

Using singular perturbation theory to determine kinetic parameters in a non-standard coupled enzyme assay

Mohit P. Dalwadi^{1,2,*}, Diego Orol², Frederik Walter²,
Nigel P. Minton², John R. King^{2,3}, Katalin Kovács²

Received: date / Accepted: date

Abstract We investigate how to characterize the kinetic parameters of an aminotransaminase using a non-standard coupled (or auxiliary) enzyme assay, where the peculiarity arises for two reasons. First, one of the products of the auxiliary enzyme is a substrate for the primary enzyme and, second, we explicitly account for the reversibility of the auxiliary enzyme reaction. Using singular perturbation theory, we characterize the two distinguished asymptotic limits in terms of the strength of the reverse reaction, which allows us to determine how to deduce the kinetic parameters of the primary enzyme for a characterized auxiliary enzyme. This establishes a parameter-estimation algorithm that is applicable more generally to similar reaction networks. We demonstrate the applicability of our theory by performing enzyme assays to characterize a novel putative aminotransaminase enzyme, CnAptA (UniProtKB Q0KEZ8) from *Cupriavidus necator* H16, for two different omega-amino acid substrates.

Keywords Asymptotic analysis · Reaction kinetics · Synthetic biology · Enzyme characterization · Aminotransaminase

1 Introduction

Enzyme assays are an important tool for characterizing enzymes. In the classic assay, a single reaction converts a substrate into a product, using an enzyme as a catalyst, and the product is measured over time to estimate the initial reaction velocity [3]. A mathematical analysis allows one to use this information to deduce the kinetic properties of the enzyme in question [18]. When characterizing an enzyme for which the reaction product is difficult to observe, an auxiliary enzyme can be introduced to convert the product of the primary enzyme reaction into a chemical that can more easily be measured (*e.g.* NADH) [19, 23]. This is called a coupled enzyme assay. However, the problem is more complicated if the reaction catalysed by the auxiliary enzyme has a product which is also a substrate for the primary enzyme.

In this paper, we are interested in understanding how to characterize a primary enzyme in a non-standard coupled enzyme assay, where one of the products of the auxiliary enzyme is a substrate for the primary enzyme and where we explicitly account for the reversibility of the auxiliary enzyme reaction. This is motivated by the characterization of the class of enzymes known as *aminotransaminases*, which catalyse the transfer of keto and amine groups between organic compounds [22], and are an important class of biocatalysts in synthetic biology for reasons we discuss below. In Figure 1 we show a schematic of the reaction network we consider in this paper. While we refer to the specific organic compounds pyruvate and alpha-L-alanine in Figure 1 and through the paper in order to be consistent with the experiments we perform, we emphasize that our analysis is more general. In particular, the general organic-group-transfer property of aminotransaminases means that ‘pyruvate’ can be generalized to ‘keto acid’, and ‘alpha-L-alanine’ can be generalized to ‘amino acid’ for an aminotransaminase primary enzyme. Since ammonia is produced when an amino acid is dehydrogenated, the presence of ammonia is a consequence of using an aminotransaminase as the primary enzyme. This means that the reaction network we show in Figure 1 is particularly applicable to coupled enzyme assays with the primary enzyme being an aminotransaminase.

The technique to deal with a standard coupled enzyme assay with one primary enzyme and one auxiliary enzyme was first outlined within a mathematical framework in [15]. We direct the interested reader to [19] for a history of the theoretical work on standard coupled enzyme assays, as well as a practical guide to performing these types of assays.

* E-mail: mohit.dalwadi@maths.ox.ac.uk

¹ Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Oxford, OX2 6GG, UK,

² Synthetic Biology Research Centre, University of Nottingham, University Park, Nottingham, NG7 2RD,

³ School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

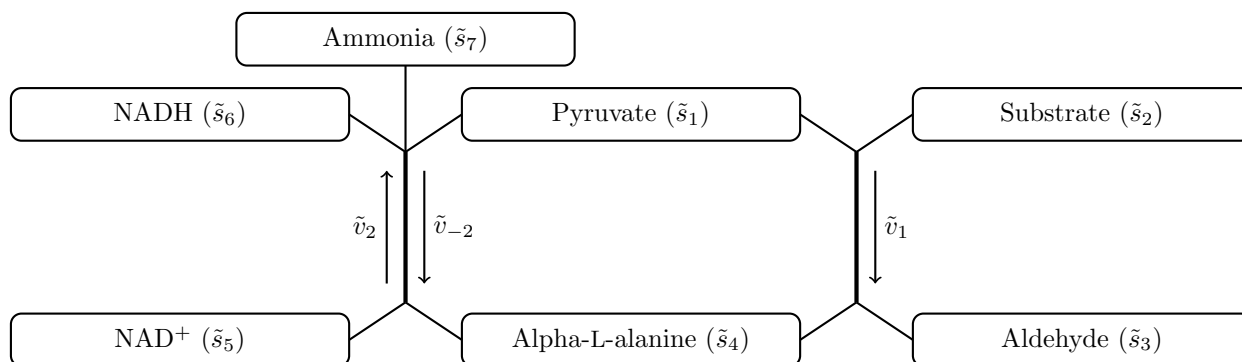


Fig. 1 A diagram of the reaction network we investigate in this paper. Each box denotes a chemical compound involved in the network, and each arrow denotes a reaction to and from the specified group of products and substrates, respectively. Each reaction is catalysed by an enzyme; the primary enzyme catalyses the reaction on the right, marked with velocity \tilde{v}_1 , and the auxiliary (secondary) enzyme catalyses both directions of the reversible reaction on the left, marked with forward velocity \tilde{v}_2 and backwards velocity \tilde{v}_{-2} . Note that our model does not rely on pyruvate being converted into alpha-L-alanine, it being valid for the primary enzyme being a general aminotransaminase. Hence, in this diagram ‘pyruvate’ can be generalized to ‘keto acid’, and ‘alpha-L-alanine’ can be generalized to ‘amino acid’.

In this paper, we investigate when the standard approach identified in [15] is appropriate for the non-standard assay we consider, and how to modify the approach when it is not. We then implement our theory by performing assays on a previously-uncharacterized enzyme for two different substrates.

Calculating enzyme properties through enzyme assays is an inherently dynamic process; the steady state is reached only when one of the substrates vanishes. In this dynamic process, it is helpful to measure the indicator chemical during a regime where the reaction velocity is approximately constant, corresponding to a linear increase of the indicator chemical in time (we henceforth refer to this as a ‘linear growth’ regime). Often this means estimating the reaction velocity through initial rate experiments, with the goal of avoiding complications due to reversible reactions. Linear growth is helpful as a sanity check that things are proceeding as they should (typically using the ‘eyeball norm’). Moreover, experimental noise can increase the error when attempting to implement an accurate fit in nonlinear systems - restricting oneself to a linear growth regime significantly reduces this issue.

For these reasons, we use singular perturbation theory [2, 11] to analyse the mathematical systems we derive, and to understand how to determine the kinetic parameter values of the primary enzyme in such a system. This provides a significant reduction of the complexity of the system, and thus allows us to minimize issues associated with experimental noise, as we aim to determine functional forms for the measurable reaction velocities. Moreover, as exhibited in recent examples [5, 6, 7, 8, 12], singular perturbation theory is particularly well suited to understanding systems of chemical reactions exhibiting a significant separation of timescales, so we will use a similar approach here.

In our analysis, we assume that the timescales of enzyme complex formation are much shorter than the other reaction timescales in our problem, and that the effective reaction velocities are governed by Michaelis–Menten-type laws. The suitability of the Michaelis–Menten equations in many scenarios, including but not limited to standard coupled enzyme assays, has been investigated in recent work by Schnell and colleagues, for example [7, 8, 24], using singular perturbation theory.

To demonstrate how our theoretical results can be used to characterize enzymes, we perform several coupled enzyme assays to characterize a putative omega-aminotransferase (also known as an aminotransaminase or an omega-amino-acid amino transferase) enzyme, CnAptA, for two omega-amino acid substrates. Omega-aminotransferases belong to class III aminotransferases; these catalyse the transamination of omega-amino acids such as beta-alanine or gamma-aminobutyric acid by transferring the amino group to a keto acid, using pyridoxal 5’-phosphate (PLP) as a co-factor [16]. Aminotransferases are increasingly important biocatalysts for the synthesis of industrially relevant chiral compounds and in the pharmaceutical industry for the production of optically pure amines and amino alcohols necessary for the synthesis of important drugs [21]. Moreover, aminotransferases are particularly important in synthetic biology because they are the key to creating biosustainable production routes to many important platform chemicals. That is, carboxylic acids make up eight of the top twelve platform chemicals selected by the US Department of Energy that can be derived from biomass [26]. As current industrial methods to produce these acids often involve fossil fuels, biological production routes provide environmentally sustainable alternatives.

The outline of this paper is as follows. In §2, we briefly recap the mathematical analysis of the standard enzyme assay, which allows us to put the main results of this paper into an appropriate context. In §3, we derive a mathematical model for the non-standard coupled enzyme assay of interest in this paper, which has the reaction network shown in Figure 1. We then explore how the non-standard assay differs from the standard one and determine how the measured reaction velocity relates to the kinetic parameters of the primary enzyme in the non-standard case. We do this using an asymptotic analysis to systematically reduce the complexity of the system, and we present the main distinguished limit (*i.e.* the dominant balance of most practical relevance) of the system in §3.1. We present the remaining distinguished

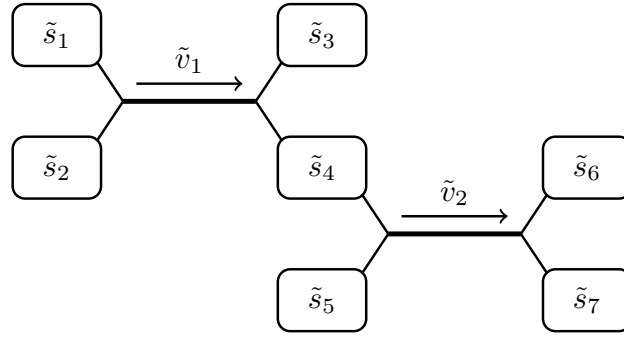


Fig. 2 Abstract network diagram for a standard enzyme assay. The primary enzyme has reaction velocity \tilde{v}_1 and the auxiliary enzyme has reaction velocity \tilde{v}_2 .

limit in §3.2, where we find that the distinguished nature of this limit is apparent only in a non-measurable chemical during the linear regime or during the depletion regime, and that the measurable system behaviour in the linear regime is captured by a sub-limit of the strong reverse reaction regime treated in §3.1. We summarise our theoretical results in §4, outlining the procedure that should be implemented to use experimental data to infer the kinetic parameters of the primary enzyme. In the same section, we compare our proposed method with a naive nonlinear fit, verifying that our method is preferable. In §5, we demonstrate that our results can be used effectively in characterizing enzymes by performing the coupled enzyme assay modelled in this paper and applying our theoretical results. Finally, in §6, we discuss our results in the context of their significance for experimental approaches involving non-standard coupled enzyme assays.

2 Standard coupled enzyme assay

We start by briefly recapping the mathematical model for the standard coupled enzyme assay, presented in broad dimensional terms in [15]. In this section, we present the relevant chemicals involved in abstract terms before relating them to the chemicals labelled in Figure 1 for the non-standard coupled enzyme assay we consider in the next section. Our goal in this section is to recap how to determine the kinetic parameters of the primary enzyme in the system shown in Figure 2, where we assume knowledge of the kinetic parameters of the auxiliary enzyme. Experimentally, we are able to measure the concentration of \tilde{s}_6 , and so our overarching goal is to understand how to use this measurement to infer the kinetic parameters that characterize the primary enzyme.

In general, the kinetic parameters used to characterize an enzyme are intrinsic to the ordinary differential equations (ODEs) assumed to govern this system, using Michaelis–Menten-type laws to quantify each reaction velocity [17]. The network in Figure 2 implies the following seven dimensional ODEs

$$\frac{d\tilde{s}_1}{d\tilde{t}} = -\tilde{v}_1, \quad \frac{d\tilde{s}_2}{d\tilde{t}} = -\tilde{v}_1, \quad \frac{d\tilde{s}_3}{d\tilde{t}} = \tilde{v}_1, \quad \frac{d\tilde{s}_4}{d\tilde{t}} = \tilde{v}_1 - \tilde{v}_2, \quad \frac{d\tilde{s}_5}{d\tilde{t}} = -\tilde{v}_2, \quad \frac{d\tilde{s}_6}{d\tilde{t}} = \tilde{v}_2, \quad \frac{d\tilde{s}_7}{d\tilde{t}} = \tilde{v}_2. \quad (1)$$

We use the convention that dimensional quantities are marked with a tilde, whereas dimensionless quantities are not. The initial conditions are

$$\tilde{s}_1(0) = \tilde{\alpha}, \quad \tilde{s}_2(0) = \tilde{\beta}, \quad \tilde{s}_3(0) = 0, \quad \tilde{s}_4(0) = 0, \quad \tilde{s}_5(0) = \tilde{\gamma}, \quad \tilde{s}_6(0) = 0, \quad \tilde{s}_7(0) = 0, \quad (2)$$

which corresponds to a system initially containing the minimal number of substrates necessary for both reactions to take place: \tilde{s}_1 , \tilde{s}_2 , and \tilde{s}_5 , then instantaneously adding the primary and auxiliary enzyme such that the entire system is well-mixed. We assume that the affinity a given substrate has for an enzyme is independent of the other substrates. Therefore, the reaction velocities follow Michaelis–Menten kinetics generalized to multiple substrates [1]

$$\tilde{v}_1 = \tilde{k}^{(1)} \left(\frac{\tilde{s}_1}{\tilde{K}_1^{(1)} + \tilde{s}_1} \right) \left(\frac{\tilde{s}_2}{\tilde{K}_2^{(1)} + \tilde{s}_2} \right), \quad \tilde{v}_2 = \tilde{k}^{(2)} \left(\frac{\tilde{s}_4}{\tilde{K}_4^{(2)} + \tilde{s}_4} \right) \left(\frac{\tilde{s}_5}{\tilde{K}_5^{(2)} + \tilde{s}_5} \right). \quad (3)$$

This problem was recently investigated through a singular perturbation analysis in [7] to determine when the Michaelis–Menten reaction-type formulation of the problem is valid by considering the intermediate enzyme complexes formed during each reaction. We will assume throughout this paper that we are in this regime, essentially assuming the reactant-stationary assumption holds, *i.e.* that the primary substrates are approximately constant as the primary intermediate complexes form.

In (3), we use the convention that a subscript refers to a relationship with a particular chemical, and a bracketed superscript refers to a particular reaction. For example, the parameter $\tilde{K}_4^{(2)}$ refers to the Michaelis constant for \tilde{s}_4 in the

reaction velocity \tilde{v}_2 . In general, $\tilde{\alpha}$ and $\tilde{\beta}$ can be viewed as control parameters that can be varied between experiments; all remaining parameters should remain constant between different experiments (at fixed temperatures and pHs), such assumptions being appropriate to the experimental approach we describe in §5. Hence, we are interested in inferring the three parameters with a bracketed superscript of 1, which are $\tilde{k}^{(1)}$, $\tilde{K}_1^{(1)}$, and $\tilde{K}_2^{(1)}$, from measurements of \tilde{s}_6 .

It is apparent from the system (1)–(3), as well as from Figure 2, that the concentrations for \tilde{s}_3 , \tilde{s}_6 , and \tilde{s}_7 decouple from the rest of the system. We include them here to keep the same notation as for §3. Moreover, we are able to reduce the system to two ODEs by noting the five following linearly independent conserved quantities:

$$\tilde{s}_1 + \tilde{s}_3 = \tilde{\alpha}, \quad \tilde{s}_2 + \tilde{s}_3 = \tilde{\beta}, \quad \tilde{s}_5 + \tilde{s}_6 = \tilde{\gamma}, \quad \tilde{s}_2 + \tilde{s}_4 + \tilde{s}_6 = \tilde{\beta}, \quad \tilde{s}_6 = \tilde{s}_7. \quad (4)$$

We delay the system reduction to discuss first the following standard assumptions made when carrying out a coupled enzyme assay, as this will aid in deciding with which ODEs to work. First, we assume that the maximum rate of reaction mediated by the auxiliary enzyme is much faster than for the primary enzyme, so that $\tilde{k}^{(2)} \gg \tilde{k}^{(1)}$. As the maximum rate of each reaction is proportional to the concentration of the corresponding enzyme, this can be ensured by using a high enough ratio of auxiliary to primary enzyme. Second, we will initially saturate the system with large quantities of \tilde{s}_5 , so that $\tilde{\gamma} \gg \tilde{K}_5^{(2)}$, thus reducing the complexity of \tilde{v}_2 in (3). Third, we will initially saturate the system with either \tilde{s}_1 or \tilde{s}_2 (so that $\tilde{\alpha} \gg \tilde{K}_1^{(1)}$ or $\tilde{\beta} \gg \tilde{K}_2^{(1)}$) if we are interested in determining $\tilde{K}_2^{(1)}$ or $\tilde{K}_1^{(1)}$, respectively. This reduces the complexity of \tilde{v}_1 in (3). This final assumption is not required to facilitate a solution of the reduced system, though it does reduce the number of fitting parameters. In §3, we only make the first two assumptions in our analysis. Not making the third assumption allows us to consider the separate cases of saturating \tilde{s}_1 and \tilde{s}_2 at the same time.

We will present the method for determining $\tilde{K}_1^{(1)}$, essentially the case with a large initial saturation of \tilde{s}_2 . The approach to determining $\tilde{K}_2^{(1)}$ is equivalent for the standard coupled enzyme assay discussed in this section, but this will not be the case for the non-standard assay we consider in §3. If we initially saturate the system with large quantities of \tilde{s}_2 , it is convenient to use the following two ODEs to describe the system:

$$\frac{d\tilde{s}_1}{d\tilde{t}} = -\frac{\tilde{k}^{(1)}\tilde{s}_1}{\tilde{K}_1^{(1)} + \tilde{s}_1}, \quad \frac{d\tilde{s}_4}{d\tilde{t}} = \frac{\tilde{k}^{(1)}\tilde{s}_1}{\tilde{K}_1^{(1)} + \tilde{s}_1} - \frac{\tilde{k}^{(2)}\tilde{s}_4}{\tilde{K}_4^{(2)} + \tilde{s}_4}, \quad \tilde{s}_1(0) = \tilde{\alpha}, \quad \tilde{s}_4(0) = 0. \quad (5)$$

This is the standard system to consider when using coupled enzyme assays, investigated in the literature in [15], as discussed in §1.

To implement our systematic asymptotic analysis, we form dimensionless variables as in Table 1, and obtain the following system

$$\frac{ds_1}{dt} = -\varepsilon a \frac{s_1}{K_1^{(1)} + s_1}, \quad \frac{ds_4}{dt} = \frac{s_1}{K_1^{(1)} + s_1} - \frac{s_4}{1 + \varepsilon s_4}, \quad s_1(0) = 1, \quad s_4(0) = 0, \quad (6)$$

where typical values and definitions of the dimensional and dimensionless parameters are given in Tables 2 and 3, respectively. The most important definition here is $\varepsilon = \tilde{k}^{(1)}/\tilde{k}^{(2)}$ which, as discussed above, is small in coupled enzyme assays. We also define $a = \tilde{K}_4^{(2)}/\tilde{\alpha}$ and $K_1^{(1)} = \tilde{K}_1^{(1)}/\tilde{\alpha}$.

In the limit of $\varepsilon \rightarrow 0$, with $t = O(1)$, (6) has leading-order solution

$$s_1 = 1, \quad s_4 = \frac{1 - e^{-t}}{K_1^{(1)} + 1}, \quad (7)$$

which corresponds to the dimensional solution

$$\tilde{s}_6 = \frac{\tilde{\alpha}\tilde{k}^{(1)}}{\tilde{K}_1^{(1)} + \tilde{\alpha}} \left(\tilde{t} + \frac{\tilde{K}_4^{(2)}}{\tilde{k}^{(2)}} \left[\exp\left\{-\tilde{k}^{(2)}\tilde{t}/\tilde{K}_4^{(2)}\right\} - 1 \right] \right), \quad (8)$$

for our measurable chemical. The long-time limit of (8) is

$$\tilde{s}_6 \sim \frac{\tilde{\alpha}\tilde{k}^{(1)}}{\tilde{K}_1^{(1)} + \tilde{\alpha}} \left(\tilde{t} - \frac{\tilde{K}_4^{(2)}}{\tilde{k}^{(2)}} \right) \quad \text{as } \tilde{t} \rightarrow \infty. \quad (9)$$

Therefore, the kinetic parameters $\tilde{k}^{(1)}$ and $\tilde{K}_1^{(1)}$ can be inferred by measuring the long-term (by which we mean $\tilde{K}_4^{(2)}/\tilde{k}^{(2)} \ll \tilde{t} \ll \tilde{\alpha}/\tilde{k}^{(1)}$, for reasons noted below) production rate of \tilde{s}_6 for different values of $\tilde{\alpha}$. A similar procedure, but with saturating \tilde{s}_1 instead of \tilde{s}_2 , allows the inference of $\tilde{k}^{(1)}$ (again) and $\tilde{K}_2^{(1)}$. The solution (9) tells us that the long-term production rate will be $\tilde{\alpha}\tilde{k}^{(1)}/(\tilde{K}_1^{(1)} + \tilde{\alpha})$, as can be seen in Figure 3, and that we require $\tilde{t} \gg \tilde{K}_4^{(2)}/\tilde{k}^{(2)}$ before \tilde{s}_6 production will appear to be at a constant rate ($\tilde{K}_4^{(2)}/\tilde{k}^{(2)} \approx 200$ s in Figure 3). However, this constant rate

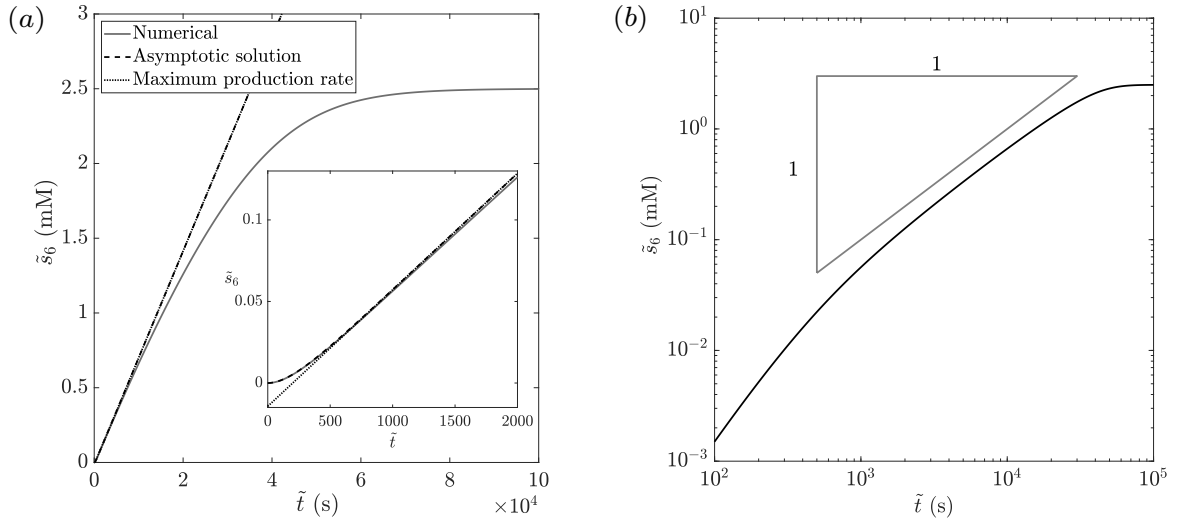


Fig. 3 The measurable chemical \tilde{s}_6 as the numerical solution to (1), the asymptotic solution (8) (prior to the depletion dynamics), and the long-time asymptotic solution (9) (also prior to the depletion dynamics). We use the parameter values in Table 2, choosing $\tilde{\alpha} = 2.5$ mM and using $\tilde{\beta} = \tilde{\gamma} = 50$ mM to focus on the effect of a limiting initial concentration of \tilde{s}_1 , rather than \tilde{s}_2 or \tilde{s}_5 .

will not occur indefinitely - given that $ds_1/dt = O(a\varepsilon)$, the production rate of \tilde{s}_6 will start to deviate from constant when $t = O(1/a\varepsilon)$, equivalent to $\tilde{t} = O(\tilde{\alpha}/\tilde{k}^{(1)}) = O(2.5 \times 10^4 \text{ s})$ in Figure 3. In the language of matched asymptotic expansions, this corresponds to an outer scaling, a timescale over which (what we term) the depletion dynamics occur. For brevity, we do not consider the depletion dynamics further for the standard enzyme assay, but we will for the non-standard enzyme assay.

3 Non-standard coupled enzyme assay

We now consider the main problem we are concerned with in this paper - the determination of the kinetic parameters $\tilde{k}^{(1)}$, $\tilde{K}_1^{(1)}$, and $\tilde{K}_2^{(1)}$ of the primary enzyme in the system shown in Figure 1, where we assume knowledge of the kinetic parameters of the auxiliary enzyme. Experimentally, we are able to measure the concentration of NADH (*i.e.* \tilde{s}_6) in the system, and so our overarching goal is to understand how to use this measurement to infer the kinetic parameters that characterize the primary enzyme for a range of different substrates. The differences between the standard coupled enzyme assay in §2 and the problem we consider in this section are that the latter involves a reversible auxiliary reaction and that \tilde{s}_1 is a product/substrate of the forward/backward auxiliary reaction, rather than the auxiliary reaction being unidirectional and \tilde{s}_1 being decoupled from the auxiliary reaction in §2. The reaction network we consider is particularly relevant to the class of enzymes known as aminotransferases¹, as discussed in the Introduction.

The governing equations for the network shown in Figure 1 consist of the following seven ODEs

$$\begin{aligned} \frac{d\tilde{s}_1}{d\tilde{t}} &= -\tilde{v}_1 + \tilde{v}_2 - \tilde{v}_{-2}, & \frac{d\tilde{s}_2}{d\tilde{t}} &= -\tilde{v}_1, & \frac{d\tilde{s}_3}{d\tilde{t}} &= \tilde{v}_1, \\ \frac{d\tilde{s}_4}{d\tilde{t}} &= \tilde{v}_1 - \tilde{v}_2 + \tilde{v}_{-2}, & \frac{d\tilde{s}_5}{d\tilde{t}} &= -\tilde{v}_2 + \tilde{v}_{-2}, & \frac{d\tilde{s}_6}{d\tilde{t}} &= \tilde{v}_2 - \tilde{v}_{-2}, & \frac{d\tilde{s}_7}{d\tilde{t}} &= \tilde{v}_2 - \tilde{v}_{-2}, \end{aligned} \quad (10)$$

where we define each variable in Table 1. As in the previous section, we use the convention that dimensional quantities are marked with a tilde, whereas dimensionless quantities are not. Here, the three reaction velocities are

$$\tilde{v}_1 = \tilde{k}^{(1)} \left(\frac{\tilde{s}_1}{\tilde{K}_1^{(1)} + \tilde{s}_1} \right) \left(\frac{\tilde{s}_2}{\tilde{K}_2^{(1)} + \tilde{s}_2} \right), \quad (11a)$$

$$\tilde{v}_2 = \tilde{k}^{(2)} \left(\frac{\tilde{s}_4}{\tilde{K}_4^{(2)} + \tilde{s}_4} \right) \left(\frac{\tilde{s}_5}{\tilde{K}_5^{(2)} + \tilde{s}_5} \right), \quad (11b)$$

$$\tilde{v}_{-2} = \tilde{k}^{(-2)} \left(\frac{\tilde{s}_1}{\tilde{K}_1^{(-2)} + \tilde{s}_1} \right) \left(\frac{\tilde{s}_6}{\tilde{K}_6^{(-2)} + \tilde{s}_6} \right) \left(\frac{\tilde{s}_7}{\tilde{K}_7^{(-2)} + \tilde{s}_7} \right), \quad (11c)$$

each following Michaelis–Menten kinetics generalized to multiple substrates [1], under the assumption that the affinity a given substrate has for an enzyme is independent of the other substrates.

Original variable	Description	Nondimensionalisation
\tilde{s}_1	Pyruvate	$\tilde{s}_1 = \tilde{\alpha}s_1$
\tilde{s}_2	Substrate	$\tilde{s}_2 = \tilde{\beta}s_2$
\tilde{s}_3	Aldehyde	$\tilde{s}_3 = \tilde{\beta}s_3$
\tilde{s}_4	Alpha-L-alanine	$\tilde{s}_4 = (\tilde{k}^{(1)}\tilde{K}_4^{(2)}/\tilde{k}^{(2)})s_4$
\tilde{s}_5	NAD ⁺	$\tilde{s}_5 = \tilde{\gamma}s_5$
\tilde{s}_6	NADH	$\tilde{s}_6 = (\tilde{k}^{(1)}\tilde{K}_4^{(2)}/\tilde{k}^{(2)})s_6$
\tilde{t}	Time	$\tilde{t} = (\tilde{K}_4^{(2)}/\tilde{k}^{(2)})t$

Table 1 Definitions of the dimensional and dimensionless variables in our system.

Typical parameter size	Notes
$\tilde{k}^{(1)} \approx 0.1 \mu\text{Ms}^{-1}$	To be determined
$\tilde{K}_1^{(1)} \approx 1 \text{ mM}$	To be determined
$\tilde{K}_2^{(1)} \approx 1 \text{ mM}$	To be determined
$\tilde{k}^{(2)} = 10 \mu\text{Ms}^{-1}$	Experimental choice
$\tilde{K}_4^{(2)} = 2 \text{ mM}$	From <i>Bacillus subtilis</i> [27], though 0.45 – 14 mM from other organisms [4, 10, 20, 25]
$\tilde{K}_5^{(2)} = 0.2 \text{ mM}$	From <i>Bacillus subtilis</i> [27], though 0.04 – 0.3 mM from other organisms [4, 10, 20, 25]
$\tilde{k}^{(-2)} = 0 - 2 \text{ mM s}^{-1}$	Value determined by choice of $\tilde{k}^{(2)}$ and assay environment.
$\tilde{K}_1^{(-2)} = 0.5 \text{ mM}$	From <i>Bacillus subtilis</i> [27], though 0.22 – 1.45 mM from other organisms [4, 10, 20, 25]
$\tilde{K}_6^{(-2)} = 0.02 \text{ mM}$	From <i>Bacillus subtilis</i> [27], though 0.02 – 0.1 mM from other organisms [4, 10, 20, 25]
$\tilde{K}_7^{(-2)} = 40 \text{ mM}$	From <i>Bacillus subtilis</i> [27], though 28 – 67 mM from other organisms [4, 10, 20, 25]
$\tilde{\alpha} = 1 - 5 \text{ mM}$	Control parameter
$\tilde{\beta} = 1 - 5 \text{ mM}$	Control parameter
$\tilde{\gamma} = 5 \text{ mM}$	Experimental choice

Table 2 Dimensional parameters and their typical size where known. For the auxiliary enzyme, we use values taken from [27] in our numerical simulations. We give an idea of the range of these parameter values by also providing data from four other organisms: *Enterobacter aerogenes* [4], *Mycobacterium tuberculosis* [10], *Phormidium lapideum* [20], and *Rhodospseudomonas capsulata* [10].

In conjunction with the governing equations (10), we use initial conditions

$$\tilde{s}_1(0) = \tilde{\alpha}, \quad \tilde{s}_2(0) = \tilde{\beta}, \quad \tilde{s}_3(0) = 0, \quad \tilde{s}_4(0) = 0, \quad \tilde{s}_5(0) = \tilde{\gamma}, \quad \tilde{s}_6(0) = 0, \quad \tilde{s}_7(0) = 0, \quad (12)$$

where $\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma} > 0$. This corresponds to a system initially containing pyruvate (\tilde{s}_1), an omega-amino acid substrate (\tilde{s}_2), and NAD⁺ (\tilde{s}_5), with no other substrates present in the reaction network, then instantaneously adding the primary and auxiliary enzyme such that the entire system is well-mixed. Hence, there are thirteen dimensional parameters. We reiterate that our goal is to infer the three parameters $\tilde{k}^{(1)}$, $\tilde{K}_1^{(1)}$, and $\tilde{K}_2^{(1)}$ that correspond to the primary enzyme, assuming that we know the remaining ten parameters, which correspond to the auxiliary enzyme and the initial chemical concentrations. We provide typical parameter values in Table 2.

As would be expected on physical grounds, the inclusion of a reverse reaction for the auxiliary enzyme causes less NADH to be made over a given time (Figure 4a). Moreover, the reverse reaction can result in a linear phase that is very short (Figure 4b), making the standard approach of calculating a steady production rate difficult to implement in practice. To this end, we will characterize the different possible types of NADH production through a systematic asymptotic analysis - see [2, 11] for general descriptions of the approach.

Before we nondimensionalize the system, we note from (10) and (12) the five following linearly independent conserved quantities:

$$\tilde{s}_1 + \tilde{s}_4 = \tilde{\alpha}, \quad \tilde{s}_2 + \tilde{s}_3 = \tilde{\beta}, \quad \tilde{s}_5 + \tilde{s}_6 = \tilde{\gamma}, \quad \tilde{s}_2 + \tilde{s}_4 + \tilde{s}_6 = \tilde{\beta}, \quad \tilde{s}_6 = \tilde{s}_7. \quad (13)$$

Thus, we may immediately reduce the number of ODEs from seven to two. However, the asymptotic analysis we will carry out is more intuitive if we keep most of the ODEs; the only ones we disregard are those for $d\tilde{s}_3/d\tilde{t}$ and $d\tilde{s}_7/d\tilde{t}$ in (10). For the former, we do this since the aldehyde produced from the primary enzyme (\tilde{s}_3) decouples from the rest of the system, so we are able to ignore this ODE henceforth, noting that \tilde{s}_3 can be determined from (13). For the latter, the concentration of ammonia is always equivalent to the concentration of NADH ($\tilde{s}_7 \equiv \tilde{s}_6$), so we will replace \tilde{s}_7 by \tilde{s}_6 henceforth.

We form the following dimensionless variables

$$\tilde{t} = (\tilde{K}_4^{(2)}/\tilde{k}^{(2)})t, \quad \tilde{s}_1 = \tilde{\alpha}s_1, \quad \tilde{s}_2 = \tilde{\beta}s_2, \quad (\tilde{s}_4, \tilde{s}_6) = (\tilde{k}^{(1)}\tilde{K}_4^{(2)}/\tilde{k}^{(2)})(s_4, s_6), \quad \tilde{s}_5 = \tilde{\gamma}s_5, \quad (14)$$

¹ Example KEGG classifications for this case would be EC 2.6.1.18 for the primary enzyme and EC 1.4.1.1 for the auxiliary enzyme.

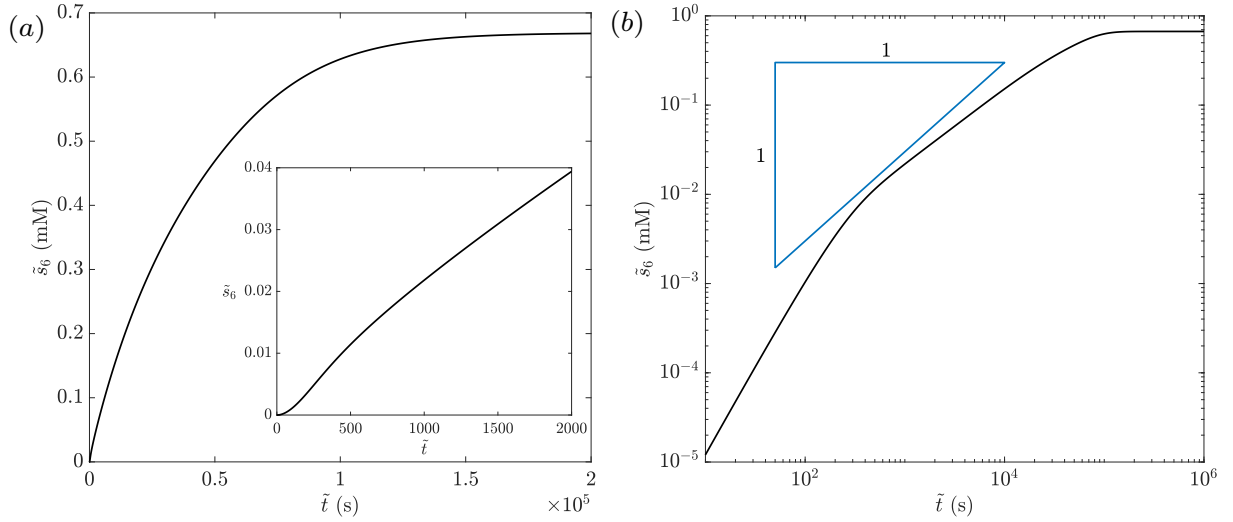


Fig. 4 The measurable chemical \tilde{s}_6 as the numerical solution to (10) in (a) a linear plot and (b) a log-log plot. We use the parameter values in Table 2, with $\tilde{k}^{(-2)} = 0.5 \text{ mM s}^{-1}$, $\tilde{\alpha} = 2.5 \text{ mM}$, and $\tilde{\beta} = 2.5 \text{ mM}$.

which we also summarize in Table 2. The timescale here (*i.e.* $\tilde{K}_4^{(2)}/\tilde{k}^{(2)}$) is that of initial alpha-L-alanine (s_4) and NADH (s_6) formation, and is 200 s for the parameters in Table 2. For the substrate dimensionless variables we scale pyruvate, beta-alanine, and NAD⁺ with their initial conditions and alpha-L-alanine and NADH with the order of magnitude of alpha-L-alanine present during the regime of linear NADH growth, which our analysis will show to be approximately $50 \mu\text{M}$. From (14), we obtain the dimensionless governing equations

$$\frac{ds_1}{dt} = -\varepsilon a (v_1 - v_2 + v_{-2}), \quad (15a)$$

$$\frac{ds_2}{dt} = -\varepsilon b v_1, \quad (15b)$$

$$\frac{ds_4}{dt} = v_1 - v_2 + v_{-2}, \quad (15c)$$

$$\frac{ds_5}{dt} = -\varepsilon c (v_2 - v_{-2}), \quad (15d)$$

$$\frac{ds_6}{dt} = v_2 - v_{-2}, \quad (15e)$$

where the three dimensionless reaction velocities are defined as

$$v_1 = \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right), \quad (16a)$$

$$v_2 = \left(\frac{s_4}{1 + \varepsilon s_4} \right) \left(\frac{s_5}{\delta + s_5} \right), \quad (16b)$$

$$v_{-2} = k^{(-2)} K \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) \left(\frac{s_6}{K_6^{(-2)} + s_6} \right) \left(\frac{s_6}{1 + \varepsilon K s_6} \right). \quad (16c)$$

The initial conditions of the dimensionless system are

$$s_1(0) = 1, \quad s_2(0) = 1, \quad s_4(0) = 0, \quad s_5(0) = 1, \quad s_6(0) = 0. \quad (17)$$

We provide the definitions of the dimensionless parameters in Table 3, from which we note that ε , δ , and K are all much smaller than 1. Additionally, we note that the value of $k^{(-2)}$, representing the ratio of the backward to forward rates controlled by the auxiliary enzyme, can vary across many orders of magnitude.

We reiterate that our goal is to understand how to relate the measurement of NADH (\tilde{s}_6) in the linear production regime to the values of the kinetic parameters governing the primary enzyme. Our approach is to comprehensively investigate how the system behaves for different magnitudes of $k^{(-2)}$, with a focus on understanding when a linear regime for NADH production is appropriate and what the measurement of NADH in this regime will tell us about the primary enzyme. To this end, we perform an asymptotic analysis that exploits the small parameters ε , δ , and K (noting that the three limits all commute), with a focus on analysing how the asymptotic size of $k^{(-2)}$ affects the system behaviour. We emphasize that the parameter ε can always be made to be small by suitable choices of the

$$\begin{aligned}
\varepsilon &= \frac{\tilde{k}^{(1)}}{\tilde{k}^{(2)}} \approx 0.01 & b &= \frac{\tilde{K}_4^{(2)}}{\tilde{\beta}} = 0.4 - 2 & K_6^{(-2)} &= \frac{\tilde{K}_6^{(-2)}\tilde{k}^{(2)}}{\tilde{K}_4^{(2)}\tilde{k}^{(1)}} \approx 1 \\
K_1^{(1)} &= \frac{\tilde{K}_1^{(1)}}{\tilde{\alpha}} \approx 0.2 - 1 & c &= \frac{\tilde{K}_4^{(2)}}{\tilde{\gamma}} = 0.8 & k^{(-2)} &= \frac{\tilde{k}^{(-2)}}{\tilde{k}^{(2)}} = 0 - 10^3 \\
K_2^{(1)} &= \frac{\tilde{K}_2^{(1)}}{\tilde{\beta}} \approx 0.2 - 1 & \delta &= \frac{\tilde{K}_5^{(2)}}{\tilde{\gamma}} = 0.04 & K &= \frac{\tilde{K}_4^{(2)}}{\tilde{K}_7^{(-2)}} = 0.05 \\
a &= \frac{\tilde{K}_4^{(2)}}{\tilde{\alpha}} = 0.4 - 2 & K_1^{(-2)} &= \frac{\tilde{K}_1^{(-2)}}{\tilde{\alpha}} \approx 0.1 - 0.5
\end{aligned}$$

Table 3 The dimensionless parameters in our system, obtained using the data in Table 2. The approximately equal signs arise from uncertainty in the kinetic parameters of the primary enzyme.

	Asymptotic timescales	Solutions	Dimensionless parameters for s_6/S_6
Strong reverse reaction	$t = O(1)$	One-dimensional ODE representation in (19), (21)	V in (20), γ in (22), and $K_6^{(-2)}$
$k^{(-2)} = O(1/K)$	$t = O(1/\varepsilon)$	Implicit quadrature representation in (40)–(41)	$a, b, K_1^{(1)}, K_2^{(1)}, K_1^{(-2)}$, and $k^{(-2)}K$
Weak reverse reaction	$t = O(1)$	Explicit analytic solutions in (45)	V in (20)
$k^{(-2)} = O(\varepsilon K)$	$t = O(1/\varepsilon)$	Implicit analytic solutions in (49)	$b, K_1^{(1)}$, and $K_2^{(1)}$

Table 4 A summary of the solutions we derive in this paper for each timescale in the two distinguished limits of the problem, as well as the dimensionless parameters that control the concentration of s_6 in each asymptotic region. We give the equation number for the definition of each parameter unless it is given in Table 3.

relative concentrations of the primary and auxiliary enzyme, which scale with $\tilde{k}^{(1)}$ and $\tilde{k}^{(2)}$, respectively, and likewise for δ by choosing $\tilde{\gamma}$, the initial concentration of \tilde{s}_5 , to be significantly larger than $\tilde{K}_5^{(2)}$, the Michaelis constant for \tilde{s}_5 in the forward auxiliary reaction. Since the smallness of these parameters reduces the complexity of the system, it is experimentally favourable to choose the relative enzyme concentrations and initial conditions to make ε and δ small. We note that the smallness of K is slightly different, however. While the typical parameter values we give in Table 2 unanimously result in small values of K , there seems no fundamental reason why it should be small. To ensure the generality of our analysis, we therefore also present results for $K = O(1)$ in Appendix C.

It turns out that there are two distinguished asymptotic limits in this system (so that all other regimes with ε , δ , and K each small are sub-cases of these two), whereby $k^{(-2)} = O(1/K) \gg 1$ and $k^{(-2)} = O(\varepsilon K) \ll 1$, respectively. We investigate the former in the next section, discussing its relevance to our inference problem, and we investigate the latter in §3.2, where we show that it can be considered a sub-limit of the former case in terms of experimental measurements. Moreover, we show that the latter regime exhibits the same constant ‘long-time’ production rate of NADH as the standard coupled enzyme assay case we summarized in §2. In Table 4, we provide a summary of the results we derive in the remainder of this section.

3.1 Strong reverse reaction: $k^{(-2)} = O(1/K)$

When $k^{(-2)} = O(1/K)$, the auxiliary-enzyme-controlled reaction in the forward direction is much stronger than for the reverse direction. This scaling results in a distinguished limit where the nonlinearity of the reverse reaction is important when $t = O(1)$, which is where the linear growth regime occurs for the standard coupled enzyme assay. We now exploit the limits $\varepsilon, \delta, K \ll 1$ with $k^{(-2)}K = O(1)$. We note that while small values of ε would result in large values of $K_6^{(-2)}$, we will treat $K_6^{(-2)} = O(1)$ (as we do with the remaining dimensionless parameters in the system) as this choice keeps more terms at leading order (*i.e.* constitutes a distinguished limit) and is the case we expect from experimental parameters.

3.1.1 Linear-phase regime

We first consider the linear-phase regime, which occurs over the timescale $t = O(1)$. By ‘linear-phase’, we mean that s_6 will exhibit linear growth for some time period which we can approximate. In the language of matched asymptotic expansions, this corresponds to linear growth in the intermediate matching region between this timescale and the longer depletion timescale (which we will show corresponds to $t = O(1/\varepsilon)$).

In the limits set out in §3.1, the leading-order version of (15) for $t = O(1)$ is

$$\frac{ds_1}{dt} = 0, \tag{18a}$$

$$\frac{ds_2}{dt} = 0, \tag{18b}$$

$$\frac{ds_4}{dt} = \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right) - s_4 + k^{(-2)} K \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) \left(\frac{s_6}{K_6^{(-2)} + s_6} \right) s_6, \quad (18c)$$

$$\frac{ds_5}{dt} = 0, \quad (18d)$$

$$\frac{ds_6}{dt} = s_4 - k^{(-2)} K \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) \left(\frac{s_6}{K_6^{(-2)} + s_6} \right) s_6, \quad (18e)$$

with initial conditions (17). We may immediately deduce that, over this timescale,

$$s_1 = 1, \quad s_2 = 1, \quad s_5 = 1, \quad s_4 + s_6 = Vt, \quad (19)$$

where we define

$$V := \frac{1}{\left(K_1^{(1)} + 1 \right) \left(K_2^{(1)} + 1 \right)}. \quad (20)$$

We can therefore reduce the system (17)–(18) to the following single ODE

$$\frac{ds_6}{dt} = Vt - s_6 - \frac{\gamma s_6^2}{K_6^{(-2)} + s_6}, \quad s_6(0) = 0, \quad (21)$$

where we define

$$\gamma := \frac{k^{(-2)} K}{K_1^{(-2)} + 1}. \quad (22)$$

Although we cannot solve (21) explicitly, we may deduce from (21) that the early-time behaviour is

$$s_6 \sim \frac{Vt^2}{2} \quad \text{as } t \rightarrow 0^+. \quad (23)$$

and the long-time behaviour is

$$s_6 \sim \frac{Vt}{1 + \gamma} + \frac{\gamma K_6^{(-2)} (1 + \gamma) - V}{(1 + \gamma)^2} \quad \text{as } t \rightarrow \infty. \quad (24)$$

These asymptotic solutions are good approximations of the numerical solution to (21) in their regions of validity (Figure 5).

It will also be helpful to note that

$$s_4 \sim \frac{\gamma Vt}{1 + \gamma} \quad \text{as } t \rightarrow \infty. \quad (25)$$

Hence when measuring s_6 experimentally, we will not observe linear growth in time immediately (rather, (23) implies that we will observe a t^2 relationship), but this linear relationship will start to develop after a lag period. To help understand further the solution of the ODE (21), we first note that we can reduce its dependence from three parameters to two by making the scaling $s_6 = K_6^{(-2)} Y$. This results in the system

$$\frac{dY}{dt} = \bar{V}t - Y - \frac{\gamma Y^2}{1 + Y}, \quad Y(0) = 0, \quad (26a)$$

where we define

$$\bar{V} = V / K_6^{(-2)}. \quad (26b)$$

Solutions to (26) generally exhibit a transient region before settling to linear growth for a wide range of parameter values (Figure 6). The long-time linear growth exhibited by the ODE (26) is

$$Y \sim \frac{\bar{V}t}{1 + \gamma} \quad \text{as } t \rightarrow \infty, \quad (27)$$

and it is of interest to understand when this linear growth becomes apparent. One way to do this is to examine the menagerie of distinguished asymptotic limits of the ODE (26), which are interesting in their own right. Moreover, the analytic solutions afforded by an asymptotic analysis can be used to fit time-series data from experimental results more easily than numerical solutions of a system of ODEs. We investigate the subsidiary distinguished asymptotic

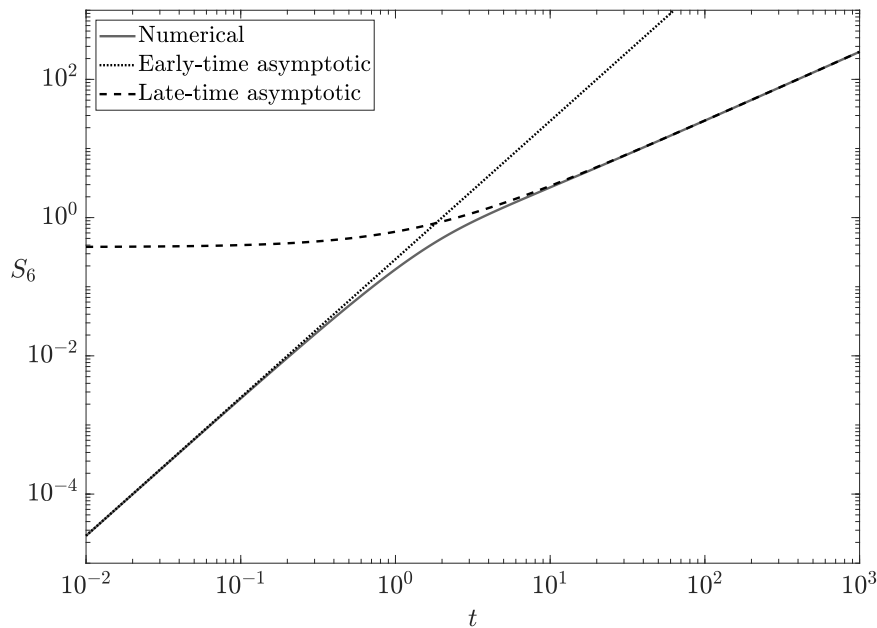


Fig. 5 A comparison of the numerical solution to the ODE (21) with the early- and late-time asymptotic solutions (23) and (24), respectively. We use parameter values $V = 0.5$, $\gamma = 1$, $K_6^{(-2)} = 1$.

limits in terms of the two parameters \bar{V} and γ in Appendix A, and we show a comparison of numerical and asymptotic solutions to (26) in Figure 6. We illustrate the four distinguished limits that we obtain from the analysis of Appendix A in Figure 7. A result from this analysis with particular importance for accurately interpreting the data is that there are two distinct linear regimes in the distinguished limit I.

While the duration of this transient region can be calculated numerically for given parameter values or approximated analytically from the asymptotic results we derive in Appendix A, it is helpful to obtain an approximate analytic estimate for when we expect to see a linear relationship for the general case. To this end, we first note that $Y \geq 0$, $\dot{Y} \geq 0$, and $Y/(1+Y) \in [0, 1)$, so we may obtain the bounds $f_1(t) \leq Y(t) \leq f_2(t)$ where f_1 and f_2 satisfy the following ODEs

$$\frac{df_1}{dt} = \bar{V}t - (1+\gamma)f_1, \quad \frac{df_2}{dt} = \bar{V}t - f_2, \quad f_1(0) = f_2(0) = 0, \quad (28)$$

which yield the following bounds

$$\frac{\bar{V}}{(1+\gamma)^2} \left[(1+\gamma)t - 1 + e^{-(1+\gamma)t} \right] \leq Y(t) < \bar{V} (t - 1 + e^{-t}). \quad (29)$$

We note that the long-time limit of the lower bound in (29) is the long-time limit (27) of the ODE (26) since the nonlinear term in the ODE becomes negligible when $Y \gg 1$. Mathematically, the lower bound of (29) appears linear when $(1+\gamma)t \gg 1$. Combining these two constraints, we expect the NADH concentration to grow linearly in time when

$$t \gg \max \left(\frac{1}{1+\gamma}, \frac{(1+\gamma)}{\bar{V}} \right). \quad (30)$$

As $1/(1+\gamma)$ is bounded above by 1, it is the second of these constraints that can lead to very large times before the linear regime is encountered. This problem is exacerbated when $\gamma \gg 1$ and $\bar{V} = V/K_6^{(-2)} \ll 1$, as can be seen by the significantly large times until linearity for Figure 6b. Moreover, we note that the two constraints in (30) approximately balance when $\bar{V} = O(\gamma^2)$ for large γ (essentially Regime III in Figure 7), and where $\bar{V} = O(1)$ for small γ . This represents a boundary between the two different constraints, where the large- \bar{V} case (on the right of Figure 7) corresponds to $t \gg 1/(1+\gamma)$. When we denote the onset of the linear regime in Figure 6, we use an asterisk when $(1+\gamma)/\bar{V} > 1/(1+\gamma)$ and a cross when $(1+\gamma)/\bar{V} < 1/(1+\gamma)$. In each case, we take for definiteness the onset to occur at 2.5 times the value of the right-hand side of (30).

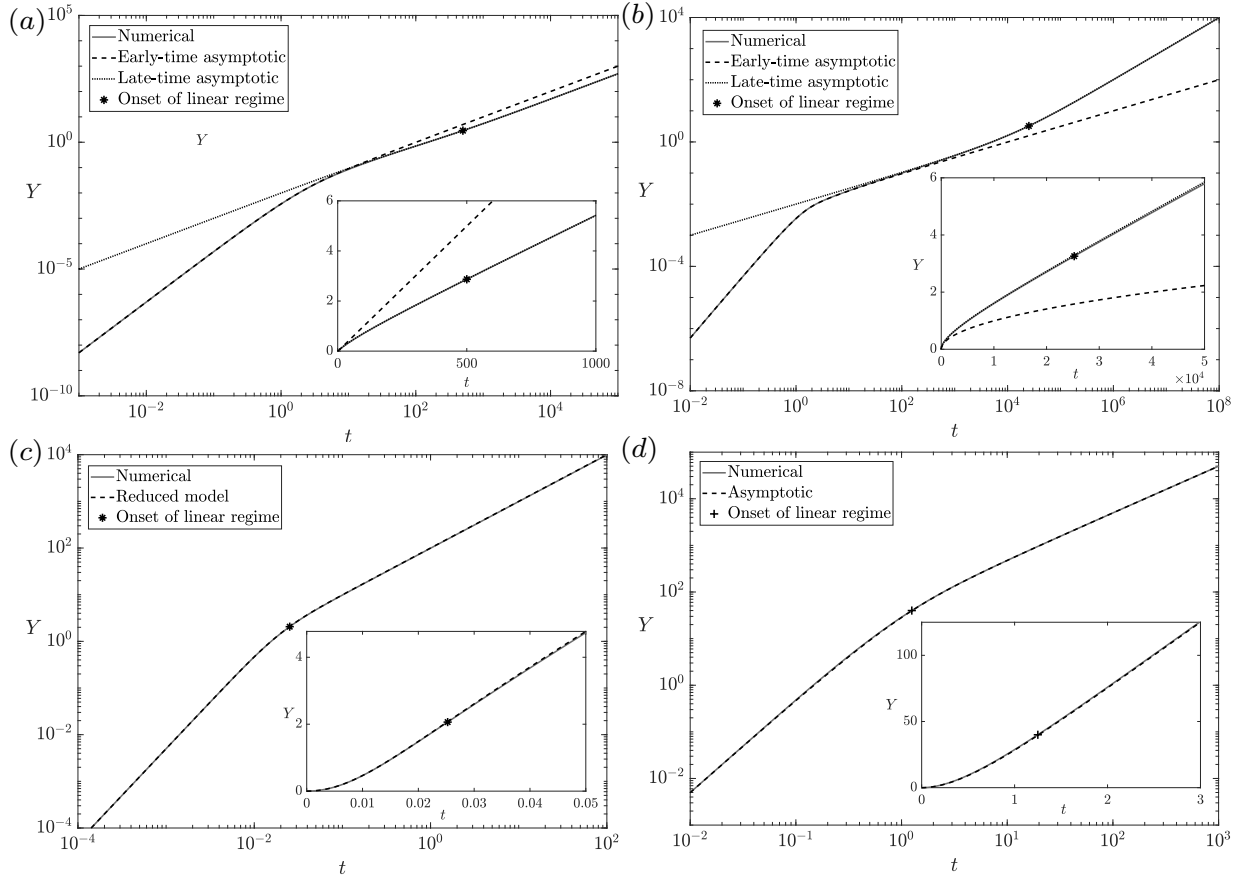


Fig. 6 A comparison of numerical and asymptotic solutions of the reduced ODE (26) for (a) $\bar{V} = 10^{-2}$, $\gamma = 10^0$, (b) $\bar{V} = 10^{-2}$, $\gamma = 10^2$, (c) $\bar{V} = 10^4$, $\gamma = 10^2$, (d) $\bar{V} = 10^2$, $\gamma = 10^0$. The onset of the linear regime is set as 2.5 times the right-hand side of (30), and is denoted by a cross/asterisk when the first/second term on the right-hand side is maximal.

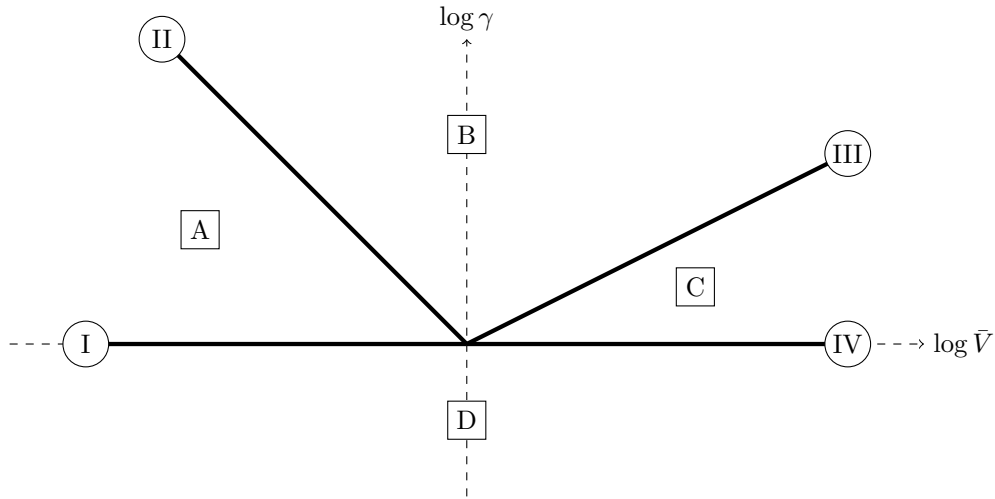


Fig. 7 The distinguished asymptotic limits of the reduced ODE (26) for asymptotically large or small values of \bar{V} and γ , denoted by dark solid lines and labelled using Roman numerals contained within a circle. Note that the distinguished limit I corresponds to $\bar{V} \ll 1$ and $\gamma = O(1)$, the distinguished limit II corresponds to $\bar{V}\gamma = O(1)$ with $\bar{V} \ll 1$ and $\gamma \gg 1$, the distinguished limit III corresponds to $\bar{V}/\gamma^2 = O(1)$ with $\bar{V} \gg 1$ and $\gamma \gg 1$, and the distinguished limit IV corresponds to $\bar{V} \gg 1$ and $\gamma = O(1)$. The sublimits of the distinguished asymptotic limits are noted using letters contained within a square.

3.1.2 Physical interpretation

To understand the physical implications of our analysis from the linear-phase regime, it is helpful to reframe our results in terms of dimensional quantities. Hence, the long-time dimensional linear behaviour of NADH, given in (24)

for the dimensionless problem, is

$$\tilde{s}_6 \sim \tilde{V} \omega \tilde{t} \quad \text{as } \tilde{t} \rightarrow \infty, \quad (31)$$

at leading order, where

$$\omega = \frac{1}{1 + \frac{\tilde{k}^{(-2)} \tilde{K}_4^{(2)}}{\tilde{k}^{(2)} \tilde{K}_7^{(-2)}} \frac{\tilde{\alpha}}{\tilde{K}_1^{(-2)} + \tilde{\alpha}}} \in (0, 1), \quad (32a)$$

$$\tilde{V} = \tilde{k}^{(1)} \frac{\tilde{\alpha}}{\tilde{K}_1^{(1)} + \tilde{\alpha}} \frac{\tilde{\beta}}{\tilde{K}_2^{(1)} + \tilde{\beta}}. \quad (32b)$$

Here, ω can be thought of as a measure of the relative strengths of the forward and backward auxiliary reactions - $\omega \rightarrow 0^+$ when the backward reaction is much stronger, but $\omega \rightarrow 1^-$ when the forward reaction is much stronger. Additionally, \tilde{V} can be thought of as the ‘natural’ observed strength of the reaction, as this is the reaction strength when the backward auxiliary reaction is unimportant, as we show in §3.2. Moreover, ω can be calculated from the experimental setup, so the standard practice of inferring the unknown parameters for a single enzyme can be used, as long as the observed reaction velocity is modified by a factor of $1/\omega$. Thus, we have determined the reaction velocity in the linear-phase regime in terms of the system parameters.

Our remaining task is to understand when we expect to observe this regime. The dimensional version of the lower constraint for the linear regime, given in (30) for the dimensionless problem, is

$$\tilde{t} \gg \max \left(\frac{\omega \tilde{K}_4^{(2)}}{\tilde{k}^{(2)}}, \frac{\tilde{K}_6^{(-2)}}{\tilde{V} \omega} \right). \quad (33)$$

Hence, this gives us the approximate time we expect to wait before observing a linear relationship between time and NADH production. The constraint (33) provides a warning against particularly low initial concentrations of the primary enzyme, pyruvate, or the substrate, as this will cause a very large lag time. One can use (33) to specify a required lag time, and determine constraints for the initial concentrations discussed above.

However, while we have identified and characterized the linear regime of NADH production, this regime is unfortunately transient; it can only last until substrate depletion starts to affect the system at leading order, which occurs when t becomes of $O(1/\varepsilon)$. We investigate this regime in the next subsection.

3.1.3 Depletion regime

To investigate the depletion regime, we introduce $T = \varepsilon t$ and analyse the limit $\varepsilon \rightarrow 0$ with $T = O(1)$. Moreover, from (24) and (25), we see that we must also make the scalings $S_4 = s_4/\varepsilon$, and $S_6 = s_6/\varepsilon$, resulting in the system

$$\varepsilon \frac{ds_1}{dT} = -a(\varepsilon v_1 - V_2 + V_{-2}), \quad (34a)$$

$$\frac{ds_2}{dT} = -bv_1, \quad (34b)$$

$$\varepsilon \frac{dS_4}{dT} = \varepsilon v_1 - V_2 + V_{-2}, \quad (34c)$$

$$\varepsilon \frac{ds_5}{dT} = -c(V_2 - V_{-2}), \quad (34d)$$

$$\varepsilon \frac{dS_6}{dT} = V_2 - V_{-2}, \quad (34e)$$

where the three reaction velocities are now defined as

$$v_1 = \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right), \quad (35a)$$

$$V_2 = \left(\frac{S_4}{1 + S_4} \right) \left(\frac{s_5}{\delta + s_5} \right), \quad (35b)$$

$$V_{-2} = k^{(-2)} K \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) \left(\frac{S_6}{\varepsilon K_6^{(-2)} + S_6} \right) \left(\frac{S_6}{1 + K S_6} \right). \quad (35c)$$

The matching conditions from the earlier timescale yield the following ‘initial conditions’

$$s_1(0) = 1, \quad s_2(0) = 1, \quad S_4(0) = 0, \quad s_5(0) = 1, \quad S_6(0) = 0. \quad (36)$$

Naively taking the limits of $\varepsilon, \delta, K \rightarrow 0$ in (34) to obtain a leading-order system would yield a duplication of information. To avoid this, we must form appropriate linear combinations of the governing equations in order to obtain enough information for a leading-order system. Using this approach, exploiting conserved quantities where possible, we obtain the following differential-algebraic system at leading-order

$$\frac{ds_2}{dT} = -b \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right), \quad (37a)$$

$$\frac{S_4}{1 + S_4} = k^{(-2)} K \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) S_6, \quad (37b)$$

$$s_1 + aS_4 = 1, \quad (37c)$$

$$s_2 + bS_4 + bS_6 = 1, \quad (37d)$$

$$s_5 + cS_6 = 1. \quad (37e)$$

Thus, we have reduced the problem to that of one ODE with four additional algebraic relationships, and the initial conditions (36). To further simplify this system, it is convenient to use (37b–d) to write

$$s_2 = f(s_1) := 1 - \frac{b}{a} \left[1 - s_1 + \frac{a}{k^{(-2)}K} \left(\frac{s_1 + K_1^{(-2)}}{s_1} \right) \left(\frac{1 - s_1}{a + 1 - s_1} \right) \right], \quad (38)$$

then to transform (37a) into a separable ODE for s_1 , obtaining

$$\frac{ds_1}{dT} = -\frac{b}{f'(s_1)} \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{f(s_1)}{K_2^{(1)} + f(s_1)} \right), \quad s_1(0) = 1. \quad (39)$$

Solving (39) allows the remaining variables to be deduced algebraically from (37b–e), (38). We are able to integrate (39) to obtain the following implicit solution

$$-bT = f(s_1) - 1 + K_2^{(1)} \log f(s_1) + bK_1^{(1)} \left(\frac{\log s_1}{a} + \frac{g(s_1)}{k^{(-2)}K(a+1)^2} \right) + K_1^{(1)} K_2^{(1)} \int_1^{s_1} \frac{f'(u)}{uf(u)} du, \quad (40a)$$

where

$$g(s_1) := \frac{K_1^{(-2)}(a+1)s_1^2 - 1}{2s_1^2} + \frac{a(a+1+K_1^{(-2)})}{a+1} \log \frac{as_1}{a+1-s_1} - (a+1+K_1^{(-2)}) \frac{1-s_1}{a+1-s_1}. \quad (40b)$$

The term with coefficient $bK_1^{(1)}$ on the right-hand side of (40a) arises from direct term-by-term integration of the ratio of the derivative of $f(s_1)$ (defined in (38)) and s_1 .

Thus, we have characterized the depletion regime. One can use (37c–d), and (40) to obtain the following functional forms for the chemical concentrations in the depletion regime

$$s_2 = f(s_1), \quad S_4 = \frac{1-s_1}{a}, \quad s_5 = 1 - c \left(\frac{1-f(s_1)}{b} - \frac{1-s_1}{a} \right), \quad S_6 = \frac{1-f(s_1)}{b} - \frac{1-s_1}{a}. \quad (41)$$

The asymptotic solution for S_6 in (41) agrees well with the numerical results for a range of values of $k^{(-2)}$ in both the linear-phase regime of the previous section and the depletion regime of this section (Figure 8).

While the transient-linear regime of the previous section is much easier to fit and has the benefit of being easier to identify from time data, this linear behaviour may only be exhibited for a short duration. Given experimental data on the NADH concentration, we can therefore estimate the three unknown parameters $k^{(1)}$, $K_1^{(1)}$, and $K_2^{(1)}$ via a nonlinear parameter fitting of (41). In practice, it is likely that either s_1 or s_2 will be saturating, corresponding to either $K_1^{(1)}$ or $K_2^{(1)}$ being small, thus reducing the number of parameters that need to be fitted at any one time. It is worth noting that, unlike in the standard assay considered in §2, the form of S_6 is not symmetric in the parameters $K_1^{(1)}$ and $K_2^{(1)}$. This is to be expected, since the symmetry in the reaction network is broken in the non-standard assay.

To identify when we may be able to use the transient-linear regime to fit data, we now estimate an upper bound for this regime. While it is possible to determine this by obtaining an early-time solution to the system (37) to calculate a correction to the long-time solution from the earlier time regime (24), the resultant expression is cumbersome and not particularly insightful. Instead, we use the following more crude approximation for an upper limit to the linear regime:

$$t \ll 1/a\varepsilon, \quad (42)$$

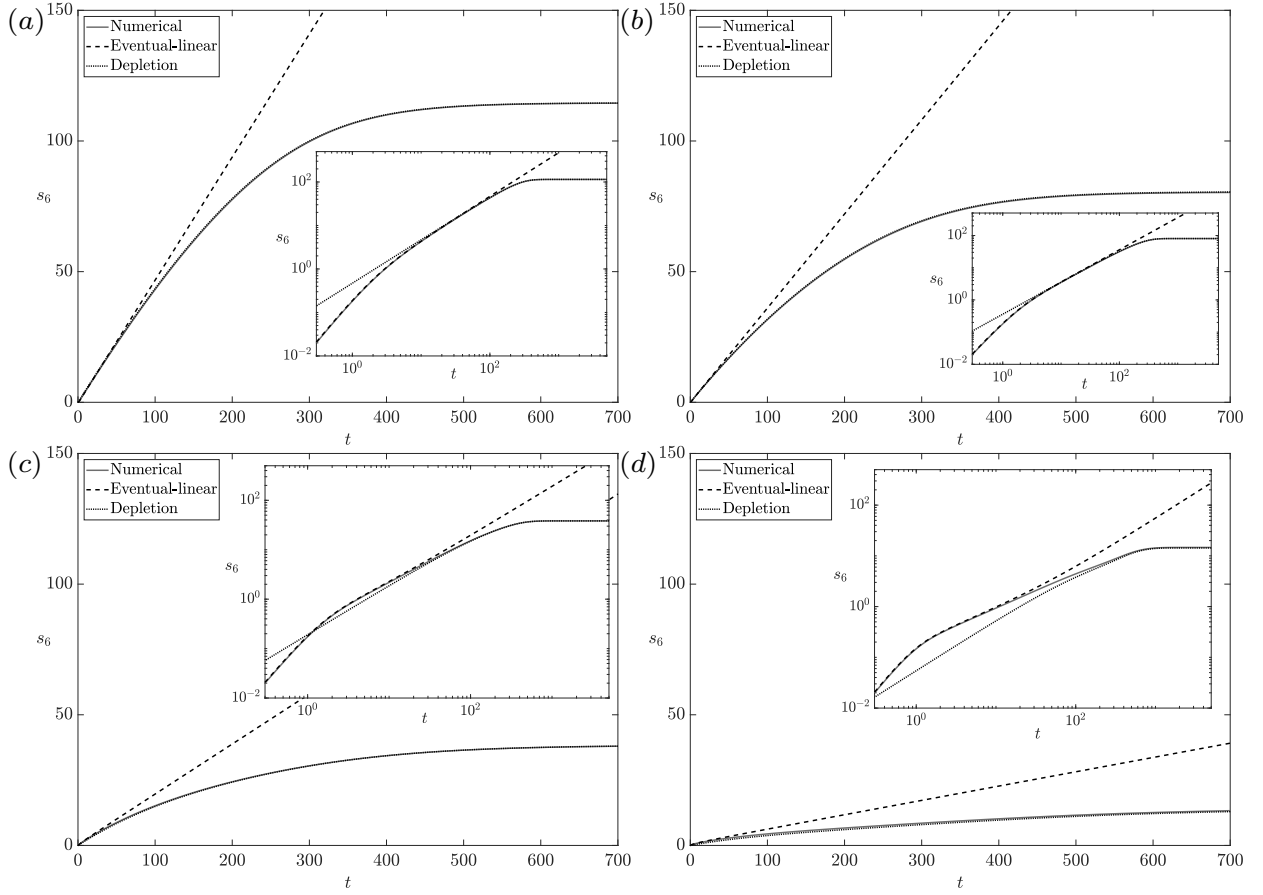


Fig. 8 A comparison of numerical solutions of the full system (15) to the asymptotic or reduced solutions in the linear-phase regime from the previous section and in the depletion regime (40)–(41). We use the parameter values in Tables 2 and 3, with $\tilde{\alpha} = \tilde{\beta} = 2.5$ mM, and varying $k^{(-2)}$ in each sub-figure such that (a) $k^{(-2)}K = 0.1$, (b) $k^{(-2)}K = 0.5$, (c) $k^{(-2)}K = 2$, and (d) $k^{(-2)}K = 10$.

obtained from the requirement to stray from linearity in this depletion regime. In terms of dimensional quantities, (42) is

$$\tilde{t} \ll \frac{\tilde{\alpha}}{\tilde{k}^{(1)}}. \quad (43)$$

Over the timescale $t = O(1/a\varepsilon)$, the levels of S_6 increase and the levels of s_2 and s_5 decrease. From (41), we see that the maximum amount of NADH produced is bounded above by an $O(1)$ quantity. Hence, we see from (34)–(35) that there will be additional depletion dynamics if $s_5 = O(\delta)$. This condition is not guaranteed to occur, and will depend on the values of the kinetic parameters, as well as on the initial conditions of the system. We do not consider this depletion possibility further, as it has no bearing on our main goal of characterizing the primary enzyme. We consider the remaining distinguished limit in the system, that of a weak reverse reaction, in the next section.

3.2 Weak reverse reaction: $k^{(-2)} = O(\varepsilon K)$

3.2.1 Linear-phase regime

The distinguished limit where the reverse reaction is weak occurs when $k^{(-2)} = O(\varepsilon K)$, and it will be helpful to introduce $\rho := k^{(-2)}/(\varepsilon K) = O(1)$ for notational purposes. This scaling results in a distinguished limit in which the nonlinearities of the primary reaction and the reverse auxiliary reaction are important when $t = O(1/\varepsilon)$. This causes the limit to be distinguished for s_4 over a timescale of $t = O(1/\varepsilon)$.

In this case, the leading-order version of (15) for $t = O(1)$ is

$$\frac{ds_1}{dt} = 0, \quad \frac{ds_2}{dt} = 0, \quad \frac{ds_4}{dt} = \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right) - s_4, \quad \frac{ds_5}{dt} = 0, \quad \frac{ds_6}{dt} = s_4, \quad (44)$$

with initial conditions (17). The solution to (44) is

$$s_1 = 1, \quad s_2 = 1, \quad s_4 = V(1 - e^{-t}), \quad s_5 = 1, \quad s_6 = V(t - 1 + e^{-t}), \quad (45)$$

where V is defined in (20). Hence, s_6 initially scales with t^2 before eventually scaling with t . Thus, when $t = O(1)$ for a weak reverse reaction, the observed NADH production is essentially equivalent to that of §2, even though pyruvate is both a substrate of the primary reaction and a product of the auxiliary reaction. This is because there is not enough alpha-L-alanine being created to produce a large enough amount of pyruvate to affect the NADH production, and the reverse reaction is too weak to siphon away the NADH that is produced by a significant amount. In dimensional terms, the observed long-time concentration of NADH in this regime is

$$\tilde{s}_6 \sim \tilde{V}\tilde{t} \quad \text{as } \tilde{t} \rightarrow \infty, \quad (46)$$

essentially the sub-limit of (31) as $\omega \rightarrow 1$, with \tilde{V} and ω defined in (32). Additionally, we are able to obtain a lower bound for this linear regime from (45), namely that $t \gg 1$ for the linear regime to hold. In dimensional terms, this corresponds to

$$\tilde{t} \gg \frac{\tilde{K}_4^{(2)}}{\tilde{k}^{(2)}}. \quad (47)$$

3.2.2 Depletion regime

As t increases further, the substrate depletion and the reversible reaction from NADH to NAD^+ start to affect the problem at leading order. To investigate this, we make the scalings $t = T/\varepsilon$ and $S_6 = s_6/\varepsilon$. Then, the leading-order version of (15) is

$$\frac{ds_1}{dT} = -a \left(\left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right) - s_4 + \rho \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) S_6 \right), \quad (48a)$$

$$\frac{ds_2}{dT} = -b \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right), \quad (48b)$$

$$0 = \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right) - s_4 + \rho \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) S_6, \quad (48c)$$

$$\frac{ds_5}{dT} = -c \left(s_4 - \rho \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) S_6 \right), \quad (48d)$$

$$\frac{dS_6}{dT} = s_4 - \rho \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) S_6. \quad (48e)$$

While (48) may appear complicated at first glance, it reduces readily to allow an implicit analytic solution for the system. To see this, we first use (48c) to reduce (48a) and deduce that

$$s_1 = 1 \quad (49a)$$

at leading order over this timescale. Using (49a) in (48b) yields a separable ODE for s_2 , with implicit solution

$$s_2 - 1 + K_2^{(1)} \log s_2 = -\frac{bT}{K_1^{(1)} + 1}, \quad (49b)$$

which can also be written in terms of the Lambert W function. This tells us that the leading-order substrate concentration will start to decrease before the pyruvate (Figure 9), even if the substrate is saturating *i.e.* $K_2^{(1)} \rightarrow 0$. Finally, forming linear combinations of (48b)–(48e) to yield conserved quantities allows us to deduce that

$$s_4 = \left(\frac{1}{K_1^{(1)} + 1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right) + \rho \left(\frac{1}{K_1^{(-2)} + 1} \right) \left(\frac{1 - s_2}{b} \right), \quad (49c)$$

$$s_5 = 1 - \frac{c}{b} (1 - s_2), \quad (49d)$$

$$S_6 = \frac{1 - s_2}{b}. \quad (49e)$$

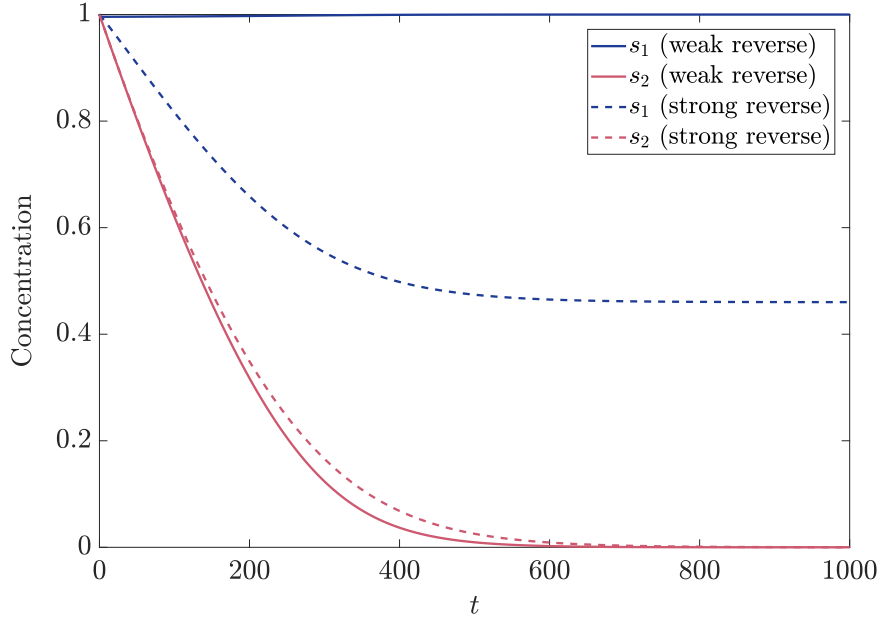


Fig. 9 A comparison of the substrate depletion between the strong and weak reverse reactions, obtained from numerical solutions to the full system (15). We use the parameter values in Table 3, with $\tilde{\alpha} = \tilde{\beta} = 2.5$ mM, $k^{(-2)} = \varepsilon K = 5 \times 10^{-4}$ for the weak reverse reaction and $k^{(-2)} = 1/K = 20$ for the strong reverse reaction.

From our analytic solutions, we are able to infer that the weak reverse reaction only affects the amount of s_4 in the system over this timescale; the other chemicals are independent of the reverse reaction despite it appearing at leading order. This change of scaling affects the asymptotic analysis of the depletion dynamics between the strong and weak reverse reaction cases; care is needed in scaling appropriately if taking $k^{(-2)} \rightarrow 0$ in (37)–(41).

We are able to deduce the long-time dynamics in terms of the implicit solution for s_2 given in (49b). As the experimental calculation for the reaction velocity is taken when the growth of NADH is linear, it is helpful to note when and how this breaks down in this long-time limit. To do this, we expand (49b) for small time to note that

$$s_6 \sim VT \left(1 - \frac{bK_2^{(1)}V}{2(1+K_2^{(1)})} T \right) \quad \text{as } T \rightarrow 0^+. \quad (50)$$

Moving back to dimensional quantities, we find that

$$\tilde{s}_6 \sim \tilde{k}^{(1)} \frac{\tilde{\alpha}}{\tilde{K}_1^{(1)} + \tilde{\alpha}} \frac{\tilde{\beta}}{\tilde{K}_2^{(1)} + \tilde{\beta}} \tilde{t} \left(1 - \frac{\tilde{k}^{(1)}}{2} \frac{\tilde{\alpha}}{\tilde{K}_1^{(1)} + \tilde{\alpha}} \frac{\tilde{K}_2^{(1)}}{(\tilde{K}_2^{(1)} + \tilde{\beta})^2} \tilde{t} \right), \quad (51)$$

for

$$\frac{K_4^{(2)}}{\tilde{k}^{(2)}} \ll \tilde{t} \ll \frac{(\tilde{K}_1^{(1)} + \tilde{\alpha})(\tilde{K}_2^{(1)} + \tilde{\beta})^2}{\alpha \tilde{k}^{(1)} \tilde{K}_2^{(1)}}. \quad (52)$$

The simplest way to make this constraint less limiting is to have as small a value of $\varepsilon = \tilde{k}^{(1)}/\tilde{k}^{(2)}$ as possible. Another way is to increase $\tilde{\beta}$, though we note that this also has the effect of making it difficult to determine $\tilde{K}_2^{(1)}$. Therefore, to minimize the correction term (and hence increase the duration of time at which the NADH growth is linear, thus reducing experimental error), we should decrease $\tilde{\alpha}$ and increase $\tilde{\beta}$. The former of these is perhaps a little surprising, though we note that we must also ensure that pyruvate depletion does not affect anything at leading order, so $\tilde{\alpha} \gg \varepsilon \tilde{K}_4^{(2)} \approx 0.05$ mM must also hold.

There will be further asymptotic regions and depletion dynamics at play here at longer timescales, such as if s_5 becomes of $O(\delta)$. However, these additional regions are of limited relevance to our analysis here.

4 Guide to using our theoretical results

We now summarize the main results from §3.1 and §3.2, and give a step-by-step guide of how to use these to characterize enzymes. We have found that a linear-growth regime for NADH is possible for a strong reverse reaction, but

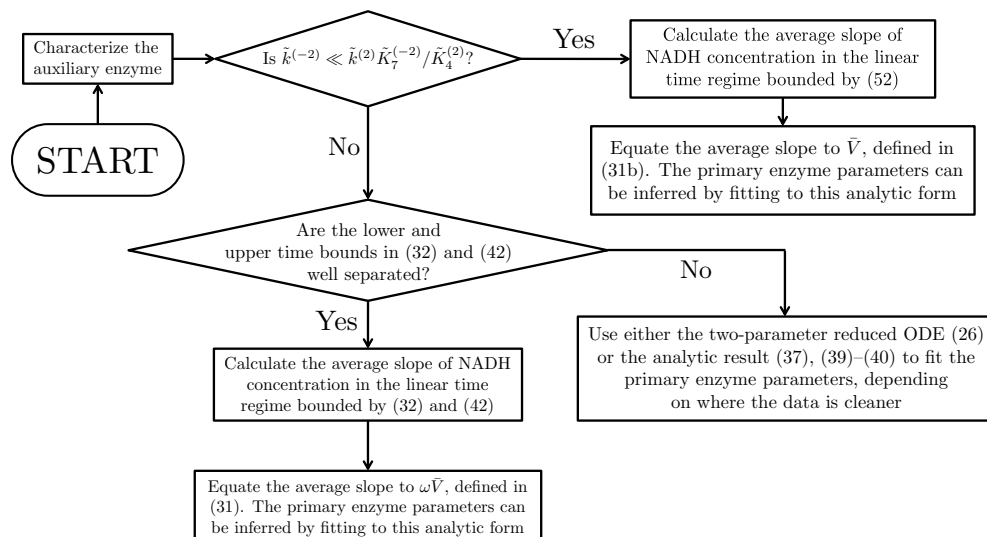


Fig. 10 A flow chart summarising how to use the main results of this paper. Further details of each step are provided in the main text of §4.

this regime can be quite short and is sandwiched between two nonlinear-growth regimes. The NADH concentration in the linear regime is given in (31) in terms of dimensional quantities. Although the linear-growth regime in the weak reverse reaction case is also sandwiched between two nonlinear-growth regimes, it lasts for a much longer duration in this case and so results from the linear regime are easier to use in practice.

Due to our analytic asymptotic results, we are able to deduce upper and lower time constraints for the linear regimes, which we give in (33) and (43) in terms of dimensional quantities for the strong reverse reaction case and in (52) for the weak reverse reaction case. We are also able to obtain an implicit closed-form solution for the NADH concentration away from this linear regime, which would allow the unknown quantities to be inferred through a nonlinear parameter fitting, if the linear regime was too short for accurate measurements to be obtained. We now provide a step-by-step guide to using our results with time data for the concentration of the indicator chemical in each case, using different values of the initial substrate concentrations $\tilde{\alpha}$ and $\tilde{\beta}$. We also present a summary of these guidelines as a flow chart in Figure 10, and note that a summary of the solutions we derive in this paper is provided in Table 4.

Step 1: Characterize the auxiliary enzyme.

Step 2: Determine $\tilde{k}^{(-2)}$, the maximum velocity of the reverse reaction for the auxiliary enzyme. If $\tilde{k}^{(-2)}$ is significantly smaller than $\tilde{k}^{(2)} \tilde{K}_7^{(-2)} / \tilde{K}_4^{(2)}$ (say, two-to-three orders of magnitude), proceed to §4.1. If not, proceed to §4.2.

4.1 Using the weak reverse reaction results

In this case, the NADH concentration should exhibit linear growth for a significant period of time, with the relevant bounds given in (52). We are able to reduce the lower bound in (52) by adding more of the auxiliary enzyme, thereby reducing the lag period. Additionally, although the upper bound in (52) contains the parameters we are trying to infer, we can see that adding less of the primary enzyme will increase it, extending the length of time until the depletion dynamics kick in.

From the experimental data in the linear growth regime, the procedure is to calculate the average slope of NADH concentration. This slope is equal to \tilde{V} , defined in (32). The kinetic parameters of the primary enzyme can be inferred by fitting the average slope obtained from the experimental data to the expression we give for \tilde{V} in (32), using nonlinear data-fitting to minimize the sum of the difference between the data and functional form.

4.2 Using the strong reverse reaction results

In this case, the NADH concentration should exhibit linear growth for some period of time, with the relevant bounds given in (33) and (43). While the upper bound of (43) can be extended by adding less of the primary enzyme, it is not necessarily true in this case that adding more of the auxiliary enzyme will decrease the lower bound of (33). If the linear regime is able to be observed for an appropriate length of time, then the appropriate approach is to calculate the average slope of NADH concentration from the experimental data in the linear growth regime. This

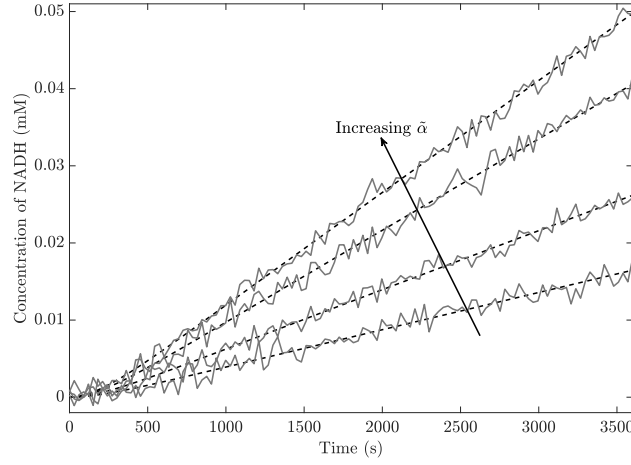


Fig. 11 An example of the ‘true’ time courses generated from the system (10)–(11), with added Gaussian noise of mean $0 \mu\text{M}$ and standard deviation $1 \mu\text{M}$. The data without noise is indicated by the dashed lines. We use the parameter values from Table 2, as well as $\tilde{k}^{(1)} = 0.02 \mu\text{M s}^{-1}$, $\tilde{k}^{(2)} = 10 \mu\text{M s}^{-1}$, $\tilde{k}^{(-2)} = 10 \mu\text{M s}^{-1}$, $\tilde{K}_1^{(1)} = 1.5 \text{ mM}$, $\tilde{K}_2^{(1)} = 1 \text{ mM}$, $\tilde{\beta} = 50 \text{ mM}$, and $\tilde{\gamma} = 50 \text{ mM}$. The four different curves correspond to $\tilde{\alpha} = 0.5, 1, 2.5, 5 \text{ mM}$.

slope is equal to $\tilde{V}\omega$, defined in (31). Here, ω is a function of the kinetic parameters of the auxiliary enzyme and the initial chemical concentrations, whereas \tilde{V} depends on the kinetic parameters of the primary enzyme and the initial chemical concentrations. Hence, the kinetic parameters of the primary enzyme can be inferred by fitting the average slope obtained from the experimental data to the expression we give for $\tilde{V}\omega$ in (32), using nonlinear data-fitting to minimize the distance between the data and functional form.

If the linear regime is *not* able to be observed for an appropriate length of time, there are two options, depending on whether better data is available for $\tilde{t} = O(\tilde{K}_4^{(2)}/\tilde{k}^{(2)})$ or for $\tilde{t} = O(\tilde{\alpha}/\tilde{k}^{(1)})$. In the former case, one should use the results from §3.1, where the data-fitting has been reduced to the single ODE (21). In terms of dimensional quantities, the ODE (21) can be re-written as

$$\frac{d\tilde{s}_6}{d\tilde{t}} = \left(\frac{\tilde{V}\tilde{k}^{(2)}}{\tilde{K}_4^{(2)}} \right) \tilde{t} - \left(\frac{\tilde{k}^{(2)}}{\tilde{K}_4^{(2)}} \right) \tilde{s}_6 - \left(\frac{\tilde{k}^{(-2)}}{\tilde{K}_7^{(-2)}} \frac{\tilde{\alpha}}{\tilde{K}_1^{(-2)} + \tilde{\alpha}} \right) \frac{\tilde{s}_6^2}{\tilde{K}_6^{(-2)} + \tilde{s}_6}, \quad \tilde{s}_6(0) = 0. \quad (53)$$

While there are four different parameter groupings in (53), the kinetic parameters of the primary enzyme are all contained within \tilde{V} , as described in (32). Therefore, the data fitting to this ODE is likely to be significantly simpler than data-fitting to the initial system of ODEs (1). If better data is instead available for $\tilde{t} = O(\tilde{\alpha}/\tilde{k}^{(1)})$, then the results of §3.1.3 should be used. Here, the functional form for the NADH concentration is given by (38), (40), and (41). In this case, the data-fitting should involve minimizing the sum of the difference between the time data and the functional form.

4.3 Comparison with a naive nonlinear solver

We now compare the asymptotic method we describe in this paper with a naive nonlinear parameter fitting. To do this, we first generate ‘true’ time courses for NADH from the system (10)–(11) for four different values of $\tilde{\alpha}$, then add Gaussian noise with mean $0 \mu\text{M}$ and standard deviation $\sigma \mu\text{M}$, and we use σ as a control parameter. An example of these time courses with $\sigma = 1$ is given in Figure 11. The asymptotic parameter fitting is carried out by fitting the asymptotic solution (31) to the true time courses using the MATLAB function ‘lsqcurvefit’ for the parameters $\tilde{k}^{(1)}$ and $\tilde{K}_1^{(1)}$. The naive nonlinear parameter fitting is carried out using the same MATLAB function ‘lsqcurvefit’ on the ODE system (10)–(11), fitting the parameter values $\tilde{k}^{(1)}$, $\tilde{K}_1^{(1)}$, and $\tilde{K}_2^{(1)}$. However, as we generate the true time courses using a large value of $\tilde{\beta}$ in order to focus on $\tilde{K}_1^{(1)}$ rather than $\tilde{K}_2^{(1)}$, we ignore the value of $\tilde{K}_2^{(1)}$ generated from this.

In general, the naive nonlinear fit is worse than the asymptotic fit in determining $\tilde{k}^{(1)}$ and better than the asymptotic fit in determining $\tilde{K}_1^{(1)}$ (Figure 12). However, the asymptotic fit appears to be more consistent in its predictions than the naive nonlinear fit, as the naive nonlinear fit can predict wildly incorrect values of $\tilde{k}^{(1)}$ even for small noise. Moreover, we note that the computational time required for the naive nonlinear fit is around 1000 times larger than for the asymptotic method.

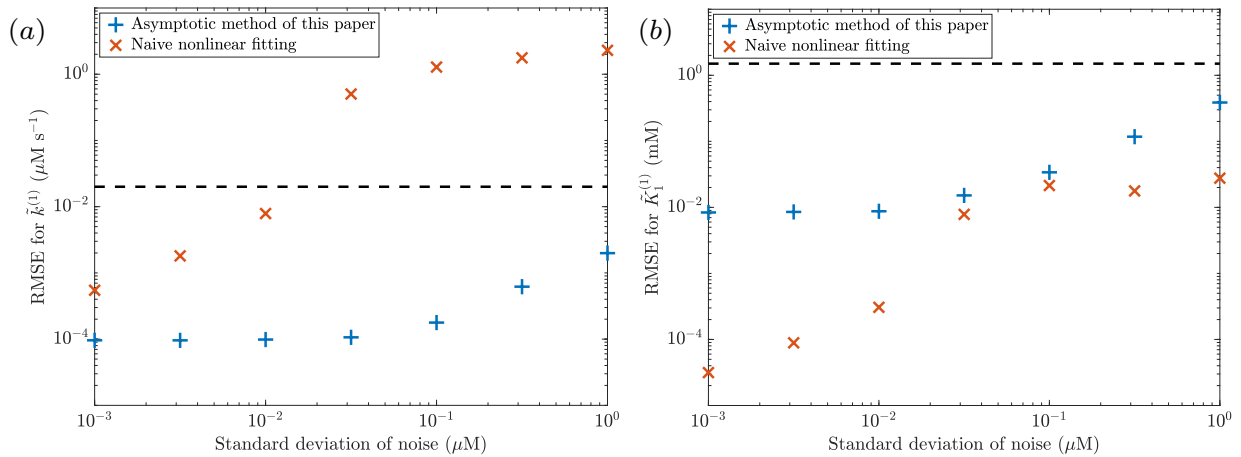


Fig. 12 A comparison between the asymptotic method presented in this paper, and a naive nonlinear fit of the system (10)–(11). We show the root mean squared error (RMSE) for both methods in estimating the parameters (a) $\tilde{k}^{(1)}$ and (b) $\tilde{K}_1^{(1)}$. The dashed line shows the value of the parameter being estimated, so a RMSE well below this line signifies good accuracy whereas a RMSE near or above the line signifies poor accuracy. We generate the ‘true’ data from the system (10)–(11), then add Gaussian noise with mean $0 \mu\text{M}$ and standard deviation indicated on the x -axis. Each marker represents the RMSE calculated from 100 total realisations. We use the parameter values from Table 2, as well as $\tilde{k}^{(1)} = 0.02 \mu\text{M s}^{-1}$, $\tilde{k}^{(2)} = 10 \mu\text{M s}^{-1}$, $\tilde{k}^{(-2)} = 10 \mu\text{M s}^{-1}$, $\tilde{K}_1^{(1)} = 1.5 \text{ mM}$, $\tilde{K}_2^{(1)} = 1 \text{ mM}$, $\tilde{\beta} = 50 \text{ mM}$, and $\tilde{\gamma} = 50 \text{ mM}$. We also use $\tilde{\alpha} = 0.5, 1, 2.5, 5 \text{ mM}$ to generate four different ‘true’ time course curves for each different standard deviation of noise.

5 Application to experimental data

The above concludes our mathematical analysis of the problem. To showcase how these results can be used in practice, we performed enzyme assays to characterize an enzyme for two different substrates using a previously uncharacterized putative omega-aminotransferase, cloned and purified in a heterologous *Escherichia coli* host. The aminotransferase is CnAptA (UniProtKB Q0KEZ8), named here for the first time in the literature and identified in the genome of *Cupriavidus necator* H16, an industrially relevant, facultative chemolithoautotrophic chassis microorganism of the Synthetic Biology Research Centre (SBRC-Nottingham). We describe the experimental protocol used to purify this enzyme in Appendix B.

5.1 Enzyme assay protocol

For these assays, the corresponding reaction network is shown in Figure 1. The primary enzyme is CnAptA, which converts pyruvate and an additional omega-amino acid substrate into a corresponding aldehyde and alpha-L-alanine. Our assay solution also included beta-nicotinamide adenine dinucleotide sodium salt (NAD^+) as an electron acceptor for the auxiliary enzyme reaction, and we monitored the formation of NADH over time, produced from the auxiliary reaction. We used a commercially available alanine dehydrogenase (A-DH) from *Bacillus cereus* (alanine dehydrogenase, recombinant; CAS-No 9029-06-5, Sigma-Aldrich Company Ltd.) as the auxiliary enzyme, and we performed the experiment at pH 10, the pH condition recommended by the supplier of this auxiliary enzyme, this condition resulting in an undetectable reverse reaction. Our analysis suggests that the primary enzyme will be easier to characterize, since the bounds on the validity of the linear regime are less severe. Hence, we can either use the sub-limit as $\omega \rightarrow 1$ of the results in the distinguished limit (31), or we can use the sub-limit directly deduced in §3.2. The approach for this is summarised in §4 and §4.1. Performing these assays at different acidities, which may be more relevant for certain bacterial cell environments, would result in a strong reverse reaction, which could be dealt with using the analytic representation for the distinguished limit given in (31). The approach for these cases is summarised in §4 and §4.2.

We performed the enzyme assays in 96 well microtiter flat bottom transparent plates (Costar[®] Sigma-Aldrich). The reaction mixture contained different concentrations of the substrate, 2.5 mM NAD^+ , 5 mM pyruvate , $5 \mu\text{l A-DH}$ [0.28 mg/ml], $5 \mu\text{M}$ of pyridoxal 5'-phosphate (PLP), and $5 \mu\text{l}$ of CnAptA [0.0245 mg/ml]. We then used an appropriate amount of sodium carbonate buffer to obtain a final volume of $200 \mu\text{l}$. The two different substrates we used were 3-aminobutanoic acid (Sigma-Aldrich 97%, CAS No 541-46-6) and 5-aminovalerate (ACROS Organics, 97%, CAS No 660-88-8).

To detect the NADH present in the solution, we measured the light absorbance at 340 nm using a TECAN Infinite[®] M1000 PRO plate reader, at $30 \text{ }^\circ\text{C}$, the absorbance being linearly proportional to the concentration of NADH in the solution. We monitored the light absorbance at 340 nm at 60 second intervals for one hour. After subtracting off the base level of absorbance at the start of each experiment and the absorbance increase monitored in the control experiment with no substrate, we converted these clean absorbance data into an NADH concentration, shown for

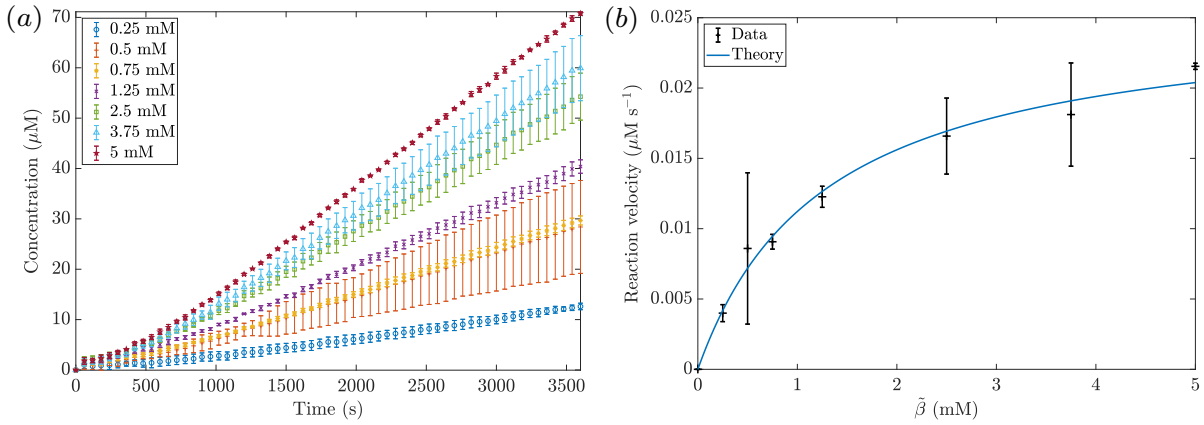


Fig. 13 (a) The time course data for NADH production with the enzyme CnAptA and substrate 3-aminobutanoic acid, using different concentrations of the latter. (b) The average reaction velocity at different substrate concentrations, and the predicted reaction velocity using our nonlinear data fitting to obtain values of $\tilde{k}^{(1)} = 0.026 \mu\text{M s}^{-1}$ and $\tilde{K}_2^{(1)} = 1.3 \text{ mM}$, each to two significant figures. The error bars on the figures correspond to one standard deviation of the data.

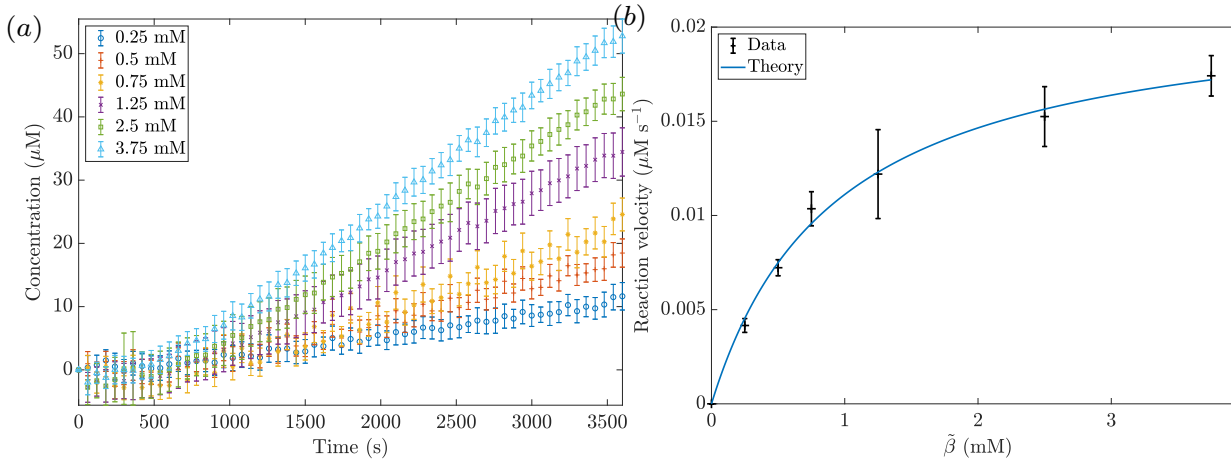


Fig. 14 (a) The time course data for NADH production with the enzyme CnAptA and substrate 5-aminovalerate, using different concentrations of the latter. (b) The average reaction velocity at different substrate concentrations, and the predicted reaction velocity using our nonlinear data fitting to obtain values of $\tilde{k}^{(1)} = 0.021 \mu\text{M s}^{-1}$ and $\tilde{K}_2^{(1)} = 0.93 \text{ mM}$. The error bars on the figures correspond to one standard deviation of the data.

each substrate in Figures 13a and 14a. We expect our TECAN data to be accurate to within a few μM , and this error margin is why there sometimes appears to be ‘negative’ concentrations for early time. As our analysis predicts, there is a clear linear regime after a lag period. We obtained a constant reaction velocity from this linear regime, using the average slope between 1800 and 3600 seconds. From this constant reaction velocity, we used a nonlinear data fit to the functional form of the solution (31)–(32) with $\omega \rightarrow 1$ (as discussed in §4 and §4.1) by minimizing the sum of the difference between the data and functional form squared, using the in-built function `fminsearch` in MATLAB, a simplex search method for finding minima [13]. From Figures 13b and 14b, we see that our nonlinear fits generally provide accurate representations of the data. For the enzyme CnAptA, our nonlinear fit predicts values of $\tilde{k}^{(1)} = 0.026 \mu\text{M s}^{-1}$ and $\tilde{K}_2^{(1)} = 1.3 \text{ mM}$ for the substrate 3-aminobutanoic acid and values of $\tilde{k}^{(1)} = 0.021 \mu\text{M s}^{-1}$ and $\tilde{K}_2^{(1)} = 0.93 \text{ mM}$ for the substrate 5-aminovalerate, assuming that the levels of pyruvate we use are saturating. In terms of the catalytic efficiencies, we see that $\tilde{k}^{(1)}/\tilde{K}_2^{(1)} = 2.0 \times 10^{-5} \text{ s}^{-1}$ for 3-aminobutanoic acid and $\tilde{k}^{(1)}/\tilde{K}_2^{(1)} = 2.3 \times 10^{-5} \text{ s}^{-1}$ for 5-aminovalerate. Therefore, we note that while 5-aminovalerate has a higher affinity (a lower $\tilde{K}_2^{(1)}$) to and catalytic efficiency for CnAptA than 3-aminobutanoic acid, the maximum reaction velocity is higher for 3-aminobutanoic acid than for 5-aminovalerate. However, these differences are fairly small - there is not a large difference in enzyme behaviour between the two substrates.

6 Discussion

We have used singular perturbation theory to determine how to infer the kinetic parameters of the primary enzyme from the measured NADH concentration in a coupled enzyme assay, with reaction network shown in Figure 1. In particular, this will allow putative aminotransferases to be quickly characterized using coupled enzyme assays. Our

analysis also allows us to obtain estimates for when we expect to see a linear growth in NADH concentration, the typical regime measured in enzyme assays since it allows a quick sanity check of experimental results and is much easier to fit to data. We have outlined how to use our theoretical results to infer the kinetic parameters of the primary enzyme from experimental data in §4, where we also compared our method with a naive nonlinear fit. Moreover, we have used our results to characterize a novel putative aminotransferase for two different substrates in §5.

Our analysis shows that there are two distinguished asymptotic limits in the system: the important one for NADH (the measurable chemical) being the strong reverse reaction case, considered in §3.1; the other is the weak reverse reaction case, considered in §3.2. We demonstrate that all the information that can be inferred from the NADH concentration in the weak reverse reaction case is a sub-limit of the strong reaction case, even though there are differences in the concentrations of alpha-L-alanine, an intermediate chemical, between these two cases. In our analysis, we exploit the smallness of three dimensionless parameters: ε , δ , and K , with a focus on analysing how the asymptotic size of $k^{(-2)}$ affects the system behaviour and treating all other parameters in the system as $O(1)$ to retain their generality. As the parameters ε and δ can always be made to be small through experimental choice, and we also present results for $K = O(1)$ in Appendix C (which shows a slight change in the depletion regimes), our work represents a comprehensive asymptotic analysis of the possible behaviours of the non-standard coupled enzyme assay we consider in this paper.

Additionally, we note that our results suggest that one cannot deduce whether the strong reaction case pertains simply by looking for a change in reaction velocity when the amount of auxiliary enzyme is varied. This can be seen by (31), where we see that the only two variables that involve the experimentally controlled auxiliary enzyme concentration, $\tilde{k}^{(2)}$ and $\tilde{k}^{(-2)}$, appear as a ratio, thus removing any dependence on this enzyme concentration. Thus, it is not always immediately clear from observations whether one is in the weak or strong reverse reaction regime, unless the auxiliary reaction has previously been completely characterized. Due to the weak reverse reaction being a sub-limit of the strong for the experimentally measurable NADH concentration, as discussed above, we advise working with the results derived in the strong reverse reaction regime in §3.1.

In the derivation of our model, we assume that the kinetic parameters of the putative enzyme are unchanging over time. This may not be valid for enzymes that are unstable or suffer a loss of activity over time, for example through incubation in a nonoptimal buffer at certain temperatures. Accounting for this change in enzyme activity would require additional information about the type and effect of the degradation occurring.

Although it would be possible to perform a brute force fitting of the kinetic parameters of the primary enzyme using the governing equations (10), such an approach would be time-consuming and would not provide the general analytic expressions we derive through our asymptotic analysis. Moreover, we are able to use our analytic expressions to estimate when we will be in the linear growth regime, and for how long we expect this regime to last, in terms of the experimental parameters. This provides a useful sanity check when it comes to interpreting the data.

Historically, the measured reaction velocity is often converted into kinetic parameters for the enzyme using a Lineweaver–Burke plot [14]. However, this procedure is generally not the best way to calculate these values as it amplifies noise in the measurements [9]. Nonlinear data fitting is orders of magnitude better than it was at the time of Lineweaver and Burke, so it is much more accurate nowadays to fit the data to analytic forms such as those we have deduced using singular perturbation theory. For the case we consider in this paper, measurements with varying initial substrate concentrations give us a two-dimensional array of data from which to fit three variables. Additionally, choosing one of these substrate concentrations to be much larger than the estimated Michaelis constant $\tilde{K}_1^{(1)}$ or $\tilde{K}_2^{(1)}$ reduces the problem to a one-dimensional array of data from which to fit two variables.

While we have considered a specific type of non-standard coupled enzyme assay here, the asymptotic techniques we use are far from restricted to this system. We hope that the techniques used in this paper will be applied to different reaction networks, to help characterize enzymes that can only be assayed in non-standard reaction network topologies.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council [grant number BB/L013940/1]; and the Engineering and Physical Sciences Research Council, jointly funding the first grant number.

A Distinguished asymptotic limits of equation (26)

In this Appendix, we investigate the distinguished asymptotic limits of (26), thus allowing us to characterize the possible behaviours of its solution.

A.1 Case I

The first distinguished limit we discuss (Region I in Figure 7) occurs when $\bar{V} \ll 1$ and $\gamma = O(1)$ for the ODE (26). Here, there are two different timescales in the problem: $t = O(1)$ and $t = O(1/\bar{V})$, for which $Y = O(\bar{V})$ and $Y = O(1)$, respectively; the limit is distinguished only over the latter.

Therefore, our brief analysis of the earlier timescale consists simply of the asymptotic solution

$$Y \sim \bar{V} (t - 1 + e^{-t}). \quad (54)$$

For the later timescale, the leading-order governing equation is

$$0 = \bar{V}t - Y - \gamma \frac{Y^2}{1+Y} \quad (55)$$

and hence we are in the full quasi-steady case. The equation (55) has solution

$$Y = \frac{\bar{V}t - 1 + \sqrt{(\bar{V}t + 1)^2 + 4\bar{V}\gamma t}}{2(\gamma + 1)}, \quad (56)$$

which matches appropriately to the earlier-time solution. We note that the large-time limit of (56) is $Y \sim \bar{V}t/(\gamma + 1)$. This is different to the large-time limit of (54) (when $t = O(1)$), which is $Y \sim \bar{V}t$. Therefore, there are two different linear regimes here, and one should take great care in this distinguished limit to ensure that one is aware of in which linear regime the measurements are being taken: failure to do so could result in a significant error of interpretation.

When $\gamma \rightarrow \infty$ in this distinguished limit, we move into sub-limit A (Figure 7). Here, the earliest timescale is the same as (54), but the later timescale splits into two, the earlier with $t = O(1/(\bar{V}\gamma))$ and $Y = O(1/\gamma)$, the later with $t = O(\gamma/\bar{V})$ and $Y = O(1)$, as could be inferred from the large- γ limit of (56). On the timescale $t = O(1/(\bar{V}\gamma))$, the solution is

$$Y = \frac{\sqrt{1 + 4\bar{V}\gamma t} - 1}{2\gamma}, \quad (57)$$

and on the timescale $t = O(\gamma/\bar{V})$

$$Y = \frac{\bar{V}t + \sqrt{(\bar{V}t)^2 + 4\bar{V}\gamma t}}{2\gamma}. \quad (58)$$

When $\gamma \rightarrow 0$ in this distinguished limit, we move into sub-limit D (Figure 7). Here, the full leading-order solution in the later timescale is simply the large-time limit of the earlier timescale. Hence, the leading-order solution is simply (54), uniformly valid for all time.

A.2 Case II

The next distinguished limit (Region II in Figure 7) occurs when $\bar{V} \ll 1$ and $\gamma \gg 1$, with $\bar{V}\gamma = O(1)$, in the ODE (26). Here, there are two different timescales in the problem: $t = O(1)$ and $t = O(\gamma^2)$, for which $Y = O(1/\gamma)$ and $Y = O(1)$, respectively; the limit is distinguished only over the earlier of these. Taking $Y = \bar{Y}/\gamma$, the leading-order equation for the earlier timescale is

$$\frac{d\bar{Y}}{dt} = \bar{V}\gamma t - \bar{Y} - \bar{Y}^2, \quad \bar{Y}(0) = 0. \quad (59)$$

The Riccati equation (59) has solution

$$\bar{Y} = (\bar{V}\gamma)^{1/3} \frac{\text{Bi}'[T(t)] + \alpha \text{Ai}'[T(t)]}{\text{Bi}[T(t)] + \alpha \text{Ai}[T(t)]} - \frac{1}{2}, \quad (60a)$$

$$\alpha = \frac{2(\bar{V}\gamma)^{1/3} \text{Bi}'[T(0)] - \text{Bi}[T(0)]}{\text{Ai}[T(0)] - 2(\bar{V}\gamma)^{1/3} \text{Ai}'[T(0)]}, \quad (60b)$$

$$T(t) = \frac{1}{4(\bar{V}\gamma)^{2/3}} + (\bar{V}\gamma)^{1/3} t, \quad (60c)$$

where $\text{Ai}(z)$ and $\text{Bi}(z)$ are the standard Airy functions. The solution (60a) exhibits the following large- t behaviour

$$\bar{Y} \sim \sqrt{\bar{V}\gamma t} \quad \text{as } t \rightarrow \infty. \quad (61)$$

The later timescale in this problem occurs when $t = O(\gamma^2) = O(1/\bar{V}^2)$ and $Y = O(1)$. From (26), this timescale yields the following leading-order equation

$$\bar{V}t = \gamma \frac{Y^2}{1+Y}, \quad (62)$$

with solution

$$Y = \frac{\bar{V}t + \sqrt{(\bar{V}t)^2 + 4\bar{V}\gamma t}}{2\gamma}. \quad (63)$$

This exhibits the large- t behaviour

$$Y \sim \bar{V}t/\gamma \quad \text{as } \bar{t} \rightarrow \infty. \quad (64)$$

Hence, we may deduce that in this regime the variable Y initially behaves as $Y \sim \bar{V}t^2/2$, then as $Y \sim \sqrt{\bar{V}t/\gamma}$, then finally as $Y \sim \bar{V}t/\gamma$, and therefore exhibits at least two points of inflection, as well as three distinct power laws.

The solutions of the distinguished limit in Case II tend to those of the sub-limit A (Figure 7) as $\bar{V}\gamma \rightarrow 0$. The early timescale, when $t = O(1)$ with solution (60a), splits into two, with solutions (54) and (57). The late timescale solution (63) is the same as in sub-limit A, as given in (58).

As $\bar{V}\gamma \rightarrow \infty$, we move into sub-limit B (Figure 7). There are two different timescales in this sub-limit: $t = O(1/(\bar{V}\gamma)^{1/3})$ and $t = O(\gamma/\bar{V})$, for which $Y = O((\bar{V}/\gamma^2)^{1/3})$ and $Y = O(1)$, respectively. For the first, $t = O(1/(\bar{V}\gamma)^{1/3})$, the solution (60a) reduces to

$$Y = \left(\frac{\bar{V}}{\gamma^2}\right)^{1/3} \frac{\text{Bi}'\left[(\gamma\bar{V})^{1/3}t\right] + \sqrt{3}\text{Ai}'\left[(\gamma\bar{V})^{1/3}t\right]}{\text{Bi}\left[(\gamma\bar{V})^{1/3}t\right] + \sqrt{3}\text{Ai}\left[(\gamma\bar{V})^{1/3}t\right]}, \quad (65)$$

For the second, $t = O(\gamma/\bar{V})$, the leading-order solution is the same as for Case II, given by (63).

A.3 Case III

The next distinguished limit (Region III in Figure 7) occurs when $\bar{V} \gg 1$ and $\gamma \gg 1$, with $\bar{V}/\gamma^2 = O(1)$, in the ODE (26). Here, the important timescale has $t = O(1/\gamma) = O(1/\bar{V}^{1/2})$, for which $Y = O(1)$. To investigate this, we make the substitution $t = \tau/\gamma$, resulting in the leading-order behaviour being described by

$$\frac{dY}{d\tau} = \frac{\bar{V}}{\gamma^2} \tau - \frac{Y^2}{1+Y}, \quad Y(0) = 0. \quad (66)$$

While we cannot solve (66) analytically, we note that the large- τ behaviour of (66) is $Y \sim \bar{V}\tau/\gamma^2 = \bar{V}t/\gamma$, and that this behaviour will dominate when $t = O(1)$. Moreover, we can show that this case reduces to the sub-limits B and C (Figure 7) as $\bar{V}/\gamma^2 \rightarrow 0$ and $\bar{V}/\gamma^2 \rightarrow \infty$, respectively.

In the former sub-limit, with $\bar{V}/\gamma^2 \rightarrow 0$, there are two timescales. For the earlier timescale scaling $Y \sim (\bar{V}/\gamma^2)^{1/3}$ and $\tau \sim (\gamma^2/\bar{V})^{1/3}$ (equivalent to $t \sim 1/(\gamma\bar{V})^{1/3}$) results in the reduction of (66) to a Riccati equation with solution (65). The problem for the later timescale can be obtained by scaling $\tau \sim \gamma^2/\bar{V}$ (equivalent to $t \sim \gamma/\bar{V}$), and this has solution (63).

For the latter sub-limit, with $\bar{V}/\gamma^2 \rightarrow \infty$, the important balance occurs when $\tau = O(1)$ (equivalent to $t \sim 1/\gamma$) and $Y = O(\bar{V}/\gamma^2)$, resulting in the leading-order system

$$\frac{dY}{d\tau} = \frac{\bar{V}}{\gamma^2} \tau - Y, \quad Y(0) = 0, \quad (67)$$

which has solution

$$Y = \frac{\bar{V}}{\gamma^2} (\tau - 1 + e^{-\tau}) = \frac{\bar{V}}{\gamma^2} (\gamma t - 1 + e^{-\gamma t}). \quad (68)$$

A.4 Case IV

The final distinguished limit (Region IV in Figure 7) occurs when $\bar{V} \gg 1$ and $\gamma = O(1)$ in the ODE (26). Here, the main balance occurs when $t = O(1)$ and $Y = O(\bar{V})$. In this case, the leading-order version of the ODE (26) is

$$\frac{dY}{dt} = \bar{V}t - Y - \gamma Y, \quad Y(0) = 0, \quad (69)$$

which has solution

$$Y = \frac{\bar{V}}{(1+\gamma)^2} \left[(1+\gamma)t - 1 + e^{-(1+\gamma)t} \right]. \quad (70)$$

In the limit of $\gamma \rightarrow \infty$, the solution (70) reduces to that in sub-limit C (Figure 7), given by (68), with the important timescale $t = O(1/\gamma)$. In the limit of $\gamma \rightarrow 0$, the solution (70) reduces to that of sub-limit D (Figure 7), given by (54). This completes our comprehensive asymptotic analysis of the distinguished limits of the ODE (26).

B Experimental protocol

B.1 Bacterial strains and cultivation conditions

We purchased the bacterial strains used in this study from New England Biolabs (*E. coli* ER2566) and DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH DSMZ (*C. necator* H16). Lysogeny broth (LB) medium was used for cultivation of both strains. We cultivated strains carrying the pTXB1 plasmid (supplied with the IMPACTTM KIT (NEB#E6901)) in the same media supplemented with ampicillin, at a final concentration of 100 $\mu\text{g/ml}$.

B.2 Cloning of the putative aminotransferase CnAptA

We amplified the coding region of CnAptA by PCR from the genomic DNA of *C. necator* H16 using 0.5 μM of FwaptA (5' taccatattggacgcccccaag-gacctg 3') and RvaptA (5' tgcaggaagagcccttgatgctcgcagcgcctt 3') primer pairs, using the Q5[®] High-fidelity 2X Master Mix (NEB, #M0492S) with the following PCR conditions: 3 minutes of denaturation at 98 °C, followed by 25 cycles of 10 seconds denaturation at 98 °C, 30 seconds of an annealing temperature at 72 °C, and 1 minute extension of at 72 °C. We also added 5% DMSO to the PCR mixture due to the high GC content of the amplicon.

We digested the resulting PCR products with the *NdeI* and *SapI* restriction enzymes and ligated into the pTXB1 expression vector supplied with the IMPACTTM KIT, which we digested with the same restriction enzymes. We introduced the resulting plasmid, pTXB1-CnAptA, into the *E. coli* ER2566 strain. The pTXB1 vector allows for the fusion of the target protein at its C-terminus to the affinity tag (chitin binding domain or CBD) via a protein splicing element. This system is suitable for the purification of the native recombinant protein in a single chromatographic step, without the use of a protease. The expression of the fusion gene is controlled by an isopropyl-beta-D-1-thiogalactopyranoside (IPTG)-inducible T7 promoter.

To overexpress the putative transaminase CnAptA, we grew the recombinant *E. coli* ER2566 strains carrying the plasmid pTXB1-CnAptA on LB medium in the presence of ampicillin (100 $\mu\text{g/ml}$), at 37 °C until the optical density at 600 nm (OD600) for the culture reached 0.5–0.6, at which point 0.4 mM IPTG was added. We then incubated this culture at 22 °C for another 10 hours and harvested it by centrifugation at 2800 $\times g$, at 4 °C.

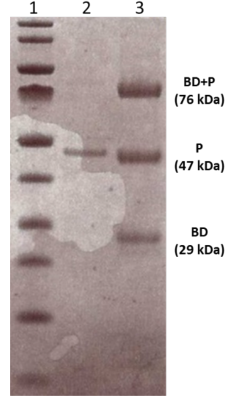


Fig. 15 The ColorPlus™ pre-stained protein ladder (10–230 kDa, Lane 1), the concentrated CnAptA protein after on column cleavage (Lane 2, 47 kDa) and 10 μ L of the resin removed after washing and cleavage of the target protein (Lane 3) was loaded onto a NuPAGE® Novex 4–12% Bis-Tris SDS-PAGE Gel and run for 2 hours at 150V in NuPAGE® MES SDS Running Buffer. Protein bands were visualised using Coomassie Blue staining. BD+P: intein tagged CnAptA, P:CnAptA after intein tag cleavage, BD: cleaved intein tag.

B.3 Purification of chitin-tagged CnAptA transaminase

We re-suspended the cell pellets from the previous centrifugation step in Tris-HCl buffer (20mM Tris-HCl, 0.5 M NaCl [pH 7.4]) and disrupted them using a Soniprep 150 MSE SANYO sonicator, at 4 °C, at 30 second intervals for 30 minutes. We removed the insoluble fraction by centrifugation at $8400 \times g$ for 30 minutes. Then, we loaded the supernatant onto the chromatography column packed with the chitin resin (Poly-Prep® prepacked columns, AG® 1-X8, chloride form #731-1550), equilibrated with the same buffer at pH 8.4, according to the suppliers protocol. We induced the on-column cleavage of the CnAptA purified enzyme by adding Tris-HCl buffer containing 1,4-Dithiothreitol (DTT) (20 mM Tris-HCl, 0.5 M NaCl [pH 8.4]), 50 mM DTT). We then performed an elution of the un-tagged enzyme by adding 12 ml of Tris-HCl buffer (20 mM Tris-HCl, 0.5 M NaCl [pH 8.4]). Finally, we concentrated the eluted protein fraction and exchanged the elution buffer to a buffer containing 20 mM Tris-HCl, 0.05 M NaCl pH 8.4, using an Aminco Ultra-15 centrifugal filter unit (molecular mass cutoff, 50 kDa; Millipore). Hence, the target protein (CnAptA) was eluted from the column in a pure form (Figure 15).

C Case in which $K = O(1)$

In this Appendix, we provide the governing system in the depletion regime for the case with a strong reverse reaction when $K = O(1)$, as the $t = O(1)$ case proceeds in the same manner as when K is small. Here, substrate depletion starts to affect the system at leading order when t becomes of $O(1/\epsilon)$. To investigate this, we introduce $T = \epsilon t$, and we must also make the scalings $S_4 = s_4/\epsilon$, and $S_6 = s_6/\epsilon$. These scalings result in the system

$$\epsilon \frac{ds_1}{dT} = -a(\epsilon v_1 - V_2 + V_{-2}), \quad (71a)$$

$$\frac{ds_2}{dT} = -bv_1, \quad (71b)$$

$$\epsilon \frac{dS_4}{dT} = \epsilon v_1 - V_2 + V_{-2}, \quad (71c)$$

$$\epsilon \frac{ds_5}{dT} = -c(V_2 - V_{-2}), \quad (71d)$$

$$\epsilon \frac{dS_6}{dT} = V_2 - V_{-2}, \quad (71e)$$

where the three reaction velocities are now defined as

$$v_1 = \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right), \quad (72a)$$

$$V_2 = \left(\frac{S_4}{1 + S_4} \right) \left(\frac{s_5}{\delta + s_5} \right), \quad (72b)$$

$$V_{-2} = k^{(-2)} K \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) \left(\frac{S_6}{\epsilon K_6^{(-2)} + S_6} \right) \left(\frac{S_6}{1 + K S_6} \right). \quad (72c)$$

The matching conditions from the earlier timescale yield the following ‘initial conditions’

$$s_1(0) = 1, \quad s_2(0) = 1, \quad S_4(0) = 0, \quad s_5(0) = 1, \quad S_6(0) = 0. \quad (73)$$

Naively taking the limit as $\epsilon \rightarrow 0$ in (34) to obtain a leading-order system would yield a duplication of information. To avoid this, we must form appropriate linear combinations of the governing equations in order to obtain enough information for a leading-order system. Using conserved quantities where possible, we obtain the following differential-algebraic system at leading-order

$$\frac{ds_2}{dT} = -b \left(\frac{s_1}{K_1^{(1)} + s_1} \right) \left(\frac{s_2}{K_2^{(1)} + s_2} \right), \quad (74a)$$

$$\frac{S_4}{1+S_4} = k^{(-2)} K \left(\frac{s_1}{K_1^{(-2)} + s_1} \right) \left(\frac{S_6}{1+KS_6} \right), \quad (74b)$$

$$s_1 + aS_4 = 1, \quad (74c)$$

$$s_2 + bS_4 + bS_6 = 1, \quad (74d)$$

$$s_5 + cS_6 = 1. \quad (74e)$$

It is still possible to follow the analysis of §3.1 and rearrange this system into a single ODE, though this is too cumbersome to be particularly enlightening.

References

1. R A Alberty. The relationship between Michaelis constants, maximum velocities and the equilibrium constant for an enzyme-catalyzed reaction. *J Amer Chem Soc*, 75(8):1928–1932, 1953.
2. C M Bender and S A Orszag. *Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory*. Springer Science & Business Media, 2013. doi: 10.1007/978-1-4757-3069-2.
3. H Bisswanger. Enzyme assays. *Persp Sci*, 1(1-6):41–55, 2014. doi: 10.1016/j.pisc.2014.02.005.
4. E K Chowdhury, T Saitoh, S Nagata, M Ashiuchi, and H Misono. Alanine dehydrogenase from *Enterobacter aerogenes*: purification, characterization, and primary structure. *Biosci Biotech Biochem*, 62(12):2357–2363, 1998.
5. M P Dalwadi, M Garavaglia, J P Webb, J R King, and N P Minton. Applying asymptotic methods to synthetic biology: Modelling the reaction kinetics of the mevalonate pathway. *J Theor Biol*, 439:39–49, 2018. doi: 10.1016/j.jtbi.2017.11.022.
6. M P Dalwadi, J R King, and N P Minton. Multi-timescale analysis of a metabolic network in synthetic biology: a kinetic model for 3-hydroxypropionic acid production via beta-alanine. *J Math Biol*, 77(1):165–199, 2018. doi: 10.1007/s00285-017-1189-3.
7. J Eilertsen and S Schnell. A kinetic analysis of coupled (or auxiliary) enzyme reactions. *Bull Math Biol*, 80(12):3154–3183, 2018. doi: 10.1007/s11538-018-0513-4.
8. J Eilertsen, W Stroberg, and S Schnell. A theory of reactant-stationary kinetics for a mechanism of zymogen activation. *Biophys Chem*, 242:34–44, 2018. doi: 10.1016/j.bpc.2018.08.003.
9. D Fell. *Understanding the control of metabolism*. Portland Press, 1997. doi: 10.1007/978-1-4757-3069-2.
10. B Hutter and M Singh. Properties of the 40 kDa antigen of *Mycobacterium tuberculosis*, a functional L-alanine dehydrogenase. *Biochem J*, 343(3):669–672, 1999.
11. J Kevorkian and J D Cole. *Perturbation methods in applied mathematics*, volume 34. Springer Science & Business Media, 2013. doi: 10.1007/978-1-4757-4213-8.
12. A Kumar and K Josić. Reduced models of networks of coupled enzymatic reactions. *J Theor Biol*, 278(1):87–106, 2011. doi: 10.1016/j.jtbi.2011.02.025.
13. J C Lagarias, J A Reeds, M H Wright, and P E Wright. Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM J Optim*, 9(1):112–147, 1998.
14. H Lineweaver and D Burk. The determination of enzyme dissociation constants. *J Amer Chem Soc*, 56(3):658–666, 1934.
15. W R McClure. Kinetic analysis of coupled enzyme assays. *Biochem*, 8(7):2782–2786, 1969. doi: 10.1021/bi00835a014.
16. P K Mehta, T I Hale, and P Christen. Aminotransferases: demonstration of homology and division into evolutionary subgroups. *Eur J Biochem*, 214(2):549–561, 1993.
17. L Menten and M I Michaelis. Die kinetik der invertinwirkung. *Biochem Z*, 49:333–369, 1913.
18. J D Murray. *Mathematical Biology I: An Introduction*, chapter 6, pages 175–217. Springer-Verlag New York, 2002. doi: 10.1007/b98868.
19. F B Rudolph, B W Baugher, and R S Beissner. Techniques in coupled enzyme assays. In *Methods in enzymology*, volume 63, pages 22–42. Elsevier, 1979. doi: 10.1016/0076-6879(79)63004-5.
20. Y Sawa, M Tani, K Murata, H Shibata, and H Ochiai. Purification and characterization of alanine dehydrogenase from a cyanobacterium, *Phormidium lapideum*. *J Biochem*, 116(5):995–1000, 1994.
21. C Sayer, M N Isupov, A Westlake, and J A Littlechild. Structural studies of *Pseudomonas* and *Chromobacterium* ω -aminotransferases provide insights into their differing substrate specificity. *Acta Crystal D: Biol Crystal*, 69(4):564–576, 2013.
22. I Slabu, J L Galman, R C Lloyd, and N J Turner. Discovery, engineering, and synthetic application of transaminase biocatalysts. *ACS Catalysis*, 7(12):8263–8284, 2017.
23. A C Storer and A Cornish-Bowden. The kinetics of coupled enzyme reactions. Applications to the assay of glucokinase, with glucose 6-phosphate dehydrogenase as coupling enzyme. *Biochem J*, 141(1):205–209, 1974. doi: 10.1042/bj1410205.
24. W Stroberg and S Schnell. On the estimation errors of K_M and V from time-course experiments using the Michaelis–Menten equation. *Biophys Chem*, 219:17–27, 2016. doi: 10.1016/j.bpc.2016.09.004.
25. R Tolxdorff-Neutzling and J-H Klemme. Metabolic role and regulation of L-alanine dehydrogenase in *Rhodospseudomonas capsulata*. *FEMS Microbiol Letters*, 13(2):155–159, 1982.
26. T Werpy and G Petersen. Top value added chemicals from biomass: volume I—results of screening for potential candidates from sugars and synthesis gas. Technical report, National Renewable Energy Lab., Golden, CO (US), 2004.
27. A Yoshida and E Freese. Enzymic properties of alanine dehydrogenase of *Bacillus subtilis*. *BBA - Enzym Biol Oxid*, 96(2):248–262, 1965.