Data Analysis in Experimental Biomedical Research

Markovich, Dmitriy

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Data Analysis in Experimental Biomedical Research

Dmitriy L. Markovich
PhD Thesis November 2017

DTU Nanotech
Department of Micro- and Nanotechnology

PhD Thesis
Doctor of Philosophy

DTU Nanotech
Department of Micro- and Nanotechnology

Data Analysis

in Experimental Biomedical Research

Dmitriy L. Markovich

Department of Micro- and Nanotechnology
Technical University of Denmark
DK-2800 Kongens Lyngby
Denmark
2017

DTU

DTU Nanotech
Department of Micro-and Nanotechnology
Technical University of Denmark

Ørsteds Plads
Building 345B
2800 Kongens Lyngby, Denmark
Phone +45 45 25 57 00
Fax: +45 45 88 77 62
info@nanotech.dtu.dk
www.nanotech.dtu.dk

Summary

This thesis covers two non-related topics in experimental biomedical research: data analysis in thrombin generation experiments (collaboration with Novo Nordisk A/S), and analysis of images and physiological signals in the context of neurovascular signalling and blood flow regulation in the brain (collaboration with University of Copenhagen).

The ongoing progress in experimental methods of thrombin generation allowed to introduce ready-to-use commercial assays for thrombin measurement. On the other hand, commercial assays use "black box" data analysis which makes it nearly impossible for researches to critically assess and compare obtained results. We reverse engineered the data analysis performed by CAT, a de facto standard assay in the field. This revealed a number of possibilities to improve its methods of data analysis. We found that experimental calibration data is described well with textbook Michaelis-Menten kinetics of the basic enzymatic reaction. We also found that a simple phenomenological model inspired by a cascade of basic enzymatic reactions describes thrombin generation data. Finally, we proved that CAT greatly overestimates several key physiological parameters that describe thrombin generation. Our work is a step towards standardization in the field of thrombin generation based on transparent and well-motivated flexible methods of data analysis.

One of the recent surprising discoveries about pericytes was their active role in the regulation of the blood flow. However, this discovery was critically acclaimed in literature, so their true regulatory function has only started to emerge. We were able to bridge the opposing points of view on this and found that pericytes are likely to be the local centers of capillary blood flow regulation. In addition, we investigated the role of octapeptide angiotensin two (ANGII) in the mechanism of hypertension on neurovascular signaling and regulation in the brain. We found that values of resting and increase in cerebral blood flow dropped upon administration of ANGII in a concentration-dependent manner, in agreement with previous findings. We also observed local constrictions of vessels in up to second order, which proves that ANGII changes their contractile tone by directly affecting smooth muscle cells as well as a subset of capillary pericytes. We developed tools for noise-robust extraction of vessel diameter traces from experimental two-photon microscopy images that could serve the broad community of biologists and neuroscientists. More importantly, we demonstrated on a number of examples how both careful data management and rigorous statistical analysis are crucial for deriving solid conclusions about the behavior of complex biological systems.

Preface

This PhD thesis was prepared at the Department of Micro-and Nanotechnology at the Technical University of Denmark in fulfillment of the requirements for acquiring a Doctor of Philosophy degree. The PhD project has been conducted in the period of November, 2014, to November, 2017

under supervision of

Henrik Flyvbjerg, associate professor, dr. scient.
Department of Micro- and Nanotechnology
Technical University of Denmark
DK-2800 Kongens Lyngby
Denmark

and co-supervision of

Kim Mortensen, senior researcher, PhD
Department of Micro- and Nanotechnology
Technical University of Denmark
DK-2800 Kongens Lyngby
Denmark

This work was funded internally by DTU Nanotech.

Department of Micro- and Nanotechnology
Technical University of Denmark
DK-2800 Kongens Lyngby
Denmark
November 14, 2017

Dmitriy L. Markovich

Acknowledgements

The work in this thesis could not have been completed without the help of, and numerous discussions with a number of people. I thank our collaborators from Novo Nordisk, Mikkel Harndahl, Klaus Bergmann and Qiansheng Ren for providing experimental data and their will to share their expertise in thrombin generation. I thank Gustav Røder, who has been our main contact person at Novo Nordisk, for giving relevant critique and doing a lot to promote our findings among his colleagues. I also thank our collaborators from the University of Copenhagen — Bodil Gesslein and Martin Lauritzen. Finally, I thank my colleagues from DTU Nanotech — Jonas Pedersen, Kim Mortensen, and Henrik Flyvbjerg for their support, help, and overall introduction to the Danish society. My supervisors Kim Mortensen and Henrik Flyvbjerg put a lot of efforts into conducting the studies and sharing ideas and experience. I learned a lot from them throughout the years, and I am grateful for that.

Contents

CHAPTER 1

Data Analysis in Thrombin Generation Experiments

1.1 Background

Thrombin, or coagulation factor II activated (FIIa), is a serine protease that is encoded by the FII gene in humans [DD87]. Its existence was hypothesized by a Baltic German physiologist Alexander Schmidt in 1872 [Sch72]. Thrombin plays the key role in the blood coagulation process (Fig. 1.1), catalyzing several coagulation-related reactions and converting the soluble fibrinogen into insoluble stands of fibrin, that together with platelets form a clot over a wound site and cause bleeding to stop.

As soon as the endothelium of the blood vessel is damaged and the blood is exposed, two simultaneous processes are started: the primary hemostasis, where the platelets form a plug at the site of injury, and the secondary hemostasis, where coagulation factors respond in a complex cascade to form fibrin strands, that strengthen the platelet clot [FF05].

Coagulation defects may cause hemorrhage or thrombosis or both, depending on the nature of the defect [Hat08]. Nowadays thrombosis is considered to be the cause of common diseases such as stroke and myocardial infarction. In Western Europe and the USA ≈3500 people per day die from thrombotic disease [She00]. Therefore, the medical importance of thrombosis can hardly be exaggerated [Sme07].

Best known among the coagulation defects are hemoplilias. Hemophilia is mostly an inherited genetic disorder that impairs the ability of the body to make blood clots, which is essential to stop bleeding. Hemophiliacs bleed as intensely as persons without it, but their bleeding time can be much longer. In the worst case even a minor injury can result in blood loss that lasts for days or weeks, and may be never healed completely. This can be fatal in areas like brain or inside joints [NHL; CDC16; NHL16]. A few examples of hemophilia are shown in Fig. 1.2.

Though there is no cure for hemophilia, it can be controlled with regular infusions of the deficient coagulation factor (the list of coagulation factors is given in Appendix A.1). This is why thrombin generation in different physiological conditions is of major interest both for research and pharmaceutical industry.

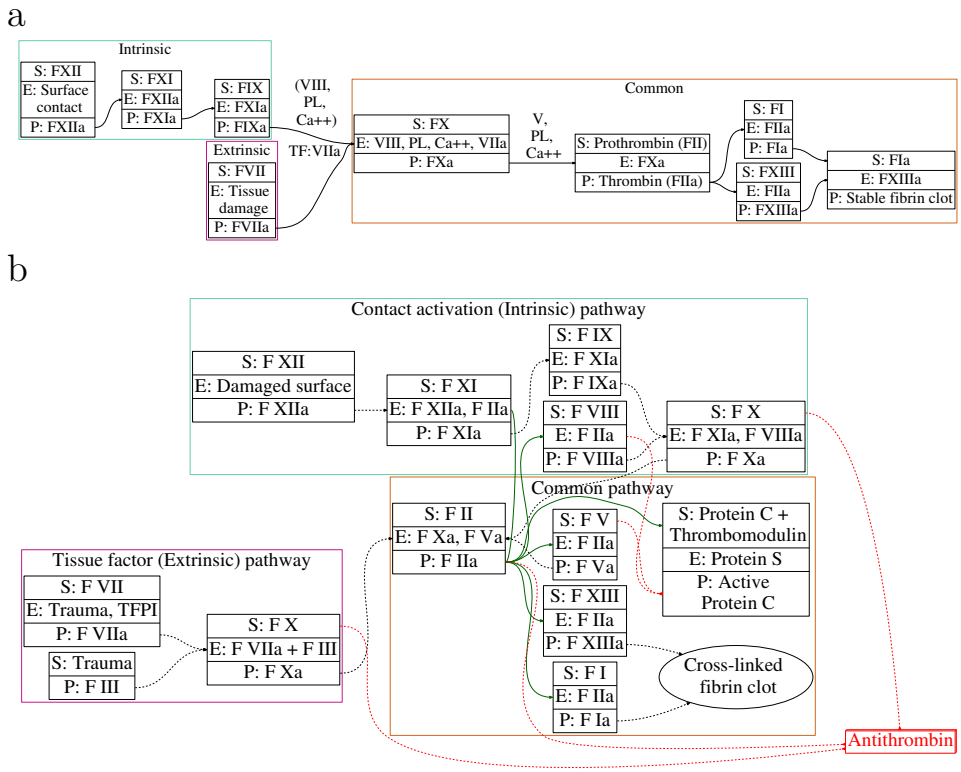


Figure 1.1: Blood coagulation cascade. **(a)** The classical coagulation cascade of secondary hemostasis has two initial pathways which lead to fibrin formation: The intrinsic activation pathway (also known as the contact pathway), and the extrinsic pathway (also known as the tissue factor pathway) which both lead to the same fundamental reactions that produce fibrin. **(b)** Modern blood coagulation pathway. It was previously thought that the two pathways of coagulation cascade were of equal importance, but it is now known that the primary pathway for the initiation of blood coagulation is the tissue factor (extrinsic) pathway [PW10; Hof02; Lon+16]. The pathways here are series of reactions, in which a zymogen (inactive enzyme precursor) of a serine protease and its glycoprotein co-factor are activated to become active components that then catalyze the next reaction in the cascade, ultimately resulting in cross-linked fibrin. In both **a** and **b**, coagulation factors are indicated by Roman numerals, with a lowercase "a" appended to indicate an active form (the list of coagulation factors is given in Appendix A.1). Each step of the cascade is a basic enzymatic reaction (Section 1.5.1). S, E, and P in the reaction scheme are shorthand notations for substrate, enzyme, and product.

Figure 1.2: Examples of hemophilia, i.e. spontaneous bleeding episodes mostly into joints and soft tissues that can lead to permanent joint damage. **(a)** Hemophilia A (FVIII deficiency): Signs of bleeding in a joint. It is the most common type of the disorder and is present in about 1 in 5000 — 10 000 male births. **(b)** Hemophilia B (FIX deficiency or "Christmas disease"): Trauma induced-bruising along a leg. **(c)** Hemophilia C (FXI deficiency): No bleeding into joints and muscles. **(d)** A knee bleed in a patient with severe hemophilia. The list of coagulation factors is given in Appendix A.1.

1.2 State-of-the-art of the measurement and application of thrombin generation

The concept of a thrombin generation test was originally described simultaneously by two groups in 1953, one using the whole blood assay [MB53], and the other using the plasma assay [PD53]. In both assays, the primary reaction tube, that served as a trigger for thrombin generation, contained plasma with or without platelets and was performing subsampling at regular intervals into secondary indicator tubes that contained a fibrinogen solution. To estimate the activity of thrombin, the clotting times of the fibrinogen solution were used, with a preceding calibration against a thrombin solution of known concentration [Bag05; Lan15].

In 1960s, several research groups focused on the theoretical aspects of thrombin generation. The pioneer idea of an enzyme cascade as a biochemical amplifier in the blood clotting mechanism was proposed by R. G. Macfarlane [Mac64]. Hemker et al. [HHL65] reported that despite the overall complexity of the clotting process, under certain conditions it behaves as a much simpler system. However, the authors also found that application of the standard methods of analysis of enzyme kinetics to the results of clotting tests is not possible without specific modifications imposed by the blood-clotting enzyme system. In continuation, Hemker et al. [Hem+67] reported that the clotting time in a thrombotest experiment bears a linear relationship to the inverse factor FX concentration under conditions in which the ratio of factor FX to the factors FII and FVII is less or equal than that in normal plasma. Also, the authors proposed a scheme for the intrinsic reaction mechanism of blood coagulation in [HK67] and a theory for the general kinetics of enzyme cascades in [HH69].

Back then, the two stage assay traditionally used for experiments, suffered from two major drawbacks — its technical difficulty and imprecision. The clots that were formed in the primary reaction tube needed to be removed with a wooden stick, and that could result in the imprecision of volume because of the obstruction of the pipette tip with fragments of clot. Moreover, the necessity to remove the clot during the procedure led to imprecision of the subsampling time. Finally, the difficulty of end-point detection of clot formation at low thrombin concentrations and substrate consumption at high thrombin concentrations were also among the sources of error in the assay. All in all, though the authors reported [MB53; PD53] that hey could identify hemophiliac patients with the test, it was very work- and time-intensive, because the measurement had to be done by continuous subsampling. A significant improvement of the experimental methods was required in the field, and in 1971 Hemker et al. showed the need for rigorous standartization of the experimental conditions and offered an improved method of the analysis of the results of the two-stage assay [Hem+71]. However, this was not enough to make a big step forward, so the field remained in stagnation for a while.

The application of thrombin generation assay was revived in 1986 by Hemker, with the introduction of several modifications that made the assay easier to perform and reduced its imprecision [HWB86]. The main differences were that the fibrinogen

solution was replaced by a chromogenic substrate and that the primary plasma sample was defibrinated, and it reduced the error of the measurement because of substrate consumption and endpoint detection, and also reduced the error related to timing and volume of subsampling. However, it was still a two-stage assay that was fairly difficult to perform without specialized equipment. The calculation of thrombin activity was complicated by the use of a chromogenic substrate, because it introduced errors to the measurement of the process of thrombin decay [HWB86].

In the meantime, the research focused on the plasma inhibitors, the antithrombins and the role of α_2-macroglobulin in the thrombin generation process. Antithrombin and α_2-macroglobulin were reported [Fis+81; TS83] to be the predominant neutralizers of thrombin. Antithrombin was the active site thrombin inhibitor, while α_2-macroglobulin neutralized thrombin activity via exosite interaction. It resulted in a continued cleavage of substrate by α_2-macroglobulin-complexed thrombin. In contrast to that, the active-site inhibited thrombin was incapable of cleaving the chromogenic substrate. Therefore, the thrombin generation curve derived from a typical small synthetic substrate wass a composite of the thrombin-time integrals produced by free thrombin and the complex of thrombin with α_2-macroglobulin — α_2-macroglobulin-thrombin. Unfortunately the major problem then was that the selected substrate exhausted quickly and did not reflect the complete amount of generated thrombin.

Simultaneously, there was also progress in modeling of thrombin generation. Khanin et al. reported [KS89] a non-linear mathematical model of the extrinsic pathway of blood coagulation and showed that is does not function if a vessel wall injury stays below a predefined threshold value. Willems et al. discussed [Wil+91] a simulation model for the production of thrombin in plasma. The good agreement between the predictions of the model and the measured activities in plasma suggested that the proposed model was essentially valid, however, some of the model parameters used were obtained from studies in purified systems and required more thorough identification.

The further development of thrombin generation measurement switched to the use of a slow reacting substrate, which permitted continuous registration of thrombin activity in the primary reaction tube, i.e. — converted the method to a much simpler one-stage assay [Hem+93]. The concept of the thrombin potential, a value that directly indicates how much of any physiological substrate can potentially be split by thrombin generated in plasma, was first introduced in [Hem+93]. Still, the method required prior defibrination of the plasma sample.

Along the way, a thorough theoretical and experimental study of the tissue factor pathway to thrombin generation was reported in [Law+94; JM94]. The studies showed that the expression or exposure of tissue factor locally at the site of the vascular injury is the rate-limiting step of the reaction which is largely a function of the initiation phase and that only a small fraction of the available factor FIX and factor FX are formed by the FVIIa-tissue factor complex during normal blood clotting.

In 1995, Hemker et al. further stressed out the importance of thrombin potential and suggested a routine test to quantitatively render the thrombin forming capacity of a plasma sample [HB95]. The test could overcome the drawbacks of clotting time

tests like no sensitivity to hypercoagulative states by determining the area under the thrombin generation curve, or the so-called endogeneous thrombin potential (ETP, see Fig. 1.4).

Later on, the replacement of the chromogenic substrate with a slow reacting fluorogenic substrate enabled continuous measurement of thrombin generation without the need for defibrination, because the signal from the fluorophore is not quenched by turbidity. Consequently, it made it possible to measure thrombin generation in platelet rich plasma (PRP) and to determine the effect of other cellular elements. The new assay was called The Calibrated Automated Thrombogram (CAT) [Hem+02] (Fig. 1.3), and its output that quantifies thrombin, *the thrombogram* (Fig. 1.4), proved to be a broad function test of the haemostatic-thrombotic mechanism of the blood [Hem+03].

Figure 1.3: Calibrated Automated Thrombogram (CAT) machine — a commercial assay for thrombin generation experiments with automated data analysis. Figure reproduced from [Gie12].

A new problem encountered with the use of fluorogenic substrates is that the thrombin activity and the fluorescent signal are not linearly related [HHD09]. Two effects account for that: Substrate consumption and the non-linearity of fluorescence intensity with the increasing concentration of the fluorophore (also called fluorescence quenching or "the inner filter effect" [PD07]). In CAT, this problem is solved by monitoring the splitting of a fluorogenic substrate and comparing it to a constant

known thrombin activity in a parallel non-clotting sample (Section 1.3). The methods of data analysis that CAT machine uses are summarized in a detailed review [HK13] by Hemker et al.

Apart from CAT, there are several commercially available thrombin generation tests, which rely on fluorogenic or chromogenic principles (manual and automated fluorogenic assay by Technoclone [Tec], chromogenic assay by Dade Behring and customized tests such as the Novel Hemostasis Assay from the Radboud University Medical Centre, Nijmegen, The Netherlands) [Lan15].

The disadvantages of commercially available assays are obvious — they are limited by the requirements of the in-built experimental protocol and the details of automated data analysis are hidden from the user. Therefore, it is not possible to fully reproduce the data analysis performed by the assay and validate the results, which is often required in research. Various fluorometers are used nowadays for thrombin generation experiments [Woo+13], but the absence of automated data analysis makes many users biased towards CAT. CAT also aims to become a standartized assay in the field, but the standartization can hardly be done by simply using a fixed instrument. On the contrary, as pointed out by Hemker himself [Hem+05] (see the caption of Fig. 1.4), it should not be based on a specific instrument, but rather on the experimental protocol and the data analysis procedure. Ideally, the experimental compounds prepared according to the same experimental protocol in different labs and measured with different fluorometers should give similar results, when a standard data analysis procedure is applied to the obtained signals.

With the new assays at hand, the theoretical research in the field of thrombin generation has also progressed a lot. A model for the stochiometric regulation of blood coagulation was reported by Hockin et al. [Hoc+02]. Maragoudaki et al. considered the mechanism of thrombin-induced angiogenesis [MTA02]. The role of platelets in thrombin generation was discussed in [MHR02]. Furthermore, a mathematical model of lipid-mediated thrombin generation was reported by Bungay et al. [BGG03], a model of tissue-factor induced thrombin generation — by K. Brummel-Ziedins et al. [Bru+05], a kinetic model for simulation of blood coagulation and inhibition in the intrinsic path — by Y. H. Qiao et al. [QLZ05], and the stochastic modeling of blood coagulation initiation — by K. Lo [LDD05]. The majority of these papers typically employ extremely complicated models with up to a hundred of free parameters.

A critical review on this approach by Hemker et al. [WHH06] confirmed the fact that might simply be anticipated from the practices of applied statistics — the large majority of thrombograms do not contain more information than can be expressed by a small number of parameters. Therefore, the information contained in a thrombogram is insufficient for the description of a physiologically realistic reaction scheme of thrombin generation. Still, more and more papers describing the possible mechanisms of thrombin generation were published in the field [Smi09; Cha+10; Wag+10] with another critical review from Hemker et al. [HKK12]. The latest trend in the field is the plasma composition-based approach to thrombin generation [Bru+14; PGA14], that also seems to apply to clinical reality [Hem15].

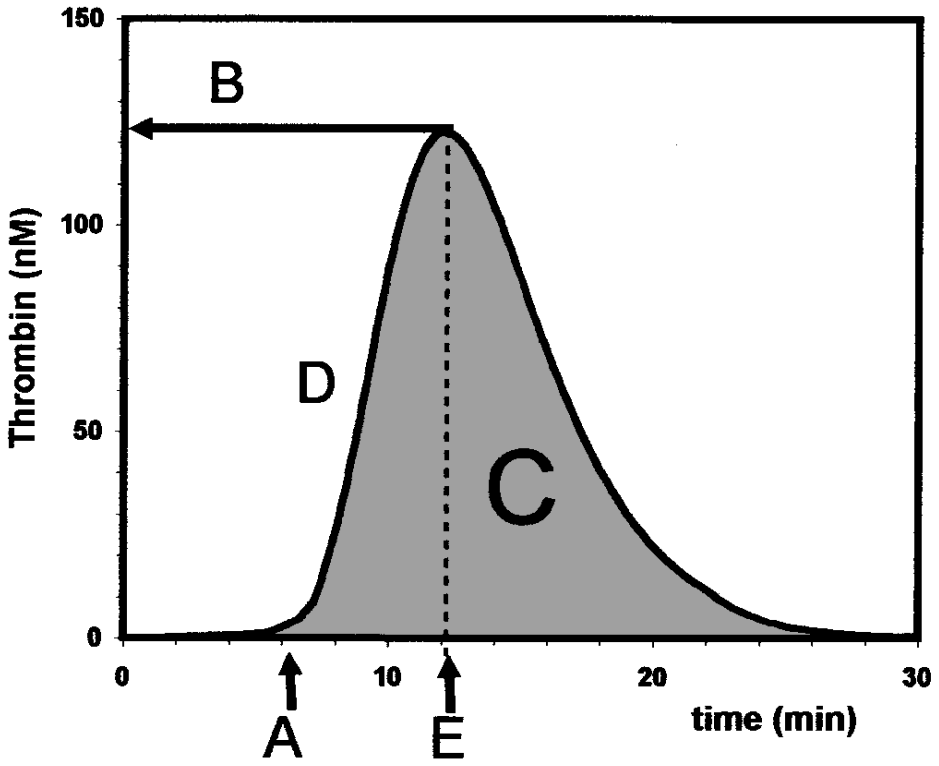
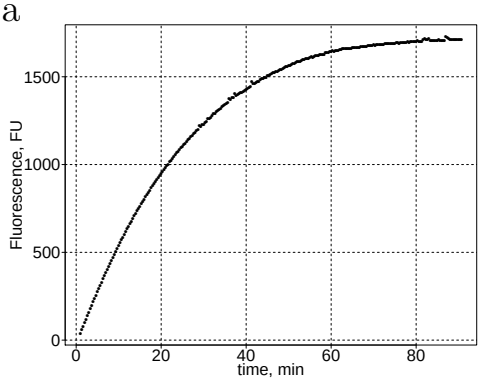


Figure 1.4: CAT thrombogram and the key parameters that characterize it. There is a lag phase at the start (A), in which only low amounts of thrombin are formed, after which the generation of thrombin starts with a sudden burst. After the burst, the concentration reaches its peak value (B) at a certain time (E), and starts decreasing afterwards. Later on, after the concentration of thrombin has effectively vanished, the curve only fluctuates around zero. The area under the thrombogram gives the ETP, or the overall workload of thrombin in the experiment (C, shaded area). The maximal rising slope of the curve (D) is the thrombin velocity. The shape of any thrombogram is not dependent on the measuring system or the experimental conditions. Figure and caption reproduced from [Hem+05].

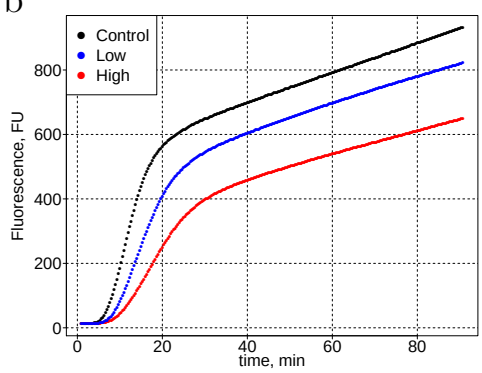
1.3 Thrombin generation experiment with CAT

A small subset of a typical dataset obtained from a thrombin generation experiment using a CAT machine is shown in Fig. 1.5a-b. It contains a calibration signal and three thrombin generation signals. Black signal is the control signal with "control" compound. The other two signals contain the "control" compound as well as a relatively low (blue) or high (red) concentration of an additional reagent, that suppresses thrombin generation.

a



b



c

CAT Parameter	Control	Low	High
Lagtime, min	5	6.33	7
Lagtime SD	0	0	0
ETP, nM · min	1064.22	919.1	701.84
ETP SD	0	0	0
Peak, nM	101.37	64.76	39.73
Peak SD,	0	0	0
ttPeak, min	11.33	14	17.33
ttPeak SD	0	0	0
VelIndex, nM / min	16.01	8.45	3.84
VelIndex SD	0	0	0
StartTail, min	35	42.33	48
StartTail SD	0	0	0
Alpha2M_Level, nM	12.32	10.44	7.99
Alpha2M_Level SD	0	0	0
TC_Initial_Slope, FU / min	63.25	63.25	63.25

d

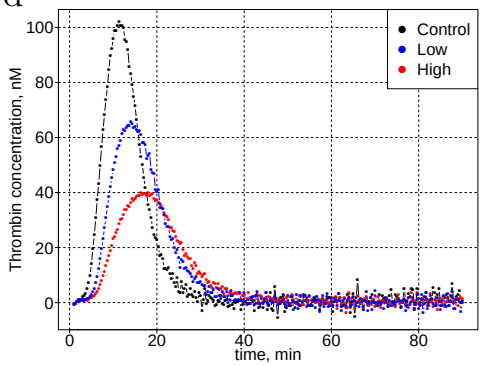


Figure 1.5: Thrombin generation experiment with CAT. **(a)** Raw calibration signal. **(b)** Raw thrombin generation signals. Blue and red colors indicate the signals with respectively relatively low and high concentrations of an additional reagent, added to the control thrombin generation compound (black). The reagent supresses thrombin generation. **(c)** The key parameters characterizing the generation of thrombin during the experiment, calculated by the data analysis algorithm of the CAT machine from the thrombograms (Fig. 1.4). **(d)** Thrombograms — time courses of thrombin concentration calculated by the CAT machine.

Automated data analysis method of the CAT machine provides the user with thrombograms (Fig. 1.5d) — the time courses of thrombin concentration, and CAT parameters (Fig. 1.5c) — the key parameters characterizing the generation of thrombin during the experiment. CAT parameters quantify the qualitative visual difference between control, low and high thrombin generation signals using several parameters: (Fig. 1.4):

- endogeneous thrombin potential (ETP) is the area under the thrombogram,
- thrombogram peak value (Peak) is the maximal concentration of thrombin during the course experiment,
- thrombin velocity index (VelIndex) is the maximal rate of change of thrombin concentration during the course of the experiment,
- time to peak (ttPeak) is the time point during the course of the experiment when Peak was observed,
- Lagtime is the end of the time interval where the difference of concentration of thrombin from the baseline is negligible,
- StartTail is the beginning of the time interval where the concentration of thrombin only fluctuates around zero.

ETP, Peak, and VelIndex decrease when the concentration of the additional reagent is increased, and vice versa for ttPeak, Lagtime and StartTail. No repetitions were made in the experiment, and therefore all the standard deviations (SD) of CAT parameters are output as zeros. Note should be made, that this convention of CAT seems rather questionable from the statistical point of view, because SD = 0 indicates an infinitely precise repetition of the experiment. Standard deviation cannot be calculated from a sample that consists of less than two observations, and it is always non-zero in the real world. The error bar on a single available observation is infinite and not zero. A better solution in this case would be to output the standard "not available" value `NA` for the missing value of SD.

1.3.1 Automated data analysis of the CAT machine

To assess the results in Fig. 1.5 critically one has to know how exactly they were obtained. This problem becomes especially relevant when a CAT machine is not available and the researcher has to perform the data analysis himself. Calculations performed by CAT are briefly described in [Gie12], closely follow the methods discussed in [Hem+93; HB95; Hem+03; HHD09; HK13], and are summarized below.

CAT starts the analysis with the calibration signal. It attempts to transform the "bending" calibration curve into a first degree polynomial, so that the initial reaction rate is maintained throughout the duration of the calibration experiment. It is achieved by fitting the raw data below a certain threshold (Fig. 1.6), but the

formula or the motivation used for it is not mentioned [Gie12]. It is very plausible that the H-transform [HHD09] is applied by CAT in this analysis. This technique claims to compensate for two effects that "bend" the calibration curve: The consumption of substrate as the reaction progresses, and the inner filter effect, i.e. the saturation of fluorescence with the increasing concentration of the fluorescent product (Section 1.7).

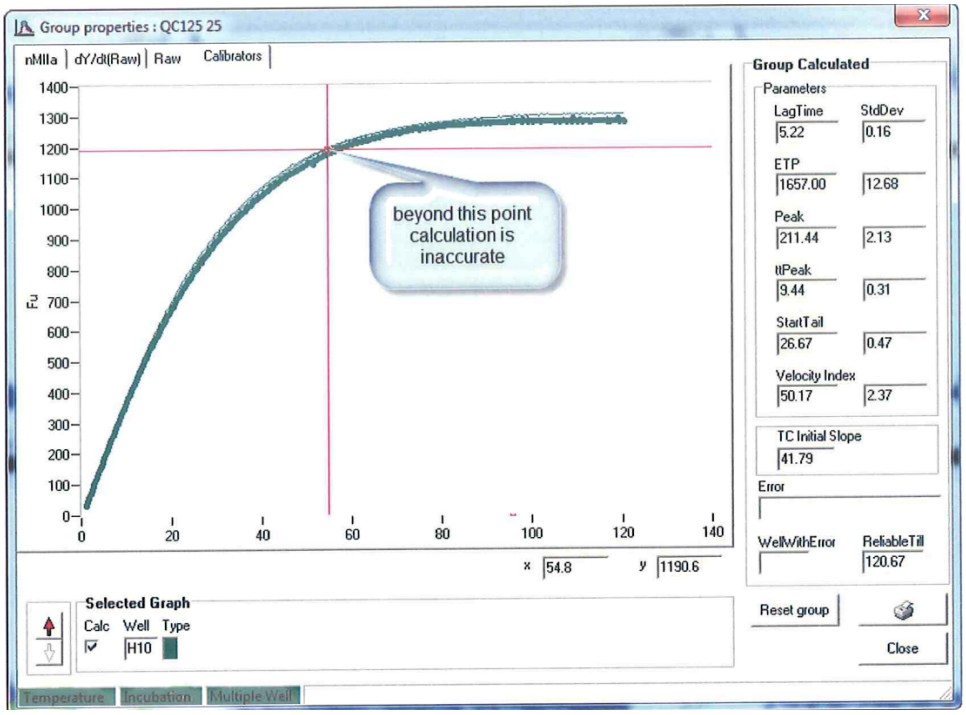


Figure 1.6: Calibrator curve measured for a prolongued period of time. At a value of 1190.6 FU the shape of the curve becomes so flat that more conversion of substrate hardly results in more fluorescence. The moment that this amount of signal occurs in the corresponding TG wells will become the ReliableTill value, in this example 120.67 minutes. Figure and caption reproduced from [Gie12].

After that, CAT uses the parameters obtained from the calibration signal to transform all thrombin generation signals assigned to be calibrated with it (Fig. 1.7).

The first derivative of the corrected thrombin generation signal is taken directly from the raw data. The initial slope of the calibration signal is used to transform the obtained first derivative from the units of fluorescence into nanomolars of thrombin, with the subtracted contribution of α_2-macroglobulin-thrombin stable complex (α_2M-T). The final result is called "thrombogram" and is provided to the user (Fig. 1.8).

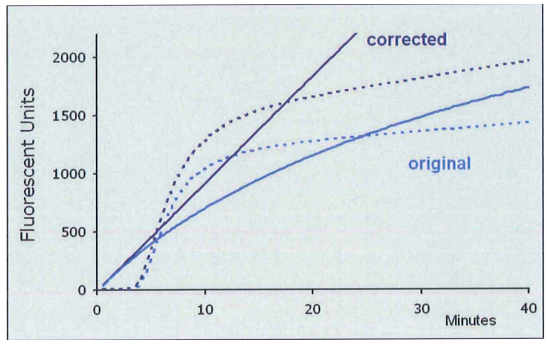


Figure 1.7: Fluorescent units in time from the first 40 minutes of a calibrator curve (solid line, light blue). Correction for inner filter effect and substrate consumption results in a straight line (solid line, dark blue). The fluorescence is measured in a thrombin generation well (dotted line, light blue) is shown before and after correction (dotted line, dark blue). Figure and caption reproduced from [Gie12].

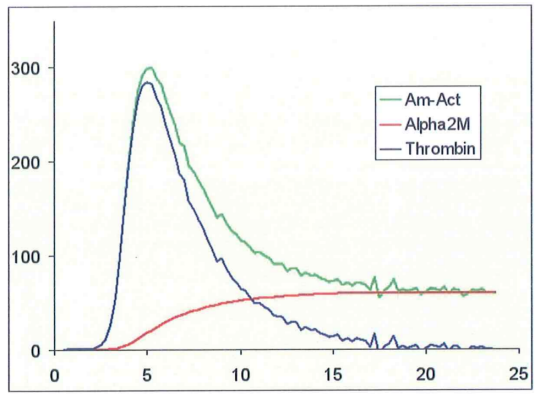


Figure 1.8: Thrombin disappears from plasma by two pathways. The bulk of thrombin is captured by antithrombin, which disables the active site of thrombin. Another part of the thrombin binds to α_2-macroglobulin. This α_2-macroglobulin-complex (α_2M-T) has no physiological function anymore but is still able to convert the small fluorogenic substrate. This means that during the reaction the substrate conversion will continue although all thrombin went to zero. Therefore the curve is analyzed by the program to see where thrombin generation has stopped. Over the whole length of curve the α_2-macroglobulin activity is then mathematically removed. The blue curve is the total amidolytic activity, which is the sum of thrombin activity and α_2M-T activity. The red curve is the α_2M-T activity, and the resulting blue curve is the corrected value when the red curve is subtracted from the green curve. Figure and caption reproduced from [Gie12].

1.4 Reverse engineering the data analysis by CAT

A detailed description of the CAT method and several issues related to the data analysis performed by it were reported in the series of papers from its creators [Hem+03; HHD09; HK13]. Following the steps described in these papers, we were able to reproduce the results of data analysis performed by the CAT machine.

1.4.1 Differentiating noisy experimental data with various stencils

Methods from the papers [Hem+03; HHD09; HK13] rely heavily on taking numerical derivatives from noisy data, and their authors usually recommend to use the basic forward or central differences. Let us first investigate taking numerical derivatives from the raw data and its consequences in more detail.

As mentioned before, it is achieved by using a certain stencil — a scheme of calculating the derivative of a sampled continuous variable. Consider $F(t)$ to represent the fluorescence signal as a function of time of the experiment t. A measured experimental signal is then represented as two vectors F_k and t_k, where $k = 1 : N$, N is the overall number of points in the signal. It is often the case with fluorometers that the sampling time $\Delta t = t_{k+1} - t_k$ is constant for all k. Then, the easiest way to calculate the first derivative of such a signal is to use the basic forward difference

$$F'_k = \frac{F_{k+1} - F_k}{\Delta t} + O(\Delta t) \qquad k = 1 : N - 1 \ , \tag{1.1}$$

where F'_k is the value of the first derivative with index number k, F_k is the "current" value of the signal, and F_{k+1} is the "next" value of the signal. This basic stencil has the first order of accuracy $O(\Delta t)$ (big O notation). It can be improved by increasing the number of points in it and there is a number of ways to introduce a stencil for calculating the first derivative with a higher order of accuracy. For example, the inspection of the thrombograms from the CAT machine shows that if the raw signal has N points in it, then the corresponding thrombogram has $N - 1$ or $N - 2$ points (points with index numbers N or N and $N - 1$ are missing in the thrombogram, Appendix A.2). It implies that the algorithm of data analysis of the CAT machine must have used either Eq. 1.1 or the following stencil

$$F'_k = \frac{F_{k+2} + F_{k+1} - 2 \cdot F_k}{3 \cdot \Delta t} + O(\Delta t^2) \qquad k = 1 : N - 2 \ , \tag{1.2}$$

which has a second order of accuracy $O(\Delta t^2)$. The same order of accuracy can be achieved by using Eq. 1.1 if the stencil is reintroduced as

$$F'_{k+1/2} = \frac{F_{k+1} - F_k}{\Delta t} + O(\Delta t^2) \qquad k = 1 : N - 1 \ , \tag{1.3}$$

so that the calculated values of the first derivative are half-step shifted from the original grid (offgrid stencil).

Increasing the number of points in the stencil does not always improve the order of accuracy. For example, the following stencil is calculated using four points in total, but still has the first order of accuracy

$$F'_k = \frac{F_{k+3} + F_{k+2} - F_{k+1} - F_k}{4 \cdot \Delta t} + O(\Delta t) \qquad k = 1 : N - 3 \ . \tag{1.4}$$

Now, let us simply take a thrombin generation signal and apply various stencils to it. The raw signal is shown in the inset plot of Fig. 1.9, as well as its first derivatives obtained using the stencils from eqs. 1.1-1.4 are shown in Fig. 1.9.

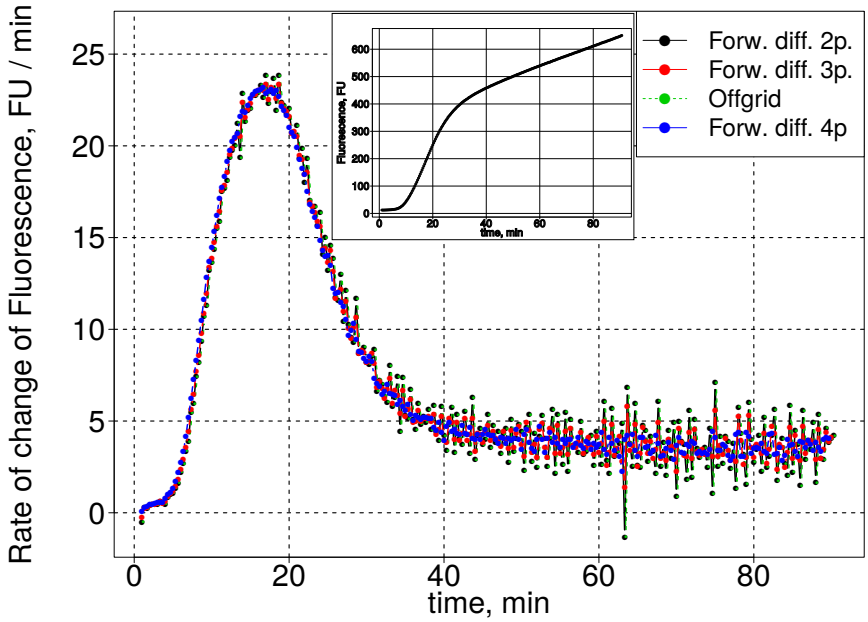


Figure 1.9: First derivative of the thrombin generation signal from the inset plot taken numerically using various stencils — Eq. 1.1 (black), Eq. 1.2 (red), Eq. 1.3 (green), and Eq. 1.4 (blue). The more points are included in the stencil, the less noisy is the first derivative.

Visually, there is little difference between the black and the green curves, and this follows from the fact that a half-step time shift of 20 / 2 = 10 s between them is hard to resolve on the timescale of the plot. The red and the blue curves seem to be more smooth than the black and the green, and that is because they utilize more points to calculate the value of the derivative.

There are two properties of stencils that are not often discussed, but are important to investigate in this context: How do they affect the correlations in the data (Section 1.4.2) and how does their magnitude response behave (Section 1.4.3).

1.4.2 Correlogram of the various stencils

Clearly, the datapoints of the raw thrombin generation signal are correlated, because it represents a process of fluorescence production in time. The value of fluorescence of every next point of the signal thus depends on the value of the current point. Therefore, let us take artificial uncorrelated data instead to illustrate the correlations in it caused by a numerical stencil.

The artificial data is very simple: a vector of independent identically distributed (iid) random variables drawn from the Gaussian distribution with $\mu = 0$ and $\sigma = 1$. The artificial data is shown in Fig. 1.10a, and the first derivative of the artificial data taken using the stencils from Eq. 1.1-1.4 — in Fig. 1.10b.

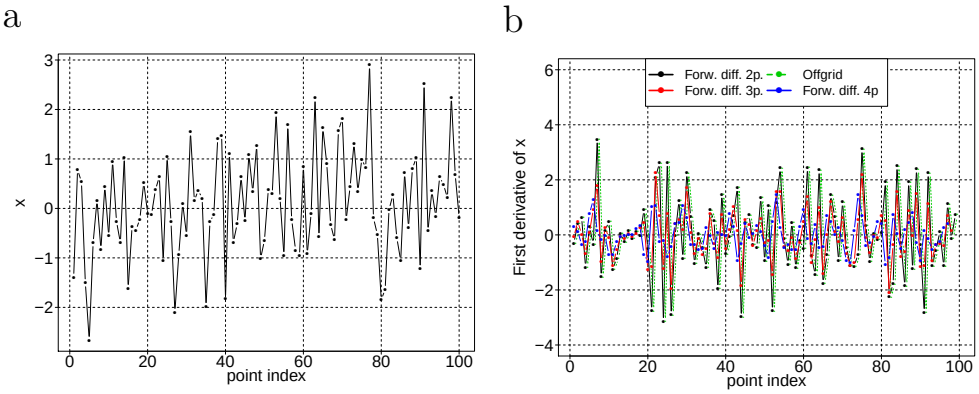


Figure 1.10: **(a)** Artificial data: a vector of 100 independent identically distributed (iid) random variables drawn from the Gaussian distribution with $\mu = 0$ and $\sigma = 1$. **(b)** First derivative of the artificial data from **a** taken numerically using various stencils — Eq. 1.1 (black), Eq. 1.2 (red), Eq. 1.3 (green), and Eq. 1.4 (blue).

There is a standard way to assess whether there is any correlation between the data points as a function of the lag index between them — to analyze the autocorrelation function of the data by plotting its correlogram. In Fig. 1.10 we used just 100 points for illustrative purposes. Reliable calculation of the correlogram requires much better statistics, therefore we generate 10000 points instead of 100 for this purpose. The correlogram of the refined artificial data and its first derivative taken by stencils from Eq. 1.1-1.4 is shown in Fig. 1.11.

As expected, the raw data is uncorrelated: its autocorrelation function equals exactly 1 at lag index 0 because of normalization, and is essentially zero for all other lag indexes. For the stencils from Eq. 1.1 and Eq. 1.3, all points with lag index 1 are sufficiently negatively correlated because there are 2 points in total in the stencil formula. For the stencils from Eq. 1.2 and Eq. 1.4, that have respectively 3 and 4 points in total in the stencil formula, the correlation vanishes after lag indexes 2 and 3, as the dashed vertical lines in Fig. 1.11 indicate. For the blue curve, even the type

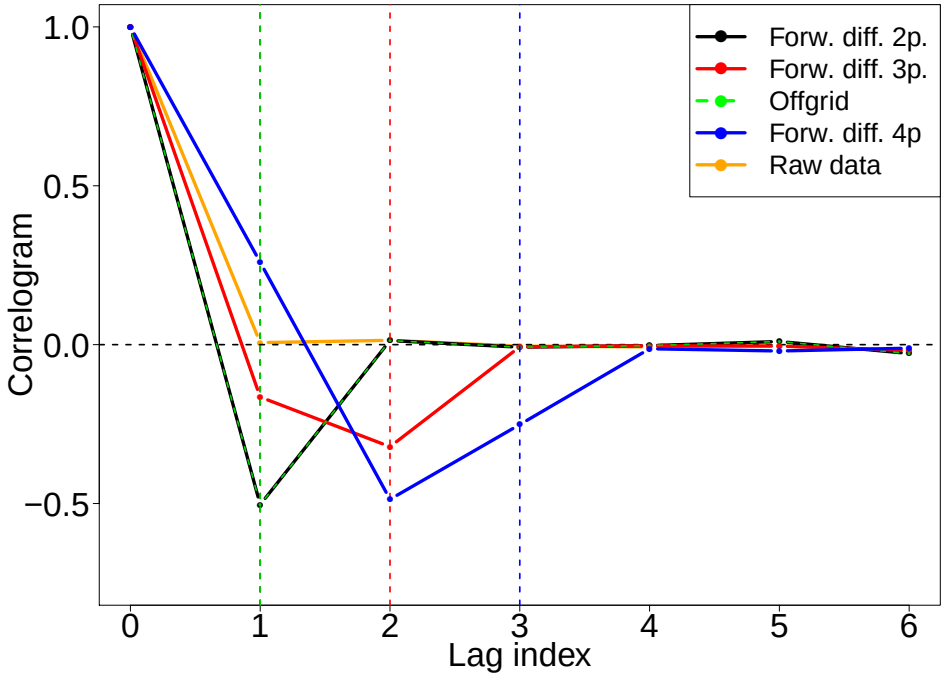


Figure 1.11: Correlograms of the first derivative of the 10000 iid random variables taken numerically using various stencils — Eq. 1.1 (black), Eq. 1.2 (red), Eq. 1.3 (green), and Eq. 1.4 (blue). The orange line shows the correlogram of the raw random variables. Dashed vertical lines of different color indicate the lag index after which the correlation in the corresponding data has vanished.

of correlation is different at different lag indexes — negative correlation at lag indexes 2 and 3 and positive correlation at lag index 1.

In conclusion, whenever a stencil is applied to originally uncorrelated data to calculate its derivative, it introduces correlations to the derivative. The type (positive or negative) and the period (lag index) of the correlations are defined by the formula of the stencil.

1.4.3 Magnitude response of the various stencils

As we have seen in Fig. 1.9, the amplitude of noise in the first derivative of a thrombin generation signal taken numerically is larger than in the raw data. Therefore, the noise-robust properties of the stencil are also important to consider in this context. To assess these properties, we have to investigate the magnitude response of the stencil. The magnitude response of the stencil F'_k is defined as

$$H(\omega) = \left| \frac{F'_k(\omega)}{F_k(\omega)} \right| \ , \tag{1.5}$$

for angular frequencies $\omega \in (0, \infty)$. The magnitude response of a stencil is normally compared to the magnitude response of the ideal differentiator. For example, let us consider a signal $F(t) = A \cdot e^{i \cdot \omega \cdot t}$. The derivative of this signal can be calculated analytically, and it is $F'(t) = i \cdot \omega \cdot F(t)$. So, the magnitude response of the ideal differentiator $H_{id}(\omega)$ is

$$H_{id}(\omega) = \left| \frac{i \cdot \omega \cdot F(t)}{F(t)} \right| = \omega \ , \tag{1.6}$$

or simply in linear proportion to the signal angular frequency. Same calculation repeated for the second derivative of $F(t)$, $F^2(t)$, will give ω^2, and in general for the derivative of order n, $F^n(t)$, the magnitude response of the ideal differentiator is ω^n.

However, Eq. 1.6 is not valid for the stencil in Eq. 1.1. Let us sample $F(t) = A \cdot e^{i \cdot \omega \cdot t}$ so that $F_k = A \cdot e^{i \cdot \omega \cdot t_k}$. We assume equidistant sampling, i.e. we assume $\Delta t = const$. Therefore, the next point $F_{k+1} = A \cdot e^{i \cdot \omega \cdot t_{k+1}} = A \cdot e^{i \cdot \omega \cdot t_k} \cdot e^{i \cdot \omega \cdot \Delta t}$. So, we plug this into Eq. 1.5 using Eq. 1.1, and we get

$$H(\omega) = \left| \frac{e^{i \cdot \omega \cdot t_k} \cdot \left(e^{i \cdot \omega \cdot \Delta t} - 1 \right)}{e^{i \cdot \omega \cdot t_k} \cdot \Delta t} \right| = \left| \frac{e^{i \cdot \omega \cdot \Delta t} - 1}{\Delta t} \right| \ . \tag{1.7}$$

With Eq. 1.6 and Eq. 1.7, we can compare the basic forward difference to the ideal differentiator in terms of their magnitude responses. In the same fashion as Eq. 1.7, we can obtain the magnitude responses of all the stencils from Section 1.4.1. For all data from Fig. 1.5, the sampling time was 20 seconds or rather the sampling frequency f_s was 0.05 Hz. According to Shannon's theorem, it allows us to perfectly reconstruct the signal for the bandlimit $B < f_s/2 = 0.025$ Hz. With f_s as the upper frequency, we make the comparison of all the stencils from Section 1.4.1 to the ideal differentiator. The comparison is presented in Fig. 1.12a.

The figure shows that all stencils suppress the high-frequency components relative to the ideal differentiator, but the stencils from Eq. 1.2 and especially Eq. 1.4 do that much more efficiently. We saw that qualitatively before in Fig. 1.9 and pointed out that this is a consequence of the fact that the stencils from Eq. 1.2 and Eq. 1.4 have more points to average over to produce the value of the first derivative. With Fig. 1.12a we are able to quantify these differences for a fixed sampling rate Δt of the raw data.

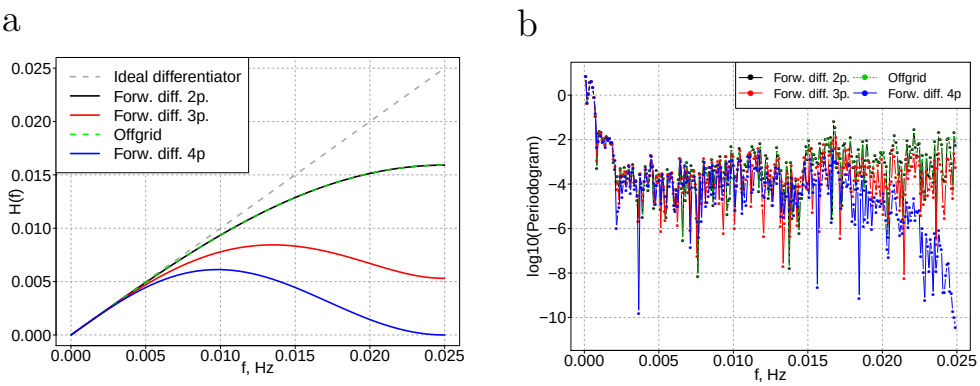


Figure 1.12: **(a)** Magnitude responses of the various stencils — Eq. 1.1 (black), Eq. 1.2 (red), Eq. 1.3 (green), and Eq. 1.4 (blue). The magnitude response of the ideal differentiator is represented by the gray dashed line. **(b)** Periodograms of the derivatives from Fig. 1.9 on the semi-log scale.

The general problem of differentiating noisy data has been investigated before in details [Pre+07; Cha11], but the majority of suggested methods are relatively difficult for implementation. Instead, in this subsection we have shown a cheap way to reduce the noise in the first derivative by simply using one (Eq. 1.4) of the so-called smooth noise-robust differentiators [Hol08]. The main purpose of these stencils is that $H(\omega)$ should be as close as possible to the response of the ideal differentiator $H_{id}(\omega)$ in the low frequency region and smoothly tend to zero towards the highest angular frequency $\omega = \pi \cdot f_s$. As we see from Fig. 1.12a, this stencil does exactly what it was designed to do — it is close to the response of the ideal differentiator in the low frequency region and smoothly approaches zero towards the highest frequency.

Finally, let us do a simple check of the theoretical result in Fig. 1.12a. Let us investigate the spectral properties of the derivatives from Fig. 1.9 by plotting their periodograms. If the theory is correct, we should observe relatively low values of the high-frequency components in the periodograms of the derivatives produced by stencils that are more noise-robust. The periodograms are shown in Fig. 1.12b. The expected difference between the stencils is clearly visible for frequencies $f \geq 0.02$ Hz.

Selected results of similar analysis for the second derivative taken from raw data are given in Appendix A.3.

1.4.4 Extraction of calibration factor from the calibration signal

Figure 1.5c shows that CAT estimated the initial slope of the calibration curve, TC_Initial_Slope to be 63.25 FU / min. This result can be checked by taking the first derivative of the calibration signal directly from the raw data (Fig. 1.13a, the value of TC_Initial_Slope given by CAT is represented by the red line).

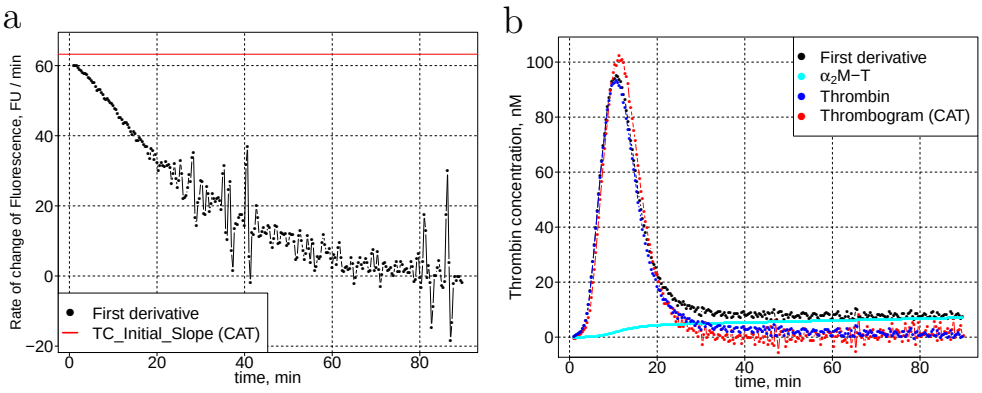


Figure 1.13: **(a)** First derivative of the calibration signal in Fig. 1.5a. The red line represents the value of TC_Initial_Slope from Fig. 1.5c. **(b)** Reverse-engineered thrombogram of the control thrombin generation signal from Fig. 1.5b. Black points represent the first derivative of the signal, blue points represent the contribution of thrombin, cyan points represent the contribution of α_2M-T, and red points represent the CAT thrombogram from Fig. 1.5d. The value of the calibration factor used to translate from FU of fluorescence to nM of thrombin is 1.71 nM · min / FU, as defined by Eq. 1.8. First derivatives were calculated from raw data using Eq. 1.4.

The first derivative of the calibration signal in Fig. 1.13a was obtained using Eq. 1.4, and the result is close to the CAT machine result — the initial rate of change of fluorescence of the calibration signal can be easily read from the plot in Fig. 1.13a, and equals approximately 60 FU / min. This value allows to calculate the calibration factor CF, defined as [Hem+03; HK13]

$$F'_{init} \cdot CF = e_0 \ , \tag{1.8}$$

where F'_{init} is the initial rate of change of fluorescence of the calibration signal (TC_Initial_Slope parameter), and e_0 is the concentration of α_2M-T in thrombin equivalent units (Appendix A.4) in its compound. Typically, the stock concentration of α_2M-T in thrombin equivalent units in a vial of Thrombin Calibrator [BV] is 650 nM, and a typical dilution by a factor of 6 is suggested by the CAT protocol. Therefore, the value of e_0 in the calibration signal is 650 / 6 = 108.33 nM, and the value of CF used by CAT can therefore be estimated as CF = 108.33 / 63.25 = 1.71 nM · min

/ FU. This value will be used to transform the first derivative of thrombin generation signals from units of fluorescence to nanomolars of thrombin (Section 1.4.5).

1.4.5 CAT parameters from the thrombin generation signal

As discussed above in Section 1.3, the derivative of the thrombin generation signal has to be calculated first, then the contribution of α_2M-T has to be subtracted from it, and finally the result has to be translated from FU of fluorescence to nM of thrombin using the calibration factor. The result of applying all these procedures to the control thrombin generation signal from Fig. 1.5b is presented in Fig. 1.13b. Calculation of the first derivative of the control thrombin generation signal is straightforward using Eq. 1.4. Correction for α_2M-T contribution comes from the fact that [Hem+93; HB95]

$$T \xrightarrow{k_{\alpha_2 M}} \alpha_2 M\text{-}T \ ; \qquad \alpha_2 M\text{-}T(t) = k_{\alpha_2 M} \cdot \int_0^t T(t')dt' \ , \tag{1.9}$$

where T is thrombin, $\alpha_2 M$-T is α_2-macroglobulin-thrombin, and $k_{\alpha_2 M}$ is the reaction rate of conversion of thrombin into α_2-macroglobulin-thrombin. Eq. 1.9 states that α_2M-T is formed from thrombin via a first order reaction with a reaction rate $k_{\alpha_2 M}$. Mathematically, it means that the concentration of α_2M-T as a function of time can be obtained by performing the integration of the concentration of thrombin as a function of time and multiplying the result by $k_{\alpha_2 M}$. It is the contribution of α_2M-T that causes the non-zero end level of the black curve in Fig. 1.13b. An easy calculation allows to roughly account for this contribution if we replace the actual $T(t)$ with the first derivative of the raw signal. Calculating its integral is straightforward using for example the trapezoidal rule of integration [Pre+07]. Next, we manually find a value of $k_{\alpha_2 M}$ that scales the result to the end-level of the first derivative. After that, the contribution of α_2M-T is subtracted from the first derivative, the result is multiplied by the value of the calibration factor CF obtained from the calibration signal, giving the rough estimate of the thrombogram (blue points in Fig. 1.13b). All the abovementioned steps were taken from [HK13] and are visualized in Fig. 1.13b.

Thrombogram obtained in Fig. 1.13b is close to the CAT thrombogram, but they are not identical. CAT machine data analysis algorithm applied a transformation to the raw thrombin generation signal based on the fit to the calibration signal, to compensate for the effects of substrate consumption and fluorescence quenching (Fig. 1.7). Plausibly, it used a different stencil for calculating the first derivatives, and also used additional smoothing [Woo+13]. The idea of the current section however was not to obtain exactly the same values of parameters as CAT machine, but rather to summarize and demonstrate in more details the principles of data analysis reported in [Hem+93; HB95; Hem+03; HHD09; HK13] and to transparently apply them to the data available in our study. The main result of the analysis described above is that following the principles from the abovementioned papers, we were able to produce the results that are comparable with the original output of the CAT machine. It means that the CAT machine is not really a black box, but rather a gray box.

1.4.6 Results and discussion

In this section, we reverse engineered the data analysis of the CAT machine relying entirely on the methods reported by its creators. Since these methods use numerical differentiation, we considered the effects of applying various stencils to raw data. We found that noise-robust stencils are much better suited for the job than their conventional variants because of the properties of their magnitude response. At the same time, the use of any numerical stencil on the originally uncorrelated data will inevitably introduce correlations in the resulting derivative. We were able to produce results that are close enough to the original output of the CAT machine, so from now on we consider it a gray box rather than a complete black box. The reverse engineering that we performed revealed a number of possibilities to improve the methods of data analysis that CAT machine relies on. We discuss that in details in subsequent sections.

1.5 Michaelis-Menten kinetics in calibration data

1.5.1 Basic enzymatic reaction

1.5.1.1 Introduction

Enzymes are typically proteins that act on some other molecules — substrates. Substrates are binded to the active site of enzyme and are transformed into products through enzymatic mechanism with two intermediate steps: the formation of enzyme / substrate complex followed by the conversion of the complex to product and enzyme. This is one of the most basic enzymatic reactions, first proposed by Michaelis and Menten [LL13]. The basic enzymatic reaction therefore reads:

$$E + S \underset{k_r}{\overset{k_f}{\rightleftharpoons}} C \xrightarrow{k_{cat}} E + P \ , \tag{1.10}$$

where S is the substrate, E is the enzyme, C is the intermediate enzyme-substrate complex, and P is the product of the reaction. k_f is the kinetic constant of forward reaction, k_r is the kinetic constant of reverse reaction, and k_{cat} is the turnover rate. Initially, the concentration of substrate s equals s_0 and the concentration of enzyme e equals e_0. Throughout the course of the reaction, E is used to convert S into P, so the final concentration of product p is exactly s_0, and the final concentration of enzyme e is e_0.

The law of mass action states that the rate of the reaction is proportional to the concentration of the reactants in the power of stochiometric coefficients. It can be applied to 1.10, accounting all the IN and OUT (see Full mass balance for an open system) terms for every reactant, and result in a system of four first-order differential equations (Eq. 1.11) for the concentrations:

$$\begin{aligned}
\frac{ds}{dt} &= -k_f \cdot e \cdot s + k_r \cdot c \ , \\
\frac{de}{dt} &= -k_f \cdot e \cdot s + (k_r + k_{cat}) \cdot c \ , \\
\frac{dc}{dt} &= k_f \cdot e \cdot s - (k_r + k_{cat}) \cdot c \ , \\
\frac{dp}{dt} &= k_{cat} \cdot c \ ,
\end{aligned} \tag{1.11}$$

where s, e, c, p denote the concentrations of S, E, C, P as functions of time. Naturally, four initial conditions are also required to solve this system. They are: nonzero initial values of substrate and enzyme concentration, and zero complex and product concentrations at $t = 0$:

$$s(0) = s_0, \qquad e(0) = e_0, \qquad c(0) = 0, \qquad p(0) = 0 \ . \tag{1.12}$$

Enzyme E is a catalyst in Eq. 1.10, that only faciliates the reaction, so its initial concentration is conserved during the course of the reaction. If we add $\frac{de}{dt}$ and $\frac{dc}{dt}$

together from Eq. 1.11 we get the conservation law of the enzyme concentration

$$\frac{de}{dt} + \frac{dc}{dt} = 0 \Rightarrow e(t) + c(t) = e_0 \ . \tag{1.13}$$

In the same fashion, we derive the substrate-complex-product conservation law

$$\frac{ds}{dt} + \frac{dc}{dt} + \frac{dp}{dt} = 0 \Rightarrow s(t) + c(t) + p(t) = s_0 \ . \tag{1.14}$$

Thorough and in-depth mathematical analysis of the system 1.11 can be found in [Mur93; SM97; Ber10]. The highlights of the reported findings of primary relevance to data analysis in thrombin generation experiments will be presented below.

One of the key results from the literature is that there is no exact analytic solution for the system 1.11 in the general case, i.e. — for arbitrary combinations of parameter values (initial concentrations and reaction rates). However, system 1.11 can be easily solved numerically on a computer (Appendix A.5) for the general case, and an example of that is presented in Fig. 1.14. In several cases, relevant for thrombin generation, system 1.11 can be approximately solved analytically (Section 1.5.1.2).

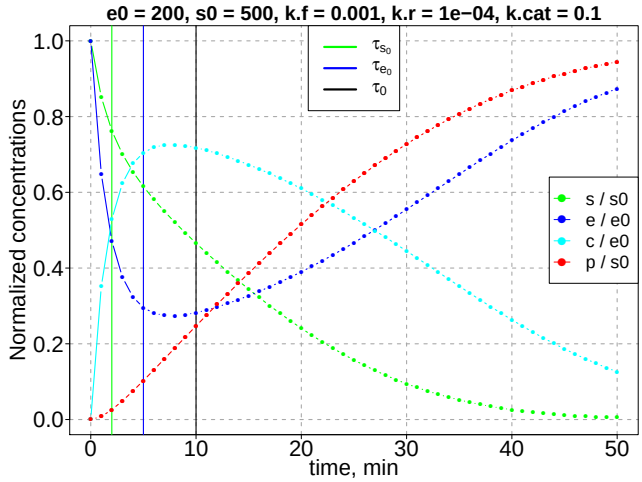


Figure 1.14: Example of a numerical solution (Appendix A.5 for details) of system 1.11: normalized concentrations of compartments from Eq. 1.10 as functions of time. The concentration of substrate (green) decreases as the product (red) is formed, whereas enzyme concentration (blue) drops initially to form the intermediate complex (cyan) and then gets back to the initial value. The main title of the plot shows the values of parameters used, in nM for e_0 and s_0 and in nM / min or 1 / min for kinetic constants. Three vertical lines visualize the three different timescales in the kinetics — substrate, enzyme, and catalytic (Section 1.5.1.2). Time step of 0.1 min was used to obtain the solution.

1.5.1.2 Early- and late-times analytic solutions of Michaelis-Menten kinetics

The system 1.11 is presented in the "raw" form, i.e. exactly as it yields from the basic enzymatic reaction 1.10. Normalization of the system gives a convenient way to explore the interrelations between the equations in it. Normalized variables for the system 1.11 can be defined as follows:

$$\widetilde{s} = \frac{s}{s0}, \qquad \widetilde{e} = \frac{e}{e0}, \qquad \widetilde{c} = \frac{s}{e0}, \qquad \widetilde{p} = \frac{p}{s0}, \qquad \tau = k_{cat} \cdot t \ , \tag{1.15}$$

using the initial concentrations of enzyme and substrate e_0 and s_0 and the turnover rate k_{cat}. The normalized version of the system 1.11 using the variables from Eq. 1.15 reads

$$\begin{aligned}
\frac{d\widetilde{s}}{d\tau} &= -\frac{k_f \cdot e_0}{k_{cat}} \cdot \widetilde{e} \cdot \widetilde{s} + \frac{e_0}{s_0} \cdot \frac{k_b}{k_{cat}} \cdot \widetilde{c} \ , \\
\frac{d\widetilde{e}}{d\tau} &= -\frac{k_f \cdot s_0}{k_{cat}} \cdot \widetilde{e} \cdot \widetilde{s} + \left(1 + \frac{k_b}{k_{cat}}\right) \cdot \widetilde{c} \ , \\
\frac{d\widetilde{c}}{d\tau} &= \frac{k_f \cdot s_0}{k_{cat}} \cdot \widetilde{e} \cdot \widetilde{s} - \left(1 + \frac{k_b}{k_{cat}}\right) \cdot \widetilde{c} \ , \\
\frac{d\widetilde{p}}{d\tau} &= \frac{e_0}{s_0} \cdot \widetilde{c} \ .
\end{aligned} \tag{1.16}$$

In this form, the normalized system reveals three different kinetic timescales, that are

$$\tau_{e_0} = \frac{1}{k_f \cdot e_0}, \qquad \tau_{s_0} = \frac{1}{k_f \cdot s_0}, \qquad \tau_0 = \frac{1}{k_{cat}} \ , \tag{1.17}$$

where τ_{e_0} is the enzyme characteristic time, τ_{s_0} is the substrate characteristic time, and τ_0 is the catalytic characteristic time. In Fig. 1.14, all three characteristic times are visualized with vertical lines. In general, the minimal characteristic time of Eq. 1.17 is the upper boundary to early-times of the kinetics, and the maximal characteristic time is the lower boundary to late-times. If we take Fig. 1.14 as an example, it means that before the green vertical line (τ_{s_0}) the kinetics is still in the early stage, and after the black vertical line (τ_0) it is in the late stage. It follows from Eq. 1.17 that if k_{cat} and k_f are fixed, then the values of the timescales depend on s_0 and e_0. Fig. 1.15 shows a few more examples of the kinetics with different values of e_0 and s_0 but with the kinetic constants and measurement time fixed.

In the regime of early-times kinetics, there is one key assumption about the concentration of substrate — since up to the green vertical line in Fig. 1.15 there is still a lot (around 90%) of substrate left, its rate of change $\frac{d\widetilde{s}}{d\tau}$ can be considered negligible:

$$\frac{d\widetilde{s}}{d\tau} = 0 \ . \tag{1.18}$$

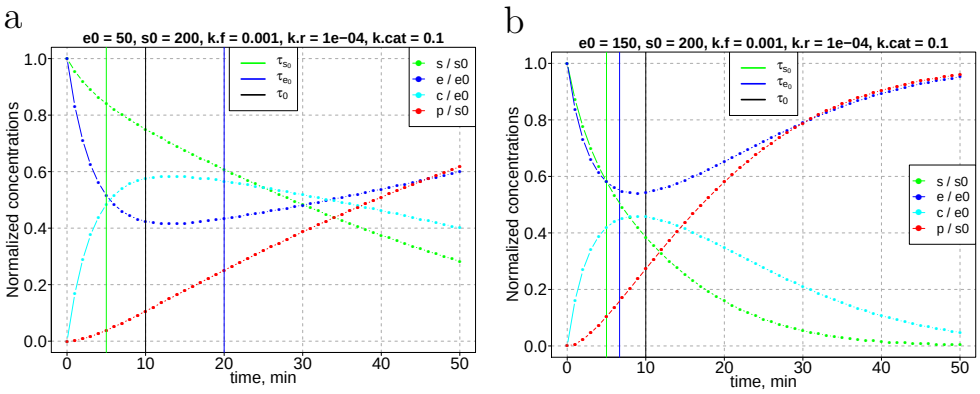


Figure 1.15: Numerical solutions of system 1.11 for two different sets of parameters (parameter values are shown in the figure titles). Characteristic times are illustrated with vertical lines.

Then, we can obtain the analytic solution for the intermediate complex concentration $c(t)$, that reads [Mur93]

$$c(t) = \frac{e_0}{1+K} \cdot \left(1 - e^{-(1+K)\cdot k_f \cdot s_0 \cdot t}\right) \quad , \tag{1.19}$$

where $K = (k_r + k_{cat})/(k_f \cdot s_0) = \frac{K_M}{s_0}$. K_M here is the well-known Michaelis constant. Using the relation $\frac{dp}{dt} = k_{cat} \cdot c$, the solution for p can be obtained

$$p(t) = \frac{k_{cat} \cdot e_0}{(1+K)} \cdot \left(t - \frac{1 - e^{-(1+K)\cdot k_f \cdot s_0 \cdot t}}{(1+K) \cdot k_f \cdot s_0}\right) \quad , \tag{1.20}$$

as well as the early-times solutions for $e(t)$ and $s(t)$ that are obtained using the conservation laws

$$e(t) = e_0 - c(t), \qquad s(t) = s_0 - p(t) - c(t) \quad . \tag{1.21}$$

Comparison of early-time analytic solutions to numerical solutions is presented in Fig. 1.16. Best correspondence between the solutions is achieved when the initial substrate concentration s_0 significantly exceeds the initial enzyme concentration e_0 (Fig. 1.16b).

On the contrary, stating $\frac{d\tilde{c}}{d\tau} = 0$ allows to obtain closed-form solutions for concentrations of all reactants at late times of the kinetics [SM97]. The formula for product concentration $p(t)$ at late times reads:

$$p(t) = s_0 \cdot \left(1 - \frac{W\left[\frac{s_0}{K_M} \cdot e^{\frac{s_0}{K_M}} \cdot e^{-\frac{k_{cat}\cdot e_0 \cdot t}{K_M}}\right]}{\frac{s_0}{K_M}}\right), \tag{1.22}$$

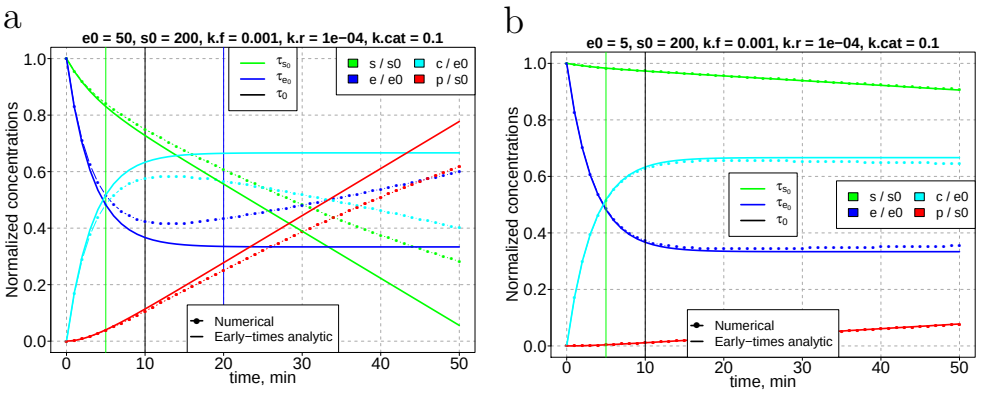


Figure 1.16: Numerical solutions of system 1.11 for two different sets of parameters (shown in the titles), compared to early-times analytic solutions. Characteristic times are illustrated with vertical lines.

where $W(x)$ is the Lambert W function [Cor+96], and the formulas for the other compartments can be derived in the same way as for the early-times solutions.

Comparison of late-times solutions to numerical solutions is presented in Fig. 1.17. Again, the best correspondence between the solutions is achieved when the initial substrate concentration significantly exceeds the initial enzyme concentration (Fig. 1.17b).

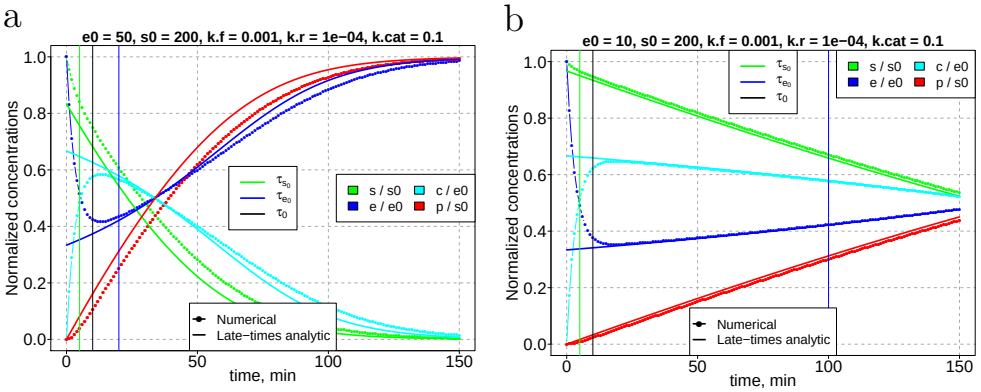


Figure 1.17: Numerical solutions of system 1.11 for two different sets of parameters (shown in titles) compared to late-times analytical solutions. Characteristic times are illustrated with vertical lines.

In conclusion, there exist two sets of analytic solutions for the system 1.11 — early-times and late-times solutions. Both sets have the best correspondence to the numerical solutions when $s_0 \gg e_0$, early-times set — for early times of the kinetics (for $0 \leq t \leq \tau_{s_0}$), and late-times set — for late times of the kinetics (for $t \geq \tau_{e_0}$).

1.5.2 Applying Michaelis-Menten kinetics to calibration data

Michaelis-Menten kinetics of the basic enzymatic reaction is often discussed in the context of the calibration signal [HHD09; HK13]. The idea of fitting the solutions of the Michaelis-Menten kinetics directly to raw data is not new as such [PD07], but to our best knowledge, it has not been reported for calibration data in thrombin generation experiments.

Calibration signal is essentially the time course of the concentration of fluorescent product as a result of activity of α_2M-T on a fluorogenic substrate in a plasma or buffer. Therefore, the straightforward model for calibration data is the basic enzymatic reaction (Eq. 1.10). Typically in calibration experiments, the ratio of the values s_0 (the concentration of Z-Gly-Gly-Arg-AMC in FluCa [Hem+03; BV], 2.5 mM in stock) and e_0 (the concentration of α_2M-T in thrombin equivalents in Thrombin Calibrator, 650 nM in stock) is $s_0/e_0 \gg 1$, so Eq. 1.22 is the natural candidate to use as a model for the calibration signal. The correspondence between the numerical and the analytic solution (Eq. 1.22) is excellent in this case (Fig. 1.18).

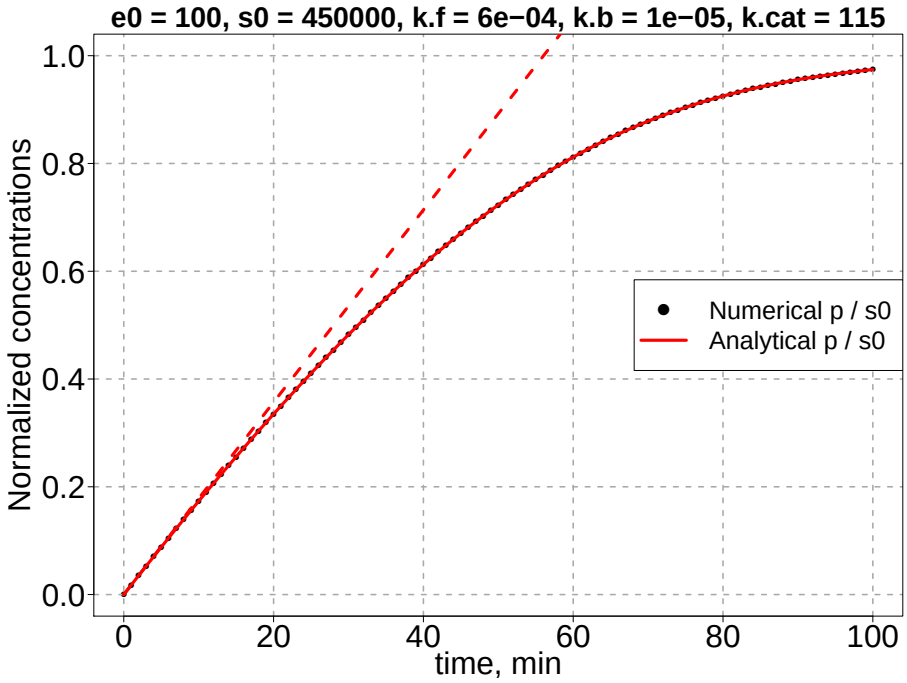


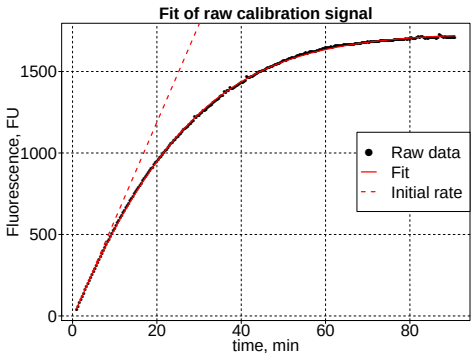
Figure 1.18: Comparison of numerical solution (black) for the product concentration and its analytic solution given by Eq. 1.22 (red) under typical conditions of the calibration experiment (title). Dashed red line visualizes the initial rate of change of fluorescence (Eq. 1.24). Note the excellent correspondence between the solutions.

One more step is required before fitting Eq. 1.22 to the raw calibration signal. It describes the concentration of the fluorescent product, whereas the data is the time course of a fluorescence signal. The link between the two is the instrument transfer function $I(t, p, \dots)$, that can in principle depend on many variables: time t, concentration of fluorescent product p, temperature, type of plasma, position of the sample on the plate, etc. In the simplest case, $I(t, p, \dots)$ is just a constant coefficient $I(t, p, \dots) = I$, so the relation between p and fluorescence signal F is linear (see also Section 1.7). This defines the statistical model for the calibration data:

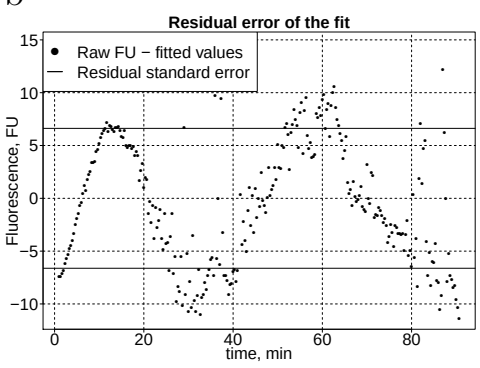
$$F_i = b + I \cdot p(t) + \epsilon_i = b + p_1 \cdot \left(1 - \frac{W\left(p_2 \cdot \exp\left(p_2 - p_3 \cdot t_i\right)\right)}{p2}\right) + \epsilon_i, i = 1 : N \quad , \quad (1.23)$$

where F_i is the dependent variable, t_i is the independent variable, N is the number of observations, b is the parameter representing constant background fluorescence, p_1, p_2, p_3 are parameters defined by Eq. 1.22 and I, and ϵ_i is the unobserved random error. ϵ_i are assumed to be normally distributed with zero mean and the residual variance σ^2. The fit of Eq. 1.23 to the calibration signal is presented in Fig. 1.19a.

a



b



c

Parameter	e_0	s_0	F'_{init}	I	k_{cat}	K_M	CF	t_0
Value	108.33	454	60.21	$3.7 \cdot 10^{-3}$	$3.98 \cdot 10^2$	$7.6 \cdot 10^5$	1.8	0.9
Units	nM	μM	FU / min	FU / nM	1 / min	nM	nM · min / FU	min

Figure 1.19: **(a)** Fit of Eq. 1.23 to the calibration signal (Fig. 1.5a). The initial rate of change of fluorescence (Eq. 1.24) is visualized with the dashed red line. **(b)** Residuals and the residual standard error of the fit. **(c)** The values of the initial rate of change of fluorescence of the calibration signal F'_{init}, the instrument transfer coefficient I, the kinetic constants k_{cat} and K_M, the calibration factor CF, and the starting time of the experiment t_0 calculated using the parameters obtained from the fit from **a** and the known values of e_0 and s_0.

Figure 1.19b shows the residuals of the fit, and the horizontal black lines correspond to the residual standard error σ. It is clearly seen from the this figure that stan-

dard errors on parameters can not be used for making statistical inference, because the residuals are not normally distributed. However, raw data are also correlated for obvious reasons — since fluorescence grows in time, each next point is dependent on the previous one, i.e. the measurements are not independent. Moreover, there is a number of artifacts in the calibration signal that could hardly be taken into account by a model, for example abrupt jumps of fluorescence, non-constant background in the signal, photobleaching of plasma and etc. Note also that the experimental point at $t = 0$ is missing, so the actual time when the reaction started may be different from the $t = 0$ provided by the CAT machine. This is why we added an additional parameter describing the experimental time zero, t_0, so that $t = t_i - t_0$. This gives a way to account for a potential effect of delay of the measurement, when $t = 0$ reported by the fluorometer is not the exact time moment when the reaction started.

At the same time, the red continuous curve in Fig. 1.19a still represents a decent mathematical approximation of the black data points, and the values of parameters as such can be used to calculate the parameters of interest of the signal. Eq. 1.22 has the property that its first derivative at $t = 0$ corresponds exactly to the rate of change of the concentration of product in the regime, when it grows linearly in time. It defines the initial rate of change of fluorescence F'_{init} analytically as

$$F'_{init} = I \cdot \left(\frac{dp}{dt}\right)_{t=0} = I \cdot \frac{k_{cat} \cdot e_0}{1 + \frac{K_M}{s_0}} = \frac{p_1 \cdot p_3}{1 + p_2} \quad , \tag{1.24}$$

via the parameters used in Eq. 1.22 or in Eq. 1.23 respectively. Not surprisingly, $\left(\frac{dp}{dt}\right)_{t=0}$ in Eq. 1.24 represents the well-known Michaelis-Menten equation when the concentration of substrate $s = s_0$. Eq. 1.24 also defines the formula for the calibration factor CF, that reads

$$CF = \frac{1 + \frac{K_M}{s_0}}{I \cdot k_{cat}} \quad . \tag{1.25}$$

Equation 1.25 shows that the samples with identical compounds and therefore the same values of K_M and k_{cat} will have the same value of CF only if they received the same initial concentration of substrate s_0 and the transfer function coefficient I is the same for the wells in the reader that contained them. Normally, the first condition always holds true in the experiments, but the second one is not *a priori* guaranteed.

The values of the initial rate of change of fluorescence of the calibration signal F'_{init}, the instrument transfer coefficient I, the kinetic constants k_{cat} and K_M, and the calibration factor CF, calculated using the parameters obtained from the fit of the calibration signal with Eq. 1.23 and the known values of e_0 and s_0, are presented in Fig. 1.19c. We investigated how the fitting that we do for calibration data could potentially affect the values of extracted parameters in details in Appendix A.8 and A.9 (see also Appendix **??**).

The value 1.8 for CF is slightly larger than the estimated CAT value for it in Section 1.4, and the difference is caused by the difference in values of the initial rate of the calibration signal. Eq. 1.24 defines the initial rate of change of fluorescence

of the calibration signal analytically, and allows to calculate it straight from the values of parameters of the fit of Eq. 1.23 to the data. Since the result in Fig. 1.19c agrees with Fig. 1.13a, it must be concluded that CAT overestimated the value of TC_Initial_Slope, and that resulted in a lower value for CF, as Eq. 1.8 suggests.

1.5.3 Titration of Thrombin Calibrator gives non-constant calibration factor

Researchers in the field of thrombin generation, who set up the calibration experiments in their own labs, quite often use the strategy outlined in [Hem+03] to get the calibration curves for the whole range of the experiment. That is, to make titrations of the Thrombin Calibrator (i.e. titrate the concentration of e_0). It normally gives series of calibration curves like the ones in Fig. 1.20.

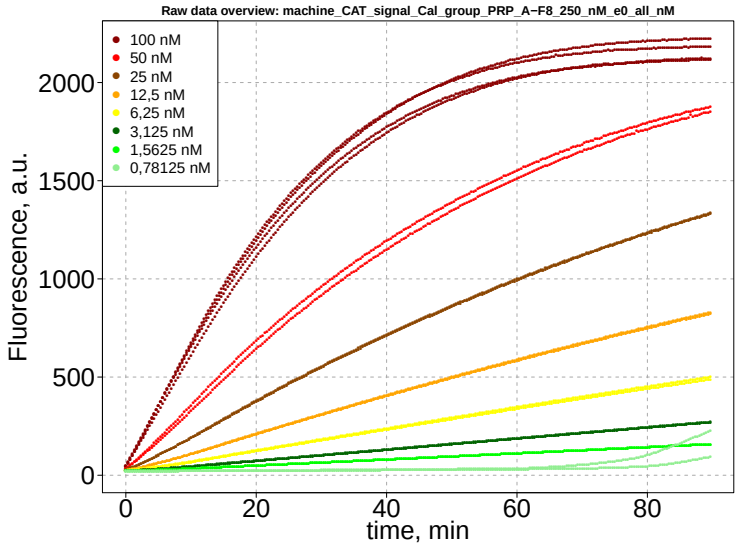


Figure 1.20: Calibration experiment with titration of Thrombin Calibrator.

In [Hem+03], the curves are "straightened" using the transformation of the raw data. Using the language of Michaelis-Menten kinetics, we can say that titration of the Thrombin Calibrator slows down the kinetics so we gradually move from late-times kinetics to the early-times kinetics. Nevertheless, we can test the applicability of Eq. 1.22 by fitting it directly to titrated series of signals.

The fits are presented in Fig. 1.21 and are convincing. The only detail that seems suspicious is the significant deviation from the experimental signal at early times ($t \leq 5$ min) of the experiment. One could think that this is an indication of the both early-times and late-times kinetics present in a single signal. However, there is no

solution derived yet for the product concentration that describes both regimes, since Eq. 1.22 and Eq. 1.20 decribe only the late- or the early-times kinetics. Nevertheless, it is possible to combine the two solutions together, so that the combined solution has both the properties of the late- and the early-times kinetics.

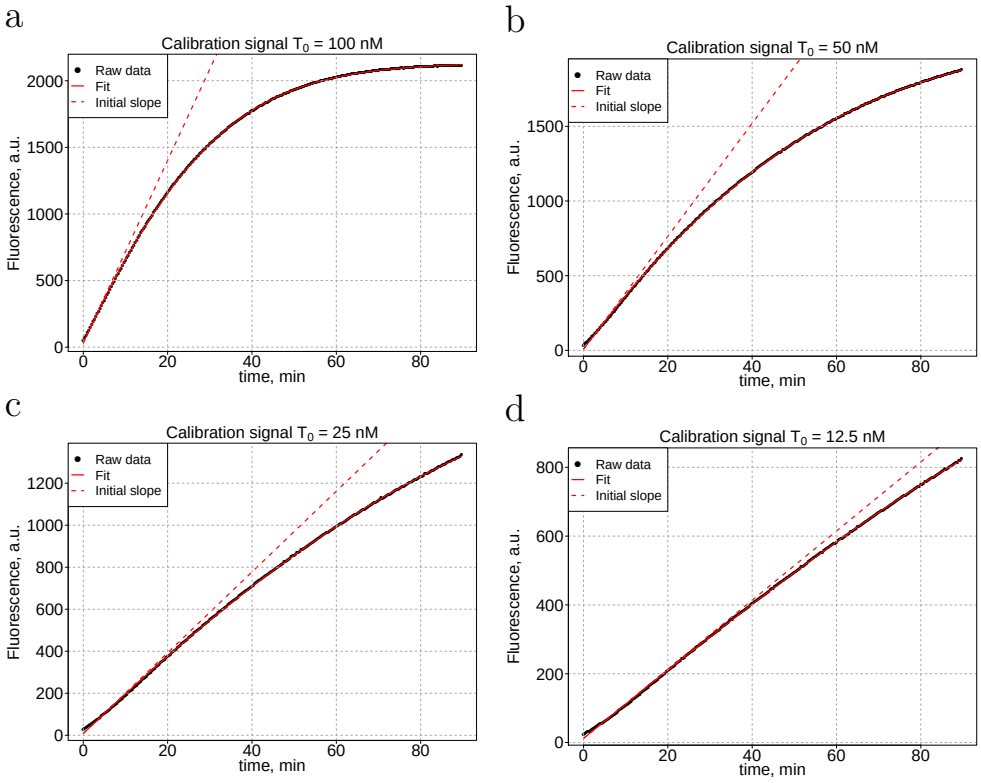


Figure 1.21: Fits of Eq. 1.22 to several selected signals from Fig. 1.20. Note the differences between the data and the fits at early times ($t \leq 5$ min) of the experiment.

There is a $1 - e^{-t}$ term in Eq. 1.20 that accounts for a slow rise of the product concentration at early times. When this term no longer plays a role, the product concentration starts growing as a first-degree polynomial. A similar first-degree polynomial is present in the Lambert W function in Eq. 1.22. To assure the slow rise of the product concentration in the beginning of the experiment, we can transform this first-degree polynomial using the $1 - e^{-t}$ term from Eq. 1.20, so that the combined solution reads

$$p(t) = s_0 \cdot \left(1 - W\left(\frac{s_0}{K_M} \cdot e^{\frac{s_0}{K_M}} \cdot e^{-\frac{k_{cat} \cdot e_0}{K_M} \cdot \left(t - \frac{1 - e^{(s_0 + K_M) \cdot k_f \cdot t}}{k_f \cdot (s_0 + K_M)}\right)}\right) \cdot \frac{K_M}{s_0}\right) \quad . \tag{1.26}$$

This solution is compared to the numerical solution for the product concentration in Fig. 1.22. The combined solution captures both the initial slow rise of the product concentration and the saturation phase. It has the necessary properties to explain the abovementioned differences between the model and the data (Fig. 1.21). However, that does not seem to be the case. The fits of Eq. 1.26 to calibration signals with T_0 = 100 nM and T_0 = 50 nM are presented in Fig. 1.23.

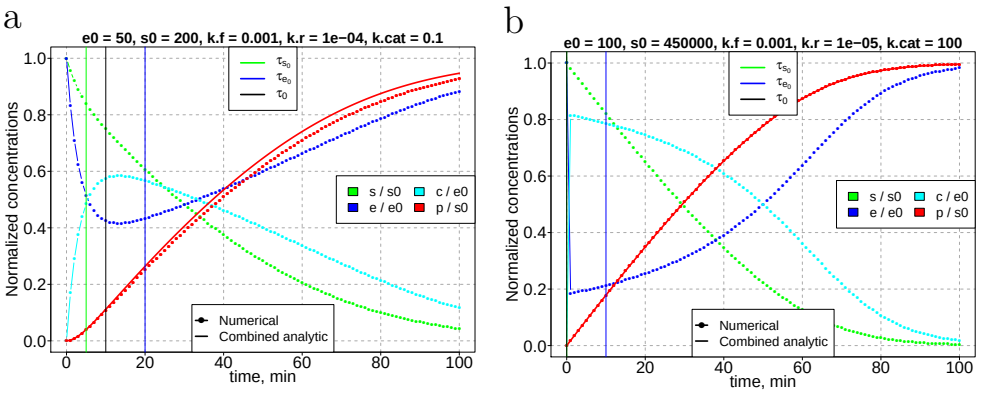


Figure 1.22: Numerical solutions compared to the combined solution for the product concentration given by Eq. 1.26. Note that the solution captures both **(a)** the early-times kinetics and **(b)** the late-times kinetics.

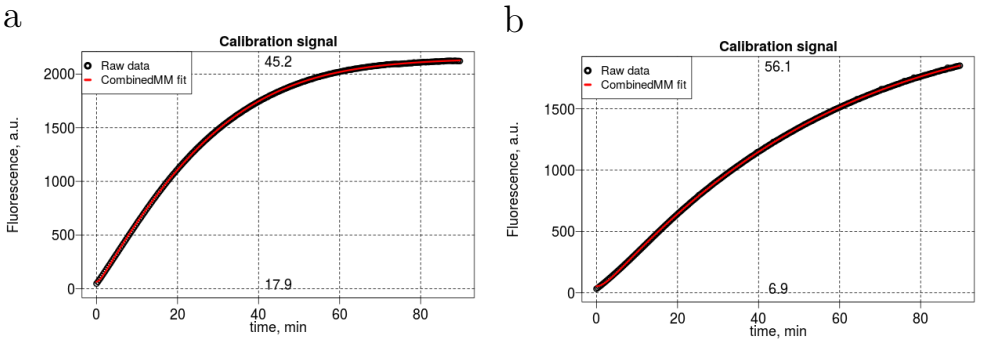


Figure 1.23: Fits of Eq. 1.26 to the calibration signals with **(a)** T_0 = 100 nM and **(b)** T_0 = 50 nM of Thrombin Calibrator.

The residual standard errors of the fits are increased because of the introduction of the additional parameter to the model that is supposed to explain a relatively low amount of variation. The values of this parameter are not consistent with the typical interrelation between k_{cat} and k_f, that is $k_{cat} \gg k_f$. Instead, it changes to $k_{cat} \simeq k_f$,

which is unrealistic. Also, the signal does not seem to grow initially as it is supposed to by the model, and it is simply unlikely in enzymology to be able to resolve the early-times of the kinetics experimentally. A much more plausible explanation is simply the temperature effect — the plasma in the sample has to warm up to 37^o C before the measurement is started, and poor temperature control is known to cause effects like this. Unfortunately, this issue leads to a conclusion that initial rate measurements might be affected by side effects a lot, and produce unreliable results when extracting the calibration factor according to Eq. 1.24.

This conclusion is confirmed here — the values of the calibration factor extracted from the titrated series shows a dependence on the initial concentration of the Thrombin calibrator (Fig. 1.24), when according to Eq. 1.8 it simply has to be a constant.

We show in Section 1.6 that this titration experiment strategy is suboptimal and there is no need to extract the initial rate from the calibration signals. Instead, the kinetic constants k_{cat} and K_M are of primary interest, and their extracted values are not affected much by the temperature effects of warming up. So, there is no need to do titration — a single prolongued measurement with a large concentration of α_2M-T gives reliable values for the kinetic constants k_{cat} and K_M, and they are all that is required for calibration.

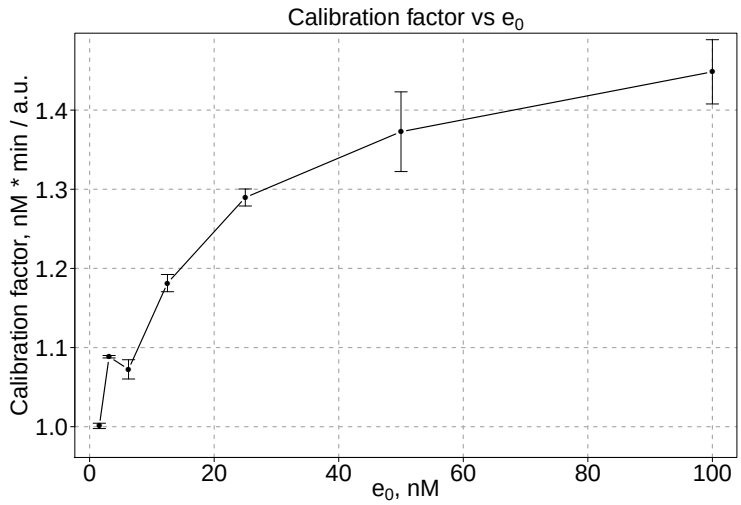


Figure 1.24: In the titration experiment, the extracted calibration factor is not supposed to be dependent on the initial concentration of Thrombin Calibrator in the sample (Eq. 1.8). However, that is clearly not the case. Section 1.6 shows that actually only the kinetic constants k_{cat} and K_M are required from the calibration signal, so there is no need to do titration and extract the initial rate.

1.5.4 Non-constant background from substrate, buffer or plasma

There are several smaller details, that are relevant to the discussion of fit of Eq. 1.22 to the experimental data from the calibration experiments. Strictly speaking, the two facts that the raw data are correlated and the non-normal distribution of the residuals (see for example Fig. 1.19b) would make a statistician reject the hypothesis that Eq. 1.22 is the proper model for the data. This is formally correct, unless we can justify the sources of errors we simply cannot account for in a simple model like Eq. 1.22.

To begin with, we rely on the assumption that the background fluorescence is constant in time. However, the experimental check of this assumption shows that it is only true for an empty plate in the fluorometer, i.e. — for a "null" sample.

Figure 1.25a shows various background measurements using an EnVision machine [Per]. The only contribution that can be considered constant in time is the contribution of an empty well. Both buffer and substrate contributions show abrupt jumps, substrate contribution increases as the experiment progresses, whereas the plasma contribution decreases (plasma aging). On the y-axis of the plot, we see the absolute value of these contributions. Figure 1.25b shows the fit of Eq. 1.22 to the calibration signal measured in buffer for 300 minutes, which is far beyond the typical range of the experiments reported in literature (around 100 min). The residual standard error of the fit is $\sigma = 9650$ a.u., and is clearly within the range of non-constant background contributions visualized in Fig. 1.25a.

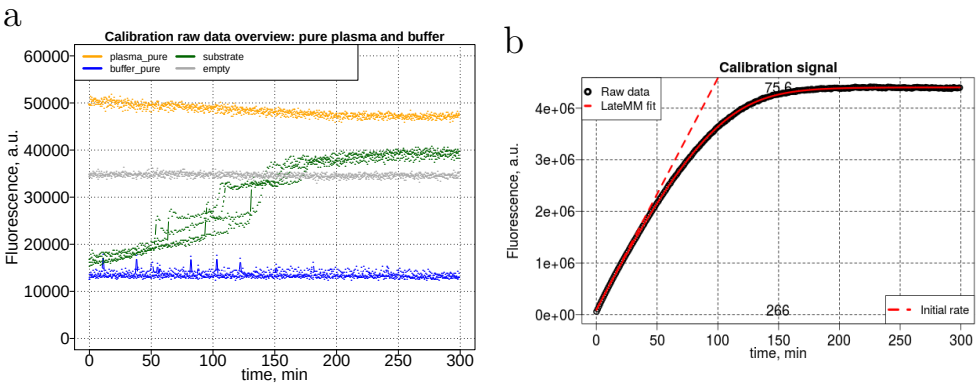


Figure 1.25: **(a)** Measurement of background from the empty wells (grey), wells containing only substrate (green), only buffer (blue), or only plasma (orange). Clearly, only the contribution of empty wells to the total background in the instrument can be considered constant in time. **(b)** Fit of Eq. 1.22 to the calibration signal measured in buffer. The residual standard error of the fit is $\sigma = 9650$ a.u., and is clearly within the range of non-constant background contributions visualized in **a**.

Calibration experiments in buffer are much "cleaner" than the ones in plasma, because there are no effects of aging and temperature on buffer. Figure 1.26 visualizes

the differences between them.

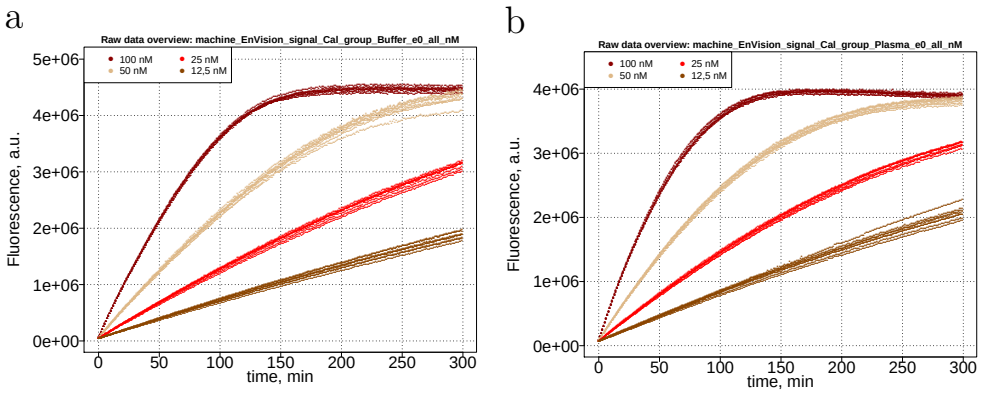


Figure 1.26: Calibration experiment ran for 300 min in a) buffer and b) plasma. The decrease of fluorescence in plasma due to temperature and aging effects is pronounced in 100 nM curves in b).

However, buffer measurements are useless for calibrating thrombin generation signals, since the latter always use plasma. Still, they are more than fine to test the applicability of Eq. 1.22 to the data, and the main conclusion is that it is indeed applicable. However, one should keep the discussion in this section in the head when evaluating the quality of fits to the raw calibration data.

1.5.5 Results and discussion

In this section, we considered the calibration data in thrombin generation experiments (see also Appendix **??**). We found that experimental data is described well with textbook Michaelis-Menten kinetics of the basic enzymatic reaction, taking into account the effects of non-constant background fluorescence from the substrate and the aging of plasma. The duration of measurement and the concentration of enzyme α_2M-T in the compound of the calibration signal define whether it is best described by the early-times or late-times approximate analytic solution of Michaelis-Menten kinetics. In any of the cases, the use of the approximate analytic solutions instead of the numerical solutions is justified because the initial concentration of enzyme is 1000 times lower than the concentration of substrate. Calibration signal contains the information about two kinetic rates — k_{cat} and K_M, that can be extracted from the fit of late-times approximate analytic solution, whereas initial rate measurements are not able to provide it. Finally, there is simply no need for the previously reported strategy of titration for the calibration experiment — a single complete time course of calibration is all that is required to calibrate any thrombin generation signal measured in the same plasma.

1.6 Phenomenological model for thrombin generation data

A large subfield in thrombin generation is the investigation of possible models explaining the blood coagulation cascade [KS89; Wil+91; JM94]. Generally in such papers, thombograms are fitted with a numerical solution of a system of differential equations defined by the model that describes the pathway to thrombin generation [WHH06; PGA14]. Several reviews [WHH06; HKK12] showed that this technique is useful, but has certain limitations, and its results should be treated with caution. Moreover, the model check in the case of several tens of parameters requires a scrupulous statistical analysis to distinguish between significant and insignificant ones, and was to our best knowledge rarely performed or discussed properly in the literature.

However, the most important practical issue for many researchers in the field is how to determine the main parameters of interest characterizing the generation of thrombin during the experiment in the best possible way. For that purpose, it is instructive to consider only the main aspects of thrombin generation mechanism and focus on the reactions primarily monitored in the experiment. The two main aspects of the problem are presented in Fig. 1.27 — the chemical reactions decsribing the formation of the fluorescent product during the time course of the experiment in Fig. 1.27a, and the mechanism of thrombin generation as a cascade of enzymatic reactions [HHL65; HH69] with damping by anti-enzymes in Fig 1.27b.

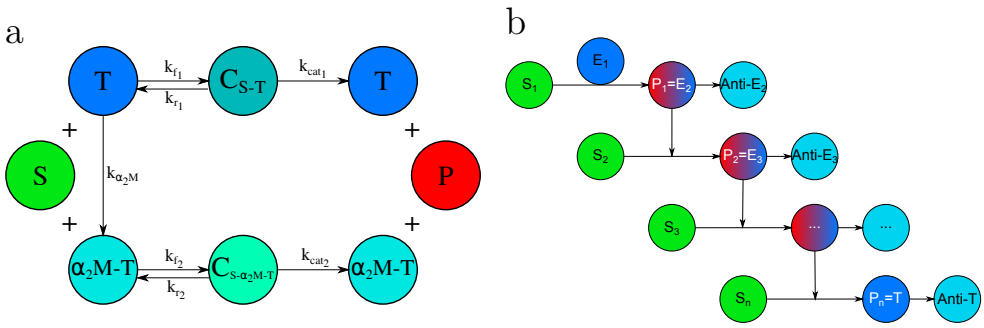


Figure 1.27: The two main aspects of thrombin generation: **(a)** The chemical reactions decsribing the formation of the fluorescent product during the time course of the experiment, with the separation of the contributions of thrombin T and α_2M-T. Both enzymes act on the same substrate S and produce the same product P via different intermediate enzyme-substrate complexes. **(b)** The mechanism of thrombin generation as a cascade of basic enzymatic reactions with damping by anti-enzymes. Enzyme E_1 acts on substrate S_1 to form product P_1. Then, P_1 serves as enzyme E_2 in the next step of the reaction, and so on until the last step. This is analogous to how blood coagulation factors become activated in the blood coagulation cascade and start the next reaction (Fig. 1.1). Damping by anti-enzymes ensures that the cascade does not go on forever producing inifite amount of final product P_n. In this mechanism, thrombin concentration is the concentration of P_n.

The two main reactions that are directly monitored in the experiment and are shown in Fig. 1.27a are the formations of the fluorescent product by thrombin and by α_2M-T. The law of mass action defines the differential equation describing the rate of change of the concentration of the fluorescent product p, that reads

$$\frac{dp}{dt} = k_{cat_1} \cdot C_{S\text{-}T} + k_{cat_2} \cdot C_{S\text{-}\alpha_2 M\text{-}T} \ . \tag{1.27}$$

As Eq. 1.27 shows, the product is formed by two stable complexes, one being the substrate-thrombin complex $C_{S\text{-}T}$, and the other being the substrate-α_2M-T complex $C_{S\text{-}\alpha_2 M\text{-}T}$. The differential equations decribing the rates of change of complexes are

$$\begin{aligned} \frac{dC_{S\text{-}T}}{dt} &= -(k_{b_1} + k_{cat_1}) \cdot C_{S\text{-}T} + k_{f_1} \cdot S \cdot T \ , \\ \frac{dC_{S\text{-}\alpha_2 M\text{-}T}}{dt} &= -(k_{b_2} + k_{cat_2}) \cdot C_{S\text{-}\alpha_2 M\text{-}T} + k_{f_2} \cdot S \cdot \alpha_2 M\text{-}T \ . \end{aligned} \tag{1.28}$$

Eqs. 1.27 and 1.28 fully describe the reaction kinetics according to the law of mass action. However, simplifying assumptions are required to use this system for data analysis. In the pseudo steady-state regime, the rate of change of concentrations of both complexes can be considered negligible, and set equal to zero. Two differential equations are then transformed into two algebraic equations

$$C_{S\text{-}T} = \frac{S \cdot T}{K_{m_1}}, \qquad C_{S\text{-}\alpha_2 M\text{-}T} = \frac{S \cdot \alpha_2 M\text{-}T}{K_{m_2}} \ , \tag{1.29}$$

and the differential equation for the product concentration is reintroduced as

$$\frac{dp}{S} = \frac{k_{cat_1}}{K_{m_1}} \cdot T \cdot dt + \frac{k_{cat_2}}{K_{m_2}} \cdot \alpha_2 M\text{-}T \cdot dt \ . \tag{1.30}$$

At this stage, a solution can be derived in the case when subtrate concentration is constant in time, i.e. when the effect of substrate consumption in the experiment is negligible. So, $S(t) = s_0$, and a general solution for p reads

$$p(t) = \frac{s_0 \cdot k_{cat_1}}{K_{m_1}} \cdot \int_0^t T(t') \cdot dt' + \frac{s_0 \cdot k_{cat_2}}{K_{m_2}} \cdot \int_0^t \alpha_2 M\text{-}T(t') \cdot dt' \ . \tag{1.31}$$

Eq. 1.9 shows that the contribution of α_2M-T can be calculated if thrombin concentration as a function of time is known. As discussed above, it is hardly possible to establish the true and generally applicable model for thrombin generation based on a certain specific reaction mechanism. A different approach to the problem would be to use a phenomenological model, that is specifically designed to capture the features of data and is based on the actual kinetics of enzyme cascade.

The simplest way to formulate a model of thrombin generation is to consider a cascade of basic enzymatic reactions leading to the formation of thrombin, shown in

Fig. 1.27b. The details of individual steps of the cascade are not considered, because the focus is the overall behaviour of the concentration of thrombin as a function of time. The assymetric shape of the thrombogram in Fig. 1.13b reminds a composite of two functions: initial rise as a polynomial t^l, and an exponential decrease after a peak. This is also the qualitative behaviour of the model of enzyme cascade, presented in Fig. 1.27b. The initial rise of the concentration E_n behaves like t^{n-1} [HH69], where n is the number of steps in the cascade. Damping by anti-enzymes via a first-order reaction describes the late-time part of the thrombogram, where it approaches zero exponentially.

A phenomenological model capturing both of the early-time and the late-time properties of the thrombogram could be expressed via a standard formula of statistics — the probability density function of gamma distribution. This function has certain normalization by default, and it is handy to use in the following form

$$T(t) = \begin{cases} \dfrac{A \cdot (t-t_0)^{k-1} \cdot e^{-(t-t_0)/\theta}}{\Gamma(k) \cdot \theta^k}, & \text{if } t > t_0 \\ 0, & \text{if } t \le t_0 \end{cases} , \tag{1.32}$$

where $T(t)$ is the concentration of thrombin, t is the time of the experiment, t_0 is the lag time, A is the total area under $T(t)$, k is the shape parameter indicating the number of reactions involved, $\Gamma(k)$ is the gamma function, and θ is the scale parameter, or the decay constant of the exponential term.

The concentration of α_2M-T can be calculated analytically from Eq. 1.32 using Eq. 1.9, and is expressed as

$$\alpha_2 M\text{-}T(t) = \begin{cases} \dfrac{k_{\alpha_2} \cdot A \cdot \gamma(k, (t-t_0)/\theta)}{\Gamma(k)}, & \text{if } t > t_0 \\ 0, & \text{if } t \le t_0 \end{cases} , \tag{1.33}$$

where k_{α_2} is the rate of formation of α_2M-T from thrombin, and $\gamma(k, (t-t_0)/\theta)$ is the lower incomplete gamma function.

Eq. 1.31 contains two separate contributions of the integrals of Eq. 1.32 and Eq. 1.33, each multiplied by a ratio of reaction rates k_{cat_i}/K_{m_i}, $i = 1{,}2$. The integrals can be calculated analytically

$$T_\Sigma(t) = \frac{A \cdot \gamma(k, (t-t_0)/\theta)}{\Gamma(k)} , \tag{1.34}$$

$$\alpha_2 M\text{-}T_\Sigma(t) = \frac{k_{\alpha_2} \cdot A \cdot (\gamma(k, (t-t_0)/\theta) \cdot (t-t_0) - \theta \cdot \gamma(k+1, (t-t_0)/\theta))}{\Gamma(k)} . \tag{1.35}$$

It was reported in [Hem+03] that $K_{m_1} \simeq K_{m_2}$ and $k_{cat_1} \simeq 2 \cdot k_{cat_2}$. The contribution of α_2M-T can therefore be expressed in thrombin equivalents (Appendix A.4),

so k_{cat_2} and K_{m_2} can be replaced with k_{cat_1} and K_{m_1}, which with the dropped subscripts become simply k_{cat} and K_m. An additional fraction $\frac{1}{2}$ can be included into the paramteter k_{α_2}, so that Eq. 1.31 transforms into

$$p(t) = \frac{s_0 \cdot k_{cat}}{K_m} \cdot (T_\Sigma(t) + \alpha_2 M\text{-}T_\Sigma(t)) \quad , \tag{1.36}$$

where $T_\Sigma(t)$ is given by Eq. 1.34 and $\alpha_2 M\text{-}T_\Sigma(t)$ — by Eq. 1.35.

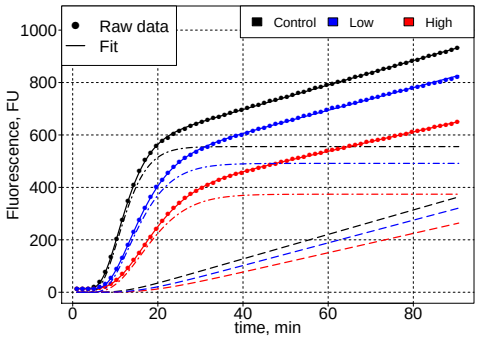
Eq. 1.36 can be easily translated into a statistical model in the same way as Eq. 1.22. The results of fitting this model to the data are depicted in Fig. 1.28a. Only one out of four points in the raw data is plotted for better visualization. The solid lines represent the fit, the two-dashed lines represent the contribution of the thrombin integral to the fit, and the long-dashed lines represent the contribution of α_2M-T integral to the fit. Colors indicate the type of signals as in Fig. 1.5b. The residuals and the residual standard errors of the fits are presented in Fig. 1.28b. Again, as in the case with calibration signal in Fig. 1.19b, residuals are not normally distributed, so the standard errors on parameters cannot be used for statistical inference. The values of parameters as such still define a decent mathematical approximation of the data via Eq. 1.36, so they can be used to obtain the values of CAT parameters.

The significant improvement in data analysis when fitting the raw data is that all noise is taken into account by the statistical model, and that the contributions of thrombin and α_2M-T are automatically separated straight in the model formula. Moreover, the fit itself is a continuous curve that is not limited by the sampling time of the fluorescence reader. It means that all relevant properties of such a continuous curve (area under curve i.e. ETP, a global maximum point i.e. Peak, ttPeakand etc.) can be calculated analytically rather than estimated numerically, in other words — with absolute accuracy. As Eq. 1.30 and Eq. 1.36 suggest in case $S = s_0$, taking the first derivative of the product concentration gives the sum of thrombin and α_2M-T concentrations multiplied by a fraction $\frac{s_0 \cdot k_{cat}}{K_m}$. Therefore, the inverse of $\frac{s_0 \cdot k_{cat}}{K_m}$ multiplied by the transfer function coefficient I gives a formula for the calibration factor

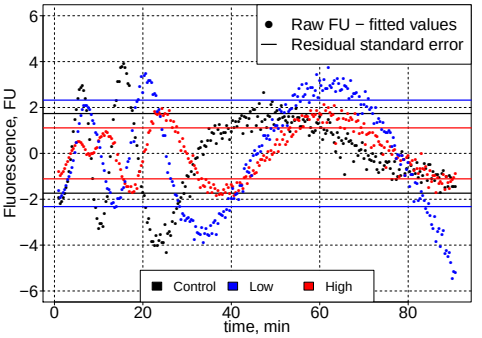
$$CF_2 = \frac{K_m}{I \cdot s_0 \cdot k_{cat}} \quad , \tag{1.37}$$

that is $\simeq 1.11$ according to the values in Fig. 1.19c. Note the 1.6 fold difference in calibration factor CF used by CAT and CF_2 derived above. This issue has been reported before in several articles [Hem+03; HK13], but the clear explanation and the detailed discussion of the problem was not presented. A calibrated thrombogram is then produced using Eq. 1.32, with the values of parameters obtained from the fit, and the calculated value of the calibration factor CF_2 using Eq. 1.37. Fig. 1.28c shows the calibrated first derivatives calculated from raw data with Eq. 1.4 as points, the same first derivatives calculated analytically with Eq. 1.36 using the values of parameters obtained from fits as solid lines, the contributions of thrombin calculated using Eq. 1.32 as two-dashed lines, and the contributions of α_2M-T calculated using Eq. 1.33 as long-dashed lines.

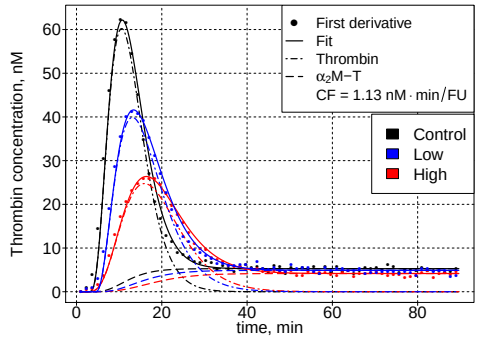
a



b



c



d

Parameter	Control	Low	High
Lagtime, min	2.96	4.52	2.13
Lagtime SD	NA	NA	NA
ETP, nM · min	628.17	555.75	422.89
ETP SD	NA	NA	NA
Peak, nM	60.27	39.91	24.77
Peak SD,	NA	NA	NA
ttPeak, min	10.58	13.11	15.93
ttPeak SD	NA	NA	NA
VelIndex, nM / min	14.21	7.96	3.36
VelIndex SD	NA	NA	NA
Alpha2M_Level, nM	5.27	4.89	4.17
Alpha2M_Level SD	NA	NA	NA

Figure 1.28: **(a)** Thrombin generation signals fitted with Eq. 1.36. Only one out of four points in the raw data is plotted for better visualization. Solid lines represent the fits, two-dashed lines represent the contributions of thrombin integrals to the fits, and long-dashed lines represent the contributions of α_2M-T integrals to the fits. **(b)** Residuals and residual standard errors of the fits. Again, residuals are not normally distributed so the standard errors on parameters obtained from fits cannot be used for statistical inference. **(c)** Calibrated first derivatives calculated from raw data with Eq. 1.4 are shown as points, the same first derivatives calculated analytically with Eq. 1.36 using the values of parameters obtained from fits — as solid lines, the contributions of thrombin using Eq. 1.32 — as two-dashed lines, and the contributions of α_2M-T using Eq. 1.9 as long-dashed lines. The value of calibration factor used to translate from FU of fluorescence to nM of thrombin is 1.11 nM · min / FU. **(d)** The calibrated main parameters of interest of the signals, calculated analytically using the values of parameters obtained from fits to raw data. Colors indicate the type of signals as in Fig. 1.5b. NA indicates that the value is not available.

The calculation of the calibrated main parameters of interest of the signals, presented in Fig. 1.28d, is a straightforward task. Lagtime is included in the model described by Eq. 1.36 as the parameter t_0, so it can be taken directly from the parameter values. ETP is the value of parameter A multiplied by the calibration factor CF_2. ttPeak is calculated as $(k-1)\cdot\theta$, and Peak is the value given by Eq. 1.32 at $t =$ ttPeak. VelIndex is value of the first derivative of Eq. 1.32 at $t = \theta\cdot(k-1-\sqrt{k-1})$, that corresponds to its global maximum point. The values of the parameters obtained from fitting the raw data in Fig. 1.28d along with the values obtained by CAT machine in Fig. 1.5c are visualized in Fig. 1.29 for clear comparison.

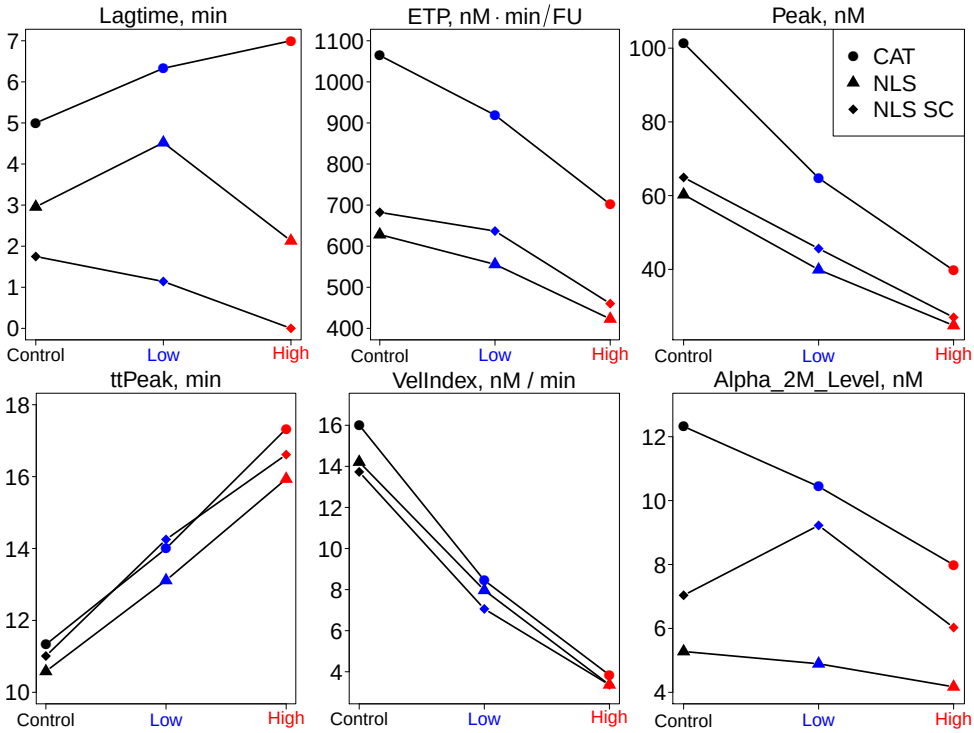


Figure 1.29: Visualization of the obtained main parameters of interest characterizing the performance of thrombin in the experiment. Circles correspond to the values provided by CAT machine (CAT), triangles — to the values obtained from nonlinear least-squares fitting of raw data with Eq. 1.36 (NLS), and diamonds — with Eq. 1.39, that takes into account the effect of substrate consumption (NLS SC). Colors indicate the type of signals as in Fig. 1.5b.

Points and triangles follow the same trend qualitatively described in Section 1.3, but the values of time-independent CAT parameters are up to 1.8 times larger mainly because of the difference in the values of CF and CF_2. Note the lagtime value for

the red signal — it seems to deviate from the growing trend for the lagtime shown by CAT. However, CAT likely estimates the lagtime parameter using the rule of three standard deviations from the baseline [HK13]. In Eq. 1.32, lagtime is a free parameter — the nonlinear least-squares fitting algorithm just finds the value that corresponds to the lowest residual standard error. The red signal rises the most slowly among all others, the value of its shape parameter k is the largest, and therefore it takes small but non-zero values for a prolongued period of time compared to the others. It explains the lower value of t_0 for the red signal, that Fig. 1.29 shows.

1.6.1 Results and discussion

In this section, we considered a simple phenomenological model for thrombin generation signal based on gamma distribution. It is inspired by the qualitative behaviour of a cascade of basic enzymatic reactions, which is close enough to the contemporary understanding of how thrombin generation happens. This model approximates the data fairly good, given its simplicity, and is fast to implement and use. We also found that kinetic equations for thrombin generation data yield a definition of the calibration factor that is a factor 1.6 different from the one used by CAT machine. This is why CAT greatly overestimates the parameters of thrombogram that do not depend on time. We found the values of the rest of the parameters to agree well with the values provided by CAT. In this way, we reported a complete benchmark for our method and proved it to work.

1.7 Effects of substrate consumption and fluorescence quenching

The effects of substrate consumption and fluorescence quenching effect are well known to be present in fluorogenic measurements and have been discussed before [HHD09; Woo+13]. As reported in Section 1.5, the assumption of the fixed concentration of enzyme used in the definition of the calibration factor (Eq. 1.8) and in the H-transform [HHD09] is quite questionable from the point of view of reaction kinetics. Moreover, this transformation claims to compensate for both effects without clearly showing their separate contributions to the experimental data.

1.7.1 Substrate consumption

It seems questionable to state [HHD09] that substrate consumption in a thrombin generation signal can be compensated with a calibration signal for a very simple reason — these two experiments represent two different systems with two different reaction kinetics. So, the effect of substrate consumption should be incorporated directly into the model for thrombin generation signal rather than imported from the model for calibration signal. In both calibration and thrombin generation experiments the concentration of substrate $s(t)$ is decreasing from initial value s_0 to zero, if the duration of the experiment is enough to capture the full time course of the reaction kinetics.

As discussed above, all steps leading to thrombin generation from the blood coagulation cascade result in a large system of differential equations, and there is no analytical solution in the general case even for a single step of the kinetics, that is the basic enzymatic reaction (Section 1.5). However, only a few details are required to obtain an extended model that accounts for the effect of substrate consumption. The law of conservation of concentration links together the concentration of fluorescent product $p(t)$, that is observed in the experiment, with the concentration of substrate $s(t)$. It reads $s_0 = s(t) + p(t) + c_1(t) + c_2(t) + T(t) + \alpha_2 M\text{-}T + \ldots$, where $\ldots$ indicate all the relevant reagents from other steps in the thrombin generation cascade in Fig. 1.27b. The peak concentrations of thrombin and α_2M-T are on the order of hundreds of nanomolars, and the same applies for the concentrations of their stable complexes. At the same time, the concentrations of substrate and product are in the millimolar range, and that is a thousand fold difference. If all the low-order summands of the conservation law are neglected, it turns into a particularly simple form $s_0 = s(t) + p(t)$. This equation allows to reintroduce Eq. 1.30 as

$$\frac{dp}{s_0 - p} = \frac{k_{cat_1}}{K_{m_1}} \cdot T \cdot dt + \frac{k_{cat_2}}{K_{m_2}} \cdot \alpha_2 M\text{-}T \cdot dt \ , \tag{1.38}$$

that has an analytical solution in the form

$$p(t) = s_0 \cdot \left(1 - e^{-\left(\frac{k_{cat_1}}{K_{m_1}} \cdot \int_0^t T(t') \cdot dt' + \frac{k_{cat_2}}{K_{m_2}} \cdot \int_0^t \alpha_2 M\text{-}T(t') \cdot dt'\right)}\right) \ . \tag{1.39}$$

The effect of substrate consumption is taken into account by the $(1 - \exp(\dots))$ term in Eq. 1.39 (see also Appendix **??**). In case the sum of the contributions in the exponent is close to zero, the first order Taylor approximation of the term will give exactly Eq. 1.31. In the same fashion as Eq. 1.31 and Eq. 1.36, Eq. 1.39 defines a statistical model for the data (see Section A.7 for fits), and the main parameters of interest obtained with this model are presented in Fig. 1.29 as diamonds. Again, the diamonds follow the triangles closely, but have slightly (5 - 10%) increased values of time-independent parameters. Note that the trend for the lagtime experiences a decrease in this case rather than a rise. The reason is the same as discussed previously — with the additional term describing the substrate consumption, the values of shape parameters are increased, and therefore the lagtimes decrease. The difference of the thrombin generation signal from linearity after the thrombin generation is over defines the required "degree of compensation" in Eq. 1.39. In the considered signals, this effect is clearly not that sufficient because even the models that do not account for substrate consumption describe the data very well. A prolongued time of the experiment would make the deviation from linearity a lot clearer visually and would highlight the limits of applicability of the simpler model in Eq. 1.36 much better. A side effect of applying a model that is too complicated for the data is best illustrated by the green diamond for `Alpha_2M_Level`. It is overcompensated, because the nonlinear least-squares fitting algorithm increased both the value of the "degree of compensation" for the substrate consumption effect and the value of $k_{\alpha_2 M}$. The particular combination of the two parameters results in a nice fit of the raw data, but the standalone comparison reveals the inconsistency of the green diamond in Fig. 1.29 with the values from two other signals. This is why a prolongued thrombin generation signal, that cannot be fitted with a simple model in Eq. 1.36, would contain valuable information about the substrate consumption effect, that could be taken into account by the extended model in Eq. 1.39.

1.7.2 Fluorescence quenching

The effect of fluorescence quenching is all about establishing the transfer function of the instrument $I(t, p, \dots)$. It can be investigated in a separate experiment, and its results are only applicable for the specific fluorometer and the specific plasma used. The CAT machine was reported before to have a negligible effect of fluorescence quenching in a large experimental window [Woo+13]. The data below is only for EnVision fluorometer, but the experimental protocol holds true for all other machines as well.

Consider a sample that consists only of a certain plasma and a certain concentration of the fluorescent product AMC. This is the so-called null experiment, because no reactions are supposed to be happening, and the level of fluorescence should ideally remain constant in time with some noise on top of it. In the real experiment, the concentration of product cannot exceed the initial concentration of substrate s_0, so the value of s_0 is the upper limit for the concentration of AMC to be added. The titration series of AMC from 0 μM to s_0 μM will therefore capture the experimental

window of interest. Fig. 1.30a shows two repetitions of the time courses with 325 μM concentration of AMC in the same plasma. The time courses were fitted with a phenomenological model that captures both the saturation and the decay phases in the data

$$F = b + a \cdot \left(1 - e^{-c \cdot t}\right) \cdot e^{-d \cdot t} \ , \tag{1.40}$$

where F is the level of fluorescence, b is the constant background fluorescence, a is the amplitude of the signal, c is the rate of saturation, and d is the rate of decay. Note that the first point is omitted, because it was generally an unreliable measurement in the experiment. The models are in satisfactory agreement with the data, and such a procedure can be repeated for all other time courses with different concentrations of AMC. To represent the data compactly, from each of the time series their group minimum and group maximum were picked up. Group minimum (`MIN` in Fig. 1.30a) is the minimum of the two repetitions excluding the first point in each of them, and it is obtained directly from the data. Group maximum (`MAX` in Fig. 1.30a) is the maximum of the two repetitions obtained from the fitted curve. In cases where Eq. 1.40 could not describe the signal adequately, `MAX` was also estimated directly from the data. This procedure gives a vector of `MIN`s and a vector of `MAX`s, that can be plotted against the concentration of AMC as means with error bars. The result is presented in Fig. 1.30b. The green line represents the weighted fit of a linear model to the subset of data $[AMC] \leq 350\mu$M, where weights are determined as inverse squared errors on the points. The continuation of the fit to higher concentrations is indicated by the dashed green line. The figure shows that the level of fluorescence is no longer in linear proportion with the concentration of AMC approximately after 300 μM.

The effect of fluorescence quenching can be taken into account by a simple model. Consider a linear proportion $\Delta F = a \cdot \Delta[AMC]$ between the elementary change in fluorescence ΔF and the elementary change in AMC concentration $\Delta[AMC]$. The boundary condition required by the data is $F([AMC] = 0) = b$, where b is the background fluorescence. This proportion defines the linear model that was investigated above. Now, the next step is to allow ΔF depend also on the current level of fluorescence F, so that $\Delta F = a \cdot \Delta[AMC] - b \cdot F$. This nonlinear proportion defines F as

$$F = b + a \cdot (1 - e^{-c \cdot [AMC]}) \ , \tag{1.41}$$

where b is the constant background fluorescence, a is the maximal level of fluorescence, c is the saturation rate, and $[AMC]$ is the concentration of AMC. The fit of Eq. 1.41 to the data is represented by the red line in Fig. 1.30b. This model describes the effect of fluorescence quenching effectively, it does not account for the individual time dependencies at the chosen AMC concentration (see Fig. 1.30a). Instead, the time dependence is represented via error bars on the mean level of fluorescence at the chosen concentration of AMC.

Eq. 1.41 provides a way to account for the fluorescence quenching effect by incorporating it into the model for the signal. At the same time, Fig. 1.30b also shows that

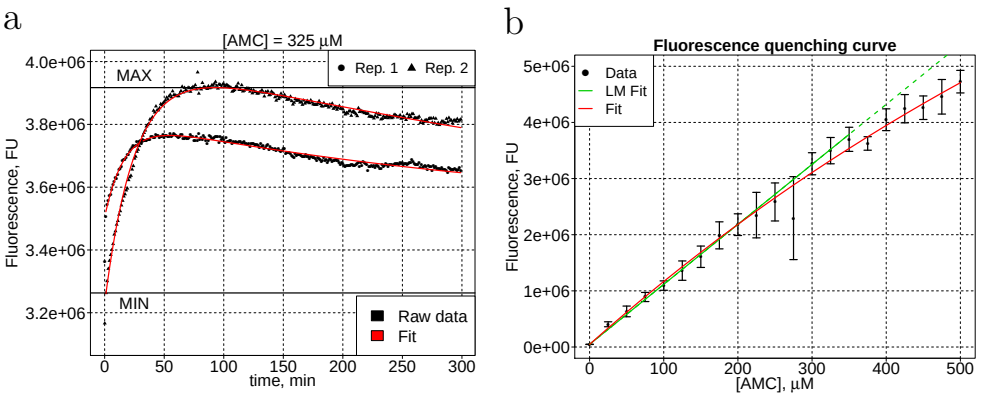


Figure 1.30: **(a)** Two repetitions of time courses of the signal with 325 μM of AMC. Curves are fitted with Eq. 1.40, the first point in both of them is excluded. Group minimum `MIN` and group maximum `MAX` are estimated from the data as shown in the figure. **(b)** The effect of fluorescence quenching — nonlinear relation between the level of fluorescence and the concentration of AMC. Green line represents the weighted fit of a linear model to the subset of data $[AMC] \leq 350\mu$M, where weights are determined as inverse squared error on the points. The continuation of the fit to higher concentrations is indicated by the dashed green line. The fit of Eq. 1.41 to the data is represented by the red line.

the nonlinearity of fluorescence quenching curve is negligible if the concentration of substrate s_0 does not exceed 300 μM, and therefore can be excluded from the model. This is the preferable and pragmatic way of assuring the proper compensation for the effect. Note, that the results in Fig. 1.30 are valid only for the specific machine and type of plasma used. For a different experimental setup, only a reproduction of the same experiment can reliably verify the applicability of the results presented above.

1.7.3 Results and discussion

In this section, we how to account for substrate consumption using kinetic equations and the phenomenological model for thrombin generation. For the data reported here, we showed that the effect of substrate consumption could be safely ignored. In general, a prolongued measurement of thrombin generation that captures the full time course of the kinetics is required to reliably test the model for correction for substrate consumption. However, it is crystal clear that the kinetics of calibration data can not be used to correct for the effect of substrate consumption in thrombin generation data, as CAT currently does. We did not have fluorescence quenching data available for CAT machine, but we investigated this effect for the EnVision fluorometer. The main conclusion is that the instrument transfer function can be indeed constant for relatively short measurement times.

1.8 Models for forward and reverse problems of thrombin generation

Equation 1.31 poses the "forward" problem of thrombin generation: how to find the product concentration $p(t)$ if the concentration of thrombin $T(t)$ and the concentration of α_2-macroglobulin-thrombin α_2M-T(t) are known. The link between $T(t)$ and α_2M-T(t) (Eq. 1.9) and their kinetic constants ([Hem+03], Section 1.6) allows to re-write Eq. 1.31 as

$$p(t) = \frac{s_0 \cdot k_{cat}}{K_m} \cdot \left(\int_0^t T(t') \cdot dt' + \frac{k_{\alpha_2 M}}{2} \cdot \int_0^t (t - t') \cdot T(t') \cdot dt' \right) \ , \tag{1.42}$$

where k_{cat} and K_m are the kinetic constants of thrombin acting on a substrate, and $k_{\alpha_2 M}$ is the rate of conversion of thrombin into α_2-macroglobulin-thrombin (Eq. 1.9). The problem that Eq. 1.42 poses is solved if thrombin concentration $T(t)$ is known. So, inserting the formula for $T(t)$ yields a model for $p(t)$ to be tested on experimental data, in a fashion described in Section 1.6. In principle, any suitable analytical formula for $T(t)$ can be tested against the experimental data using Eq. 1.42, and several models can describe the data equally well.

There is also a "backward" problem of thrombin generation: how to find the concentration of thrombin $T(t)$ if the concentration of product $p(t)$ is known? In the same fashion as Eq. 1.42, Eq. 1.30 for the case of negligible substrate consumption ($S = s_0$) can be re-written as

$$\frac{dp}{s_0}(t) = \frac{k_{cat}}{K_m} \cdot \left(T(t) + \frac{k_{\alpha_2 M}}{2} \cdot \int_0^t T(t') \cdot dt' \right) \cdot dt \ . \tag{1.43}$$

Differentiating Eq. 1.43 with respect to t will give a linear differential equation with constant coefficients and a source term

$$\frac{dT}{dt}(t) + \frac{k_{\alpha_2 M}}{2} \cdot T(t) = \frac{K_m}{s_0 \cdot k_{cat}} \cdot \frac{d^2 p}{dt^2}(t) \ . \tag{1.44}$$

Thrombin $T(t)$ is the cause of the effect $p(t)$, so the proper solution of Eq. 1.44 in terms of $p(t)$ is the following advanced solution

$$T(t) = -\frac{K_m}{k_{cat} \cdot s_0} \cdot \int_t^\infty e^{\frac{k_{\alpha_2 M} \cdot (t' - t)}{2}} \cdot \frac{d^2 p}{dt^2}(t') dt' \ . \tag{1.45}$$

Equation 1.45 allows to calculate $T(t)$ from the second derivative of product concentration $\frac{d^2 p}{dt^2}$ if the value of $k_{\alpha_2 M}$ and kinetic constants k_{cat} and K_m are known. Since k_{cat} and K_m are typically known from the calibration experiment (Section 1.5), it is only $k_{\alpha_2 M}$ that needs to be determined additionally to make use of Eq. 1.45.

The difference between the two approaches is the underlying assumption: the forward problem uses a guess for $T(t)$ and then validates it on the data. The backward

problem uses the data to calculate $T(t)$ "as is". Which of the two approaches is better depends on the goal of the analysis, but in practice the difference in thrombograms between the two is negligible with a proper model choice (see also Appendix **??**).

1.8.1 Models for the forward problem

1.8.1.1 Based on gamma distribution

The simplest phenomenological model for the forward problem of thrombin generation based on gamma distribution (Eq. 1.36, later referred to FM_1 for short) was already presented in Section 1.6. We reproduce the fitting of this model to the control signal here (Fig. 1.31), and keep the the signal fixed for the rest of the section for clear comparison with other models.

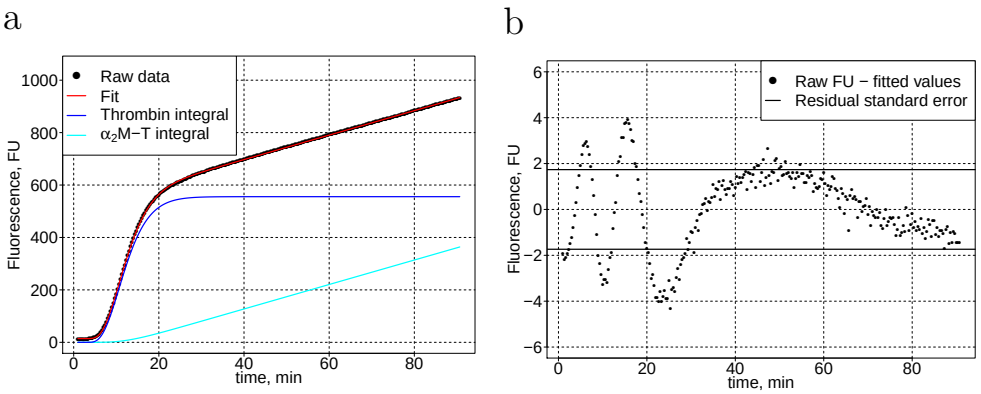


Figure 1.31: Same as Fig. 1.28 but for control signal only.

As discussed in Section 1.6, there is an overall small but clear pattern in the residuals (Fig. 1.31b) of FM_1. It can be easily improved by fitting a slightly more complex model inspired by FM_1 and Fig. 1.1b. There are actually two pathways to thrombin generation (intrinsic and extrinsic), but FM_1 represents only a single enzymatic pathway. It is therefore straightforward to add a second enzymatic pathway to FM_1 to make a more flexible model FM_2.

Let us refer to Eq. 1.32 as $F(t, t_0, A, k, \theta)$. Then, the two-pathway model FM_2 makes the following guess for $T(t)$

$$T(t) = F(t, t_0, A_1, k_1, \theta) + F(t, t_0, A_2, k_2, \theta) \ , \tag{1.46}$$

where each of the pathways is described by a personal set of parameters with corresponding indexes. We only demand the pathways to have the same lagtime t_0 and the same decay rate θ for the model to be qualitatively consistent with how thrombin generation is assumed to happen according to Fig. 1.1b. The fit of FM_2 to the control signal and the residuals of the model are presented in Fig. 1.32.

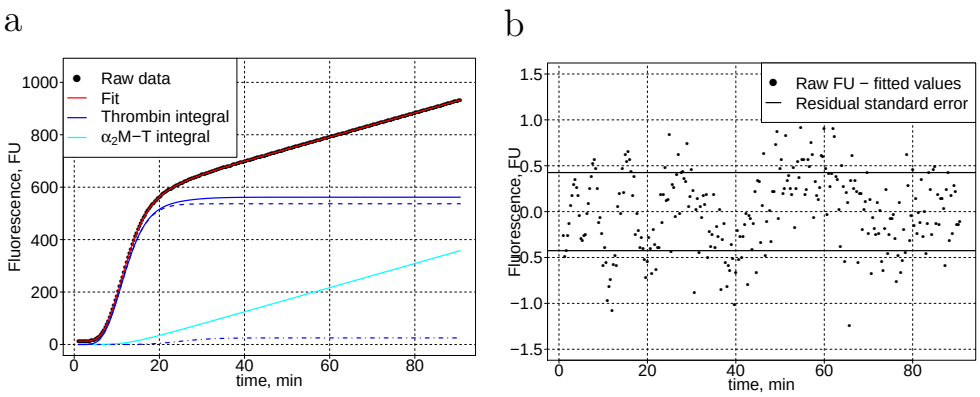


Figure 1.32: Same as Fig. 1.31 but for model FM_2 (Eq. 1.46). The contributions of both pathways to the thrombin integral (blue solid line) are indicated with dashed blue lines. Note how the contributions differ in endlevels or "importances" for the overall fluorescence, being in this case consistent with the idea of a primary and a secondary pathway to thrombin generation (Fig. 1.1b).

Model FM_2 is definitely an improvement over FM_1, because there is hardly any pattern in the residuals in Fig. 1.32b. The contributions of two pathways (Fig. 1.32a) are consistent with the idea of a primary and a secondary pathway to thrombin generation (Fig. 1.1b), because one of them largely dominates over the other. However, this is not typically the case with other thrombin generation signals we analyzed but did not show here, so FM_2 should not be believed to unravel the true underlying mechanism of thrombin generation. It is perfectly suitable though for what it was designed — practical applications, where the key parameters characterizing thrombin generation are of interest.

Both FM_1 and FM_2 are easily transformed for the case of non-negligible substrate consumption (Section 1.7, Eq. 1.39). The fit of FM_1 for substrate consumption is presented in Fig. A.7a-b. For the particular case of the control signal (and from our experience also often in practice if the total measurement time is less than 120 minutes) the models for substrate consumption do not fit the data significantly better than the models for negligible substrate consumption. Still, for consistency, the fit of FM_2 for substrate consumption is presented in Fig. 1.32.

1.8.1.2 Based on rational function

Gamma distribution proved to be useful in the context of the forward problem of thrombin generation (Section 1.6, 1.7, 1.8.1.1), but it is not the only option. Any mathematical function with similar properties has the potential of doing equally well for thrombin generation signals. For example, the following function

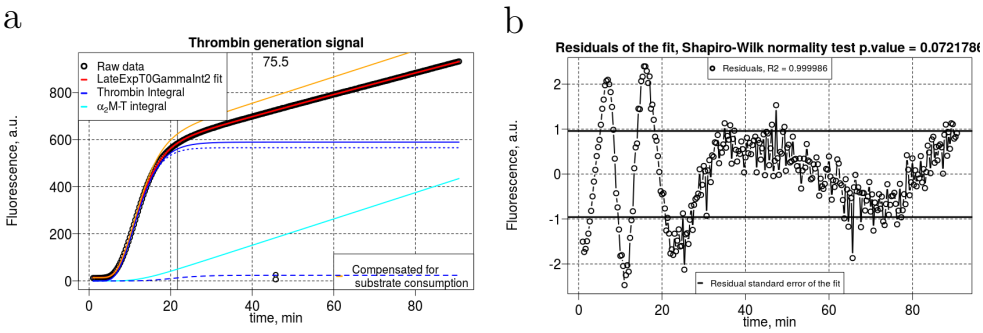


Figure 1.33: Same as Fig. 1.32 but for model FM_2 for the case of non-negligible substrate consumption. No significant improvement over the model for negligible substrate consumption.

$$T(t) = A \cdot t^k \cdot e^{-R(t)} \quad , \tag{1.47}$$

where $R(t)$ is a rational function that satisfies $\lim_{t\to\infty} R(t) = t$ is also a suitable candidate for thrombin concentration (see also Appendix **??**). The prefactor t^k ensures that $T(0) = 0$ and the early times behaviour t^k of the enzymatic cascade (Section 1.6). We used the following $R(t)$

$$R(t) = \frac{p_0 + p_1 \cdot t + p_2 \cdot t^2 + p_3 \cdot t^3}{1 + q1 \cdot t + q_2 \cdot t^2} \quad , \tag{1.48}$$

to provide enough flexibility for the model (later referred to as FM_3) and tested it on the control signal assuming negligible substrate consumption (Fig. 1.34).

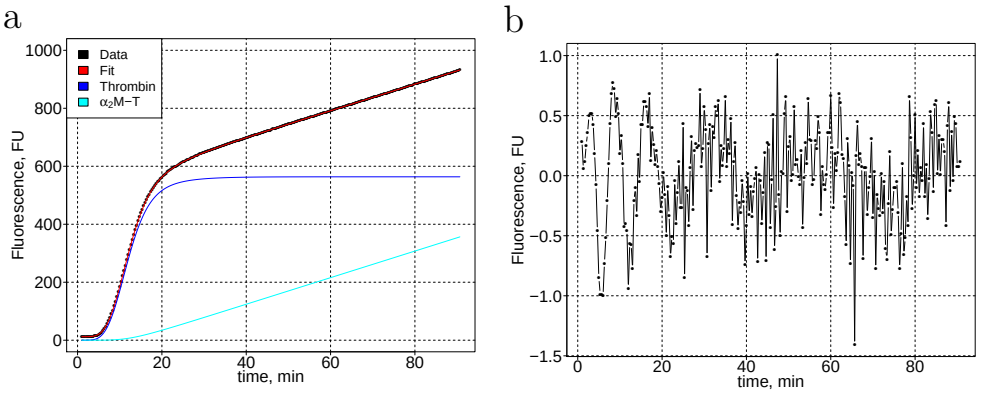


Figure 1.34: Same as Fig. 1.32 but for model FM_3 (Eq. 1.47, 1.48) with negligible substrate consumption. This model performs as good as FM_2 (Fig. 1.32).

We found that FM_3 (9 free parameters, residual standard error 0.41 on 261 degrees of freedom) performs as good as FM_2 (8 free parameters, residual standard error 0.43 on 262 degrees of freedom). However, the fitted value of t_0 in FM_2 was exactly zero, so the same result could be achieved with FM_2 with a fixed parameter $t_0 = 0$. It means that FM_2 saves two degrees of freedom preserving the quality of the fit. It gives relatively easy-to-use closed from analytic expressions for $T(t)$ and its derivatives or integrals that are a built-in part of any contemporary statistical software. Because of the simplicity of the fitted analytic expressions, it also takes significantly less time to fit FM_2 to data than FM_3. Essentially, FM_2 is much better designed for practical use in a high-throughput setting.

1.8.2 Models for the inverse problem

1.8.2.1 Raw data

Equation 1.45 requires $\frac{d^2p}{dt^2}(t)$ and the value for $k_{\alpha_2 M}$ from the raw fluorescence of the thrombin generation signal (see also Appendix **??**). The second derivative can be calculated numerically directly from the raw data with the help of a stencil (Appendix A.3), however the result will inevitably suffer from the amplification of noise in the signal (Fig 1.37a shows this for control signal).

The value of $k_{\alpha_2 M}$ can be roughly estimated using the part of the thrombin generation signal where the thrombin was fully deactivated by antithrombin. In this case, the contribution from the thrombin pathway to the signal is constant, and according to Eq. 1.9 the contribution of α_2M-T pathway grows as a first degree polynomial. It means that if we introduce an offset t_{off} that indicates the start of such a regime in the experimental signal, then we can fit it with the following formula:

$$F = a + b \cdot (t - t_{off}) \quad , \tag{1.49}$$

where t_{off} is is fixed and a and b are free parameters. The value of the parameter a is the sum of three values — the background fluorescence F_0 (experimental signals do not start from zero fluorescence), the constant contribution from the thrombin pathway into the signal A, and the corresponding starting contribution from the α_2M-T pathway $\mathrm{k}_{\alpha_2 M} \cdot A$, i.e.

$$a = F_0 + A + k_{\alpha_2 M} \cdot A \ . \tag{1.50}$$

We can estimate F_0 by simply saying that $F_0 = F(t = 0)$, i.e. background fluorescence equals the fluorescence of the first experimental data point. We use Eq. 1.9 in the next step — if A is the contribution from the thrombin pathway, then $\mathrm{k}_{\alpha_2 M} \cdot A \cdot t$ is the contribution from the α_2M-T pathway, and the latter implies that $b = A \cdot k_{\alpha_2 M}$. Consequently, we calculate $k_{\alpha_2 M}$ and A as follows

$$k_{\alpha_2 M} = \frac{b}{a - b - F_0}, \qquad A = \frac{b}{k_{\alpha_2 M}} \ . \tag{1.51}$$

Figure 1.35a shows three different examples of experimental signals fitted with Eq. 1.49. Both contributions from the thrombin pathway and from the α_2M-T pathway are visualized in the figure, that also illustrates the main source of imprecision in the procedure — the contribution from α_2M-T pathway equals zero until $t = t_{off}$. Another issue is that t_{off} is a free parameter — how do we choose the best value for it? According to Fig. 1.35a, the faster the process of thrombin generation takes place, the lower value of t_{off} can be used in the fit. The dependence of the extracted value of $k_{\alpha_2 M}$ on the offset value t_{off} used is presented in Fig. 1.35b for the black thrombin generation signal. The figure clearly shows a significant decrease of $k_{\alpha_2 M}$ with the increase of t_{off}. It is associated with the increase of the value of parameter a obtained from the fit, because the signal grows in time. So, if the value of t_{off} is too small, $k_{\alpha_2 M}$ will be overestimated and vice versa. Unfortunately, there is no guideline in the raw data to the optimal value of t_{off}.

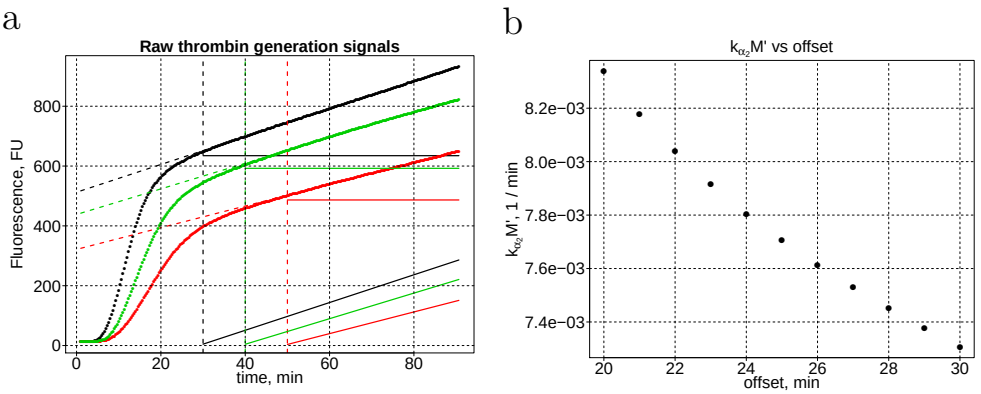


Figure 1.35: **(a)** Examples of typical fluorescence intensity as function of time in a CAT assay. First-degree polynomials (see Eq. 1.49) were fitted to subsets of experimental data in the range $t \geq t_{off}$ and are indicated by the dashed lines. The color of the lines corresponds to the color of the signals. Vertical dashed lines indicate the values of t_{off} used for the fits. Horizontal solid lines visualize the contribution from the thrombin pathway to the signal, and solid lines — the contribution from the α_2-macroglobulin-thrombin pathway. Both contributions were calculated using the values obtained from the fits. **(b)** Dependence of the extracted value of $k'_{\alpha_2 M}$ on the offset value t_{off} used for the black thrombin generation signal.

However, there is a straightforward alternative way to do it that is much more reliable. Equation 1.45 must satisfy the boundary condition $T(0) = 0$ and therefore

$$0 = T(0) = \int_0^\infty e^{\frac{k_{\alpha_2 M} \cdot (t'-t)}{2}} \cdot \frac{d^2 p}{dt^2}(t') dt' \ . \tag{1.52}$$

Equation 1.52 can be solved numerically (Fig. 1.36) and its root will be the proper value of $k_{\alpha_2 M}$ that will be used again in Eq. 1.45 to calculate the full time course $T(t)$

(Fig. 1.37b). Note that no model whatsoever has been fitted to the data to produce the resulting thrombogram and we used the data completely raw.

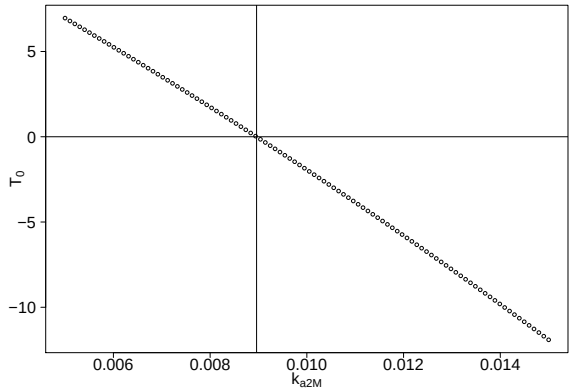


Figure 1.36: Equation 1.52 as a function of $k_{\alpha_2 M}$ and black signal data (Fig. 1.37a) in the vicinity of its root. Vertical solid line is the root of the equation.

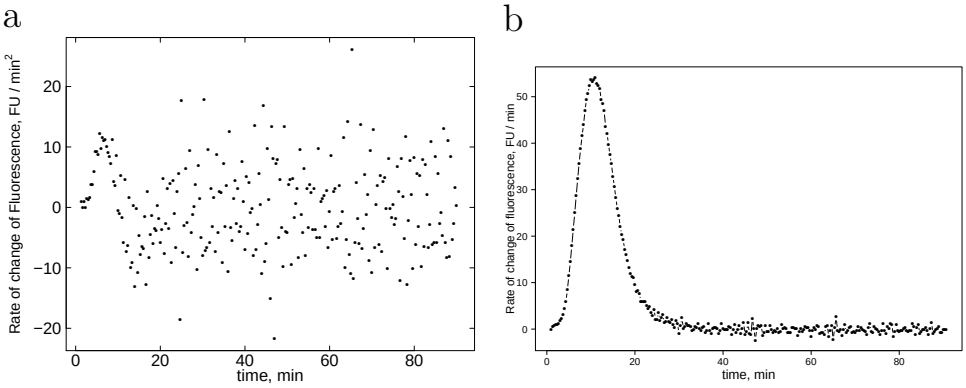


Figure 1.37: (a) Second derivative of the black signal (Fig. 1.35a) taken numerically (Appendix A.3) from the raw data. **(b)** Uncalibrated thrombogram obtained with Eq. 1.45 using the second derivative from **a** and the value of $k_{\alpha_2 M}$ from Fig. 1.36.

The major problem of this approach is that raw data and its derivatives are noisy and therefore the resulting thrombogram is also noisy. On the contrary, the very same procedure on noiseless data would give noiseless results. So the next step is obvious: we need a model for the fluorescence signal that will capture the overall structure of the data good enough and make it noiseless.

1.8.2.2 Based on Padé approximant

We have already seen (Figure 1.35) that the raw thrombin generation signals shown in Fig. 1.5 seem to approach a first-degree polynomial asymptotically. To carefully check that, we fit the control signal at late times (for example $t \geq 50$ min) with a first-degree polynomial. The fitted first-degree polynomial is shown in Fig. 1.38a as a solid red line in the time interval fitting was done in. The dashed red line shows the continuation of the fit to early times. The residuals of the fit are shown in Fig. 1.38b.

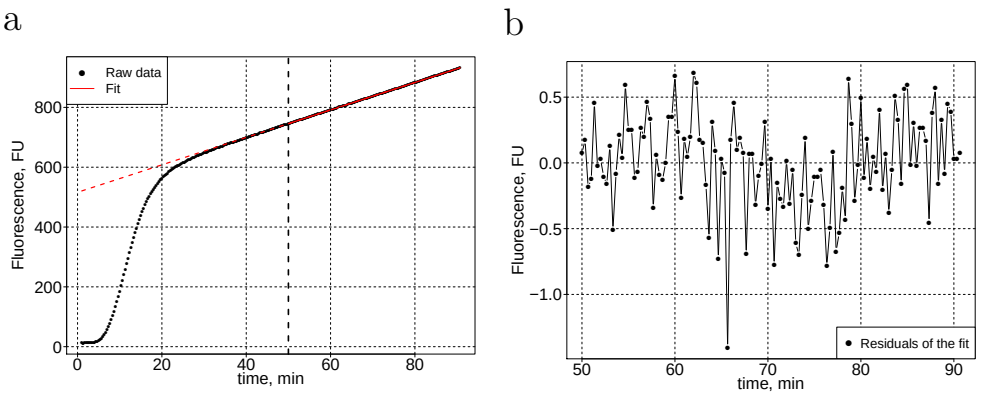


Figure 1.38: **(a)** Raw thrombin generation signal (black circles). Fit of a first degree polynomial to the subset of the raw signal for $t \geq 50$ min (solid red line). Continuation of the fit to early times shown by the dashed red line. **(b)** Residuals of the fit.

The residuals in Fig. 1.38b are very small compared to the fitted values of the signal ($\leq 1\%$), so the signal at late times is indeed well described by a first-degree polynomial. To fit the signal at all times, we need a function that describes the initial rise of the signal and its saturation to a first-degree polynomial after the initial rise. The desired asymptotic behavior is achieved ad hoc, by writing the fluorescence signal $F(t)$ as follows

$$F(t) = a + b \cdot t - e^{-R(t)} \ , \tag{1.53}$$

where $R(t)$ is a rational function of t which is proportional to t for $t \to \infty$, as we already encountered in Section 1.8.1.2 (see also Appendix **??**). This last condition is secured by having the highest power of t in the numerator in $R(t)$ exceed the highest power of t in the denominator of $R(t)$ by one. The general structure of such a function is therefore the following

$$R(t) = \frac{P^{n+1}(t)}{Q^n(t)} \ , \tag{1.54}$$

where $P^{n+1}(t)$ and $Q^n(t)$ are polynomials of power $n+1$ and n, respectively, and n is a non-negative integer. In other words, $\ln(a + b \cdot t - F(t))$ is the well-known Padé

approximant of a rational function of the order [n+1/n]. Several examples of the Padé approximants of various orders are given below

$$\begin{aligned} R[1/0](t) &= \frac{p_0 + p_1 \cdot t}{1} \ , \qquad R[2/1](t) = \frac{p_0 + p_1 \cdot t + p_2 \cdot t^2}{1 + q_1 \cdot t} \ , \qquad \dots \ , \\ R[5/4](t) &= \frac{p_0 + p_1 \cdot t + p_2 \cdot t^2 + p_3 \cdot t^3 + p_4 \cdot t^4 + p_5 \cdot t^5}{1 + q_1 \cdot t + q_2 \cdot t^2 + q_3 \cdot t^3 + q_4 \cdot t^4} \ . \end{aligned} \tag{1.55}$$

By increasing the order of the approximant, we increase the number of free parameters available for fitting the signal. However, what thrombin generation signal $F(t)$ actually represents is the concentration of the fluorescent product $p(t)$ as a function of time. At $t = 0$ the initial concentration of product $p(t = 0)$ equals exactly 0, and so does its first derivative $p'(t = 0)$ and all other derivatives up to and after a certain order, defined by the number of the chemical reactions in the coagulation mechanism. For Eq. 1.53 it means that less parameters are free, and we can fix some of the parameters in the numerator of $R(t)$, as Eq. 1.56 shows.

$$\begin{aligned} F(0) = a - e^{-p_0} = 0 \ &\Rightarrow \ p_0 = -\ln a \ , \\ F'(0) = b + R'(0) \cdot e^{-R(0)} = 0 \ &\Rightarrow \ p_1 = -q_1 \cdot \ln a - \frac{b}{a} \ , \\ F''(0) = e^{-R(0)} \cdot (R''(0) - R'(0)^2) = 0 \ &\Rightarrow \ p_2 = -q_2 \cdot \ln a - \frac{q_1 \cdot b}{a} + \frac{b^2}{2a^2} \ , \\ F^n(0) = \dots = 0 \ &\Rightarrow \ p_n = \ \dots \end{aligned} \tag{1.56}$$

In practice, we limited ourselves to fixing three parameters — p_0, p_1, p_2, and using the [5/4] approximant to guarantee us enough flexibility to fit the signal. The final phenomenological function that we use is therefore the following

$$F(t) = a_0 + a + b \cdot t - e^{-R[5/4]_{0-2}(t)} \ , \tag{1.57}$$

where a_0 is the additional free parameter that accounts for the background in the signal, and $R[5/4]_{0-2}(t)$ is the [5/4] approximant from Eq. 1.55 with p_0, p_1, p_2 given by Eq. 1.56. We fit Eq. 1.57 directly to the raw thrombin generation signal (Fig. 1.39).

Again, the residuals in Fig. 1.39b are small compared to the fitted values and the fit is therefore in good agreement with the data. Still, we carefully examine the residuals to find out how good is our good fit.

First, we check whether the residuals are consistent with the hypothesis of being drawn from a normal distribution, as assumed by the least-squares fitting algorithms. The Shapiro-Wilk test tests exactly that and outputs the p-value for the null hypothesis that data are drawn from the normal distribution. The p-value of the test for the residuals from Fig. 1.39b is 0.11, so even with significance level $\alpha = 0.05$, the null hypothesis is not rejected. The Shapiro-Wilk test is insensitive to non-randomness in the data, so the independence of the residuals has to be assessed separately. We do that by analyzing the autocorrelation plot (correlogram) of the residuals, which is shown in Fig. 1.40a.

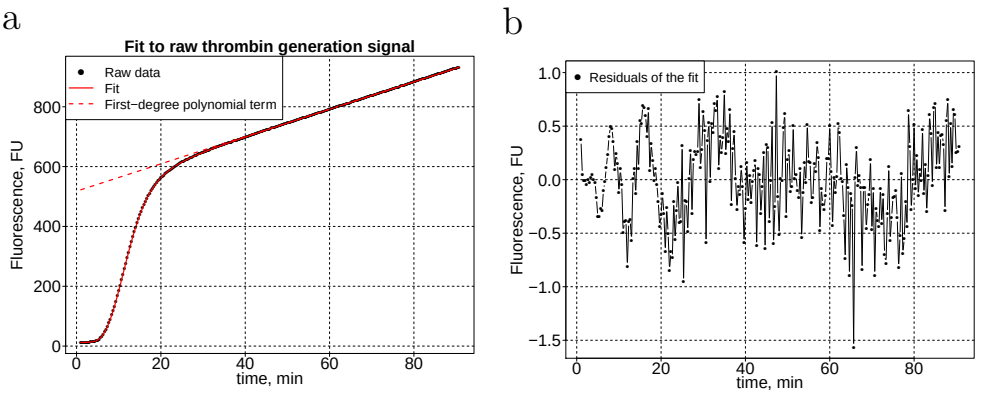


Figure 1.39: **(a)** Fit of Eq. 1.57 (solid red line) to the raw signal (black points). The first-degree polynomial term of Eq. 1.57 shown as the dashed red line. **(b)** Residuals of the fit showing short-range positive correlations superposed with a periodic signal with one-minute period.

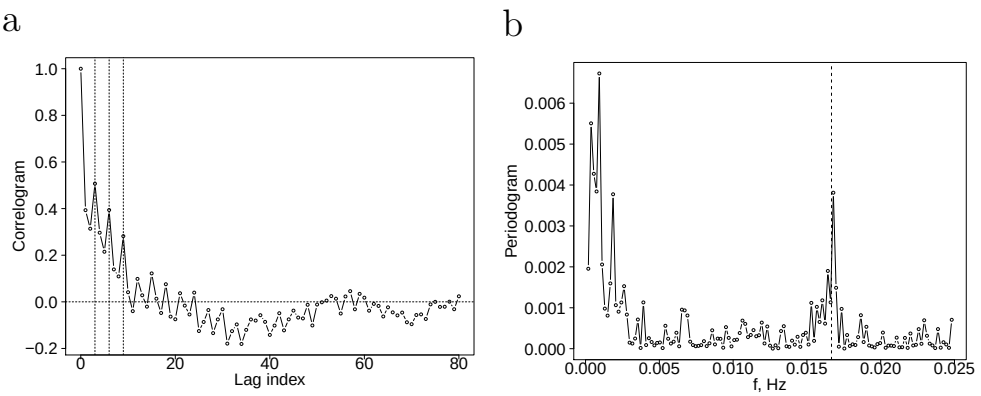


Figure 1.40: **(a)** Correlogram of the residuals. Vertical dashed lines at lag indexes 3, 6, and 9 indicate peaks with period 3. **(b)** Periodogram of the residuals. Vertical dashed line drawn at the frequency 1/60 Hz (one minute periodicity).

The spikes in the correlogram at lag indexes 3, 6, and 9 indicate that for some reason the points that are separated exactly by integer multiple of 3 sampling intervals are more correlated than the others. The period-three component of the autocorrelation function is seen out to lag index 35, approximately, after which it is too weak to show among noise or loss of phase coherence. Since the sampling interval of the raw signal was 20 seconds, the lag of 3 sampling intervals equals exactly one minute.

To check whether this one-minute periodicity stands out clearly in the spectral properties of the residuals, we transform them into the frequency domain using the

Fast Fourier Transform. Figure 1.40b shows the periodogram of the residuals, where the vertical dashed line represents the frequency 1/60 Hz that corresponds to one minute periodicity. Clearly, there is a spike right on that frequency in the periodogram. The detailed analysis of the cause of the spike is presented in Section 1.9.

1.8.2.3 Padé approximant fitted to the first derivative of the raw data

For thrombin generation experiments, the key interest of the analysis is the first derivative of the raw signal, because it is directly linked to the thrombin concentration, or the thrombogram. Therefore to improve the fit, we can try fitting to the first derivative of the signal instead of the raw data (see also Appendix **??**). We take the first derivative of the raw signal numerically using a stencil. The properties of various stencils were investigated in Section 1.4.1. For this subsection, we will use the off-grid stencil

$$F'_{k+1/2} = \frac{F_{k+1} - F_k}{\Delta t} \quad . \tag{1.58}$$

The first derivative of the raw signal is obtained using Eq. 1.58. Figure 1.41a shows the fit to it, Fig. 1.41b shows the residuals of the fit. The correlogram and the periodogram of the residuals from Fig. 1.41b are shown in Fig. 1.42.

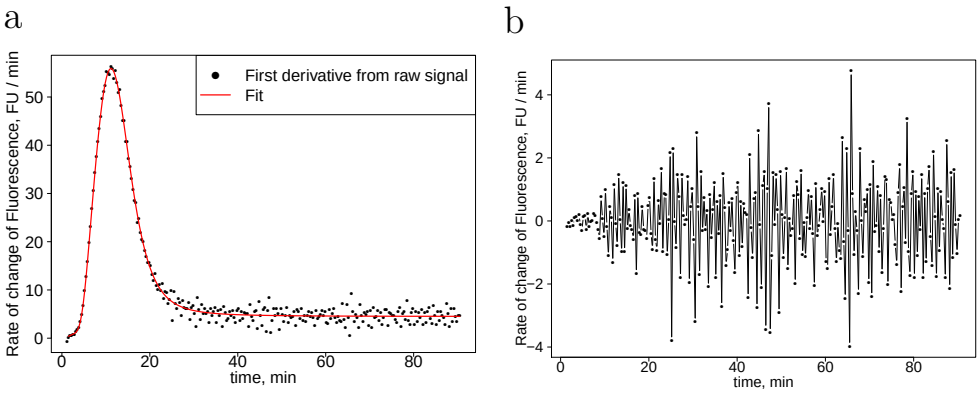


Figure 1.41: (a) Fit of the first derivative of Eq. 1.57 (solid red line) to the first derivative of the signal taken numerically using Eq. 1.58 (black points). **(b)** Residuals of the fit. Note the period-three component in the residuals.

Fig. 1.42a shows the similar periodicity of the correlogram as the Fig. 1.40a. Fig. 1.42b shows, as expected, the same strange spike around the frequency 1/60 Hz. However, at lower frequencies the periodogram of the residuals looks similar to periodogram of the additive white Gaussian noise (AWGN). The AWGN has the property that the values of the periodogram are exponentially distributed around their expected values [BF04], and we can check this property in the periodogram shown in Fig. 1.42b.

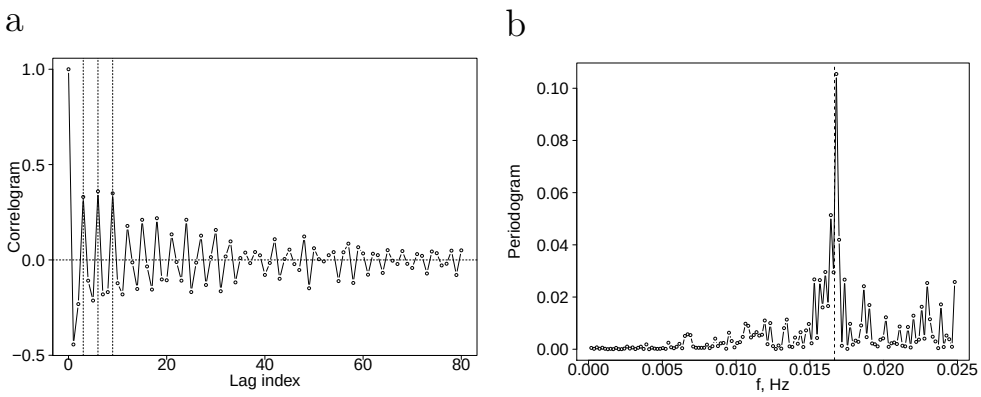


Figure 1.42: **(a)** Correlogram of the residuals from Fig. 1.41b. Vertical dashed lines drawn at lag indexes 3, 6, and 9 to indicate peaks with period 3. **(b)** Periodogram of the residuals. Vertical dashed line represents the frequency 1/60 Hz (one-minute periodicity).

The Kolmogorov-Smirnov test can be used to compare the distribution of the values of the periodogram to the reference probability distribution. In Fig. 1.43a, we select subsets of the periodogram $0 \leq f \leq f_{max}$, where various f_{max} are indicated by dashed gray lines. Along each of the gray lines, we show the p-values of the Kolmogorov-Smirnov test with exponential distribution as the reference distribution. For frequencies up to 0.01 Hz, the distribution of the values of the periodogram is still consistent with the hypothesis of being exponential. It means, that the periodogram in Fig. 1.40b (the residuals of the fit to raw signal) might be the integrated white noise plus the spike at 1/60 Hz and therefore — be proportional to $\frac{1}{f^2}$ at lower frequencies. This property is checked in Fig. 1.43b, where the periodogram from Fig. 1.40b is plotted on the log-log plot. For the purposes of better visualization, we also show the blocking average of the data (mean is indicated by the blue crosses, error bars are standard errors on the mean). In the region indicated by the two dashed red lines, we fitted the raw data with a first-degree polynomial (solid red line). The slope of the line is -1.87, which is fairly close to -2 that we anticipated above.

The key result here is that fitting to the first derivative of the raw signal instead of the raw data itself could be preferred because of the properties of the residuals of the fit. As Fig. 1.43 shows, the residuals of the fit to the first derivative of the data resemble the AWGN much more than the residuals of the fit to the raw data. However, taking derivatives numerically makes even uncorrelated data correlated and is still an approximation to the true derivative of the data, so the data after the application of any stencil is no longer raw. For the practical purpose of the analysis — to estimate the phenomenological parameters of the curve, the differences between the two approaches are subtle, because the quality of the fit is sufficiently high in both cases.

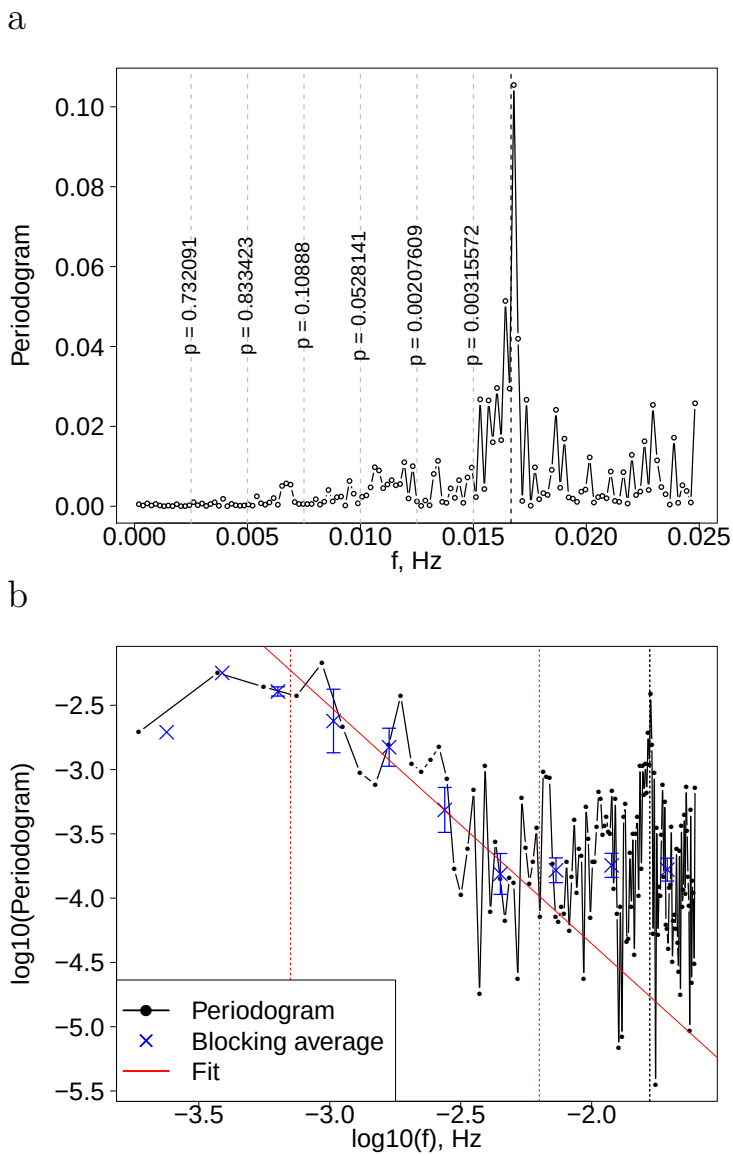


Figure 1.43: **(a)** Periodogram of the residuals from Fig. 1.42b. Dashed gray lines indicate various f_{max} that define the subsets of the periodogram $0 \leq f \leq f_{max}$. Along each of the gray lines, the p-values of the Kolmogorov-Smirnov test with exponential distribution as the reference distribution are shown. Vertical black dashed line represents the frequency 1/60 Hz that corresponds to one minute periodicity. **(b)** The periodogram from Fig. 1.40b as the log-log plot with base 10 (black points). The blocking average of the data (mean is indicated by the blue crosses, error bars are standard errors on the mean). The fit of a first-degree polynomial (solid red line) was performed to the subset of data between the two vertical dashed red lines.

1.8.2.4 Based on pure Padé approximant

Examining Eq. 1.57, we asked ourselves whether it can be simplified. The asymptotic first degree polynomial of Eq. 1.57 can be as well captured by $R(t)$ itself, so it might be possible to achieve the same quality of fit with a simpler function (see also Appendix **??**). Therefore, we fitted the following model to the control signal (Fig. 1.44)

$$F(t) = a_0 + \frac{p_2 \cdot t^2 + p_3 \cdot t^3 + p_4 \cdot t^4 + p_5 \cdot t^5}{1 + q_2 \cdot t^2 + q_3 \cdot t^3 + q_4 \cdot t^4} \quad . \tag{1.59}$$

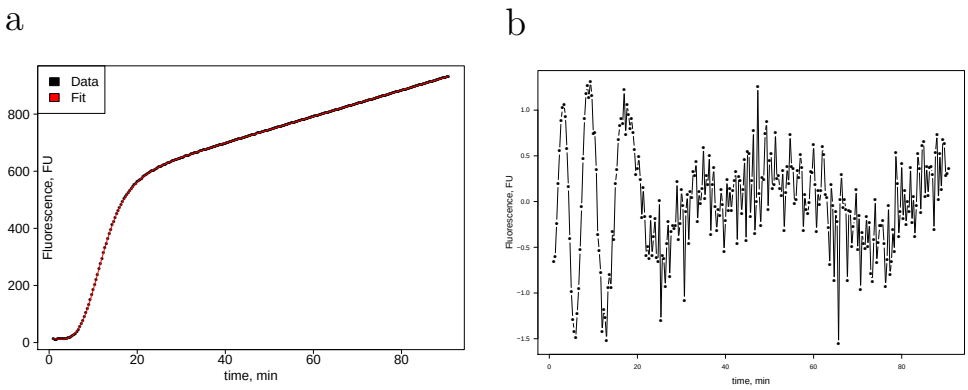


Figure 1.44: **(a)** Fit of Eq. 1.59 to the control signal. **(b)** Residuals of the fit.

The quality of the fit (8 free parameters, residual standard error 0.57 on 262 degrees of freedom) from Fig. 1.44 is only slightly worse than that of Fig. 1.39 (10 free parameters, residual standard error 0.44 on 260 degrees of freedom). It proves once again that with enough free parameters different models can fit same data equally fine: does not matter whether Eq. 1.57 or Eq. 1.59 are used as long as each of them describes the data well.

1.8.3 Comparison of forward and inverse thrombograms

In the sections above, we have considered a number of models both for forward and inverse problems of thrombin generation. Each model gives a resulting thromboram, either as a closed-form analytic expression (Eq. 1.46) or via Eq. 1.45. For example, models Eq. 1.46 (8 free parameters, residual standard error 0.43 on 262 degrees of freedom) and Eq. 1.57 (10 free parameters, residual standard error 0.44 on 260 degrees of freedom) are not significantly different, but will their thrombograms be significantly different? We show the answer in Fig. 1.45.

There is a little difference in the resulting thrombograms (Fig. 1.45a), that is obviosuly partly explained by the fact that we fitted different models to the data.

However, there is typically some numerical approximation behind the scenes when using Eq. 1.45, that could potentially account for some of the total difference. Figure 1.45b shows that numerical approximation affects the resulting thrombogram only negligibly when the model for the data is fixed.

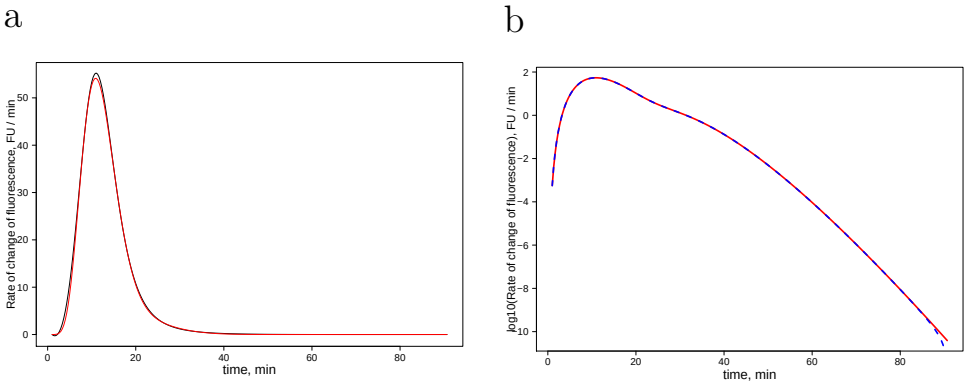


Figure 1.45: **(a)** Uncalibrated thrombogram from Eq. 1.57 (solid black line) compared to the uncalibrated thrombogram from Eq. 1.46 (solid red line). Little difference is observed between the two thrombograms. **(b)** Red thrombogram from **a** compared to the thrombogram from Eq. 1.46 used as if for inverse problem (Section 1.8.2.1, dashed blue line) on the logarithmic scale. This figure proves that the difference in **a** is associated almost entirely with the choice of model and negligibly with the choice of problem.

The idea in Fig. 1.45b is very simple — let us use the forward model FM_2 (Eq. 1.46) as the inverse model. That is, we take the fit (solid red line) from Fig. 1.32, calculate its $\frac{d_2 p}{dt^2}$ analytically, find the proper value of using Eq. 1.52 and plug it all into Eq. 1.45 to calculate the thrombogram, exactly as we already did in Section 1.8.2.1. And then we compare the result to the blue solid line from Fig. 1.32. The two thrombograms have to be plotted on logarithmic scale to see the differences! It proves that the difference in thrombograms comes almost entirely from the difference in data approximation by chosen models (Eq. 1.57 and Eq. 1.46), and negligibly from the choice of the problem (forward or inverse). This, however, leads once again to an important conclusion: If there is no difference between the forward and the inverse problem when the model for the data is fixed, the forward problem is much better suited for practical application, because it fits the thrombogram right away and requires no additional calculations afterwards. We only have to make sure that the quality of the fit is sufficient to represent the structure of the data.

1.8.4 Results and discussion

In this section, we have formulated the forward and the inverse problem of thrombin generation (see also Appendix **??**). We have considered several models for each of the problems and found that with enough free parameters different models can fit the same data equally fine. We also considered fitting to the first derivative taken numerically from the raw data instead of the raw data itself. Though a stencil inevitably introduces additional correlation to the derivative, we found the residuals of such fit to be closer to AWGN at least in the low-frequency range. We compared the resulting thrombograms from different model and found only little difference, associated the model choice, but not with the choice of problem (forward or inverse).

1.9 Strange spikes in raw data from CAT machines

1.9.1 Periodogram of the raw signal has a strange spike at 1/60 Hz

In the previous section we found a strange spike that stands out strong in the periodogram of the residuals of the fit (Fig. 1.40b). We also verified that it is seen directly in the periodogram of the raw data. To do that, we first "reflected" the raw signal around the last experimental time point to make it periodic without pronounced discontinuity. Next, we applied FFT to this periodic signal and obtained its periodogram, which is shown by the black curve in Fig. 1.46 both on the linear and the double logarithmic scale. This periodogram consists of two contributions: One from the noiseless part of the signal and the other from the noise. We fitted the raw signal with Eq. 1.36 to get its noiseless part, and plotted it against the raw data in Fig. 1.46 (red curve), and its periodogram — against the periodogram of the raw data. The red curve can be represented as the sum of the blue curve, the cyan curve, and the background fluorescence, where blue represents the contribution from thrombin, and cyan — the contribution from α_2-macroglobulin thrombin.

The periodogram of the fit helps us to interpret the periodogram of the raw signal. The broad spikes of the periodogram come from the blue part, whereas the periodogram of the fit for the frequencies above 0.05 is almost solely represented by the cyan curve. Therefore, the strange spike at the frequency 1/60 Hz is clearly from the contribution of the noise, because its value of the periodogram is 10^3 times larger than that of the contribution from the noiseless part.

Examining the data from the CAT machine closer, we found a property of it that might explain the strange spike at 1/60 Hz that we see. The property is that when the sampling time of the machine is set to 20 seconds, it introduces significant round-off errors. The measurements of the CAT machine read

```
  "No"   "1"   "2"   "3"   "4"   "5" ...  "269"  "270"
"time"     1  1.33  1.67     2  2.33 ...  90.33  90.67
  "fl" 12.34 12.11 12.18 12.35 12.52 ... 930.58 932.15
```

whereas with the higher precision they would read

```
  "No"   "1"             "2"             "3"   "4"             "5" ...
"time"     1 1.3333333730697 1.6666667461395     2 2.3333334922790 ...
  "fl" 12.34           12.11           12.18 12.35           12.52 ...
```

In Fig. 1.46, we used the timing with higher precision. If we repeat the analysis in Fig. 1.46 for the timing with round-off errors, we get a significant difference in the behavior of the fitted curves at the frequency 1/60 Hz, as Fig. 1.47 shows.

This round-off error might happen on the software level of the CAT machine, but it is most likely the result of parsing from the raw ".dat" format of the CAT machine to Excel, because of the default setting of Excel to round off to two decimals.

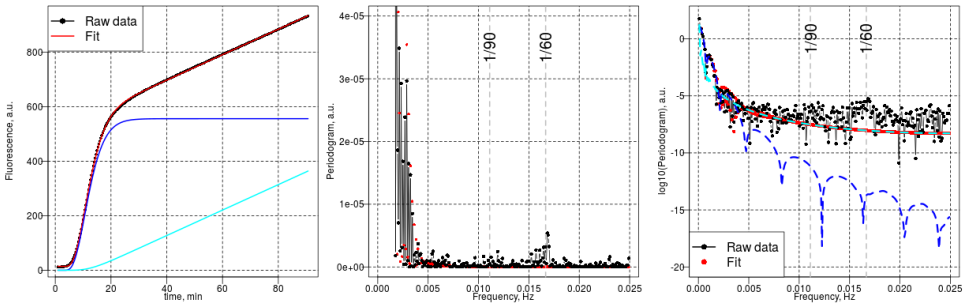


Figure 1.46: **Left**: Raw data (black curve) and the fit of a phenomenological function to it (red curve). The red curve is the sum of the blue curve, the cyan curve, and the background fluorescence where blue represents the contribution from thrombin, and cyan — the contribution from α_2-macroglobulin thrombin. **Middle**: Periodograms of the raw data and the fit from the left panel. The limits of the y-axis are fixed throughout the report to illustrate the strange spike around 1/60 Hz. textbfRight: Periodograms of the raw data and the fit from the left panel on the double logarithmic scale with base 10. Vertical grey lines indicate the frequencies 1/60 Hz and 1/90 Hz. Note the strange spike at the frequency 1/60 Hz that looks to be man-made.

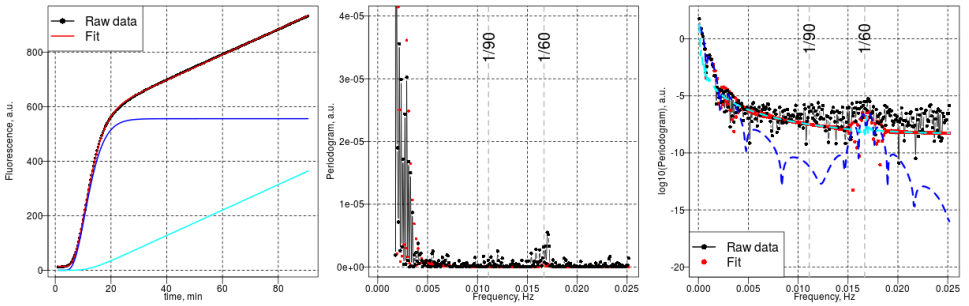


Figure 1.47: Same as in Fig. 1.46 but for the signal with the round-off errors. Note how the behavior of the red, blue and cyan curves is different from Fig. 1.46 at the frequency 1/60 Hz.

1.9.2 Strange spikes in all data from CAT machines

We checked whether the abovementioned strange spike at 1/60 Hz is property of the CAT machine in general, by using all available data from our collaborators. We found all data from CAT machines to have strange spikes, most often at the frequencies 1/90 and 1/60 Hz. A few illustrative examples are presented in Fig. 1.48.

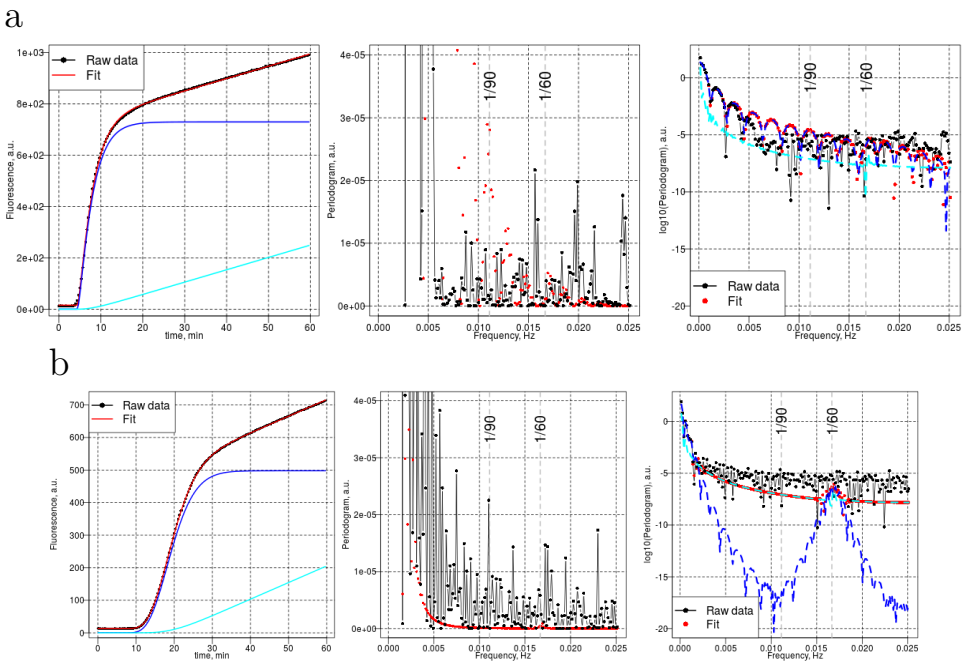


Figure 1.48: Same as Fig. 1.46 but for other datasets we obtained. Note the spikes at 1/90, and 1/60 Hz.

1.9.3 Results and discussion

We found all data from CAT machines to have strange spikes, most often at the frequencies 1/90 and 1/60 Hz. We believe that the strange spikes were most likely caused by the parsing of raw data from CAT machine into Excel.

1.10 Thrombin Analyzer: A web application for data analysis in thrombin generation experiments

The work in this project was a collaboration with Novo Nordisk A/S. Our collaborators did not have access to CAT machines and needed to perform data analysis themselves. They soon realized that they needed help with it and this is how we got on board. One of the things our collaborators wanted was a transparent alternative to CAT: An open-source method that is capable of taking raw data from an arbitrary fluorometer, analyze it and output the results of the analysis. This is a step towards standartization in the field of thrombin generation based on a transparent open-source method of data analysis rather than on a black box commercial assay (see also [Woo+13]) that guarantees easy money for its creators.

Though CAT was introduced by Hemker [Hem+03], he claims that the software for data analysis used in the assay was developed by a third party [HK13]. However, he still published the methods he used for data analysis in [HK13] that were the most likely source of inspiration for the third-party software developers of CAT machine.

Even though the methods of data analysis used by the CAT machine might be gray box rather than a true black box, it still limits the experimental protocol for the user, which is a problem in research. The data analysis software of the assay can not be used independently and this is yet another problem.

We investigated the data analysis by CAT machine and various ways to improve it and make it transparent in the previous sections. Applying the results of our work is relatively straightforward for a highly-educated person with academic background, but it is likely to be tricky for a laboratory technician. This is why we created an easy-to-use solution for this: A web application for transparent automated data analysis in thrombin generation experiments that we called Thrombin Analyzer [Mar]. It allows to analyze individual signals as well as datasets directly in the web browser. The application has a detailed description and a tutorial available online, so they will not be discussed here.

You might think that there is not much need for automated data analysis if the experimental dataset consists of three thrombin generation signals (Fig. 1.5). However, in reality researchers deal with up to 384 signals in a single experiment and tens of runs per day when a massive drug discovery campaign is going on. Hundreds of thousands compounds with unknown properties will be tested in search for the desired result. There is simply no way a person can manually analyze and digest that much information, but that is easy for computers.

For appetizers, Fig. 1.49 shows an example of a dataset from a real industrial setting visualized in Thrombin Analyzer. Suppose that the upset researcher (the majority of signals in the dataset simply did not show thrombin generation at all) who made the dataset saw that there might be something in compound 379 and wants to analyze it. He selects the signal, chooses the model `Auto` and in a few seconds looks at the summary of the fitted model chosen by the computer (Fig. 1.50a). It seems like the measurement time was a little short for this thrombin generation signal to really

get to the first degree polynomial asymptote, but still the chosen simple model based on gamma distribution does a fair (and very fast) job. Having read the Thrombin Analyzer tutorial, the researcher decides to check if the signal could be fitted better with a two-cascade model based on gamma distribution (Fig. 1.50b). Indeed, this is much better! And it might also help to save someone's life ...

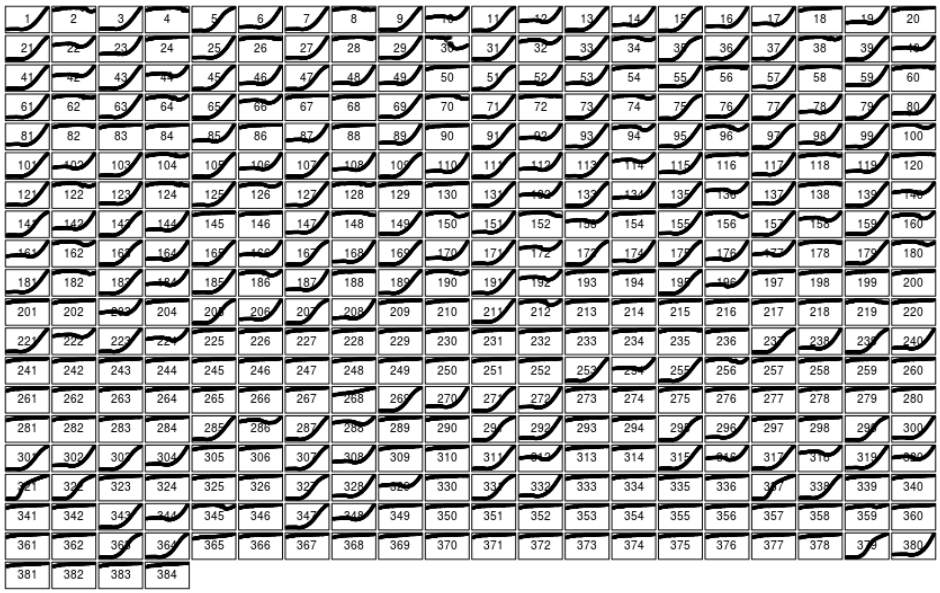


Figure 1.49: An example of a thrombin generation dataset from a real industrial setting, visualized in Thrombin Analyzer.

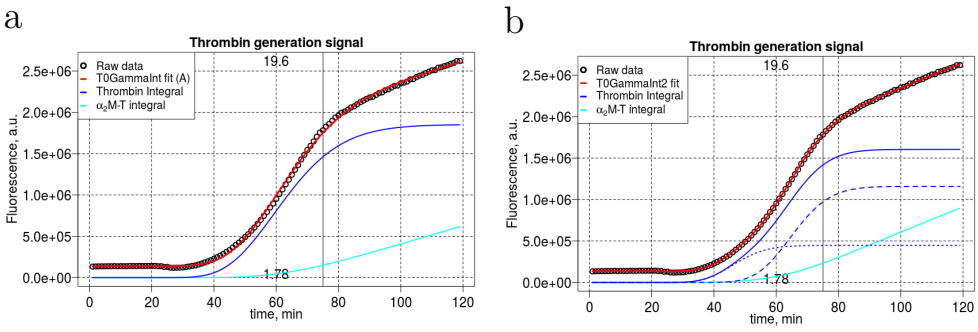


Figure 1.50: **(a)** Signal 379 from Fig. 1.49 analyzed automatically. **(b)** Signal 379 analyzed manually with a more complex model. All analyses performed in Thrombin Analyzer.

1.11 Conclusion

Thrombin is the enzyme that plays the key role in the blood coagulation process. Regular infusions of thrombin help to deal with hemophilia, which is currently best known among coagulation disorders. Therefore, thrombin generation in different physiological conditions is of major interest both for research and pharmaceutical industry.

The field of thrombin generation emerged in early 1950s and is still actively developing. Over the years, the ongoing progress in experimental methods of thrombin generation allowed to introduce ready-to-use commercial assays for thrombin measurement. However, the true mechanism of thrombin generation is still a mystery: Considered initially as a two-pathway mechanism and later as a single-pathway one, it is currently shifting towards the plasma composition-based model. This is why experimental methods in the field dominate largely over theoretical methods. However, experimental data are only useful when analyzed properly, and this is where we had something to offer to the field — a transparent method of data analysis in thrombin generation experiments.

The users of commercial assays have to deal with obvious disadvantages — experimental protocols are often strictly limited, and the details of automated data analysis are typically hidden. Therefore, it is not possible to fully reproduce the data analysis performed by the assay and validate the results, which is often required in research. CAT aims to become a standartized assay in the field, but the standartization can hardly be done by simply using a fixed instrument. On the contrary, it should be based on the experimental protocol and the data analysis procedure. Ideally, the experimental compounds prepared according to the same experimental protocol in different labs and measured with different instruments should give similar results, when a standard data analysis procedure is applied to the obtained signals.

We reverse engineered the data analysis of CAT, relying entirely on the methods reported by its creators. Since these methods use numerical differentiation, we considered the effects of applying various stencils. We found that noise-robust stencils are much better suited for the job than their conventional variants because of the properties of their magnitude response. At the same time, the use of any numerical stencil on the originally uncorrelated data will inevitably introduce correlations in the resulting derivative. We were able to produce results that are close enough to the original output of CAT, so from now on we consider it a gray box rather than a complete black box. However, the reverse engineering that we performed revealed a number of possibilities to improve the methods of data analysis that CAT relies on.

We found that experimental calibration data is described well with textbook Michaelis-Menten kinetics of the basic enzymatic reaction, taking into account the effects of non-constant background fluorescence from the substrate and the aging of plasma. The duration of measurement and the concentration of enzyme α_2M-T in the compound of the calibration signal define whether it is best described by the early-times or late-times approximate analytic solution of Michaelis-Menten kinetics. In any of the cases, the use of the approximate analytic solutions instead of the nu-

merical solutions is justified because the initial concentration of enzyme is 1000 times lower than the concentration of substrate. Calibration signal contains the information about two kinetic rates — k_{cat} and K_M, that can be extracted from the fit of late-times approximate analytic solution, whereas initial rate measurements are not able to provide it. We showed that there is simply no need for the previously reported strategy of titration for the calibration experiment — a single complete time course of calibration is all that is required to calibrate any thrombin generation signal measured in the same plasma.

As for thrombin generation data, we first considered a simple phenomenological model based on gamma distribution. It is inspired by the qualitative behaviour of a cascade of basic enzymatic reactions, which is close enough to the contemporary understanding of how thrombin generation happens. This model approximates the data fairly good, given its simplicity, and is fast to implement and use. We also found that kinetic equations for thrombin generation data yield a definition of the calibration factor that is a factor 1.6 different from the one used by CAT machine. This is why CAT greatly overestimates the parameters of thrombogram that do not depend on time. We found that the values of the rest of the parameters agree well with the values provided by CAT. In this way, we reported a complete benchmark for our method and proved it to work.

We investigated the effects of substrate consumption and fluorescence quenching. Using kinetic equations, we showed how to account for substrate consumption by plugging in the phenomenological model for thrombin generation into an additional $(1 - \exp(\dots))$ term. For the reported thrombin generation data, we showed that the effect of substrate consumption was negligible and could be safely ignored. In general, a prolongued measurement of thrombin generation that captures the full time course of the kinetics is required to reliably test this model for correction for substrate consumption. However, it is crystal clear that the kinetics of calibration data can not be used to correct for the effect of substrate consumption in thrombin generation data, as CAT currently does. We presented the experimental protocol that allows to investigate the effect of fluorescence quenching in a predefined experimental window. We did not have fluorescence quenching data available for CAT machine, but we investigated this effect for the EnVision fluorometer. We showed that the contribution of this effect can be made negligible for an EnVision fluorometer, if the amount of substrate used in the experiment is less than 300 μM. The main conclusion however is that the instrument transfer function can be indeed considered constant for relatively short measurement times.

We formulated the forward and the inverse problem of thrombin generation. We have considered several models for each of the problems and found that with enough free parameters different models can fit the same data equally fine. We also considered fitting to the first derivative taken numerically from the raw data instead of the raw data itself. Though a stencil inevitably introduces additional correlation to the derivative, we found the residuals of such fit to be closer to additive white gaussian noise at least in the low-frequency range. We compared the resulting thrombograms from different models and found only little difference, associated with the model

choice, but not with the choice of problem (forward or inverse).

We found all data from CAT machines to have strange spikes, most often at the frequencies 1/90 and 1/60 Hz. We believe that the strange spikes were most likely caused by the parsing of raw data from CAT machine into Excel.

Essentialy, we investigated the data analysis by CAT in details and considered various ways to improve it and make it transparent. Applying the results of our work is relatively straightforward for a highly-educated person with academic background, but it is likely to be tricky for a laboratory technician. This is why we created an easy-to-use solution for this problem: A web application for transparent automated data analysis in thrombin generation experiments that we called Thrombin Analyzer. It allows to analyze individual signals as well as datasets directly in the web browser. We see it as a ready-to-use tool for a step towards standardization in the field of thrombin generation based on transparent and well-motivated methods of data analysis. Fortunately for us, this is definitely something that the field is currently lacking.

CHAPTER 2

Pericytes, hypertension and neurovascular coupling

2.1 Background

The discovery of pericytes was originally made by scientists whose primary interest was the nature of capillary contractility. Eberth is often credited for describing the existence of pericytes in 1871 [J71]. However, their actual discovery is commonly associated with Rouget [C73; C79], who described a population of contractile cells surrounding the endothelial cells of small blood vessels. Later on, these cells we called "Rouget cells" by Zimmerman, who introduced the name "pericyte" in 1923 [K23]. A reproduction of one of the original preparations of pericytes by Zimmerman is shown in Fig. 2.1.

From 1920 to 1950 a number of publications reported on pericytes. Some of them disagreed that pericytes were contractile cells of the capillaries [Sim86]. The differences between the findings were partly caused by experimental reasons, but they also provided a strong motivation to study the heterogeneity of pericytes and to remove the existing confusion about cell identities. Nowadays, it has become clear that different cell types occupy the periendothelial compartments, but their correct identification still remains challenging [MI10].

There are several currently accepted criteria for a cell to be considered a pericyte: it should be embedded within the vascular basement membrane and should be present in microvessels. A detailed discussion of the problematics of the exact definition of pericytes is given in [AGC11], as well as a detailed illustration of the pericyte anatomy that is reproduced in Fig. 2.2. Figure 2.3 from [Kis+17a] visualizes pericytes and other cellular elements in the neurovascular unit.

One of the recent surprising discoveries about pericytes was their active role in the regulation of the blood flow [M+06; Hal+14], because it was classically thought to be the result of the relaxation of smooth muscle cells around arterioles. At the same time, this idea was also dismissed in [Hil+15], where the authors stated that the regulation of cerebral blood flow is mediated by arteriolar smooth muscle cells. Both ideas are discussed in detail in Section 2.1.1, and a further clarification of the true role of pericytes in blood flow regulation in the brain is given in Sections 2.5 and 2.6.

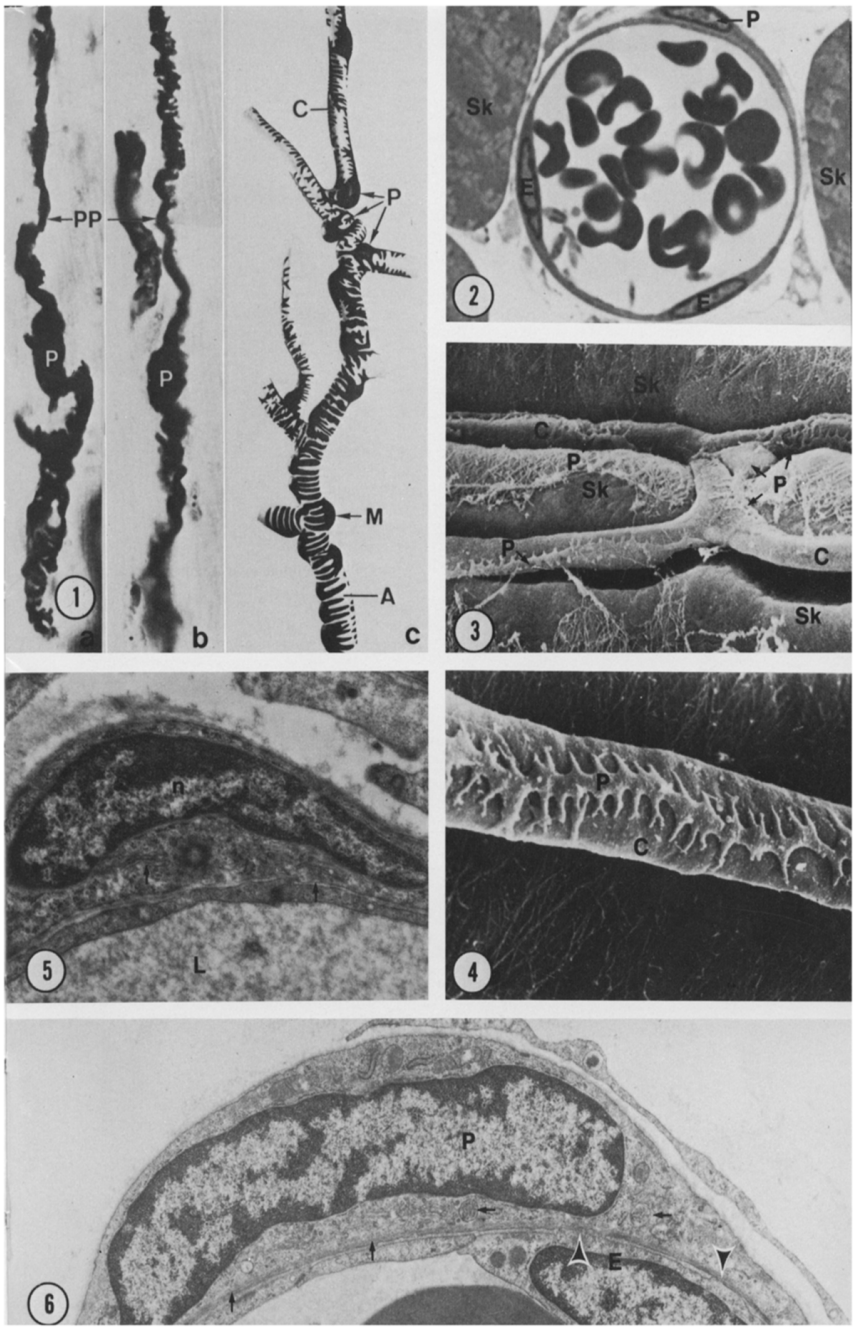


Figure 2.1: Light micrographs of cardiac peficytes from Zimmermann's original Golgi-Kopsch preparations **(a, b)**, and **(c)** a model, also by Zimmermann, of the intermediate cells (M) and pericytes (P) along an arteriole (A) and capillaries (C). The long primary processes of capillary pericytes are indicated in **a** and **b** as PP. Figure and caption reproduced from [R74].

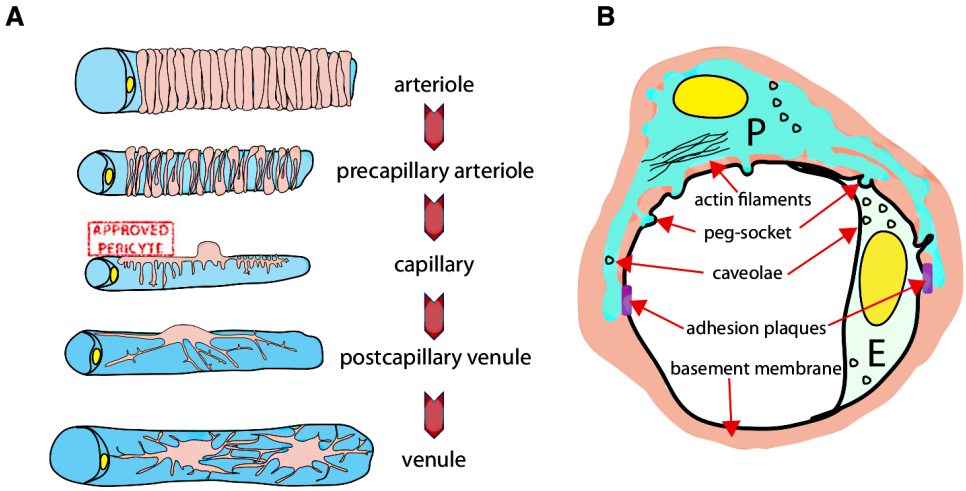


Figure 2.2: **(a)** A continuum of mural cell cyto-architecture from arteriole to venule. A single vascular smooth muscle cell layer around arterioles and precapillary arterioles encircles the entire abluminal side of the endothelium. Vascular smooth muscle cells around arterioles have a flattened, spindle-shaped appearance with few cytoplasmic processes, whereas around precapillary arterioles the cell bodies are distinctly protruding and extend several processes encircling the endothelium. Pericytes investing capillaries have a nearly rounded cell body that gives rise to a few primary processes running on the endothelium in the length of the capillary. The primary processes give rise to secondary perpendicular processes. The tips of secondary processes attach firmly to the endothelium. On postcapillary venules the mural cell body flattens and gives rise to many slender, branching processes. Vascular smooth muscle cell covering venules have a relatively big, stellate shape cell body with many branching processes, which, unlike arteriolar vascular smooth muscle cells, do not wrap circularly around the endothelium. **(b)** Ultrastructural characteristics of pericytes and pericyte-endothelial interactions. Pericytes are rather anonymous in transmission electron microscopy. The mature capillary pericyte (P) has a discoid nucleus that is surrounded by a small amount of cytoplasm containing protein-producing organelles and mitochondria. Microtubules stretch along the primary and secondary cytoplasmic extensions. Intermediate filaments composed of desmin and vimentin are mostly concentrated within in the primary extensions. Dense bands of microfilaments containing actin, myosin, and tropomyosin are concentrated beneath the plasma membrane, in particular the inner surface membrane facing the endothelium. The outer, abluminal pericyte surface often shows numerous caveolae. Despite being separated by the shared BM, pericytes and endothelial cells (E) make numerous direct contacts of different type: schematically depicted are peg-socket contacts and adhesion plaques. Figure and caption reproduced from [AGC11].

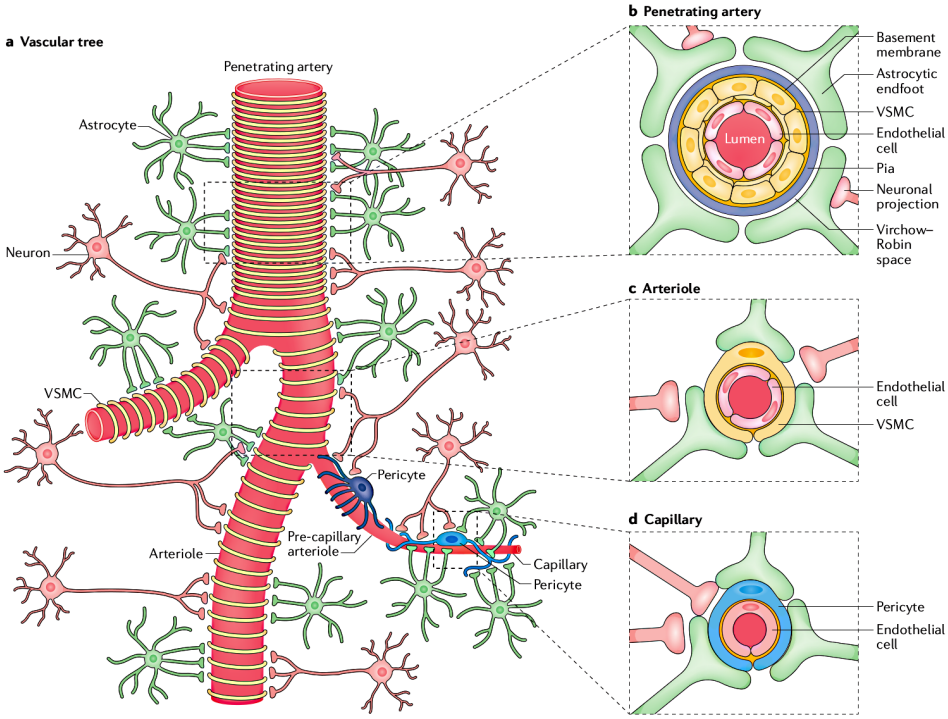


Figure 2.3: A schematic representation of the neurovascular unit showing cellular elements regulating cerebral blood flow along the vascular tree. **(a)** Different cell types of the neurovascular unit (NVU), including neurons, astrocytes, mural cells — vascular smooth muscle cells (VSMCs) and pericytes — and endothelial cells, regulate cerebral blood flow at different levels of the vascular tree. The cellular composition of the NVU differs along the vascular tree, but the principal cellular components all remain represented, as illustrated here. **(b)** At the level of penetrating arteries, the NVU is composed of endothelial cells making up the inner layer of the vessel wall, covered by a thin extracellular basement membrane and ringed by one to three layers of VSMCs, and ensheathed by pia. The Virchow-Robin space containing the cerebrospinal fluid is between pia and the glia limitans formed by astrocytic endfeet. Both VSMCs and astrocytes are innervated by local neurons. **(c)** Arterioles differ in that there is only one layer of VSMCs, and astrocyte coverage and innervation of the vessel wall and the endothelial inner layer display continuity with penetrating arteries, and brain capillaries, above and below the arteriole level, respectively. In addition to VSMCs, pre-capillary arterioles may also contain transitional pericytes, a cell type between pericytes and VSMCs. **(d)** At the capillary level, the NVU is composed of endothelial cells that share a common basement membrane with pericytes. Pericytes stretch their processes along and around capillaries and make direct interdigitated or "peg-socket"-like contacts with endothelial cells. Pericytes and endothelial cells are covered by astrocyte endfeet. Both astrocytes and pericytes are innervated by local neurons, as shown for astrocytes and VSMCs in the upper segments of the vascular tree. Figure and caption reproduced from [Kis+17a].

2.1.1 Discussion in literature: the role of pericytes in blood flow regulation

The main result reported in [Hal+14] is that capillary pericytes regulate cerebral blood flow in health and disease. It has been critically assessed by [Hil+15], where the authors claimed that regional blood flow in the normal and ischemic brain is controlled by arteriolar smooth muscle cell contractility and not by capillary pericytes. Later on, a commentary [Att+16] was published claiming that the differences between [Hal+14] and [Hil+15] are "apparent rather than real, and depend on how pericytes are defined", and that "an objective definition of different sub-classes of pericyte along the capillary bed is needed to develop novel therapeutic approaches for stroke and disorders caused by pericyte malfunction".

Both claims [Hal+14; Hil+15] were criticized biologically-wise and experimentally-wise [Mar+15; OHe+16; Fil+16]. We contribute to this discussion data analysis-wise, bringing out a few important points to consider when doing the analysis of the response dynamics of vessel diameters (Section 2.3).

The quality of data in [Hal+14; Hil+15] is worth a separate comment here. First, in [Hal+14] (Fig. 2.17) the authors did not show the data after the stimulation when rescaling the diameter traces with respect to their amplitudes. This will make a slower responding vessel artificially faster, if its trace actually peaked even slightly after the end of stimulation, which in our study happened quite often. In Figure 7 from [Hil+15] (Fig. 2.4), the authors show the image that seems to be counter-intuitive to their own claims. It shows virtually no response in a vessel segment of a left branch that does not express the smooth muscle actin. However, one can clearly see a different top branch of the same vessel that also does not express smooth muscle actin, but reacts to stimulation! It suggests that the authors might have analyzed the essentially non-responsive branch of the vessel, no matter whether it expressed smooth muscle actin or not. Finally, the authors of both papers analyzed very local parts of the vessel segments using line selections without discussing whether they accurately represent the response of the selections as a whole (Fig. 2.15).

In summary, the discussion of the issue in literature is still open: The true role for pericytes in the regulation of capillary blood flow still remains controversial and does not seem to be well-established. We bridged the two opposing points of view and further clarified the role of pericytes in Sections 2.5 and 2.6.

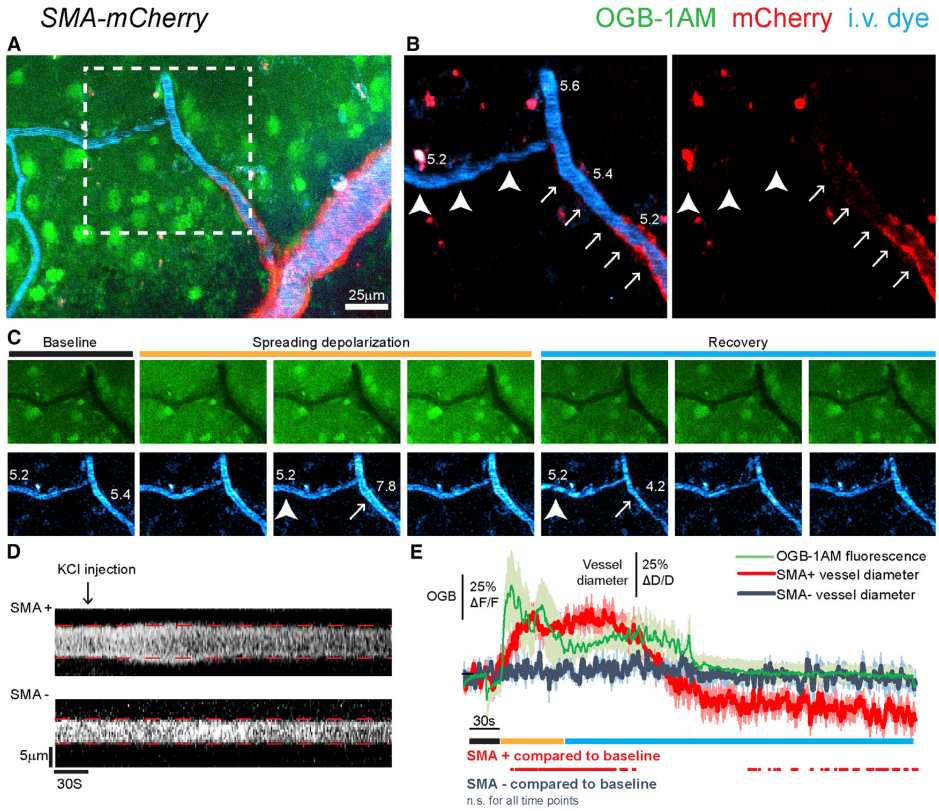


Figure 2.4: Cortical Spreading Depolarization Induces Vasomotion of Smooth Muscle but Not Pericyte-Covered Microvessels. **(a)** In vivo image captured from the cortex of an SMA-mCherry (red) transgenic mouse with neurons and astrocytes labeled with the Oregon Green Bapta-1AM (OGB-1AM) calcium indicator (green) and vessels labeled via intravascular dye (blue). **(b)** Region depicted in **a** of the transition from precapillary SMA-covered vessel (arrows) to capillary pericyte-covered vessel (arrowheads); numbers indicate baseline vessel diameters. **(c)** Time sequence showing the OGB-1AM fluorescence intensity and vessels at baseline, during KCl-induced spreading depolarization, and during the recovery period; numbers indicate vessel diameters. **(d)** Line-scan time series images showing KCl-induced change in vessel diameter in an SMA+ vessel, but not in the downstream SMA- vessel. **(e)** Quantification of the neuronal and vascular responses to spreading depolarization (gold bar) with the average neural OGB-1AM fluorescence signal (green), the average change in diameter of precapillary vessels with a smooth muscle layer (SMA+, red), and capillaries lacking a smooth muscle layer (SMA, blue). Traces show mean values ± s.e.m., n = 10 SMA+ vessels, and 10 SMA- vessels from 3 mice. Asterisk indicates significance compared to baseline determined by a 99% confidence interval. Figure and caption reproduced from [Hil+15].

2.1.2 Pericyte coverage, structure, and distribution

Generally, the coverage of pericytes depends on the animal and the organ [Sim86]. For cerebral gray matter of mice, the reported mean pericyte index was around 5% [CS85]. Also, the authors found that out of all profiles examined in the study 77% included pericyte processes, only 10% of all profiles also had a portion of the pericyte nucleus visible. Lysosome-like granules, although only found in 12.5% of all profiles with pericytes, were almost always seen in association with the pericyte nucleus (83% of the time).

A number of recent studies also reported on the issue. Figure 1 in [M+06] (Fig. 2.11) clearly demonstrates well-resolved pericyte somatas with outgoing processes that might cover the entire capillary wall in the rat retina. The central nervous system (CNS) vasculature is generally regarded as being the most pericyte covered, with a 1:1–3:1 ratio between endothelial cells and pericytes, and an approximately 30% coverage of the abluminal surface [Mat+10; Sim86]. A conservative estimate of the pericyte coverage of the endothelium abluminal surface in normal tissues ranges between 70% and 10% [AGC11].

In addition to coverage, it is important to discuss the labeling efficiency of various dyes and receptors to identify pericytes in the cerebral cortex. A relatively low number of studies reported on the matter, with the most comprehensive work being [A+15]. It was reported in [A+15] that the combination of the platelet-derived growth factor receptor β (PDGFRβ) as a pericyte marker and the tdTomato dye shows $\simeq$99% sensitivity to pericytes. For the combination of the neural/glial antigen 2 (NG2) and the tdTomato dye the sensitivity varied from 0% to 86%, and this might be similar to the dye DsRed used in [Hal+14] and our current setting.

The authors of [A+15] also thoroughly studied pericyte structure and distribution in the cerebral cortex of transgenic mice were with the help of high-resolution imaging. They reported structural differences between pericytes adjacent to arterioles versus the ones broadly distributed in the capillary bed that may underlie differential roles in control of vessel caliber. Their summary of the findings on mural cell organization in the cerebrovasculature is reproduced in Fig. 2.5.

A recent review [Kis+17a] discusses similar issues in the context of the idea of control of capillary blood flow by pericytes. Though the prevailing view now is that pericytes contribute to the regulation of capillary diameter and CBF, it has been seriously challenged by several recent studies. The authors conclude that although a growing body of evidence supports the role of pericytes in CBF regulation, there is also a growing need to better define pericytes, particularly at the molecular level. The summary of their discussion on pericyte definition and the question of contractility is presented in Fig. 2.6.

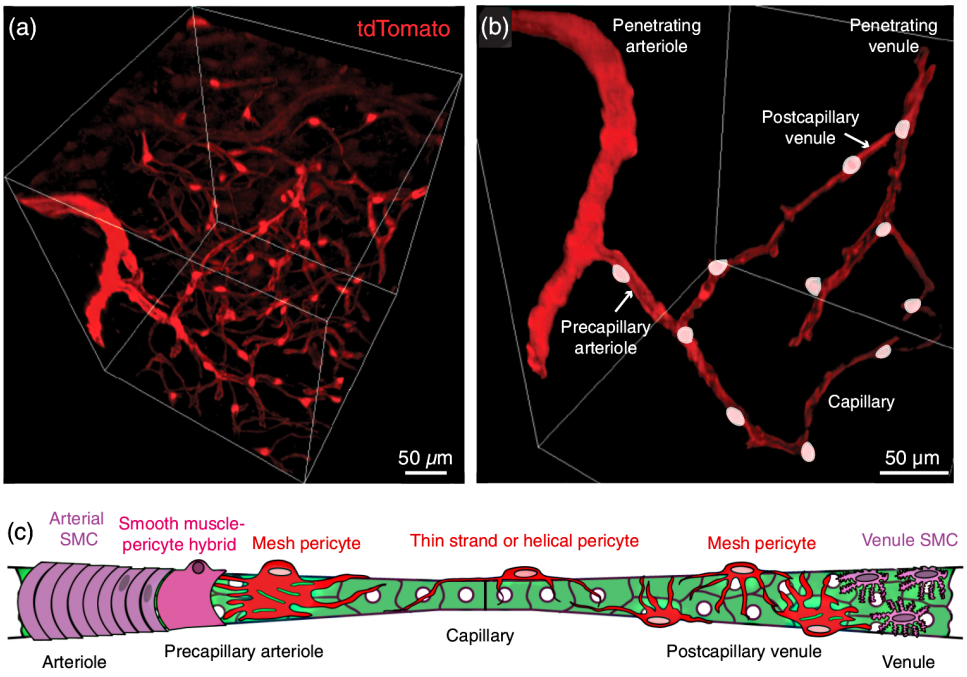


Figure 2.5: Summary of mural cell organization in the cerebrovasculature: **(a)** A tissue volume from a PDGFRβ-tdTomato mouse imaged after optical clearing. **(b)** Distribution of pericytes and pericyte-smooth muscle hybrids in an arteriole-capillary-venule loop. The cell bodies are labeled in white. **(c)** Schematic showing the continuum of mural cell types along the cerebral vasculature. Smooth muscle cells form concentric rings on arterioles. Hybrid smooth muscle-pericyte cells reside on precapillary arterioles and interlock with mesh pericytes at the arteriole-capillary interface, which occurs as penetrating arterioles ramify into the capillary bed. Pericytes in capillary beds typically exhibit long processes that traverse the microvasculature in single strands or pairs that twist in a helical fashion. Mesh pericytes become more prevalent again as capillaries turn into postcapillary venules. Stellate-shaped smooth muscle cells cover the walls of parenchymal venules. Figure and caption reproduced from [A+15].

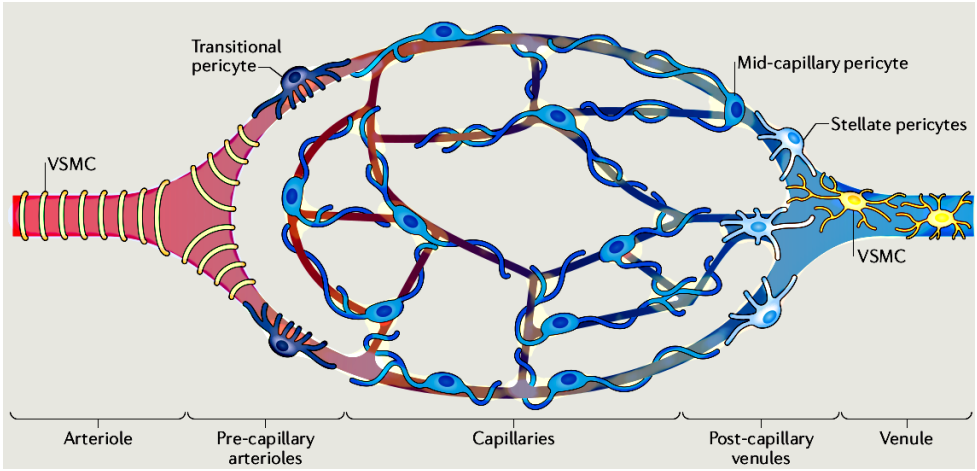


Figure 2.6: Pericyte definition and the question of contractility Zimmermann suggested that pericytes represent a heterogeneous cell population with several subtypes that differ by morphology, location within the vascular tree and function. A recent morphological study suggested that pericyte subtypes include "mid-capillary" pericytes in the vast majority of the capillary bed, "transitional" pericytes close to vascular smooth muscle cells (VSMCs) and "stellate" pericytes on post-capillary venules, and that all differ from arteriolar ringed VSMCs. Whether different pericyte subtypes have different functions, such as, for example, regulation of blood–brain barrier (BBB) permeability versus control of blood flow, remains to be determined. Transitional pericytes on pre-capillaries and post-capillaries express α-smooth muscle actin (α-SMA), a contractile protein typically found in VSMCs. α-SMA was also found in mid-capillary pericytes in the brain and retina by some studies but not by others. In addition, mid-capillary pericytes express the contractile proteins myosin and vimentin. A recent single-cell RNA sequencing study has shown that mouse brain pericytes express several contractile proteins, including skeletal muscle actin, vimentin, desmin, calponin and non-muscle myosin variants, and low α-SMA levels. Some studies failed to show pericyte contractility in vivo, but multiple independent studies reported that pericytes can dilate or contract capillaries in response to different stimuli and / or regulate cerebral blood flow in vivo. When it comes to a definition of a pericyte, several outstanding questions still remain: can a pericyte be molecularly defined by a single-cell RNA sequencing of a genome-deep quantitative transcriptional profiling? Do different pericyte subtypes exist in the same organ, and can they be molecularly distinguished from each other in the brain as suggested by morphological studies? Do pericytes have organotypic features, as do endothelial cells? And are pericytes in the brain different from pericytes in peripheral organs? Figure and caption reproduced from [Kis+17a].

2.2 Imaging in the depth of mouse brain cortex

2.2.1 Two-photon microscopy

The majority of biological tissues scatter light strongly. This poses a challenge for traditional fluorescence microscopy techniques, that use linear (one-photon) absorption processes for contrast generation — at depths larger than 100 microns strong and multiple light scattering blurs the images. Multiple-photon excited fluorescence microscopy techniques, extensively developed over the past decades, overcome this challenge using the higher-order nonlinear light-matter interactions involving multiple photons. Therefore, the technique allows imaging down to several hundred microns deep in various organs of living animals [HD05].

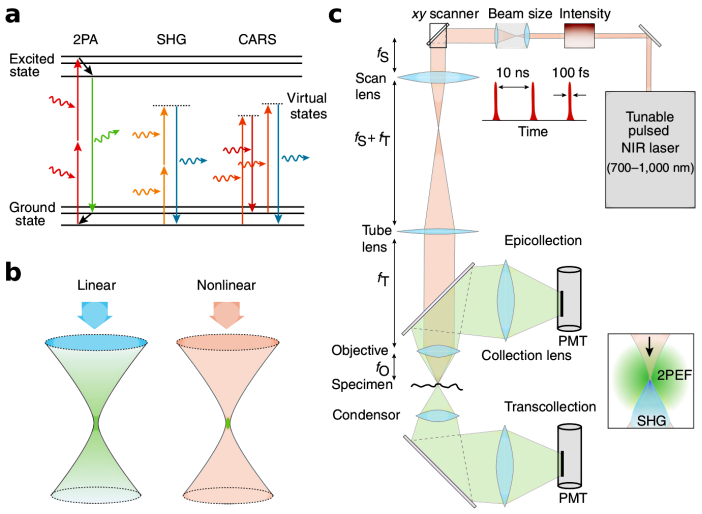


Figure 2.7: Nonlinear optical microscopy. **(a)** Jablonski diagram, illustrating two-photon absorption (2PA), second-harmonic generation (SHG) and coherent anti-Stokes Raman scattering (CARS). **(b)** Spatial confinement of signal generation with nonlinear excitation. Visible ("blue-ish") light is used for excitation in single-photon microscopy, whereas near-infrared ("red-ish") light is used in two photon-excited fluorescence lase-scanning microscopy. In single-photon microscopy an entire cone of fluorescence light (green) is generated, whereas nonlinear signal production is localized to the vicinity of the focal spot. **(c)** Generic nonlinear laser-scanning microscope. A laser source provides near-infrared ultrashort pulses; intensity and beam size are adjusted before coupling the laser beam to the microscope. The focal lengths of the scan lens (f_S), the tube lens (f_T) and the objective (f_O) are indicated. Two-photon excited fluorescence, which is isotropically emitted (inset), can be collected in epi- and / or trans-collection mode, using whole-area detection by photomultiplier tubes. Figure and caption reproduced from [HD05].

In comparison to traditional techniques, all nonlinear microscopy techniques require expensive pulsed laser systems to achieve sufficient excitation rates. However, this is worth because of the major advantages it gives. The first one is that generated photons have higher energies than the excitation light making emission "bluer" than the excitation. The second advantage is that when focusing the laser beam through a microscope objective, multiphoton absorption is spatially confined to the perifocal region (Fig. 2.7). Because the density of scattered excitation photons is generally too low to generate significant signal, the localization of excitation is maintained even in strongly scattering tissue. It means that nonlinear microscopy is far less sensitive to light scattering than traditional microscopy (Fig. 2.8). This is of vital importance for deep imaging, because it means that all fluorescence photons are known to originate from near the focus and thus are providing the useful signal.

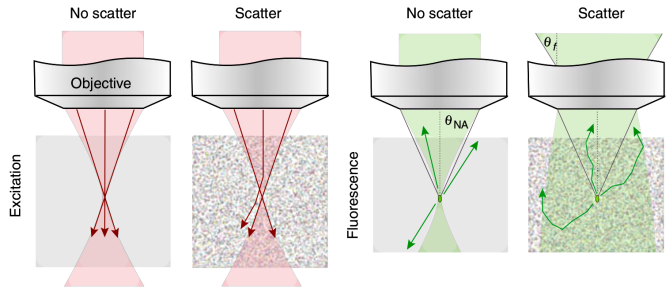


Figure 2.8: Signal generation and fluorescence collection in clear tissue (no scatter) and in scattering tissue (scatter). In clear tissue all excitation light reaches the focus, but in scattering tissue, scattering (even by a small angle) causes light rays to miss the focus and be lost to signal generation. This leads to a roughly exponential decrease in excitation with depth. In clear tissue only fluorescence light rays initially emitted into the collection cone can be detected, but in scattering tissue fluorescence light is (multiply) scattered and may even "turn around". Fluorescence light apparently originates from a large field of view but a larger fraction than in the nonscattering case is actually within the angular acceptance range Θ_f of the objective. Figure and caption reproduced from [HD05].

2.2.2 Timeline of the experiment and mouse groups

For each anesthetized (Appendix B.1.1) mouse in the study, vessel responses to stimulation (Appendix B.1.2) were recorded with time resolution using two-photon microscopy (Appendix B.1.3). Each recording of a stimulation was 400 frames long, with a sampling frequency 6.54 Hz ($\simeq$0.153 s per frame). Several fields of view, possibly containing overlapping vessel segments (see e.g. Fig. 2.12), were recorded sequentially. Each such field of view was recorded in several experimental phases. The timeline for each experiment was typically the following: a **Control phase**, where all fields of

view are imaged without treatment; an **ANGII phase** where the mouse is treated with ANGII (Section 2.6) and all fields of view are re-imaged; and a **Rinsed phase**, where ANGII treatment is rinsed out and all fields of view are re-imaged. We have obtained seven data sets, as summarized in Table 2.9. The data sets are characterized by two parameters: the presence/absence of gel on the brain and the concentration of applied ANGII ranging from zero to toxic doses. A toxic dose is expected to cause systemic effects, from which the direct effect of ANGII on the local vessels is hard to isolate. The effect of ANGII might be delayed by the presence of gel on top of the brain, because of slow diffusion through it. Timelines of the experiment for individual groups of mice from Table 2.9 are visualized in Fig. 2.10.

Group	Description	Mouse IDs
1. Time control w gel	No ANGII	65, 67, 69, 70
2. ANGII w gel	10 uM of ANGII	304, 307, 309, 312, 315, 316
3. ANGII w/o gel	10 uM of ANGII, no gel	57, 285, 286, 288, 298
4. ANGII from start	10 uM of ANGII	80, 83, 185, 188, 190, 191
5. High ANGII w/o gel #1	100 uM of ANGII (low-responsive)	119, 126, 133, 135
6. Toxic ANGII w/o gel	1-10 mM of ANGII (systemic effect)	127, 129, 130
7. High ANGII w/o gel #2	100 uM of ANGII (low-responsive)	175, 177, 178, 180

Table 2.9: Overview of the available data sets. Unfortunately, the mice in Groups 5 and 7 did not respond well to stimulation in the control phase of the experiment, so we did not use them for the studies of response dynamics.

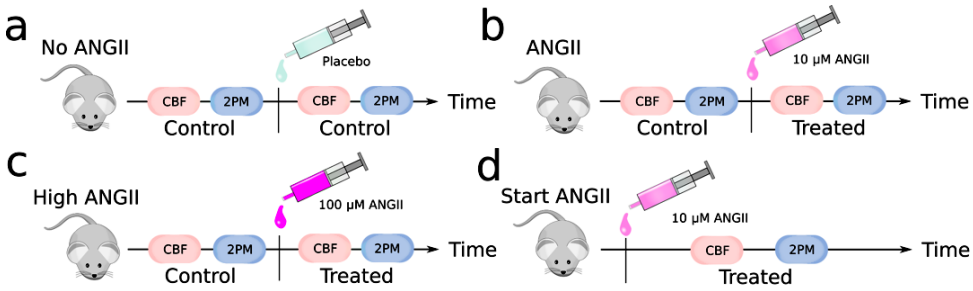


Figure 2.10: **(a)** A group of live, anaesthetized mice were submitted to laser Doppler flowmetry measurements of cerebral blood flow (CBF, Appendix B.1.4) and two-photon microscopy (2PM) of the whisker-barrel cortex through a small craniotomy. A non-active solution (placebo) was superfused topically on the cortex via the craniotomy followed by a repetition of both CBF and 2PM measurements, to control for temporal effects. **(b)** Same as a but 10μM angiotensin II (ANGII) was superfused topically. **(c)** Same as **b** but 100 μM ANGII was applied. **(d)** Same as **b** but ANGII was applied before data collection.

2.2.3 Pericyte classification

Figure 2.2 gives an artistic representation of pericyte anatomy. However, for image analysis it is mostly important to classify a chosen vessel region in terms of absence or presence of pericytes. This classification was adopted from [M+06] (Fig. 2.11).

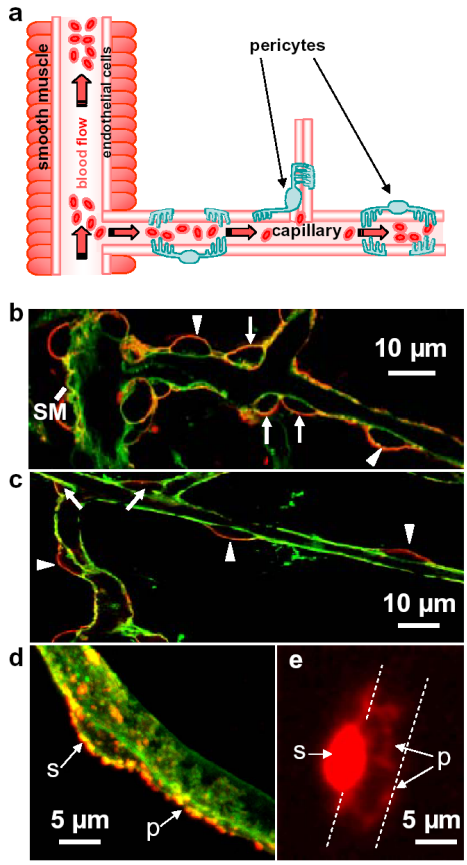


Figure 2.11: Pericyte anatomy confers flow regulating capability downstream of arterioles. **(a)** Potential blood flow control sites in cerebral vasculature: arteriolar smooth muscle, and pericytes on capillaries. **(b)** Cerebellar molecular layer arteriole (left), surrounded by smooth muscle, giving off a capillary. Capillary labelled with isolectin B4 (green); pericytes labelled for NG2 (red) are on the straight part of capillaries (arrow heads) and at junctions (arrows). **(c)** Retinal capillaries. **(d)** Soma (s) of cerebellar pericyte gives off processes (p) running along/ around capillary. **(e)** Dye fill of retinal pericyte reveals processes running around capillary (dashed lines). Figure reproduced from [M+06].

2.3 Image analysis

2.3.1 Raw experimental data

The raw experimental data consisted of stacks of images (Appendix B.1.3) that visualized the response to stimulation (Appendix B.1.2) of various vessel segments (Fig. 2.11) at different times during the experiment (Section 2.2.2). Each vessel segment was classified using its branching order, biological type, and position from the root of the network branch (Fig. 2.12).

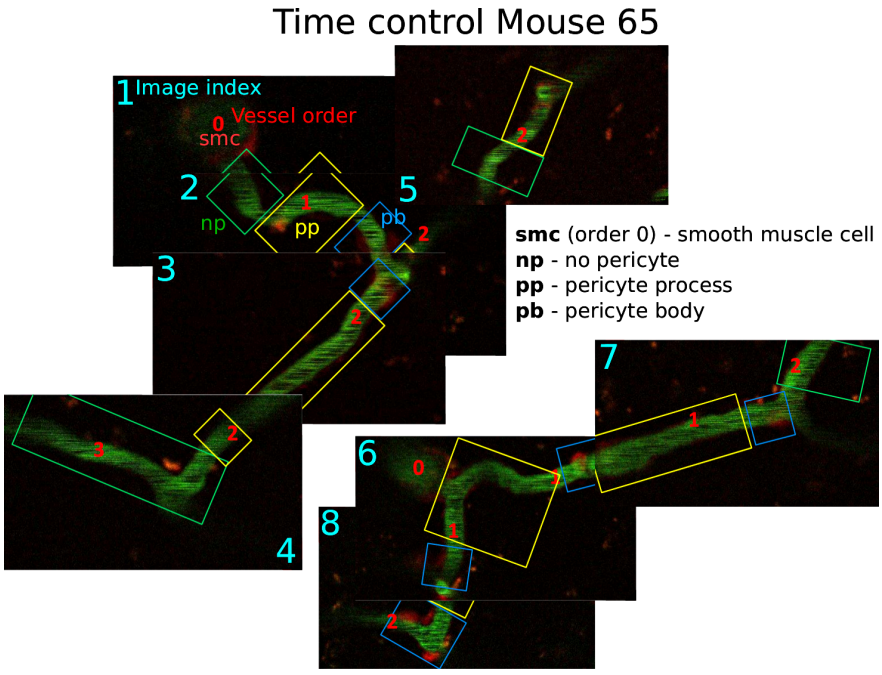


Figure 2.12: Raw images of the fields of view recorded from a single mouse (Mouse 65 from the time control group). Cyan indices define the sequence of their imaging in the control phase of the experiment. Note how the images partly overlap (for example 1 and 2, 6 and 8) to trace out branches of two vessel networks. Red indices indicate the branching order of each vessel. The penetrating arterioles (also called smooth muscle cells, Fig. 2.11) are considered as the roots of the imaged branching networks and therefore have vessel order 0. Colored selections correspond to the classification of the regions of interest based on presence of pericyte cell body or somata (pb), presence of pericyte process (pp), and absence of the pericytes (np), as in Fig. 2.11.

2.3.2 Diameter extraction from raw data

The extraction of diameter traces from raw two-photon microscopy images is nowadays typically performed using custom scripts of varying complexity [Hal+14; KN14]. They extract the borders of the vessel lumen from the so-called linescan images employing the Chan-Vese active contours algorithm [CV01]. The main disadvantage of this approach is that regions of interest of vessels are often much larger than a single line. Therefore, an accurate trace of the averaged vessel diameter of a real biological vessel region would require averaging over many lines, unless its very local behavior is of primary interest for the study. The specific question of how many lines are required for averaging to accurately represent the response dynamics of a large vessel region was never brought up in the literature, so it is considered in Fig. 2.15.

The second major disadvantage is the necessity to implement and fine-tune the many parameters of the Chan-Vese algorithm or especially its custom modification. Fig. 2.15 illustrates that almost identical results could be obtained using much less parameters (a thresholding scheme and a radius of the Gaussian blur spatial filter) and working by default with areas instead of lines, i.e. to effectively average over many lines in a single measurement.

2.3.2.1 Areascan method

The schematic problem of diameter extraction is presented in Fig. 2.13.

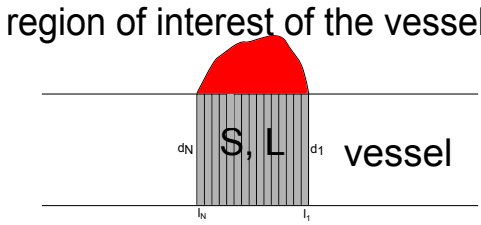


Figure 2.13: There is a region of interest of the vessel (gray) of total area S and total length L, found e.g. with the help of the pericyte marker (red), that has to be characterized in terms of its averaged diameter. The region is a complex shape, but it can be splitted into N elementary rectangles ($[l_1, d_1]$, $[\dots]$, $[l_N, d_N]$) of arbitrary small lengths. Averaging over the diameters of elemetal rectangles gives the averaged diameter of the region of interest as a whole.

Mathematical interpretation of Fig. 2.13 is that if either the actual diameter of each of the N elementary rectangles d_i or its area S_i and length l_i are known, then the averaged diameter trace d of the region of interest is calculated as

$$d = \langle d_i \rangle = \left\langle \frac{S_i}{l_i} \right\rangle \qquad i = 1 : N \ , \tag{2.1}$$

Linescan method follows the first strategy given by Eq. 2.1 — it extracts the diameters of several lines across the region of interest of the vessel. Areascan method follows the second strategy — it extracts the areas of several elementary rectangles of known lengths.

Now, let us split the region of interest of length L into $N \to \infty$ elementary rectangles of the same lengths $dl = \frac{L}{N} \to 0$. Then, Eq. 2.1 can be re-written as

$$d = \left\langle \frac{S_i}{l_i} \right\rangle = \frac{\langle S_i \rangle}{dl} = \frac{\sum_{i=1}^{i=N} S_i}{dl \cdot N} = \frac{S}{L}, \qquad i = 1 : N, \tag{2.2}$$

Thus, there are essentially only two values required to measure the averaged diameter trace of the region of interest of the vessel — its total length L and its total area S. Therefore, the areascan method offers a huge reduction in workload in comparison to the linescan method, because the latter has to extract the diameters of many individual lines and average over them afterwards.

In practice, the areascan method is to be applied to image stacks to extract the diameter traces. Real experimental images suffer from several difficulties the method also has to be robust to: the diffusion of the dye that illuminates the vessel through its volume, small "jumps" of the images during or before stimulus, and the passing red blood cells that sometimes block the blood flow.

The illustration of the method in practice is given in Fig. 2.14. High-frequency image noise is removed by applying a low-pass Gaussian blur filter (Fig. 2.14b). Next, thresholding is applied to separate the inner boundaries of the vessel from the background (Fig. 2.14c). Finally, the area (the number of pixels) of the vessel within the selection is calculated from the thresholded image, and divided over the length of the selection to get the value of the diameter. The whole process is repeated for the image stack to get the diameter trace. Figure 2.14d shows the extracted traces from the three consecutive recordings of the same vessel. This is the null experiment: the responses were proofed to be reproducible under constant experimental conditions.

2.3.2.2 Areascan method compared to linescan

Figure 2.15 presents the validation of the areascan method against the linescan method. It shows that areascan reproduces the results of the linescan (Fig. 2.15b), and illustrates the danger of averaging over too low number of lines with the linescan method (Fig. 2.15c).

2.3.3 Results and discussion

In this section, we developed a simple, fast, and noise robust areascan method for the extraction of a diameter time trace of a region of a vessel. We showed that it reproduces the results of the conventionally used linescan method for a narrow line-like area, and performs much more accurate on regions that require a lot of linescan measurements of the diameter trace (Fig. 2.15).

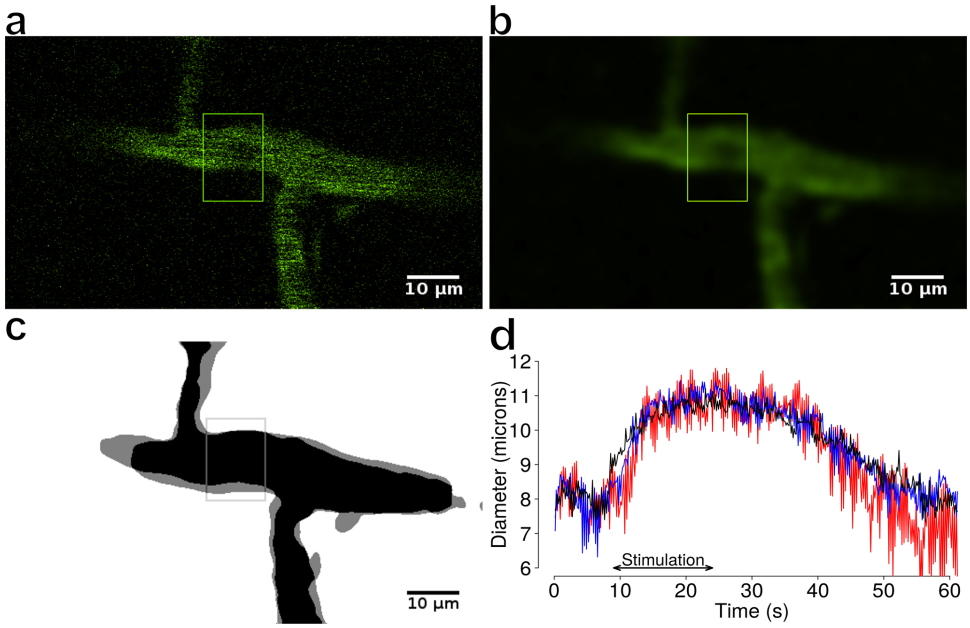


Figure 2.14: **(a)** A two-photon image at ≃150 microns depth of a vessel of the barrel cortex in the brain of a NG2-DsRed mouse. The vessel lumen is labeled with FITC-dextran (green). The selected region of interest of the vessel is given by the green rectangle. **(b)** Gaussian blur low-pass filter applied to the image from a to reduce the high-frequency noise to easily distinguish the inner boundaries of the vessel from the dark background. This keeps the original temporal resolution of the image stack, as it is of critical importance for the careful analysis of response dynamics. **(c)** Thresholding applied to image from **b**. The thresholded image represents the inner boundaries of the vessel. The black vessel was obtained at 0 s, the grey one — at 20 s. This is an overlay of the vessel in rest (black) with the vessel in dilation (grey). For each image of the stack, the area of the selected region of interest was calculated as the number of pixels that fall within the selection. Dividing by the length of the selection along the vessel, we obtain the value of the averaged diameter within it. **(d)** Time courses of the extracted diameter from **c** showing how the mouse responded to stimulation. The three lines represent the three consecutive recordings of the same vessel from **a** to illustrate the reproduction of the responses. The order of the recordings: black, red, blue. Sometimes the extracted traces do not show a clear response. This is why we introduced an empirical criterion to automatically detect such non-responsive diameter traces (Appendix B.2).

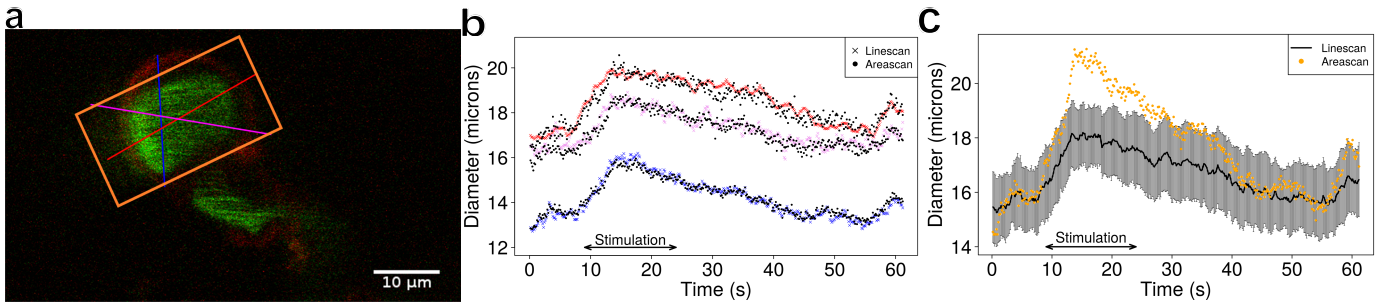


Figure 2.15: Areascan method reproduces the results of the linescan method, but saves the necessity to average over multiple lines to accurately capture the dynamics of large vessels as a whole. **(a)** A two-photon image at ≃150 microns depth of a vessel of the barrel cortex in the brain of a NG2-DsRed mouse that was not included in the study. The vessel lumen is labeled with FITC-dextran (green). NG2 promoter and DsRed dye are used to label pericyte somatas and processes (red). Colored lines represent the selections for linescan method [Hal+14] for the imaged arteriole. Small rectangles (up to 5 pixels wide) around the same lines were used as selections for areascan method. The orange rectangle is a specific selection for areascan method that covers the arteriole as a whole rather than its small local regions. **(b)** Extracted diameters for linescan selections from a obtained with the linescan method (colored crosses) and the areascan method (filled black circles). The results of the two methods agree very well with each other. **(c)** Linescan method extracted diameters from a averaged (mean ± s.e.m., solid black lines) and compared to the extracted diameter from the orange selection from a. Note the major differences between the two methods as the orange trace reaches its peak diameter. Note also that the peak diameter falls off long before the stimulation ends — such a response is unclear from the biological point of view, and this is why this particular mouse and alike mice were removed from the studies of dynamics.

2.4 Arterioles and first order capillaries show identical dynamics

Initially, we focused on the control experimental phase (Section 2.2.2) and consider responses from mice studied in experiments with gel (Groups 1 and 2 from Table 2.9). These groups were selected, because the presence of the gel on top of the mouse brain in these experiments improved the stability of images. From those groups we selected mice with ID numbers 65, 67, 69, 70, 304, 307, 309, 312 and 315. Then we selected fields of view in which *penetrating arterioles (order 0) and first order capillaries (order 1) were imaged simultaneously* (e.g. as in Fig. 2.12, image index 1). For all time-lapse movies satisfying this constraint, we averaged the responsive time traces normalized to their resting diameter (Appendix B.2) from vessels of order 0, respectively, order 1. Comparison of such traces demonstrates a striking data collapse in the entire period of the 15 second stimulation, indicating that *penetrating arterioles and first order capillaries have identical averaged response dynamics* (Fig. 2.16)!

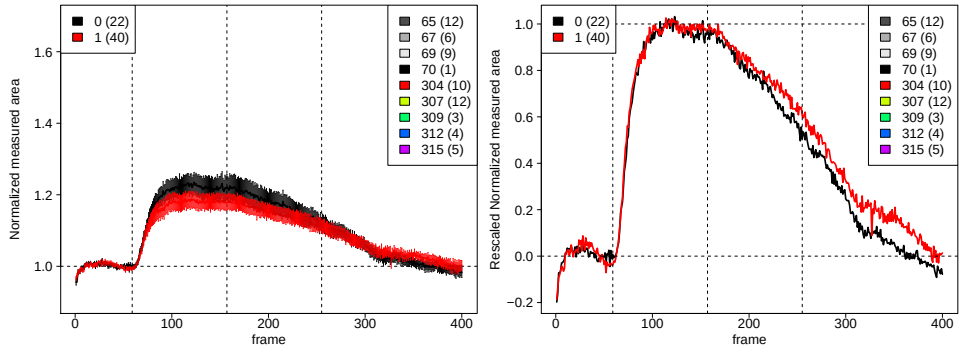


Figure 2.16: *Left*: Averaged responsive (Appendix B.2) time traces normalized to their resting diameter from simultaneously imaged vessels of order 0 (black) and 1 (red), respectively. The numbers of contributing traces for each order are indicated in the top left legend. Note that, in some cases, we included multiple segments of first-order vessels in our study for each zeroth-order vessel included, which explains why their numbers are not identical. Error bars indicate s.e.m. The top right legend shows the number of trace contributions from each mouse. Note that the coloring is not used here but present for Fig. 2.16 to be consistent with later plots. *Right*: Same as *Left*, but each averaged trace was rescaled with its maximum value with its resting area (averaged vessel area preceding stimulation) subtracted. The data collapse demonstrates that the responses, on average, are identical! Error bars are omitted to avoid clutter. Same plot but with splits according to biological classification is shown in Fig. B.2.

2.4.1 Dynamics challenge current views

The striking agreement between averaged dynamics of penetrating arterioles and first-order capillaries (Fig. 2.16) is in conflict with the data reported in [Hal+14]. Figure 3 of that reference is reproduced below in Figure 2.17. Panel e shows averaged responses normalized to their maximum values for vessels of, respectively, order 0 and order 1 (color coding is identical to that in Figure 2.16a). Note that the time traces are shown only for the first 15 seconds after the onset of the stimulation. Those data suggest that there is a delay of ∼1.5 seconds of the zeroth order vessels compared to the first order vessels (Fig. 2.17f). Based on this, it is argued that pericytes play an active role in the regulation of the blood-flow dynamics.

Notice also the "plateau" in the averaged response of the first order capillaries, which roughly coincides with the initiation of the response for the zeroth order penetrating arterioles.

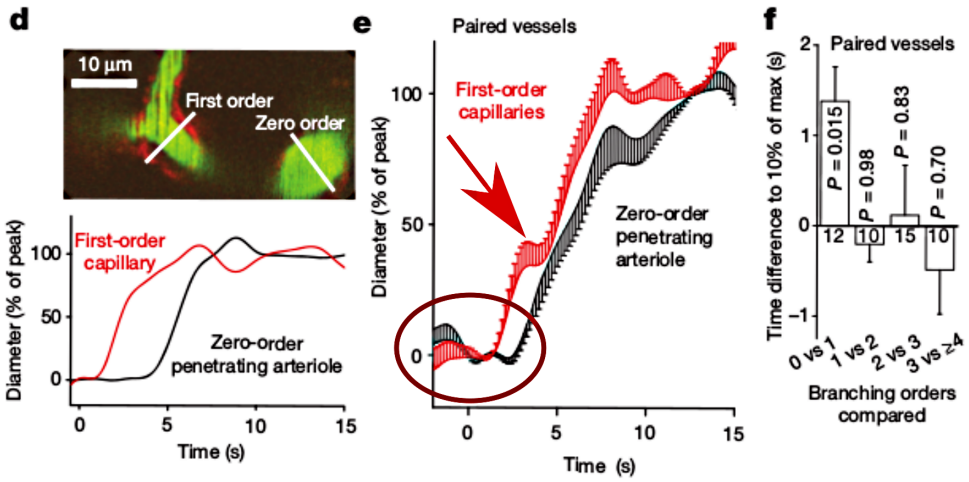


Figure 2.17: Reproduction of Figure 3 from [Hal+14]. A significant delay in the responses of orders 0 compared to orders 1+ (greater than or equal 1) is reported in panels d-f. The circle in panel e indicates the clear difference in the timing of averaged responses between penetrating arterioles and first order capillaries. The arrow in panel e indicates the plateau during the early part of the response to stimulation for first order capillaries. Notice the absence of the plateau in the response for zeroth order penetrating arterioles.

2.4.2 Possible explanation of differences

Figure 2.17d shows an example of a response from a pair of simultaneously imaged zeroth order penetrating arteriole and first order capillary from [Hal+14]. In this case, one sees a dramatic delay of ∼3 seconds. However, we re-analyzed this particular time-lapse movie. The results are shown in Figure 2.18. Clearly, the response of the first order vessel seems ahead of the zeroth order.

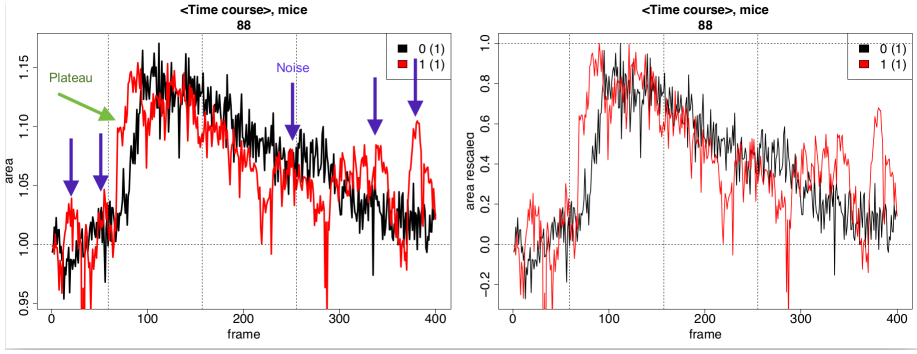


Figure 2.18: *Left*: Responses of a pair of first order capillary (red) and zeroth order penetrating arteriole (black) from Mouse 88 from [Hal+14]. A delay of ∼3 seconds (=20 frames) of the response of zero-order vessel compared to the first-order vessel was reported (Fig. 2.17d, in which data were smoothed and only shown in a 15 second window after the stimulation). Large amplitude fluctuations are seen in the late part of the response from the first order vessel (purple arrows labeled "noise"). Furthermore, a plateau is seen on the initial rise after the stimulation (green arrow). *Right*: Same as *Left* but each response was rescaled with respect to its peak amplitude and the resting area was subtracted.

In the very last part of the response curves, the stimulation is well over and an approach back to the baseline diameter is expected. This is exactly the case for the trace of the zeroth order penetrating arteriole. For the first order capillary, however, there are fluctuations that are too prolonged to be caused by measurement noise. The amplitude of these fluctuations (up to 10% over the baseline) is comparable to the peak amplitude of the trace (15% over the baseline)! Also, the fluctuation amplitude is identical to the level of the plateau seen on the initial rise after stimulation. Fundamentally, when a fluctuation of such large amplitude can happen at random, one should hesitate to trust any part of the signal in the time trace. For additional examples, see Figs. B.3 and B.4.

We do not observe such fluctuations in any of the data contributing to Figure 2.16. Therefore this partly explains the discrepancy between the data from [Hal+14] and our results, which neither show such fluctuations nor a delay of the zero order vessels.

2.4.3 Timing of paired responses in zeroth and first orders is identical

The analysis on averaged responses (Fig. 2.16) conceals potential variations, e.g., due to mouse-to-mouse differences. Here, we extend the analysis to individual vessel responses and use that to substantiate and expand the above results on timing.

To quantify the relationship between distance along the branching tree and timing of responses, we used (i) the "time-to-peak" and (ii) the "time-to-20%-of-peak" (Appendix B.4). The latter quantity captures the initial onset of the response, while still being relatively insensitive to fluctuations (very similar to the 10%-criterion used previously in [Hal+14]). These two measures of timing are investigated as a function of vessel order, individual mouse number, and phase of the experiment below.

Throughout in the following analysis, we averaged over possible repetitions and/or classifications (*pb*, *pp*, or *np*) of a vessel segment, prompted by the absence of timing effects with respect to these variables (see Fig. B.2).

In Fig. 2.16, we established that there is no difference between penetrating arterioles (Order 0) and first order capillaries (Order 1) in terms of their averaged response dynamics. Analysis on the single-vessel level confirms this (Fig. 2.19). For each vessel branch in each mouse, we averaged the time-to-peak, respectively time-to-20%-of-the-peak, over all repetitions of simultaneously imaged vessels of Order 0 and Order 1. These times are plotted as a function of the corresponding vessel order (Fig. 2.19).

To interpret the data, one must construct a linear model (LM) that parameterizes the possible effects in the data. Inspection of Figure 2.19 suggests that there are obvious mouse-to-mouse differences in response speed. Furthermore, the response timing might depend on the vessel order. Information regarding the significance of these two factors may initially be examined using, so-called, two-way analysis of variance (ANOVA) that examines the influence of two independent variables (here mouse individuality and vessel order) on one continuous dependent variable (time-to-peak or time-to-20%-of-the-peak). The two-way ANOVA assesses the main effect of each independent variable but also if there is any interaction between them.

We used the two-way ANOVA to assess if two factors — vessel order (Vo) and individual mouse number (M) are statistically significant to account for the variation in the data. Vo accounts for any differences between the orders we have (0, 1, and 2), while M accounts for the possible individuality of mice (for example if one particular mouse responds faster than the rest, like mouse 309).

In this case, the output of the analysis (top right inset of the plots) shows that the effect on time-to-peak by *vessel order is insignificant* ($p \approx 0.2$), while there is a clearly significant effect ($p \approx 2 \cdot 10^{-5}$) of mouse-to-mouse differences, as anticipated (see top right insert in Fig. 2.19 for details). For the time-to-20%-of-the-peak, both factors are insignificant, the analysis shows. Note that this agrees with the data collapse demonstrated in Fig. 2.16.

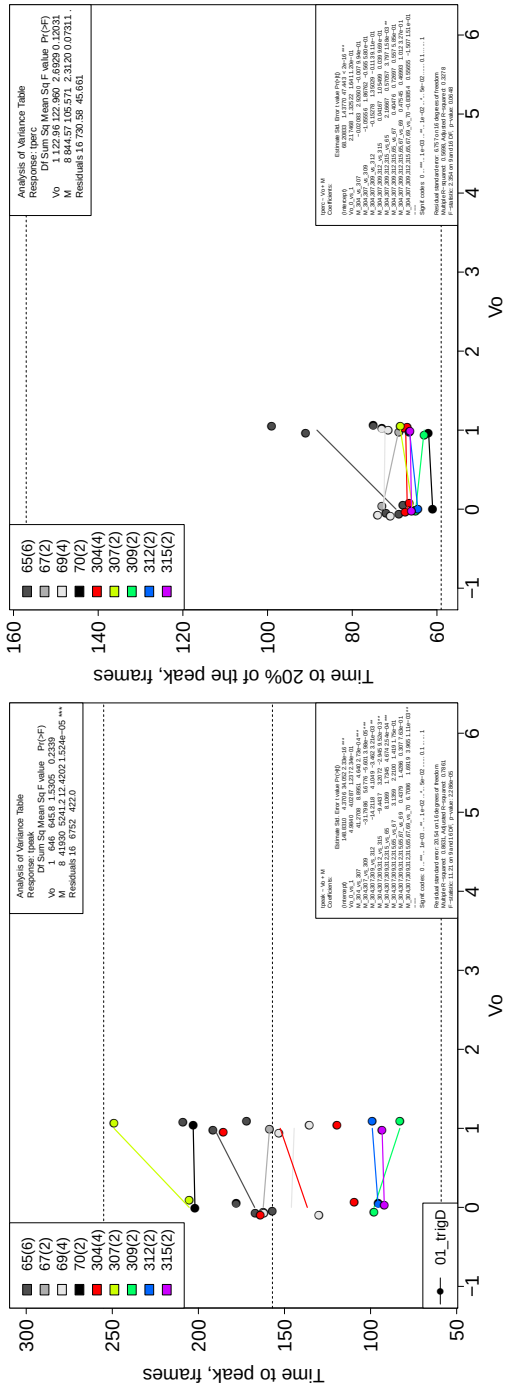


Figure 2.19: *Left*: Time-to-peak of all vessel responses imaged simultaneously of nine mice plotted as a function of vessel order (colored dots as given by the legend with numbers in parentheses representing the total number of data points from an individual mouse). For each branch and for each vessel order, we averaged the time-to-peak over the repetitions and the classifications (*pb*, *pp*, or *np*). For each mouse, lines connect the average time-to-peak value in consecutive orders. Time is measured in units of frames. Vessel order (Vo) is considered a non-significant factor by the analysis, which agrees with the results of Fig. 2.16. Top right inset shows the results of the analysis of variance. Bottom right inset shows the results of fitting a linear model to the data. *Right*: Same as *Left* but for time-to-20%-of-the-peak.

2.4.4 Response dynamics of first-order vessels

Initially, to investigate the effect of distance along the branching tree on the timing of responses, we classified fields of view containing first order vessels in terms of the other vessels imaged simultaneously with them. For example, if a first order vessel is imaged only with a zero order vessel, it is likely to represent an early segment of the vessel. On the other hand, if it is imaged only with second (and potentially higher order) vessels, it is likely to represent a late segment. Finally, some vessels, of course, may be imaged with both zeroth and second order vessels. Figure 2.20 shows that the *onset* of first order vessel segment responses does not depend on their qualitative distance along the branching tree, but that their *time to peak* does. This highlights the need for a discrimination between the onset of a response (which reflects the time of arrival of a cellular signal) and the time to peak (which reflects also the inherent response dynamics of the contractile cell-vessel complex).

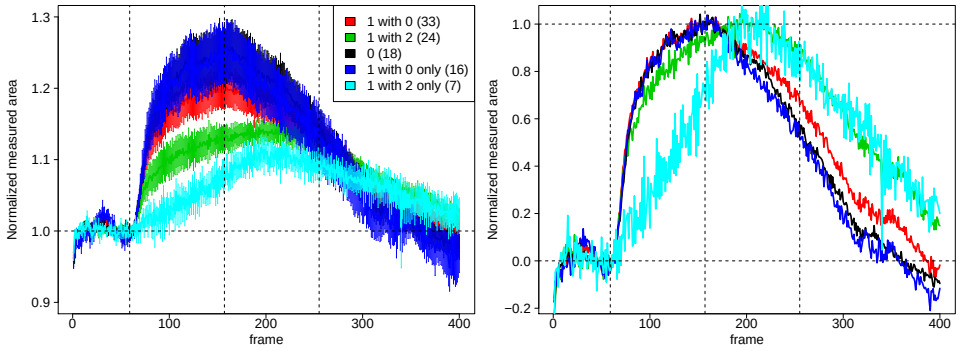


Figure 2.20: *Left*: Averaged responses of vessels of order 0 imaged simultaneously with order 1 serves as a reference (black). Averaged response of order 1 imaged together with order 0 (and potentially also with some higher orders) is shown in red. Also, we averaged the responses of vessels of order 1 that were imaged with vessels of order 0 *only* (blue). Similarly, we averaged over the first-order vessel responses that were imaged with vessels of order 2 only (cyan). The idea is that first-order vessels that appear in images with second-order vessels only are positioned later in the branching network as compared to first-order vessels that appear in images with vessels of order 0. *Right*: Same as *Left* but each curve was rescaled by its peak amplitude. Notice that the cyan curve peaks significantly later than the blue. Comparison of the blue curve with the cyan curve suggests that the early segments peak earlier than the late segments, however without an actual delay in the onset of the response.

2.4.5 Responses become slower as function of vessel order

To reconcile this type of analysis with the finding that not all first-order capillaries respond equally fast, we consider now all responses to and including order 3. We no longer require the vessels to be simultaneously imaged. We include all such responses from the combined Mouse Groups 1 (Time control) and 2 (ANGII w. gel). Initially, we restrict the analysis to the first phase of the experiment in which the two groups are identical. We repeat the same analysis as in Fig. 2.19 for these data and present it in Fig. 2.21.

We find that vessel order indeed affects the time-to-peak, but the effect on the initial dynamics, parametrized as the time-to-20%-of-the-peak, is less clear (Fig. 2.21).

ANOVA analysis only assesses if a factor has a statistically significant effect on a response, but it does not reveal anything about the relationship between the individual levels of that factor, e.g. the vessel orders. To assess such differences we fit a linear model to the data and analyze the outcome of that. Such a linear model may include factors that take into account the mouse-to-mouse variation and the vessel order, but it may also include interaction effects, as above, which allows, e.g. a particular mouse to be affected by the vessel order in a way that differs from the effects on the rest.

We assume a linear model $t \sim \mathrm{Vo}+\mathrm{M}$, where t either is the time-to-peak or time-to-20%-of-the-peak, Vo is the vessel order, and M is the mouse number, and fit it to data. Instead of comparing all possible levels of factors, which is complicated because not all comparisons are statistically independent, one may compare the so-called *orthogonal contrasts* in the data. Briefly, in terms of time-to-peak, such an analysis reveals that first order capillaries are delayed with respect to penetrating arterioles, while third order capillaries are delayed with respect to the preceding orders. With respect to time-to-20%-of-the-peak, the third order capillaries are delayed with respect to all other orders included above, i.e. orders 0,1, and 2.

This indicates that the onset of response of low-order vessels happens simultaneously, which suggests that the cellular signaling arrives simultaneously to these vessels, while some delay may be seen, if one assesses the time-to-peak, which however quantifies the combined ability of the pericyte-vessel complex to dilate.

Note here that we found a possible effect of time *within* the first phase to be non-significant. Thus, it did not matter for the response timing that some vessels were imaged later in the experiment within that experimental phase.

LM calculated the overall mean of time-to-peak measurements to be 152.8498 and to be significantly different from zero. After that, it assessed the contrast of Vo factor levels 0 and 1 and found it to be non significant, as well as the contrast between 0,1 and 2. If we compare M 304 (red) to M 307 (yellow) we will see that yellow points are significantly above the red for all vessel orders. This was also recognized by the LM, which found the contrast M_304_vs_307 to be significant.

LM suggests that we could leave out levels 1 and 2 from the Vo factor and levels 65, 67, 69 from the M factor. We are mostly interested in Vo, and the summary of LM means no difference between the vessel orders. This is in perfect agreement with Fig. 2.16 and establishes a solid ground to move forward with the analysis.

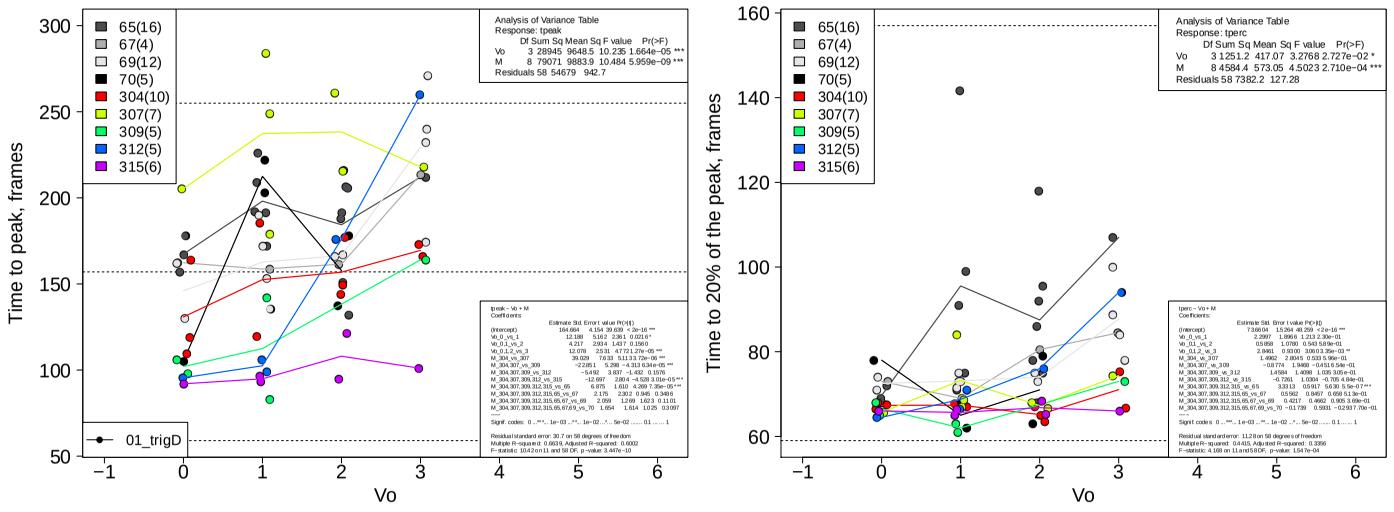


Figure 2.21: Same as Fig. 2.19 but for all time control and all ANGII w. gel mice without specific selection of images. Responses from the first phase only are included, which eliminates the difference between the groups. Vo is a highly significant effect on time-to-peak ($p \approx 2 \cdot 10^{-5}$) and significant, but much less so, on time-to-20%-of-the-peak ($p \approx 3 \cdot 10^{-2}$).

2.4.6 Analysis of individual vessel branches

Recent data [Hil+15] suggest that only a subset of pericytes contain smooth-muscle actin enabling contractility. This is qualitatively in agreement with the results of the previous section, namely that first order capillaries seem to be affected differently by ANGII as compared to higher order capillaries. Reference [Hil+15] indicates that the probability of expressing smooth muscle actin decreases with increasing distance along the branching tree. However, for a given vessel-branch, smooth-muscle actin-expression at a certain depth in the branching tree guarantees smooth-muscle actin-expression also at shallower depths in the tree. Thus, smooth muscle actin expressing pericytes may exist at all levels of the branching tree but only until a certain depth that depends on the mouse and/or the vessel branch. An active response is only possible, if the pericytes have the ability to contract.

This suggests a working hypothesis in which (i) the response dynamics are of an active nature and similar for vessels associated with either smooth muscle cells (zeroth order vessels) or smooth-muscle action-expressing pericytes (potentially of any order); and (ii) the response dynamics of vessels associated with pericytes that do not express smooth-muscle actin are similar but passive and thus distinctly different from the active responses in (i). Furthermore, for given mouse or possibly vessel branch, the active responses exist until a certain branching depth after which the responses become "slower" due to lack of ability to dilate actively.

Here, we again focused on the control phase of mice groups 1 and 2 in Table 2.9. For each mouse, a number of vessel branches were studied (e.g. Figure 2.12 shows two vessel branches for Mouse 65). All recordings from fields of view associated with a given vessel branch were combined, and the responses were studied as function of vessel order and pericyte classification.

Each of the Figures 2.22 and 2.23 (see also Fig. B.5) shows the branch-wise responses from an individual mouse. Notice how the responses shown in the right columns of panels are consistent with the hypothesis about dynamics changing at one branch-dependent level in tree. "Fast" responses are seen at the upper levels (low order numbers) in the branching tree, and after a certain level, responses become "slower". Notice, also that this transition may happen at different levels in the branching tree, even between different branches in an individual mouse (Fig. 2.22). Finally, Figure 2.23 shows that the "fast" response may be found at all levels in a branch.

Obviously, for lack of specific markers for smooth muscle actin, we cannot establish rock solid proofs, but our data is consistent with the hypothesis of Ref. [Hil+15].

Thus, the *current data supports the idea that pericytes play an active role in the blood flow regulation in the capillaries*, but the evidence disagrees with what was previously published (Ref. [Hal+14]): We do not find a delay of the response of zeroth order penetrating arterioles relative to first (or higher) order capillaries.

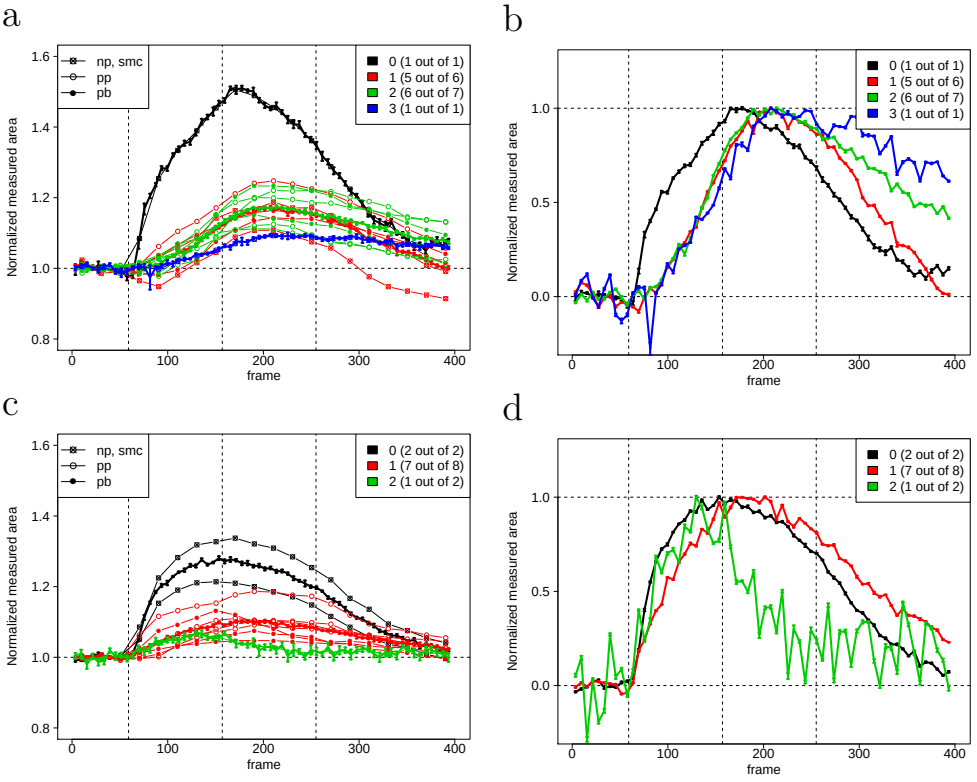


Figure 2.22: **(a)** All responses from the first phase of the experiment that belong to a single vessel are plotted (Mouse 65, here the "top-left" vessel branch from Fig. 2.12). Different colors represent the responses of vessels of different orders (see legend). Thick solid lines represent the averaged response of each order (6-frame temporal averaging was used for these averages). All individual time traces, that contribute to each of the order groups, are plotted with thin lines (using 20-frame temporal averaging). The different symbols indicate their classification (see legend). The amount of responsive time traces that contributed to each average is indicated in the legend. **(b)** Same as **a** showing only each group average rescaled with respect to its peak amplitude. **(c-d)** Same as **(a-b)** but for the "bottom-right" vessel branch from Fig. 2.12.

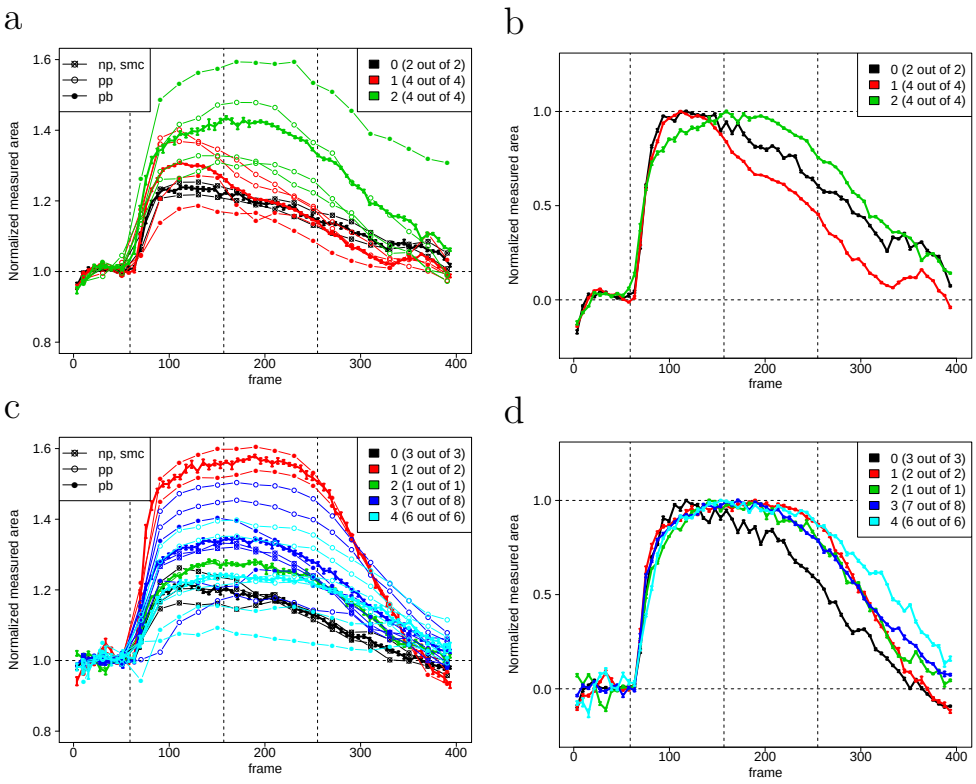


Figure 2.23: Same as in Fig. 2.22, but for Mouse 304. The figure shows that the "fast" response may be found at all levels in a branch.

2.4.7 Results and discussion

In this section, we performed a thorough analysis of dynamics of penetrating arterioles and capillaries. We found penetrating arterioles and first order capillaries to have identical response dynamics both when comparing averaged paired responses (Fig. 2.16) and individual mouse-wise contributions (Fig. 2.19). This finding disagrees with the previously published claim that there is a significant delay in the responses of orders 0 compared to orders 1+ (Fig. 2.17), and we showed that the discrepancy between the two could be at least partly explained by the fluctuations in the first order vessels in the data used in [Hal+14]. However, our current data supports the idea that pericytes play an active role in the blood flow regulation in the capillaries, and it is only the evidence for it that we found to be different (Section 2.4.6).

2.5 Pericytes are local centers of blood flow regulation

The true role for pericytes in the regulation of capillary blood flow still remains controversial and highly-debated in the literature and does not seem to be well-established. We were able to find issues (Fig. 2.17 and Fig. 2.4) with data analysis in both [Hal+14] and [Hil+15], that might have led to the misinterpreted results about the relative response dynamics of arterioles and capillaries. In this section, our analysis showed that pericytes are likely to be the local centers of capillary blood flow regulation rather than the initiators of the response to stimulus (see also Appendix **??**). Moreover, though the relative timing of the responses of capillaries and arterioles is typically mouse-specific, on average we find no significant differences between them.

We selected segments in the two-photon microscopy time-lapse movies that covered biologically different regions of the vessel network and classified them manually as arterioles, somata and non-somata regions with corresponding vessel orders, as shown in Fig. 2.24a. Each segment produced a trace of the averaged local diameter of the vessel that we picked up using a custom *ImageJ* macro. The extracted raw diameter traces are shown in Fig. 2.24b. The resting diameters of the segments decrease as we go deeper into the branching network, from $\simeq 20$ μm for the large arterioles to $\simeq 5-6$ μm at the branching point with vessel order 2.

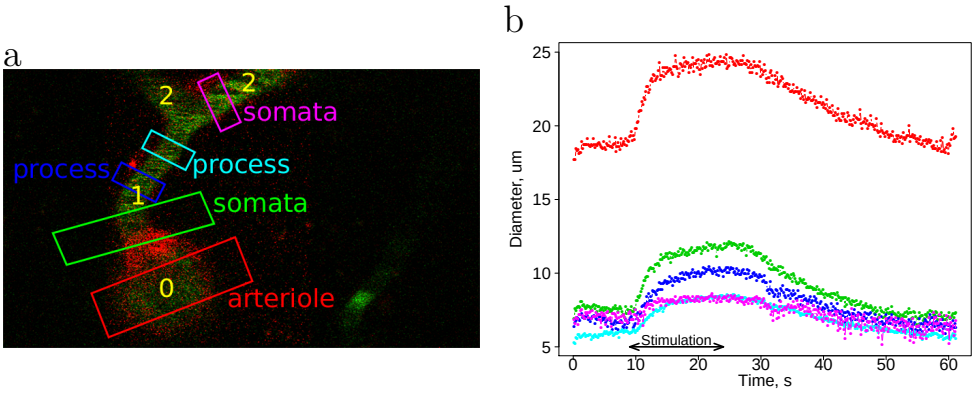


Figure 2.24: Early stimulation-induced local responses of somata-associated capillary segments. **(a)** Two-color two-photon microscopy image from the whisker-barrel region in the brain of an anesthetized mouse. The lumen of blood vessels appear green while smooth muscle and pericytes appear red. We define the penetrating arteriole as order zero and capillaries branching of from this as orders 1,2,... (yellow numbers). We selected five local segments of vessels for analysis (colored regions), based on their location and association with pericyte somata, processes, or absence of pericytes. **(b)** We recorded a time-lapse movie at 6.5 Hz for 400 frames during electrical stimulation of the whisker pad (stimulation period of $\sim$ 15 s indicated by arrow). For each segment in **a**, the vessel diameter traces were determined as a function of time (colored points).

As Fig. 2.24b shows, the raw traces of vessel diameter are noisy, and this is problematic for accurate analysis of the dynamics of the traces. This is why we smoothed the traces first using a local regression model, and then rescaled them with respect to their resting and amplitude diameter, as shown in Fig. 2.25a. We extracted the time to 50% of the peak ($T_{50\%}$) from each trace as the measure of the dynamics of the response and show them separately in Fig. 2.25b.

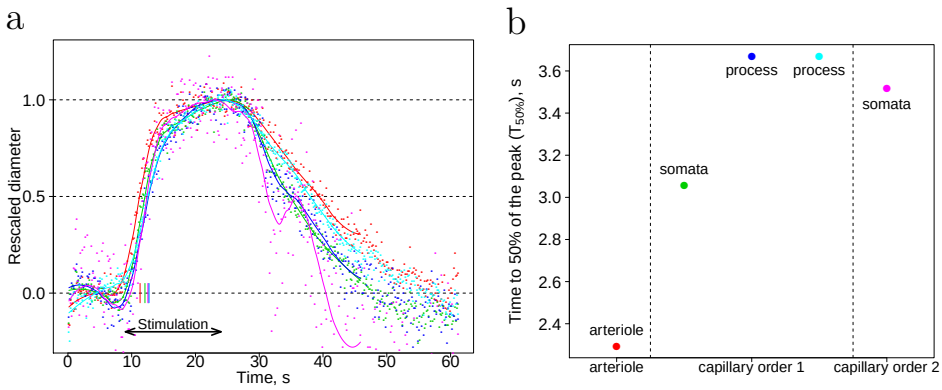


Figure 2.25: **(a)** For each time-resolved vessel diameter trace in Fig. 2.24b, we obtained a smoothed trace (colored solid lines, Methods). Then, we subtracted the average vessel diameter prior to stimulation and normalized the resulting diameter trace with respect to its peak value. Notice that each trace achieves its peak at the end of the stimulation period, roughly. For each response, the time to 50% of the peak, $T_{50\%}$, is indicated (colored vertical lines). **(b)** $T_{50\%}$ from **a**, shown as function of the position of each segment in the branching network and its classification, suggest that local vessel segments associated with pericyte somata might exhibit faster responses than their immediate non-somata capillary neighbors (compare green to blue and magenta to cyan, respectively). In this particular case, the response of the penetrating arteriole precedes all others.

Finally, we selected the cases where a somata region had a non-somata local neighbour (like green and violet segments in Fig. 2.24a) to compare their dynamics in terms of their $T_{50\%}$. We show the histogram of the differences in $T_{50\%}$ between the somata center and the penetrating arteriole (Fig. 2.26a) and the local non-somata neighbours (Fig. 2.26b). The histograms show that somata are on average slightly slower ($\sim$0.33 s, 33 obs., p-value = 0.14, not statistically significant) than arterioles and that somata are significantly ($\sim$0.86 s, 94 obs., p-value = $2e^{-5}$) faster than their local non-somata neighbors. At the same time, we found that the response can be leaded at any microvascular level: 1st order capillaries lead the response $\simeq$60% of the time followed by penetrating arterioles ($\simeq$30%). The numbers are in fine agreement with recent findings [Cai+17].

Note that Fig. 2.26 reports differences of relative dynamics and does not take

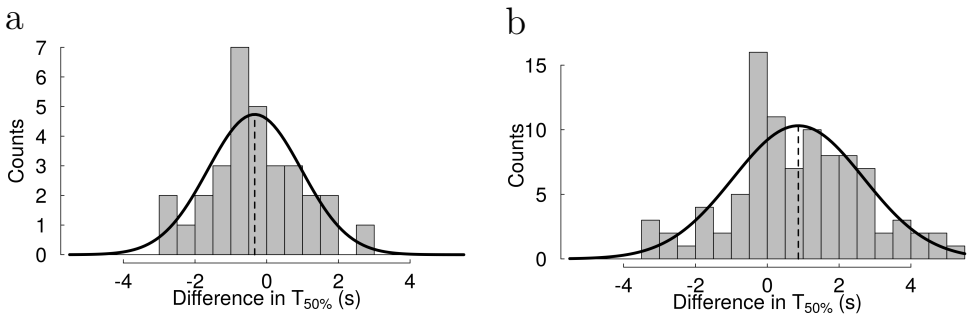


Figure 2.26: Pericyte somatas respond to stimulus significantly faster than their immediate non-somata neighbors. **(a)** The histogram of differences in T50% between arterioles and fastest non-arterioles imaged simultaneously, obtained from all mice selected for the study (Fig. 2.27). The data is consistent with the null hypothesis of true mean equals zero (33 observations, Wilcoxon signed rank test with continuity correction (W) p-value = 0.14), meaning that there is no significant difference between the two groups. **(b)** Same as **a**, but for differences between the non-somata immediate neighbors and somata centers. The data is not consistent with the null hypothesis of true mean equals zero (94 observations, W p-value = $2e^{-5}$), so we accept that there is a significant difference between the two groups. Vessel regions associated with pericyte somata respond to stimulus significantly faster than their immediate non-somata neighbours.

into account how large the contributing regions actually are. Arterioles have resting diameters that are significantly larger than those of capillaries (Fig. 2.28a), so the latter still being somewhat (though insignificantly) faster on average (Fig. 2.26b) might be simply because arterioles, that serve as large blood supply hubs, have a higher potential to dilate than capillaries and it takes relatively longer for them to reach this limit. At the same time, the difference in resting diameters of immediate somata and non-somata regions is insignificant (Fig. 2.28b, p-value = 0.73), so the result in Fig. 2.26b is definitely not caused by the different sizes of the compared regions.

So, the difference in timing between the somata centers and local non-somata regions suggests that centers react to stimulus significantly faster, i.e. regulate the the blood flow locally. At the same time, none of our previous analyses (Section 2.4) of the relative response dynamics of arterioles and capillaries was able to find significant differences between the two, so we are still not able to confirm the findings from [Hal+14]. This is why we are only able to conclude that *the role of pericytes is likely to be the local centers of capillary blood flow regulation, rather than the initiators of the response to stimulus.*

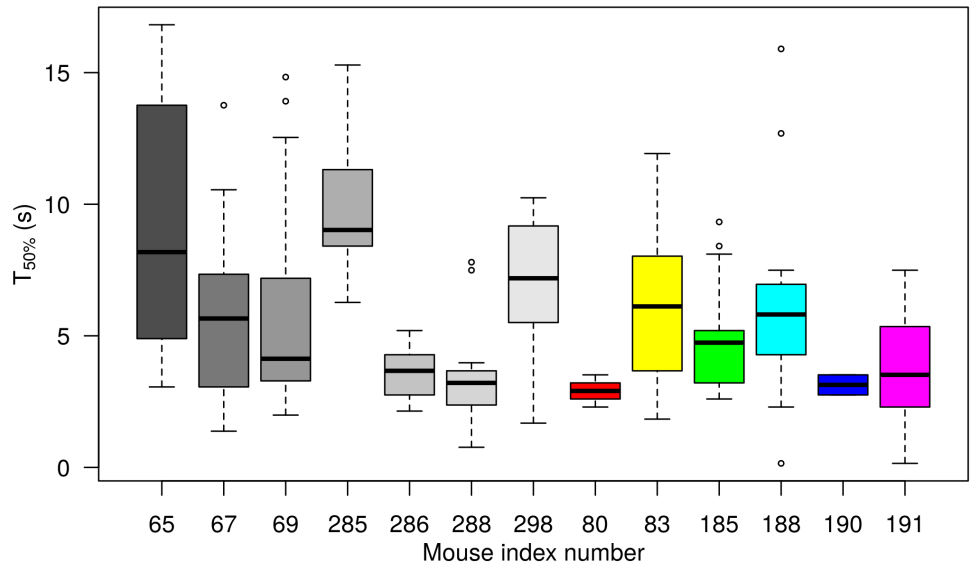


Figure 2.27: Boxplots of $T_{50\%}$ of all mice selected for the study of dynamics. Grayscale-colored boxplots indicate the No ANGII (control) group, and rainbow-colored — the ANGII (treatment) group. This is an illustrative example of a large mouse-to-mouse variation in our study, that prevented us from using the raw $T_{50\%}$ data when comparing dynamics.

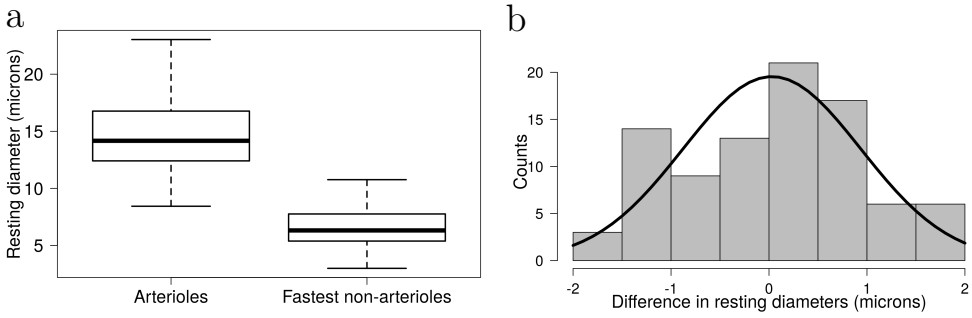


Figure 2.28: **(a)** Boxplots of resting diameters of traces contributing to Fig. 2.26a. Clearly, the resting diameters of arterioles are significantly larger than of capillaries. **(b)** Same as **a**, but for resting diameter differences for data from Fig. 2.26b. As expected, there is no significant difference (Wilcoxon signed rank test with continuity correction p-value = 0.73) between somatas and non-somata immediate neighbours.

2.6 Effects of angiotensin II (ANGII)

Hypertension has profound effects on the brain and the cerebral circulation is a major target of these actions. Acute or chronic elevations in arterial pressure attenuate cerebrovascular reactivity, disrupt the blood-brain barrier, and alter cerebrovascular autoregulation [Kon78; Kaz+03; Pir+13; al13]. Additionally, hypertension is associated with an impaired regulation of the cerebral circulation during neural activity [Kaz+03; Cal+13; BOG15]. Dysfunctional regulation of this neurovascular coupling (NVC) has important implications, since it perturbs the delivery of substrates to active brain cells and impairs the removal of potential deleterious by-products of cerebral metabolism [Kis+17b]. Different cell types within the neurovascular unit (NVU), including astrocytes, vascular smooth muscle cells (VSMCs), and endothelial cells, contribute to NVC. And lately, pericytes have also been recognized as important players in the vascular response to neural activity due to their ability to dilate capillaries in functional hyperaemia [Hal+14; KN14; Hil+15; Kis+17a].

Notably, the octapeptide hormone angiotensin II (ANGII) is a key mediator in the development and the maintenance of hypertension [BOG15]: ANGII is centrally involved in the pathogenesis of many forms of human hypertension, which is evidenced by the prevalence and efficacy of antihypertensive therapies targeting the "classic" renin-ANG system (RAS) that produces circulating ANGII. However, a tissue RAS that locally produces ANGII is present in all major organs, including the blood vessels and the brain. Available evidence, though limited, suggests that dysfunctional local ANGII signalling, rather than hypertension itself, is the culprit behind impaired neurovascular regulation [Kaz+03; Cal+13].

2.6.1 ANGII does not affect the difference in timing between somata centers and non-somata local neighbours

One group of mice in our study received ANGII right from the start of the experiment (Fig. 2.10). In this way we ensured the best possible physiological condition of the ANGII-treated mice for the study of dynamics. If ANGII had an effect on the response dynamics of arterioles and capillaries (Section 2.5), we could expect the results of the analysis of difference in timing to be dependent on the group of mice it was applied to. We show the results of this analysis in Fig. 2.29.

For both No ANGII and ANGII groups of mice, the mean difference of the sample is significantly non-zero, indicating the same fact that somata centers are faster than their local non-somata regions (Fig. 2.26a). However, it is clear from the one sample tests that there is no statistically significant difference between the two groups of mice, so we conclude that *ANGII had no effect on the difference in timing between somata centers and penetrating arterioles, as well as between somata centers and non-somata local neighbours.*

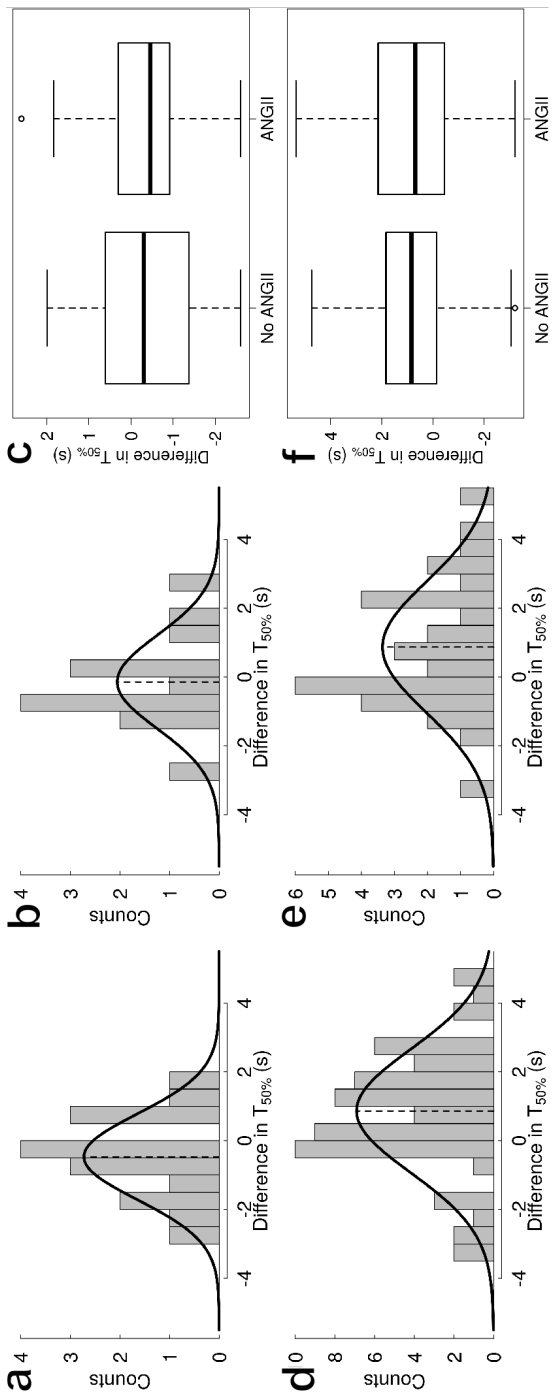


Figure 2.29: **(a)** Same as Fig. 2.26a, but for No ANGII mice only (Wilcoxon signed rank test with continuity correction (W) p-value = 0.15). **(b)** Same as Fig. 2.26a, but for ANGII mice only (W p-value = 0.6). **(c)** Boxplots of **a** and **b** compared to each other. The data is consistent with the null hypothesis true difference in means equals zero (Wilcoxon rank sum test with continuity correction (WS) p-value = 0.7). **(d)** Same as Fig. 2.26f, but for No ANGII mice only (W p-value = $3.2e^{-4}$). **(e)** Same as Fig. 2.26f, but for ANGII mice only (W p-value = 0.026). **(f)** Boxplots of **d** and **e** compared to each other. The data is consistent with the null hypothesis true difference in means equals zero (WS p-value = 0.65).

2.6.2 ANGII decreases resting CBF and CBF increase

Cerebral blood flow (CBF) measurements have proofed to be a reliable tool to study the neuronal function of the brain [LVD93]. The effects of ANGII on the mouse somatosensory cortex at the level of CBF were considered in [Kaz+03]. For the topical application of ANGII, they showed that application of just 50 nanomolars (nM) of ANGII causes a dramatic change in the CBF increase (Fig. 2.30b) over its baseline, but does not affect the resting CBF significantly. They also demonstrated that concentrations of ANGII higher than 1 micromolar (μM) cause a significant decrease in the value of the resting CBF (Fig. 2.30b).

We used laser Doppler flowmetry near the craniotomy to measure the cerebral blood flow (CBF) in the neocortex. In our setting, absolute values of such measurements were arbitrary. As proof of their stability and reliability, we show the averaged raw CBF signal for Mouse 65 (Fig. 2.10) in several phases during the timeline of experiment (Fig. 2.30a). We observed how CBF remained constant before stimulus, clearly responded to it, and went back to the pre-stimulation value repeatedly after the initial train. Therefore, our measurements were well-suited for monitoring relative changes in CBF — resting CBF and CBF increase (Fig. 2.30b).

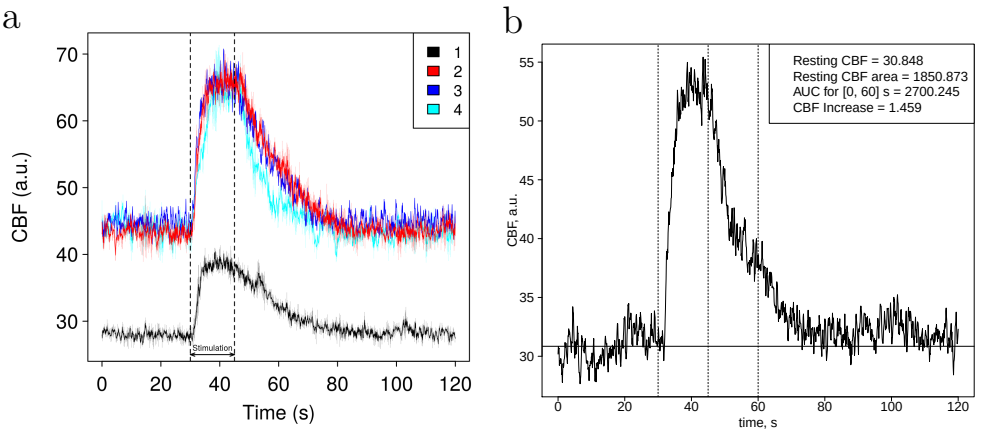


Figure 2.30: **(a)** Cerebral blood flow (CBF) measurements (mean ± s.e.m.) during the timeline of the experiment (Fig. 2.10). After the initial train (1, black), CBF went up and stayed virtually constant for one hour (2-4). We have therefore limited ourselves to using only the data from the experimental phases 2-4, where the overall physical condition of mice remained stable. **(b)** Resting CBF is the mean of the pre-stimulation part of the signal, from 0 to 30 seconds (legend). CBF increase is calculated as the ratio of the area under curve from 0 to 60 seconds S_{AUC} and the resting CBF area $S_{\mathrm{rest}} = 60 \cdot$ Resting CBF (legend). It is a quantification of how much total response is happening during the stimulation compared to the "resting" state of the mouse.

Initially, we considered all groups of mice (Fig. 2.10) that did not show systemic effects (Table 2.9): Time control, ANGII mice w gel, ANGII mice w/o gel, and High ANGII w/o gel. We considered the value of the resting CBF normalized to its value in the initial experimental phase, respectively, the CBF increase over the resting CBF, both in each of the three experimental phases (Fig. 2.31).

For High ANGII mice, there is a clear drop in resting CBF that is consistent with previously reported results [Kaz+03]. Note that though resting CBF does not seem to get back to the initial value in the rinse phase of the experiment, such a recovery was not shown in [Kaz+03] for high concentration of ANGII and, to our best knowledge, in the literature in general.

CBF increase for our data is consistent with being constant for all stages of the experiment, and that contradicts the findings from [Kaz+03] for low concentrations of ANGII. This is likely due to the fact that we are dealing with much larger concentrations of ANGII (10 - 100 μM versus 50 nM).

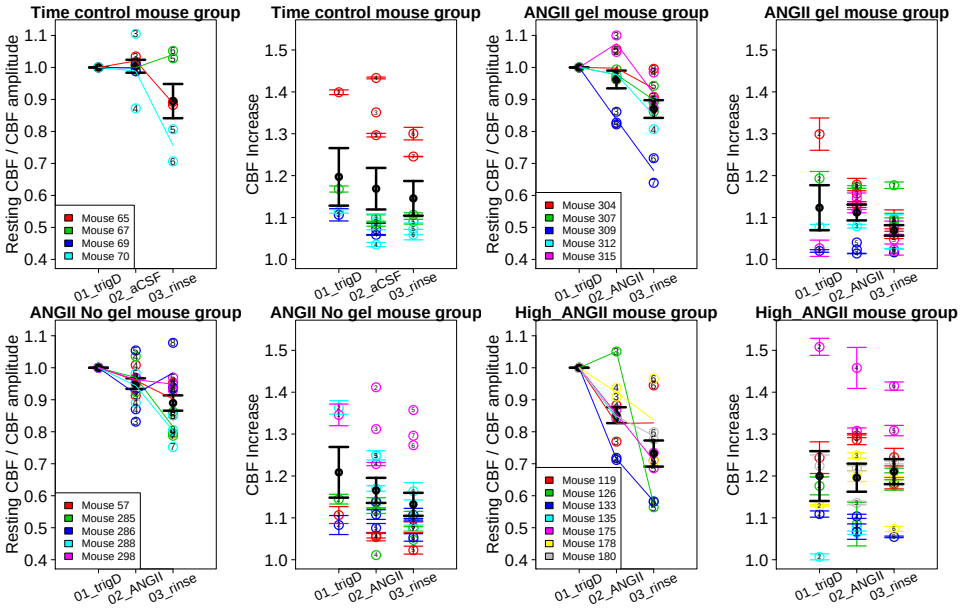


Figure 2.31: Normalized resting CBF and CBF increase for four groups of mice. *Top left*: time control mice, *Top right*: ANGII mice with gel, *Bottom left*: ANGII mice without gel, *Bottom right*: High-ANGII mice without gel. Black symbols represent the average over all mice within a group as mean ± s.e.m. Indexes represent the phases in the experiment. Resting CBF shows a significant decrease in the ANGII phase of the experiment, but fails to recover the initial value in the final phase of the experiment. CBF increase is consistent with being constant in all groups and phases of the experiment.

Next, we re-grouped the mice (Fig. 2.10) to emphasize the effect of ANGII on CBF (Fig. 2.32). Same as in Fig. 2.31, resting CBF decreased upon administration of ANGII in a concentration-dependent manner (Fig. 2.32a), in agreement with previous findings [Kaz+03]. Topical administration of ANGII did not have an effect on the CBF increase when measured relative to their values in the control phase of the experiment (Fig. 2.31). However, the resting values themselves decreased upon administration of ANGII, so CBF increase must have dropped accordingly. To take into account the effect of changing resting CBF upon topical administration of ANGII (Fig. 2.32a), we normalized each value for the CBF increase in both experimental phases, Control and ANGII, with respect to its resting value in the control phase (Fig. 2.32b). In the absence of ANGII superfusion we found that the CBF increase dropped five per cent on average (black circles) between the two experimental phases separated in time by about an hour (Fig. 2.30). Note that this decrease is independent of how the CBF increase was normalized, as the corresponding values for resting CBF remained constant for this mouse group (Fig. 2.32a). However, CBF increases dropped upon administration of ANGII in a concentration-dependent manner (significant for High ANGII; significant also for ANGII if a single outlier is disregarded). Thus, *ANGII also modulates neurovascular coupling in a concentration-dependent manner*. This confirms and extends previous work demonstrating this for a single low concentration of ANGII [Kaz+03].

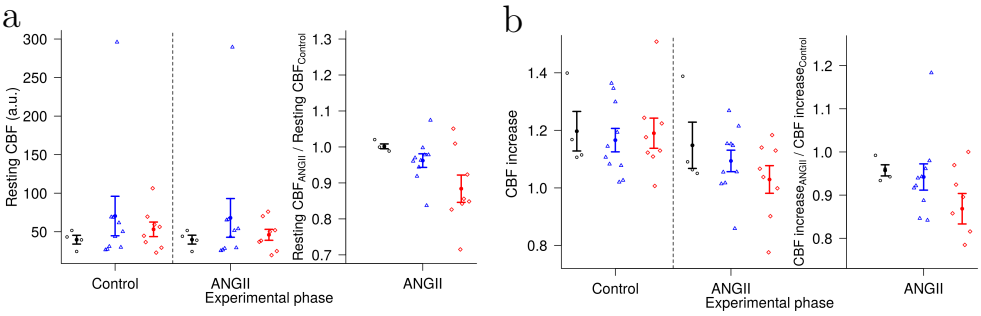


Figure 2.32: **(a)** Raw laser Doppler flowmetry measurements of resting CBF in Control and ANGII phases for each mouse in three mouse groups (Fig. 2.10a-c): No ANGII (black, 4 mice), ANGII (blue, 10 mice), and High ANGII (red, 8 mice). Mean ± s.e.m. for each group is indicated (filled symbol with error bars). For the ANGII phase, additionally, we normalized each resting CBF value its value in the Control phase. Differences between group averages of such normalized resting CBF values are mutually significant (linear model fitted to normalized resting CBF as a function of ANGII concentration (LM) p-value = 0.01). **(b)** Same as **a**, but for CBF increase relative to the resting value in the control phase. All group averages decreased significantly between the two experimental phases. Differences between group averages were also significant (same as in **a** but for CBF increase, LM p-value = 0.03).

2.6.3 ANGII constricts resting diameters of vessels up to order 2

In this section we investigated the role of ANGII as a potential vasocontrictor at the level of resting diameters (see also Appendix **??**). The idea behind the analysis was very simple: if ANGII is really a vasoconstrictor, then upon its application the resting diameter of the vessel must decrease compared to its initial value.

For No ANGII mice (Fig. 2.10), we found that during experiments vessel diameters increased for pial arteries (30%) and penetrating arterioles ($\simeq$20%) (black circles in Fig. 2.33a, b). Relative to such time-dependent increased vessel diameters, however, topical administration of ANGII resulted in decreased diameters of microvessels in a concentration-dependent manner (blue and red circles in Fig. 2.33a, b). Interestingly, we also observed increased resting diameters (10%) during experiments of capillary segments co-localising with pericyte somata in 1st and 2nd orders, in the absence of ANGII (Fig. 2.33c in black circles), while diameters of capillaries of 3rd and higher order remained unchanged on average (black circles in Fig. 2.33d). Upon administration of ANGII in high concentrations, we observed constrictions of 1st and 2nd order capillaries co-localising with pericyte somata (red circles in Fig. 2.33c), while lower concentrations of ANGII did not significantly affect capillary diameters in these orders (blue circles in Fig. 2.33c). Neither concentration of ANGII affected capillaries of 3rd and higher order (blue and red circles in Fig. 2.33d).

This shows that *ANGII directly changes the contractile tone of microvessels, presumably by directly affecting smooth muscle cells (SMCs) enwrapping pial arteries and penetrating arterioles; as well as by affecting a subset of capillary pericytes directly and not through systemic effects. Thus, ANGII is a local vasoconstrictor.*

While all SMCs express the contractile proteins, e.g. α-SMA, required for regulation of microvessel diameter, only a subset of capillary pericytes do so. The probability for a capillary pericyte to possess a contractile protein apparatus, however, is not uniform but appears to decrease with branching order [Hil+15]. This likely explains why we observed both temporal and ANGII-coupled regulation of capillary diameters in the 1st and the 2nd branching order but not significantly in higher orders.

This establishes the potency of ANGII as a local in vivo vasoconstrictor at the capillary level, in qualitative agreement with previous in vitro studies [MCA97].

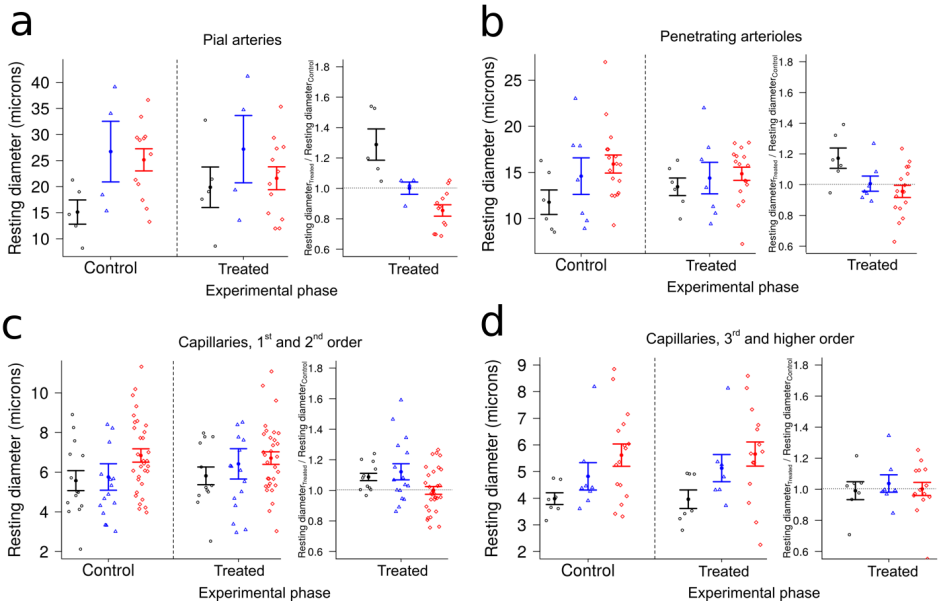


Figure 2.33: **(a)** Same as Fig. 2.32a, but for resting diameters of pial arteries. No ANGII (black, 5 observations), ANGII (blue, 4 observations), and High ANGII (red, 12 observations). Normalized resting diameters decreased with concentration of ANGII (No ANGII vs. ANGII, Wilcoxon signed rank test with continuity correction (W) p-value = 0.03; ANGII vs. High ANGII, W p-value = 0.04; No ANGII vs. High ANGII, W p-value = $3e^{-4}$). **(b)** Same as **a**, but for penetrating arterioles. No ANGII (black, 6 observations), ANGII (blue, 7 observations), and High ANGII (red, 16 observations). Application of ANGII results in decreased vessel diameter with no detectable concentration dependence, however (No ANGII vs. ANGII, W p-value = 0.03; No ANGII vs. High ANGII, W p-value = 0.008). **(c)** Same as **a**, but for capillary segments of 1st and 2nd order co-localising with pericytes somata. No ANGII (black, 12 observations), ANGII (blue, 16 observations), and High ANGII (red, 31 observations). Application of a high concentration of ANGII resulted in decreased diameters while a low concentration had no detectable effect (No ANGII vs. High ANGII, W p-value = 0.02). **(d)** Same as **a**, but for capillary segments of 3rd and higher order co-localising with pericyte somata. No ANGII (black, 7 observations), ANGII (blue, 7 observations), and High ANGII (red, 14 observations). No differences in resting values across treatments were detected.

2.6.4 ANGII prevents pericyte response

In Section 2.6.1 we found that ANGII did not affect the dynamics of penetrating arterioles and pericyte somata regions and their immediate neighbours. For penetrating arterioles, we observed that peak diameters normalized to their individual resting diameters decreased with topical superfusion of ANGII (Fig. 2.34a). This suggests that the myogenic mechanism in the walls of small arterioles is responsible for part of the decrease in CBF under functional hyperaemia (Fig. 2.32b). In contrast, we did not observe any such ANGII-induced changes for peak diameters of first order capillaries (Fig. 2.34b).

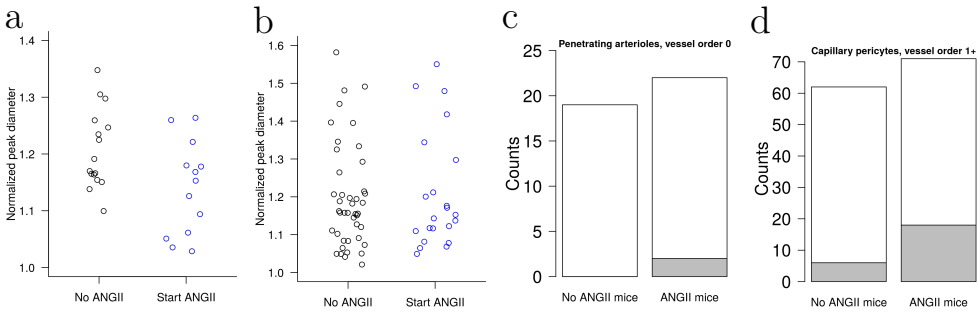


Figure 2.34: **(a)** Peak diameters of penetrating arterioles normalized to their resting diameters. No ANGII (black, 16 observations from 7 mice) and Start ANGII (blue, 13 observations from 6 mice). Such normalized peak values decreased significantly with superfusion of 10 μM ANGII (Wilcoxon rank sum test (W) p-value = 0.03). **(b)** Same as **a**, but for capillary pericytes and their immediate non-somata neighbours of vessel order 1. No significant difference for Start ANGII and No ANGII mice (W p-value = 0.97). **(c)** The number of responsive (white) and non-responsive (grey) diameter traces measured with two-photon microscopy of penetrating arterioles during stimulation of the whisker pad for two groups of mice: No ANGII mice (7 mice, 19 observations) and Start ANGII mice (6 mice, 22 observations). ANGII did not change the ratio of responsive and non-responsive traces (Fisher's exact test for count data (F) p-value = 0.49). **(d)** Same as **c**, but for capillary segments of all orders co-localising with pericytes somata. Here, ANGII significantly changed the ratio of responsive and non-responsive traces (F p-value = 0.024).

Another potential effect of ANGII could be that is changes the ratio of active (responsive, Appendix B.2) and passive (non-responsive) pericytes. We observed clear and reproducible dilations at the levels of penetrating arterioles (Fig. 2.15), and at the capillary level (Fig. 2.14d). The dilations persisted throughout the period of stimulation, demonstrating that the stimulation did not cause vessels to reach their maximum diameter prematurely. We counted the cases of absence of dilations in arterioles and capillaries (Fig. 2.34c, d). Interestingly, while penetrating arterioles responded consistently, the number of responding capillary segments co-localising

with pericyte somata was decreased significantly (Fisher's exact test for count data p-value = 0.024) in the ANGII mouse group (71 obs. from 6 mice) compared to No ANGII (62 obs. from 7 mice). This demonstrates that *ANGII not only reduces the resting diameters of pericytes but interacts with pericytes to prevent them from dilating microvessels as response to sensory stimulation, suggesting a possible role for pericytes in ANGII-impaired NVC.*

Then, we proceeded to investigate possible transitions between capillary regions co-localising with different subtypes of pericytes. Previous studies have indicated that such a transition exists between capillary regions co-localising with α-SMA$^+$, respectively, α-SMA$^-$ pericytes and that the location of this transition varies from branch to branch in the microvasculature [Hil+15; A+15]. For lack of specific markers for α-SMA, we used the presence and absence of dilations to discriminate between contractile and non-contractile subtypes of pericytes. Therefore, for each capillary branch shooting off from a penetrating arteriole, we determined the branching orders of responsive and non-responsive capillary vessel segments that co-localised with a pericyte cell body (Fig. 2.35a).

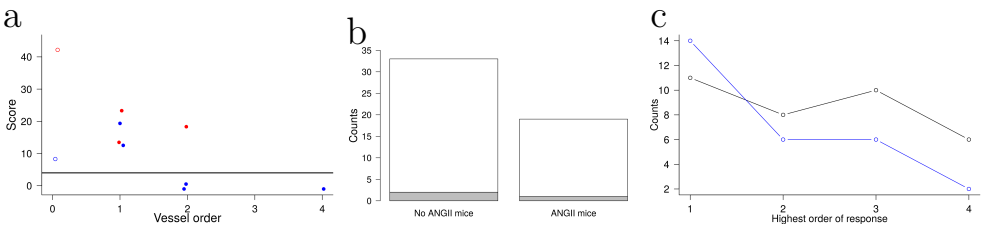


Figure 2.35: (a) Response scores for penetrating arterioles (open circles) and capillary segments co-localising with pericyte somata (filled circles) as a function of branching order in two vascular branches (red and blue) of a mouse. Responses were classified as responsive if their score exceeded 4 (solid black line). **(b)** Based on data as in **a**, for each branch, we detected the lowest order at which a non-responsive segment was observed. Then, we detected any occurrence of a responsive segment in a higher branching order. The number of branches with that response pattern (grey) versus the number of branches with other response patterns is shown (white). Superfusion of the cortex with ANGII did not change this ratio. **(c)** For each capillary branch as in **a**, we detected the highest branching order with a responsive segment. Superfusion with ANGII decreased the median of the highest branching order (Wilcoxon rank sum test with continuity correction p-value = 0.046).

In 94% of branches (Fig. 2.35b), we found data to be consistent with the existence of a transition between α-SMA$^+$ (responding capillary segments) and α-SMA$^-$ (non-responding capillary segments) pericytes, in agreement with recent imaging data [Hil+15; A+15]. Here, both branches with no detections of non-responsive segments (red data in Fig. 2.35a) and branches with a responsive to non-responsive transition (blue data in Fig. 2.35a) were considered consistent. While there was no

difference in the number of branches consistent with such transitions between control and ANGII mice (Fig. 2.35b), the location of this transition shifted towards lower branching orders with superfusion of ANGII (Fig. 2.35c). Those observations can be explained by *the potential ability of ANGII to interfere with cell-to-cell communication along microvessels*, as demonstrated in vitro [MCA97]. Rather than affecting responses of individual pericytes at random locations, ANGII may interrupt signalling at some position along a microvessel to prevent downstream pericytes from responding. This interpretation is consistent with signal propagation along microvessels observed recently in vivo [Cai+17].

2.6.5 Results and discussion

In this section, we investigated a number of potential effects of ANGII (see also Appendix **??**). We showed that ANGII did not affect the difference in timing of the responses of the vessels (Fig. 2.29). At the same time, we found that resting CBF and CBF increase dropped upon administration of ANGII in a concentration-dependent manner (Fig. 2.32), in agreement with previous findings [Kaz+03]. These results confirm that ANGII modulates neurovascular coupling. We also observed local constrictions of penetrating arterioles, and first and second order capillaries co-localising with pericyte somata after administration of high concentrations of ANGII (Fig. 2.33), which proves that ANGII changes the contractile tone of microvessels by directly affecting SMCs as well as a subset of capillary pericytes. We also demonstrated that ANGII interacts with pericytes to prevent them from dilating microvessels as response to sensory stimulation, by assessing the ratio of active pericytes over passive ones in two groups of mice (Fig. 2.35). In addition, we showed that the mismatch in NVC, suggested to occur during hypertension, can be at least partially explained by ANGII-induced decreases in dilations of arterioles and absence of pericyte-generated capillary dilatation. Such ANGII-induced absences of dilation correlate with distance along a vascular branch, suggesting that ANGII interrupts cell-to-cell communication along microvessels.

2.7 Conclusion

Pericytes are contractile cells surrounding the endothelial cells of small blood vessels. The contractility of pericytes has been a subject of discussion in literature since their discovery. It is clear nowadays that pericytes can be contractile (active) and passive, and their exact classification is still an open issue. Pericytes also play an active role in the regulation of the blood flow, however their true regulatory function has only started to emerge recently.

Biological experiments are among the hardest to reproduce and this is why the results of similar investigation reported in different studies are often drastically different from each other. The difference is also partly attributed to the (un)intentionally murky communication of the methods of data analysis used in the study. It applies both to selection of valid data from the overall available dataset and especially to careful and transparent statistical analysis. This is exactly where we had a lot to offer to the field.

We developed a simple, fast, and noise robust areascan method for the extraction of a diameter time trace of a region of a vessel. It reproduces the results of the conventionally used linescan method for a narrow line-like area, and performs much better on regions that require a lot of linescan measurements for accurate extraction of the diameter trace.

With that tool in hand, we performed a thorough analysis of dynamics of penetrating arterioles and capillaries. We found penetrating arterioles and first order capillaries to have identical response dynamics both when comparing averaged paired responses and individual mouse-wise contributions. This finding disagrees with the previously published claim that there is a significant delay in the responses of orders 0 compared to orders 1+, and we showed that the discrepancy between the two could be at least partly explained by the fluctuations in the first order vessels in the data used in that study. However, our current data supports the idea that pericytes play an active role in the blood flow regulation in the capillaries, and it is only the evidence for it that we found to be different. We were able to conclude that the role of pericytes is likely to be the local centers of capillary blood flow regulation, rather than the initiators of the response to stimulus.

A number of mice in our study received the octapeptide angiotensin two (ANGII), that is known to play a central role in the mechanisms of hypertension. Still, the impact of hypertension on neurovascular signaling and regulation in the brain has not been extensively investigated, and mechanistic insights have only begun to emerge in the last decade. It is unclear whether hypertension itself impairs neurovascular regulation or if ANGII signaling is the real culprit behind the apparent hypertensive neurovascular dysfunction.

So we investigated a number of potential effects of ANGII. We showed that ANGII did not have effect on the difference in timing between somata centers and penetrating arterioles, as well as between somata centers and non-somata local neighbors. We also established that ANGII did not modulate the responses: we found no significant differences in response amplitude of vessel dilations at any level of the microvascula-

ture.

At the same time, we found that resting CBF decreased upon administration of ANGII in a concentration-dependent manner, in agreement with previous findings. We took into account the drop in resting CBF to demonstrate that CBF increases also dropped upon administration of ANGII in a concentration-dependent manner. These results confirm that ANGII modulates neurovascular coupling, and extend previous work demonstrating this for a single concentration of ANGII.

We also observed local constrictions of pial arteries, penetrating arterioles, and first and second order capillaries co-localising with pericyte somata after administration of high concentrations of ANGII, which proves that ANGII changes the contractile tone of microvessels by directly affecting smooth muscle cells as well as a subset of capillary pericytes. We also demonstrated that ANGII interacts with pericytes to prevent them from dilating microvessels as response to sensory stimulation, by assessing the ratio of active pericytes over passive ones in two groups of mice. Both for No ANGII mice and ANGII mice, we found that responses of pericytes are correlated with position along vessel branches in the capillary microvasculature: segments were either responsive throughout the branch or non-responsive from some position along a branch and throughout, which is consistent with earlier findings. In addition, we showed that the mismatch in NVC, suggested to occur during hypertension, can be at least partially explained by ANGII-induced decreases in dilations of arterioles and absence of pericyte-generated capillary dilatation. Such ANGII-induced absences of dilation correlate with distance along a vascular branch, suggesting that ANGII interrupts cell-to-cell communication along microvessels.

The tools developed in this thesis for noise-robust extraction of vessel diameter traces from experimental two-photon microscopy images could serve the broad community of biologists and neuroscientists. More importantly, we demonstrated on a number of examples how both careful data management and rigorous statistical analysis are crucial for deriving solid conclusions about the behavior of complex biological systems. Future datasets for similar scientific problems could make use of the analyses presented here.

APPENDIX A

Thrombin and blood coagulation

A.1 Coagulation factors

F No.	Name	Activated form
FI	Fibrinogen	Fibrin (FIa)
FII	Prothrombin	Thrombin (FIIa)
FIII	Thromboplastin (Tissue factor)	FIIIa
FIV	Calcium	FIVa
FV	Proaccelerin	FVa
FVI	Calcium	FVIa
FVII	Proconvertin	FVIIa
FVIII	Antihemophilic	FVIIIa
FIX	Christmas	FIXa
FX	Stuart-Prower	Prothrombinase (FXa)
FXI	Plasma thromboplastin antecedent	FXIa
FXII	Hageman	FXIIa
FXIII	Fibrin stabilizing factor (Laki - Lorand)	FXIIIa

Table A.1: Coagulation factors and their names.

A.2 Thrombograms from CAT

Raw experimental data from CAT is displayed below. `N` is the index number of the experimental data point, `Time` is the registered time after the start of the measurement, `Signal` is the raw fluorescence measured by CAT, and `Thrombogram` is the resulting output from CAT. Note that `Signal` column has 270 observations, while the `Thrombogram` column has only 268 observations.

N	Time	Signal	Thrombogram
"1"	0.95	12.65	0
"2"	1.29	12.63	0
"3"	1.62	12.67	0.44
...	...	...	...
"266"	...	...	1.48
"267"	...	...	4.78
"268"	90.21	1141.92	6.62
"269"	90.54	1143.93	
"270"	90.88	1145.45	

Same as above, but from a different experiment. Now, `Signal` column has 900 observations, while the `Thrombogram` column has 899 observations.

N	Time	Signal	Thrombogram
"1"	0	8.89	1.21
"2"	0.33	8.92	1.51
"3"	0.67	8.98	0.74
...	...	...	...
"898"	299	920.48	0
"899"	299.33	918.69	0
"900"	299.67	916.93	

A.3 Taking the second derivative from raw experimental data

One of the key parameters quantifying thrombograms in thrombin generation experiments is the velocity of thrombin, that is roughly the second derivative of the experimental signal. It can be calculated using the basic second-order forward difference stencil

$$F''_k = \frac{F_{k+2} - 2 \cdot F_{k+1} + F_k}{\Delta t^2} + O(\Delta t) \qquad k = 1 : N - 2, \tag{A.1}$$

and the noise-robust stencil

$$F''_k = \frac{F_{k+4} - 2 \cdot F_{k+2} + F_k}{4 \cdot \Delta t^2} + O(\Delta t) \qquad k = 1 : N - 4. \tag{A.2}$$

Both Eq. A.1 and Eq. A.2 are compared to the ideal differentiator in Fig. A.2a. The application of both stencils to the calibration signal (Fig. 1.5a) is presented in Fig. A.2b.

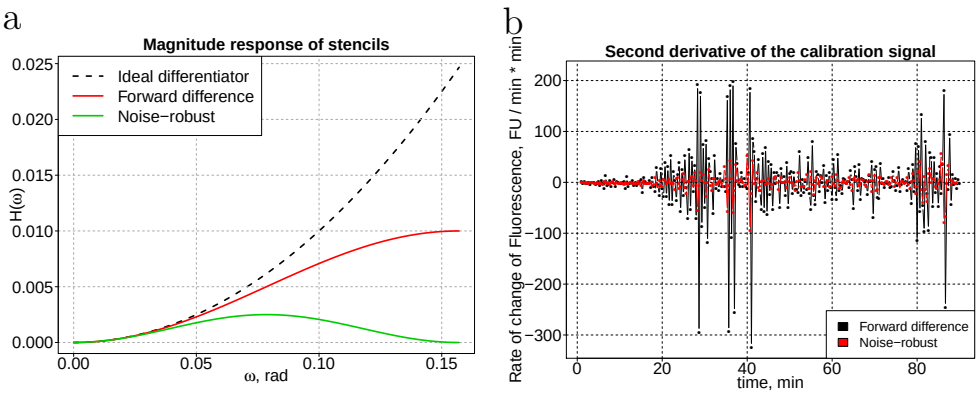


Figure A.2: Second derivative of noisy experimental data. **(a)** Magnitude response of stencils Eq. A.1 (red) and Eq. A.2 (green). **(b)** Visual demonstration of the enhancement of noise reduction for the second derivative of the raw calibration signal (Fig. 1.5a) calculated with the noise-robust stencil (red).

A.4 Thrombin equivalent units

Consider the following problem. There are two enzymes, E_0 and E_1 that act on the same substrate S and produce the same product P. The kinetic rates of the enzymes are however different. The following chemical reactions summarize the problem

$$S + E_0 \underset{k_{r_0}}{\overset{k_{f_0}}{\rightleftharpoons}} C_0 \xrightarrow{k_{cat_0}} E_0 + P, \qquad S + E_1 \underset{k_{r_1}}{\overset{k_{f_1}}{\rightleftharpoons}} C_1 \xrightarrow{k_{cat_1}} E_1 + P \ . \tag{A.3}$$

Let us focus on the late-times behaviour of the system in the typical case $s \gg e_{0,1}$. Assuming the steady-state regime for the intermediate complexes $C_{0,1}$ we get (see also Eq. 1.29) the following for their concentrations

$$c_0 = \frac{s \cdot e_0}{K_{M_0}} \qquad c_1 = \frac{s \cdot e_1}{K_{M_1}} \ . \tag{A.4}$$

We want both reactions to produce the same product. According to the law of mass action we write

$$\frac{dp}{dt} = k_{cat_0} \cdot c_0 = k_{cat_1} \cdot c_1 \ . \tag{A.5}$$

Inserting Eq. A.4 into Eq. A.5 we get

$$k_{cat_0} \cdot \frac{s \cdot e_0}{K_{M_0}} = k_{cat_1} \cdot \frac{s \cdot e_1}{K_{M_1}} e_0 = \frac{k_{cat_1} \cdot K_{M_0}}{k_{cat_0} \cdot K_{M_1}} \cdot e_1 \ . \tag{A.6}$$

Equation A.6 states that as long as we know the ratio of the kinetic constants of the two reactions we can easily calculate the concentration of E_0 in equivalent units of the concentration of E_1. In the particular case of thrombin generation E_0 is α_2M-T-T and E_1 is thrombin. It was reported in [Hem+03] that for those two enzymes $K_{M_0} = K_{M_1}$ and $2 \cdot k_{cat_0} = k_{cat_1}$. Therefore it follows from Eq. A.6 that $e_0 = 2 \cdot e_1$. The latter means that the concentration of α_2M-T-T in the equivalent units of thrombin concentration is simply half of its actual concentration.

A.5 Basic enzymatic reaction: numerical solution with R

System Eq. 1.11 can not be solved analytically, so it is easier to study it numerically. Listing A.1 gives the R [R C15] code to do it.

```
library(deSolve);  ## load deSolve library to use ode() function
time <- seq(0, 250, by = 0.1); ## define the time vector in minutes
## define the values of kinetic constants and initial conditions
parameters <- c(k.f = 0.001, k.r = 0.0001, k.cat = 0.1,
                s0 = 500, e0 = 200, c0 = 0, p0 = 0);
## define boundary conditions
state <- c(s = parameters[["s0"]], e = parameters[["e0"]],
           c = parameters[["c0"]], p = parameters[["p0"]]);
## define function that calculates the rates of changes
MichMent <- function(t, state, parameters) {
  with(as.list(c(state, parameters)), {
      ## define the rates of changes
      ds <- -k.f * e * s + k.r * c;
      de <- -k.f * e * s + (k.r + k.cat) * c;
      dc <- k.f * e * s - (k.r + k.cat) * c;
      dp <- k.cat * c;
      return(list(c(ds, de, dc, dp)));
    } );  # End of with()
}  ## End of MichMent
out <- deSolve::ode(y=state,times=time,func=MichMent,parms=parameters,
           atol = 1e-4, rtol = 1e-4);  ## output result to variable out
```

Listing A.1: Numerical solution of the system Eq. 1.11 with R and deSolve [SPS10].

A.6 Fluorescence measurements in empty wells

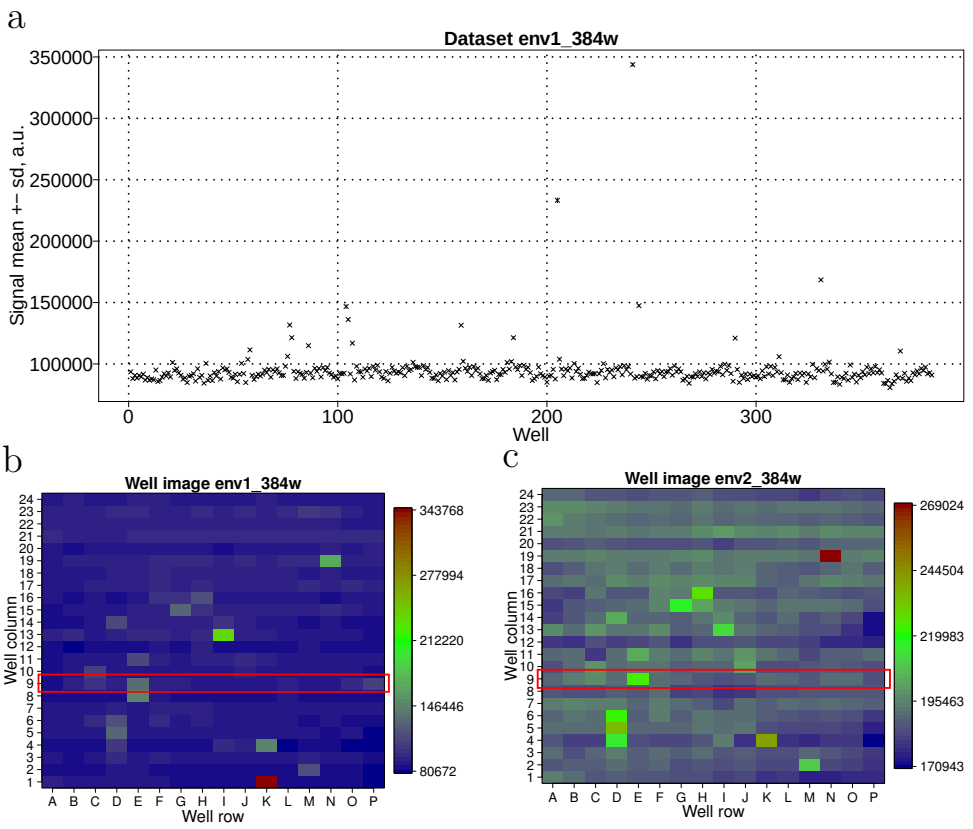


Figure A.3: Fluorescence measurements in empty wells. **(a)** Fluorometer performed a single measurement in each of the 384 empty wells on the plate, going from well to well consecutively. This was repeated for 10 runs. Each run took around 7 minutes. Mean ± standard deviation of the measurements for each of the 384 wells of the plate, well index number indicated on the x-axis. **(b)** Mean fluorescence for each of the wells on the plate, viewed from top. The selected row has the largest mean value of background fluorescence. **(c)** Same as **c**, but for a different plate and a different fluorometer. The spread of background fluorescence is less than for **b**, but the minimal value is much larger.

A.7 Substrate consumption in thrombin generation signals

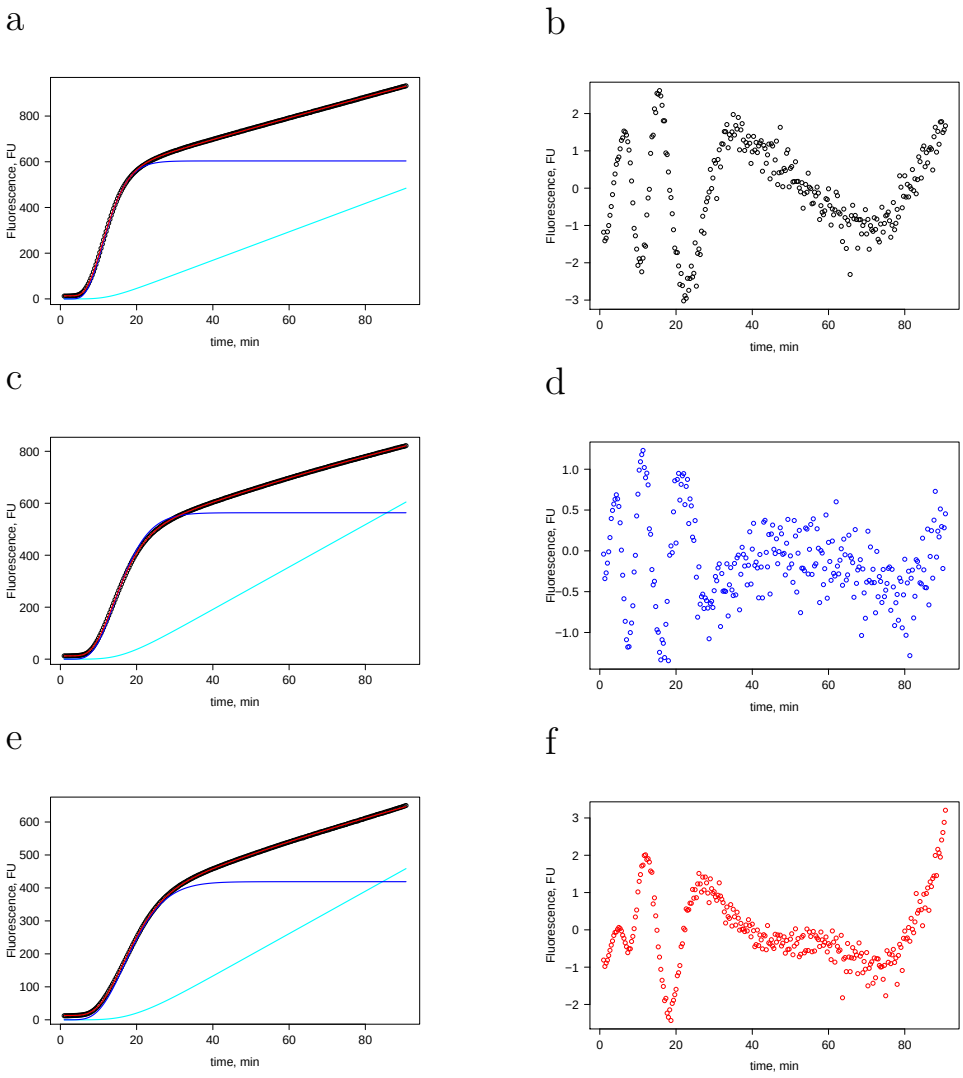


Figure A.4: Fits (red lines) of Eq. 1.39 to thrombin generation signals (black open circles) from Fig. 1.5, the coloring of which **(a, c, e)** is preserved in the plots of the residuals of the fits **(b,d,f)**. The part of Eq. 1.39 that describes the effect of substrate consumption is statistically significant, but its effect on the main physiological parameters is still negligible compared to the incorrect results given by the CAT machine (Fig. 1.29).

A.8 Autoregressive model for fluorescence signal

During the course of the thrombin generation or calibration experiment more and more fluorophore is continuously formed as a result of chemical reaction(s). It means that the measurements are not independent, but correlated: we observe a process of growth in time, so at every next time step we expect to get more fluorescence than for the previous time step. Time series analysis deals with correlated data, essentially by analyzing it either in the time domain or in the frequency domain.

There is no doubt that correlation is present in calibration or thrombin generation signals. The real question is whether accounting for it has an effect on the meaningful parameters that are extracted from the signals.

Let us take a calibration signal and fit it with Eq. 1.23 (Fig. A.5 red solid line, later abbreviated M_1). The simplest way to account for correlation would be to fit the same model but as a first order autoregressive process, i.e.

$$y_t = \alpha \cdot y_{t-1} + z_t \ , \tag{A.7}$$

where y is the signal, z is additive white Gaussian noise, and α is a constant coefficient that describes the coupling between the current value of the signal y_t and its value at the previous time step y_{t-1}. This fit is given in Fig. A.5 as the dashed green line (M_2). Visually, there is no difference between the two fits.

Complimentary to the autoregressive model, there is also a moving average model that reads e.g.

$$\hat{y}_t = \frac{y_{t-1} + y_{t-2} + y_t}{3} \ . \tag{A.8}$$

The number of previous points to average over is called the order of the moving average model, and it equals two for Eq. A.8. A model that combines both Eq. A.7 and Eq. A.8 is called autoregressive moving average model. Such a model of order 2 was also fitted to the calibration signal and is visualized in Fig. A.5 with cyan dots (M_3). Again, there is no visual difference between the fits.

The standard way to compare several models is to use the ANOVA. The log of the two ANOVA tests performed in R is given in Listing A.2. Based on the AIC, M_2 definitely explains more variation than M_1, and at the same time it also performs better than M_3. As for the meaningful physiological parameter CF_2 that the whole calibration experiment is made for — there is negligible difference between its value given by M_1 and M_2 (Table A.7).

In conclusion, autoregressive models perform significantly better for the correlated fluorescence signals than the models that do not take correlation into account. However, the difference between the two in terms of the values of meaningful physiological parameters they give is negligible. As soon as the extraction of physiological parameters is of primary interest for the analysis, taking correlation into account can be safely disregarded.

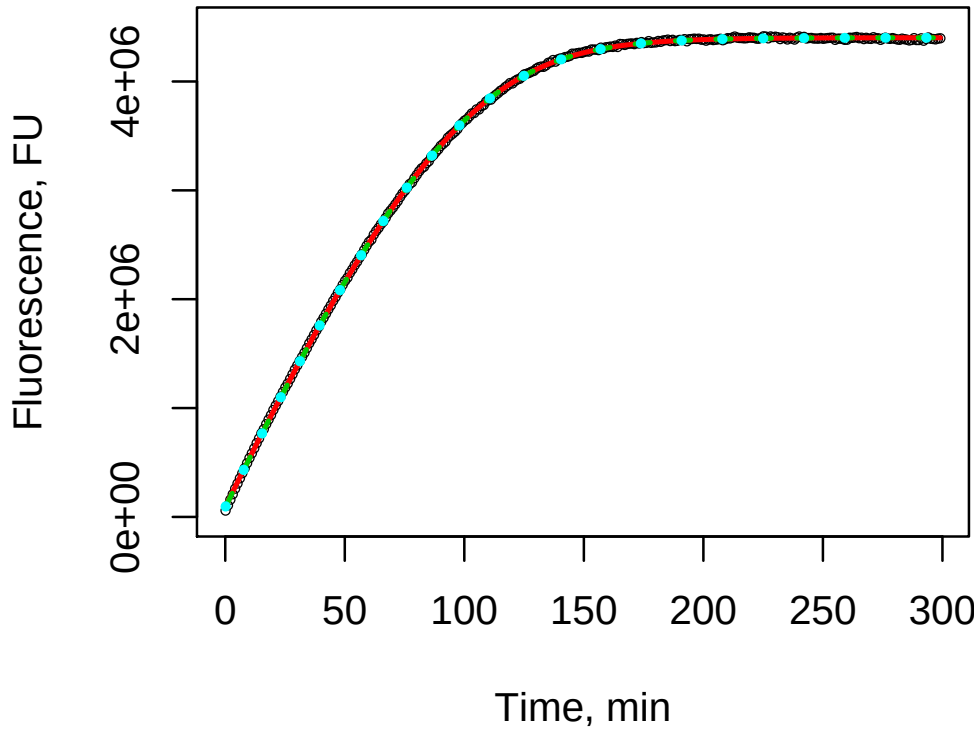


Figure A.5: Fitting autoregressive models to a calibration signal (black open circles). Solid red line is the fit of Eq. 1.23 to the data (M_1). Dashed green line is the fit of Eq. 1.23 as a first-order autoregressive process (Eq. A.7) (M_2), and cyan dotted line is the fit of Eq. 1.23 with both a first-order autoregressive process (Eq. A.7) and a moving average component (Eq. A.8) (M_3). There is no visual differences between the fits.

```
Model df      AIC      BIC    logLik   Test  L.Ratio p-value
   M1 15 6382.091 6437.647 -3176.045
   M2 16 6245.929 6305.190 -3106.965 1 vs 2 138.1616  <.0001
Model df      AIC      BIC    logLik   Test  L.Ratio p-value
   M2 16 6245.929 6305.19 -3106.965
   M3 17 6272.615 6335.58 -3119.308 1 vs 2 24.68638  <.0001
```

Listing A.2: ANOVA tests performed in R for models M_1 (Eq. 1.23), M_2 (Eq. 1.23 and Eq. A.7) and M_3 (Eq. 1.23 and Eq. A.7 and Eq. A.8). M_2 performs the best according to AIC.

A.9 Fitting a fluorescence signal with repetitions

It is widely known that repetitions of the same experiment allow to average data and therefore smooth it. In thrombin generation, once the measurement process was started it can strictly speaking never be rolled back and repeated under the same conditions. The technician will have to use a new plate to host a new sample of the same compound, but there is no guarantee that the plates as well as the wells on the plates are identical to previous ones in terms of the fluorescence signal they produce (Fig. A.3). The differences in the compounds of the samples come from the imprecise titration of chemicals, typically performed by hand. So, the repetitions of the same fluorescence signal are actually not the same, but are more alike to individuals within a group that show similar behavior. Therefore, averaging over such individuals should be done carefully.

Figure A.6a shows 8 repetitions of the same calibration signal. The simplest analysis is to average the signals (mean $\mu \pm$ standard deviation σ), and fit the mean fluorescence with weights of normal distribution ($1/\sigma^2$, Fig. A.6b) exactly as individual repetition would be fitted (Fig. A.6c). The standard deviation of the datapoints increases with the value of the mean fluorescence, so the weights and hence the "importance" of datapoints to the fit decreases. It means that early-times points are much more important for the fit than late-times points, and this is why the fit captures the first much better.

However, the fit is not supposed to simply favor early-times datapoints. Their standard deviation is low, but their ratio μ/σ is also low compared to the rest (Fig. A.6d). The values of the fitted parameters and their standard errors (Eq. 1.23, Fig. A.6c) are presented in column M_2 of Table A.7. The values of the calibration factor CF_2 (Eq. 1.37) are also given in the table, because CF_2 is ultimately the value the whole calibration experiment is required for.

The first natural thing to compare the fit to averaged fluorescence would be the fit to a single repetition, e.g. the first one. It is given in Table A.7 as column M_1. All points are equally weighted in this fit, so there is no focus on early-times points anymore and hence the value of CF_2 is quite different from that given by M_2.

There is an modern alternative to simply averaging over repetitions, called a . In this particular case the random effect of the model is a random repetition-specific value added on top of the "average" value for each of the fitted parameters. This model treats each signal individually, keeping the fixed effect the same for all of them. The parameters describing this same fixed effect are given in column M_3 of Table A.7, and they are very different from those of the fit to mean fluorescence in M_2. Finally, column M_4 presents the parameters obtained from a fit of a mixed effect model with first-order autoregression (Appendix A.8).

In conclusion, a mixed effect model produces the parameters that are a lot more alike to the ones obtained from the fits to individual signals, and both are very different from the ones obtained from the fit to averaged data. Autocorrelation plays a negligible role in the analysis (Appendix A.8).

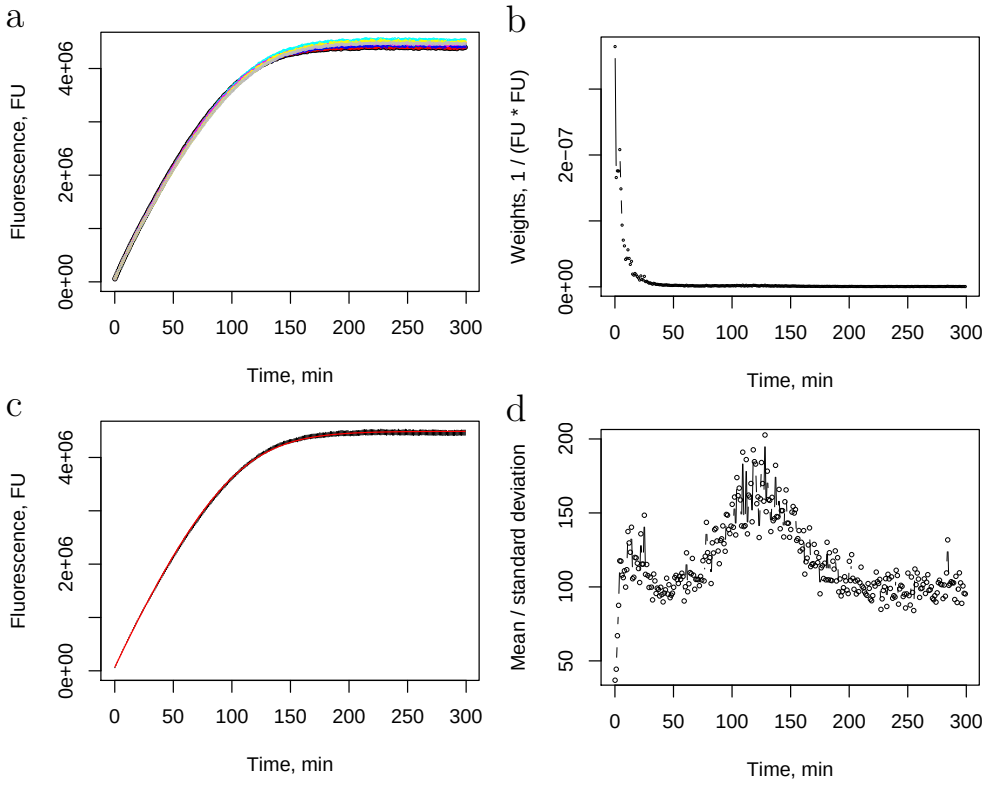


Figure A.6: Fitting a fluorescence signal with repetitions. **(a)** Eight repetitions of the same calibration signal, indicated by different colors. The fluorometer and the plate were the same for all the signals, so the variability in fluorescence measurements comes from individual differences between the wells (e.g. Fig. A.3) within the plate or the minor differences in sample compounds due to titration of chemicals performed by hand. **(b)** The weights of the averaged datapoints (mean $\mu \pm$ standard deviation σ) taken as if for normal distribution $1/\sigma^2$. Early-times points have much larger weights than the rest. **(c)** Fit (red line) of Eq. 1.23 to mean fluorescence μ with weights from **b**, data (black segments) shown as $\mu \pm \sigma$. The fit captures early-times points much better than the late-times points, because of the weights from **b**. (d) The ratio μ/σ for data from **c**. For early-times datapoints, the ratio is relatively low even though σ is also low.

	M_1	M_2	M_3	M_4
b	$9.1e^4 \pm 2.5e^3$	$6.6e^4 \pm 8.5e^2$	$9.7e^4 \pm 1.5e^3$	$9.4e^4 \pm 2.3e^3$
p_1	$4.31e^6 \pm 2.8e^3$	$4.43e^6 \pm 4e^3$	$4.38e^6 \pm 1.4e^4$	$4.38e^6 \pm 1.4^4$
p_2	$3.17 \pm 3e^{-2}$	$2.47 \pm 4e^{-2}$	$3.15 \pm 1.2e^{-2}$	$3.12 \pm 2.4e^{-2}$
p_3	$4.34e^{-2} \pm 2.6e^{-4}$	$3.6e^{-2} \pm 3.8e^{-4}$	$4.17e^{-2} \pm 2.94e^{-4}$	$4.15e^{-2} \pm 3.3e^{-4}$
CF_2	$5.34e^{-4} \pm 3e^{-8}$	$6.27e^{-4} \pm 7e^{-8}$	$5.48e^{-4} \pm 4e^{-8}$	$5.51e^{-4} \pm 5e^{-8}$

Table A.7: Parameters from fits to calibration signals with repetitions (Fig. A.6). Methods of fitting are indicated as follows: M_1 — fit of Eq. 1.23 to the first repetition only, M_2 — fit of Eq. 1.23 to mean fluorescence with weights (Fig. A.6c), M_3 — fit of Eq. 1.23 in a mixed effects model, M_4 — same as M_3 but with first-order autocorrelation structure in the model (Appendix A.8, Eq. A.7). The mixed effect model M_3 produces the parameters that are a lot more alike to the ones obtained from M_1, and both are very different from M_2. Autocorrelation (M_4) plays a negligible role in the analysis (Appendix A.8), so it can be safely disregarded when the analysis is only aimed at extracting the physiological parameters like calibration factor. The most important information that is contained in the calibration signal are the kinetic rates K_M and k_{cat}. They are contained in parameters p_2 and p_3 (Eq. 1.23). These parameters describe the late times part of the calibration signal. Therefore, weighted fits that put less importance on late time points of the signal (like M_2) are likely to produce more erroneous values of these parameters than unweighted ones.

APPENDIX B

Pericytes, hypertension and neurovascular coupling

B.1 Specifications of the experiment

B.1.1 Animals and preparation

All procedures involving animals were approved by the Danish National Ethics Committee carried out in accordance to the guidelines set forth in the European Council's Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes and are in compliance with the ARRIVE guidelines. Experiments used xxx 10-15 weeks old male Tg(Cspg4-DsRed.T1)1Akik [MGI] on a C57BL/6J background (called NG2 DsRed in the text and figures) which were purchased from the Jackson Laboratory (Bar Harbor, USA) and breed in house.

The animals were anesthetized with xylazine (10 mg kg^{-1} intraperitoneal) and ketamine (60 mg kg^{-1} intraperitoneal) and maintained during surgery with supplemental doses of ketamine (30 mg kg^{-1} every 20-30 minutes intraperitoneal). Upon completion of all surgical procedures, anesthesia was switched to alpha chloralose (50 mg kg^{-1} h^{-1} intravenous). Mice were prepared for experiments by cannulation of the trachea for mechanical ventilation (MiniVent model 845, Harvard Apparatus, Holliston, MA, USA). Catheters were placed into the left femoral artery and vein and perfused with physiological saline. The end-expiratory CO_2 (Capnograph type 340, Harvard Apparatus) and blood pressure (Pressure Monitor BP-1; World Precision Instruments) were monitored continuously in combination with blood gases in arterial blood samples (p_{O2} 115-130 mmHg; p_{CO_2} 35-40 mmHg; pH 7.35-7.45; ABL 700 Series; Radiometer Medical) to ensure the animals were kept under physiological conditions. The temperature was measured and maintained at 37°C during the experiment with a rectal thermometer regulated heating pad (TC-1000 Temperature Controller; CWE). The skull was exposed and glued to a metal plate using cyanoacrylate gel (Loctite Adhesives) and the plate was fixed in the experimental setup. A craniotomy was drilled with a diameter of approximately 4 mm with the centre 0.5 mm behind and 3 mm to the right of the bregma over the sensory barrel cortex region. The dura was removed and the preparation covered with 0.75% agarose (type III-A, low EEO;

Sigma-Aldrich) and moistened with artificial cerebrospinal fluid (pH 7.4 at 34^{o}C). The craniotomy was covered with a glass coverslip. When experiments were complete, mice were euthanized by intravenous injection of anesthesia (pentobarbital, 200 mg ml^{-1} and lidocaine hydrochloride, 20 mg ml^{-1}) followed by decapitation.

B.1.2 Stimulation

The mouse sensory barrel cortex was activated by stimulation of the contralateral ramus infraorbitalis of the trigeminal nerve using a set of custom-made bipolar electrodes inserted percutaneously. The cathode was positioned at the hiatus infraorbitalis, and the anode was inserted into the masticatory muscles. Thalamocortical infraorbitalis stimulation was performed at an intensity of 1.5mA (ISO-flex; AMPI) and lasting 1 ms, in trains of 15 s at 3 Hz. The stimulation was controlled by a sequencer file running within Spike2 software (version 7.02; Cambridge Electronic Design).

B.1.3 Technical details of the two-photon imaging setting

Fluorescein isothiocyanate-dextran (FITC-dextran, 2% w/v, MW 70,000, 50 μl, Sigma-Aldrich) was administered into the femoral vein to label the blood plasma. In vivo imaging of blood vessel diameter and pericyte location was performed using a commercial two-photon microscope (SP5, Leica), a MaiTai HP Ti:Sapphire laser (Millennia Pro; SpectraPhysics, mean output power 10 mW), and a 1.0 N.A. water-immersion objective (Leica). Tissue was excited at 900 nm wavelength, and the emitted light was filtered to collect red and green light from DsRed (pericytes) and FITC-dextran (vessel lumens). Penetrating arterioles were identified unequivocally in vivo in two ways, first by tracing their connections back to pial arterioles with obvious smooth muscle around them, and second by observing the direction of flow of the red blood cells from the arterioles into the parenchymal capillaries. Z-stack images were taken of the area of interest. XY-time series were taken to image pericytes and blood vessels during stimulation, with a frame size 5123300 pixels (170 ms per frame; pixel size was 93 to 201 nm depending on the magnification used, with a mean value of 155 nm; pixel dwell time was 1.1 ms).

B.1.4 Cerebral blood flow measurements

Cerebral blood flow (CBF) was recorded before, after and in between two-photon imaging recordings continuously using a laser-Doppler flowmetry (LDF) probe glued to the scull and angled to the two-photon imaging recording site (wavelength 780 nm; 250 um fiber separation allowing CBF measurement to a depth of 1 mm; Perimed AB, Järfälla, Sweden). Laser Doppler signal can't be recorded simultaneously to two-photon imaging since the outgoing two-photon laser light will interfere with the lasser-Dopler detector. The LDF signal was digitally recorded (PeriFlux 4001 Master, Perimed AB), sampled at 10 Hz, analog- to-digital (A/D) smoothed with a time

constant of 0.1 s using the Spike2 software (Cambridge Electronic Design). The LDF method does not measure CBF in absolute terms, but is valid in determining relative changes in CBF during moderate flow increases [FL96].

B.2 How to eliminate non-responding measurements

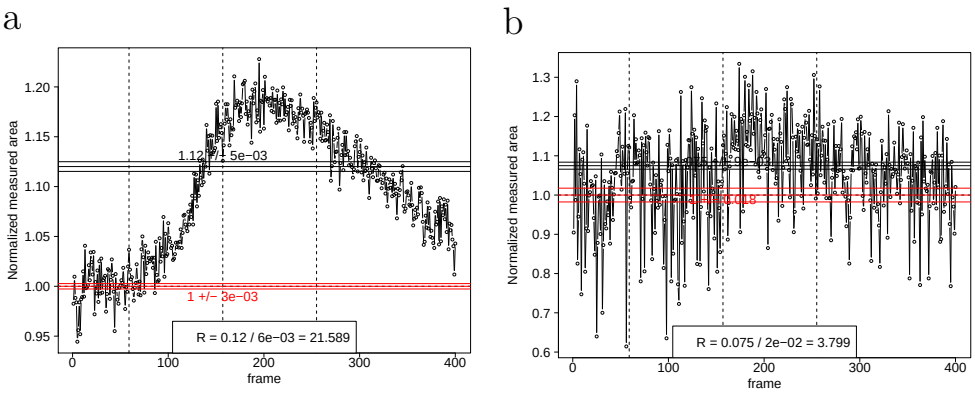


Figure B.1: We use an empirical criterion to define the responsive and the non-responsive time traces. In this figure, each trace was normalized with respect to its resting diameter — the mean of the pre-stimulation diameter of the vessel. We calculate the mean m_1 and the standard error on the mean s_1 from the pre-stimulation part of the response (from frame 1 to frame 59, indicated by red text and solid red lines). Next we do the same (m_2, s_2 indicated by black text and solid black lines) for the stimulation and post-stimulation part (frame 59 to frame 255). We calculate the difference $m_2 - m_1$, the standard error on the difference $\sqrt{s_1^2 + s_2^2}$ and their ratio R, the results are given in the bottom box. The time trace is considered responsive if $R \geq 4$, so **(a)** is responsive and **(b)** is not. The vertical dashed lines indicate the period of the 15 second stimulus starting at frame 59 and 15 second relaxation after the end of the stimulus, ending at frame 255. We do it to capture the response in case the time course peaks after the end of the actual stimulus, as in **a**.

B.3 Responses of arterioles vs. capillaries

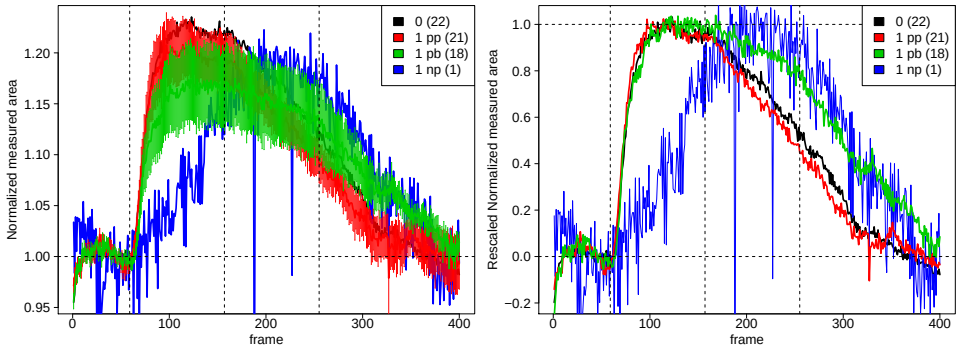


Figure B.2: Same as Fig. 2.16 except that first order vessels are split into three groups of, respectively, segments with pericyte bodies (pb), pericytes processes (pp), and absence of pericytes (np). There are no apparent differences between pericyte body and pericyte process regions in terms of speed of response or amplitude, while regions without pericytes seem different (the difference most likely is due to the fact that this particular region is from a slower responding mouse see Section 2.4.6).

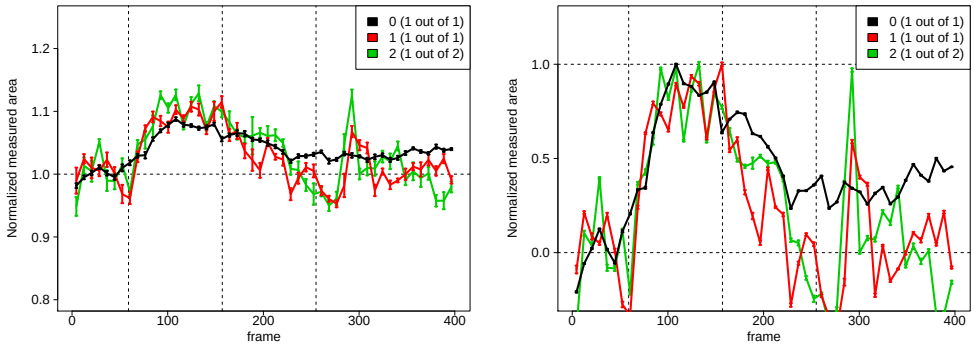


Figure B.3: *Left:* Responses of a pair of first order capillary (red), second order capillary (green) and zeroth order penetrating arteriole (black) from Mouse 140 from [Hal+14]. *Right*: Same as *Left* but each response was rescaled with respect to its peak amplitude and the resting area was subtracted. Note a fluctuation of large amplitude occuring at frame 300, which is comparable to the maximum response amplitude of the time trace.

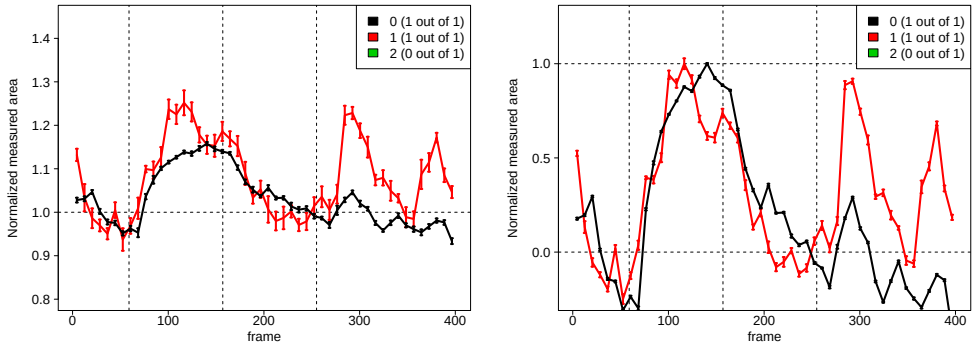


Figure B.4: *Left:* Responses of a pair of second order capillary (green) and zeroth order penetrating arteriole (black) from Mouse 166 from [Hal+14]. *Right*: Same as *Left* but each response was rescaled with respect to its peak amplitude and the resting area was subtracted. Note fluctuations of large amplitude occuring at frames 300 and 400, which are comparable to the maximum response amplitude of the time trace.

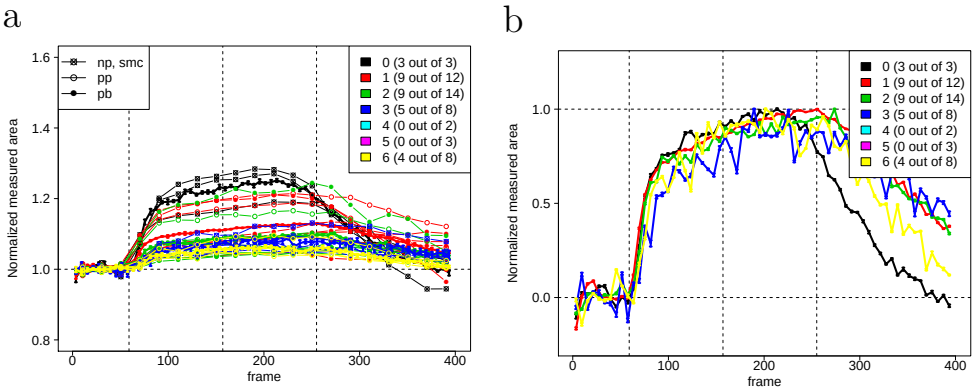


Figure B.5: Same as in Fig. 2.22, but for Mouse 307.

B.4 How to extract time-to-peak and time-to-x%-of-the-peak

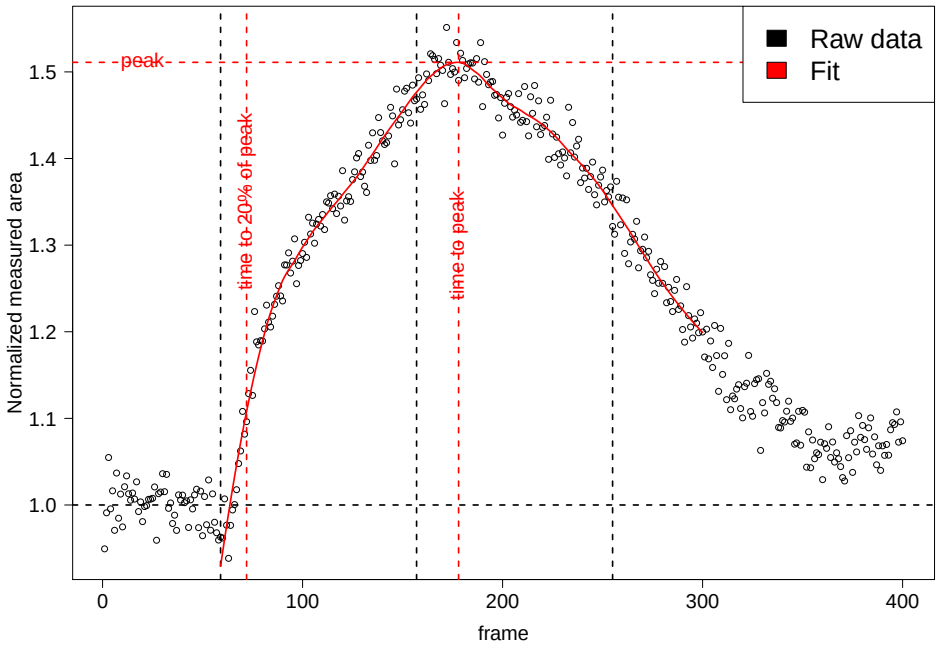


Figure B.6: Analysis of individual time traces. Normalized raw trace (black points) is fitted with a locally weighted polynomial regression method (LOESS) in the range from starting frame to frame 300 (red line). Peak, time-to-peak and time-to-20%-of-the-peak (value of 20% chosen as an example) are determined from the fitted model and shown in the plot.

Glossary

α_2-macroglobulin α_2-macroglobulin is a large plasma protein found in the blood. It is mainly produced by the liver, and also locally synthesized by macrophages, fibroblasts, and adrenocortical cells. In humans it is encoded by the A2M gene. α_2-macroglobulin acts as an antiprotease and is able to inactivate an enormous variety of proteinases. It functions as an inhibitor of coagulation by inhibiting thrombin [Sot+84]. 5

activity Activity is moles of substrate converted per unit time = rate · reaction volume. In SI it is measured in 1 katal = 1 mol s^{-1} [IUP]. 5, 136, 138, 141

additive white Gaussian noise Additive white Gaussian noise (AWGN) is a basic noise model used in Information theory to mimic the effect of many random processes that occur in nature. The modifiers denote specific characteristics:

- Additive because it is added to any noise that might be intrinsic to the information system.
- White refers to the idea that it has uniform power across the frequency band for the information system. It is an analogy to the color white which has uniform emissions at all frequencies in the visible spectrum.
- Gaussian because it has a normal distribution in the time domain with an average time domain value of zero.

Wideband noise comes from many natural sources, such as the thermal vibrations of atoms in conductors (referred to as thermal noise or Johnson-Nyquist noise), shot noise, black body radiation from the earth and other warm objects, and from celestial sources such as the Sun [McC]. 57, 124

ANGII Angiotensin II is a hormone that may act on the central nervous system to regulate renal sympathetic nerve activity, renal function, and, therefore, blood pressure [Pub]. 82, 104

ANOVA Analysis of variance (ANOVA) is a collection of statistical models used to analyze the differences among group means and their associated procedures (such as "variation" among and between groups), developed by statistician and evolutionary biologist Ronald Fisher. In the ANOVA setting, the observed variance in a particular variable is partitioned into components attributable to

different sources of variation. In its simplest form, ANOVA provides a statistical test of whether or not the means of several groups are equal, and therefore generalizes the t-test to more than two groups. ANOVAs are useful for comparing (testing) three or more means (groups or variables) for statistical significance. It is conceptually similar to multiple two-sample t-tests, but is more conservative and is therefore suited to a wide range of practical problems [Gel05]. 92, 124

big O notation a mathematical notation that describes the limiting behavior of a function when the argument tends towards a particular value or infinity. It is a member of a family of notations invented by Paul Bachmann, Edmund Landau, and others, collectively called Bachmann–Landau notation or asymptotic notation. In computer science, big O notation is used to classify algorithms by how they respond to changes in input size, such as how the processing time of an algorithm changes as the problem size becomes extremely large [Cor01]. 13

blood coagulation Blood coagulation is the process by which blood changes from liquid to gel, forming a clot. It potentially results in hemostasis, the cessation of blood loss from a damaged vessel, followed by repair. The mechanism of coagulation involves activation, adhesion, and aggregation of platelets along with deposition and maturation of fibrin. Disorders of coagulation are disease states which can result in bleeding (hemorrhage or bruising) or obstructive clotting (thrombosis) [Lil+09]. 1, 2, 136

chromogenic substrate Chromogenic substrate is a substrate that changes color when modified by a specific enzyme [IUP]. Foe example, chromogenic substrates used to detect thrombin activity are small molecules that access the active site of thrombin bound to α_2-macroglobulin. 5

clotting time Clotting time is the time required for a sample of blood to coagulate in vitro under standard conditions [Sil14]. 4

coagulation factor Coagulation factor is any of several plasma components that are involved in the blood coagulation [dic]. 1, 2, 117, 140

correlogram In the analysis of data, a correlogram is an image of correlation statistics. For example, in time series analysis, a correlogram, also known as an autocorrelation plot, is a plot of the sample autocorrelations versus the time lags. The correlogram is a commonly used tool for checking randomness in a data set. This randomness is ascertained by computing autocorrelations for data values at varying time lags. If random, such autocorrelations should be near zero for any and all time-lag separations. If non-random, then one or more of the autocorrelations will be significantly non-zero. [Cha89]. 15, 16, 55

endothelium Endothelium is a thin layer of simple squamos cells that lines the interior surface of blood vessels and lymphatic vessels, forming the interface between circulating blood or lymph in the lumen and the rest of the vessel wall [Dor]. 1

enzyme Enzyme is a macromolecular biological catalyst, responsible for thousands of metabolic processes that sustain life. Enzymes are highly selective catalysts, greatly accelerating both the rate and specificity of metabolic chemical reactions, from the digestion of food to the synthesis of DNA. Most enzymes are proteins, although some catalytic RNA molecules have been identified. Enzymes act by converting starting molecules (substrates) into different molecules (products). Almost all chemical reactions in a biological cell need enzymes in order to occur at rates sufficient for life [SBT02]. 2, 22, 136, 140, 141, 143

Fast Fourier Transform A fast Fourier transform (FFT) algorithm computes the discrete Fourier transform (DFT) of a sequence, or its inverse. Fourier analysis converts a signal from its original domain (often time or space) to a representation in the frequency domain and vice versa. An FFT rapidly computes such transformations by factorizing the DFT matrix into a product of sparse (mostly zero) factors. As a result, it manages to reduce the complexity of computing the DFT. Fast Fourier transforms are widely used for many applications in engineering, science, and mathematics. The basic ideas were popularized in 1965, but some algorithms had been derived as early as 1805 [HJB85]. 57

fibrin Fibrin is a fibrous protein involved in the clotting of blood. It is formed by the action of the serine protease thrombin on fibrinogen which causes the latter to polymerize. The polymerized fibrin forms a hemostatic plug or clot over a wound site [Keh03]. 1, 2, 136

fibrinogen Fibrinogen is a soluble glycoprotein that helps in the formation of blood clots [HCS58]. 1, 137

Full mass balance for an open system IN - OUT + GENERATION - CONSUMPTION = ACCUMULATION. 22

gaussian blur In image processing, a Gaussian blur (also known as Gaussian smoothing) is the result of blurring an image by a Gaussian function. It is a widely used effect in graphics software, typically to reduce image noise and reduce detail. The visual effect of this blurring technique is a smooth blur resembling that of viewing the image through a translucent screen. Mathematically, applying a Gaussian blur to an image is the same as convolving the image with a Gaussian function. Since the Fourier transform of a Gaussian is another Gaussian, applying a Gaussian blur has the effect of reducing the image's high-frequency components; a Gaussian blur is thus a low pass filter [SS01]. 86, 87

gene Gene is the molecular unit of heredity of a living organism. Living beings depend on genes, as they specify all proteins and functional RNA chains. Genes hold the information to build and maintain the cells of the organism and pass genetic traits to offspring [Sla14]. 1

glycoprotein Glycoprotein is a protein that contains oligosaccharide chains (glycans) covalently attached to polypeptide side-chains [IUP]. 2, 137

hemophilia Hemophilia is a genetic disorder that impairs the ability of the body to form blood clots, a process needed to stop bleeding when a blood vessel is broken [NHL]. 1, 3

hemorrhage Hemorrhage is blood escaping from the circulatory system [Hea]. 1, 136

hemostasis Hemostasis is a process which causes bleeding to stop, meaning to keep blood within a damaged blood vessel [dic]. 1, 136

hypertension Hypertension (HTN or HT), also known as high blood pressure (HBP), is a long-term medical condition in which the blood pressure in the arteries is persistently elevated. High blood pressure usually does not cause symptoms. Long-term high blood pressure, however, is a major risk factor for coronary artery disease, stroke, heart failure, peripheral vascular disease, vision loss, and chronic kidney disease [DP]. 104

independent identically distributed In probability theory and statistics, a sequence or other collection of random variables is independent and identically distributed (i.i.d.) if each random variable has the same probability distribution as the others and all are mutually independent. The abbreviation i.i.d. is particularly common in statistics (often as iid, sometimes written IID), where observations in a sample are often assumed to be effectively i.i.d. for the purposes of statistical inference. The assumption (or requirement) that observations be i.i.d. tends to simplify the underlying mathematics of many statistical methods (see mathematical statistics and statistical theory). However, in practical applications of statistical modeling the assumption may or may not be realistic. To test how realistic the assumption is on a given data set, the autocorrelation can be computed, lag plots drawn or turning point test performed [Cla]. 15

inhibitor Inhibitor is a substance that is responsible for decrease of a property of chemical reaction (rate of reaction, activity) [IUP]. 5

Kolmogorov-Smirnov test In statistics, the Kolmogorov–Smirnov test (K–S test or KS test) is a nonparametric test of the equality of continuous, one-dimensional probability distributions that can be used to compare a sample with a reference probability distribution (one-sample K–S test). The Kolmogorov–Smirnov

statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution. The null distribution of this statistic is calculated under the null hypothesis that the sample is drawn from the reference distribution (in the one-sample case) [Dan90]. 58

law of mass action Consider a *closed system* of reactants A and B *under isochoric conditions* (volume is constant) that transforms into products P and Q without intermediate steps: $aA + bB \rightarrow pP + qQ$, where a, b, p, q are stochiometric coefficients. The law of mass action [GW64] states that the rate of the reaction is proportional to concentration of the reactants in the power of stochiometric coefficients, so the reaction rate r for this system is: $r = \frac{-1}{a} \cdot [A]_t = \frac{-1}{b} \cdot [B]_t = \frac{1}{p} \cdot [P]_t = \frac{1}{q} \cdot [Q]_t = k \cdot [A]^m [B]^n$, where $[X]$ denotes the concentration of substance X, and $_t$ stands for the first time derivative, k is the reaction constant, and m and n are experimentally determined parameters. 22

linear model In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data. Such models are called linear models [Sea67]. 92

log-log plot In science and engineering, a log-log plot is a two-dimensional graph of numerical data that uses logarithmic scales on both the horizontal and vertical axes. Monomials — relationships of the form $y = ax^k$ — appear as straight lines in a log-log plot, with the power and constant term corresponding to slope and intercept of the line, and thus these graphs are very useful for recognizing these relationships and estimating parameters. Any base can be used for the logarithm, though most common are 10, e, and 2 [Bou]. 58

magnitude response Magnitude response is the quantitative measure of the output spectrum of a system or device in response to a stimulus, and is used to characterize the dynamics of the system. It is a measure of magnitude and phase of the output as a function of frequency, in comparison to the input. In simplest terms, if a sine wave is injected into a system at a given frequency, a linear system will respond at that same frequency with a certain magnitude and a certain phase angle relative to the input. For a linear system, doubling the amplitude of the input will double the amplitude of the output. If the linear system is time-invariant, then the frequency response also will not vary with time. Thus for linear time-invariant systems, the frequency response can be seen as applying the system's transfer function to a purely imaginary number argument representing the frequency of the sinusoidal excitation [Feu90]. 14, 17, 18, 119

michaelis constant The Michaelis constant is the substrate concentration at which the reaction rate is half of the maximum rate achieved by the system at saturating substrate concentration [LL13]. 25

neurovascular coupling Neurovascular coupling is the dynamic functional change in cerebral blood flow that occurs in response to local neuronal activity [Kis+17a]. 104

neurovascular unit Neurovascular unit is a functionally integrated group of different cell types in the brain composed of vascular cells (endothelial cells, vascular smooth muscle cells and pericytes), glia (astrocytes, microglia and oligodendrocytes) and neurons [Kis+17a]. 71, 74, 104

Padé approximant In mathematics a Padé approximant is the "best" approximation of a function by a rational function of given order — under this technique, the approximant's power series agrees with the power series of the function it is approximating. The technique was developed around 1890 by Henri Padé, but goes back to Georg Frobenius who introduced the idea and investigated the features of rational approximations of power series. The Padé approximant often gives better approximation of the function than truncating its Taylor series, and it may still work where the Taylor series does not converge. For these reasons Padé approximants are used extensively in computer calculations. They have also been used as auxiliary functions in Diophantine approximation and transcendental number theory, though for sharp results ad hoc methods in some sense inspired by the Padé theory typically replace them [BG96]. 54

pathway In biochemistry, a metabolic pathway is a linked series of chemical reactions occurring within a cell. The reactants, products, and intermediates of an enzymatic reaction are known as metabolites, which are modified by a sequence of chemical reactions catalyzed by enzymes [CN08]. 2

periodogram In signal processing, a periodogram is an estimate of the spectral density of a signal. The term was coined by Arthur Schuster in 1898. The periodogram has been superseded for most applications since 1898, but it is the most common tool for examining the detailed characteristics of filters and window functions, and it is used in conjunction with multi-block windowing to implement filter banks [Eng08]. 18, 57

platelet Platelet is a component of blood whose function (along with the coagulation factors) is to stop bleeding by clumping and clotting blood vessel injuries [Lak72]. 1

platelet rich plasma Platelet-rich plasma (PRP) is a concentrate of platelet-rich plasma protein derived from whole blood, centrifuged to removed red blood cells. It has a greater concentration of growth factors than whole blood, and has been used to encourage a brisk healing response across several specialties, in particular dentistry, orthopedics and dermatology. As a concentrated source of blood plasma and autologous conditioned plasma, PRP contains several different growth factors and other cytokines that can stimulate healing of soft tissue

and joints. Main indication in sports medicine and orthopedics are acute muscle strains, tendinopathy and muscle-fascial injuries and osteoarthritis [Smi16]. 6

protein Protein is a macromolecule, consisting of one or more long chains of amino acid residues. Proteins perform: catalyzing metabolic reactions, replicating DNA, responding to stimuli, and transporting molecules from one location to another. Proteins differ from one another primarily in their sequence of amino acids, which usually results in folding of the protein into a specific three-dimensional structure that determines its activity [CN08]. 22, 137, 138, 141, 142

rational function In mathematics, a rational function is any function which can be defined by a rational fraction, i.e. an algebraic fraction such that both the numerator and the denominator are polynomials [Haz01]. 54

semi-log In science and engineering, a semi-log graph or semi-log plot is a way of visualizing data that are related according to an exponential relationship. One axis is plotted on a logarithmic scale. This kind of plot is useful when one of the variables being plotted covers a large range of values and the other has only a restricted range — the advantage being that it can bring out features in the data that would not easily be seen if both variables had been plotted linearly [Bou]. 18

serine protease Serine protease is an enzyme that cleaves peptide bonds in proteins. In humans, it is responsible for co-ordinating various physiological functions, such as digestion, immune response, blood coagulation and reproduction [Hed02]. 1, 2, 137

Shannon's theorem Sampling is a process of converting a signal (for example, a function of continuous time and/or space) into a numeric sequence (a function of discrete time and/or space). Shannon's version of the theorem states:

- If a function $x(t)$ contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $1/2B$ seconds apart.
- A sufficient sample-rate is therefore $2B$ samples / second, or anything larger. Equivalently, for a given sample rate f_s, perfect reconstruction is guaranteed possible for a bandlimit $B < f_s/2$.

[Sha49]. 17

Shapiro-Wilk test a test of normality in frequentist statistics, published in 1965 by Samuel Sanford Shapiro and Martin Wilk [SW65]. The null hypothesis for this test is that the data are normally distributed. The Prob < W (the test statistic) value listed in the output is the p-value. If the chosen alpha level is

0.05 and the p-value is less than 0.05, then the null hypothesis that the data are normally distributed is rejected. If the p-value is greater than 0.05, then the null hypothesis is not rejected. 55

stencil In mathematics, especially the areas of numerical analysis concentrating on the numerical solution of partial differential equations, a stencil is a geometric arrangement of a nodal group that relate to the point of interest by using a numerical approximation routine. Stencils are the basis for many algorithms to numerically solve partial differential equations (PDE) [FF15]. 13, 14

subsampling Subsampling is a procedure by which a small, representative sample (subsample) is taken from a larger sample. A subsample may be a portion of the sample obtained by selection or division, an individual unit of the lot taken as part of the sample, the final unit of multistage sampling. The term "subsample" is used either in the sense of a "sample of a sample" or as a synonym for "unit". In practice, the meaning is usually apparent from the context or is defined. Good sub-sampling technique becomes important when the large sample is not homogeneous [IUP]. 4

substrate Substrate is typically the chemical species being observed in a chemical reaction, which reacts with reagent to generate a product [IUP]. 22

thresholding Thresholding is the simplest method of image segmentation. The simplest thresholding methods replace each pixel in an image with a black pixel if the image intensity is less than some fixed constant, or a white pixel if the image intensity is greater than that constant [SS01]. 86, 87

thrombosis Thrombosis is the formation of a blood clot inside a blood vesel, obstructing the flow of blood through the circulatory system [FF08]. 1, 136

time series analysis A time series is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time. Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Methods for time series analysis may be divided into two classes: frequency-domain methods and time-domain methods. In the time domain, correlation and analysis can be made in a filter-like manner using scaled correlation, thereby mitigating the need to operate in the frequency domain. The parametric approaches to time series analysis assume that the underlying stationary stochastic process has a certain structure which can be described using a small number of parameters (for example, using an autoregressive or moving average model). In these approaches, the task is to estimate the parameters of the model that describes the stochastic process [Shu88]. 124

tissue factor Tissue factor is a protein present in subendothelial tissue and leukocytes, that is necessary for the initiation of thrombin formation from the zymogen prothrombin [GJ79]. 2

turbidity the cloudiness or haziness of a fluid caused by large numbers of individual particles that are generally invisible to the naked eye, similar to smoke in air [Man+07]. 6

zymogen Zymogen is an inactive enzyme precursor. It requires a biochemical change for it to become an active enzyme [dic]. 2, 142

Bibliography

[A+15] Hartmann D. A. et al. “Pericyte Structure and Distribution in the Cerebral Cortex Revealed by High-resolution Imaging of Transgenic Mice”. In: *Neurophotonics* 2.5 (2015).

[AGC11] Armulik A., Genove G., and Betsholtz C. “Pericytes: developmental, physiological, and pathological perspectives, problems, and promises”. In: *Developmental Cell* 21.2 (2011), pages 193–215.

[al13] P. Toth et al. “Age-related autoregulatory dysfunction and cerebromicrovascular injury in mice with angiotensin II-induced hypertension”. In: *J. Cereb. Blood Flow Metab.* 33 (2013), pages 1732–1742.

[Att+16] D. Attwell et al. “What is a pericyte?” In: *Journal of Cerebral Blood Flow and Metabolism* 36.2 (2016), pages 451–455.

[Bag05] T. Baglin. “The measurement and application of thrombin generation”. In: *British Journal of Haematology* 130 (2005), pages 653–661.

[Ber10] M. N. Berberan-Santos. “Exact and approximate solutions for a century-old problem: A general treatment of Henri-Michaelis-Menten enzyme kinetics”. In: (2010).

[BF04] K. Berg-Sørensen and H. Flyvbjerg. “Power spectrum analysis for optical tweezers”. In: *Review of Scientific Instruments* 75.3 (2004), pages 594–612.

[BG96] G. A. Baker and P. Graves-Morris. “Padé Approximants”. In: *Cambridge U.P.* (1996).

[BGG03] S. D. Bungay, P. A. Gentry, and R. D. Gentry. “A mathematical model of lipid-mediated thrombin generation”. In: *Mathematical Medicine and Biology* 20 (2003), pages 105–109.

[BOG15] S. Bloch, D. Obari, and H. Girouard. “Angiotensin and neurovascular coupling: beyond hypertension”. In: *Microcirculation* 22 (2015), pages 159–167.

[Bou] M. Bourne. *Graphs on Logarithmic and Semi-Logarithmic Paper*. http://www.intmath.com/exponential-logarithmic-functions/7-graphs-log-semilog.php.

[Bru+05] K. Brummel-Ziedins et al. “The plasma hemostatic proteome: thrombin generation in healthy individuals”. In: *Journal of Haemostasis and Thrombosis* 3.7 (2005), pages 1472–1481.

[Bru+14] K. E. Brummel-Ziedins et al. “Modeling thrombin generation: plasma composition based approach”. In: *Journal of Thrombolysis* 37 (2014), pages 32–44.

[BV] Thrombinoscope BV. http://www.thrombinoscope.com/.

[C73] Rouget C. “Memoire sur le developpement, la structure et les proprietes physiologiques des capillaires sanguins et lymphatiques”. In: *Arch. Physiol. Normale Pathol.* 5 (1873), pages 603–661.

[C79] Rouget C. “Sur la contractilit6 des capillaires sanguins”. In: *C. R. Acad. Sci.* 88 (1879), pages 916–918.

[Cai+17] C. Cai et al. “Stimulation-induced rises in cerebral blood flow and local capillary vasoconstriction depend on conducted vascular responses in brain capillaries”. In: *TBA* TBA (2017), TBA.

[Cal+13] N. Calcinaghi et al. “Multimodal Imaging in Rats Reveals Impaired Neurovascular Coupling in Sustained Hypertension”. In: *Stroke* 44 (2013), pages 1957–1964.

[CDC16] CDC. *Hemophilia Facts*. https://www.cdc.gov/ncbddd/hemophilia/facts.html. 2016.

[Cha+10] M. S. Chatterjee et al. “Systems Biology of Coagulation Initiation: Kinetics of Thrombin Generation in Resting and Activated Human Blood”. In: *PLoS Computational Biology* 6.9 (2010), pages 1–24.

[Cha11] R. Chartrand. “Numerical Differentiation of Noisy, Nonsmooth Data”. In: *ISRN Applied Mathematics* (2011).

[Cha89] C. Chatfield. *The Analysis of Time Series: An Introduction*. New York, NY: Chapman & Hall, 1989.

[Cla] Aaron Clauset. *A brief primer on probability distributions*. http://tuvalu.santafe.edu/~aaronc/courses/7000/csci7000-001_2011_L0.pdf.

[CN08] D. L. Cox and M. M. Nelson. *Lehninger principles of biochemistry*. New York: W.H. Freeman, 2008.

[Cor+96] R. Corless et al. “On the Lambert W function”. In: *Advances in Computational Mathematics* 5 (1996), pages 329–359.

[Cor01] T. H. Cormen. *Introduction to Algorithms*. 2001.

[CS85] B. L. Coomber and P.A. Stewart. “Morphometric analysis of CNS microvascular endothelium”. In: *Microvascular Research* 30.1 (1985), pages 99–115.

[CV01] T. F. Chan and L. A. Vese. “Active Contours without Edges”. In: *IEEE Transactions On Image Processing* 10.2 (2001), pages 266–277.

[Dan90] W. W. Daniel. “Kolmogorov–Smirnov one-sample test”. In: *Applied Nonparametric Statistics* (1990), pages 319–330.

[DD87] S. J. Degen and E. W. Davie. “Nucleotide sequence of the gene for human prothrombin”. In: *Biochemistry* 26.19 (1987), pages 6165–77.

[dic] Merriam-Webster dictionary. https://www.merriam-webster.com/.

[Dor] W.A. Newman Dorland. https://www.dorlands.com/index.jsp.

[DP] Centers for Disease Control and Prevention. *High Blood Pressure Fact Sheet*. https://www.cdc.gov/dhdsp/data_statistics/fact_sheets/fs_bloodpressure.htm.

[Eng08] S. Engelberg. *Digital Signal Processing: An Experimental Approach*. Springer, 2008.

[Feu90] D. L. Feucht. *Handbook of Analog Circuit Design*. Biomathematics Texts. Elsevier Science, 1990. ISBN: 978-1-4832-5938-3.

[FF05] B. Furie and B. C. Furie. “Thrombus formation in vivo”. In: *J Clin Invest.* 115(12) (2005), pages 3355–3362.

[FF08] B. Furie and B. C. Furie. “Mechanisms of thrombus formation”. In: *New England Journal of Medicine* 359.9 (2008), pages 938–949.

[FF15] B. Fornberg and N. Flyer. *Brief Summary of Finite Difference Methods*. Society for Industrial and Applied Mathematics, 2015.

[Fil+16] J. A. Filosa et al. “Beyond Neurovascular Coupling, Role Of Astrocytes In The Regulation Of Vascular Tone”. In: *Neuroscience* 323 (2016), pages 96–109.

[Fis+81] A. M. Fischer et al. “Respective roles of antithrombin III and alpha 2 macroglobulin in thrombin inactivation”. In: *Thrombosis and Haemostasis* 45.1 (1981), pages 51–54.

[FL96] M. Fabricius and M. Lauritzen. “Laser-Doppler evaluation of rat brain microcirculation: comparison with the [14C]-iodoantipyrine method suggests discordance during cerebral blood flow increases”. In: *J Cereb Blood Flow Metab.* 16.1 (1996), pages 156–161.

[Gel05] A. Gelman. “Analysis of variance? Why it is more important than ever”. In: *The Annals of Statistics* 33 (2005), pages 1–53.

[Gie12] P. Giesen. *The Thrombogram Guide*. Thrombinoscope BV, 2012.

[GJ79] M. Gouault-Helimann and F. Josso. “Initiation in vivo of blood coagulation. The role of white blood cells and tissue factor”. In: *Nouv Presse Med.* 8.40 (1979), pages 3249–3253.

[GW64] C.M. Guldberg and P. Waage. “Studies Concerning Affinity”. In: *C. M. Forhandlinger: Videnskabs-Selskabet i Christiana* (1864).

[Hal+14] C. N. Hall et al. “Capillary pericytes regulate cerebral blood flow in health and disease”. In: *Nature* 508 (2014), pages 55–60.

[Hat08] C. Hatton. “Haematology (Lecture Notes)”. In: *Cambridge, MA: Blackwell Publishers* (2008), pages 145–166.

[Haz01] M. Hazewinkel. *Rational function, Encyclopedia of Mathematics*. Springer Science and Business Media B.V. / Kluwer Academic Publishers, 2001.

[HB95] H. C. Hemker and S. Beguin. “Thrombin Generation in Plasma: Its Assessment Via the Endogeneous Thrombin Potential”. In: *Thrombosis and Haemostasis* 74.1 (1995), pages 134–138.

[HCS58] Ph. D. Hall, E. Cecil, and H. S. Slayter. “The Fibrinogen Molecule: Its Size, Shape, and Mode of Polymerization”. In: *The Journal of Biophysical and Biochemical Cytology* 5.1 (1958), pages 11–16.

[HD05] Fritjof Helmchen and Winfried Denk. “Deep tissue two-photon microscopy”. In: *Nature Methods* 2.12 (2005), pages 932–940.

[Hea] Healthline. http://www.healthline.com/symptom/hemorrhage.

[Hed02] L. Hedstrom. “Serine protease mechanism and specificity”. In: *Chem. Rev.* 102.12 (2002), pages 4501–4524.

[Hem+02] H. C. Hemker et al. “The Calibrated Automated Thrombogram (CAT): a universal routine test for hyper-and hypocoagulability”. In: *Pathophysiology of Haemostasis and Thrombosis* 32 (2002), pages 249–253.

[Hem+03] H. C. Hemker et al. “Calibrated Automated Thrombin Generation Measurement in Clotting Plasma”. In: *Pathophysiology of Haemostasis and Thrombosis* 33 (2003), pages 4–15.

[Hem+05] H. C. Hemker et al. “Thrombin generation, a function test of the haemostatic-thrombotic system”. In: *Thromb Haemost* 96 (2005), pages 553–61.

[Hem+67] H. C. Hemker et al. “Kinetic aspects of the interaction of blood clotting enzymes II. The relation between clotting time and plasma concentration in prothrombin-time estimations”. In: *Thrombosis et diathesis haemorrhagica* 17.31 (1967), pages 349–357.

[Hem+71] H. C. Hemker et al. “The Evaluation of the Two-Stage Prothrombin Assay”. In: *Thrombosis et diathesis haemorrhagica* 25.3 (1971), pages 545–554.

[Hem+93] H. C. Hemker et al. “Continuous Registration of Thrombin Generation in Plasma, Its Use for the Determination of the Thrombin Potential”. In: *Thrombosis and Haemostasis* 70.4 (1993), pages 617–624.

[Hem15] H. C. Hemker. “Comment on Pike et. al. Thrombin generation: biochemical possibilities and clinical reality”. In: *Thrombosis and Hemostasis* 126.3 (2015), pages 288–289.

[HH69] H. C. Hemker and P. W. Hemker. "H. The kinetics of enzyme cascade systems. General kinetics of enzyme cascades". In: *Proc. R. Soc. Lond. B. Biol. Sci.* 173.32 (1969), pages 411–420.

[HHD09] H. C. Hemker, P. W. Hemker, and R. A. Dieri. "The technique of measuring thrombin generation with fluorescent substrates: IV. The H-transform, a mathematical procedure to obtain thrombin concentrations without external calibration". In: *Thrombosis and Haemostasis* 101 (2009), pages 171–177.

[HHL65] H. C. Hemker, P. W. Hemker, and E. A. Loeliger. "Kinetic aspects of the interaction of blood clotting enzymes I. Derivation of basic formulas". In: *Thrombosis et diathesis haemorrhagica* 15.13 (1965), pages 155–175.

[Hil+15] R. A. Hill et al. "Regional Blood Flow in the Normal and Ischemic Brain Is Controlled by Arteriolar Smooth Muscle Cell Contractility and Not by Capillary Pericytes". In: *Neuron* 87 (2015), pages 95–110.

[HJB85] M. T. Heideman, D. H. Johnson, and C. S. Burrus. "Gauss and the history of the fast Fourier transform". In: *Archive for History of Exact Sciences* 34.3 (1985), pages 265–277.

[HK13] H. C. Hemker and R. M. W. Kremers. "Data management in Thrombin Generation". In: *Thrombosis Research* 131 (2013), pages 3–11.

[HK67] H. C. Hemker and M. J. P. Kahn. "Reaction Sequence of Blood Coagulation". In: *Nature* 215 (1967), pages 1201–1202.

[HKK12] H. C. Hemker, S. Kerdelo, and R. M. W. Kremers. "Is there value in kinetic modeling of thrombin generation? No (unless...)" In: *Journal of Thrombosis and Haemostasis* 10 (2012), pages 1470–1477.

[Hoc+02] M. F. Hockin et al. "A Model for the Stoichiometric Regulation of Blood Coagulation". In: *The Journal of Biological Chemistry* 277 (2002), pages 18322–18333.

[Hof02] A. V. Hoffbrand. "Essential Haematology". In: *Oxford: Blackwell Science* (2002), pages 241–243.

[Hol08] Pavel Holoborodko. *Smooth Noise Robust Differentiators*. http://www.holoborodko.com methods/numerical-derivative/smooth-low-noise-differentiators/. 2008.

[HWB86] H. C. Hemker, G. M. Willems, and S. Beguin. "A Computer Assisted Method to Obtain the Prothrombin Activation Velocity in Whole Plasma Independent of Thrombin Decay Processes". In: *Thrombosis and Haemostasis* 56.1 (1986), pages 9–17.

[IUP] Compendium of Chemical Terminology IUPAC. http://goldbook.iupac.org/.

[J71] Eberth C. J. "Handbuch der Lehre von der Gewegen des Menschen und der Tiere." In: *Leipzig* 1 (1871).

[JM94] K. C. Jones and K. G. Mann. “A Model for the Tissue Factor Pathway to Thrombin II. A Mathematical Study.” In: *The Journal of Biological Chemistry* 269.37 (1994), pages 23367–23373.

[K23] Zimmerman K. “Die Feinere Bau der Blutcapillaren”. In: *Z. Anat. Entwickl.* 68 (1923), pages 29–109.

[Kaz+03] K. Kazama et al. “Angiotensin II attenuates functional hyperemia in the mouse somatosensory cortex”. In: *Am J Physiol Heart Circ Physiol* 285 (2003), H1890–H1899.

[Keh03] B. E. Kehrel. “Blood platelets: biochemistry and physiology”. In: *Hamostaseologie* 23.4 (2003), pages 149–158.

[Kis+17a] K. Kisler et al. “Cerebral blood flow regulation and neurovascular dysfunction in Alzheimer disease”. In: *Nature Reviews Neuroscience* 18 (2017), pages 419–434.

[Kis+17b] K. Kisler et al. “Pericyte degeneration leads to neurovascular uncoupling and limits oxygen supply to brain”. In: *Nature Neuroscience* 20 (2017), pages 406–416.

[KN14] T. E. Kornfield and E. A. Newman. “Regulation of Blood Flow in the Retinal Trilaminar Vascular Network”. In: *The Journal of Neuroscience* 34.34 (2014), pages 11504–11513.

[Kon78] H. A. et al. Kontos. “Responses of cerebral arteries and arterioles to acute hypotension and hypertension”. In: *Am J Physiol* 234 (1978), pages 371–383.

[KS89] M. A. Khanin and V. V. Semenov. “A Mathematical Model of the Kinetics of Blood Coagulation”. In: *J. Theor. Biol.* 136 (1989), pages 127–134.

[Lak72] K. Laki. “Our ancient heritage in blood clotting and some of its consequences”. In: *Annals of the New York Academy of Sciences* 202 (1972), pages 297–307.

[Lan15] M. D. Lance. “A general review of major global coagulation assays: thrombelastography, thrombin generation test and clot waveform analysis”. In: *Thrombosis Journal* 13.1 (2015).

[Law+94] J. H. Lawson et al. “A Model for the Tissue Factor Pathway to Thrombin I. An Empirical Study.” In: *The Journal of Biological Chemistry* 269.37 (1994), pages 23357–23366.

[LDD05] K. Lo, W. S. Denney, and S. L. Diamond. “Stochastic Modeling of Blood Coagulation Initiation”. In: *Pathophysiol Haemost Thromb* 34 (2005), pages 80–90.

[Lil+09] D. Lillicrap et al. *Practical Hemostasis and Thrombosis*. Wiley-Blackwell, 2009.

[LL13] Michaelis L. and Menten N. L. “Die Kinetik der Invertinwirkung”. In: *Biochem Z* 49 (1913), pages 333–369.

[Lon+16] A. T. Long et al. “Contact system revisited: An interface between inflammation, coagulation, and innate immunity”. In: *Journal of Thrombosis and Haemostasis* 14(3).3 (2016), pages 427–437.

[LVD93] U. Lindauer, A. Villringer, and U. Dirnagl. “Characterization of CBF response to somatosensory stimulation: model and influence of anesthetics.” In: *Am J Physiol.* 264.4 (1993), pages 1223–1228.

[M+06] Peppiatt C. M. et al. “Bidirectional Control of CNS Capillary Diameter by Pericytes”. In: *Nature* 443.7112 (2006), pages 700–704.

[Mac64] R. G. Macfarlane. “An Enzyme Cascade in the Blood Clotting Mechanism, and its Function as a Biochemical Amplifier”. In: *Nature* 202 (1964), pages 498–499.

[Man+07] A. G. Mann et al. “The association between drinking water turbidity and gastrointestinal illness: a systematic review”. In: *BMC Public Health* 7.256 (2007), pages 1–7.

[Mar] Dmitriy Markovich. https://dmitriymarkovich.shinyapps.io/Thrombin_Analyzer/.

[Mar+15] L. Y. Di Marcoa et al. “Vascular dysfunction in the pathogenesis of Alzheimer’s disease—Areview of endothelium-mediated mechanisms andensuing vicious circles”. In: *Neurobiology of Disease* 82 (2015), pages 593–606.

[Mat+10] T. M. Mathiisen et al. “The Perivascular Astroglial Sheath Provides a CompleteCovering of the Brain Microvessels: An ElectronMicroscopic 3D Reconstruction”. In: *Glia* 58 (2010), pages 1094–1103.

[MB53] R. G. Macfarlane and R. Biggs. “A Thrombin Generation Test: The Application in Haemophilia and Thrombocytopenia”. In: *J Clin Pathol.* 6.1 (1953), pages 3–8.

[MCA97] T. Matsugi, Q. Chen, and D. R. Anderson. “Contractile Responses of Cultured Bovine Retinal Pericytes to Angiotensin II”. In: *Arch Ophthalmol.* 115.10 (1997), pages 1281–1285.

[McC] K. McClaning. *Radio Receiver Design.*

[MGI] MGI. http://www.informatics.jax.org/allele/MGI:3796063.

[MHR02] D. M. Monroe, M. Hoffman, and H. R. Roberts. “Platelets and Thrombin Generation”. In: *Arteriosclerosis, Thrombosis, and Vascular Biology* 22 (2002), pages 1381–1389.

[MI10] Krueger M. and Bechmann I. “CNS pericytes: concepts, misconcep- tions, and a way out”. In: *Glia* 58 (2010), pages 1–10.

[MTA02] M. E. Maragoudakis, N. E. Tsopanoglou, and P. Andriopoulou. "Mechanism of thrombin-induced angiogenesis". In: *Biochemical Society Transaction* 30.2 (2002), pages 173–177.

[Mur93] J. D. Murray. *Mathematical biology*. Biomathematics Texts. Springer, 1993. ISBN: 3-540-57204-X.

[NHL] NHLBI. https://www.nhlbi.nih.gov/health/health-topics/topics/hemophilia.

[NHL16] NHLBI. *How Is Hemophilia Diagnosed?* https://www.nhlbi.nih.gov/health/health-topics/topics/hemophilia/diagnosis. 2016.

[OHe+16] P. O'Herron et al. "Neural correlates of single-vessel haemodynamic responses in vivo". In: *Nature* 534 (2016), pages 378–394.

[PD07] M. O. Palmier and S. R. Van Doren. "Rapid determination of enzyme cascade kinetics from fluorescence: Overcoming the inner filter effect". In: *Analytical Biochemistry* 371 (2007), pages 43–51.

[PD53] W. R. Pitney and J. V. Dacie. "A Simple Method of Studying the Generation of Thrombin in Recalcified Plasma: Application in the Investigation of Haemophilia". In: *J Clin Pathol.* 6.1 (1953), pages 9–14.

[Per] PerkinElmer. http://www.perkinelmer.com/category/plate-readers.

[PGA14] K. P. Papadoupoulos, M. Gavaises, and C. Atkin. "A simplified mathematical model for thrombin generation". In: *Medical Engineering and Physics* 36 (2014), pages 196–204.

[Pir+13] P. W. Pires et al. "The effects of hypertension on the cerebral circulation". In: *Am J Physiol Hear. Circ Physiol* 304 (2013), pages 1598–1614.

[Pre+07] W. H. Press et al. *Numerical Recipes*. Cambridge University Press, 2007.

[Pub] Pubchem. *Angiotensin II*. https://pubchem.ncbi.nlm.nih.gov/compound/angiotensin_ii.

[PW10] C. J. Pallister and M. S. Watson. "Haematology". In: *Scion Publishing* (2010), pages 336–347.

[QLZ05] Y. H. Qiao, J. L. Liu, and Y. J. Zeng. "A kinetic model for simulation of blood coagulation and inhibition in the intrinsic path". In: *Journal of Medical Engineering and Technology* 29.2 (2005), pages 70–74.

[R C15] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2015. URL: https://www.R-project.org/.

[R74] Weibel E. R. "On pericytes, particularly their existence on lung capillaries". In: *Microvasc. Res.* 8 (1974), pages 218–235.

[SBT02] L. Stryer, J. M. Berg, and J. L. Tymoczko. *Biochemistry*. San Francisco: W.H. Freeman, 2002.

[Sch72] A. Schmidt. “Neue Untersuchungen ueber die Fasserstoffesgerinnung”. In: *Pflueger’s Archiv fuer die gesamte Physiologie* 6 (1872), pages 413–538.

[Sea67] H. L. Seal. “The historical development of the Gauss linear model”. In: *Biometrika* 54.1/2 (1967), pages 1–24.

[Sha49] C. E. Shannon. “Communication in the presence of noise”. In: *Proc. Institute of Radio Engineers* (1949).

[She00] R. W. Shepard. “Pharmacology: antiplatelet and antithrombin therapy in acute coronary syndromes”. In: *J Cardiovasc Nurs.* 15.1 (2000), pages 54–61.

[Shu88] R. H. Shumway. *Applied statistical time series analysis*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[Sil14] L. Silvestri. *Saunders comprehensive review for the NCLEX-RN examination*. St. Louis, MO: Elsevier, 2014.

[Sim86] D. E. Sims. “The Pericyte—a Review”. In: *Tissue and Cell* 18.2 (1986), pages 153–174.

[Sla14] J. M. W. Slack. “Genes — a very short introduction”. In: *Oxford University Press* (2014).

[SM97] S. Schnell and C. Mendoza. “A closed form solution for time-dependent enzyme kinetics”. In: *J. Theor. Biol.* 187.207 (1997).

[Sme07] E. de Smedt. *Advanced Thrombinoscopy*. 2007.

[Smi09] S. A. Smith. “The cell-based model of coagulation”. In: *Journal of Veterinary Emergency and Critical Care* 19.1 (2009), pages 3–10.

[Smi16] P. A. Smith. “Intra-articular Autologous Conditioned Plasma Injections Provide Safe and Efficacious Treatment for Knee Osteoarthritis: An FDA-Sanctioned, Randomized, Double-blind, Placebo-controlled Clinical Trial”. In: *The American Journal of Sports Medicine* 44.4 (2016), pages 378–394.

[Sot+84] L. Sottrup-Jensen et al. “Primary structure of human alpha 2-macroglobulin. V. The complete structure”. In: *J Biol Chem.* 259.13 (1984), pages 8318–8327.

[SPS10] Karline Soetaert, Thomas Petzoldt, and R. Woodrow Setzer. “Solving Differential Equations in R: Package deSolve”. In: *Journal of Statistical Software* 33.9 (2010), pages 1–25. ISSN: 1548-7660. DOI: 10.18637/jss.v033.i09. URL: http://www.jstatsoft.org/v33/i09.

[SS01] L. G. Shapiro and G. C. Stockman. *Computer Vision*. Prentice Hall, 2001.

[SW65] S. S. Shapiro and M. B. Wilk. “An analysis of variance test for normality (complete samples)”. In: *Biometrika* 52.3-4 (1965), pages 591–611.

[Tec] Technoclone. http://www.technoclone.com/welcome.

[TS83] J. Travis and G. S. Salvesen. “Human Plasma Protein Inhibitors”. In: *Annual Reviews in Biochemistry* 52 (1983), pages 655–709.

[Wag+10] R. J. Wagenwoord et al. “The paradoxical stimulation by a reversible thrombin inhibitor of thrombin generation in plasma measured with thrombinography is caused by α_2-macroglobulin-thrombin”. In: *Journal of Thrombosis and Haemostasis* 8 (2010), pages 1281–1289.

[WHH06] R. Wagenvoord, P. W. Hemker, and H. C. Hemker. “The limits of Simulation of the Clotting System”. In: *Journal of Thrombosis and Haemostasis* 4 (2006), pages 1331–1338.

[Wil+91] G. M. Willems et al. “Simulation Model for Thrombin Generation in Plasma”. In: *Haemostasis* 21 (1991), pages 197–207.

[Woo+13] S. A. Woodle et al. “Determining the impact of instrument variation and automated software algorithms on the TGT in hemophilia and normalized plasma”. In: *Thrombosis Research* 132 (2013), pages 374–380.

Published by:
DTU Nanotech
Department of Micro- and Nanotechnology
Technical University of Denmark
Ørsteds Plads, building 345C
DK-2800 Kgs. Lyngby