

A theory of reactant-stationary kinetics for a mechanism of zymogen activation

Justin Eilertsen^a, Wylie Stroberg^a, Santiago Schnell^{a,b,c,1}

^a*Department of Molecular & Integrative Physiology, University of Michigan Medical School, Ann Arbor, MI 48109, USA*

^b*Department of Computational Medicine & Bioinformatics, University of Michigan Medical School, Ann Arbor, MI 48109, USA*

^c*Brehm Center for Diabetes Research, University of Michigan Medical School, Ann Arbor, MI 48105, USA*

Abstract

A theoretical analysis is performed on the nonlinear ordinary differential equations that govern the dynamics of a reaction mechanism of zymogen activation. The reaction consists of a primary non-observable zymogen activation reaction that it is coupled to an indicator (observable) reaction. The product of the first reaction is the enzyme of the indicator reaction, and both reactions are governed by the Michaelis–Menten reaction mechanism. Using singular perturbation methods, we derive asymptotic solutions that are valid under the quasi-steady-state and reactant-stationary assumptions. In particular, we obtain closed form solutions that are analogous to the Schnell–Mendoza equation for Michaelis–Menten type reactions. These closed-form solutions approximate the evolution of the observable reaction and provide the mathematical link necessary to measure the enzyme activity of the non-observable reaction. Conditions for the validity of the asymptotic solutions are also derived, and we demonstrate that these asymptotic expressions are applicable under reactant-stationary kinetics.

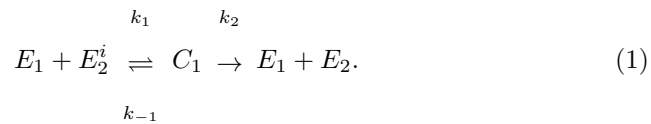
Keywords: Coupled enzyme assay, time course experiments, timescale separation analysis, singular perturbation analysis, Schnell–Mendoza equation,

Email address: schnells@umich.edu (Santiago Schnell)

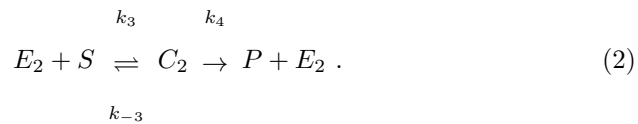
¹Corresponding author

1. Introduction

Many enzyme catalyzed reactions that occur in physiological processes require an activation step in which a precursor of a zymogen (inactive enzyme precursor or pro-enzyme) is converted to an active enzyme. This process, known generally as zymogen activation [1], is typically the first step in a cascade of coupled enzyme catalyzed reactions [2]. The activation step of the zymogen is itself an enzyme catalyzed reaction, and the inactive enzyme precursor is activated by a functional enzyme. The active enzyme can be generated by enzyme-catalyzed proteolysis or enzyme activation by phosphorylation [3]. For example, the digestive enzyme trypsin, which is the activate form of trypsinogen, is activated by the enzyme enterokinase; trypsin can then bind with trypsinogen to convert remaining trypsinogen into trypsin [2]. Likewise, plasminogen is activated by streptokinase to form plasmin (an enzyme), which then degrades fibrin (a substrate) to break down clots in blood coagulation [4]. Denoting the *active* enzyme, zymogen, *activated* enzyme, and intermediate complex of the activation reaction as E_1 , E_2^i , E_2 , and C_1 respectively, the preliminary zymogen activation step coupled with its secondary enzyme-catalyzed reaction can be expressed with the following reaction mechanism:

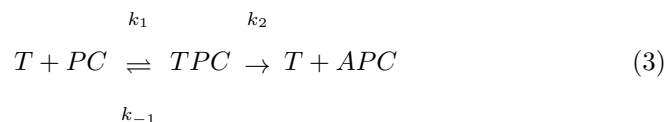


Regardless of the reactants, the zymogen activation step simply produces E_2 . The secondary reaction occurs when E_2 and substrate S bind to synthesize the final product P :

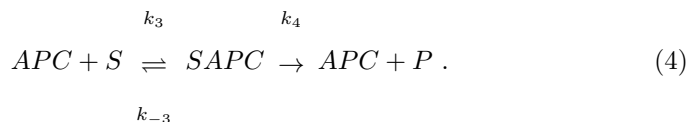


In the above chemical pathways, k_1 , k_{-1} , k_3 , k_{-3} are rate constants, and k_2 , k_4 are catalytic constants.

As mentioned previously, the reaction mechanism of zymogen activation (1)–(2) occurs naturally in coagulation cascades [5]. As a distinct example, the activation of protein C (PC) by thrombin (T) follows a reaction consistent with (1):



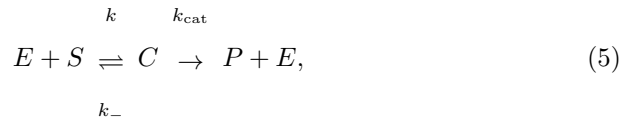
where “ APC ” denotes the activated form of PC . Assuming S is specific to APC and does not bind with T , the secondary observable reaction follows the form of (2):



Another interesting aspect of coupled enzyme catalyzed reactions with a
 5 zymogen activation step (1)–(2) is the quantification of the catalytic conversion
 of zymogen in vitro. Formally, the quantification of enzyme activity through
 measurements obtained from an in vitro assay is mathematically known as an
inverse problem. If the activation step (1) is not detectable experimentally (i.e.,
 non-observable), then the secondary reaction step (2) is selected to be an easily
 10 observable reaction. This is done with the goal of measuring the enzyme activity
 of the non-observable reaction through analysis of progress curves generated by
 the observable reaction. In this case, the secondary reaction step (2) is known
 as the *indicator* reaction. Traditionally, coupled enzyme assays are designed so
 that the product of the non-observable reaction is a substrate for the secondary
 15 enzyme in the indicator reaction (see [6] for specific applications). While this
 type of assay is well-studied [7, 8, 9, 10], in vitro assays that consist of a zymogen
 activation step have not been analyzed with the same degree of interest.

The kinetics of the non-observable zymogen activation step is measured by decoupling the analysis of progress curves by adding excessive concentrations
 20 E_1 . The assumption is then made that the first reaction (the activation step) is pseudo-first order (PFO) [5]. However, it has been demonstrated that an excessive concentration of E_1 is not sufficient to guarantee the validity of PFO model. Instead, it is necessary that initial concentration of zymogen for (1) be
 25 much less than the Michaelis constant of the primary reaction [11]. Thus, from an experimental point of view, it is difficult to ensure the validity of the PFO model when the Michaelis constant is unknown.

Since PFO models are difficult to validate when the Michaelis constant is unknown, it is more convenient to rely on the quasi-steady-state (QSS) models when quantifying enzyme activity in vitro. If appropriate experimental conditions are employed, then the MM reaction mechanism (5),



will obey QSS kinetics, and the rate of substrate depletion for the reaction is described by the MM equation

$$\dot{s} = -\frac{V}{K_M + s}s, \quad (6)$$

where s is the concentration of S , $K_M \equiv (k_- + k_{\text{cat}})/k$ is the Michaelis constant,
 30 and $V \equiv k_{\text{cat}}e_0$ is the limiting rate of the reaction (5). Note that the zymogen activation step in (1)–(2) is a single-enzyme, single-substrate reaction. Once the QSS model is established, the inverse problem is carried out in two stages. First, experimental data is produced in the form of a progress curve for either
 35 s or p (we have used lower case letters to denote the concentrations of S and P respectively). Second, the experimental data is then used to estimate both K_M and V by optimally *fitting* the model (6) through the utilization of either a deterministic (i.e., such as Levenburg-Marquardt) or a stochastic (Markov Chain Monte Carlo) algorithm. In general, one seeks to estimate kinetic constants with

an expression that contains the fewest number of parameters: this is why the
40 MM equation is more attractive than the complete set of mass action equations.
The MM equation is known as a *reduced model*; it is *reduced* in the sense that
it contains fewer variables (s versus s and c) and fewer parameters (K_M and V
versus k , k_- and k_{cat}).

The inverse problem presents a unique challenge for both experimentalists
45 and theorists in coupled enzyme assays like (1)-(2). First, the parameters that
govern the enzyme activity of the non-observable reaction must somehow be
determined from the indicator reaction, since progress curves from a typical in
vitro laboratory experiment can only be generated for the indicator reaction.
Second, a reduced model for the model mechanism of zymogen activation (1)-
50 (2) must be developed. The reduced model should: (1) decrease the number of
variables, and (2) lessen the number of parameters needed to describe the time
course of the reaction mechanism (1)-(2).

1.1. Goals of this paper

The primary goal of this paper is the derivation of a reduced model that
55 can be utilized to quantify the enzyme activity for an experimental assay of the
model mechanism of zymogen activation (1)-(2). Central to the derivation will
be the application of *slow manifold projection*. This is challenging for coupled
reactions, since the time to completion of the indicator reaction can occur be-
fore, after, or at approximately the same time as the non-observable reaction.
60 Furthermore, it is unlikely that the relative speeds and completion time of the
non-observable reaction will be known. Thus, there is a need derive a reduced
model that is general enough so that its validity is certain regardless of which re-
action is faster. Finally, we will seek a model that admits a closed form solution.
This will eliminate the need to generate *explicit* progress curves for substrate
65 depletion of the primary reaction since the time course of substrate (i.e., E_2^i) is
unknown in coupled enzyme assays.

1.2. Structure of this paper

As mentioned previously, the theoretical reduction analysis of zymogen activation reactions has been limited to PFO models [12, 1, 13, 4, 2]. Such models
70 have limited validity in time course experiments [11], and the aim of this work is first and foremost to take a necessary “first step” in the nonlinear analysis of such reactions. First, we will introduce proper scaling techniques that can be employed in a general methodology to more complicated reactions. In Section 3 we will show how to estimate timescales based on these scaling methods,
75 and we will formulate a reduced model from the analysis of these timescales (Section 4). The reduced model admits closed-form solutions in the form of a Schnell–Mendoza equation [14], and conditions for the validity of the model will be established. In addition, we will exploit the geometry of the mathematical structure [15, 16] in extreme situations when the speeds of the reactions are
80 significantly disparate. This will allow us to “simplify” the reduced model and obtain asymptotic solutions that are essentially less complicated (in form) than both the general reduced model and the system of mass action equations. Finally, in Section 7, we conclude with a brief discussion of the results and their relevance in possible future work involving the inverse problem.

85 2. Derivation of the governing equations for the reaction mechanism of zymogen activation (1)–(2)

We first consider the mass action formulation of the zymogen activation reaction mechanism (1)–(2). In reaction (1), the zymogen E_2^i is effectively a substrate. To distinguish mathematically between substrates and enzymes in (1)–(2), we will change notation by replacing E_2^i with S_1 in (1), and S with S_2

in (2). Applying the law of mass action yields seven rate equations

$$\dot{e}_1 = -k_1 e_1 s_1 + (k_{-1} + k_2) c_1, \quad (7a)$$

$$\dot{s}_1 = -k_1 e_1 s_1 + k_{-1} c_1, \quad (7b)$$

$$\dot{c}_1 = k_1 e_1 s_1 - (k_{-1} + k_2) c_1, \quad (7c)$$

$$\dot{e}_2 = k_2 c_1 - k_3 e_2 s_2 + (k_{-3} + k_4) c_2, \quad (7d)$$

$$\dot{s}_2 = -k_3 e_2 s_2 + k_{-3} c_2, \quad (7e)$$

$$\dot{c}_2 = k_3 e_2 s_2 - (k_{-3} + k_4) c_2, \quad (7f)$$

$$\dot{p} = k_4 c_2, \quad (7g)$$

where lowercase letters represent concentrations of the corresponding uppercase species. Typically, laboratory enzyme assays present the following initial conditions

$$(e_1, s_1, c_1, e_2, s_2, c_2, p) |_{t=0} = (e_1^0, s_1^0, 0, 0, s_2^0, 0, 0). \quad (8)$$

We will subsequently refer to (8) as *experimental initial conditions*. By examining the system of rate equations (7), the reaction mechanism of zymogen activation (1)–(2) obeys three conservation laws:

$$e_1(t) + c_1(t) = e_1^0, \quad (9a)$$

$$s_1(t) + c_1(t) + c_2(t) + e_2(t) = s_1^0, \quad (9b)$$

$$s_2(t) + c_2(t) + p(t) = s_2^0. \quad (9c)$$

The solution trajectory to (7) must lie on the intersection of the hyperplanes defined in (9). This implies the presence of conserved quantities which can be used to reduce the dimension of the problem. Using (9a) and (9b) to decouple the enzyme concentrations, the redundancies in the system (7) are eliminated

to yield

$$\dot{s}_1 = -k_1 (e_1^0 - c_1) s_1 + k_{-1} c_1, \quad (10a)$$

$$\dot{c}_1 = k_1 (e_1^0 - c_1) s_1 - (k_{-1} + k_2) c_1, \quad (10b)$$

$$\dot{s}_2 = -k_3 (s_1^0 - s_1 - c_1 - c_2) s_2 + k_{-3} c_2, \quad (10c)$$

$$\dot{c}_2 = k_3 (s_1^0 - s_1 - c_1 - c_2) s_2 - (k_{-3} + k_4) c_2, \quad (10d)$$

90 where $e_1(t)$, $e_2(t)$ and $p(t)$ are readily calculated once $s_1(t)$, $c_1(t)$, $s_2(t)$ and $c_2(t)$ are known.

3. Rate expressions for the non-observable enzyme catalyzed reaction

The rate equations (10a)–(10b) are uncoupled from (10c)–(10d), and have the same structure to those of the single-substrate, single-enzyme reaction that
 95 follows the MM mechanism. Therefore, it is possible to derive rate equations to model the reaction mechanism of zymogen activation (1)–(2), and estimate its kinetic parameters using the general theory of the reactant-stationary assumption (RSA, [17]).

3.1. Review of the single substrate, single enzyme MM reaction

100 It has long been established from the analysis of single-enzyme, single-substrate reactions that there there can be a rapid buildup of c_1 during an initial fast transient of the non-observable reaction. After the rapid buildup, c_1 is assumed to be in a QSS, and the rate of depletion of c_1 approximately equals its rate of formation:

$$\dot{c}_1 \approx 0 \quad \text{for } t > t_{c_1}. \quad (11)$$

105 The timescale t_{c_1} is the time associated with the initial transient buildup of c_1 , and is independent of the initial concentration of E_1 :

$$t_{c_1} = \frac{1}{k_1(K_{M_1} + s_1^0)}. \quad (12)$$

In the above equation, $K_{M_1} = (k_{-1} + k_2)/k_1$ is the Michaelis constant for the zymogen activation step (1). The quasi-steady-state assumption (QSSA, 11),

in combination with (10a)–(10b), leads to the derivation of the well-known rate expressions

$$c_1 = \frac{e_1^0}{K_{M_1} + s_1} s_1 \quad (13a)$$

$$\dot{s}_1 = -\frac{V_1}{K_{M_1} + s_1} s_1. \quad (13b)$$

In (13b), $V_1 \equiv k_2 e_1^0$ is the limiting rate of the zymogen reaction. Note that the mass action equations (10a)–(10b) are reduced to a differential-algebraic equation systems with a single differential equation for s_1 in (13a)–(13b).

110 Since equations (13a) and (13b) are only valid after the initial fast transient, t_{c_1} , it is necessary to define a boundary condition for s_1 at $t = t_{c_1}$. We will assume that there is a negligible decrease in s_1 during the initial buildup of c_1 . This is equivalent to the initial experimental condition for the initial rate or time course experiments. The assumption that the depletion of s_1 is negligible
115 over the fast transient is known as the RSA. Formally, the RSA is

$$s_1(t < t_{c_1}) \approx s_1^0. \quad (14)$$

The RSA provides an initial condition for (10a) under the variable transformation $\hat{t} \mapsto t - t_{c_1}$. The mathematical expression (13b) is the MM equation, and the system (13a)–(13b) governs the dynamics of the substrate s_1 and complex c_1 of the non-observable reaction under the QSS and RSA. The *explicit*
120 closed-form solution of (13b), with the initial condition (14), is known as the Schnell–Mendoza equation [14], and is written in terms of the Lambert- W function:

$$\frac{s_1(\hat{t})}{K_{M_1}} = W[\sigma_1 \exp(\sigma_1 - \eta_1 \hat{t})], \quad \sigma_1 = \frac{s_1^0}{K_{M_1}}, \quad \eta_1 = \frac{V_1}{K_{M_1}}. \quad (15)$$

Asymptotically, Schnell and Mendoza [14] have provided a piecewise solution for the MM reaction in terms of a fast transient solution for s_1 (16a) and a QSS solution for s_1 (16b):

$$s_1 \simeq s_1^0, \quad t \leq t_{c_1} \quad (16a)$$

$$s_1 \simeq K_{M_1} W[\sigma_1 \exp(\sigma_1 - \eta_1 \hat{t})], \quad t > t_{c_1}. \quad (16b)$$

In addition, from the earlier work of Segel [18], we have the corresponding approximation for c_1 :

$$c_1 \simeq \bar{c}_1 [1 - \exp(-t/t_{c_1})] \quad \text{with} \quad \bar{c}_1 = \frac{e_1^0}{K_{M_1} + s_1^0} s_1^0, \quad t < t_{c_1}, \quad (17a)$$

$$c_1 \simeq \frac{e_1^0}{K_{M_1} + s_1} s_1, \quad t \geq t_{c_1}. \quad (17b)$$

Collectively, equations (16a)–(17b) constitute an asymptotic solution that serves as an accurate approximation to the full time course of (10) when the appropriate qualifiers (i.e, the RSA and the QSSA) are obeyed.

The time it takes for the majority of the substrate s_1 to deplete is given by t_{s_1} . Although there are several methods for estimating the significant timescales of chemical reactions [19], we employ the heuristic method proposed by Segel [18], and approximate the depletion time to be effectively the total depletion of s_1 (the total depletion is s_1^0) divided by the maximum rate of substrate of depletion after t_{c_1} :

$$t_{s_1} = \frac{\Delta s_1}{\max_{0 \leq t} |\dot{s}_1|} = \frac{K_{M_1} + s_1^0}{V_1}. \quad (18)$$

Generally speaking, t_{s_1} is a reasonable measure of how long it takes for a significant amount of s_1 to deplete, although its precise interpretation depends on the magnitude of σ_1 .

3.2. Geometrical picture of the enzyme catalyzed reaction, and conditions for the validity of asymptotic solutions of the rate equations

While the asymptotic solutions are useful in that they can be employed to make certain predictions about the behavior of the reaction, asymptotic theory fails to yield a visual or geometric understanding of the dynamical behavior of the zymogen activation reaction mechanism (1)–(2). To paint a complete picture of the mathematical structure behind the reaction mechanism, we turn to dynamical systems theory, and analyze this problem from phase space. From this perspective, after the initial buildup of c_1 , the phase space trajectory of the

non-observable reaction (10a)–(10b) *hugs* a slow manifold, \mathcal{M}_ε , and is asymptotic to \mathcal{M}_ε in the approach to equilibrium. The time it takes for the trajectory to reach the slow manifold is approximately t_{c_1} , and the time it takes for the trajectory to equilibrium is approximately t_{s_1} . The condition for the validity of the asymptotic solution resides in *how well* the c_1 -nullcline approximates the slow manifold, and also *how straight* the phase space trajectory is in its approach to the slow manifold during the initial *fast* transient. The former of these conditions is known as the QSSA, and the latter is of course the geometrical interpretation of the RSA. We note that if the trajectory is close to \mathcal{M}_ε , then the complex C_1 is assumed to be in a QSS and the difference between the rate of C_2 depletion is approximately equal to the rate C_2 formation.

It was originally proposed that (16a)–(16b) was valid if $t_{c_1} \ll t_{s_1}$. However, although timescale separation is necessary, the validity of (16a)–(16b) is actually determined by the validity of the RSA. To determine the criteria for the validity of the RSA, Segel [20] proposed that if one assumes little change in s_1 during the approach to the slow manifold (an almost *straight* phase space trajectory towards the slow manifold), then it should hold that

$$\max_{t \geq 0} |\dot{s}_1| \cdot t_{c_1} \ll s_1^0. \quad (19)$$

Since $|\dot{s}_1| \leq k_1 s_1^0 e_1^0$, the strict inequality given in (19) translates to

$$\frac{e_1^0}{K_{M_1} + s_1^0} \equiv \varepsilon \ll 1. \quad (20)$$

Through scaling analysis, Segel [18] went on to show that the RSA determines single-handedly the validity of the asymptotic solutions² (16) and (17). Introducing the dimensionless variables $\hat{s}_1 = s_1/s_1^0$ and $\hat{c}_1 = c_1/\bar{c}_1$, Segel and Slemrod [20] demonstrated that, with respect to the dimensionless timescale

²Segel did not refer to the condition $\varepsilon \ll 1$ as the RSA [see 17, for more details].

$\tau = t/t_{c_1}$, equations (10a)-(10a) scale as

$$\frac{d\hat{s}_1}{d\tau} = \varepsilon \left[-\hat{s}_1 + \frac{\sigma_1}{\sigma_1 + 1} \hat{c}_1 \hat{s}_1 + \frac{\kappa_1(1 + \kappa_1)^{-1}}{\sigma_1 + 1} \hat{c}_1 \right], \quad (21a)$$

$$\frac{d\hat{c}_1}{d\tau} = \hat{s}_1 - \frac{\sigma_1}{\sigma_1 + 1} \hat{c}_1 \hat{s}_1 - \frac{1}{\sigma_1 + 1} \hat{c}_1, \quad (21b)$$

where $\kappa_1 = k_{-1}/k_2$. Moreover, (10a)-(10a) become

$$\frac{d\hat{s}_1}{dT} = (\kappa_1 + 1)(\sigma_1 + 1) \left[-\hat{s}_1 + \frac{\sigma_1}{\sigma_1 + 1} \hat{c}_1 \hat{s}_1 + \frac{\kappa_1(1 + \kappa_1)^{-1}}{\sigma_1 + 1} \hat{c}_1 \right], \quad (22a)$$

$$\varepsilon \frac{d\hat{c}_1}{dT} = (\kappa_1 + 1)(\sigma_1 + 1) \left[\hat{s}_1 - \frac{\sigma_1}{\sigma_1 + 1} \hat{c}_1 \hat{s}_1 - \frac{1}{\sigma_1 + 1} \hat{c}_1 \right], \quad (22b)$$

when the time is scaled with respect to the depletion timescale $T = t/t_{s_1}$. Thus, it is apparent from the dimensionless equations (21a)-(22b) that if $\varepsilon \ll 1$, then not only will the RSA hold, but the QSSA also holds. In fact the RSA (i.e., $\varepsilon \ll 1$) is more restrictive than separation of timescales. After some algebraic calculations, the separation of timescales ($t_{c_1}/t_{s_1} \ll 1$) is equivalently expressed as

$$\frac{e_1^0}{K_{M_1} + s_1^0} \ll \left(1 + \frac{K_{S_1}}{K_1} \right) \left(1 + \frac{s_1^0}{K_{M_1}} \right), \quad (23)$$

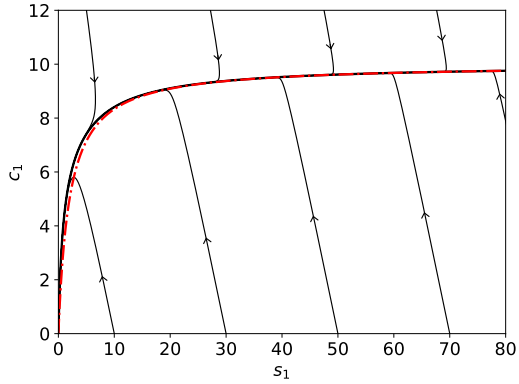
where $K_{S_1} = k_{-1}/k_1$, and $K_1 = k_2/k_1$. For the RSA to be valid, the condition

$$\frac{e_1^0}{K_{M_1}} \ll \left(1 + \frac{s_1^0}{K_{M_1}} \right), \quad (24)$$

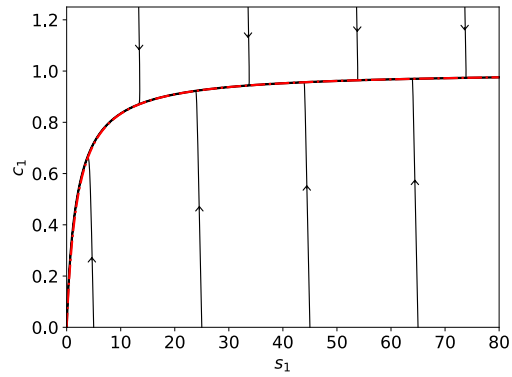
must be satisfied; this is more stringent than condition (23), and hence dictates the conditions under which equation (13b) or (15) can be applied. For this reason, the MM expressions are considered valid under the RSA (see FIGURES 1a and 1b) rather than the QSSA [21].

3.3. Scaling analysis of the indicator reaction

The scaling analysis of the indicator reaction requires knowledge of fast and slow timescales as well as knowledge of reasonable upper and lower bounds of s_2 and c_2 . We will start by trying to estimate a depletion timescale for the



(a)



(b)

Figure 1: Geometrical picture of the single-substrate, single-enzyme non-observable reaction (1) representing the zymogen activation step. (a) Phase space dynamics with $e_1^0 = 10$, $k_1 = 1$, $k_2 = 1$ and $k_{-1} = 1$. (b) Phase space dynamics with $e_1^0 = 1$, $s_1^0 = 78$, $k_1 = 1$, $k_2 = 5$ and $k_{-1} = 1$. As $\varepsilon \rightarrow 0$, the accumulation of c_1 is more rapid, and the c_1 -nullcline (dashed red curve) becomes a better approximation to the slow manifold, \mathcal{M}_ε , which is the thick black curve. The slow manifold curve is a graphical representation of the steady-state kinetic rate equation. The thin black curves are trajectories (numerical solutions of the mass action equations (10)) starting from different initial conditions, and represent the fast-transient kinetics of the reaction.

indicator reaction. An accurate depletion timescale should give us a reasonable estimation of the completion time for the indicator reaction. In the case of the reaction mechanism of zymogen activation (1)–(2), the completion of the indi-

180 cator reaction can be faster, as fast, or slower than the non-observable reaction. For the non-observable reaction, the depletion timescale is expressed in terms of s_1^0 , e_1^0 , and K_{M_1} :

$$t_{s_1} = \frac{K_{M_1} + s_1^0}{V_1}. \quad (25)$$

The quantity e_1^0 is the total amount of available enzyme for the non-observable reaction. The construction of a homologous depletion timescale for the indicator
185 reaction is problematic since the total amount of available enzyme e_2^A ,

$$e_2^A(t) = s_1^0 - s_1 - c_1, \quad (26)$$

is a time-dependent quantity. If we start by assuming the QSSA is valid, then the mass action equations reduce to

$$c_2 \simeq \frac{e_2^A(t)}{K_{M_2} + s_2} s_2, \quad (27a)$$

$$s_2 \simeq -\frac{V_2(t)}{K_{M_2} + s_2} s_2, \quad (27b)$$

where $V_2(t) \equiv k_4 e_2^A(t)$. The general solution to (27b) is given in terms of a Lambert- W function:

$$s_2 = K_{M_2} W \left[\sigma_2 e^{\sigma_2 - \int_0^t V_2(s) ds / K_{M_2}} \right]. \quad (28)$$

The term “ s ” in (28) has been employed as a dummy variable, and $\sigma_2 \equiv s_2^0 / K_{M_2}$. We will employ a mean-field approach to derive a depletion timescale for the
190 indicator reaction. Let us first assume that we know the depletion timescale for the indicator reaction; we will denote this timescale as T_{s_2} . The *mean* available enzyme over the time course of the indicator reaction is given by

$$\langle e_2^A \rangle = \frac{1}{T_{s_2}} \int_0^{T_{s_2}} e_2^A(t) dt. \quad (29)$$

If the completion of the indicator reaction occurs long before the completion of the non-observable reaction, then we expect $\langle e_2^A \rangle \ll s_1^0$. In contrast, if the

195 completion of the indicator reaction occurs long after the completion of the non-observable reaction, then we expect $\langle e_2^A \rangle \approx s_1^0$. In any case, we can define the depletion timescale as

$$T_{s_2} = \frac{K_{M_2} + s_2^0}{k_4 \langle e_2^A \rangle}, \quad (30)$$

which should yield a reasonable estimate for the slow timescale *if* the depletion of s_2 is influenced by a slow manifold. Note that $K_{M_2} \equiv (k_{-3} + k_4)/k_3$ is the
200 Michaelis constant of the indicator reaction.

Next, we want to scale the mass action equations that model the indicator reaction with respect to the quantities $\mathbb{T} = t/\hat{t}$, s_2^0 and $\max(e_2^A)$, where $\max(e_2^A)$ is the maximum amount of e_2^A over the course of the indicator reaction:

$$\max(e_2^A) \equiv \max_{t \leq T_{s_2}} (s_1^0 - s_1 - c_1). \quad (31)$$

Utilizing $\max(e_2^A)$ as an upper bound on the available enzyme dictates a natural
205 scaling of c_2 :

$$c_2 \leq \frac{\max(e_2^A)}{K_{M_2} + s_2^0} s_2^0 \equiv \hat{c}_2. \quad (32)$$

The remaining upper bounds provide us with the following ensemble of dimensionless variables,

$$\bar{s}_2 s_2^0 = s_2, \quad \bar{c}_2 \hat{c}_2 = c_2, \quad \bar{e}_2^A \max(e_2^A) = e_2^A, \quad \mathbb{T} \hat{t} = t, \quad (33)$$

where \hat{t} denotes an arbitrary timescale. Substitution of (33) into the mass action equation yields

$$\frac{d\bar{s}_2}{d\mathbb{T}} = \frac{\max(e_2^A)}{\langle e_2^A \rangle} \frac{\hat{t}}{T_{s_2}} (1 + \kappa_2)(1 + \sigma_2) \left[\left(\frac{\sigma_2}{1 + \sigma_2} \bar{c}_2 - \bar{e}_2^A \right) \bar{s}_2 + \frac{\alpha}{1 + \sigma_2} \bar{c}_2 \right], \quad (34a)$$

$$\lambda \frac{d\bar{c}_2}{d\mathbb{T}} = \frac{\max(e_2^A)}{\langle e_2^A \rangle} \frac{\hat{t}}{T_{s_2}} (1 + \kappa_2)(1 + \sigma_2) \left[\left(\bar{e}_2^A - \frac{\sigma_2}{1 + \sigma_2} \bar{c}_2 \right) \bar{s}_2 - \frac{1}{1 + \sigma_2} \bar{c}_2 \right]. \quad (34b)$$

In the above expressions, the dimensionless quantities σ_2 , κ_2 and α are:

$$\sigma_2 \equiv s_2^0/K_{M_2}, \quad \kappa_2 \equiv k_{-3}/k_4, \quad \alpha \equiv \kappa_2/(1 + \kappa_2). \quad (35)$$

The parameter λ is defined as

$$\lambda \equiv \frac{\max(e_2^A)}{K_{M_2} + s_2^0}, \quad (36)$$

210 and is unique in that if it is sufficiently small, then it mathematically characterizes the indicator reaction as a singularly perturbed differential equation for which model reduction is possible through means of projecting onto the slow manifold.

4. Asymptotic analysis of the reaction mechanism (1)–(2)

Now that we have a good idea as to how the mass action equations of the indicator reaction scale, we want to try and find closed-form asymptotic solutions to the mass action equations or, at the very least, try and reduce the dimension of the mass action differential equations. The exact form of the scaled mass action equations will depend on the slow timescales of both the observable and non-observable indicator reactions. Thus, given that the respective depletion timescales of the indicator and non-observable reactions are T_{s_2} and t_{s_1} , we will analyze

$$\frac{d\bar{s}_2}{dT} = \frac{\max(e_2^A)(1 + \kappa_2)(1 + \sigma_2)}{\langle e_2^A \rangle \delta_S} \left[\left(\frac{\sigma_2}{\sigma_2 + 1} \bar{c}_2 - \bar{e}_2^A \right) \bar{s}_2 + \frac{\alpha}{\sigma_2 + 1} \bar{c}_2 \right], \quad (37a)$$

$$\lambda \frac{d\bar{c}_2}{dT} = \frac{\max(e_2^A)(1 + \kappa_2)(1 + \sigma_2)}{\langle e_2^A \rangle \delta_S} \left[\left(\bar{e}_2^A - \frac{\sigma_2}{\sigma_2 + 1} \bar{c}_2 \right) \bar{s}_2 - \frac{1}{\sigma_2 + 1} \bar{c}_2 \right], \quad (37b)$$

215 where δ_S is the ratio of the substrate depletion timescales, $\delta_S \equiv T_{s_2}/t_{s_1}$, and $T = t/t_{s_1}$. Based on the scaling given in (37a) and (37b), we will derive an estimate for T_{s_2} as well as solutions for three particular cases: **Case 1:** the indicator reaction is faster than the non-observable reaction ($\delta_S \ll 1$). **Case 2:** the indicator reaction is roughly the same speed as the non-observable reaction
 220 ($\delta_S \approx 1$). **Case 3:** the indicator reaction is much slower than the non-observable reaction ($\delta_S \gg 1$).

4.1. Case 1: The indicator reaction is faster than the non-observable reaction
 $(\delta_S \ll 1)$

If the indicator reaction is fast, then the completion of the non-observable reaction will occur long after the completion of the indicator reaction, and the slow timescale is t_{s_1} . To start the analysis, we will rescale the mass action equations that govern the non-observable reaction with respect to $\hat{T} = t/T_{s_2}$:

$$\frac{d\hat{s}_1}{d\hat{T}} = \delta_S(1 + \kappa_1)(1 + \sigma_1) \left[-\hat{s}_1 + \frac{\sigma_1}{\sigma_1 + 1} \hat{c}_1 \hat{s}_1 + \frac{\kappa_1(1 + \kappa_1)^{-1}}{\sigma_1 + 1} \hat{c}_1 \right], \quad (38a)$$

$$\varepsilon \frac{d\hat{c}_1}{d\hat{T}} = \delta_S(1 + \kappa_1)(1 + \sigma_1) \left[\hat{s}_1 - \frac{\sigma_1}{\sigma_1 + 1} \hat{c}_1 \hat{s}_1 - \frac{1}{\sigma_1 + 1} \hat{c}_1 \right]. \quad (38b)$$

By inspection of (38a), if $\delta_S \ll 1$, then s_1 will be a slow variable over the T_{s_2} timescale, and thus we will expect s_1 to be essentially constant over the time course of the indicator reaction. In addition, let us assume that $T_{s_2} \gg t_{c_1}$, in which case c_1 will be on the order of its maximum value on the T_{s_2} timescale. Combining these observations leads to the approximation

$$s_1 \simeq s_1^0, \quad t \leq T_{s_2} \quad (39a)$$

$$c_1 \simeq \varepsilon s_1^0, \quad t \leq T_{s_2} \quad (39b)$$

for the non-observable reaction over the timescale T_{s_2} . Equations (39a) and
225 (39b) suggest that $e_2^A \ll s_1^0$ over the T_{s_2} timescale. Furthermore, since the changes in s_1 and c_1 are comparatively minimal when $t_{c_1} \leq t \leq T_{s_2}$, the production of e_2^A is effectively constant over the T_{s_2} timescale:

$$\dot{e}_2^A \approx \varepsilon k_2 s_1^0 \equiv \varpi. \quad (40)$$

Integration of (40) yields the following approximation of e_2^A on the T_{s_2} timescale

$$e_2^A \approx \int_0^t \varpi \, du = \varpi t, \quad (41)$$

230 where “ u ” in (41) has been utilized to denote a dummy variable. The approximate average value $\langle e_2^A \rangle$ on T_{s_2} is easily obtained through straightforward

integration

$$\langle e_2^A \rangle = \frac{\varpi}{T_{s_2}} \int_0^{T_{s_2}} t \, dt = \frac{1}{2} T_{s_2} \varpi, \quad (42)$$

and insertion of (42) into (30) yields an estimate for T_{s_2} :

$$T_{s_2} = \sqrt{\frac{2(K_{M_2} + s_2^0)}{k_4 \varpi}} \equiv T_{s_2}^*. \quad (43)$$

We can write (43) in a slightly more convenient form by defining the *limiting*
 235 depletion timescale $t_{s_2}^*$ as

$$t_{s_2}^* \equiv \frac{K_{M_2} + s_2^0}{V_2}, \quad (44)$$

which allows us to express $T_{s_2}^*$ as

$$T_{s_2}^* = \sqrt{2t_{s_1}^* t_{s_2}^*}. \quad (45)$$

Note that $V_2 = k_4 s_1^0$ is defined as the limiting rate of the indicator reaction.

$T_{s_2}^*$ should provide an accurate estimate for total completion time of the
 indicator reaction as long as the non-observable reaction is comparatively slow.

240 For a generic (and linear) dynamical system of the form

$$\dot{x} = -ax, \quad x(0) = x_0, \quad (46)$$

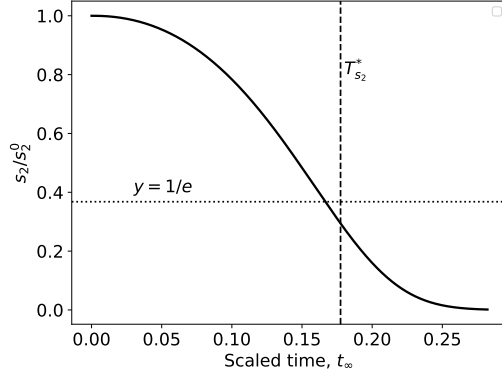
the depletion or *characteristic* timescale is $1/a$. Analogously, we will look for
 a timescale that is indicative of the time it takes for the initial quantity (i.e.,
 x_0 in the context of (46)) to deplete to an amount that is less than or equal to
 x_0/e . Following suit from the linear theory, we will consider the timescale $T_{s_2}^*$

245 to be a sufficient depletion timescale as long as

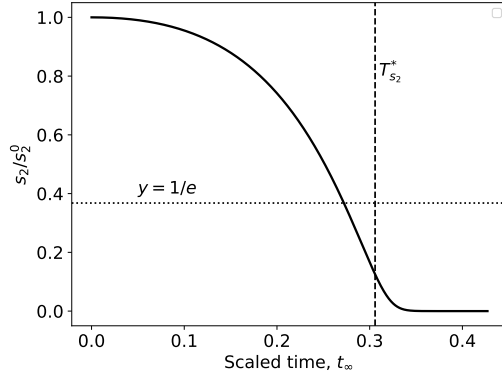
$$s_2(T_{s_2}^*) \leq s_2^0/e \approx 0.37s_2^0. \quad (47)$$

Numerical solutions of the mass action equations confirm the validity of the
 timescale $T_{s_2}^*$ when the indicator reaction is much faster than the non-observable
 reaction provided $t_{c_1} \ll T_{s_2}^*$ (see FIGURES 2a and 2b).

Next, we develop an asymptotic solution to the mass action equations that
 250 will be valid when $T_{s_2}^*$ is an accurate depletion timescale, and the concentrations



(a)



(b)

Figure 2: The accuracy of the timescale $T_{s_2}^*$ when the indicator reaction (2) is fast ($\delta_S \ll 1$). The solid black curves are numerical solutions to the mass action equations of the complete reaction (10). The dashed line marks the timescale $T_{s_2}^*$ and the dotted line represents the quantity $1/e$. (a) The constants (without units) used in the numerical simulation are: $e_1^0 = 1$, $s_1^0 = 100$, $k_1 = 1$, $k_2 = 1$ and $k_{-1} = 1$. $s_2^0 = 10$, $k_3 = 10$, $k_4 = 100$ and $k_{-3} = 10$. (b) The constants (without units) used in the numerical simulation are: $e_1^0 = 1$, $s_1^0 = 100$, $k_1 = 1$, $k_2 = 1$ and $k_{-1} = 1$. $s_2^0 = 100$, $k_3 = 10$, $k_4 = 100$ and $k_{-3} = 10$. In both cases, we see that the timescale $T_{s_2}^*$ yields an accurate approximation to the completion time of the indicator reaction. Time has been mapped to the t_∞ scale: $t_\infty(t) = 1 - 1/\ln(t + e)$.

s_1 and c_1 remain on the order of their maximum values for the duration of the

indicator reaction. To begin, let us assume that s_2^0 is large enough so that

$$\max_{t \leq T_{s_2}^*} (e_2^A) \ll s_2^0, \quad (48)$$

in which case we can assume $\lambda \ll 1$. Then, from Tikhonov's theorem, we obtain

$$c_2 \simeq \frac{e_2^A}{K_{M_2} + s_2} s_2 \quad (49)$$

255 as a leading order approximation. Insertion of this approximation into the mass action equation for s_2 yields

$$\dot{s}_2 \simeq -\frac{k_4 e_2^A}{K_{M_2} + s_2} s_2. \quad (50)$$

Finally, substitution of $e_2^A \approx \varpi t$ into (50) gives us

$$\dot{s}_2 \simeq -\frac{k_4 \varpi t}{K_{M_2} + s_2} s_2 \quad (51)$$

as our final asymptotic approximation to the differential equations governing the temporal depletion of s_2 . Equation (51) has a closed-form solution in the
260 form of the Schnell–Mendoza equation

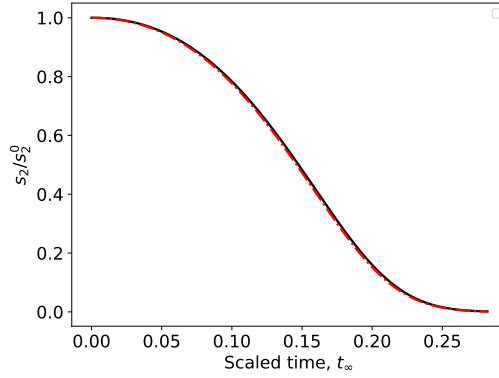
$$s_2 = K_{M_2} W \left[\sigma_2 \exp \left(\sigma_2 - \frac{k_4 \varpi t^2}{2K_{M_2}} \right) \right], \quad (52)$$

and provides an accurate approximation to the mass action model (see FIGURES 3a and 3b).

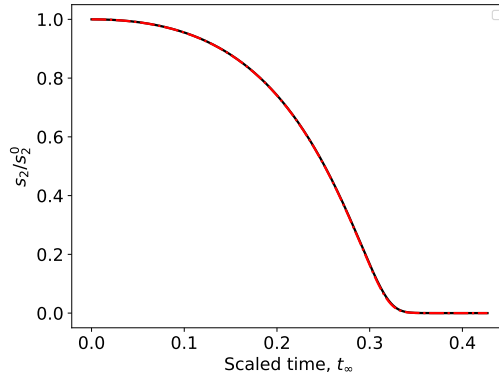
4.2. Case 2: The indicator reaction is roughly the same speed as the non-observable reaction ($\delta \approx 1$)

265 If the non-observable reaction and the indicator reaction both complete at roughly the same time, then it is appropriate to use either t_{s_1} or T_{s_2} as the depletion timescale for the complete reaction. Of course, given our earlier definition of the timescale T_{s_2}

$$T_{s_2} = \frac{K_{M_2} + s_2^0}{k_4 \langle e_2^A \rangle}, \quad (53)$$



(a)



(b)

Figure 3: The leading order asymptotic solution (52) of the substrate concentration for the indicator reaction matches the numerical solution when the indicator reaction is faster than the non-observable reaction ($\delta_S \ll 1$). The solid black curves are numerical solutions to the mass action equations of the complete reaction (7) and the broken red curves are numerical solutions to the asymptotic differential equation (51). (a) The constants (without units) used in the numerical simulation are: $e_1^0 = 1, s_1^0 = 100, k_1 = 1, k_2 = 1$ and $k_{-1} = 1$. $s_2^0 = 10, k_3 = 10, k_4 = 100$ and $k_{-3} = 10$. (b) The constants (without units) used in the numerical simulation are: $e_1^0 = 1, s_1^0 = 100, k_1 = 1, k_2 = 1$ and $k_{-1} = 1$. $s_2^0 = 100, k_3 = 10, k_4 = 100$ and $k_{-3} = 10$. Time has been mapped to the t_∞ scale: $t_\infty(t) = 1 - 1/\ln(t + e)$.

we can formulate a nonlinear algebraic equation that will allow us to compute
 270 an estimate for the depletion timescale when the reactions are equivalent in

speed. First,

$$\langle e_2^A \rangle = \frac{1}{T_{s_2}} \int_0^{T_{s_2}} (s_1^0 - s_1 - c_1) dt, \quad (54)$$

and thus we see that T_{s_2} should satisfy

$$\int_0^{T_{s_2}} (s_1^0 - s_1 - c_1) dt = \frac{K_{M_2} + s_2^0}{k_4}. \quad (55)$$

Second, under the RSA, the concentration c_1 is expressible (algebraically) in terms of s_1 . Therefore, the integrand given in (55) can be expressed as³

$$\int_0^{T_{s_2}} (s_1^0 - s_1 - c_1) dt \approx \int_0^{T_{s_2}} \frac{(K_{M_1} + s_1)\Delta s_1 - e_1^0 s_1}{K_{M_1} + s_1} dt, \quad (56)$$

275 where $\Delta s_1 = s_1^0 - s_1$. Third, the definite integral on the right hand side of (56) is straightforward to compute analytically; evaluating it will yield a nonlinear equation in terms of the variable T_{s_2} , and the solution to (55) can be approximated numerically. Using the average $\langle e_2^A \rangle$ provides an accurate estimate of the depletion timescale (see FIGURE 4).

From a practical point of view, the utility in numerically estimating T_{s_2} is rather minimal. The objective here will be to construct a criteria from which a reduced model can be extracted from the mass action equations that will be valid without any *a priori* knowledge of the intrinsic timescales of the indicator reaction (or the non-observable reaction). To achieve this, let us first revisit the generic scaling introduced in the previous section:

$$\frac{d\bar{s}_2}{dT} = \frac{\max(e_2^A)(1 + \kappa_2)(1 + \sigma_2)}{\langle e_2^A \rangle \delta_S} \left[\left(\frac{\sigma_2}{1 + \sigma_2} \bar{c}_2 - \bar{e}_2^A \right) \bar{s}_2 + \frac{\alpha}{1 + \sigma_2} \bar{c}_2 \right], \quad (57a)$$

$$\lambda \frac{d\bar{c}_2}{dT} = \frac{\max(e_2^A)(1 + \kappa_2)(1 + \sigma_2)}{\langle e_2^A \rangle \delta_S} \left[\left(\bar{e}_2^A - \frac{\sigma_2}{1 + \sigma_2} \bar{c}_2 \right) \bar{s}_2 - \frac{1}{1 + \sigma_2} \bar{c}_2 \right]. \quad (57b)$$

280 Bearing in mind the assumption $\delta_S \approx 1$, it is sufficient (but not necessary) to bound λ in order to assemble a dynamical model that can be reduced (asymptotically) through slow manifold projection. The upper bound on λ , which we

³The production of e_2^A over the t_{c_1} timescale is negligible under the RSA.

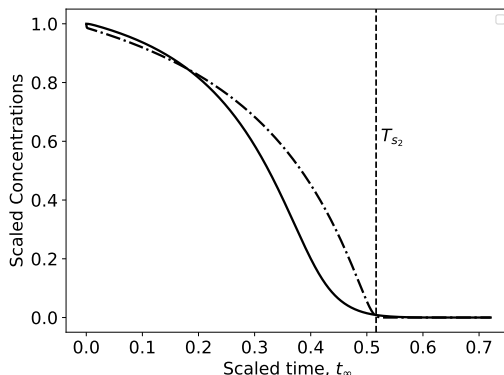


Figure 4: The averaging method for the estimation of the depletion timescale T_{s_2} for the indicator substrate is still valid when the non-observable and indicator reactions occur at roughly the same speed ($\delta_S \approx 1$). The solid black curve is the numerically-computed depletion curve of s_2 (10c) and the dotted/dashed black curve is the numerically-integrated depletion curve of s_1 (10a). In this numerical simulation $k_3 = 1, k_4 = 1, k_{-3} = 10, s_2^0 = 70$, and $k_1 = 10, k_2 = 15, k_{-1} = 1, e_1^0 = 1$ and $s_1^0 = 70$. Both substrates have been scaled as s_2/s_2^0 and s_1/s_1^0 . Time has been mapped to the t_∞ scale: $t_\infty(t) = 1 - 1/\ln(t + e)$.

denote as λ^{\max} , is given by

$$\lambda \leq \lambda^{\max} \equiv \frac{s_1^0}{K_{M_2} + s_2^0}. \quad (58)$$

The parameter λ^{\max} is the natural small parameter when the indicator is very slow. Furthermore, if the non-observable reaction completes very quickly relative to the non-observable reaction, and $\delta_S \ll 1$, then the average available enzyme should be on the order of s_1^0 :

$$\langle e_2^A \rangle = \frac{1}{T_{s_2}} \int_0^{T_{s_2}} e_2^A dt \approx s_1^0. \quad (59)$$

Thus, if $s_2^0 \gg s_1^0$, then the approximation

$$\dot{s}_2 \simeq -\frac{k_4 e_2^A}{K_{M_2} + s_2} s_2 \quad (60)$$

will be valid if $\lambda^{\max} \ll 1$. Furthermore, (60) admits a closed-form solution using separation of variables that consists of composite Lambert- W functions (we do

not present this expression here, although we remark that it is straightforward to obtain through careful integration). If the RSA is valid, then

$$\dot{s}_2 \simeq - \left(\frac{(K_{M_1} + s_1)\Delta s_1 - e_1^0 s_1}{K_{M_1} + s_1} \right) \left(\frac{k_4}{K_{M_2} + s_2} \right) s_2 \quad (61)$$

is the final form of our reduced differential equation for \dot{s}_2 when the reactions are comparable in speed.

295 *4.3. Case 3: The indicator reaction is much slower than the non-observable reaction ($\delta_S \gg 1$)*

We now consider the case when $\delta_S \gg 1$. As mentioned in the previous subsection, a very slow indicator reaction suggests that s_2 will be slow over the timescale t_{s_1} . Consequently, we can approximate s_2 as

$$s_2 = s_2^0, \quad t < t_{s_1}. \quad (62)$$

300 Furthermore, because the non-observable reaction has effectively completed when $t = t_{s_1}$, we can approximate $\Delta s_1 \approx s_1^0$ when $t \geq t_{s_1}$. This yields

$$\dot{s}_2 \simeq - \frac{k_4 s_1^0}{K_{M_2} + s_2} s_2, \quad t \geq t_{s_1}, \quad (63)$$

which should be valid if λ^{\max}

$$\lambda^{\max} \equiv \frac{s_1^0}{K_{M_2} + s_2^0} \quad (64)$$

is small. Equation (63) can be integrated directly to yield a Schnell–Mendoza equation for s_2 :

$$s_2 = K_{M_2} W [\sigma_2 \exp(\sigma_2 - \eta_2(t))], \quad t \geq t_{s_1}. \quad (65)$$

305 The validity of the approximate solution (62) can be established by the mathematical formulation of the RSA for the indicator reaction. If $s_2 \approx s_2^0$ over the interval $[0, t_{s_1}]$, then

$$\max_{t \leq t_{s_1}} |\dot{s}_2| \cdot t_{s_1} \ll s_2^0. \quad (66)$$

The inequality given in (66) translates to

$$\delta_S \gg (\sigma_2 + 1)(\kappa_2 + 1), \quad (67)$$

with $\max \dot{s}_2 = k_3 s_1^0 s_2^0$. Thus, we have a RSA that is applicable to slow indicator

310 reactions:

$$\frac{V_1}{V_2} \gg \frac{K_{M_1}}{K_{M_2}}(1 + \sigma_1)(1 + \kappa_2). \quad (68)$$

Equation (68) establishes a region of validity for the solution to the mass action equations during the initial build-up of c_2 when $t \leq t_{s_1}$. Interestingly, (68) is analogous to the term used to measure the strength of fully competitive enzyme reactions with alternative substrates [22, 23]. Numerical simulations (see FIGURE 5) confirm the validity of $t_{s_2}^*$ and (63).

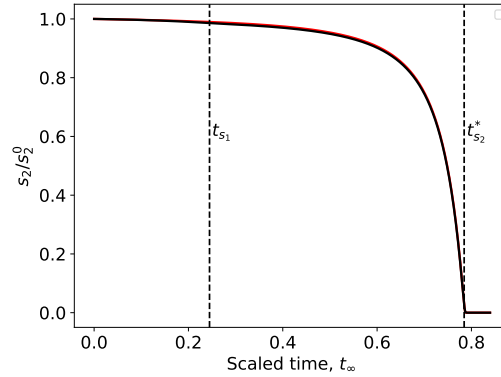


Figure 5: Validity of the timescale $t_{s_2}^*$ and the reduced ordinary differential equation given by (63) for the substrate depletion of the indicator reaction when the indicator reaction is much slower than the non-observable reaction ($\delta_S \gg 1$). The solid black curve is the numerical solution to the mass action equations (7) and the solid red curve corresponds to the numerical solution to (63) extended to $t \geq 0$. In this numerical simulation $k_3 = 0.1, k_4 = 1, k_{-3} = 10, s_2^0 = 10000$, and $k_1 = 25, k_2 = 100, k_{-1} = 1, e_1^0 = 1$ and $s_1^0 = 100$. The respective values of λ^{\max} and δ_S are ≈ 0.009 and ≈ 0.01 . Time has been mapped to the t_∞ scale: $t_\infty(t) = 1 - 1/\ln(t + e)$.

315

5. Applicability of the QSSA for slow indicator reactions

In the context of coupled reactions, timescales are categorized as *fast* if they are short in comparison to the depletion timescale (t_{s_1}) of the primary reaction. So far, we have not discussed the transient timescale of the indicator reaction. This timescale is a “fast” timescale for the indicator reaction that it is analogous to t_{c_1} , the transient timescale of the indicator reaction. This timescale appears when initial conditions start on the c_2 -nullcline (i.e., in QSS), then the system may exhibit a *fast* transient (analogous to t_{c_1}) over which the phase plane trajectory *swiftly* travels towards the c_2 -nullcline, especially if the QSSA holds over the timescale of interest. However, the QSSA is not necessarily valid with respect to the t_{s_1} timescale, and it is possible to derive a timescale that must be *short* in comparison to both t_{s_1} and T_{s_2} in order to apply QSSA over the time course of the non-observable reaction when the indicator reaction is substantially slower. Carefully rescaling the mass action equation for c_2 with respect to $T = t/t_{s_1}$ yields

$$\lambda \frac{d\bar{c}_2}{dT} = \frac{\max(e_2^A)(1 + \kappa_2)(1 + \sigma_2)}{\langle e_2^A \rangle \delta_S} \left[\left(\bar{e}_2^A - \frac{\sigma_2}{1 + \sigma_2} \bar{c}_2 \right) \bar{s}_2 - \frac{1}{1 + \sigma_2} \bar{c}_2 \right]. \quad (69)$$

If the indicator reaction is slow, and $\delta_S \gg 1$, then it is necessary that the inequality

$$\lambda \cdot \frac{\langle e_2^A \rangle}{\max(e_2^A)} \cdot \frac{1}{(1 + \kappa_2)(1 + \sigma_2)} \cdot \delta_S = \frac{t_{c_2}}{t_{s_1}} \ll 1, \quad t_{c_2} \equiv \frac{1}{k_3(K_{M_2} + s_2^0)} \quad (70)$$

holds in order to impose the QSSA on the t_{s_1} timescale. The timescale t_{c_2} has no obvious physical interpretation in the context of experimental initial conditions: it simply arises naturally as a result of the scaling analysis.

To gain an understanding of the behavior of the indicator reaction over t_{c_2} ,

we rescale time with respect to $T^* = t/t_{c_2}$:

$$\frac{d\bar{s}_2}{dT^*} = \lambda \left[\left(\frac{\sigma_2}{1 + \sigma_2} \bar{c}_2 - \bar{e}_2^A \right) \bar{s}_2 + \frac{\alpha}{1 + \sigma_2} \bar{c}_2 \right], \quad (71a)$$

$$\frac{d\bar{c}_2}{dT^*} = \left(\bar{e}_2^A - \frac{\sigma_2}{1 + \sigma_2} \bar{c}_2 \right) \bar{s}_2 - \frac{1}{1 + \sigma_2} \bar{c}_2. \quad (71b)$$

We see from the scaled equations (71) that t_{c_2} defines a *stagnation* timescale when experimental initial conditions are prescribed and $\lambda \ll 1$. If the timescale t_{c_2} is short, then the indicator reaction is essentially stationary over t_{c_2} . This is because s_2 scales as a slow variable over t_{c_2} , and the phase space trajectory should stay *near* the c_2 -nullcline over short timescales. Thus, if t_{c_2} is small (i.e., $t_{c_2} \ll \min\{t_{s_2}, t_{s_1}\}$), then this timescale translates to a scale over which the indicator reaction exhibits a “slow response”. In fact, any timescale “ t^* ” that satisfies $t^* \ll \min\{t_{s_1}, T_{s_2}\}$ qualifies as a stagnation timescale.

In addition to the exposition of t_{c_2} as a stagnation timescale (when $\lambda \ll 1$), the separation of t_{c_2} and t_{s_1} also retains a biophysical interpretation. After the initial fast transient of the non-observable reaction, the production rate of e_2^A is roughly

$$\dot{e}_2^A \simeq \frac{V_1}{K_{M_1} + s_1^0} s_1^0 \equiv \max(\dot{e}_2^A). \quad (72)$$

If we demand that the total production of available enzyme be negligible over t_{c_2} , then it is sufficient to require

$$\max(\dot{e}_2^A) \cdot t_{c_2} \ll s_1^0. \quad (73)$$

The inequality, (73), is equivalent to $t_{c_2}/t_{s_1} \ll 1$, and we see that the QSSA can be imposed when production of e_2^A is asymptotically negligible over t_{c_2} .

Moreover, the relationship between λ , t_{c_2} and T_{s_2} is now evident:

$$\frac{t_{c_2}}{T_{s_2}} < \lambda. \quad (74)$$

The strict inequality in (74) follows from the fact that

$$\frac{t_{c_2}}{T_{s_2}} = \frac{\bar{\lambda}}{(1 + \sigma_2)(1 + \kappa_2)}, \quad (75)$$

where $\bar{\lambda}$ is given by

$$\bar{\lambda} \equiv \frac{\langle e_2^A \rangle}{K_{M_2} + s_2^0}. \quad (76)$$

355 Furthermore, since $\langle e_2^A \rangle \leq \max(e_2^A)$, we see that

$$\bar{\lambda} \leq \lambda, \quad (77)$$

from which (74) follows. We note that the parameter λ^{\max} is easily derived using Segel's heuristic approach [18]:

$$\max |\dot{s}_2| \cdot t_{c_2} \ll s_2^0 \rightarrow \lambda^{\max} \ll 1. \quad (78)$$

Since it is clear that

$$\bar{\lambda} \leq \lambda \leq \lambda^{\max}, \quad (79)$$

it follows that the RSA (i.e., $\lambda^{\max} \ll 1$) ensures separation of relevant timescales.

360 Consequently, the RSA for the indicator reaction is a universal qualifier for the validity of the reduced model with respect to the timescale T_{s_2} .

6. Estimation of lag times

Under the QSSA, enzyme catalyzed reactions usually express a lag time. The lag time is normally defined as the time it takes for the rate of product
365 generation to reach its maximum (steady-state) value. This coincides with the time it takes for c_2 to reach its maximum value; it is straightforward to calculate under the limiting circumstances.

6.1. Estimation of the lag time for fast indicator reactions

Let us start by considering the case when the indicator reaction is very fast;
370 we will assume s_2 is given by

$$s_2 = K_{M_2} W \left[\sigma_2 e^{\sigma_2 - \nu t^2 / 2K_{M_2}} \right]. \quad (80)$$

If $\sigma_2 \ll 1$, then (80) is approximately

$$s_2 \simeq s_2^0 e^{\sigma_2 - \nu t^2 / 2K_{M_2}}. \quad (81)$$

Next, notice that under the QSSA we have

$$c_2 \simeq -\frac{1}{k_4} \frac{ds_2}{dt}, \quad t \geq 0. \quad (82)$$

Differentiating both sides of (82), we see that \dot{c}_2 vanishes when \ddot{s}_2 vanishes:

$$\frac{dc_2}{dt} \simeq -\frac{1}{k_4} \frac{d^2s_2}{dt^2}. \quad (83)$$

Inserting (81) into the right hand side of (83), and setting the left hand side to zero yields

$$t = \sqrt{\frac{K_{M_2}}{V_2}} \cdot t_{s_1} \equiv t_{c_2}^*. \quad (84)$$

For the case of the fast indicator reaction, the timescale $t_{c_2}^*$ is identically the *lag time* when $\sigma_2 \ll 1$ (see FIGURE 6).

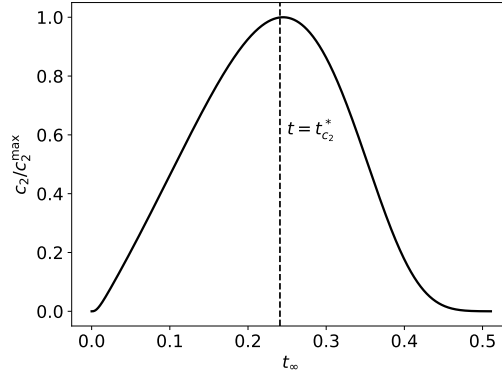


Figure 6: Validity of the timescale $t_{c_2}^*$. The curve represents the numerical solution to the mass action equations (10) (thick black line) with $s_1^0 = 100, c_1^0 = 1, k_1 = 1, k_2 = 1, k_{-1} = 1, s_2^0 = 1, k_3 = 1, k_{-3} = 1$ and $k_4 = 100$. The dashed line corresponds to $t_{c_2}^*$ and is the time it takes for c_2 to reach its maximum value. The total concentration c_2 has been scaled by c_2/c_2^{\max} .

6.2. Estimation of the lag time for slow indicator reactions

For slow indicator reactions will can employ the RSA

$$\max_{t \leq t_{s_1}} |\dot{s}_2| \cdot t_{s_1} \ll s_2^0, \quad (85)$$

380 which allows us to linearize the mass action equation for c_2 :

$$\dot{c}_2 = k_3(e_2^A - c_2)s_2^0 - (k_{-3} + k_4)c_2, \quad t \leq t_{s_1}. \quad (86)$$

Furthermore we will assume that $\max(c_2)$ is $\lambda^{\max}s_2^0$ when the indicator reaction is slow. In this case, the timescale t_{s_1} will serve as a good approximation to the lag time when σ_1 very large. However, when σ_1 is small, the asymptotic solution to the MM equation reduces to

$$s_1 = K_{M_1} W \left[\sigma_1 e^{\sigma_1 - \eta_1 t} \right] \simeq s_1^0 e^{\sigma_1 - \eta_1 t}. \quad (87)$$

385 It follows from (87) that the timescale t_{s_1} is characteristic when σ_1 is small; this means roughly 1/3 of s_1^0 still needs to be converted to product when $t = t_{s_1}$. Consequently, we need an estimate for the time it takes for the non-observable reaction to complete when σ_1 is small. To do this we set

$$s_1^0 e^{\sigma_1 - \eta_1 t} = \epsilon, \quad \epsilon \equiv t_{c_1}/t_{s_1}, \quad (88)$$

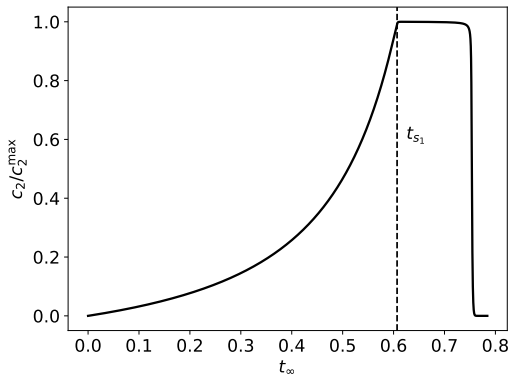
and solve for t . This yields

$$t = -t_{s_1} \ln \epsilon \equiv t_{s_1}^*, \quad (89)$$

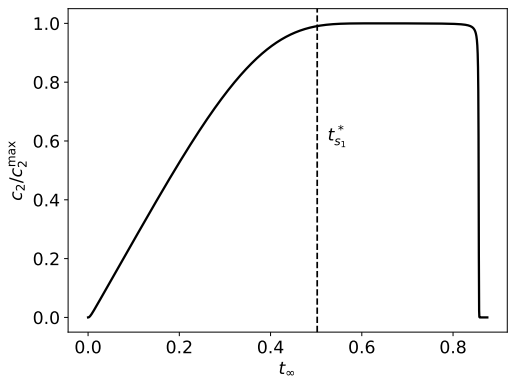
390 and is a much better estimate of the lag time when σ_1 is small. A similar analysis can be carried out when σ_1 is of order unity, but we will not dive into the details of this calculation here. Numerical results confirm the accuracy of the lag time estimates t_{s_1} and $t_{s_1}^*$ when the indicator reaction is slow (see FIGURES 7a– 7b).

395 7. Discussion

The primary contributions of this paper are the estimation of scaling variables and timescale for a reaction mechanism of zymogen activation (1)–(2). The identification of specific parameters through scaling has yielded necessary and sufficient conditions for the QSSA and RSA, whereas previous nonlinear studies
400 of the coagulation cascade with zymogen activation reactions have employed the



(a)



(b)

Figure 7: The validity of t_{s_1} and $t_{s_1}^*$. The former is the approximate lag time when σ_1 is large (dashed line in panel (a)) and the latter is the lag time when σ_1 is small (dashed line of panel (b)). The solid black curves are the numerical solutions to the mass action equations (7) of the complete reaction. (a) The constants (without units) used in the numerical simulation are: $e_1^0 = 1, s_1^0 = 100, k_1 = 100, k_2 = 10$ and $k_{-1} = 1$. $s_2^0 = 5000, k_3 = 1, k_4 = 1$ and $k_{-3} = 10$. (b) The constants (without units) used in the numerical simulation are: $e_1^0 = 1, s_1^0 = 1, k_1 = 1, k_2 = 100$ and $k_{-1} = 1$. $s_2^0 = 1000, k_3 = 1, k_4 = 1$ and $k_{-3} = 10$. Time has been mapped to the t_∞ scale: $t_\infty(t) = 1 - 1/\ln(t + e)$.

QSSA without justification [24]. Moreover, previous analyses [12, 1, 13, 4, 2] have only employed PFO kinetic models, and do not provide insight as to how to properly estimate kinetic timescales via nonlinear methods, even though reaction mechanism of zymogen activation is inherently nonlinear. This work

405 outlines a clear procedure for estimating depletion timescales, and serves as a template for the analysis of more complicated reactions. We give a brief summary of the results of the analysis in what follows.

Scaling analysis of the mass action equations that model the kinetics of a reaction mechanism of zymogen activation (1)–(2) has revealed two small parameters:

$$\lambda^{\max} = \frac{s_1^0}{K_{M_2} + s_2^0} \ll 1,$$

$$\varepsilon = \frac{e_1^0}{K_{M_1} + s_1^0} \ll 1.$$

The QSSA is valid over the respective depletion timescales of the indicator and non-observable reactions when both λ^{\max} and ε are sufficiently small.

In addition, simple asymptotic solutions to the mass action equations were derived that are valid when the indicator reaction is very fast (or very slow) in comparison to the non-observable reaction. If the indicator reaction is fast, then the time course of the indicator substrate s_2 is accurately approximated by

$$s_2 = K_{M_2} W \left[\sigma_2 \exp \left(\sigma_2 - \frac{k_4 \varpi t^2}{2K_{M_2}} \right) \right],$$

where W denotes the Lambert- W function. In contrast, if the indicator reaction is very slow, then the time course of s_2 can be approximated by

$$s_2 = K_{M_2} W \left[\sigma_2 \exp \left(\sigma_2 - \frac{V_2 t}{K_{M_2}} \right) \right].$$

410 Note that the above two expressions are analogous to the Schnell–Mendoza equation [14].

It should be pointed out that the condition $\lambda^{\max} \ll 1$, which can be ensured by requiring an excess of the initial amount of substrate s_2 (i.e., requiring that s_2^0 be large enough so that $s_1^0 \ll s_2^0$), is sufficient but not necessary for the
 415 validity of the reduced model presented in (61). In general, it is desirable that s_2^0 be much larger than the maximum of amount of e_2^A over the timescale of

the indicator reaction. If the indicator reaction is fast, then the maximum amount of available enzyme will be small; thus, the requirement that $s_1^0 \ll s_2^0$ is unnecessary if $\max(e_2^A) \ll K_{M_2}$ (see FIGURE 7). Of course, the integrity of the reduced model does not diminish if $s_1^0 \ll s_2^0$.

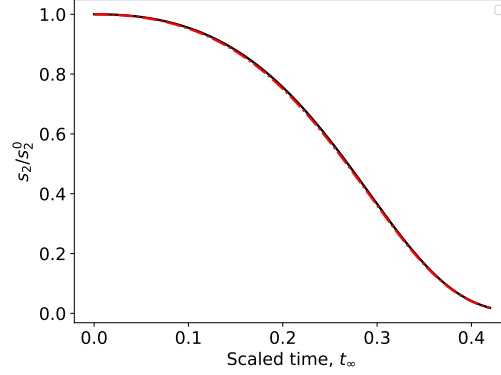


Figure 8: The condition that $\lambda \ll 1$ is necessary for slow manifold projection, \mathcal{M}_λ , while the condition $\lambda^{\max} \ll 1$ is merely sufficient. The solid black curve is the numerical solution to the mass action equations (10) and the broken red curve is to the numerical solution to (61). In this simulation $k_3 = 1, k_4 = 100, k_{-3} = 10, s_2^0 = 1$, and $k_1 = 1, k_2 = 1, k_{-1} = 1, e_1^0 = 1$ and $s_1^0 = 100$. $s_1^0/s_2^0 = 100$ and $\lambda^{\max} \approx 1$. However, $\max e_2^A \approx 1.543$ and therefore $\lambda \approx 0.014 \ll 1$. Time has been mapped to the t_∞ scale: $t_\infty(t) = 1 - 1/\ln(t + e)$.

420

Finally, three reduced models have been derived that can be utilized in the analysis of the inverse problem. Our analysis seems to suggest that a fast indicator reaction is the most favorable case for parameter estimation. If the indicator reaction has sufficient speed, then theoretically these two expressions,

$$s_2 = K_{M_2} W \left[\sigma_2 \exp \left(\sigma_2 - \frac{V_2 V_1 t^2}{2K_{M_2}(s_1^0 + K_{M_1})} \right) \right]$$

and

$$\dot{s}_2 = - \left(\frac{(K_{M_1} + s_1) \Delta s_1 - e_1^0 s_1}{K_{M_1} + s_1} \right) \left(\frac{k_4}{K_{M_2} + s_2} s_2 \right),$$

can be utilized simultaneously to estimate the four unknown parameters: V_1, V_2, K_{M_1} , and K_{M_2} . However, the complete understanding of the inverse problem is

beyond the scope of this paper, and we hope to investigate the parameter estimation for the reaction mechanism of zymogen activation (1)–(2) in subsequent
425 future work.

Acknowledgements

We are grateful to Dr. Enrico Di Cera (Saint Louis University School of Medicine) for his suggestions to explore this problem. We are also grateful to Dr. Antonio Baici (University of Zurich) for helpful discussions about this
430 work during the 2017 Beilstein Enzymology Symposia (Rüdesheim, Germany), and the anonymous reviewers for their comments during the revision of this manuscript. This work is partially supported by the University of Michigan Protein Folding Diseases Initiative, and Beilstein-Institut zur Förderung der Chemischen Wissenschaften through its Beilstein Enzymology Symposia. Dr.
435 Stroberg is a fellow of the Michigan IRACDA program (NIH/NIGMS grant: K12 GM111725).

References

- [1] B. Havsteen, M. Garcia-Moreno, E. Valero, M. Manjabacas, R. Varn, The kinetics of enzyme systems involving activation of zymogens, *Bull. Math. Biol.* 55 (1993) 561–583.
440
- [2] R. Varn, B. Havsteen, M. Garca, A. Vsquez, J. Tudela, F. Cnovas, Kinetics of the trypsinogen activation by enterokinase and/ or trypsin: Coupling of a reaction in which the trypsin acts on one of its substrates, *J. Mol. Catal.* 66 (1991) 409–419.
- [3] D. L. Purich, *Enzyme kinetics: Catalysis & control*, Academic Press, London, UK, 2010.
445
- [4] R. Varón, A. Román, F. García, F. G. Carmona, Transient phase kinetics of activation of human plasminogen, *Bull. Math. Biol.* 48 (1986) 149–166.

- [5] O. D. Dang, A. Vindigni, E. Di Cera, An allosteric switch controls the
450 procoagulant and anticoagulant activities of thrombin, *Proc. Natl. Acad. Sci. USA* 92 (1995) 5977–5981.
- [6] F. B. Rudolph, B. W. Baugher, R. S. Beissner, Techniques in coupled
enzyme assays, *Methods Enzymol.* 63 (1979) 22–42.
- [7] B. A. C. Storer, A. Cornish-bowden, P. O. Box, B. Birmingham, The ki-
455 netics of coupled enzyme reactions, *Biochem. J.* 141 (1974) 205–209.
- [8] W. W. Cleland, Optimizing coupled enzyme assays, *Anal. Biochem.* 99
(1979) 142–145.
- [9] W. R. McClure, Kinetic analysis of coupled enzyme assays, *Biochemistry*
8 (1969) 2782–2786.
- 460 [10] J. S. Easterby, Coupled enzyme assays: A general expression for the tran-
sient, *Biochim Biophys Acta.* 293 (1973) 552–558.
- [11] S. Schnell, C. Mendoza, The condition for pseudo-first-order kinetics in
enzymatic reactions is independent of the initial enzyme concentration,
Biophys. Chem. 107 (2004) 165–174.
- 465 [12] M. E. Fuentes, E. Valero, M. García-Moreno, E. Vique, R. Varón, Kinetic
analysis of the mechanism of plasminogen activation by streptokinase, *J.*
Math. Chem. 42 (2007) 753–774.
- [13] R. Varn, B. Havsteen, Kinetics of the transient-phase and steady-state of
the monocyclic enzyme cascades, *J. theor. Biol.* 144 (1990) 397–413.
- 470 [14] S. Schnell, C. Mendoza, Closed form solution for time-dependent enzyme
kinetics, *J. Theor. Biol.* 187 (1997) 207–212.
- [15] A. H. Nguyen, S. J. Fraser, Geometrical picture of reaction in enzyme
kinetics, *J. Chem. Phys.* 91 (1989) 186–193.

- [16] M. R. Roussel, S. J. Fraser, Geometry of the steady-state approximation:
475 Perturbation and accelerated convergence methods, *J. Chem. Phys.* 93
(1990) 1072–1081.
- [17] S. M. Hanson, S. Schnell, Reactant stationary approximation in enzyme
kinetics, *J. Phys. Chem. A* 112 (2008) 8654–8658.
- [18] L. A. Segel, On the validity of the steady state assumption of enzyme
480 kinetics, *Bull. Math. Biol.* 50 (1988) 579–593.
- [19] S. K. Shoffner, S. Schnell, Approaches for the estimation of timescales in
nonlinear dynamical systems: Timescale separation in enzyme kinetics as
a case study, *Math. Biosci.* 287 (2017) 122–129.
- [20] L. A. Segel, M. Slemrod, The quasi-steady-state assumption: a case study
485 in perturbation, *SIAM Rev.* 31 (1989) 446–477.
- [21] S. Schnell, Validity of the Michaelis-Menten equation – Steady-state, or
reactant stationary assumption: that is the question, *FEBS J.* 281 (2014)
464–472.
- [22] S. Schnell, C. Mendoza, Time-dependent closed form solutions for fully
490 competitive enzyme reactions, *Bull. Math. Biol.* 62 (2000) 321–336.
- [23] S. Schnell, C. Mendoza, Enzyme kinetics of multiple alternative substrates,
J. Math. Chem. 27 (2000) 155–170.
- [24] M. Khanin, V. Semenov, A mathematical model of the kinetics of blood
coagulation, *J. theor. Biol.* 136 (1989) 127–134.