Dynamic Modeling of the Human Coagulation Cascade using Reduced Order Effective Kinetic Models

Adithya Sagar and Jeffrey D. Varner*

School of Chemical and Biomolecular Engineering

Cornell University, Ithaca NY 14853

Running Title: Modeling blood coagulation

To be submitted: *Processes*

*Corresponding author:

Jeffrey D. Varner,

Associate 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

Introduction

Developing mathematical models of biochemical networks is a significant facet of systems biology. Modeling approaches differ in their degree of detail, and the choice of approach is often determined by prior knowledge, or model requirements [1]. Ordinary differential equation (ODE) models are a common tool for modeling biochemical systems because of their ability to capture dynamics. However, ODEs often come with difficult (and sometimes impossible) parameter identification problems. For example, Gadkar et al., showed that even with near-perfect information and high frequency sampling, it was often impossible to identify (to within a specified uncertainty) all the parameters in typical signal transduction models [2]. ODE models can also require significant mechanistic knowledge of the underlying biology. However, Bailey suggested more than a decade ago, that achieving qualitative or even quantitative understanding of biological systems should not require complete structural and parametric knowledge [3]. Since Bailey's complex biology with no parameters hypothesis, Sethna and coworkers have shown that model behavior and predictive ability is typically sensitive to only a few parameter combinations, a characteristic seemingly universal to multi-parameter signaling models referred to as *sloppiness* [4]. Thus, accurate model predictions from a sloppy model may be possible, even if the parameters or mechanistic structure is not well understood, if the critical directions in parameter space are properly constrained. Alternatively, data-driven approaches [5] or logical model formulations [6] are emerging paradigms that constrain model structure (and hence the identifiability challenge) by the availability of training and validation data. These are interesting techniques because the data naturally informs model structure, which can be expanded as more data becomes available. While not mechanistic, these models can simulate surprisingly rich system responses in the absence of mechanism [REFHERE]. However, modeling dynamics using logical models is not straightforward. Thus, a potential third approach could be a combination of ODEs and logical models, where mechanistic in-

1

formation was used when available, and missing or incomplete knowledge encoded using causal local rules.

In this study, we developed a hybrid approach which combined ODEs and logical rules to model a well studied biochemical network, the human coagulation system. Coagulation is an archetype proteolytic cascade involving both positive and negative feedback [7–9]. Coagulation is mediated by a family proteases in the circulation, called factors and a key group of blood cells, called platelets. The central process in coagulation is the conversion of prothrombin (fII), an inactive coagulation factor, to the master protease thrombin (FIIa). Thrombin generation involves three phases, initiation, amplification and termination [10, 11]. Initiation requires a trigger event, for example vessel injury, which leads to the activation of factor VII (FVIIa). Two converging pathways, the extrinsic and intrinsic cascades, then process and amplify this initial coagulation signal. The extrinsic cascade is generally believed to be the main mechanism of thrombinogenesis in the blood [12–14]. Initially, thrombin is produced upon cleavage of prothrombin by fluid phase activated factor X (FXa), which itself has been activated by FVIIa [15]. Picomolar amounts of thrombin then activate the cofactors factors V and VIII (fV and fVIII) and platelets, leading to the formation of the tenase and prothrombinase complexes on activated platelets. These complexes amplify the early coagulation signal by further activating FXa, and directly converting prothrombin to thrombin. There are several control points in the cascade that inhibit thrombin formation, and eventually terminate thrombin generation. Tissue Factor Pathway Inhibitor (TFPI) inhibits FXa formation catalyzed by TF/FVIIa, while antithrombin III (ATIII) neutralizes several of the proteases generated during coagulation, including thrombin. Thrombin itself also inadvertently plays a role in its own inhibition; thrombin, through interaction with thrombomodulin, protein C and endothelial cell protein C receptor (EPCR), converts protein C to activated protein C (APC) which attenuates the coagulation response by proteolytic cleavage of fV/FVa and fVIII/FVIIIa. Termination occurs after

2

either prothrombin is consumed, or thrombin formation is neutralized by inhibitors such as activated protein C (APC) or antithrombin III (ATIII).

There are many mathematical models of the human coagulation cascade in the literature. Previous coagulation models have typically been formulated as systems of nonlinear ordinary differential equations, often using mass-action kinetics to describe the rates of biochemical conversions [REFHERE]. Previous ODE coagulation models from our laboratory [16, 17], were built upon the pioneering studies of Jones and Mann [18], and later Hockin *et al.* [19], who developed and subsequently refined a highly mechanistic model of the extrinsic coagulation cascade. Other laboratories have also expanded upon Hockin *et al.*, for example Chatterjee *et al.,* augmented the Hockin model to include the intrinsic pathway, and the role of thrombin in clot formation [20]. More recently, modeling studies have explored other biochemical or biophysical aspects of coagulation, such as the biochemical networks inside platelets [21], multi-scale models of clot formation [22, 23], and the transport of material inside growing clots [24]. However, most of these previous studies share the common feature of a mechanism based description of coagulation. Such a description is possible because blood, while enormously complex, can be systematically interrogated. However, other biochemical networks, such as the intracellular networks which mediate cancer, are more difficult to probe and model. Towards this need, we have proposed a hybrid approach which combined ODEs and logical rules that can be used to model biochemical processes where mechanistic information is difficult or impossible to obtain. We have used this approach to model the coagulation cascade, because of the wealth of training and validation data available for this system. Our hypothesis is that coagulation represents a best case, if our approach fails in this proof-of-concept study, it will never be acceptable in model difficult cases. We proposed a reduced order coagulation model that consisted of five differential equations augmented with several logical rules. We identified the reduced order coagulation model parameters from in vitro extrinsic data

3

sets of thrombin generation initiated by TF/FVIIa in the presence of ATIII, with and without the protein C pathway. We then validated the model using thrombin data sets from multiple laboratories, for both normal and hemophilia coagulation. The reduced order hybrid approach produced a surprisingly predictive coagulation model, suggesting this framework could potentially be used to model other biochemical networks important to human health.

Results

Formulation of reduced order coagulation models. We developed a reduced order coagulation model to test our hybrid kinetic modeling approach (Fig. 1). A trigger event initiates thrombin formation (FIIa) from prothrombin (fII) through a lumped initiation step. This step loosely represents the initial fluid phase activation of thrombin by activated FXa. Once activated, thrombin catalyzes its own formation (amplification step), and inhibition via the conversion of protein C to activated protein C (APC). APC and tissue factor pathway inhibitor (TFPI) inhibit initiation, while antithrombin III (ATIII) inhibits amplification. All initiation and inhibition processes, as well as the dependence of amplification upon other coagulation factors, was approximated using our rule-based approach (Fig. 2). Individual regulatory contributions to the activity of pathway enzymes were integrated into control coefficients using an integration rule. These control coefficients then modified the rates of model processes at each time step. Hill-like transfer functions $0 \leq f(\mathcal{Z}) \leq 1$ quantified the influence of components upon a target process. Components were either individual inhibitor or activator levels or some function of levels, e.g., the product of factor levels. In this study, $\mathcal{Z}$ corresponded to the abundance of individual inhibitors or activators, with the exception of the dependence of amplification upon specific coagulation factors (modeled as the product of factors). When a process was potentially sensitive to multiple inputs, logical integration rules were used to select which transfer functions influenced the process at any given time. In our proof of concept model, we used a winner takes all strategy; the maximum or minimum transfer function was selected at any given time step. However, other integration rules are certainly possible. Taken together, while the reduced order coagulation model encodes significant biological complexity, it is highly compact, consisting of only five differential equations. Thus, it will serve as an excellent proof of principle example to study the reduction of a highly complex human subsystem.

Identification of model parameters using particle swarm optimization. A critical

challenge for any dynamic model is the estimation of kinetic parameters. We estimated

kinetic and control parameters simultaneously from eight *in vitro* time-series coagulation

data sets with and without the protein C pathway. The residual between model simula-

tions and experimental measurements was minimized using particle swarm optimization

(PSO). A population of particles (N = 20) was initialized with randomized kinetic and con-

trol parameters and allowed to search for parameter vectors that minimized the residual.

However, not all parameters were varied simultaneously. We partitioned the parame-

ter estimation problem into two subproblems based upon the biological organization of

the training data; (i) estimation of parameters associated with thrombin formation in the

absence of the protein C pathway and (ii) estimation of parameters associated with the

protein C pathway. Only those parameters associated with each subproblem were varied

during the optimization procedure for that subproblem e.g., thrombin parameters were *not*

varied during the protein C subproblem. The PSO procedure was run for 20 generations

for each subproblem, where each generation was 1200 iterations. The best particle from

each generation was used to generate the particle population for the next generation. We

rotated the subproblems, starting with subproblem 1 in the first generation.

The reduced order coagulation model captured the role of initial prothrombin abun-

dance, and the decay of the thrombin signal following from ATIII activity (Fig. 3). However,

we systematically under-predicted the thrombin peak and the strength of ATIII inhibition

in this training data set. On the other hand, with fixed thrombin parameters, we captured

peak thrombin values and the decay of the thrombin signal (at least for the 150% fII case)

in the presence of both ATIII and the protein C pathway (Fig. 4). Lastly, we were unable

to capture global differences in initiation time *across* separate data sets with a single en-

semble of model parameters. These differences likely result from normal experimental

variability. For example different thrombin generation experiments within our training data

6

(at the same physiological factor levels) have significantly different initiation times (supplemental results). However, this also highlights a potential shortcoming of the initiation module within the model. To capture the variability in initiation time *across* training data sets, we included a constant time-delay parameter (T_D) for each data group. The delay parameter was constant within a data set, but allowed to vary *across* training data sets. Introduction of the delay parameter allowed the model to simulate multiple training data sets using a single ensemble of model parameters. Taken together, the model identification results suggested that our kinetic-rules based approach could reproduce a panel of thrombin generation data sets conducted at physiological factor and inhibitors concentrations. However, it was unclear whether the reduced order model could predict new data, without updating the model parameters.

Validation of the reduced order coagulation model. We tested the predictive power of the reduced order coagulation model with validation data sets not used during model training. Two validation data sets were used, thrombin generation for various prothrombin and ATIII concentrations with the protein C pathway, and thrombin generation in normal versus hemophilic plasma in the presence of the protein C pathway. Lastly, we compared the qualitative output of the model to rFVIIa addition in the presence of hemophilia. The hemophilia case was an especially difficult test as it was taken from a different study which used a plasma-based *in vitro* assay involving platelets instead of phospholipid vesicles (PCPS). All kinetic and control parameters were fixed for the validation simulations. The only globally adjustable parameter T_D, was fixed within each validation data set but allowed to vary between data sets. The reduced order model predicted the thrombin generation profile for ratios of prothrombin and ATIII in the absence of the protein C pathway (Fig. 5). Simulations near the physiological range (fII,ATIII) = (100%, 100%) or (125%,75%) tracked the measured thrombin values (Fig. 5B and C). On the other hand, predictions for factors levels outside of the physiological range (fII,ATIII) = (50%, 150%)

7

162 or (150%, 50%), while qualitatively consistent with measured thrombin values, did show

163 significant deviation from the measurements (Fig. 5A and D). Likewise, simulations of

164 thrombin generation in normal versus hemophilia (missing both fVIII and fIX) were con-

165 sistent with measured thrombin values (Fig. 6). We modeled the dependence of thrombin

166 amplification on factor levels using a product rule ($\mathcal{Z} = fV \times fX \times fVIII \times fIX$), which

167 was then was integrated using a $\min$ integration rule into the control variable governing

168 amplification. Thus, in the absence of fVIII or fIX, the amplification control variable evalu-

169 ated to zero, and the only thrombin produced was from initiation (Fig. 6B). However, the

170 decay of the thrombin signal was underpredicted in the normal case (Fig. 6A), while the

171 activated thrombin level was overpredicted in hemophilia simulations, although thrombin

172 generation was far less than normal (Fig. 6B). Taken together, the reduced order model

173 performed well in the physiological range of factors, even with unmodeled components

174 such as platelet activation in the hemophilia data set. Thus, our kinetic-rules framework

175 predicted the output of a physiologically important cascades such as coagulation, despite

176 significant unmodeled components.

177   The model ensemble predicted a direct correlation between thrombin generation and

178 rFVIIa addition in hemophilia (Fig. 7). In the current model, we cannot distinguish be-

179 tween different initiation sources, e.g., TF/FVIIa and rFVIIa, as we have only a single

180 lumped initiation source (trigger). Thus, we simulated the addition of rFVIIa in hemophilia

181 by removing fVIII and fIX from the model, and modulating the initial level of trigger. Sim-

182 ulations with a baseline level of trigger were consistent with the previous hemophilia sim-

183 ulations, where the only thrombin being produced was from initiation (Fig. 7, 1$\times$ trigger).

184 As we increased the strength of the trigger event, the thrombin peak time and the maxi-

185 mum value of thrombin increased (Fig. 7, 50$\times$ trigger). However, as the trigger strength

186 increased, the thrombin generated quickly decayed. For example, for the highest trigger

187 strength simulated (50$\times$trigger), 95% of the thrombin was gone by 20 min after initiation.

8

We performed flux analysis to understand how the reduced order coagulation model balanced initiation, amplification and inhibition of thrombin formation for normal coagulation and hemophilia.util.. Analysis of the reaction flux through the reduced order network for thrombin generation in normal, hemophilia and rFVIIa-treated hemophilia identified three distinct operational modes (Fig. 8). The reduced order network includes four lumped reactions, initiation, amplification, thrombin-induced APC generation and total thrombin inhibition (including both APC and ATIII action). Directly after the addition of a trigger (e.g., TF/FVIIa or rFVIIa), the lumped initiation flux was the largest for all three cases. However, within a few minutes enough thrombin was generated by the initiation mechanism to induce the amplification stage. During amplification, thrombin catalyzes its own formation and inhibition by generating activated protein C (APC), a potent inhibitor of the coagulation cascade. For normal coagulation, amplification and thrombin inhibition are the dominate reactions by 6 min after initiation (Fig. 8, left). After 10 min, the dominate reaction has shifted to thrombin inhibition (both ATIII and APC action). In the current proof-of-principle model, APC inhibits upstream of amplification thus APC activity will slow the rate of thrombin formation directly from the initiation trigger. On the other hand, ATIII inhibits thrombin directly, as well as upstream coagulation factors. In hemophilia (missing both fVIII and fIX), the amplification reaction does not occur and the only thrombin produced is from initiation (Fig. 8, center). Initiation is quickly inhibited by APC, and the thrombin level stabilizes and eventually decays because of ATIII activity. Lastly, when $50\times$rFVIIa is used to induce thrombin formation in hemophilia (absence of fVIII/fIX), initiation mechanisms dominate for up to 6 min following initiation (Fig. 8, right). However, similar to hemophilia alone, without amplification the thrombin signal is quickly extinguished by the combined action of ATIII and APC generated by thrombin.

9

Sensitivity analysis of the reduced order coagulation model. We conducted a global sensitivity analysis to estimate which parameters controlled the performance of the reduced order model. We calculated the sensitivity of the time to maximum thrombin (peak time) and the thrombin exposure (area under the thrombin curve) for different levels of prothrombin, and protein C (Fig. 9). Globally, 41% of the parameters shifted in importance between the (fII,PC) = (50%, 0%) and (150%,100%) cases for the peak thrombin time (Fig. 9A). The majority of these shifts were involved increased prothrombin and the protein C pathway, while only 5% were directly associated with increased prothrombin alone. The rate constant for thrombin amplification was the most important parameter controlling the peak thrombin time. While this parameter was differentially important for different prothrombin levels, and in the presence or absence of the activated protein C pathway, it's sensitivity was consistently above all other parameters in the model. The saturation constant governing thrombin amplification was the second most important parameter, followed by the initiation control gain parameter. Other important parameters influencing the thrombin peak time included the control gain for activated protein C formation, as well as the rate constant controlling ATIII inhibition of thrombin activity. On the other hand, only 27% of the model parameters were differentially sensitive between the (%fII,%PC) (50%, 0%) and (150%,100%) cases for the thrombin exposure (Fig. 9B). Of these parameters, all of the shifts were associated the interplay between thrombin formation and the protein C pathway. The rate constant controlling ATIII inhibition was the most important parameter controlling the thrombin exposure. While this parameter was less important in the presence of protein C for 150% prothrombin levels, it was significantly above all other parameters. Similar to the peak time, for 150% prothrombin, the control gain for activated protein C formation was differentially important along with the rate constant controlling amplification. However, the influence of the amplification parameter was much less in the thrombin exposure case versus peak time.

10

Discussion

Controlling hemorrhage has been a driving force to understand the molecular basis of blood coagulation. Replacement therapy improves bleeding times for most hemophilia patients. However, repeated fVIII/fIX administration can lead to neutralizing inhibitors. Alternatively, rFVIIa has been used as a universal hemostatic agent to initiate clotting with and without fVIII or fIX inhibitors [25]. Despite its wide adoption, the mechanism and the utility of rFVIIa in trauma-associated hemorrhage remains controversial [**?**]. In this study, we formulated a model of the human coagulation cascade and studied thrombin formation in normal and rFVIIa treated and untreated hemophilic plasma. The model described 193 proteins or protein complexes connected by 301 interactions. The coagulation network architecture was based on literature and a previous mathematical model from our laboratory [**?**]. Simulations of thrombin dynamics were independently validated using blood drawn from patients with coronary artery disease (reported here) and previous TF/FVIIa and FVIIa *in-vitro* studies [**?**]. While the human coagulation cascade is perhaps an ideal model system to develop network analysis tools, not all the model parameters were identifiable given the current training data. Instead, an *ensemble* of models (N = 437) was estimated using nine cell-based coagulation training sets [**?**]. We used this family of consistent coagulation models in all simulations instead of a single best-fit but uncertain model. Additionally, using sensitivity analysis on the ensemble of models, we identified structurally sensitive components of the coagulation architecture as a function of condition. We also demonstrated that the ensemble robustly constrained model predictions of independent validation sets, despite having many poorly constrained parameters. Thus, the model ensemble displayed *sloppy* behavior similar to that observed by Sethna and coworkers for other signal transduction networks [**?**]. By studying a family of models, we perhaps partially addressed the uncertainty stemming from the many poorly characterized model parameters. However, many other factors could influence our results. For

11

example, missing structural interactions or biophysical factors such as blood flow could play a large role. Thus, while the results presented here may be a valuable first step, more studies are required.

Malfunctions in coagulation can have serious or potentially even fatal consequences. For example, aggressive clotting is involved with Coronary Artery Diseases (CADs), which collectively account for 38% of all deaths in North America [26]. Coagulation management during surgery can also be challenging, particularly because of the increasing clinical use of antithrombotic drugs [27]. Insufficient coagulation due to genetic disorders such as hemophilia can also result in recurrent bleeding. The coagulation factors VIII (fVIII) and IX (fIX) are deficient in Hemophilia A and B, respectively [28–30]. People with mild hemophilia have 5-40% of the normal clotting factor levels while severe hemophiliacs have <1% [30]. Hemophilia can be controlled with regular infusions of the deficient clotting factors. However, clotting factor replacement sometimes leads to the formation of *in-vivo* fVIII and fIX inhibitors [31]. Activated Prothrombin Complex Concentrates (aPCCs) from pooled plasma have been used to treat hemophilia. However, aPCCs can have safety issues [32–34]. Alternatively, recombinant factor VIIa (rFVIIa) has been used to treat bleeding disorders [25, 35] including hemophilia with and without factor VIII/IX inhibitors [36–39]. However, rFVIIa requires frequent administration (every 2-3 hr) because of its short half-life in the circulation. Many questions also remain about its mechanism [15, 40**? ? ?** –43] and effective dose range [31]. Thus, despite its wide adoption, the utility of rFVIIa in trauma-associated hemorrhage remains controversial [**?**].

While the performance of the proof of principle reduced order model was impressive, there are several questions that should be explored further. [FINISH]

Materials and Methods

Formulation and solution of the model equations.   We used ordinary differential equations (ODEs) to model the time evolution of proteins (x_i) in our reduced order coagulation 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. The system material balances were subject to the initial conditions $\mathbf{x}\left(t_o\right) = \mathbf{x}_o$, which were specified by the experimental setup.

Each reaction rate was written as the product of two terms, a kinetic term ($\bar{r}_j$) and a regulatory term (v_j):

$$r_j\left(\mathbf{x}, \epsilon, \mathbf{k}\right) = \bar{r}_j v_j \tag{2}$$

We used multiple saturation kinetics to model the reaction term $\bar{r}_j$:

$$\bar{r}_j = k_j^{max} \epsilon_i \left(\prod_{s \in m_j^-} \frac{x_s}{K_{js} + x_s} \right) \tag{3}$$

where k_j^{max} denotes the maximum rate for reaction j, ϵ_i denotes the scaled enzyme activity which catalyzes reaction j, and K_{js} denotes the saturation constant for species s in reaction j. The product in Eqn. (3) was carried out over the set of *reactants* for reaction j (denoted as m_j^-).

The control term v_j depended upon the combination of factors which influenced the activity of enzyme i. For each enzyme, we used a rule-based approach to select from competing control factors (Fig. 2). If an enzyme was activated by m metabolites, we modeled this activation as:

$$v_j = \max\left(f_{1j}\left(\mathcal{Z}\right), \ldots, f_{mj}\left(\mathcal{Z}\right)\right) \tag{4}$$

where $0 \leq f_{ij}\left(\mathcal{Z}\right) \leq 1$ was a regulatory transfer function that calculated the influence of metabolite i on the activity of enzyme j. Conversely, if enzyme activity was inhibited by a m metabolites, we modeling this inhibition as:

$$v_j = 1 - \max\left(f_{1j}\left(\mathcal{Z}\right), \ldots, f_{mj}\left(\mathcal{Z}\right)\right) \tag{5}$$

Lastly, if an enzyme had both m activating and n inhibitory factors, we modeled the regulatory term as:

$$v_j = \min\left(u_j, d_j\right) \tag{6}$$

where:

$$u_j = \max_{j^+}\left(f_{1j}\left(\mathcal{Z}\right), \ldots, f_{mj}\left(\mathcal{Z}\right)\right) \tag{7}$$

$$d_j = 1 - \max_{j^-}\left(f_{1j}\left(\mathcal{Z}\right), \ldots, f_{nj}\left(\mathcal{Z}\right)\right) \tag{8}$$

The quantities j^+ and j^- denoted the sets of activating and inhibitory factors for enzyme j. If a process has no modifying factors, we set $v_j = 1$. There are many possible functional forms for $0 \leq f_{ij}\left(\mathcal{Z}\right) \leq 1$. However, in this study, each individual transfer function took the form:

$$f_i\left(\mathbf{x}\right) = \frac{\kappa_{ij}^{\eta} \mathcal{Z}_j^{\eta}}{1 + \kappa_{ij}^{\eta} \mathcal{Z}_j^{\eta}} \tag{9}$$

14

where $\mathcal{Z}_j$ denotes the abundance of the j factor (e.g., metabolite abundance), and κ_{ij} and η are control parameters. The κ_{ij} parameter was species gain parameter, while η was a cooperativity parameter (similar to a Hill coefficient). Applying the general framework to the reduced coagulation network resulted in five ordinary differential equations:

$$\frac{dx_1}{dt} = -\left(r_{init}v_{init} + r_{amp}v_{amp}\right) \tag{10}$$

$$\frac{dx_2}{dt} = r_{amp}v_{amp} + r_{init}v_{init} - r_{inh,ATIII}v_{inh,ATIII} \tag{11}$$

$$\frac{dx_3}{dt} = -r_{apc}v_{apc} \tag{12}$$

$$\frac{dx_4}{dt} = r_{apc}v_{apc} \tag{13}$$

$$\frac{dx_5}{dt} = -r_{inh,ATIII}v_{inh,ATIII} \tag{14}$$

where $\mathbf{x} = (fII, FIIa, PC, APC, ATIII)^T$. The terms r_*v_* in the balance equations denote corrected kinetic expressions for initiation, amplification and inhibition processes. The rate of initiation $\bar{r}_{init}$ was modeled as:

$$\bar{r}_{init} = k_{init}\left(trigger\right) \frac{x_1}{K_{init,fII} + x_1} \tag{15}$$

where $k_{init}, \; K_{init,fII}$ are the rate and saturation constants governing initiation, respectively. The rate of initiation was modified by v_{init}, the control parameter governing initiation. Initiation was sensitive to level of trigger (activator), and TFPI (inhibitor):

$$v_{init} = \min\left(f_{init}^-\left(TFPI\right), f_{init}^+\left(trigger\right)\right) \tag{16}$$

where the transfer functions f took the form of Eqn (9). The rate of thrombin amplification was given by:

$$\bar{r}_{amp} = k_{amp}\left(x_2\right) \frac{x_1}{K_{amp,fII} + x_1} \tag{17}$$

15

where k_{amp}, $K_{amp,fII}$ denote the rate and saturation constants governing amplification, respectively. The amplification control term, which modified amplification rate, was modeled as a combination of multiple inhibition terms and one activation term:

$$v_{amp} = \min \left(f_{amp}^{-} \left(TFPI \right), f_{amp}^{-} \left(x_4 \right), f_{amp}^{+} \left(\mathscr{Z}_{amp} \right) \right) \tag{18}$$

where $\mathscr{Z}_{amp} = fV \times fX \times fVIII \times fIX$. Although, $f_{amp}^{+} \left(\mathscr{Z}_{amp} \right)$ is an activating term, we included it in the min integration rule; the factors in $\mathscr{Z}_{amp}$ were essential for amplification (if any of these factors was missing the amplification reaction would not occur). Thus, the factors in $\mathscr{Z}_{amp}$ were required components, a classification that we implemented by the min selection rule. The rate activated protein C formation was given by:

$$\bar{r}_{apc} = k_{APC,formation} \left(TM \right) \frac{x_3}{K_{formation,PC} + x_3} \tag{19}$$

where $k_{APC,formation}$, $K_{formation,PC}$ denote the rate and saturation constants governing activated protein C formation, respectively and TM denotes the thrombomodulin abundance. We modeled the control term which governed APC formation as a single thrombin dependent activation term:

$$v_{apc} = \max \left(f_{apc}^{+} \left(x_2 \right) \right) \tag{20}$$

Lastly, we included direct irreversible inhibition of FIIa by ATIII:

$$\bar{r}_{inh,ATIII} = k_{ATIII,inhibition} \left(x_5 x_2^{\gamma} \right) \tag{21}$$

where γ was estimated to be $\gamma = 1.26$. For ATIII inhibition of FIIa, the control variables $v_{inh,ATIII}$ was taken to be unity. The model equations were encoded using the Python programming language and solved using the ODEINT routine of the SciPy module [44].

The model files can be downloaded from http://www.varnerlab.org.

Estimation of model parameters from experimental data. Model parameters were estimated by minimizing the difference between simulations and experimental thrombin measurements (squared residual):

$$\min_{\mathbf{k}} \sum_{\tau=1}^{\mathcal{T}} \sum_{j=1}^{\mathcal{S}} \left(\frac{\hat{x}_j \left(\tau \right) - x_j \left(\tau, \mathbf{k} \right)}{\omega_j \left(\tau \right)} \right)^2 \tag{22}$$

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 respect to time, while the inner summation is with respect to state. We minimized the model residual using Particle swarm optimization (PSO) [45]. PSO uses a *swarming* metaheuristic to explore parameter spaces. A strength of PSO is its ability to find the global minimum, even in the presence of potentially many local minima, by communicating the local error landscape experienced by each particle collectively to the swarm. Thus, PSO acts both as a local and a global search algorithm. For each iteration, particles in the swarm compute their local error by evaluating the model equations using their specific parameter vector realization. From each of these local points, a globally best error is identified. Both the local and global error are then used to update the parameter estimates of each particle using the rules:

$$\begin{align} \mathbf{\Delta}_i &= \theta_1 \mathbf{\Delta}_i + \theta_2 \mathbf{r}_1 \left(\mathcal{L}_i - \mathbf{k}_i \right) + \theta_3 \mathbf{r}_2 \left(\mathcal{G} - \mathbf{k}_i \right) \tag{23} \\ \mathbf{k}_i &= \mathbf{k}_i + \mathbf{\Delta}_i \tag{24} \end{align}$$

where $\left(\theta_1, \theta_2, \theta_3 \right)$ are adjustable parameters, $\mathcal{L}_i$ denotes local best solution found by particle i, and $\mathcal{G}$ denotes the best solution found over the entire population of particles.

The quantities r_1 and r_2 denote uniform random vectors with the same dimension as the number of unknown model parameters ($\mathcal{K} \times 1$). In thus study, we used $(\theta_1, \theta_2, \theta_3) = (1.0, 0.05564, 0.02886)$. The quality of parameter estimates was measured using goodness of fit (model residual). The particle swarm optimization routine was implemented in the Python programming language. All plots were made using the Matplotlib module of Python [46]. The optimization code can be downloaded from http://www.varnerlab.org.

Global sensitivity analysis of model performance. We conducted a global sensitivity analysis to estimate which parameters controlled the performance of the reduced order model. We computed the total sensitivity index of each parameter relative to two performance objectives, the peak thrombin time and the area under the thrombin curve (thrombin exposure). [FINISH ME].

Acknowledgements

References

1. Kholodenko B, Yaffe MB, Kolch W (2012) Computational approaches for analyzing information flow in biological networks. Sci Signal 5: re1.

2. Gadkar KG, Varner J, Doyle FJ (2005) Model identification of signal transduction networks from data using a state regulator problem. Syst Biol (Stevenage) 2: 17–30.

3. Bailey JE (2001) Complex biology with no parameters. Nat Biotechnol 19: 503-4.

4. Machta BB, Chachra R, Transtrum MK, Sethna JP (2013) Parameter space compression underlies emergent theories and predictive models. Science 342: 604-7.

5. Cirit M, Haugh JM (2012) Data-driven modelling of receptor tyrosine kinase signalling networks quantifies receptor-specific potencies of pi3k- and ras-dependent erk activation. Biochem J 441: 77-85.

6. Morris MK, Saez-Rodriguez J, Clarke DC, Sorger PK, Lauffenburger DA (2011) Training signaling pathway maps to biochemical data with constrained fuzzy logic: quantitative analysis of liver cell responses to inflammatory stimuli. PLoS Comput Biol 7: e1001099.

7. Butenas S, Mann KG (2002) Blood coagulation. Biochemistry (Mosc) 67: 3-12.

8. Schenone M, Furie BC, Furie B (2004) The blood coagulation cascade. Curr Opin Hematol 11: 272-7.

9. Adams RLC, Bird RJ (2009) Review article: Coagulation cascade and therapeutics update: relevance to nephrology. part 1: Overview of coagulation, thrombophilias and history of anticoagulants. Nephrology (Carlton) 14: 462-70.

10. Goldhaber SZ, Colman RW, Clowes AW, editors (2006) Hemostasis and Thrombosis: Basic Principles and Clinical Practice. Lippincott Williams and Wilkins.

11. Brummel KE, Paradis SG, Butenas S, Mann KG (2002) Thrombin functions during tissue factor-induced blood coagulation. Blood 100: 148-152.

12. Mann K, Nesheim M, Church W, Haley P, Krishnaswamy S (1990) Surface-dependent

406   reactions of vitamin k-dependent enzyme complexes. Blood 76: 1-16.

407 13. Roberts H, Monroe D, Oliver J, Chang J, Hoffman M (1998) Newer concepts of blood

408   coagulation. Haemophilia 4: 331-334.

409 14. Mann K (1999) Biochemistry and physiology of blood coagulation. Thromb Haemost

410   82: 165-174.

411 15. Butenas S, Brummel KE, Branda RF, Paradis SG, Mann KG (2002) Mechanism of

412   factor VIIa-dependent coagulation in hemophilia blood. Blood 99: 923-930.

413 16. Luan D, Zai M, Varner JD (2007) Computationally derived points of fragility of a human

414   cascade are consistent with current therapeutic strategies. PLoS Comput Biol 3:

415   e142.

416 17. Luan D, Szlam F, Tanaka KA, Barie PS, Varner JD (2010) Ensembles of uncertain

417   mathematical models can identify network response to therapeutic interventions. Mol

418   Biosyst 6: 2272-86.

419 18. Jones KC, Mann KG (1994) A model for the tissue factor pathway to thrombin. ii. a

420   mathematical simulation. J Biol Chem 269: 23367-73.

421 19. Hockin MF, Jones KC, Everse SJ, Mann KG (2002) A model for the stoichiometric

422   regulation of blood coagulation. J Biol Chem 277: 18322-33.

423 20. Chatterjee MS, Denney WS, Jing H, Diamond SL (2010) Systems biology of coagu-

424   lation initiation: kinetics of thrombin generation in resting and activated human blood.

425   PLoS Comput Biol 6.

426 21. Stalker TJ, Traxler EA, Wu J, Wannemacher KM, Cermignano SL, et al. (2013) Hier-

427   archical organization in the hemostatic response and its relationship to the platelet-

428   signaling network. Blood 121: 1875-85.

429 22. Leiderman K, Fogelson A (2014) An overview of mathematical modeling of thrombus

430   formation under flow. Thromb Res 133 Suppl 1: S12-4.

431 23. Bannish BE, Keener JP, Fogelson AL (2014) Modelling fibrinolysis: a 3d stochastic

432 multiscale model. Math Med Biol 31: 17-44.

24. Voronov RS, Stalker TJ, Brass LF, Diamond SL (2013) Simulation of intrathrombus fluid and solute transport using in vivo clot structures with single platelet resolution. Ann Biomed Eng 41: 1297-307.

25. Hedner U (2008) Factor viia and its potential therapeutic use in bleeding-associated pathologies. Thromb Haemost 100: 557–562.

26. GKHansson (2005) Inflammation, Atherosclerosis and Coronary Artery Disease. N Engl J Med 352: 1685 - 1695.

27. Tanaka KA, Key NS, Levy JH (2009) Blood coagulation: hemostasis and thrombin regulation. Anesth Analg 108: 1433-46.

28. Tuddenham E, Cooper D (1994) The molecular genetics of haemostasis and its inherited disorders., volume 25 of *Oxford monographs in medical genetics*. Oxford University Press.

29. Mannucci MP, Tuddenham EGD (2001) The hemophilias - from royal genes to gene therapy. N Engl J Med 344: 1773 - 1780.

30. Mitchell J, Phillott A (2008) Haemophilia and inhibitors 1: diagnosis and treatment. Nursing Times 104: 26–27.

31. Tomokiyo K, Nakatomi Y, Araki T, Teshima K, Nakano H, et al. (2003) A novel therapeutic approach combining human plasma-derived factors viia and x for haemophiliacs with inhibitors: evidence of a higher thrombin generation rate in vitro and more sustained haemostatic activity in vivo than obtained with factor viia alone. Vox Sanguinis 85: 290-299.

32. Chavin SI, Siegel DM, Rocco TA, Olson JP (1988) Acute mycardial infarction during treatment with an activated prothrombin complex concentrate in a patient with factor viii deficiency and a factor viii inhibitor. Am J Med 85: 245-249.

33. Allen GA, Hoffman M, Roberts HR, Monroe DM (2006) Manipulation of prothrombin

458 concentration improves response to high-dose factor VIIa in a cell-based model of
459 haemophilia. Br J Haematology 134: 314 - 319.

34. Schneiderman J, Rubin E, Nugent DJ, Young G (2007) Sequential therapy with activated prothrombin complex concentrates and recombinant FVIIa in patients with severe haemophilia and inhibitors: update of our previous experience. Haemophilia 13: 244-248.

35. Talbot M, Tien HC (2009) The use of recombinant factor viia in trauma patients. J Am Acad Orthop Surg 17: 477-81.

36. Lusher JM, Roberts HR, Davignon G, Joist JH, Smith H, et al. (1998) A randomized, double-blind comparison of two dosage levels of recombinant factor viia in the treatment of joint, muscle and mucocutaneous haemorrhages in persons with haemophilia a and b, with and without inhibitors. rfviia study group. Haemophilia 4: 790–798.

37. Shapiro AD, Gilchrist GS, Hoots WK, Cooper HA, Gastineau DA (1998) Prospective, randomised trial of two doses of rfviia (novoseven) in haemophilia patients with inhibitors undergoing surgery. Thromb Haemost 80: 773–778.

38. Shapiro AD (2000) Recombinant factor viia in the treatment of bleeding in hemophilic children with inhibitors. Semin Thromb Hemost 26: 413–419.

39. Shapiro AD (2008) Single-dose recombinant activated factor vii for the treatment of joint bleeds in hemophilia patients with inhibitors. Clin Adv Hematol Oncol 6: 579–586.

40. Komiyama Y, Pedersen AH, Kisiel W (1990) Proteolytic activation of human factors ix and x by recombinant human factor viia: Effects of calcium, phospholipids and tissure factor. Biochemistry 29: 9418-9425.

41. Payne MA, Neuenschwander PF, Johnson AE, Morrissey JH (1996) Effect of soluble tissue factor on the kinetic mechanism of factor viia: Enhancement of p-guanidinobenzoate substrate hydrolysis. Biochemistry 35: 7100-7106.

484 42. Monroe DM, Hoffman M, Oliver JA, Roberts HR (1997) Platelet activity of high-dose

485 factor viia is independent of tissure factor. Br J Haematol 99: 542-547.

486 43. Hoffman M, Monroe III DM, Roberts HR (1998) Activated factor VII activates factors

487 IX and X on the surface of activated platelets: thoughts on the mechanism of action

488 of high-dose activated factor VII. Blood Coag Fibrin 9: S61-S65.

489 44. Jones E, Oliphant T, Peterson P (2001–). SciPy: Open source scientific tools for

490 Python. http://www.scipy.org/.

491 45. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of the

492 International Conference on Neural Networks. pp. 1942 - 1948.

493 46. Hunter JD (2007) Matplotlib: A 2d graphics environment. Computing in Science and

494 Engineering 9: 90 - 95.

495 47. Butenas S, van't Veer C, Mann KG (1999) "normal" thrombin generation. Blood 94:
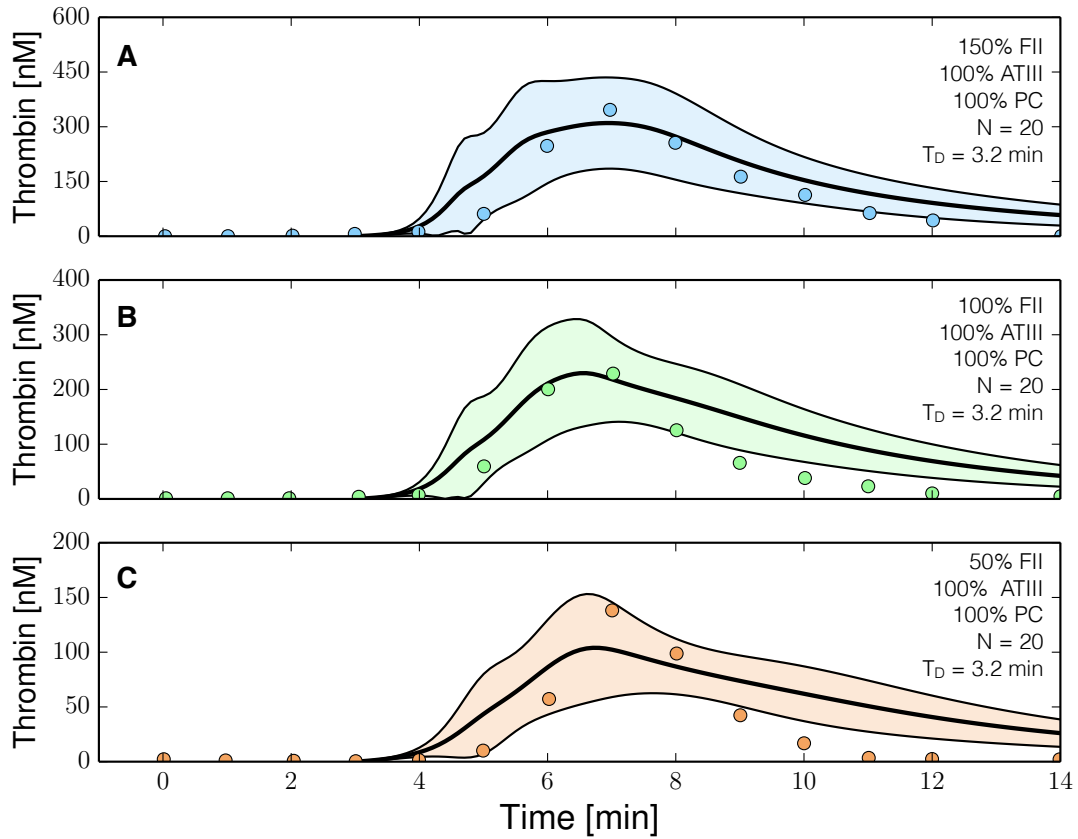
496 2169-78.
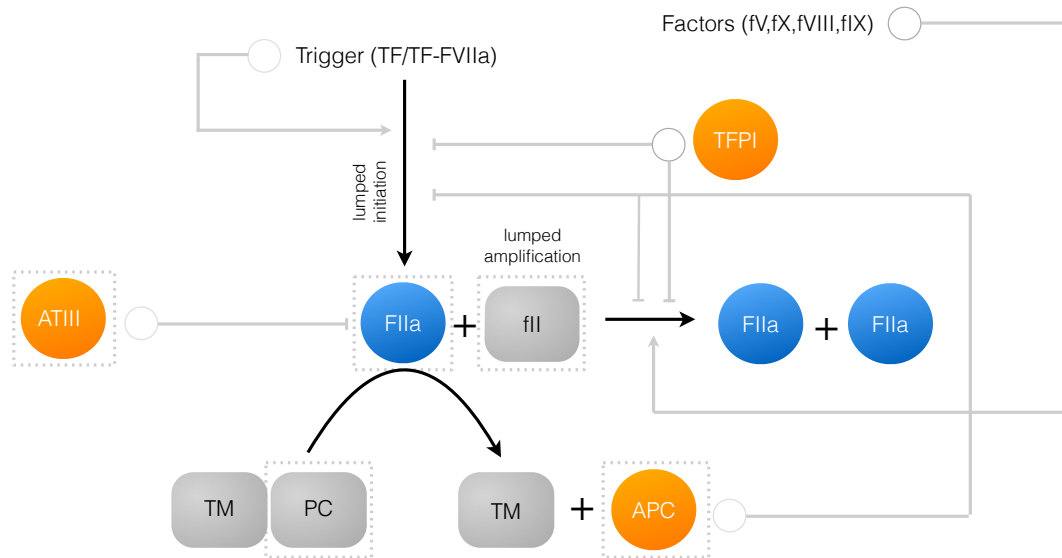
Fig. 1: Schematic of the connectivity of the reduced order coagulation model. A trigger compound e.g., TF/FVIIa initiates thrombin production (FIIa) from prothrombin (fII). Once activated, thrombin catalyzes its own activation (amplification step), as well as its own inhibition via the conversion of protein C to activated protein C (APC). APC and tissue factor pathway inhibitor (TFPI) inhibit initiation and amplification, while antithrombin III (ATIII) directly inhibits thrombin. All inhibition steps, and trigger-induced initiation were modeled using a rule-based approach. Likewise, the dependence of amplification on other coagulation factors was also modeled using a rule-based approach. The abundance of the highlighted species (in the dashed boxes) was governed by an ordinary differential equation. All other species were assumed to be constant.
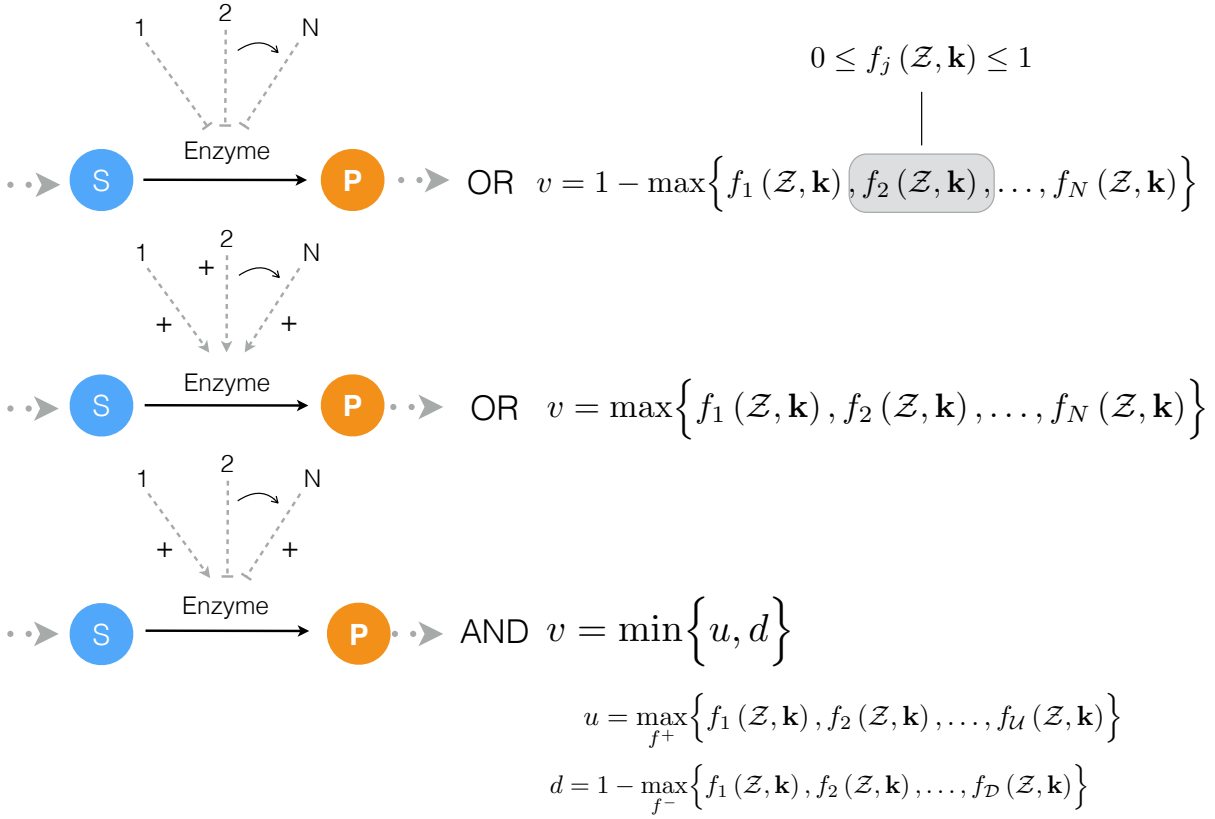
Fig. 2: Schematic of rule based effective control laws. Traditional enzyme kinetic expressions e.g., Michaelis–Menten or multiple saturation kinetics are multiplied by an enzyme activity control variable $0 \leq v_j \leq 1$. Control variables are functions of many possible regulatory factors encoded by arbitrary functions of the form $0 \leq f_j(\mathcal{Z}) \leq 1$. At each simulation time step, the v_j variables are calculated by evaluating integration rules such as the max or min of the set of factors $f_1, \ldots, f_n$ influencing the activity of enzyme E_j.
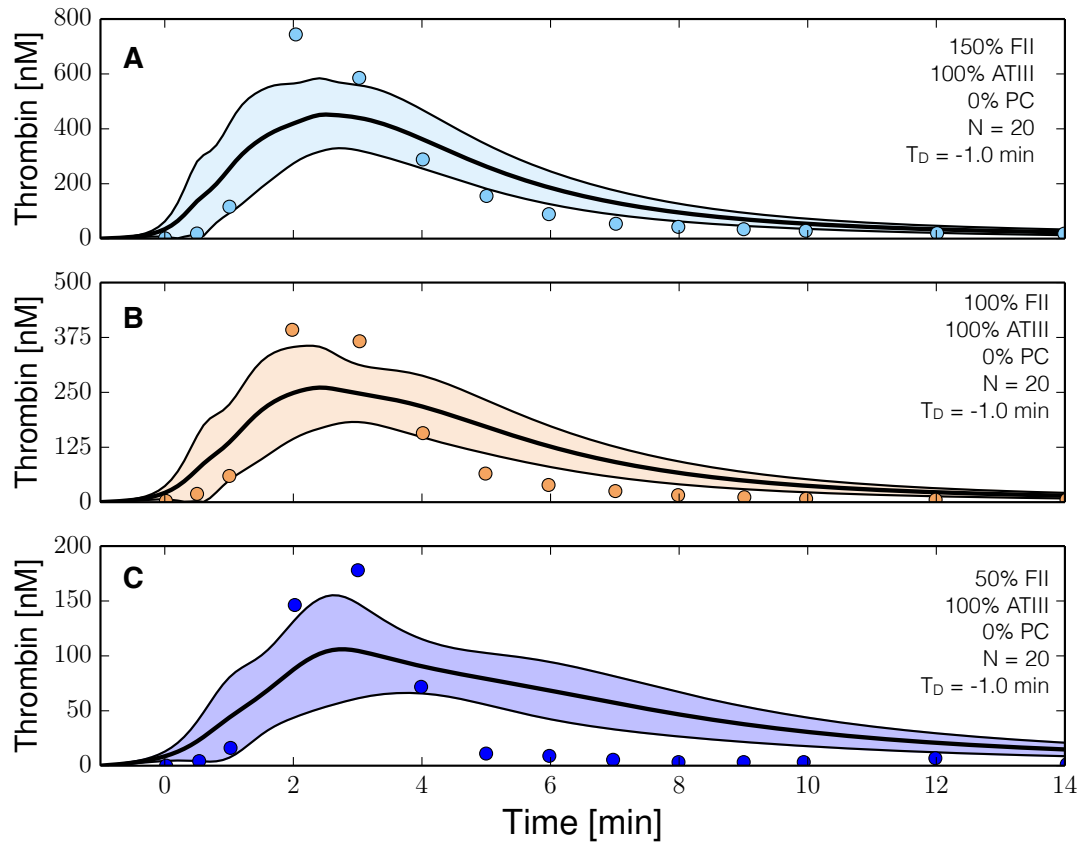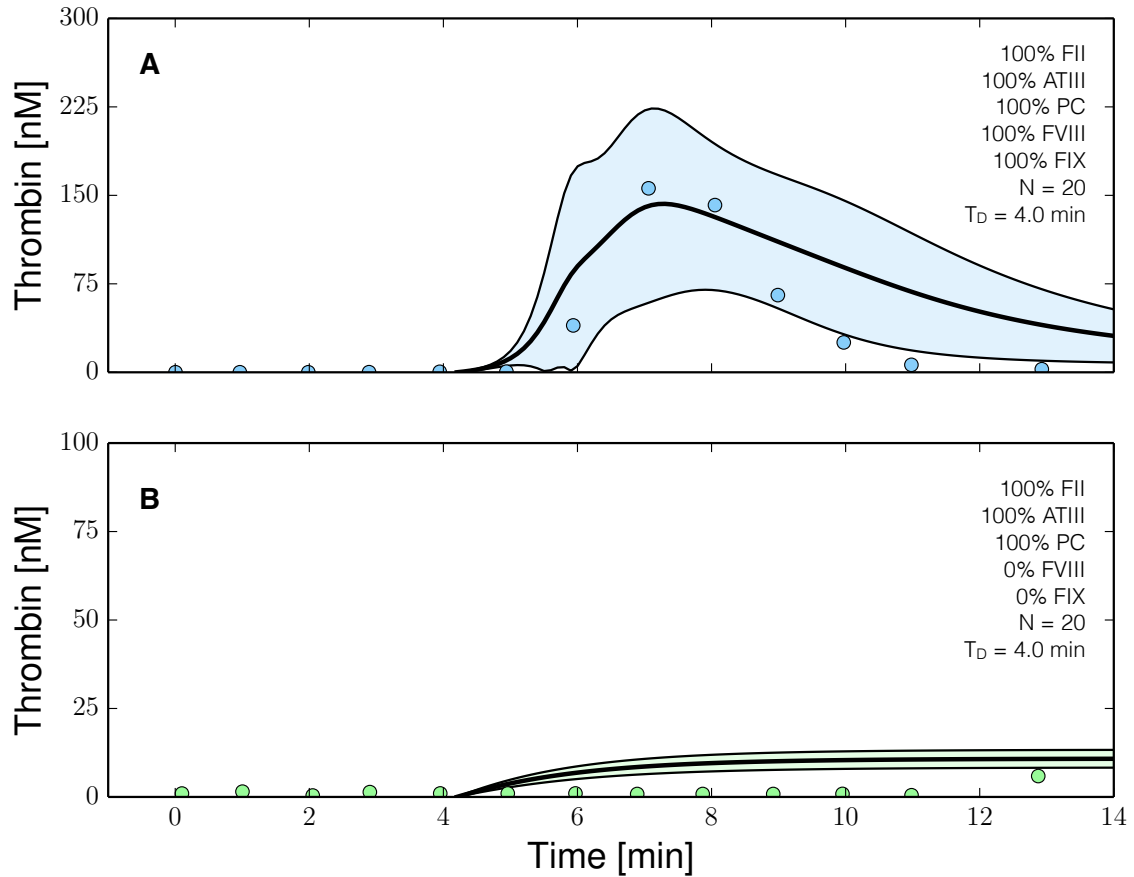
Fig. 3: Reduced order coagulation model training simulations. Reduced order coagulation model parameters were estimated using particle swarm optimization (PSO) with and without the protein C pathway as a function of prothrombin. Solid lines denote the simulated mean value of the thrombin profile for N = 20 independent particles, points denote experimental data. The shaded region denotes the 99% confidence estimate of the mean simulated thrombin value (uncertainty in the model simulation). (A,B,C) training results for 150%, 100% and 50% of physiological prothrombin levels in the absence of the protein C pathway. Thrombin generation was initiated using 5 pmol/L FVIIa-TF in the presence of 200 μmol/L of phospholipid vesicles (PCPS). All factors and control proteins were at their physiological concentration unless others denoted. The experimental training data was reproduced from the study of Butenas et al. [47].
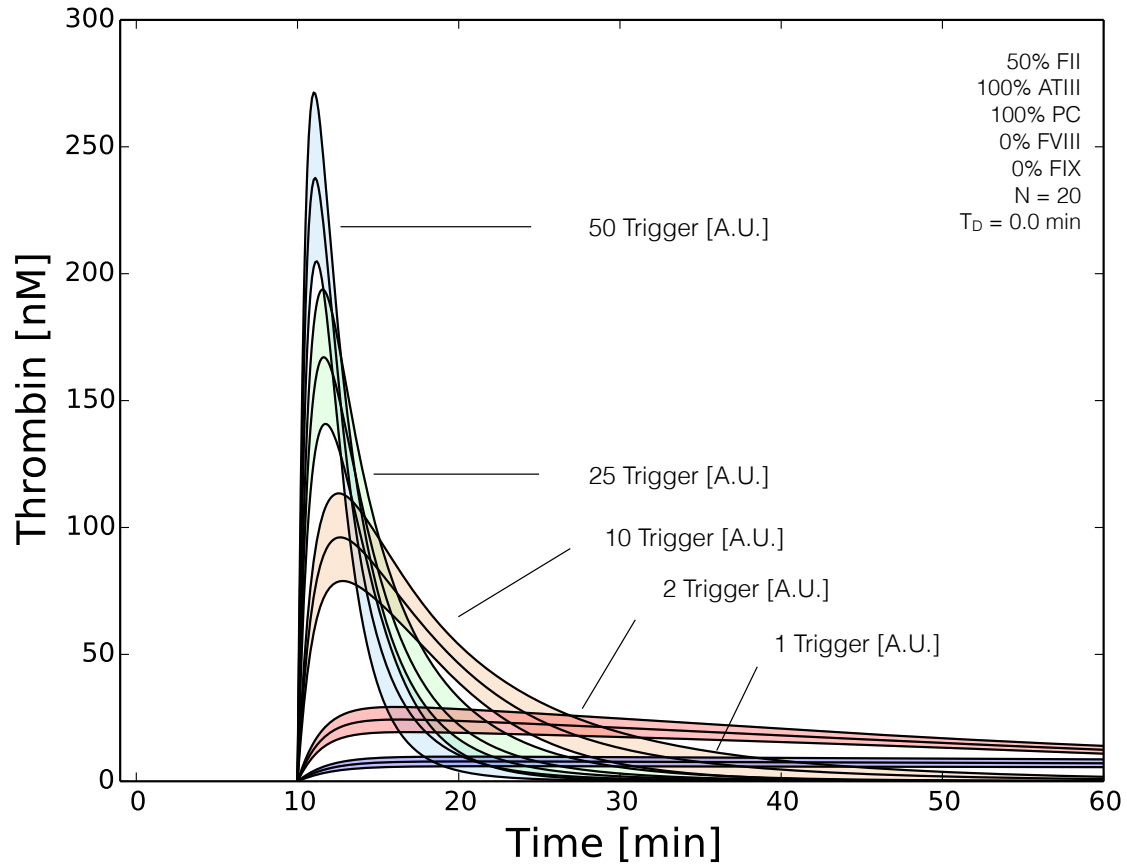
Fig. 4: Reduced order coagulation model training simulations. Reduced order coagulation model parameters were estimated using particle swarm optimization (PSO) with and without the protein C pathway as a function of prothrombin. Solid lines denote the simulated mean value of the thrombin profile for N = 20 independent particles, points denote experimental data. The shaded region denotes the 99% confidence estimate of the mean simulated thrombin value (uncertainty in the model simulation). (A,B,C) training results for 150%, 100% and 50% of physiological prothrombin levels in the presence of the protein C pathway. Only APC pathway parameters were allowed to vary in the simulations on the right. Thrombin generation was initiated using 5 pmol/L FVIIa-TF in the presence of 200 μmol/L of phospholipid vesicles (PCPS). All factors and control proteins were at their physiological concentration unless others denoted. The experimental training data was reproduced from the study of Butenas et al. [47].
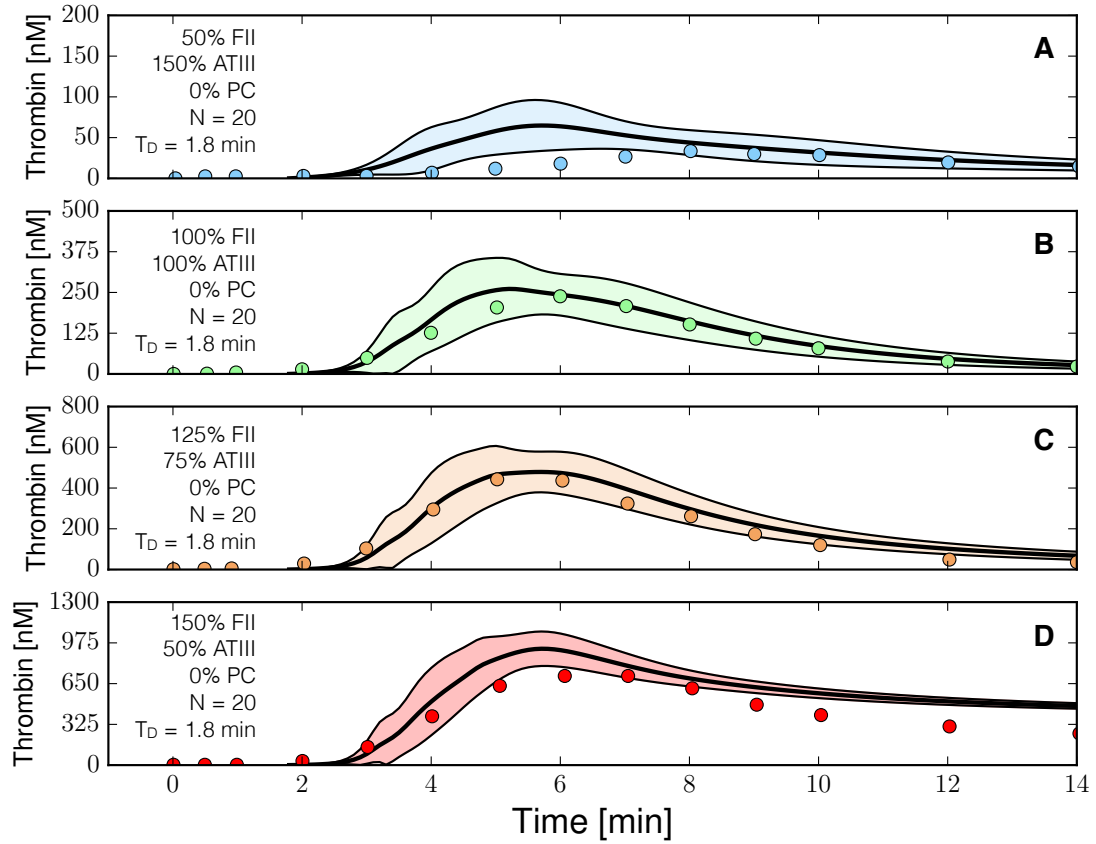
Fig. 5: Reduced order coagulation model predictions versus experimental data for normal coagulation. The reduced order coagulation model parameter estimates were tested against data not used during model training. Simulations of different levels of prothrombin and ATIII were compared with experimental data in the absence of the protein C pathway. Solid lines denote the simulated mean value of the thrombin profile for N = 20 independent particles, points denote experimental data. The shaded region denotes the 99% confidence estimate of the mean simulated thrombin value (uncertainty in the model simulation). (A,B,C,D) prediction results for (FII,ATIII): (50%,150%), (100%, 100%), (125%, 75%) and (150%, 50%) of physiological prothrombin and ATIII levels in the absence of the protein C pathway. Thrombin generation was initiated using 5 pmol/L FVIIa-TF in the presence of 200 μmol/L of phospholipid vesicles (PCPS). All factors and control proteins were at their physiological concentration unless others denoted. The experimental validation data was reproduced from the study of Butenas et al. [47]

Fig. 6: Reduced order coagulation model predictions versus experimental data with and without FVIII and FIX. The reduced order coagulation model parameter estimates were tested against data not used during model training. Simulations of normal thrombin formation with ATIII and the protein C pathway were compared with thrombin formation in the absence of fVIII and fIX. Solid lines denote the simulated mean value of the thrombin profile for N = 20 independent particles, points denote experimental data. The shaded region denotes the 99% confidence estimate of the mean simulated thrombin value (uncertainty in the model simulation). (A,B) prediction results for normal thrombin generation and thrombin generation in hemophilia. All factors and control proteins were at their physiological concentration unless others noted. The experimental validation data was reproduced from the study of Allen et al. [33].
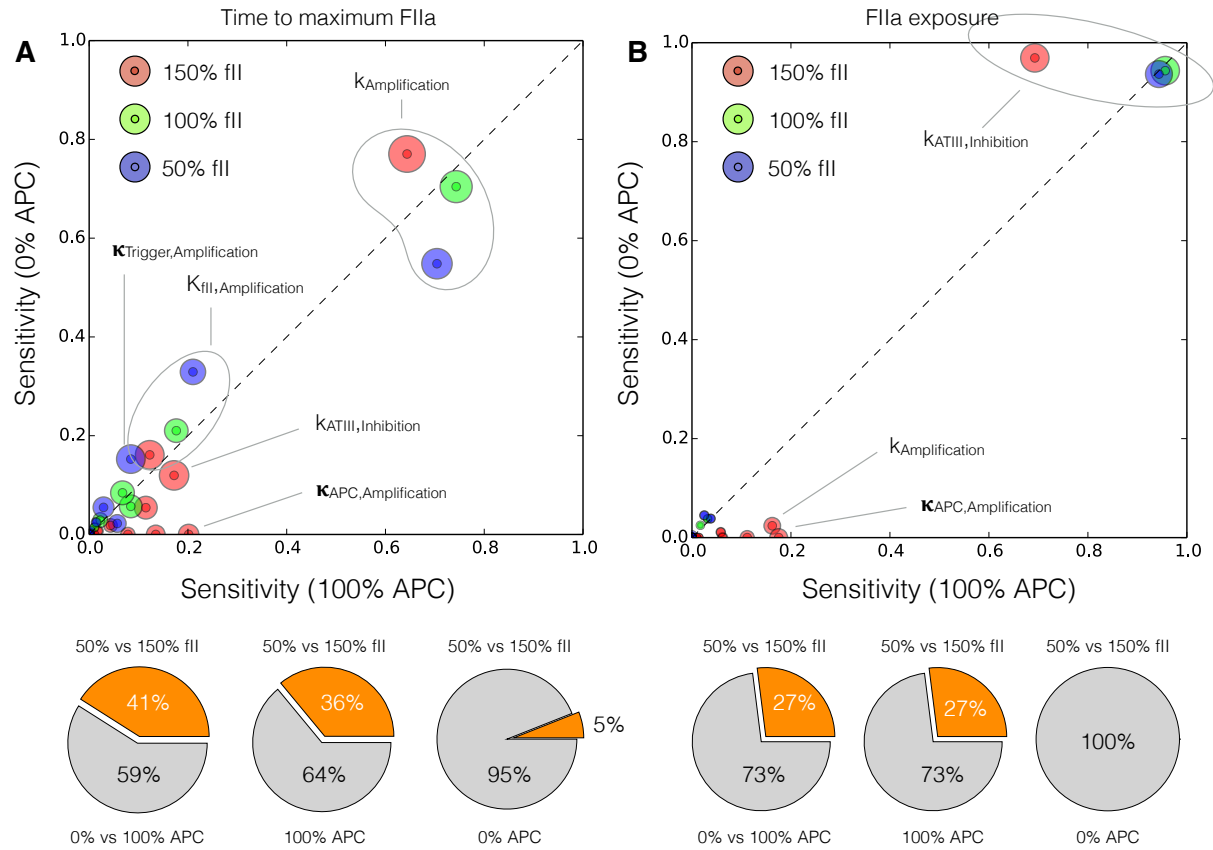
Fig. 7: Reduced order coagulation model predictions of rFVIIa administration. Simulations of thrombin formation in the presence of ATIII and the protein C pathway were conducted for a range of trigger values (1x - 50x) in the absence of fVIII and fIX. Solid lines denote the simulated mean value of the thrombin profile for N = 20 independent particles. The shaded region denotes the 99% confidence estimate of the mean simulated thrombin value (uncertainty in the model simulation). All factors and control proteins were at their physiological concentration unless others noted.
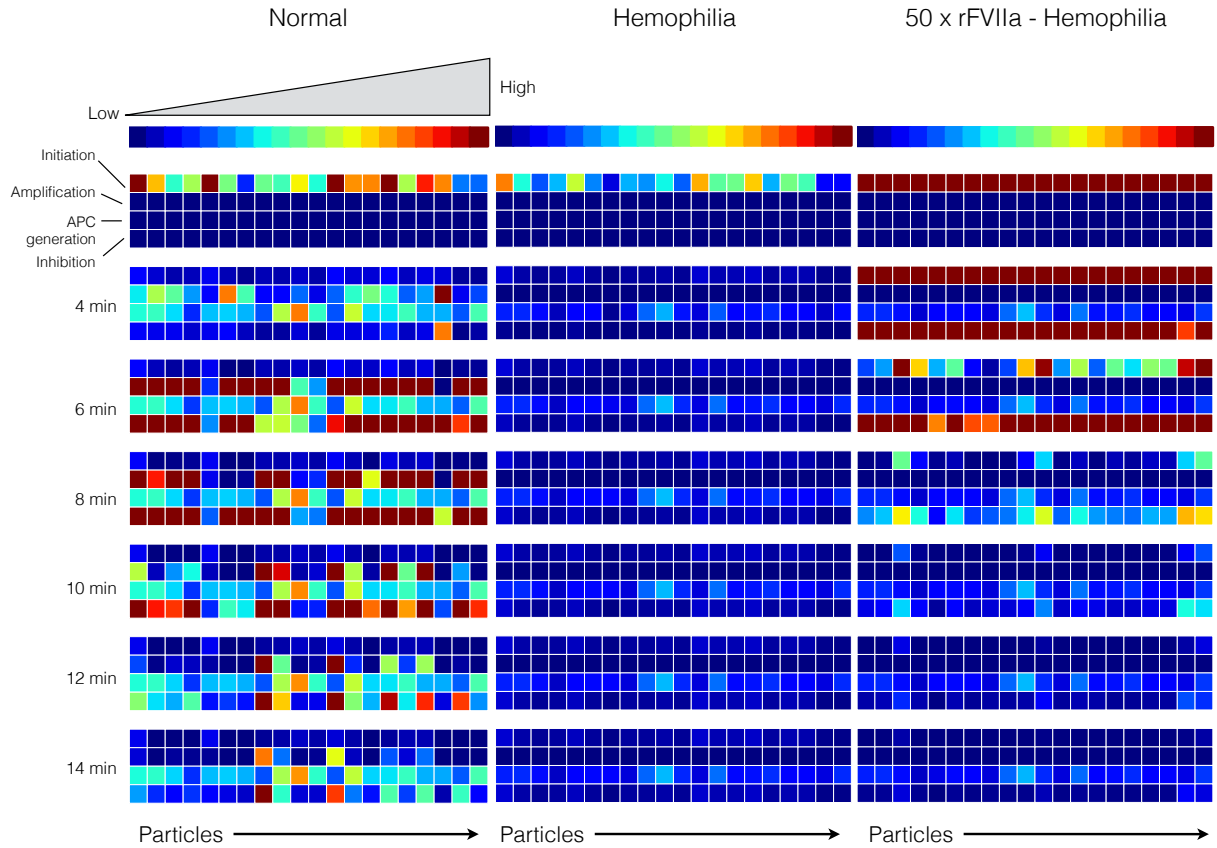
Fig. 8: Reaction flux distribution as a function of time for thrombin generation under normal (left), hemophilia (center) and rFVIIa treated hemophilia (right). Reaction flux was calculated for each particle at $T = 0, 4, 6, 8, 10, 12, 14$ min after the initiation of coagulation. Reaction fluxes were calculated for each particle in the parameter ensemble (N = 20). Blue colors denote low flux values while red colors denote high flux values.

Fig. 9: Global sensitivity analysis of the reduced order coagulation model with with respect to the model parameters. A: Sensitivity analysis of the thrombin peak time for different prothrombin levels (150%,100% and 50% of the physiological value) as a function of activated protein C. B: Sensitivity analysis of the thrombin exposure for different prothrombin levels (150%,100% and 50% of the physiological value) as a function of activated protein C. Points denote the mean total sensitivity value, while the area around each point denotes the uncertainty in the sensitivity value. The gray dashed line denotes the 45° degree diagonal, if sensitivity values are equal for different conditions they will lie on the diagonal. Sensitivity values significantly above or below the diagonal indicate differentially important model parameters.