Centrum voor Wiskunde en Informatica
REPORT*RAPPORT*

Mathematical Modelling in Blood Coagulation; Simulation and Parameter Estimation

W.J.H. Stortelder, P.W. Hemker, H.C. Hemker

Modelling, Analysis and Simulation (MAS)

MAS-R9720 September 30, 1997

Report MAS-R9720
ISSN 1386-3703

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Mathematical Modelling in Blood Coagulation; Simulation and Parameter Estimation

W. Stortelder[†] and P.W. Hemker
CWI,
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands.

H.C. Hemker
Department of Biochemistry,
University Maastricht,
P.O. Box 616, 6200 MD Maastricht, The Netherlands.

ABSTRACT
This paper describes the mathematical modelling of a part of the blood coagulation mechanism. The model includes the activation of factor X by a purified enzyme from Russel's Viper Venom (RVV), factor V and prothrombin, and also comprises the inactivation of the products formed.

In this study we assume that in principle the mechanism of the process is known. However, the exact structure of the mechanism is unknown, and the process still can be described by different mathematical models. These models are put to test by measuring their capacity to explain the course of thrombin generation as observed in plasma after recalcification in presence of RVV. The mechanism studied is mathematically modelled as a system of differential-algebraic equations (DAEs). Each candidate model contains some freedom, which is expressed in the model equations by the presence of unknown parameters. For example, reaction constants or initial concentrations are unknown. The goal of parameter estimation is to determine these unknown parameters in such a way that the theoretical (i.e., computed) results fit the experimental data within measurement accuracy and to judge which modifications of the chemical reaction scheme allow the best fit.

We present results on model discrimination and estimation of reaction constants, which are hard to obtain in another way.

1991 Mathematics Subject Classification: 65C20, 65K10, 90C31, 92C40, 65L05.
Keywords and phrases: Mathematical simulation, blood coagulation, parameter estimation, model discrimination, interactive modelling.
Note: Work carried out under the project 'Modelling of processes in Chemistry' (no. MAS2.6).

1. Introduction

One of the problems encountered in the study of a complicated biochemical process like thrombin generation in plasma, is that neither the reaction mechanism nor the reaction constants and initial concentrations are precisely known. The knowledge on the reaction mechanism of the process is obtained mainly through experiments on isolated parts of the system. The elements of the system, i.e. the clotting factors and their interactions, are separated from blood plasma and their interaction is studied under circumstances that are necessarily not precisely identical to those under which they cooperate in plasma. In fact it is not even known whether the reaction scheme that we deduce from such experiments is indeed the one operative in plasma. There may exist unknown factors or reactions, and reactions that have been shown to be possible in principle may not occur in reality. An example of this is the fact that factor X_a can activate factors V and VIII under experimental circumstances, but that this reaction does not seem to play a role in clotting plasma [MT90]. Also the reaction conditions

[†] Financially supported by the Dutch Technology Foundation (STW).

in plasma are different from those used for the study of the interaction of isolated factors. They may even be unsuitable for the study of such interactions. The kinetic parameters of activation of factor V by thrombin, e.g., can not be measured directly in plasma because the presence of natural thrombin inhibitors renders it impossible to achieve a fixed enzyme concentration.

In this article we introduce mathematical model validation and parameter estimation as a possible solution to these problems. In this procedure, on basis of the existing biochemical knowledge, a probable reaction mechanism is postulated. This is transformed into a set of differential-algebraic equations, which contains unknown parameters. These parameters correspond with the reaction constants and initial concentrations of the reactants, both approximatively known from previous experiments and used as an initial guess for the parameters to be estimated. Then, one or more results of the reaction process are monitored, e.g. the course of thrombin concentration in plasma in time after triggering of the coagulation process, and the parameters in the model are adapted to obtain an optimal fit. Different hypothetical reaction mechanisms can be tested in parallel to see which one results in a better fit. If the best fit leads to improbably large discrepancies between the computed and the experimental results, the model is adapted and the validation process is repeated.

In this paper we briefly indicate this process of model derivation and validation. In fact, the process consists of checking a long sequence of improving models, adapted during the process for a wide range of reasons. The final model should not only lead to a satisfactory fit, but should also be simple, in accordance with established facts, and –preferably– it should not contain an unreasonably large number of parameters. In order to validate the many models and to estimate the corresponding parameters, an interactive software package for parameter estimation on a fast computer is an indispensable tool. Such a computer program, called *spIds* [EHS95] and partially constructed by two of the authors, was available to carry out the necessary computations.

The model we consider in this paper describes thrombin formation, a part of the blood coagulation process, by a system of differential-algebraic equations. The variation in time of the concentrations of each reactant is described by a (differential) equation. The chain of reactions which leads to thrombin starts with the activation of factor X by RVV, followed by the activation of factor V, the production of prothrombinase in the presence of phospholipid and the activation of prothrombin. We also take into account the inactivation of the factor Xa by anti-thrombin III (ATIII) and the inactivation of thrombin by ATIII and α_2-macroglobulin (α_2M).

A description of the experiments used is given in Section 2, followed by a derivation of the reaction mechanism in Section 3. The step from reaction mechanism to mathematical equations is given in Section 4. The parameter estimation process is briefly described in Section 5. The results and conclusions are given in Section 6 and 7, respectively.

2. Experimental Data

In order to obtain the required data, four experiments were performed, which resulted in four series of measurements. The output of the system used for our tests was the course of thrombin-like amidolytic activity. This activity is caused by two types of molecules: thrombin itself and the thrombin-α_2 macroglobulin complex (briefly denoted as II_a and II_a–α_2M respectively, in the reaction scheme, Figure 1).

The data were obtained as follows. To 240 μl of defibrinated plasma, in which the clotting factors are contained, we add 3.6 μl of a suspension of procoagulant phospholipids (1 μM) and 80.4 μl of a solution of RVV. This concentration of RVV was halved in the subsequent experiments. The thrombin generation process was started at $t = 0$ by addition of 36 μl of $CaCl_2$ (100 mM). At different time intervals, more frequently in the initial phase of the reaction and less frequently at the end, we took 0.01 ml samples from the reaction mixture and added it to 0.49 ml of a solution of the chromogenic substrate S2238 (0.5 mM) in a buffer that contains the Ca_2^+ chelating agent EDTA in order to stop further thrombin generation. Thrombin and α_2M-thrombin split the yellow-coloured para-nitroaniline from S2238. After 2 min. this reaction is stopped by adding citric acid and the colour is measured

and used to determine the thrombin activity in the sample. Time measurements for the thrombin generation are made automatically and samples are taken until a stable end level of amidolytic activity is observed. This takes about 15 minutes.

3. Reaction Mechanism

At this point we first present a commonly accepted reaction sequence for thrombin generation in Figure 1. Thereafter we describe three possible variants as found in [Hem93]. In this section the reaction mechanism and its alternatives are given in a schematic way. In Section 4 we give a more precise description by deriving differential equations. This is followed by an overview of the motivation and selection criteria involved in choosing one set of equations in favour of its alternatives.

In the reaction schemes the coagulation factors are denoted by their Roman numbers, the subscript 'a' indicates their activated form, 'PL' and 'PT' denote phospholipid and prothrombinase, respectively. 'ATIII' and 'α_2M' (anti-thrombin III and α_2-macroglobulin) are responsible for inactivation of the factors II_a and X_a.

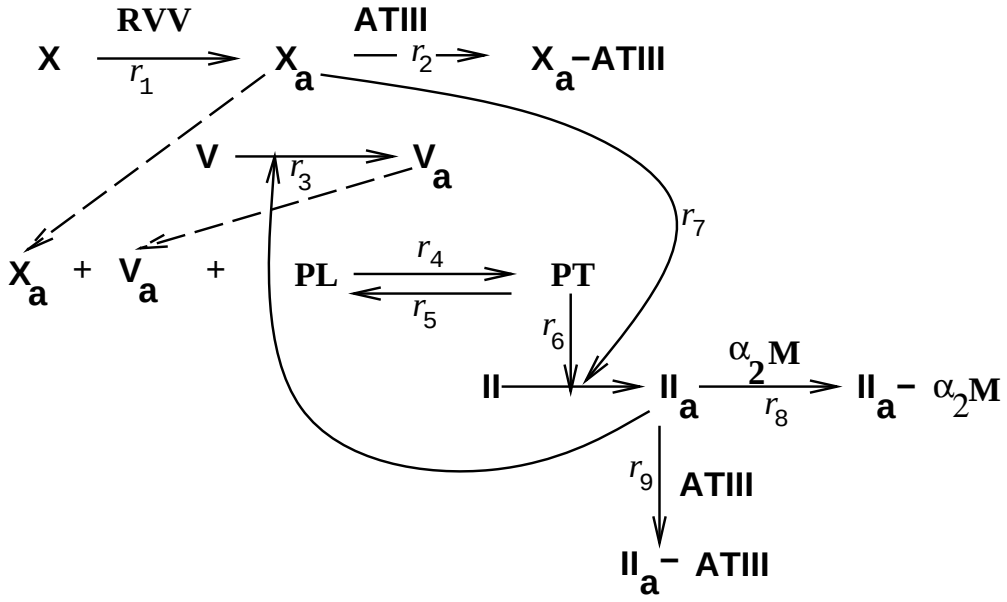

Figure 1: The reaction scheme for the part of the blood coagulation studied.

In the scheme of Figure 1, the activation of X by RVV, (reaction r_1), leads to X_a, followed by its inactivation by ATIII (r_2). Next, factor V is activated by II_a (r_3). The factors X_a, V_a and PL produce PT in a reversible association (r_4 and r_5). Subsequently, thrombin (II_a) is formed out of prothrombin (II), either in the presence of PT (r_6) or of X_a (r_7). Finally, II_a is inactivated either by α_2M or by ATIII (r_8 and r_9, respectively).

In this study we show that the above scheme is suitable to explain the experimental results. It summarises the present common knowledge, but it is not necessarily complete and/or unique. We also investigate a number of possible alternatives. One such alternative concerns the formation of prothrombinase (PT), not in a trimolecular reaction but as a sequence of bimolecular reactions (Figure 2). Two other alternatives are given in the Figures 3 and 4. In the former we account for the existence of the intermediate meizothrombin that in itself has amidolytic activity [BTH$^+$95], in the latter we account for the existence of an intermediate form of the α_2M-thrombin complex [MFG92].

All proposed alternatives are more complex than the reaction mechanism we start with in Figure 1. By "more complex" we mean that it has more state variables and more intermediate reactions,

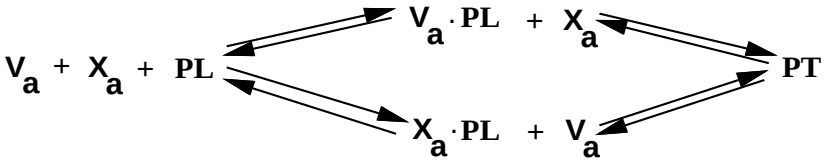

Figure 2: The alternative reaction scheme to account for prothrombin formation.

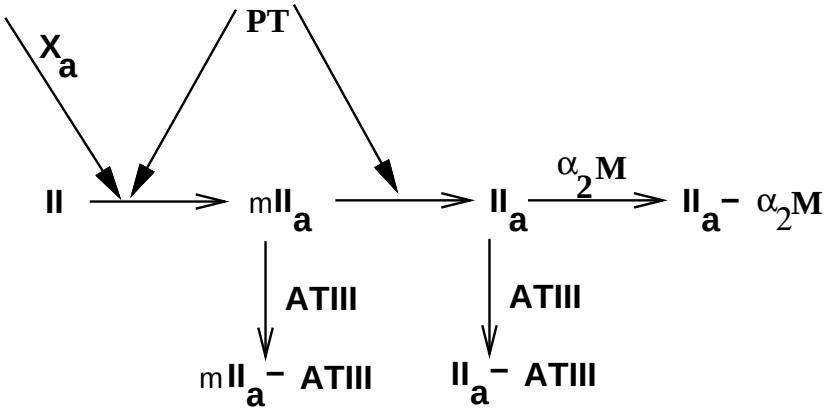

Figure 3: The alternative reaction mechanism for the formation of thrombin by the introduction of an intermediate reactant, meizothrombin (mII_a).

which implies that they are likely to fit better because there are more degrees of freedom available. In Section 5 we will derive model equations from the reaction schemes and judge by statistical tests if an increase of the complexity of the model leads to a significant improvement of the fit between the calculated model responses and the observed data.

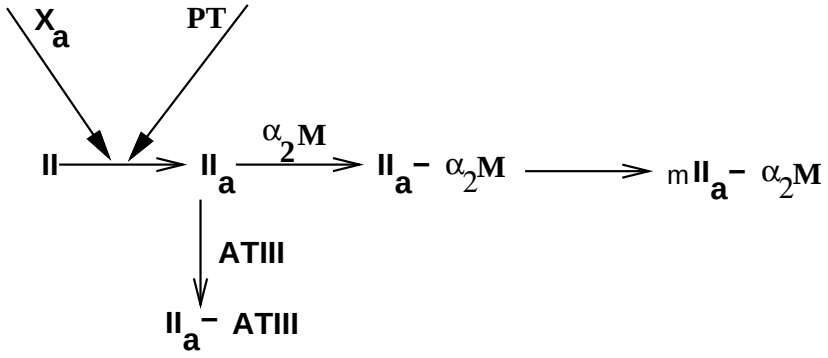

Figure 4: The alternative reaction mechanism for thrombin inactivation by $\alpha_2 M$. Here we assume that $II_a\alpha_2 M$ transforms further into an amidolytic less active form, $mII_a\alpha_2 M$.

4. Model Equations

From the four reaction schemes as they are introduced in Section 3, mathematical model equations were derived. It is obvious that the schemes presented lead to different sets of equations. But also from a single reaction scheme various sets of alternative mathematical model equations can be derived. As an example we consider the reaction r_1, which is present in all four reaction schemes. The concentrations of the chemical species are given in nM and indicated by '[ ]'; the time, t, is given in minutes. The dimension of the reaction constants are derived from these units. The change in time of the concentration of factor X can be given by the well-known Michaelis-Menten relation:

$$\frac{d[X]}{dt} = -r_1 = -\frac{kcat_X \cdot [X] \cdot [RVV]}{km_X + [X]} \ . \tag{4.1}$$

Although we know from literature that this relation is likely to be valid, it may be replaced by closely related expressions. In cases where $km_X \gg [X]$ or $km_X \ll [X]$, expression (4.1) transforms respectively into the alternatives

$$r_1 = kk_1 \cdot [X] \cdot [RVV] \ , \tag{4.2}$$

with $kk_1 \approx kcat_X/km_X$ or

$$r_1 = kk_2 \cdot [RVV] \ , \tag{4.3}$$

with $kk_2 \approx kcat_X$. Both alternatives have one parameter less than the Michaelis-Menten relation and, depending on the ratio $km_X/[X]$, they can replace (4.1) without loss of accuracy. A third possible alternative reads:

$$r_1 = kk_3 \cdot [X] \ , \tag{4.4}$$

which follows from (4.2), when RVV-dependence is negligible. Similar alternatives exist for the other reactions. Together, this leads to a large number of candidate models.

From all these candidates we select that model (or subset of models, if the statistical tests do not lead to a decisive answer) which, (i) is in accordance with established knowledge in the field, (ii) is devoid of irrelevant steps (cf. the Michaelis-Menten reaction mentioned above), and (iii) fits the phenomena observed.

In Section 5 we will highlight the process of parameter estimation and deal with model validation. In the last part of the present section we give the set of model equations which was chosen from the candidates on the basis of the criteria (i)-(iii). This set is one of the possible mathematical representations for the scheme given in Figure 1. and as such it is an example of the many possible systems of DAEs. In addition, it describes the connection with the experiments.

The selected system of equations reads:

$$\frac{d[X]}{dt} = -r_1 \ , \tag{4.5}$$

$$\frac{d[Xa]}{dt} = r_1 - r_2 - r_4 + r_5 \ , \tag{4.6}$$

$$\frac{d[V]}{dt} = -r_3 \ , \tag{4.7}$$

$$\frac{d[Va]}{dt} = r_3 - r_4 + r_5 \ , \tag{4.8}$$

$$\frac{d[PL]}{dt} = -r_4 + r_5 \ , \tag{4.9}$$

$$\frac{d[PT]}{dt} = r_4 - r_5 \ , \tag{4.10}$$

$$\frac{d[II]}{dt} = -r_6 - r_7 \ , \tag{4.11}$$

$$\frac{d[IIa]}{dt} = r_6 + r_7 - r_8 - r_9 \ , \tag{4.12}$$

$$\frac{d[IIa\alpha_2M]}{dt} = r_9 \ , \tag{4.13}$$

$$AmAct = [II_a] + 0.556 \cdot [IIa\alpha_2M] \ , \tag{4.14}$$

$$r_1 = \frac{kcat_X \cdot [X] \cdot [RVV]}{km_X + [X]} \ , \tag{4.15}$$

$$r_2 = ki_{Xa} \cdot [Xa] \ , \tag{4.16}$$

$$r_3 = \frac{kcat_V \cdot [V] \cdot [IIa]}{km_V + [V]} \ , \tag{4.17}$$

$$r_4 = k_{PT} \cdot [Va] \cdot [Xa] \cdot [PL] \ , \tag{4.18}$$

$$r_5 = k_{PL} \cdot [PT] \ , \tag{4.19}$$

$$r_6 = \frac{kcat_{II} \cdot [II] \cdot [PT]}{km_{II} + [II]} \ , \tag{4.20}$$

$$r_7 = \frac{kcat_2 \cdot [II] \cdot [Xa]}{km_2 + [II]} \ , \tag{4.21}$$

$$r_8 = ki_{IIa\alpha_2M} \cdot [IIa] \ , \tag{4.22}$$

$$r_9 = ki_{IIaATIII} \cdot [IIa] \ . \tag{4.23}$$

The concentration of RVV is supposed to be constant during each experiment. However, it should be noted that [RVV] differs for the different experiments. The inactivation of II_a and X_a in the presence of ATIII and α_2M is modelled by first order reactions (r_2, r_8 and r_9). This implies that the concentrations of these inhibitors do not occur in the equations.

The available measurements concern the amidolytic activity, which is expressed as the equivalent amount of thrombin (nM). This means that, in addition to the equations describing the chemistry, an equation for the amidolytic activity should be added. This equation is given in (4.14). It takes into account that the amidolytic activity does not only depend on the activity of thrombin (IIa), but also on the activity of the thrombin inactivated by α_2M (IIaα_2M). It is known from [Hem93] that the inactivated form shows an activity of 55.6% of the active thrombin.

In addition to the system of nine differential equations (4.5)-(4.13), we need the same number of initial conditions. At the start ($t = 0$), the initial concentrations of all state variables are zero, except for [PL], [II], [V] and [X].

5. Parameter estimation and model validation

The system of equations (4.5)-(4.23) contains 13 reaction constants. None of these constants nor the initial concentrations of the coagulation factors [II], [V] and [X] are known exactly, but they are assumed to be constant for each experiment. These 13 reaction constants, plus the three unknown initial conditions, are the quantities we want to determine; the unknown parameters. We summarise these parameters in Table 2. From the current literature we know upper and lower bounds for the concentrations of the clotting factors in normal plasma: i.e. [750nM,2200nM] for II, [10nM,30nM] for V and [70nM,200nM] for X.

The parameters are determined in such a way that the model responses fit the measurements in a least squares sense. Besides the parameters, confidence regions for the parameters are derived. For more details about the numerical solution of the model equations, minimisation of the least squares criterion and the confidence regions the reader is referred to [Hem72] or [Sto96].

To get more insight in our process of model discrimination, we compare each of the four options, (4.1)-(4.4), in combination with the reactions r_2 to r_9 from Figure 1 as they are described in (4.16)-(4.23). The expressions for r_2 to r_9 are obtained by a similar process of selection and validation as we will describe below.

Under the assumption of (4.16)-(4.23) we immediately reject option (4.4), because it implies that RVV has no influence on the reaction scheme, which is not in agreement with the experiments.

Under the assumption of (4.16)-(4.23), with one of the options (4.1), (4.2) or (4.3) we compare the corresponding model performances shown in Table 1. From this table it is obvious that the first

r_1	# par	df	$S(\widehat{\theta})$
(4.1)	16	104	6460.4
(4.2)	15	105	70203.8
(4.3)	15	105	100511.0

Table 1: Comparison for the three remaining options (4.1), (4.2) and (4.3). We show the number of parameters (# par), the degrees of freedom (df: the number of measurements minus the number of free parameters) and the least squares sum ($S(\widehat{\theta})$).

alternative performs better than the other two, if we take only $S(\widehat{\theta})$ into account. In order to decide if one model performs *significantly* better than another, we use the $\mathcal{F}$-ratio test (see Appendix A). To apply this test to the three remaining options for r_1, we take the reaction scheme from Figure 1 and r_2 to r_9 as in (4.16)-(4.23). The relevant data for the $\mathcal{F}$-ratio test are given in Table 1. The test of a significant difference between (4.1) and (4.2) consists of computing the quantity (cf. (A.2))

$$X_{(4.1),(4.2)} = \frac{6460.4/104}{70203.8/105} = 0.0929 \ , \tag{5.1}$$

and to check whether this number is between $1/\mathcal{F}_{\frac{\alpha}{2}}(105,104)$ and $\mathcal{F}_{\frac{\alpha}{2}}(104,105)$, i.e. 0.602 and 1.661, respectively, for $\alpha = 0.01$. The ratio (5.1) exceeds the bounds of this region, which means that the model with (4.1) accounts significantly better for the phenomena observed. Therefore, r_1 from (4.2) is rejected. Similarly (4.3) is rejected, because it performs even poorer, as can be seen from Table 1.

Also, the other models which are derived from alternative schemes described in the Figures 2, 3 and 4, have been tested. All these alternatives give rise to models with more state variables and more parameters. However, following the same strategy none of them turned out to perform significantly better.

6. Results

An initial estimate for the parameters consists of an educated guess from the existing biochemical literature ([Hem93] and references therein). These initial values are given in Table 2. The final estimates, and the corresponding confidence regions are also listed in this table. For details on the statistics, the reader is referred to [Sto96]. The sum of squared residuals for the initial estimates was 2.40×10^7, after minimisation it was reduced to 6.287×10^3.

The measurements (120 in total and 30 for each experiment) and the model responses for the final estimates of the parameters are given in Figure 5. The plots show a very acceptable fit between the computed and measured values, i.e. a fit within the measurement accuracy, which means that the model gives a sufficiently accurate description of the measured quantities.

The independent and dependent confidence regions as they are listed in the fourth and fifth column of Table 2 show that by far not all the parameters can be estimated within reasonable accuracy. From the singular value decomposition of the covariance matrix of the parameters (see e.g. [Hem72]), we can deduce that with the current model and the available measurements 5 parameters (or combinations of parameters) can be estimated with acceptable accuracy. By making use of other chromogenic substrates, additional measurements for V_a and X_a can be obtained in order to estimate more parameters more accurately.

The parameter km_2 tends to become small during the parameter estimation procedure and the idea came up to replace the corresponding reaction, r_7 (cf. (4.21)), with $kk_5 \cdot [Xa]$, in order to reduce the

para-meter	initial est. (θ_{ini})	final est. ($\widehat{\theta}$)	independent confidence regions ($\delta\theta^*$)	dependent confidence regions ($\delta\theta^\dagger$)
$kcat_X$	5.00×10^{3}	2.391×10^{2}	5.301×10^{3}	1.963×10^{1}
km_X	4.00×10^{2}	2.365×10^{1}	5.776×10^{2}	6.335×10^{0}
ki_{X_a}	2.50×10^{-1}	4.531×10^{0}	1.408×10^{1}	3.667×10^{-1}
k_{PT}	1.00×10^{-1}	1.229×10^{2}	3.117×10^{5}	4.152×10^{1}
k_{PL}	1.00×10^{1}	8.014×10^{2}	2.032×10^{6}	2.711×10^{2}
$kcat_V$	1.40×10^{1}	7.844×10^{0}	2.166×10^{3}	1.862×10^{0}
km_V	7.20×10^{1}	1.497×10^{2}	4.261×10^{4}	3.666×10^{1}
$kcat_{II}$	2.00×10^{3}	4.387×10^{1}	8.678×10^{2}	2.956×10^{0}
km_{II}	2.10×10^{2}	6.225×10^{1}	2.147×10^{2}	2.073×10^{1}
$kcat_2$	2.30×10^{0}	1.240×10^{1}	2.596×10^{2}	9.150×10^{-1}
km_2	5.80×10^{1}	6.148×10^{-2}	2.937×10^{1}	1.630×10^{1}
$ki_{IIaATIII}$	1.30×10^{0}	7.859×10^{-1}	5.794×10^{-1}	4.423×10^{-2}
$ki_{IIa\alpha_2M}$	1.50×10^{0}	1.762×10^{-1}	4.611×10^{-2}	2.673×10^{-2}
X_{ini}	1.33×10^{2}	8.125×10^{1}	1.729×10^{3}	7.556×10^{0}
V_{ini}	1.67×10^{1}	6.712×10^{0}	1.663×10^{2}	5.821×10^{-1}
II_{ini}	1.33×10^{3}	5.093×10^{2}	2.677×10^{2}	2.112×10^{1}
$S(\theta)$	2.40×10^{7}	6.287×10^{3}		

Table 2: Initial guess and final estimates for the parameters and their confidence regions.

number of parameters by one. The corresponding model gave negative results for the concentration of factor II, which is a consequence of adapting r_7 (the inequality $[II] \gg km_2$ did not hold on the whole time interval), and was therefore rejected.

The term r_7 is inevitable, because without this term the production of thrombin will not even start. This can be seen from the reaction scheme of Figure 1 and the fact that the initial concentrations of II_a and V_a are zero. Before the start of the experiments the expectation of the biochemists was that the activation of prothrombin (II) would be mainly performed by prothrombinase (PT) and that the contribution of X_a would be marginal here. In other words: r_7 would be small compared to r_6 and therefore (after initiating the reaction) could be neglected after a few seconds. By investigating the separate contributions to the thrombin production for r_6 and r_7 during the simulations, we found that the contribution of r_7 is about 50% of the production by r_6 and therefore not negligible. This conclusion should, however, be strictly limited to the case of RVV as a factor X activator and be extrapolated to other experimental setups.

Although the results of Table 2 may look poor with respect to the confidence regions, it appears that with the current data we were able to discriminate between many models in a systematic way and to come up with a model which fits the observations satisfactorily.

7. Conclusions

In this paper we compare a number of possible reaction schemes which describe part of the blood coagulation mechanism. For each scheme mathematical model equations have been derived and parameters have been adapted in order to obtain a best fit with a set of experimental data. Depending on the complexity of the model, and the quality of the fit, judged by the statistical criteria, we were able to discriminate between many candidate models. The final model is compact, meets the established knowledge in the field and fits the measurements satisfactorily. A large number of more compact models were rejected on the account of the measurements. More sophisticated models were rejected because the increase of complexity did not account for a sufficient improvement of the fit.

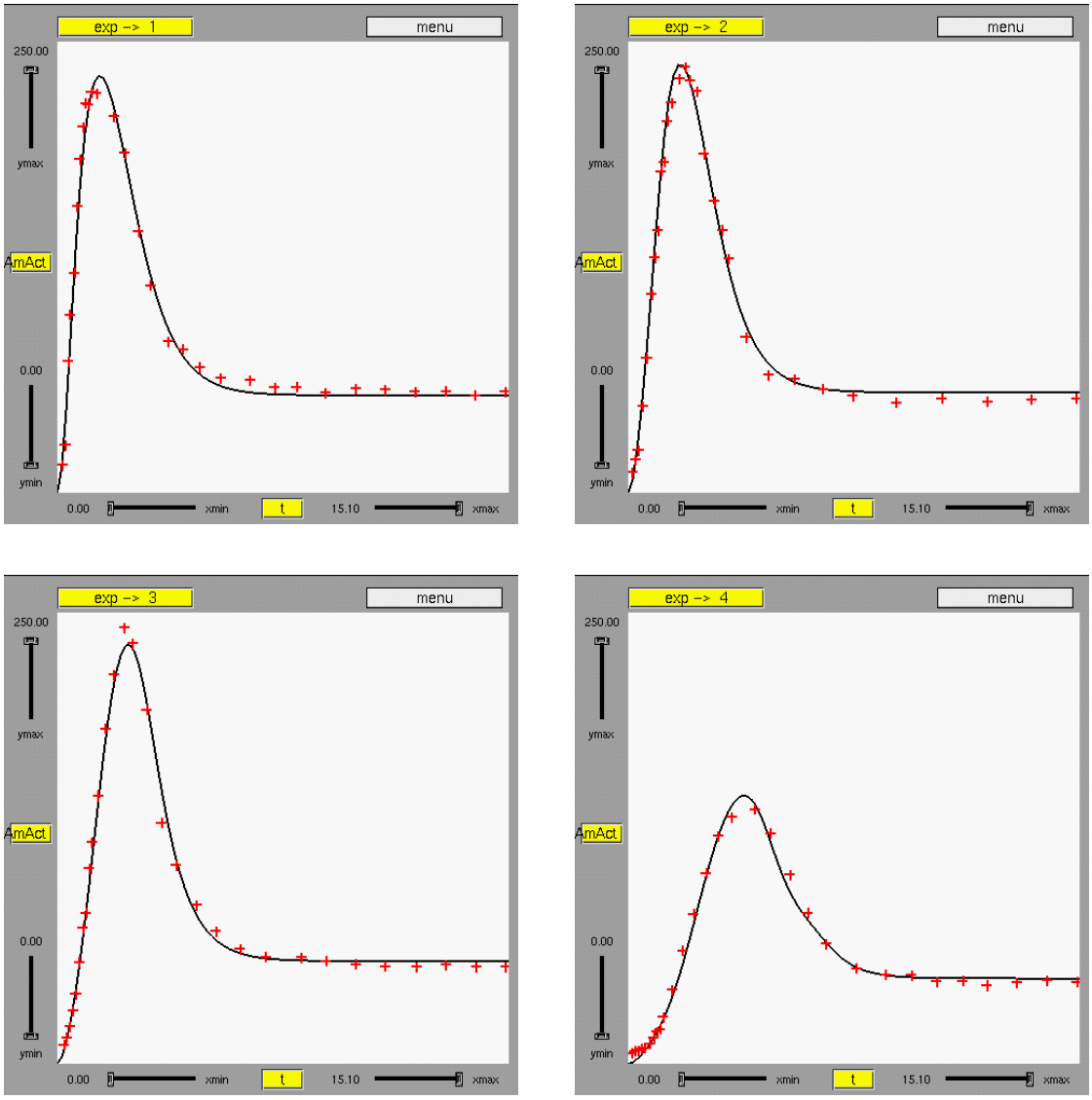

Figure 5: Plots of the measurements ('+') and the model responses for the final estimates of the parameters from Table 2 over the 4 experiments with decreasing concentrations of RVV.

With the final model selected not only its parameter estimates are presented, which are optimal in a least squares sense with respect to the available data, but also the corresponding confidence regions. Additional experiments can make the confidence regions smaller, but they may also lead to a more complex model in favour of one of the alternatives which had to be rejected in this paper.

In this sense the presented model can be a good starting point for ongoing research and may show its value when more experimental data are available.

Acknowledgement:

P. Devilée (from the biochemistry department of the university of Maastricht) is gratefully acknowledged for the numerous discussions, his work on preparing the data and his patience to deal with a layman experimenting in the lab.

References

[BTH+95] E.G. Bovill, R.P. Tracy, T.E. Hayes, R.J. Jenny, F.H. Bhushan, and K.G. Mann. Evidence that meizothrombin is an intermediate product in the clotting of whole blood. *Arterioscler. Thromb. Vasc. Biol.*, 15(6):754–758, 1995.

[EHS95] C.T.H. Everaars, P.W. Hemker, and W. Stortelder. Manual of spIds, a software package for parameter identification in dynamic systems. Technical Report NM-R9521, CWI, Amsterdam, 1995.

[Hem72] P.W. Hemker. Numerical methods for differential equations in system simulation and in parameter estimation. In H.C. Hemker and B. Hess, editors, *Analysis and Simulation of Biochemical Systems*, pages 59–80. North Holland Publ. Comp., 1972.

[Hem93] H.C. Hemker. Thrombin generation, an essential step in haemostasis and thrombosis. In A.L. Bloom and D. Thomas, editors, *Haemostasis and thrombosis, 3E*, pages 477–491, 1993.

[MFG92] L.B. Marshall, N.L. Figler, and S.L. Gonias. Identification of alpha 2-macroglobulin conformational intermediates by electron microscopy and image analysis. Comparison of alpha 2-macroglobulin-thrombin and alpha 2-macroglobulin reacted with cis-dichlorodiammineplatinum(II) and trypsin. *Journ. Biol. Chem.*, 267(9):6347–6352, 1992.

[MGB74] A.M. Mood, F.A. Graybill, and D.C. Boes. *Introduction to the Theory of Statistics.* McGraw-Hill, Inc., 1974.

[MT90] D.D. Monkovic and P.B. Tracey. Activation of human factor V by factor Xa and thrombin. *Biochemistry*, 29(5):1118–1128, 1990.

[Rat83] D.A. Ratkowsky. *Nonlinear Regression Modeling.* Marcel Dekker, Inc., New York, 1983.

[Sto96] W. Stortelder. Parameter estimation in dynamic systems. *Mathematics and Computers in Simulation*, 42(2/3):135–142, 1996.

A. $\mathcal{F}$-ratio test

We refer to [Rat83] for an introduction to the statistical tests which should be performed and help the modeller to decide what parameters can be skipped or what model should be chosen. When we have two models with approximately the same fit, the model with the fewest parameters is favourite for further investigation. What we mean by approximately will be made more precise below.

Suppose we have two solutions coming from different models

$$y(t,\theta) \qquad \text{and} \qquad z(t,\phi). \tag{A.1}$$

We use n_y and m_θ for the dimension of y and θ, respectively. The dimensions of z and ϕ will be denoted by n_z and m_ϕ. In general, different models describing the same physical process have different numbers of state variables or parameters. The only restriction is that the vectors y and z both contain the state variables for which measurements are available.

From the two models we get optimal parameters and corresponding sums of squares: $\widehat{\theta}$, $\widehat{\phi}$, $S(\widehat{\theta})$ and $S(\widehat{\phi})$. From the normal assumption with respect to the measurement error, and assuming that the optimal parameters are close to the true parameters, we know that:

$$X = \frac{S(\widehat{\theta})/(N-m_\theta)}{S(\widehat{\phi})/(N-m_\phi)} \sim \mathcal{F}(N-m_\theta, N-m_\phi)\ . \tag{A.2}$$

Where $\mathcal{F}(p,q)$ denotes the Fisher's F distribution with p and q degrees of freedom respectively. From the characteristics of an F distribution we know:

$$\mathbf{E}(X) = \frac{N-m_\phi}{N-m_\phi-2}\ , \qquad \text{(for: } N-m_\phi > 2\text{)}$$

and

$$P\left(\frac{1}{\mathcal{F}_{\frac{\alpha}{2}}(N-m_\phi, N-m_\theta)} \le X \le \mathcal{F}_{\frac{\alpha}{2}}(N-m_\theta, N-m_\phi)\right) = 1-\alpha\ ,$$

where $\mathcal{F}_{\frac{\alpha}{2}}(N-m_\theta, N-m_\phi)$ is the upper $\alpha/2$ quantile for Fisher's distribution (see e.g. [MGB74]). Notice that the expectation of X does not depend on m_θ. When the two models have about the same performance, X will be close to its expectation. Whenever X exceeds $F_{\alpha/2}(N-m_\theta, N-m_\phi)$, which means that $S(\widehat{\theta})$ is significantly larger than $S(\widehat{\phi})$, the model which corresponds to $y(t,\theta)$ is rejected in favour of the model which corresponds to $z(t,\phi)$. Of course, the opposite happens if X is less then $1/\mathcal{F}_{\frac{\alpha}{2}}(N-m_\phi, N-m_\theta)$. In all other cases no conclusion can be drawn, although if there is a model with less parameters, then this one is favourite for reasons of compactness.