

Parameter inference to motivate asymptotic model reduction: an analysis of the gibberellin biosynthesis pathway

Leah R Band^{a,b}, Simon P Preston^b

^a*Division of Plant and Crop Sciences, School of Biosciences, University of Nottingham, Sutton Bonington, LE12 5RD.*

^b*School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD.*

Abstract

Developing effective strategies to use models in conjunction with experimental data is essential to understand the dynamics of biological regulatory networks. In this study, we demonstrate how combining parameter estimation with asymptotic analysis can reveal the key features of a network and lead to simplified models that capture the observed network dynamics. Our approach involves fitting the model to experimental data and using the Profile Likelihood to identify small parameters and cases where model dynamics are insensitive to changing particular individual parameters. Such parameter diagnostics provide understanding of the dominant features of the model and motivate asymptotic model reductions to derive simpler models in terms of identifiable parameter groupings.

We focus on the particular example of biosynthesis of the plant hormone gibberellin (GA), which controls plant growth and has been mutated in many current crop varieties. This pathway comprises two parallel series of enzyme-substrate reactions, which have previously been modelled using the law of mass action [23]. Considering the GA20ox-mediated steps, we analyse the identifiability of the model parameters using published experimental data; the analysis reveals the ratio between enzyme and GA levels to be small and motivates us to perform a quasi-steady state analysis to derive a reduced model. Fitting the parameters in the reduced model reveals additional features of the pathway and motivates further asymptotic analysis which produces a hierarchy of reduced models. Calculating the Akaike information criterion and parameter confidence intervals enables us to select a parsimonious model with identifiable parameters. As well as demonstrating the benefits of combining parameter estimation and asymptotic analysis, the analysis shows how GA biosynthesis is limited by the final GA20ox-mediated steps in the pathway and generates a simple mathematical description of this part of the GA biosynthesis pathway.

Keywords: asymptotic analysis, profile likelihood, plant hormones, hormone biosynthesis

1. Introduction

Ordinary differential equation (ODE) models of biological regulatory networks usually contain many unknown parameters. Such parameters are typically very challenging to measure experimentally and are often estimated indirectly by calibrating the model to experimental data such as a time course of measurements under an experimental perturbation. Such model calibration deepens understanding of the network dynamics and enables the model to make more accurate predictions (predicting, for instance, the dynamics of unseen components or the behaviour of the network under alternative perturbations). Furthermore, the calibration enables inference to be made about the parameter values themselves, which are often of scientific interest.

The question of whether or not parameters can be estimated well, or at all, from experimental data is called “identifiability”. There are many good reviews of approaches for determining identifiability of ODE models (see e.g. [22]) but relatively much less on how models with non-identifiable parameters can be reduced to obtain iden-

tifiable models, which is the focus of this paper.

A characteristic common to many ODE models is that model dynamics and data agree for parameter values that range over many orders of magnitude, and hence the parameters are not precisely identifiable from data. This phenomena is often referred to as “sloppiness” [4, 14] (which we will define more precisely below). In some circumstances, sloppiness can be mitigated with careful experimental design, i.e., the choice over what experiments to perform, including what to measure and when [30]. There are also circumstances in which sloppiness causes no issues; for example, even though parameter estimates may have very high variance, predictions for particular variables can still have low variance [4, 14].

In contrast to these approaches, here we consider how to reduce a sloppy model to a simpler one with identifiable parameters. We show how analysing parameter sensitivity provides understanding of the model dynamics that enables us to identify suitable asymptotic model reductions. Although the full model encapsulates the key mechanisms thought to generate the data, the reduced model is an approximation that describes the experimental data nearly

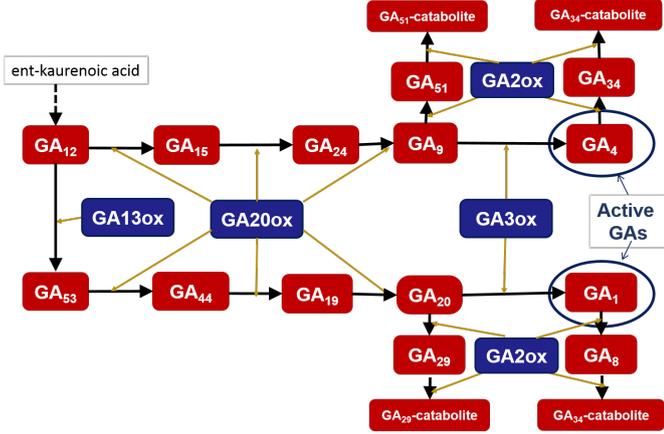


Figure 1: The gibberellin (GA) biosynthesis and degradation pathway. GA metabolites are shown in red and enzymes are shown in purple.

as well. The reduced model reveals the level of complexity required to capture the dominant dynamics on a particular time scale, and suggests how these dynamics depend on groupings of the original model parameters. Deriving the reduced model retains mechanistic interpretation and physical meaning of the parameter groupings, and fitting the reduced model reveals how well these parameter groupings can be identified from the data. Deriving reduced models is helpful not only for parameter inference but also for computational reasons: complex multiscale models are often created by coupling submodels for different components of a large system (e.g. [7, 3]) so having smaller model components helps to reduce the overall complexity and computational cost.

The particular application we consider is the biosynthesis of the plant hormone gibberellin (GA), which regulates organ growth, flower development and seed germination, and plays a key role in adaptation to environmental stress [8, 11]. GA biosynthesis has been shown to be predominantly regulated at the later steps in the pathway whereby GA_{12} is converted to active forms GA_4 and GA_1 [12]. Here, the GA biosynthesis pathway comprises two parallel series of enzyme-substrate reactions [15, 17, 33] (shown in Fig. 1): GA_{12} either undergoes a series of reactions catalysed by members of the GA20ox family, to produce GA_9 which is then converted to the active GA_4 via the GA3ox enzymes, or GA_{12} is converted to GA_{53} (via GA13ox) and then undergoes a series of GA20ox-mediated steps to produce GA_{20} which is converted to the active GA_1 by GA3ox. The active GAs and their precursors can also be degraded via the GA2ox family of enzymes. Thus, the level of bioactive GA depends on the balance between synthesis and degradation, which are controlled via the levels of GA20ox, GA3ox and GA2ox enzymes; these enzymes exhibit distinct spatial and temporal patterns and are regulated by environmental signals and other plant hormones [15, 17, 18, 25, 33].

The GA pathway has played a major role in improving crop productivity: during the Green Revolution in the 1960s, high-yielding semi-dwarf cereal crop varieties were developed, many of which were later found to have genetic mutations in the GA pathway [26]. Understanding the mechanisms that control GA levels is a prime target for further increasing yields [8, 16].

In this study, we focus on the GA20ox-mediated steps of the GA biosynthesis pathway and describe how parameter estimation, parameter sensitivity analysis and asymptotic analysis can generate a simpler description of this biological network, providing understanding of the underlying dynamics and identifying the parameter groupings that govern the network dynamics. We first describe our methods for a general case (§2), describing fitting criteria, derivation of confidence intervals and how the profile likelihood can be used to motivate asymptotic model reduction. We then apply our methods to analyse the dynamics of the GA20ox-mediated steps in the GA biosynthesis pathway (§3). Beginning with the model suggested by Middleton et al. [23], based on the law of mass action, we develop a series of reduced models, assessing each model by comparing simulation results with the experimental data from Appleford et al. [2]. We demonstrate how our methods motivate the derivation of a simple description of the observed dynamics and reveal the key network interactions. Our conclusions are discussed in §4.

2. Methods

2.1. Models and fitting criteria

The dynamics of a biological network are often modelled by a system of ODEs of the form

$$\frac{dy}{dt} = \mathbf{f}(\mathbf{y}, \boldsymbol{\theta}) \quad \text{with } \mathbf{y} = \mathbf{y}_0 \quad \text{at } t = 0, \quad (1)$$

where $\mathbf{y} = \mathbf{y}(t; \boldsymbol{\theta}) = (y_1(t; \boldsymbol{\theta}), \dots, y_s(t; \boldsymbol{\theta}))$ is a vector containing variables representing concentrations of the various interacting species included in the model; $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)$ is a vector of model parameters; and \mathbf{y}_0 is the initial value of \mathbf{y} at time $t = 0$. We suppose that only a subset of the y_1, \dots, y_s – here taken without loss of generality to be the first $m \leq s$ – are measurable experimentally, and that these are measured only at a finite number of time points and subject to measurement error. Supposing that variable y_j , say, is measured at the n_j time points t_{j1}, \dots, t_{jn_j} (not necessarily the same for different j) and that y_{ji}^{Data} denotes the experimental measurement of variable y_j at time t_{ji} , then the full set of experimental data, comprising $N = n_1 + \dots + n_m$ data points, is

$$\{y_{ji}^{\text{Data}}; i = 1, \dots, n_j, j = 1, \dots, m\}. \quad (2)$$

Fitting a model (1) to experimental data (2) involves determining $\boldsymbol{\theta}$ (and possibly some elements of \mathbf{y}_0 if they themselves are unknown) so that the model solution is close, by

some criterion, to the data. One approach is to minimise the sum-of-squares criterion,

$$S(\boldsymbol{\theta}) = \sum_{j=1}^m \sum_{i=1}^{n_j} (y_{ij}^{\text{Data}} - y_j(t_{ji}; \boldsymbol{\theta}))^2, \quad (3)$$

to find

$$\hat{\boldsymbol{\theta}} = \operatorname{argmin} S(\boldsymbol{\theta}), \quad (4)$$

which is the “least-squares estimate” of $\boldsymbol{\theta}$. Another, more statistically minded, approach is to define an error model relating the deterministic model (1) to the data (2), construct the “likelihood function”, $L(\boldsymbol{\theta})$ (see e.g. [9]) then estimate $\boldsymbol{\theta}$ by maximising $L(\boldsymbol{\theta})$ (or equivalently the log-likelihood function, $\ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$). For example, under a Gaussian error model

$$y_{ij}^{\text{Data}} = y_j(t_{ji}; \boldsymbol{\theta}) + \varepsilon_{ij}, \quad (5)$$

where the errors $\varepsilon_{ij} \sim N(0, \sigma^2)$ are independent Gaussian random variables (i.e. Normally distributed) with mean 0 and variance σ^2 , the log-likelihood function is

$$\ell(\boldsymbol{\theta}) = -\frac{N}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} S(\boldsymbol{\theta}), \quad (6)$$

where N is the number of data points and $S(\boldsymbol{\theta})$ is given by (3). For any σ^2 , $\ell(\boldsymbol{\theta})$ is maximised with respect to $\boldsymbol{\theta}$ by minimising $S(\boldsymbol{\theta})$. In other words, under a Gaussian error model, (5), which we adopt throughout this paper, the “maximum-likelihood estimate”, $\hat{\boldsymbol{\theta}} = \operatorname{argmax} \ell(\boldsymbol{\theta})$ is identical to the least-squares estimate of $\boldsymbol{\theta}$ (4). The advantage of the likelihood framework, however, is that it provides distributional results for making statistical inference about $\boldsymbol{\theta}$, as described below.

2.2. Confidence intervals for parameter values

The asymptotic distribution (i.e., as $N \rightarrow \infty$) of $\hat{\boldsymbol{\theta}}$, under mild conditions that commonly hold, is

$$\hat{\boldsymbol{\theta}} \sim N_p(\boldsymbol{\theta}, \mathbf{I}^{-1}(\boldsymbol{\theta})), \quad (7)$$

where $N_p(\cdot, \cdot)$ denotes the multivariate Gaussian distribution, and $\mathbf{I}(\boldsymbol{\theta}) = -E\mathbf{H}$ is the so-called Fisher information matrix, where E denotes expectation and

$$\mathbf{H}(\boldsymbol{\theta}) = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \quad (8)$$

is the Hessian matrix of ℓ with respect to $\boldsymbol{\theta}$ [29]. Result (7) can be used to construct individual or joint confidence regions for the components of $\boldsymbol{\theta}$. For example, from (7), by basic properties of the multivariate Gaussian distribution [21], θ_j is distributed as

$$\hat{\theta}_j \sim N(\theta_j, [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{jj}),$$

then, approximating $\mathbf{I}(\boldsymbol{\theta})$ with $-\mathbf{H}(\hat{\boldsymbol{\theta}})$, we can derive the bounds of a $100(1 - \alpha)\%$ confidence interval as

$$\hat{\theta}_j \pm \sqrt{\chi_1^2(\alpha) [-\mathbf{H}^{-1}(\hat{\boldsymbol{\theta}})]_{jj}}, \quad (9)$$

where $\chi_1^2(\alpha)$ is the α upper quantile of the χ_1^2 distribution.

The term “sloppiness”, introduced above, is now sometimes used very loosely to mean “difficulty in estimating parameters”, though the term was originally introduced with the more particular meaning that the spectrum of eigenvalues of \mathbf{H} span many orders of magnitude, and include some very small values. Small eigenvalues make \mathbf{H} ill-conditioned, inflating elements of \mathbf{H}^{-1} and hence inflating the confidence intervals for the individual parameters. We show later in Fig. 7 the eigenvalue spectra of models of different complexity, and in Tables 2 and 3 examples of confidence intervals computed using (9). For details of the numerical calculations of \mathbf{H} in this paper, see Appendix A.

A different way to investigate the information about particular parameters contained within the data is by the “profile likelihood” [28, 32]. This approach explores parameter space in the neighbourhood of $\hat{\boldsymbol{\theta}}$ by varying a single parameter and fitting the model with respect to the remaining parameters. As described below, this provides insight into the shape of the likelihood in the neighbourhood of the maximum. To calculate the profile likelihood, consider a particular parameter of interest, θ_j , denote the other parameters, $\boldsymbol{\theta}_{-j}$, and write the log-likelihood as $\ell(\theta_j, \boldsymbol{\theta}_{-j})$. The log-profile-likelihood is defined as

$$\ell_P(\theta_j) = \ell(\theta_j, \hat{\boldsymbol{\theta}}_{-j}), \quad (10)$$

where $\hat{\boldsymbol{\theta}}_{-j} = \operatorname{argmax}_{\boldsymbol{\theta}_{-j}} \ell(\theta_j, \boldsymbol{\theta}_{-j})$, i.e., (10) is the log likelihood maximised over all the parameters except the particular one of interest. By Wilks’ theorem [31], under a null hypothesis that θ_j takes a particular value, θ_j^0 , say, then asymptotically

$$2 \left(\ell(\hat{\boldsymbol{\theta}}) - \ell_P(\theta_j^0) \right) \sim \chi_1^2, \quad (11)$$

which can be rearranged to give the confidence interval

$$\left\{ \theta_j : 2 \left(\ell(\hat{\boldsymbol{\theta}}) - \ell_P(\theta_j) \right) < \chi_1^2(\alpha) \right\}. \quad (12)$$

The confidence interval can be visualised by marking the $\chi_1^2(\alpha)$ threshold on a plot of $-2\ell_P(\theta_j)$ versus θ_j ; see for example, Fig. 10.

An intuitive view of the Gaussian- and profile-likelihood-based confidence intervals, (9, 12), is that they both involve investigating the curvature of $\ell(\boldsymbol{\theta})$ in the vicinity of $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$: flatness indicates that the data are not informative about $\boldsymbol{\theta}$, whereas high curvature indicates that the data are very informative about $\boldsymbol{\theta}$. The Gaussian-based confidence intervals amount to approximating $\ell(\boldsymbol{\theta})$ by a quadratic with its curvature matched at the maximum, $\hat{\boldsymbol{\theta}}$. This entails an assumption of symmetry that may not be warranted for the model and data (see, e.g., Fig 11). The profile-likelihood confidence intervals involve no such assumption and are often more accurate than the Gaussian-based confidence intervals for nonlinear models and data sets of typical size [28, 32]. Another advantage of profile-likelihood confidence intervals is that they are invariant

to the choice of parameterisation, e.g. whether or not the parameters are defined on a log scale (as in this paper), whereas Gaussian-based ones are not.

A point sometimes not highlighted is that the asymptotic results (7, 11) and hence confidence intervals derived from them, (9, 12), require that θ is a point in the interior of parameter space rather than a point on the boundary. This is not always the case: for example, some of the parameters that are non-negative on physical grounds may be exactly zero. In Model II later $\ell(\theta)$ is maximised when one or more of the parameters approaches zero, in which case we do not compute confidence intervals. Another point we note is that in the derivations of the confidence intervals above we have assumed that σ^2 is known, whereas here we substitute σ^2 with its estimate $\hat{\sigma}^2 = N^{-1}S(\hat{\theta})$, where $S(\hat{\theta})$ is computed from the fitted model. This has little effect on confidence intervals and saves the expressions for the confidence intervals from becoming more complicated. Details of how to compute the confidence intervals without making this substitution are given in [29, p196].

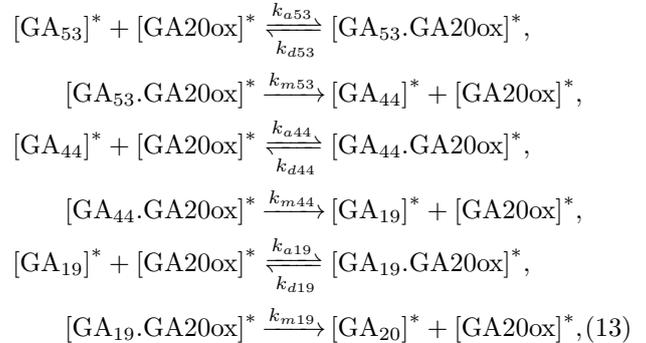
2.3. Profile likelihood motivating model reduction

As we demonstrate in our case study below, the presence of non-identifiable parameters suggests that the observed dynamics can be captured by a simpler model. To derive the simpler model, the profile likelihood can be used to suggest an appropriate asymptotic model reduction. For example, if the confidence interval is unbounded in one direction, we take the corresponding limit in which $\theta_i \rightarrow 0$ or $\theta_i \rightarrow \infty$. In some cases, the model dynamics may depend on θ_j only in a ratio with other parameters, in which case plotting how the other fitted parameters, $\hat{\theta}_{-j}$, vary as θ_j is varied suggests an appropriate model reduction. Taking an asymptotic limit may also be appropriate for an identifiable parameter if the estimated value and profile likelihood show this parameter to be large or small. Having derived a reduced model, we compute an information criterion (AICc, see Appendix B) to select amongst the candidate models. A schematic of this model reduction approach is shown in Fig. 2.

3. Modelling the GA20ox-mediated steps in the GA biosynthesis pathway

The GA biosynthesis pathway features two parallel series of oxidation steps mediated by the GA20ox enzymes (Fig. 1). Each of these steps involves the GA species first binding to GA20ox with a reversible reaction, and the resulting complex then dissociating into the next GA species in the pathway and GA20ox [23] (Fig. 3a). Thus, the con-

version of GA₅₃ to GA₂₀ (Fig. 3b) involves the reactions



where k_{dj} , k_{aj} and k_{mj} denote rate constants for $j = 53, 44, 19$, square brackets denote concentrations and asterisks denote dimensional quantities. Equivalent reactions govern the conversion of GA₁₂ to GA₉ (Fig. 3c); however, to maintain conciseness, we give reactions and equations for only the GA₅₃ pathway throughout.

3.1. Middleton et al's model: Model I

To describe the reactions (13), we initially adopt the model by Middleton et al. [23] who modelled the conversion of GA₁₂ to GA₉ using the law of mass action. Thus, reactions (13) can be described by a system of coupled

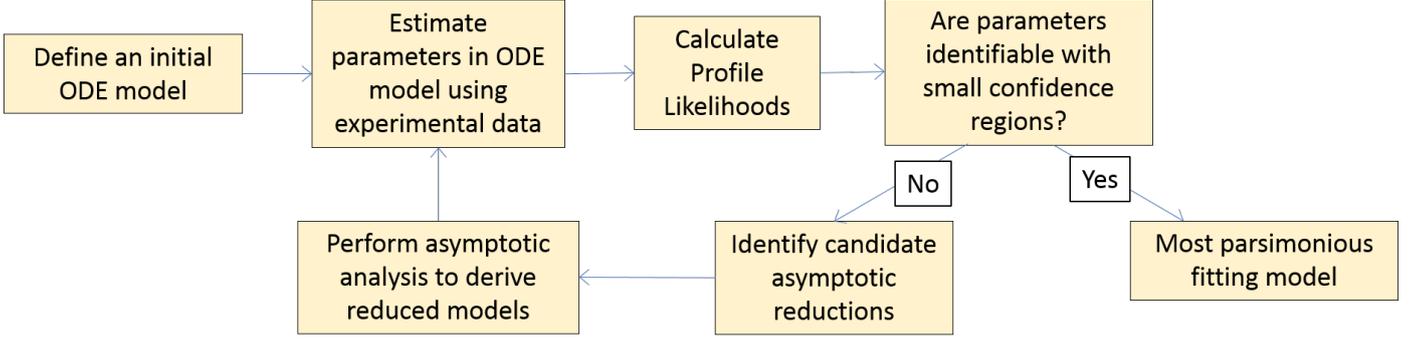


Figure 2: Schematic showing the key steps in using parameter inference to identify and derive reduced versions of sloppy models.

ordinary differential equations (ODEs):

$$\frac{d[\text{GA}_{53}]^*}{dt^*} = -k_{a53}[\text{GA}_{53}]^*[\text{GA}_{20\text{ox}}]^* + k_{d53}[\text{GA}_{53}\cdot\text{GA}_{20\text{ox}}]^*,$$

$$\frac{d[\text{GA}_{44}]^*}{dt^*} = k_{m53}[\text{GA}_{53}\cdot\text{GA}_{20\text{ox}}]^* - k_{a44}[\text{GA}_{44}]^*[\text{GA}_{20\text{ox}}]^* + k_{d44}[\text{GA}_{44}\cdot\text{GA}_{20\text{ox}}]^*,$$

$$\frac{d[\text{GA}_{19}]^*}{dt^*} = k_{m44}[\text{GA}_{44}\cdot\text{GA}_{20\text{ox}}]^* - k_{a19}[\text{GA}_{19}]^*[\text{GA}_{20\text{ox}}]^* + k_{d19}[\text{GA}_{19}\cdot\text{GA}_{20\text{ox}}]^*,$$

$$\frac{d[\text{GA}_{20}]^*}{dt^*} = k_{m19}[\text{GA}_{19}\cdot\text{GA}_{20\text{ox}}]^*,$$

$$\frac{d[\text{GA}_{53}\cdot\text{GA}_{20\text{ox}}]^*}{dt^*} = k_{a53}[\text{GA}_{53}]^*[\text{GA}_{20\text{ox}}]^* - (k_{d53} + k_{m53})[\text{GA}_{53}\cdot\text{GA}_{20\text{ox}}]^*,$$

$$\frac{d[\text{GA}_{44}\cdot\text{GA}_{20\text{ox}}]^*}{dt^*} = k_{a44}[\text{GA}_{44}]^*[\text{GA}_{20\text{ox}}]^* - (k_{d44} + k_{m44})[\text{GA}_{44}\cdot\text{GA}_{20\text{ox}}]^*,$$

$$\frac{d[\text{GA}_{19}\cdot\text{GA}_{20\text{ox}}]^*}{dt^*} = k_{a19}[\text{GA}_{19}]^*[\text{GA}_{20\text{ox}}]^* - (k_{d19} + k_{m19})[\text{GA}_{19}\cdot\text{GA}_{20\text{ox}}]^*,$$

$$\begin{aligned} \frac{d[\text{GA}_{20\text{ox}}]^*}{dt^*} = & (k_{d53} + k_{m53})[\text{GA}_{53}\cdot\text{GA}_{20\text{ox}}]^* \\ & + (k_{d44} + k_{m44})[\text{GA}_{44}\cdot\text{GA}_{20\text{ox}}]^* \\ & + (k_{d19} + k_{m19})[\text{GA}_{19}\cdot\text{GA}_{20\text{ox}}]^* \\ & - k_{a53}[\text{GA}_{53}]^*[\text{GA}_{20\text{ox}}]^* \\ & - k_{a44}[\text{GA}_{44}]^*[\text{GA}_{20\text{ox}}]^* \\ & - k_{a19}[\text{GA}_{19}]^*[\text{GA}_{20\text{ox}}]^*, \end{aligned} \quad (14)$$

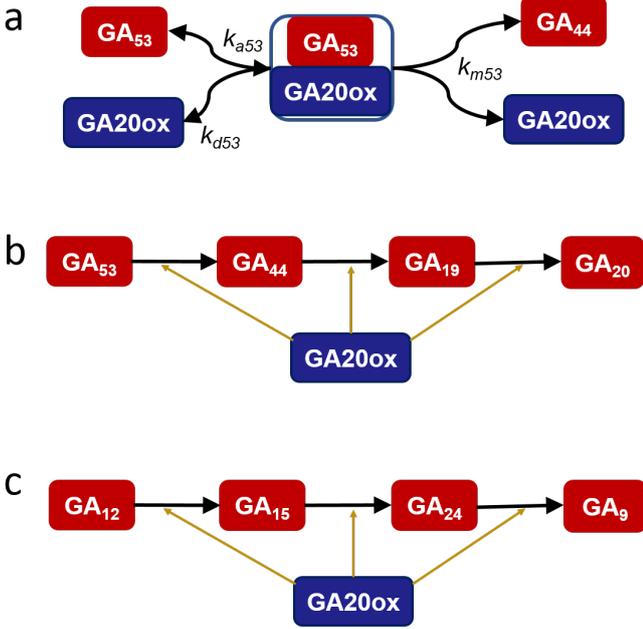


Figure 3: (a) The reactions involved in a single oxidation step, shown for the conversion of GA_{53} to GA_{44} . (b) The series of $\text{GA}_{20\text{ox}}$ -mediated reactions that convert GA_{53} to GA_{20} . (c) The series of $\text{GA}_{20\text{ox}}$ -mediated reactions that convert GA_{12} to GA_9 .

where t^* denotes time.

To estimate the parameters in this model, we follow [23] in comparing the model dynamics to data from Appleford et al. [2] who performed in vitro studies to investigate the dynamics of the conversion of GA_{53} to GA_9 and GA_{12} to GA_{20} . Considering the GA_{53} conversion, for ex-

ample, their solution initially contained fixed concentrations of GA_{53} and $\text{GA}_{20\text{ox}}$, which we denote by s_0 and e_0 respectively, and the concentrations of GA_{53} , GA_{44} , GA_{19} and GA_{20} were measured at subsequent time points (see data points in Fig. 4). Thus, the data corresponds to the initial conditions $[\text{GA}_{53}]^* = s_0$, $[\text{GA}_{20\text{ox}}]^* = e_0$, $[\text{GA}_{44}]^* = [\text{GA}_{19}]^* = [\text{GA}_{20}]^* = [\text{GA}_{53}\cdot\text{GA}_{20\text{ox}}]^* = [\text{GA}_{44}\cdot\text{GA}_{20\text{ox}}]^* = [\text{GA}_{19}\cdot\text{GA}_{20\text{ox}}]^* = 0$. In the experiments [2], the initial concentration s_0 is given, which following [23] we take to be known initial conditions. The initial $\text{GA}_{20\text{ox}}$ concentration e_0 is an unknown and is a parameter to be estimated.

Fitting the model parameters to the data (as described in §2.1 and Appendix A), we see that equations (14) (hereafter referred to as Model I), can represent the observed dynamics for the GA_{53} and GA_{12} pathways; see blue lines in Fig. 4 and 5 respectively (consistent with the results in [23] for the GA_{12} pathway). The profile likelihood for each parameter (Fig. 6) shows that the fitted model remains in agreement with the data for wide ranges of parameter values. This conclusion is consistent with the eigenvalue spectrum of the Hessian varying over several orders of magnitude (Fig. 7), and hence the model being sloppy. The model's sloppiness motivates us to seek a simpler model for the observed dynamics. The profile likelihood suggests that to fit the data the estimated initial $\text{GA}_{20\text{ox}}$ concentration, e_0 , needs to be small compared to the prescribed initial GA_{53} concentration, s_0 . Thus, motivated by the profile likelihoods and previous studies of enzyme-substrate reactions, we will consider the limit in which $e_0 \ll s_0$.

3.2. Model reduction, Model II

Applying the law of mass action is a standard way of physically describing enzyme-substrate reactions; however, the resulting Model I is complex with the model dynamics depending on 10 parameter values. Given the profile likelihood revealed that Model I only fits the data if e_0/s_0 is small, we now consider a model reduction by taking the limit $e_0/s_0 \rightarrow 0$. This quasi-steady-state assumption is the basis of the Briggs-Haldane derivation of the Michaelis-Menten equation [24], and, for a single substrate-enzyme reaction, has been shown to generate a reduced model in which parameters are identifiable using standard time course data [6]. In the limit as $e_0/s_0 \rightarrow 0$, the complexes form rapidly and the complex concentrations are in equilibrium after an initial transient time scale.

3.2.1. Nondimensionalisation

To derive a reduced model, we first non-dimensionalise the governing equations. We consider the GA concentrations relative to the initial GA_{53} concentration, s_0 , the complex and enzyme concentrations relative to the initial enzyme concentration, e_0 and time relative to the rate at

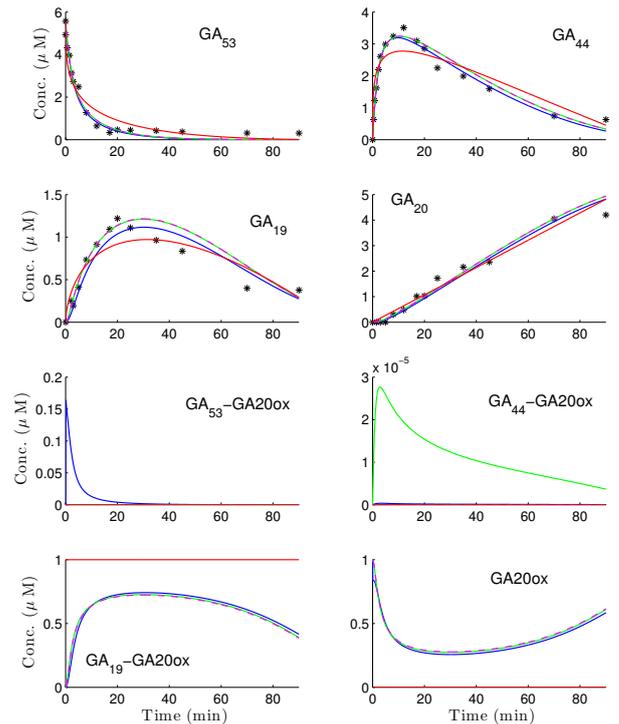


Figure 4: Fits of model dynamics to the GA_{53} pathway data; blue lines = Model I; green lines = Model II; purple dashed lines = Model III; red lines = Model IV. For these data, the known initial concentration of GA_{53} is $s_0 = 5.56 \mu\text{M}$. We note that there is no noticeable difference between the predicted Model II and Model III dynamics and that in this case, Model IV does not capture the dynamics well (as is reflected in the $S(\hat{\theta})$ and AICc values in Table 1). In the upper panels, the predicted GA_{53} , GA_{44} , GA_{19} and GA_{20} for Models II-IV are redimensionalised via (15) to show comparison with the experimental data.

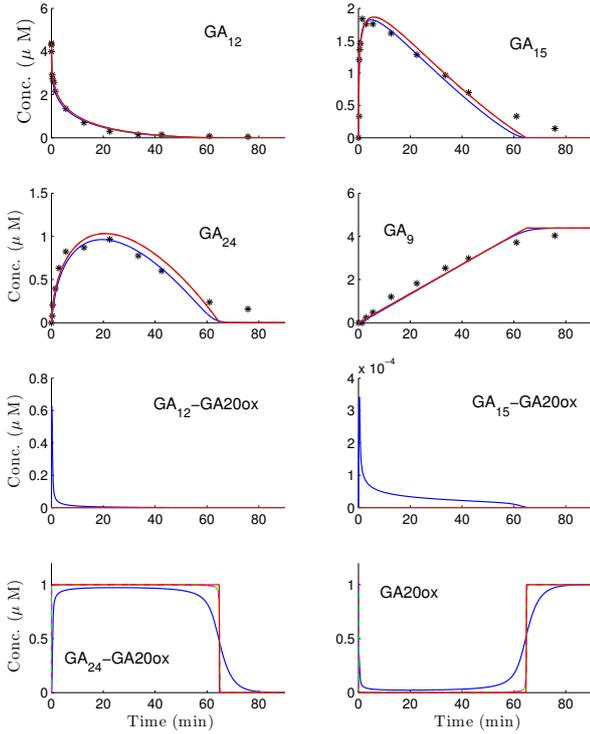


Figure 5: Fits of model dynamics to the GA_{12} pathway data; blue lines = Model I; green lines = Model II; purple dashed lines = Model III; red lines = Model IV. For these data, the known initial concentration of GA_{53} is $s_0 = 4.38 \mu\text{M}$. We note that there is no noticeable difference between the predicted dynamics of Model II, Model III or Model IV and as a result the green, purple and red lines cannot be easily distinguished. In the upper panels, the predicted GA_{12} , GA_{15} , GA_{24} and GA_9 for Models II-IV are redimensionalised via (15) to show comparison with the experimental data.

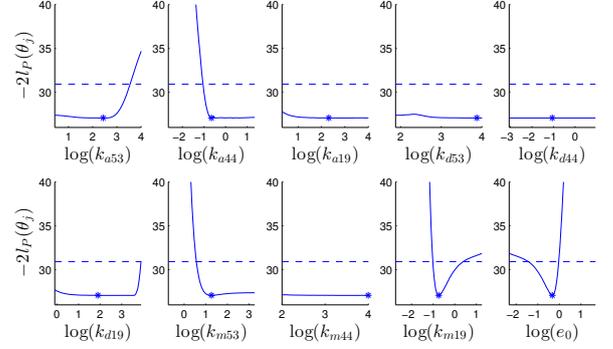


Figure 6: Profile likelihoods for the full dimensional model (Model I) fit to the GA_{53} pathway data. Each panel shows the profile likelihood for the parameter (solid lines), the estimated parameter value (asterisk) and the confidence interval threshold for $\alpha = 0.95$ (dashed lines). We note that for this GA_{53} pathway $s_0 = 5.56 \mu\text{M}$ and therefore the $\log(e_0)$ confidence region suggests that $e_0/s_0 \ll 1$ which motivates the asymptotic reduction to Model II. Similar profile likelihoods are obtained when fitting with the GA_{12} pathway data (results not shown).

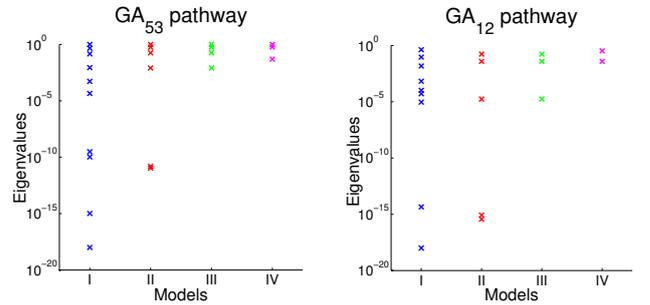


Figure 7: Eigenvalues of $-\mathbf{H}(\hat{\theta})$, for each model ($\lambda_k / \max(\lambda_k)$) for $k = 1, 2, \dots, p$, where \mathbf{H} is defined in (8).

which GA_{53} binds to GA20ox ; thus

$$\begin{aligned}
& ([\text{GA}_{53}]^*, [\text{GA}_{44}]^*, [\text{GA}_{19}]^*, [\text{GA}_{20}]^*) \\
& = s_0([\text{GA}_{53}], [\text{GA}_{44}], [\text{GA}_{19}], [\text{GA}_{20}]), \\
& ([\text{GA}_{53} \cdot \text{GA20ox}]^*, [\text{GA}_{44} \cdot \text{GA20ox}]^*, \\
& [\text{GA}_{19} \cdot \text{GA20ox}]^*, [\text{GA20ox}]^*) = \\
& e_0([\text{GA}_{53} \cdot \text{GA20ox}], [\text{GA}_{44} \cdot \text{GA20ox}], \\
& [\text{GA}_{19} \cdot \text{GA20ox}], [\text{GA20ox}]), \\
& t^* = \frac{1}{k_{a53}e_0}t. \tag{15}
\end{aligned}$$

We introduce the following dimensionless parameter groupings:

$$\begin{aligned}
\epsilon &= \frac{e_0}{s_0}, \quad (\bar{\alpha}_{44}, \bar{\alpha}_{19}) = \frac{1}{k_{a53}}(k_{a44}, k_{a19}), \\
(\bar{\lambda}_{53}, \bar{\lambda}_{44}, \bar{\lambda}_{19}) &= \frac{1}{k_{a53}s_0}(k_{m53}, k_{m44}, k_{m19}), \\
(\bar{\kappa}_{53}, \bar{\kappa}_{44}, \bar{\kappa}_{19}) &= \frac{1}{k_{a53}s_0}(k_{d53} + k_{m53}, \\
& k_{d44} + k_{m44}, k_{d19} + k_{m19}). \tag{16}
\end{aligned}$$

On nondimensionalising in this fashion, the governing equations, (14), become

$$\begin{aligned}
\frac{d[\text{GA}_{53}]}{dt} &= -[\text{GA}_{53}][\text{GA20ox}] \\
& + (\bar{\kappa}_{53} - \bar{\lambda}_{53})[\text{GA}_{53} \cdot \text{GA20ox}], \tag{17a}
\end{aligned}$$

$$\begin{aligned}
\frac{d[\text{GA}_{44}]}{dt} &= \bar{\lambda}_{53}[\text{GA}_{53} \cdot \text{GA20ox}] \\
& - \bar{\alpha}_{44}[\text{GA}_{44}][\text{GA20ox}] \\
& + (\bar{\kappa}_{44} - \bar{\lambda}_{44})[\text{GA}_{44} \cdot \text{GA20ox}], \tag{17b}
\end{aligned}$$

$$\begin{aligned}
\frac{d[\text{GA}_{19}]}{dt} &= \bar{\lambda}_{44}[\text{GA}_{44} \cdot \text{GA20ox}] \\
& - \bar{\alpha}_{19}[\text{GA}_{19}][\text{GA20ox}] \\
& + (\bar{\kappa}_{19} - \bar{\lambda}_{19})[\text{GA}_{19} \cdot \text{GA20ox}], \tag{17c}
\end{aligned}$$

$$\frac{d[\text{GA}_{20}]}{dt} = \bar{\lambda}_{19}[\text{GA}_{19} \cdot \text{GA20ox}], \tag{17d}$$

$$\begin{aligned}
\epsilon \frac{d[\text{GA}_{53} \cdot \text{GA20ox}]}{dt} &= [\text{GA}_{53}][\text{GA20ox}] \\
& - \bar{\kappa}_{53}[\text{GA}_{53} \cdot \text{GA20ox}], \tag{17e}
\end{aligned}$$

$$\begin{aligned}
\epsilon \frac{d[\text{GA}_{44} \cdot \text{GA20ox}]}{dt} &= \bar{\alpha}_{44}[\text{GA}_{44}][\text{GA20ox}] \\
& - \bar{\kappa}_{44}[\text{GA}_{44} \cdot \text{GA20ox}], \tag{17f}
\end{aligned}$$

$$\begin{aligned}
\epsilon \frac{d[\text{GA}_{19} \cdot \text{GA20ox}]}{dt} &= \bar{\alpha}_{19}[\text{GA}_{19}][\text{GA20ox}] \\
& - \bar{\kappa}_{19}[\text{GA}_{19} \cdot \text{GA20ox}], \tag{17g}
\end{aligned}$$

$$\begin{aligned}
\epsilon \frac{d[\text{GA20ox}]}{dt} &= \bar{\kappa}_{53}[\text{GA}_{53} \cdot \text{GA20ox}] \\
& + \bar{\kappa}_{44}[\text{GA}_{44} \cdot \text{GA20ox}] \\
& + \bar{\kappa}_{19}[\text{GA}_{19} \cdot \text{GA20ox}] \\
& - [\text{GA}_{53}][\text{GA20ox}] \\
& - \bar{\alpha}_{44}[\text{GA}_{44}][\text{GA20ox}] \\
& - \bar{\alpha}_{19}[\text{GA}_{19}][\text{GA20ox}], \tag{17h}
\end{aligned}$$

with initial conditions,

$$[\text{GA}_{53}] = 1, \quad [\text{GA}_{44}] = [\text{GA}_{19}] = [\text{GA}_{20}] = 0, \tag{18a}$$

$$[\text{GA20ox}] = 1, \quad [\text{GA}_{53} \cdot \text{GA20ox}] = 0, \tag{18b}$$

$$[\text{GA}_{44} \cdot \text{GA20ox}] = [\text{GA}_{19} \cdot \text{GA20ox}] = 0. \tag{18c}$$

3.2.2. Model reduction

Having non-dimensionalised, we reduce the model by taking the limit $e_0/s_0 \rightarrow 0$. Considering $t^* = O(1)$, equations (17e-g) give expressions for the complex concentrations in the limit as $\epsilon \rightarrow 0$

$$\begin{aligned}
[\text{GA}_{53} \cdot \text{GA20ox}] &= \kappa_{53}[\text{GA}_{53}][\text{GA20ox}], \\
[\text{GA}_{44} \cdot \text{GA20ox}] &= \kappa_{44}[\text{GA}_{44}][\text{GA20ox}], \\
[\text{GA}_{19} \cdot \text{GA20ox}] &= \kappa_{19}[\text{GA}_{19}][\text{GA20ox}], \tag{19}
\end{aligned}$$

where

$$\kappa_{53} = \frac{1}{\bar{\kappa}_{53}}, \quad \kappa_{44} = \frac{\bar{\alpha}_{44}}{\bar{\kappa}_{44}}, \quad \kappa_{19} = \frac{\bar{\alpha}_{19}}{\bar{\kappa}_{19}}, \tag{20}$$

(i.e. $\kappa_j = k_{aj}s_0/(k_{dj} + k_{mj})$ the ratio between the typical complex association rate and the complex dissociation rate). Summing (17e-h) shows that the total enzyme concentration is conserved; thus the initial conditions, (18b-c), require that the enzyme and complex concentrations sum to one for all times, and from (19), we obtain

$$\begin{aligned}
[\text{GA20ox}] &= \\
& \frac{1}{1 + \kappa_{53}[\text{GA}_{53}] + \kappa_{44}[\text{GA}_{44}] + \kappa_{19}[\text{GA}_{19}]}. \tag{21}
\end{aligned}$$

Substituting (19, 21) into (17a-d) and returning to dimensional time, t^* , (for ease of fitting with the data), we obtain

$$\begin{aligned}
\frac{d[\text{GA}_{53}]}{dt^*} &= \\
& \frac{-\lambda_{53}[\text{GA}_{53}]}{1 + \kappa_{53}[\text{GA}_{53}] + \kappa_{44}[\text{GA}_{44}] + \kappa_{19}[\text{GA}_{19}]}, \\
\frac{d[\text{GA}_{44}]}{dt^*} &= \\
& \frac{\lambda_{53}[\text{GA}_{53}] - \lambda_{44}[\text{GA}_{44}]}{1 + \kappa_{53}[\text{GA}_{53}] + \kappa_{44}[\text{GA}_{44}] + \bar{\kappa}_{19}[\text{GA}_{19}]}, \\
\frac{d[\text{GA}_{19}]}{dt^*} &= \\
& \frac{\lambda_{44}[\text{GA}_{44}] - \lambda_{19}[\text{GA}_{19}]}{1 + \kappa_{53}[\text{GA}_{53}] + \kappa_{44}[\text{GA}_{44}] + \kappa_{19}[\text{GA}_{19}]}, \\
\frac{d[\text{GA}_{20}]}{dt^*} &= \\
& \frac{\lambda_{19}[\text{GA}_{19}]}{1 + \kappa_{53}[\text{GA}_{53}] + \kappa_{44}[\text{GA}_{44}] + \kappa_{19}[\text{GA}_{19}]}. \tag{22}
\end{aligned}$$

where parameter groupings

$$\lambda_j = \frac{k_{a53}e_0\bar{\lambda}_j}{\kappa_j} = \frac{k_{mj}(k_{dj} + k_{mj})}{k_{aj}s_0}, \tag{23}$$

for $j = 53, 44, 19$, have dimensions min^{-1} . In summary, by assuming that the enzyme concentrations are small relative to the GA concentrations, the governing equations,

Model	p	GA ₅₃ pathway N=54		GA ₁₂ pathway N=46	
		$S(\hat{\theta})$	AICc	$S(\hat{\theta})$	AICc
I	10	2.7266	-136.12	1.5584	-129.42
II	6	2.9559	-143.09	1.6505	-138.91
III	4	2.9560	-148.06	1.6506	-144.09
IV	3	7.5705	-99.62	1.6605	-146.22

Table 1: Comparison of model fits. The Akaike Information Criterion (AICc) provides a method of model selection by assessing goodness of fit whilst penalising over-complicated models that have more parameters (see definition in Appendix B)

(17), reduce to four ODEs for the GA concentrations (22) coupled to four expressions for the quasisteady complex and enzyme concentrations (19, 21) for $t^* = O(1)$, with initial conditions

$$[GA_{53}] = 1, \quad [GA_{44}] = [GA_{19}] = [GA_{20}] = 0. \quad (24)$$

These equations and expressions depend on 6 parameter groupings (20, 23). The reduced model cannot satisfy initial conditions (18b-c). As in the Michaelis-Menten approximation of a single enzyme-substrate reaction, the complex concentrations evolve to their quasisteady state on an initial time scale $0 \leq t^* \ll 1$ [24]; however, to remain focussed on data-fitting, we consider the derivation of this inner solution to be beyond the scope of this study.

Fitting Model II to the experimental data [2] (as described in Appendix A), we see that the reduced model captures well the observed dynamics (Fig. 4 and 5), suggesting that the complex concentrations are indeed quasisteady on the time scale of the experiment. Since Model II is derived as an asymptotic reduction of Model I, it fits the data slightly less well than Model I for both pathways in the sense of having a larger residual sum-of-squares, $S(\hat{\theta})$, but it involves four fewer parameters and by AICc Model II is favoured over Model I (see Table 1 and Appendix B).

The profile likelihoods show different behaviour for the two pathways (Figs. 8 and 9). For the GA₅₃ pathway, the profile likelihood indicates that the model dynamics fit the data well when κ_{53} and κ_{44} are small; however, the confidence intervals of the remaining four parameters are finite (and do not include zero) (Fig. 8). For the GA₁₂ pathway, we similarly see that κ_{12} and κ_{15} are small, but in this case, the other parameters have unbounded confidence intervals and hence any of these may be very large (Fig. 9). For both pathways, the eigenvalue spectra of the Hessian is sloppy (Fig. 7). Given the similar features in the profile likelihoods of the two pathways, we consider a model reduction by taking κ_{53} and κ_{44} (or equivalently κ_{12} and κ_{15}) to be small. We note that the profile likelihoods (Figs. 8 and 9) also show that the fitted values of the other parameters do not change as κ_{53} , κ_{44} , κ_{12} and κ_{15} varied from small values, and thus, it appears that these four parameters have a negligible influence on the model dynamics on this time scale.

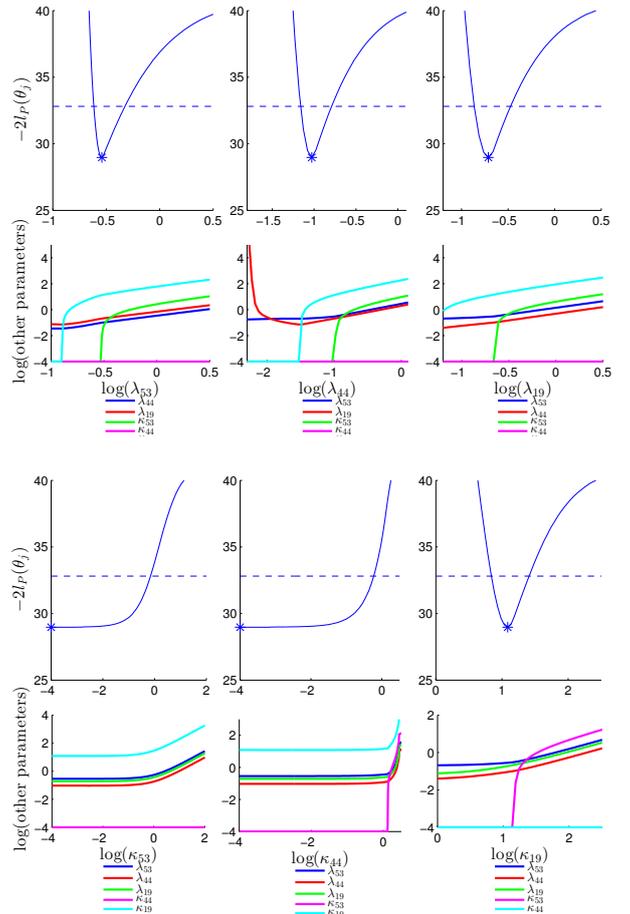


Figure 8: Profile likelihoods for Model II fit to the GA₅₃ pathway data. The first and third row of panels show the profile likelihood for each parameter (solid lines), the estimated parameter value (asterisk) and the confidence interval threshold for $\alpha = 0.95$ (dashed lines). The second and fourth row of panels show how the other parameter values vary along these Likelihood profiles. In the numerical optimization, we have imposed a lower bound of 10^{-4} for the value of each parameter.

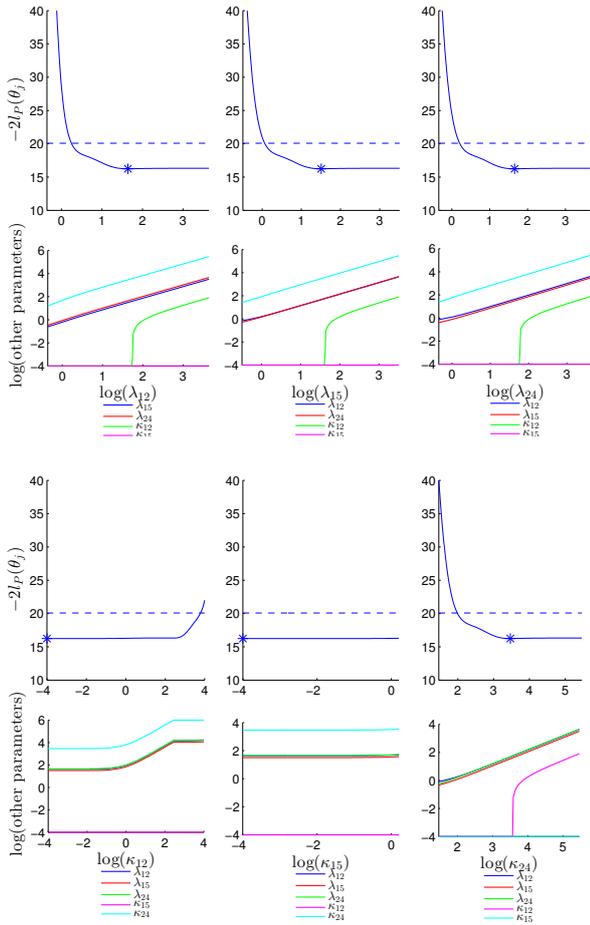


Figure 9: Profile likelihoods for Model II fit to the GA_{12} pathway data. The first and third row of panels show the profile likelihood for each parameter (solid lines), the estimated parameter value (asterisk) and the confidence interval threshold for $\alpha = 0.95$ (dashed lines). The second and fourth row of panels show how the other parameter values vary along these Likelihood profiles. In the numerical optimization, we have imposed a lower bound of 10^{-4} for the value of each parameter.

3.3. A further model reduction: Model III

The profile likelihoods (Fig. 8 and 9) show that the model fits the data equally well for any small values of the parameters κ_{53} and κ_{44} for the GA_{53} pathway and κ_{12} and κ_{15} for the GA_{12} pathway. This motivates a further model reduction (Model III) in which we suppose that $\epsilon \ll \kappa_{53} \ll 1$, $\kappa_{44} = O(\kappa_{53})$ as $\kappa_{53} \rightarrow 0$. Here the dynamics depend on four parameters, and the governing equations are

$$\begin{aligned} \frac{d[\text{GA}_{53}]}{dt^*} &= \frac{-\lambda_{53}[\text{GA}_{53}]}{1 + \kappa_{19}[\text{GA}_{19}]}, \\ \frac{d[\text{GA}_{44}]}{dt^*} &= \frac{\lambda_{53}[\text{GA}_{53}] - \lambda_{44}[\text{GA}_{44}]}{1 + \kappa_{19}[\text{GA}_{19}]}, \\ \frac{d[\text{GA}_{19}]}{dt^*} &= \frac{\lambda_{44}[\text{GA}_{44}] - \lambda_{19}[\text{GA}_{19}]}{1 + \kappa_{19}[\text{GA}_{19}]}, \\ \frac{d[\text{GA}_{20}]}{dt^*} &= \frac{\lambda_{19}[\text{GA}_{19}]}{1 + \kappa_{19}[\text{GA}_{19}]}, \end{aligned} \quad (25)$$

which are coupled to an expression for the enzyme concentration

$$[\text{GA}_{20\text{ox}}] = \frac{1}{1 + \kappa_{19}[\text{GA}_{19}]}, \quad (26)$$

and expressions (19) for the complex concentrations, and subject to initial conditions (24). In this limit, the dissociation rates of complexes $\text{GA}_{53}\text{-GA}_{20\text{ox}}$ and $\text{GA}_{44}\text{-GA}_{20\text{ox}}$ are much larger than their association rates so that their concentrations are asymptotically small.

Fitting the parameters for Model III, we find no noticeable difference between the predicted Model III and Model II dynamics (see Fig. 4 and 5); the fitted parameters and residual sum-of-squares, $S(\hat{\theta})$, are similar to those of Model II; however, since Model III has fewer parameters than Model II, Model III has a lower AICc (Table 1). For both pathways, the range of eigenvalues of the Hessian is much smaller for Model III than Model II, showing that Model III is less sloppy than Model II (Fig. 7). The profile likelihoods show that the parameters are identifiable for the GA_{53} pathway and we find reasonable agreement between the Gaussian- and profile-likelihood-based confidence intervals in this case (Fig. 10, Table 2). In contrast, the profile-likelihood based confidence intervals are unbounded for the GA_{12} pathway (Fig. 11). Thus, the profile likelihoods suggest that dynamics of the GA_{53} pathway appear to be parsimoniously described by Model III, whereas those of the GA_{12} pathway may be captured by a simpler model. The profile likelihoods for the GA_{12} pathway show that any of the four parameters may be large, and there appears to be a linear relationships between the parameters along each profile likelihood (Fig 11), which suggests that the model dynamics depend on ratios between certain parameter values.

3.4. A final model reduction: Model IV

Motivated by the profile likelihoods of the GA_{12} pathway, we consider the limit in which $\kappa_{19} \rightarrow \infty$, and introduce $\tilde{\lambda}_j = \lambda_j/\kappa_{19}$ for $j = 53, 44, 19$ which we take to be

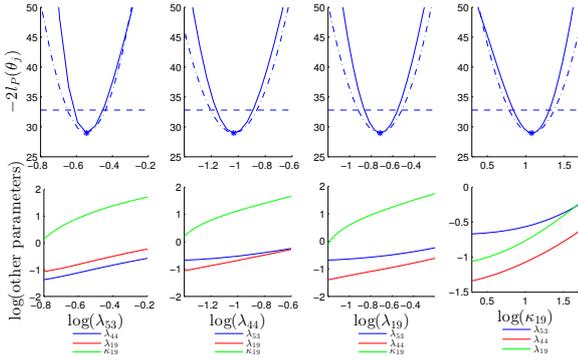


Figure 10: Profile likelihoods for Model III fit to the GA₅₃ pathway data. Upper panels show the profile likelihood for each parameter (solid lines), the approximation to the likelihood entailed by the Gaussian-based confidence intervals (dot-dash), the estimated parameter value (asterisk) and the confidence interval threshold for $\alpha = 0.95$ (dashed lines). Lower panels show how the other parameter values vary along these likelihood profiles.

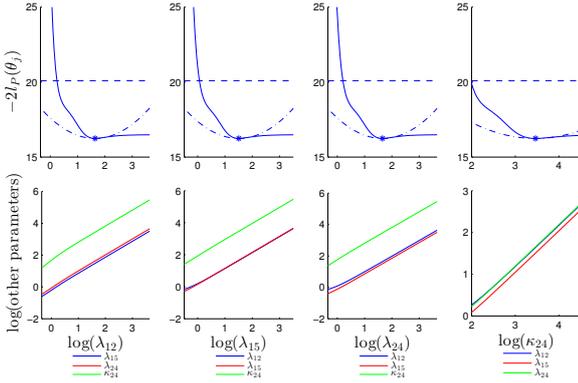


Figure 11: Profile likelihoods for Model III fit to the GA₁₂ pathway data. Upper panels show the profile likelihood for each parameter (solid lines), the approximation to the likelihood entailed by the Gaussian-based confidence intervals (dot-dash), the estimated parameter value (asterisk) and the confidence interval threshold for $\alpha = 0.95$ (dashed lines). Lower panels show how the other parameter values vary along these likelihood profiles.

Parameter	Estimated value	Confidence interval PL	Confidence interval Gaussian
$\log(\lambda_{53})$	-0.54	[-0.62, -0.45]	[-0.65, -0.44]
$\log(\lambda_{44})$	-1.03	[-1.16, -0.89]	[-1.21, -0.86]
$\log(\lambda_{19})$	-0.72	[-0.86, -0.55]	[-0.92, -0.52]
$\log(\kappa_{19})$	1.08	[0.84, 1.31]	[0.78, 1.37]

Table 2: Estimated parameters and their profile-likelihood (PL) and Gaussian-based confidence intervals (with $\alpha = 0.95$) for Model III and the GA₅₃ pathway. These values correspond with the results shown in Fig. 10.

$O(1)$ parameters as $\kappa_{19} \rightarrow \infty$. Letting $\delta = 1/\kappa_{19}$, Model III can be rewritten as

$$\frac{d[\text{GA}_{53}]}{dt^*} = \frac{-\tilde{\lambda}_{53}[\text{GA}_{53}]}{\delta + [\text{GA}_{19}]}, \quad (27a)$$

$$\frac{d[\text{GA}_{44}]}{dt^*} = \frac{\tilde{\lambda}_{53}[\text{GA}_{53}] - \tilde{\lambda}_{44}[\text{GA}_{44}]}{\delta + [\text{GA}_{19}]}, \quad (27b)$$

$$\frac{d[\text{GA}_{19}]}{dt^*} = \frac{\tilde{\lambda}_{44}[\text{GA}_{44}] - \tilde{\lambda}_{19}[\text{GA}_{19}]}{\delta + [\text{GA}_{19}]}, \quad (27c)$$

$$\frac{d[\text{GA}_{20}]}{dt^*} = \frac{\tilde{\lambda}_{19}[\text{GA}_{19}]}{\delta + [\text{GA}_{19}]}, \quad (27d)$$

and

$$[\text{GA}_{20\text{ox}}] = \frac{\delta}{\delta + [\text{GA}_{19}]^*}. \quad (28)$$

We now consider the limit as $\delta \rightarrow 0$. In this case, the leading-order equations depend on the magnitude of $[\text{GA}_{19}]$ compared to δ , therefore to derive the governing equations, we consider a series of distinct timescales in which different terms are present at leading order (see [10, 19] for examples where similar approaches have been used, and [20] for a textbook description). As detailed below, five timescales emerge from the asymptotic analysis; to validate this analysis, Figs. 12 and 13 compare the asymptotic solutions to numerical simulations of Model III, (27), with $\delta = 0.01$.

3.4.1. Matched asymptotic analysis to evaluate the leading-order dynamics in the limit $\delta \rightarrow 0$

Timescale 1, $t^ = O(\delta^2)$:* At the beginning of the experiment, only GA₅₃ and GA_{20ox} are present. Therefore at early times $[\text{GA}_{19}] \ll \delta$, the amount of enzyme bound with $[\text{GA}_{19}]$ is asymptotically small and the concentration of free enzyme $[\text{GA}_{20\text{ox}}] = 1$ (see (28)). We rescale

$$[\text{GA}_{44}] = \delta[\text{GA}_{44}]^\dagger, \quad [\text{GA}_{19}] = \delta^2[\text{GA}_{19}]^\dagger, \quad (29)$$

$$[\text{GA}_{20}] = \delta^3[\text{GA}_{20}]^\dagger,$$

(using daggers to denote the rescaled variables) and find that the concentrations evolve on a fast time scale, T , such that $t^* = \delta^2 T$. At leading order, equations (27) become

$$\frac{d[\text{GA}_{53}]}{dT} = 0, \quad \frac{d[\text{GA}_{44}]^\dagger}{dT} = \tilde{\lambda}_{53}[\text{GA}_{53}],$$

$$\frac{d[\text{GA}_{19}]^\dagger}{dT} = \tilde{\lambda}_{44}[\text{GA}_{44}]^\dagger, \quad \frac{d[\text{GA}_{20}]^\dagger}{dT} = \tilde{\lambda}_{19}[\text{GA}_{19}]^\dagger, \quad (30)$$

which have solution

$$[\text{GA}_{53}] = 1, \quad [\text{GA}_{44}]^\dagger = \tilde{\lambda}_{53}T,$$

$$[\text{GA}_{19}]^\dagger = \frac{\tilde{\lambda}_{53}\tilde{\lambda}_{44}T^2}{2}, \quad [\text{GA}_{20}]^\dagger = \frac{\tilde{\lambda}_{53}\tilde{\lambda}_{44}\tilde{\lambda}_{19}T^3}{6}, \quad (31)$$

(satisfying initial conditions (24)). As shown in Fig. 12a, these analytical expressions agree with the dynamics of Model III, (27), at early times.

Timescale 2, $t^* = O(\delta^{3/2})$: As the GA_{19} concentration increases, we move to a different asymptotic regime in which $[\text{GA}_{19}] = O(\delta)$ and significant amounts of both free and GA_{19} -bound enzyme are present. To ensure matching between consecutive timescales, we require the long-time behaviour of each variable to match to the short-term dynamics on the subsequent timescale. Thus, considering the long-time dynamics of Timescale 1, if T is scaled with δ^α as $T \rightarrow \infty$ (for unknown α) then from (31), we find that $[\text{GA}_{44}]^\ddagger$ is scaled with δ^α , $[\text{GA}_{19}]^\ddagger$ with $\delta^{2\alpha}$ and $[\text{GA}_{20}]^\ddagger$ with $\delta^{3\alpha}$. We find a suitable balance setting $\alpha = -1/2$. Thus, the rescaled time, \hat{t} , is $t^* = \delta^{3/2}\hat{t}$ and the relevant scalings are

$$\begin{aligned} [\text{GA}_{44}] &= \delta^{1/2}[\text{GA}_{44}]^\ddagger, & [\text{GA}_{19}] &= \delta[\text{GA}_{19}]^\ddagger, \\ [\text{GA}_{20}] &= \delta^{3/2}[\text{GA}_{20}]^\ddagger. \end{aligned} \quad (32)$$

The resulting rescaled problem reads

$$\begin{aligned} \frac{d[\text{GA}_{53}]}{d\hat{t}} &= 0, \\ \frac{d[\text{GA}_{44}]^\ddagger}{d\hat{t}} &= \frac{\tilde{\lambda}_{53}[\text{GA}_{53}]}{1 + [\text{GA}_{19}]^\ddagger}, \\ \frac{d[\text{GA}_{19}]^\ddagger}{d\hat{t}} &= \frac{\tilde{\lambda}_{44}[\text{GA}_{44}]^\ddagger}{1 + [\text{GA}_{19}]^\ddagger}, \\ \frac{d[\text{GA}_{20}]^\ddagger}{d\hat{t}} &= \frac{\tilde{\lambda}_{19}[\text{GA}_{19}]^\ddagger}{1 + [\text{GA}_{19}]^\ddagger}. \end{aligned} \quad (33)$$

Applying initial conditions $[\text{GA}_{53}]^\ddagger = 1$, $[\text{GA}_{44}]^\ddagger = [\text{GA}_{19}]^\ddagger = [\text{GA}_{20}]^\ddagger = 0$, numerical solutions reveal that these equations, (33), faithfully reproduce the dynamics of Model III at early times (Fig. 12b).

Given we do not have an analytical solution on this time scale, we elucidate the long-time behaviour of the dynamics by supposing each variable takes the form $\alpha \hat{t}^\eta$ as $\hat{t} \rightarrow \infty$ (where α and η are unknown constants for each variable); substituting into the governing equations (33), we find that

$$\begin{aligned} [\text{GA}_{44}]^\ddagger &= \left(\frac{6\tilde{\lambda}_{53}^2}{\tilde{\lambda}_{44}}\right)^{1/3} \hat{t}^{1/3}, \\ [\text{GA}_{19}]^\ddagger &= \left(\frac{9\tilde{\lambda}_{53}\tilde{\lambda}_{44}}{2}\right)^{1/3} \hat{t}^{2/3}, \\ [\text{GA}_{20}]^\ddagger &= \tilde{\lambda}_{19}\hat{t} \quad \text{as } \hat{t} \rightarrow \infty. \end{aligned} \quad (34)$$

These long-time dynamics are in good agreement with numerical solutions of the governing equations on this time scale, (33) (Fig. 12b).

Timescale 3, $t^* = O(1)$: Once $[\text{GA}_{19}] = O(1)$, the majority of the enzyme is bound with GA_{19} and the concentration of free enzyme $[\text{GA}_{20\text{ox}}] = O(\delta)$, see (28). Rescaling the long-time dynamics on the previous timescale, (34), we see that on a $t^* = O(1)$ timescale the four GA concentrations are $O(1)$; thus, the leading-order dynamics are

governed by

$$\frac{d[\text{GA}_{53}]}{dt^*} = \frac{-\tilde{\lambda}_{53}[\text{GA}_{53}]}{[\text{GA}_{19}]}, \quad (35a)$$

$$\frac{d[\text{GA}_{44}]}{dt^*} = \frac{\tilde{\lambda}_{53}[\text{GA}_{53}] - \tilde{\lambda}_{44}[\text{GA}_{44}]}{[\text{GA}_{19}]}, \quad (35b)$$

$$\frac{d[\text{GA}_{19}]}{dt^*} = \frac{\tilde{\lambda}_{44}[\text{GA}_{44}] - \tilde{\lambda}_{19}[\text{GA}_{19}]}{[\text{GA}_{19}]}, \quad (35c)$$

$$\frac{d[\text{GA}_{20}]}{dt^*} = \tilde{\lambda}_{19}. \quad (35d)$$

From (35d), we obtain an explicit solution for the GA_{20} concentration:

$$[\text{GA}_{20}] = \tilde{\lambda}_{19}t^*. \quad (36)$$

As the enzyme and complex concentrations sum to one at all times, this explicit solution (36) reveals that the Timescale 3 equations (35) are only valid for $t^* < 1/\tilde{\lambda}_{19}$. For this time period, $t^* \in [0, 1/\tilde{\lambda}_{19}]$, we see that numerical solutions of (35) faithfully reproduce the solutions of Model III, (27) (Fig. 12c).

Matching to the subsequent time scale requires the dynamics in the limit $t^* \rightarrow 1/\tilde{\lambda}_{19}$. Supposing each variable takes the form $\alpha(1/\tilde{\lambda}_{19} - t^*)^\eta$, we find suitable balances in the governing equations, (35), with

$$\begin{aligned} [\text{GA}_{53}] &= \frac{(\tilde{\lambda}_{44} - \tilde{\lambda}_{53})(\tilde{\lambda}_{19} - \tilde{\lambda}_{53})}{\tilde{\lambda}_{44}} \left(\frac{1}{\tilde{\lambda}_{19}} - t^*\right), \\ [\text{GA}_{44}] &= \frac{\tilde{\lambda}_{53}(\tilde{\lambda}_{19} - \tilde{\lambda}_{53})}{\tilde{\lambda}_{44}} \left(\frac{1}{\tilde{\lambda}_{19}} - t^*\right), \\ [\text{GA}_{19}] &= \tilde{\lambda}_{53} \left(\frac{1}{\tilde{\lambda}_{19}} - t^*\right), \quad \text{as } \left(\frac{1}{\tilde{\lambda}_{19}} - t^*\right) \rightarrow 0. \end{aligned} \quad (37)$$

These analytical solutions accurately represent the Timescale 3 dynamics in the limit as $t^* \rightarrow 1/\tilde{\lambda}_{19}$ (Fig. 13).

Timescale 4: The Timescale 3 equations show that as $t^* \rightarrow 1/\tilde{\lambda}_{19}$, all the GA_{53} is converted to GA_{20} , $[\text{GA}_{20}] \rightarrow 1$ and the concentrations of GA_{53} , GA_{44} and GA_{19} become small. Matching with the previous timescale, (37), we rescale

$$\begin{aligned} t^* &= \frac{1}{\tilde{\lambda}_{19}} + \delta\tau, & [\text{GA}_{53}] &= \delta[\text{GA}_{53}]^\diamond, \\ [\text{GA}_{44}] &= \delta[\text{GA}_{44}]^\diamond, & [\text{GA}_{19}] &= \delta[\text{GA}_{19}]^\diamond, \\ [\text{GA}_{20}] &= 1 - \delta[\text{GA}_{20}]^\diamond, \end{aligned} \quad (38)$$

and the governing equations become

$$\begin{aligned} \frac{d[\text{GA}_{53}]^\diamond}{d\tau} &= \frac{-\tilde{\lambda}_{53}[\text{GA}_{53}]^\diamond}{1 + [\text{GA}_{19}]^\diamond}, \\ \frac{d[\text{GA}_{44}]^\diamond}{d\tau} &= \frac{\tilde{\lambda}_{53}[\text{GA}_{53}]^\diamond - \tilde{\lambda}_{44}[\text{GA}_{44}]^\diamond}{1 + [\text{GA}_{19}]^\diamond}, \\ \frac{d[\text{GA}_{19}]^\diamond}{d\tau} &= \frac{\tilde{\lambda}_{44}[\text{GA}_{44}]^\diamond - \tilde{\lambda}_{19}[\text{GA}_{19}]^\diamond}{1 + [\text{GA}_{19}]^\diamond}, \\ \frac{d[\text{GA}_{20}]^\diamond}{d\tau} &= -\frac{\tilde{\lambda}_{19}[\text{GA}_{19}]^\diamond}{1 + [\text{GA}_{19}]^\diamond} \end{aligned} \quad (39)$$

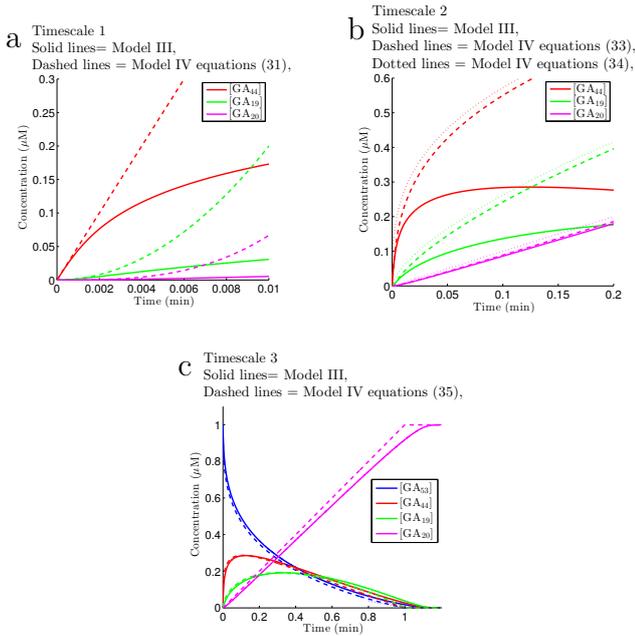


Figure 12: Comparison of predicted dynamics of Model IV for Timescales 1-3 with numerical solutions of Model III (27). In each panel, solutions of Model III are shown in solid lines and we prescribe $\tilde{\lambda}_{53} = 0.5$, $\tilde{\lambda}_{44} = 0.8$, $\tilde{\lambda}_{19} = 1$ and $\delta = 0.01$. (a) Analytical solutions for Timescale 1 (31) (dashed lines). (b) Numerical solutions of governing equations for Timescale 2 (33) (dashed lines), and analytical solutions for the long time behaviour of the dynamics on Timescale 2 (34) (dotted lines). (c) Numerical solutions of governing equations for Timescale 3 (35) (dashed lines), where initial conditions were applied at $t^* = 10^{-4}$ using the analytical solution of the long time behaviour of Timescale 2 (34).

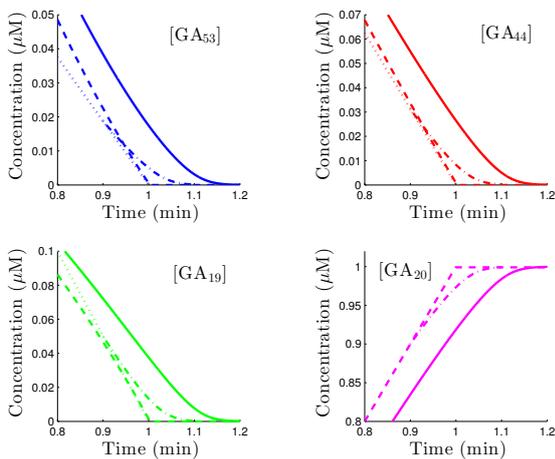


Figure 13: Comparison of predicted dynamics of Model IV for Timescales 3-4 with numerical solutions of Model III (27) for $\tilde{\lambda}_{53} = 0.5$, $\tilde{\lambda}_{44} = 0.8$, $\tilde{\lambda}_{19} = 1$ and $\delta = 0.01$. Each panel shows numerical solutions of Model III (27) (solid lines), numerical solutions of the governing equations for Timescale 3 (35) (dashed lines), analytical solutions of the Timescale 3 dynamics in the limit as $t^* \rightarrow 1/\lambda_{19}$ (37) (dotted lines) and numerical solutions of the governing equations for Timescale 4 (39) with initial conditions applied at $t^* = 0.9$ using the Timescale 3 solutions in the limit as $t^* \rightarrow 1/\lambda_{19}$ (dash-dot lines).

Parameter	Estimated value	Confidence interval, PL	Confidence interval, Gaussian
$\log(\tilde{\lambda}_{53})$	-2.11	$[-2.22, -2.00]$	$[-2.26, -1.95]$
$\log(\tilde{\lambda}_{44})$	-2.40	$[-2.56, -2.27]$	$[-2.60, -2.21]$
$\log(\tilde{\lambda}_{19})$	-2.02	$[-2.06, -1.98]$	$[-2.07, -1.96]$
$\log(\tilde{\lambda}_{12})$	-1.83	$[-1.91, -1.75]$	$[-1.94, -1.72]$
$\log(\tilde{\lambda}_{15})$	-1.96	$[-2.06, -1.86]$	$[-2.10, -1.82]$
$\log(\tilde{\lambda}_{24})$	-1.81	$[-1.84, -1.79]$	$[-1.84, -1.78]$

Table 3: Estimated parameters and their profile-likelihood (PL) and Gaussian-based confidence intervals (with $\alpha = 0.95$) for Model IV. These values correspond with the results shown in Fig. 14.

Solving the Timescale 4 governing equations, (39), numerically, we see good agreement with the numerical solutions of Model III (Fig. 13). As $\tau \rightarrow \infty$, the numerical solutions show that $[GA_{53}]^\diamond \rightarrow 0$, $[GA_{44}]^\diamond \rightarrow 0$, $[GA_{19}]^\diamond \rightarrow 0$ and $[GA_{20}]^\diamond \rightarrow 0$.

Timescale 5: It remains to characterise the dynamics for $t^* = O(1)$ and $t^* > 1/\tilde{\lambda}_{19}$. Considering the long-time dynamics of Timescale 4 (Fig. 13), the leading-order solution for Timescale 5 is simply given by

$$[GA_{53}] = [GA_{44}] = [GA_{19}] = 0, \quad [GA_{20}] = 1. \quad (40)$$

3.4.2. Fitting Model IV

The dominant ($t^* = O(1)$) dynamics are given by the asymptotic solutions in Timescales 3 and 5, which can be solved subject to prescribing an initial non-zero concentration of GA_{19} . For a given parameter set, we simulate the Timescale 3 equations for $t^* \in [0, 1/\tilde{\lambda}_{19}]$ (35), prescribing initial conditions using the long time dynamics of Timescale 2 (34), applied at $t^* = 10^{-4}$, and then prescribe the Timescale 5 equations for $t > 1/\tilde{\lambda}_{19}$ (40). Simulating Model IV in this way, we fit the three model parameters to the experimental data (Figs. 4 and 5). The profile likelihoods (Fig. 14) show that for Model IV each parameter is identifiable using the data, with narrow Gaussian and profile-likelihood-based confidence intervals that are in good agreement (Table 3).

As expected given the parameter estimates for Model III, the value of $S(\hat{\theta})$ and the AICc (Table 1) show that the Model IV approximation ($\kappa_{19} \rightarrow \infty$) captures the observed dynamics of the GA_{12} pathway, but is not appropriate for the GA_{53} pathway. Thus, our study recommends that Model IV is the appropriate choice for the GA_{12} pathway, whereas one would choose Model III for the GA_{53} pathway.

4. Discussion

We have described how parameter estimation can motivate asymptotic reduction of a network model and lead to a simpler model that represents well the observed dynamics. We demonstrated the use of these methods to analyse a recently published model of GA biosynthesis [23].

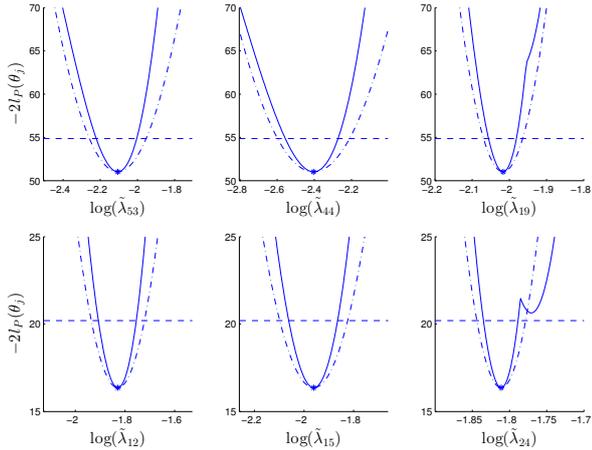


Figure 14: Profile likelihood for Model IV fit to the GA₅₃ pathway data (upper panels) and the GA₁₂ pathway data (lower panels). Each panel show the profile likelihood for the parameter (solid line), the approximation to the likelihood entailed by the Gaussian-based confidence intervals (dot-dash line), the estimated parameter value (asterisk) and the confidence interval threshold for $\alpha = 0.95$ (dashed lines).

When fitting sloppy systems biology models to experimental data, calculating the profile likelihood provides helpful insight into the nature of parameter space in the neighbourhood of the best estimated parameter set. The profile likelihood reveals when taking a parameter to be increasingly small or large is consistent with the model fitting well the observed dynamics, or where the model dynamics are insensitive to particular parameter values. This, together with plots of how the other fitted parameters change as the prescribed one is varied, helps identify an appropriate asymptotic model reduction. The simpler model can then be fitted to the data to assess whether it is an appropriate approximation. As asymptotically reduced models depend on groupings of the original model parameters, the reduced model has fewer parameters and is a more parsimonious representation of the observed dynamics. Asymptotic model reduction often involves careful consideration of distinct time scales (for instance, as in the reduction of Model III to Model IV), and therefore automating the model-reduction step is unlikely to be possible in many cases.

The profile likelihood also enables calculation of confidence intervals for each parameter. These tend to be preferable in practise compared with confidence intervals based on the approximate Gaussian distribution of $\hat{\theta}$, especially so for strongly non-linear models where the Gaussian approximation is poor (see e.g. Fig. 11). However, we have found that in the final reduced models, where parameters are identifiable, the two types of confidence interval do agree quite closely (e.g. Fig. 10).

We have focussed here on using the profile likelihood for parameter inference and model reduction. Other approaches for considering dependence between model pa-

rameters with a view to model reduction include investigating the eigenvectors of \mathbf{H} that correspond to the smallest eigenvalues [22], or adopting a Bayesian paradigm [13] and investigating structure amongst posterior parameter samples. In our view, though, the profile likelihood offers the most direct way to understand the effect of varying the value of individual parameters, and hence identifying good candidate model reductions.

Our study focussed on analysing a model of the GA biosynthesis pathway. Given the importance of GA for plant development, a GA pathway model is likely to feature in future plant models and the existing GA pathway model [23] has already featured in a multiscale model of plant growth regulation [3] and a study on hormone cross-talk [1]. In assessing the required level of complexity for capturing the dynamics of GA biosynthesis, our analysis will simplify future plant models. We recommend that Model III with the parameter estimates given in Table 2 could be used to simulate the conversion of GA₅₃ to GA₂₀ whereas Model IV with the parameter estimates given in Table 3 captures the conversion of GA₁₂ to GA₉.

For both pathways, the analysis revealed the typical enzyme concentrations to be significantly smaller than the typical GA concentrations; as a result, the dynamics can be captured by Model II in which complex and enzyme concentrations equilibrate rapidly and are at a quasi-steady state after an initial transient time scale. The concentrations of complexes or free enzyme can be approximated as simply a function of the GA concentrations (19, 21). Fitting the parameter groupings under this Model II approximation further revealed that for GA₅₃ and GA₄₄ (and equivalently GA₁₂ and GA₁₅) the ratios between the typical complex association and dissociation rates (κ_{53} and κ_{44} respectively) are small, resulting in small GA₅₃-GA_{20ox} and GA₄₄-GA_{20ox} complex concentrations. We note that these parameter groupings, κ_{53} and κ_{44} , arose through the first asymptotic model reduction to Model II and therefore the second asymptotic limit is only revealed once Model II is fitted to the data. We find a further approximation is applicable to the GA₁₂ pathway; here, fitting Model III to the data showed that binding of GA₂₄ to GA_{20ox} to form GA₂₄-GA_{20ox} occurs much more rapidly than GA₂₄-GA_{20ox} dissociation, resulting in the majority of the enzyme being bound in a complex with GA₂₄. Thus, the modelling suggests that dissociation of GA₂₄-GA_{20ox} is a major limiting step in the production of bioactive GA₄.

Acknowledgements This work was supported by the Leverhulme Trust (Early Career Fellowship to Dr Leah Band, ECF-2012-681) and the BBSRC (grant number BB/M012837/1).

Appendix A. Numerical optimisation

The likelihood function, $\ell(\boldsymbol{\theta})$, for nonlinear models often has many local optima, so gradient-based numerical optimisation algorithms tend to get trapped in local optima and fail to find the global optimum, $\hat{\boldsymbol{\theta}}$. A simple strategy to overcome this is to repeat the optimisation many times using different widely spaced starting values for $\boldsymbol{\theta}$. Following Raue *et al.* [27], we choose starting values using Latin hypercube sampling, which involves specifying a range for each parameter, dividing it into p equally sized intervals and then constructing p parameter sets such that for each parameter a value is selected from each interval once and only once. For each optimisation, we use Matlab function *lsqnonlin*'s gradient-based trust-region-reflective algorithm, which requires the derivative

$$\frac{\partial \ell}{\partial \boldsymbol{\theta}} = \frac{1}{\sigma^2} \sum_{j=1}^m \sum_{i=1}^{n_j} (y_{ij}^{\text{Data}} - y_j(t_{ji})) \frac{\partial y_j(t_{ji})}{\partial \boldsymbol{\theta}}.$$

in which $\partial \mathbf{y} / \partial \boldsymbol{\theta}$ can be computed by solving numerically the ‘‘sensitivity equations’’

$$\frac{d}{dt} \left(\frac{d\mathbf{y}}{d\boldsymbol{\theta}} \right) = \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}} + \frac{\partial \mathbf{f}}{\partial \mathbf{y}} \frac{d\mathbf{y}}{dt} \quad \text{with} \quad \left. \frac{d\mathbf{y}}{d\boldsymbol{\theta}} \right|_{t=0} = 0.$$

For the models and data in this paper, this method consistently identifies the optimum $\hat{\boldsymbol{\theta}}$.

A useful by-product of computing $\partial \mathbf{y} / \partial \boldsymbol{\theta}$ is that it can be used to compute a cheap approximation to \mathbf{H} [4],

$$\mathbf{H} \approx - \sum_{j=1}^m \sum_{i=1}^{n_j} \frac{1}{\sigma^2} \frac{\partial y_j(t_{ji})}{\partial \boldsymbol{\theta}} \frac{\partial y_j(t_{ji})}{\partial \boldsymbol{\theta}^\top},$$

which we used for calculations of \mathbf{H} throughout the paper.

Appendix B. Information criterion for model selection

A common way to select amongst candidate models is via the Akaike Information Criterion, or the ‘‘corrected version’’ (AICc) defined as [5]

$$\text{AICc} = 2p - 2\ell(\hat{\boldsymbol{\theta}}) + \frac{2pN}{N - p - 1}. \quad (\text{B.1})$$

This information criterion trades off the quality of model fit, measured by $\ell(\hat{\boldsymbol{\theta}})$, with the number of parameters involved: when computed for multiple models, the model to be favoured is the one with smallest AICc. We compute the AICc values for each of the models we consider.

References

[1] Allen, H.R., Ptashnyk, M., 2017 Mathematical modelling and analysis of the Brassinosteroid and Gibberellin signalling pathways and their interactions. *J. Theor. Biol.* arXiv:1703.07584v2 [q-bio.SC]

[2] Appleford, N.E.J., Evans, D.J., Lenton, J.R., Gaskin, P., Croker, S.J., Devos, K.M., Phillips, A.L., Hedden, P., 2006. Function and transcript analysis of gibberellin-biosynthetic enzymes in wheat. *Planta* 223, 568-582.

[3] Band, L.R., Ubeda-Tomas, S., Dyson, R.J., Middleton, A.M., Hodgman, T.C., Owen, M.R., Jensen, O.E., Bennett, M.J., and King, J.R. (2012). Growth-induced hormone dilution can explain the dynamics of plant root cell elongation. *Proc. Natl Acad. Sci. USA* 109:7577-7582.

[4] Brown, K.S., Sethna, J.P., 2003. Statistical mechanical approaches to models with many poorly known parameters. *Phys. Rev. E* 68, 021904.

[5] Burnham, K.P., Anderson, D.R., 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd Edition Springer-Verlag, New York, USA.

[6] Chen, W.W., Niepel, M., Sorger, P.K., 2010. Classic and contemporary approaches to modeling biochemical reactions. *Genes Dev.* 24, 1861-1875.

[7] Chew, Y.H., Wenden, B., Flis, A., Mengin, V., Taylor, J., Davey, C.L., Tindal, C., Thomas, H., Ougham, H.J., de Reffye, P., Stitt, M., Williams, M., Muetzelfeldt, R., Halliday, K.J., Millar, A.J., 2014. Multiscale digital Arabidopsis predicts individual organ and whole-organism growth. *Proc. Natl Acad. Sci. USA*. 111:E4127-36.

[8] Claeys, H., De Bodt, S., Inzé, D., 2014. Gibberellins and DELLAs: central nodes in growth regulatory networks. *Trends Plant Sci.* 19, 231-239.

[9] Cox, D.R., 2006. Principles of statistical inference. Cambridge University Press.

[10] Dunster, J.L., King, J.R., 2017. Mathematical modelling of thrombin generation: asymptotic analysis and pathway characterisation. *IMA Journal of Applied Mathematics*, 82, 60-96.

[11] Fleet, C.M., Sun, T.P., 2005. A DELLAcate balance: the role of gibberellin in plant morphogenesis. *Current Opin. Plant Biol.* 8, 77-85.

[12] Fleet, C.M., Yamaguchi, S., Hanada, A., Kawaide, H., David, C.J., Kamiya, Y., Sun, T.P., 2003. Overexpression of AtCPS and AtKS in Arabidopsis confers increased ent-kaurene production but no increase in bioactive gibberellins. *Plant Physiol.* 132, 830-839.

[13] Girolami, M., 2008. Bayesian inference for differential equations. *Theor. Comp. Sci.* 408, 4 - 16.

[14] Gutenkunst, R.N., Waterfall, J.J., Casey, F.P., Brown, K.S., Myers, C.R., Sethna, J.P., 2007. Universally sloppy parameter sensitivities in systems biology models. *Plos Comp. Biol.* 3, e189.

[15] Hedden, P., Phillips, A.L., 2000. Gibberellin metabolism: new insights revealed by the genes. *Trends Plant Sci.* 5, 523-530.

[16] Hedden, P., 2003. The genes of the Green Revolution. *Trends Genet.* 19(1), 5-9.

[17] Hedden, P., Thomas, S.G., 2012. Gibberellin biosynthesis and its regulation. *Biochem. J.* 444(1), 11-25.

[18] Hiramatsu, T., King, R.W., Helliwell, C.A., Koshiok, M., 2005. The involvement of Gibberellin 20-Oxidase genes in phytochrome-regulated petiole elongation of Arabidopsis. *Plant Physiol.* 138, 1106-1116.

[19] Jabbari, S.J., Cartman, S.T., King, J.R., 2015. Mathematical modelling reveals properties of TcdC required for it to be a negative regulator of toxin production in *Clostridium difficile*. *J. Math. Biol.* 70, 773-804.

[20] Kevorkian J., Cole .J.D., 1981. Perturbation methods in applied mathematics. Springer, New York

[21] Lindgren, B., 1993. Statistical Theory. CRC Press.

[22] Miao, H., Xia, X., Perelson, A. S., Wu, H. (2011). On identifiability of nonlinear ODE models and applications in viral dynamics. *SIAM review*, 53(1), 3-39.

[23] Middleton, A.M., Ubeda-Tomas, S., Griffiths, J., Holman, T., Hedden, P., Thomas, S.G., Phillips, A.L., Holdsworth, M.J., Bennett, M.J., King, J.R., Owen, M.R., 2012. Mathematical modelling elucidates the role of transcriptional feedback in gibberellin signalling. *Proc. Natl Acad. Sci. USA* 109, 7571-7576.

- [24] Murray, J.D., 2002. *Mathematical Biology, I: An introduction*. 3rd Edition Springer.
- [25] Nelissen, H., Rymen, B., Jikumaru, Y., Demuyne, K., Van Lijsebettens, M., Kamiya, Y., Inzé, D., Beemster, G.T.S., 2012. A local maximum in gibberellin levels regulates Maize leaf growth by spatial control of cell division. *Curr. Biol.* 22, 1183-1187.
- [26] Peng, J., Richards, D.E., Hartley, N.M., Murphy, G.P., Devos, K.M., Flintham, J.E., Beales, J., Fish, L.J., Worland, A.J., Pelica, F., Sudhakar, D., Christou, P., Snape, J.W., Gale, M.D., Harberd, N.P., 1999. Green revolution genes encode mutant gibberellin response modulators. *Nature* 400, 256-261.
- [27] Raue, A., Schilling, M., Bachmann, J., Matteson, A., Schelke, M., Kaschek, D., Hug, S., Kreutz, C., Harms, B.D., Theis, F.J., Klingmüller, U., Timmer, J., 2013. Lessons learned from quantitative dynamical modeling in systems biology. *PloS one*, 8, e74335.
- [28] Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., Timmer, J., 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* 25, 1923-1929.
- [29] Seber, G. A. F., & Wild, C. J., 2003. *Nonlinear regression*. John Wiley & Sons.
- [30] Tönsing, C., Timmer, J., Kreutz, C. 2014. Cause and cure of sloppiness in ordinary differential equation models. *Phys. Rev. E* 90, 023303.
- [31] Wilks, S. S. 1938. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Stat.* 9: 6062.
- [32] Venzon, D.J., Moolgavkar, S.H. 1988. A method for computing profile-likelihood-based confidence intervals. *J. Roy. Stat. Soc. C-App*, 37: 87-94.
- [33] Yamaguchi, S., 2008. Gibberellin metabolism and its regulation. *Annu. Rev. Plant Biol.*, 59, 225-251.