using different levels of zymosan (0.1, 0.01, and 0.001 grams of zymosan) were compared with the corresponding experimental data (**A**–**F**). The solid red line shows the simulation with the best-fit parameter, the solid black lines show the simulated mean value of C3a or C5a for 50 independent particles. The shaded region denotes 99 % confidence interval on the simulated mean concentration of C3a or C5a (uncertainty in the model simulation). All initial concentrations of complement proteins are at human serum levels unless otherwise noted.

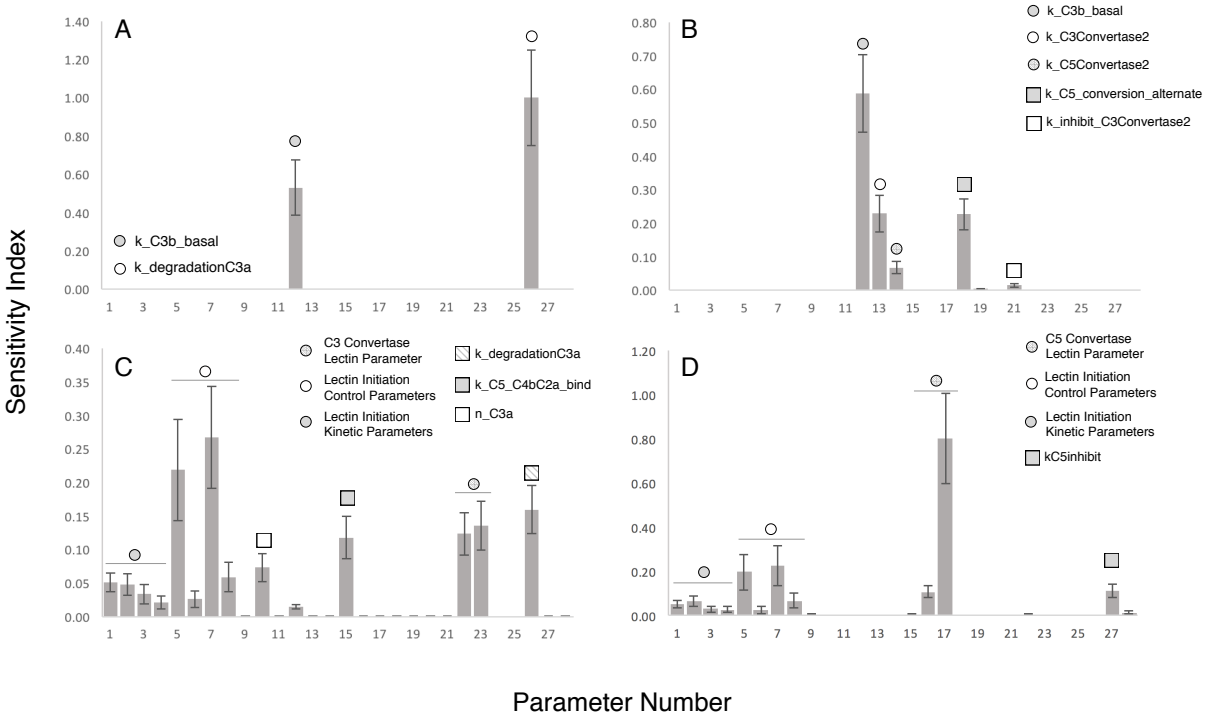

Fig. 4: Sobol's sensitivity analysis of the reduced order complement model with respect to the modeling parameters. Sensitivity analysis was conducted on the four cases we used to train our model: (A) C3a at 0 zymosan, (B) C5a 0 zymosan, (C) C3a 1 g zymosan, and (D) C5a 1 g zymosan. The bars denote total 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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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