



University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Thomas R.B. Grandjean, Mike J. Chappell, James T.W. Yates, Kevin Jones, Gemma Wood, Tanya Coleman

Article Title: Compartmental modelling of the pharmacokinetics of a breast cancer resistance protein

Year of publication: 2011

Link to published article:

<http://dx.doi.org/10.1016/j.cmpb.2010.08.018>

Publisher statement: None (Pre-print)

Compartmental Modelling of the Pharmacokinetics of a Breast Cancer Resistance Protein

Thomas R. B. Grandjean*, Mike J. Chappell*, James T. W. Yates**,
Kevin Jones**, Gemma Wood** and Tanya Coleman**

** School of Engineering, University of Warwick, Coventry, CV4 7AL, UK*

(Tel: +44 (0)24765 24309; e-mail: M.J.Chappell@warwick.ac.uk)

***AstraZeneca R&D, Alderley Park, Cheshire, UK*

(Tel: +44 (0)1625 512996; e-mail: James.Yates@AstraZeneca.com)

Abstract: A mathematical model for the pharmacokinetics of Hoechst 33342 following administration into a culture medium containing a population of transfected cells (HEK293 hBCRP) with a potent Breast Cancer Resistance Protein inhibitor, Fumitremorgin C (FTC), present is described. This non-linear compartmental model has seven macroscopic sub-units, with fourteen rate parameters. It describes the relationship between the concentration of Hoescht 33342 and FTC, initially spiked in the medium, and the observed change in fluorescence due to Hoescht 33342 binding to DNA. Structural identifiability analysis has been performed using two methods, one based on the similarity transformation/exhaustive modelling approach and the other based on the differential algebra approach. The analyses demonstrated that all models derived are uniquely identifiable for the experiments/observations available. A kinetic modelling software package, namely FACSIMILE (MPCA Software, UK), was used for parameter fitting and to obtain numerical solutions for the system equations. Model fits gave very good agreement with in-vitro data provided by AstraZeneca across a variety of experimental scenarios.

Keywords: Breast Cancer Resistance Protein, BCRP, Compartmental models, Drug kinetics.

1. INTRODUCTION

In this paper a mathematical model for the pharmacokinetics of Hoechst 33342 following administration into a culture medium containing a population of transfected cells (HEK293 hBCRP) is described.

Cancer cells develop mechanisms that allow them to resist the action of anti-cancer compounds. This can reduce the exposure of the diseased tissue and so have consequences for the efficacy of a compound. One such important mechanism is efflux transport by the Breast Cancer Resistance Protein (BCRP) [1]. It is therefore important to know whether a novel drug is a substrate for BCRP.

This article reports the modelling of the kinetics of an assay that uses transfected cells that express BCRP. The assay indirectly measures the binding potential of a drug or similar molecule to BCRP by observing the effects on the kinetics of Hoechst 33342 [2], a BCRP substrate. When Hoechst 33342 binds to DNA the resulting complex fluoresces. This allows the relative levels of Hoechst 33342 bound to DNA to be measured under different experimental conditions. Mathematical modelling of *in vitro* pharmacokinetic assays has proven useful elsewhere [3,4].

With this experimental set up it is challenging to measure the binding affinity of a drug for BCRP because the only known quantities in the system are the initial extracellular concentrations of Hoechst 33342 and the drug of interest, as well as a fluorescence time series.

This paper describes the modelling of this system along with parameter estimates based upon experimentally obtained data. The intention here is to derive a compartmental model to characterise substrate binding to DNA and in addition account for the effect of transportation of the substrate out of the cell. Such a model, once validated, should permit the prediction of the dosage levels required in order to achieve the levels of absorption desired once bound to DNA.

2. THE MODEL

In the model, compartments represent different parts of the cell. Based upon what is known of the system a seven compartment model illustrated in *Fig. 1* is used initially to describe the flow of the substrate and inhibitor within and out of the cell.

Insert Figure 1

The extracellular Hoechst 33342 (SO) diffuses into the cell (SI) and may then bind to the DNA in the nucleus (NS) resulting in fluorescence. Hoechst 33342 is also transported out of the cell by a BCRP transporter (TS).

Similarly extracellular inhibitor (IO) diffuses into the cell and is also transported out by the BCRP transporter (TI).

The seven compartments and the inter-compartmental rate transfers are summarised in *Table 1* (square brackets denoting concentration, all quantities are in term of relative fluorescence units – RFU).

Insert Table 1

2.1 System Equations

The corresponding set of nonlinear ordinary differential equations characterising the proposed model was derived using classical mass-balance principles is given by the following:

$$\frac{d[SO]}{dt} = -k_{S_{in}}[SO] + k_{S_{out}}[SI] - k_{S_{tran}}[TS] \quad (1)$$

$$\frac{d[SI]}{dt} = k_{S_{in}}[SO] - k_{S_{out}}[SI] - k_S^+[SI](T_0 - [TS] - [TI]) + k_S^-[TS] - k_N^+[SI]^n(N_0 - [NS]) + k_N^-[NS] \quad (2)$$

$$\frac{d[IO]}{dt} = -k_{I_{in}}[IO] + k_{I_{out}}[II] - k_{I_{tran}}[TI] \quad (3)$$

$$\frac{d[II]}{dt} = k_{I_{in}}[IO] - k_{I_{out}}[II] - k_I^+[II](T_0 - [TI] - [TS]) + k_I^-[TI] \quad (4)$$

$$\frac{d[TS]}{dt} = k_S^+[SI](T_0 - [TS] - [TI]) - (k_S^- + k_{S_{tran}})[TS] \quad (5)$$

$$\frac{d[TI]}{dt} = k_I^+[II](T_0 - [TI] - [TS]) - (k_I^- + k_{I_{tran}})[TI] \quad (6)$$

$$\frac{d[NS]}{dt} = k_N^+[SI]^n(N_0 - [NS]) - k_N^-[NS] \quad (7)$$

Where T_0 is the total concentration of transporter molecules, N_0 is the number of binding sites on the DNA, and n is the order of the nucleus binding reaction, which represents the number of binding sites per molecule of DNA.

There is also an eighth equation, which defines the constant k , the observation gain, given by

$$y = \text{Fluorescence} = k \cdot [NS] \quad (8)$$

3. EXPERIMENTAL DATA

Data were gathered at AstraZeneca, Alderley Park, UK. The multi-well plate provided 96 experiments with varying initial amounts of marker compound (ranging between 0.5-10 μ Mol) and inhibitor levels (ranging between 0-100 μ Mol), see *Table 2* for more details.

Insert Table 2

Hoechst 33342 accumulation was measured using a Polarstar Optima fluorescence plate reader. Excitation filter = 355nm Emission filter = 460-10nm, fluorescence was measured for 120-135 cycles, 1 cycle = 1 complete 96-well read, 30secs, Gain = 1200. Fluorescence data were captured by the Optima software for analysis.

The measured outputs are in relative fluorescence units (RFU). The inhibitor was added at time $t=0$ s (to compartment [I0]) and the compound marker is added five minutes later at time $t=300$ s (compartment [S0]).

The compound marker was added later to provide a measurement of the residual background noise fluorescence present during the experiments. The average background RFU for the first five minutes is subtracted individually from each time series (to allow for potential inhibitor independence) and the data are time shifted to begin at time $t=300$ s. A sample plot of the data is shown in *Fig. 2*.

Insert Figure 2

3.1 Reduced Model

The data provided lent themselves naturally to only modelling the marker compound as some time series were collected with no inhibitor present. The inhibitor compartments and corresponding rate constants were therefore initially set to zero, reducing the system to a four compartment system with ten parameters, illustrated in *Fig. 3*.

Insert Figure 3

The corresponding set of nonlinear ordinary differential equations characterising the reduced model with no inhibitor was derived using classical mass-balance principles and is given by:

$$\frac{d[S0]}{dt} = -k_{S_{in}}[S0] + k_{S_{out}}[SI] - k_{S_{tran}}[TS] \quad (9)$$

$$\frac{d[SI]}{dt} = k_{in}[SO] - k_{out}[SI] - k_{\bar{S}}^+[SI](T_0 - [TS]) + k_{\bar{S}}^-[TS] - k_N^+[SI]^n(N_0 - [NS]) + k_N^-[NS] \quad (10)$$

$$\frac{d[TS]}{dt} = k_{\bar{S}}^+[SI](T_0 - [TS]) - (k_{\bar{S}}^- + k_{trans})[TS] \quad (11)$$

$$\frac{d[NS]}{dt} = k_N^+[SI]^n(N_0 - [NS]) - k_N^-[NS] \quad (12)$$

$$y = \text{Fluorescence} = k \cdot [NS] \quad (13)$$

It was assumed that any kinetic parameters estimated for this sub-model would then also be applicable for use in the full model with inhibitor present as the rate constants concerned should not be affected by the presence of the inhibitor.

4. STEADY STATE ANALYSIS

From the data, we can observe that the fluorescence is not reaching steady state values in the time span of the experiments, see *Fig. 2*. Albeit perhaps for the compound concentration of 10 μ Mol; the graph may be suggesting that the binding is approaching a steady state value. Although the purpose of the modelling is to investigate the transient behaviour, a steady state analysis was performed so that it can potentially be used at a later stage to validate the model. It will identify the levels at which each compartment eventually settles and can be a useful method to obtain fundamental information on the system, the basic relationships between the compartments and for initial guesses for parameter estimation for subsequent fitting (i.e. saturation levels). Steady state analysis is performed by setting all the derivatives in the system equations to zero and solving for each compartment variable. Due to the complex non-linear nature of the equations, this was performed using symbolic mathematical packages capable of solving simultaneous equations, namely Mathematica and Maple, which both yielded the same solutions.

4.1 Reduced Model

The steady state analysis solution for the reduced model of form (9) – (13) is shown below.

$$[SO] = \frac{(k_{S_{out}}(k_{S^-} + k_{S^+}[SI]) + k_{S_{tran}}) + k_{S^+}T_0k_{S_{tran}} \cdot [SI]}{k_{S_{in}}(k_{S^+}[SI] + k_{S^-} + k_{S_{tran}})} \quad (14)$$

$$[SI] = [SI] \quad (15)$$

$$[TS] = \frac{k_{S^+}T_0[SI]}{k_{S^+}[SI] + k_{S^-} + k_{S_{tran}}} \quad (16)$$

$$[NS] = \frac{N_0k_{N^+}[SI]^n}{k_{N^-} + k_{N^+}[SI]^n} \quad (17)$$

$$\quad (17)$$

Equation 15 indicates that the equations under-determine the system and there is effectively one degree of freedom (DOF). For a given value of $[SI]$, the other three compartment concentrations can be computed.

4.2 Full Model

The steady state analysis solution for the full model of from (1) – (8) is shown in *Equations 18-24* below.

$$[SO] = \frac{k_{S_{out}}[SI] + k_{S_{tran}}[TS]}{k_{S_{in}}} \quad (18)$$

$$[SI] = [SI] \quad (19)$$

$$[IO] = \frac{-k_{I^+}k_{I_{tran}}(k_{S^-}^2 + 2k_{S^-} + k_{S_{tran}}^2)[TS]}{k_{I_{in}}k_{S^+}k_{I^+}(k_{S^-} + k_{S_{tran}})[SI]} - \frac{k_{S^+}^2k_{I_{out}}(k_{S^-} + k_{S_{tran}})(k_{I_{tran}} + [SI])}{k_{I_{in}}k_{S^+}k_{I^+}(k_{S^-} + k_{S_{tran}})[TS]} \\ - \frac{k_{S^+}(k_{S^-} + k_{S_{tran}})(k_{I_{out}}(k_{I^-} + k_{I_{tran}}) + k_{I_{tran}}k_{I^+}([TS] - T_0))[TS]}{k_{I_{in}}k_{S^+}k_{I^+}(k_{S^-} + k_{S_{tran}})[TS]} \quad (20)$$

$$[II] = \frac{k_{S^+}(k_{I^-} + k_{I_{tran}})([TS] - T_0) + k_{I_{tran}}(k_{S^-} + k_{S_{tran}})[TS]}{k_{I^+}(k_{S^-} + k_{S_{tran}})[TS]} \quad (21)$$

$$[TS] = [TS] \quad (22)$$

$$[TI] = (T_0 - [TS]) - \frac{(k_{S^-} + k_{S_{tran}})[TS]}{k_{S^+}[SI]} \quad (23)$$

$$[NS] = \frac{N_0k_{N^+}[SI]^n}{k_{N^-} + k_{N^+}[SI]^n} \quad (24)$$

$$\quad (24)$$

Similarly, *Equations 19* and *22* indicate that the equations under-determine the system and there are effectively two DOFs. For given values of $[SI]$ and $[TS]$, the other five compartment concentrations can be computed.

The output of main interest is obviously the observation compartment, which is the same function for the reduced model of form (9) – (13) and the full model of form (1) – (8) (*equations 17* and *24*). The steady state analysis shows that the ultimate level of binding to the DNA is a function of the total number of binding sites N_0 , the intracellular concentration of compound $[SI]$ and the substrate to DNA binding affinity, i.e. the binding association and dissociation rate constants, k_N^- and k_N^+ . Although this is a fairly intuitive result, it describes the exact relationship and may be useful at a later stage for parameter estimation and model validation.

5. STRUCTURAL IDENTIFIABILITY ANALYSIS

Structural identifiability arises from the inverse problem of inferring from the known, or assumed, properties of a biomedical or biological system a suitable model structure and estimates for the corresponding rate constants and other parameters. Structural identifiability analysis considers the uniqueness of the unknown model parameters from the input-output structure corresponding to proposed experiments to collect data for parameter estimation (under an assumption of the availability of perfect, noise-free data) [5-6]. This is an important, but often overlooked, theoretical prerequisite to experiment design, system identification and parameter estimation, since numerical estimates for unidentifiable parameters are effectively meaningless. If parameter estimates are to be used to inform about intervention or inhibition strategies, or other critical decisions, then it is essential that the parameters be uniquely identifiable. Such analysis is highly relevant to large-scale, highly complex systems, which are typical in Chemical kinetics and Systems Biology.

Numerous techniques for performing a structural identifiability analysis on linear parametric models exist and this is a well-understood topic [6-7]. In comparison, there are relatively few techniques available for nonlinear systems (the Taylor series approach [8], similarity transformation based approaches [9,10] and differential algebra techniques [11,12]) and significant computational problems can arise for these, even for relatively simple models. However, recently a structural identifiability analysis was successfully applied to a large-scale nonlinear mathematical model (43 state variables and 81 parameters) of a highly complex biomedical system [13].

Structural identifiability analysis has been performed on the Hoechst 33342 pharmacokinetic models developed using all of the three methods cited in order to ascertain whether the unknown system parameters can be identified uniquely or otherwise for the observation available.

5.1 Taylor series approach

This approach is normally used for systems with impulsive inputs and can be applied to both linear and non-linear systems. Given a compartmental model in the following general form:

$$\dot{\mathbf{q}}(t, \mathbf{p}) = \mathbf{F}(\mathbf{q}(t, \mathbf{p}), \mathbf{u}(t), \mathbf{p}), \quad (25)$$

$$\mathbf{q}(0, \mathbf{p}) = \mathbf{q}_0(\mathbf{p}), \quad (26)$$

$$\mathbf{y}(t, \mathbf{p}) = \mathbf{h}(\mathbf{q}(t, \mathbf{p}), \mathbf{p}), \quad (27)$$

where \mathbf{p} is the r dimensional vector of unknown parameters. The n dimensional vector $\mathbf{q}(t, \mathbf{p})$ is the state vector,

such that $\mathbf{q}_0(\mathbf{p})$ is the initial state. The m dimensional vector $\mathbf{u}(t)$ is the input vector and $\mathbf{y}(t, \mathbf{p})$ is the

observation vector. For the reduced model of form (9) – (13), these are given as:

$$\mathbf{p} = [kS_{in}, kS_{out}, k_N^+, k_N^-, k_S^+, k_S^-, kS_{tran}, SO_0, k, N_0, T_0],$$

$$\mathbf{q}(t, \mathbf{p}) = [[SO], [SI], [TS], [NS]],$$

$$\mathbf{q}_0(\mathbf{p}) = [SO_0, 0, 0, 0],$$

$$\mathbf{u}(t) = D\delta(t),$$

$$\mathbf{y}(t, \mathbf{p}) = [0, 0, 0, k[NS]],$$

where SO_0 is the initial concentration of $[SO]$. The components of the observation vector $y_i(t, \mathbf{p})$ are expanded

as a Taylor series about the initial condition.

$$y_i(t, \mathbf{p}) = y_i(0, \mathbf{p}) + \dot{y}_i(0, \mathbf{p})t + \ddot{y}_i(0, \mathbf{p})\frac{t^2}{2!} + \dots + y_i^{(k)}(0, \mathbf{p})\frac{t^k}{k!} + \dots$$

(28)

where

$$y_i^{(k)}(0, \mathbf{p}) = \left. \frac{d^k y_i}{dt^k} \right|_{t=0}$$

$$(k = 1, 2, \dots). \quad (29)$$

The Taylor series coefficients $y_i^{(k)}(0, \mathbf{p})$ are measurable and unique for a particular output. Unfortunately, due to the structural complexity of the system, this method did not converge to any solutions for either the reduced model of form (8) – (13) or the full model of form (1) – (8) in both Mathematica or Maple.

5.2 Similarity transformation/Exhaustive modelling approach

Given a linear model structure, this approach generates all the linear models that have the same input/output behaviour. It has been successfully applied to non-linear models by mapping the state equations to a linear set [13,14]. Taking a compartmental model in the general form (25) – (27), this approach entails establishing an Observability rank criterion. This is performed by defining a function \mathbf{H} given by

$$\mathbf{H}(\mathbf{q}, \mathbf{p}) = (\theta_1(\mathbf{q}, \mathbf{p}), \dots, \theta_n(\mathbf{q}, \mathbf{p}))^T \quad (30)$$

Where $\theta_n(\mathbf{q}, \mathbf{p})$ is the *Lie derivative* of \mathbf{h} , given by

$$L_f h(\mathbf{q}) = \frac{\partial h}{\partial \mathbf{q}}(\mathbf{q}) \cdot \mathbf{F}(\mathbf{q}) \quad (31)$$

Where \mathbf{h} is the observation from *Equation 27* and \mathbf{F} the system coordinate functions from *Equation 25*. If the resultant functions $\theta_1, \dots, \theta_n$ are linearly independent then the system (25) – (27) is said to satisfy the

Observability Rank Condition and it is possible to construct a smooth mapping from the state corresponding to a parameter vector $\bar{\mathbf{p}}$, indistinguishable from \mathbf{p} , to the state corresponding to \mathbf{p} . The smooth map ϕ satisfies the

following condition:

$$\mathbf{H}_{\bar{\mathbf{p}}}(\phi(\mathbf{q})) = \mathbf{H}_{\mathbf{p}}(\mathbf{q}) \quad (32)$$

Equations can then be derived from the initial conditions \mathbf{q}_0 , the model structure \mathbf{F} and the observation function

\mathbf{h} by using the following

$$\phi(\mathbf{q}_0(\bar{\mathbf{p}})) = \mathbf{q}_0(\bar{\mathbf{p}}) \tag{33}$$

$$\mathbf{F}(\phi(\mathbf{q}(t, \bar{\mathbf{p}})), \mathbf{p}) = \frac{\partial \phi}{\partial \mathbf{q}}(\mathbf{q}(t, \bar{\mathbf{p}})) \mathbf{F}(\mathbf{q}(t, \bar{\mathbf{p}}), \bar{\mathbf{p}}) \tag{34}$$

$$\mathbf{h}(\phi(\mathbf{q}(t, \bar{\mathbf{p}})), \mathbf{p}) = \mathbf{h}(\mathbf{q}(t, \bar{\mathbf{p}}), \bar{\mathbf{p}}) \tag{35}$$

The matrix for the reduced model of form (8) – (13) has full rank for appropriate $\theta_1 - \theta_4$, which for $n = 1$ is the

following

$$\theta_1 = [0, 0, 0, k]$$

$$\theta_2 = [0, k k_N^+ N_0, 0, k k_N^-]$$

$$\theta_3 = [k k_N^+ N_0 k S_{in}, -k k_N^+ N_0 k S_{out} - k k_N^{+2} N_0^2 - k k_N^- k_N^+ N_0, k k_N^+ N_0 k_N^-, \\ k k_N^- k_N^+ N_0 - k k_N^+ k S_{in} d + k k_N^{-2}]$$

$$\begin{aligned}
\theta_4 = & \left[-k k_N^+ N_0 kS_{in}^2 - k k_N^{+2} N_0^2 kS_{in} - k k_N^+ N_0 k_S^+ T_0 kS_{in} - k k_N^+ N_0 kS_{out} kS_{in} \right. \\
& - k k_N^- k_N^+ N_0 kS_{in}, \\
& k k_N^{-2} k_N^+ N_0 + k k_N^+ N_0 kS_{out}^2 + 2k k_N^{+2} N_0^2 kS_{out} + 2 k k_N^- k_N^+ N_0^2 + k k_N^{+3} N_0^3 \\
& + k k_N^- k_N^+ N_0 k_S^+ T_0 + k k_N^+ N_0 k_S^- k_S^+ T_0 + 2k k_N^+ N_0 kS_{out} k_S^+ T_0 + k k_N^+ N_0 kS_{out} kS_{in} \\
& + k k_N^+ N_0 k_S^{+2} T_0^2 + 2k k_N^{+2} N_0^2 k_S^+ T_0 + k k_N^- k_N^+ N_0 kS_{out} \\
& - 3k k_N^{+2} N_0 kS_{in} d, -k k_N^{+2} N_0^2 k_S^- - k k_N^+ N_0 k_S^{-2} - k k_N^+ N_0 k_S^- k_S^+ T_0 \\
& + k k_N^+ N_0 kS_{in} kS_{tran} + k k_N^+ N_0 k_S^+ kS_{in} d, -2 k k_N^{-2} k_N^+ N_0 - k k_N^- k_N^+ N_0^2 - k k_N^{-3} \\
& + k k_N^+ kS_{in}^2 d - k k_N^- k_N^+ N_0 k_S^+ T_0 - k k_N^- k_N^+ N_0 kS_{out} + 2 k k_N^{+2} N_0 kS_{in} d \\
& \left. + k k_N^+ kS_{out} kS_{in} d + 3 k_N^- k_N^+ kS_{in} d + k_N^+ k_S^+ T_0 kS_{in} d \right]
\end{aligned}$$

Analysis of the resulting equations demonstrates that the reduced model of form (8) – (13) was structurally globally identifiable if the initial compound dose is known. This is the case as both the concentration and the exact extracellular volume are known. In the case that the initial dose is unknown, only five parameters can be shown to be globally identifiable (kS_{in} , kS_{out} , k_S^- , kS_{tran} , and k_N^-) and the model is structurally unidentifiable.

Similarly for the full model of form (1) – (8); as the initial doses are known, it is structurally globally identifiable, and if it was not the case, only a limited number of parameters are globally identifiable (nine in total, including the same five for the reduced model of form (9) – (13), i.e. kS_{in} , kS_{out} , k_S^- , kS_{tran} , k_N^- , kI_{out} , kI_{tran} , kI_{in} , and k_I^-).

5.3 Differential algebra approach

In order to verify the identifiability analysis and gauge its level of applicability, this relatively new approach was also applied on both models. Given a compartmental model in the general form (25) – (27) where the observation $y = h(q(t, p), p)$ is linear one can take two parameter vectors, p and \bar{p} , that produce the same output for all t , and thus produce the same derivatives of the observation for all t , i.e.

$$y^{(n)}(t, \mathbf{p}) = y^{(n)}(t, \bar{\mathbf{p}}), \forall t. \quad (32)$$

Assuming one can generate an expression $g(y, \dots, y^{(n-1)}, \mathbf{p})$ derived from the model equations (25) – (27)

purely in terms of the observation vector $\mathbf{y}(t, \mathbf{p})$ and its derivatives then consider

$$g(y, \dots, y^{(n-1)}, \mathbf{p}) = g(y, \dots, y^{(n-1)}, \bar{\mathbf{p}}). \quad (33)$$

If equating the monomials of this function produces only one solution for the unknown parameters, then the system is globally identifiable. This approach yielded exactly the same results for the reduced model of form (9) – (13) as for the similarity transformation approach; confirming the model is structurally globally identifiable for a known dose. The same five parameters can also be shown to be globally identifiable (kS_{in} , kS_{out} , k_5^- , kS_{tran} , and $k_{\bar{N}}$) when the dose is unknown. The differential algebra approach has not yet successfully been applied to the full model of form (1) – (8). It has not yet been possible to generate the input output relationship $g(y, \dots, y^{(n-1)}, \mathbf{p})$ due to computational difficulties.

The analyses demonstrate that all models derived are uniquely identifiable for the experiments/observations available. This permits subsequent numerical parameter estimation to be performed with greater confidence.

6. DATA ANALYSIS

6.1 Software

Considering the number of parameters to be estimated and the complicated nature of the system equations, it was necessary to use an appropriate and numerically robust kinetic modelling software package. The software package FACSIMILE (MCPA Software, UK) was used since it provides a powerful means of solving ordinary differential equations encountered in biomedical engineering and can easily cope with the robust numerical solution given by the system equations. More importantly, it also contains a powerful parameter-fitting option, whereby specified parameters can be adjusted to obtain the best fit to observed data.

6.2 Parameter Estimation

During the optimisation process of the model parameter estimation described above, FACSIMILE measures the statistical goodness of the fit. This is achieved by calculating the residual sum of squares (RSS) and the standard

deviation of the natural logarithms (SDLN) of the estimated parameters as a measure of the confidence level (FACSIMILE works in terms of internal parameters that are the natural logarithms of the given model parameters). The RSS gives the difference between the model and the experimental data, i.e. the sum of the square of the error at each time point. FACSIMILE outputs a combined RSS value, i.e. sums all the individual time series RSS together, which is given as

$$RSS = \sum_{i=1}^r \sum_{j=1}^n \left(\frac{y_{obs,i}(j) - y_{sim,i}(t_j)}{\sigma_i} \right)^2 \quad (34)$$

where $y_{sim,i}(t_j)$ is the i^{th} model output at the j^{th} sampling time (t_j); $y_{obs,i}(j)$ is the corresponding experimental data point; and $\sigma_i = 0.01$ is an estimate for the standard error for the i^{th} output. FACSIMILE weighs each data set by this default value and thus negatively weighs the residuals by σ^2 , hence the RSS is effectively multiplied by a factor 10^4 [16]. The RSS provides an overall measure of how close the fit is to the experimental data.

The confidence level of the estimated parameter value is a statistical measure of how well the model and the data define the parameter. It is given as the standard deviation of the natural logarithms (SDLN) of the estimated parameters. The SDLN values are estimated using the variance-covariance matrix of the total number of parameters and the number of well-determined parameters and can be considered as a percentage error.

7. RESULTS

To obtain parameter estimates for the full system the data with Hoechst 33342 alone were fitted to the reduced model of form (9) – (13). These were then used as initial estimates to fit the full system.

7.1 Reduced model of form (9) – (13)

During the numerous fits undertaken, analysing the graphs showed that for 0.5 μ M and 10 μ M initial substrate concentration, the fits consistently underestimated the data. Considering that each time series is a different experiment it was reasonable to assume that there could be some variance in the total number of binding sites on the nucleus. A higher number of binding sites would allow the time series to reach higher values and give a better visual fit. Instead of using one parameter of N_0 common to all twelve time series, five individual

parameters were set, one for each initial substrate concentration. *Table 3* provides the parameter estimates for the best fit obtained for the twelve data sets used where no inhibitor is present.

Insert Table 3

Fig. 4 illustrates the best fit associated with *Table 3*, obtained for the twelve data sets used where no inhibitor is present.

Insert Figure 4

Fig. 4 is split into five separate charts (a-e), one for each initial substrate concentration.

7.2 Full Model of form (1) – (8)

It did not prove possible to obtain a complete set of well determined parameters within FACSIMILE for the full model. This was remedied by fixing the 16 independent values of N_0 and T_0 , the order of the nucleus binding reaction – n , the extracellular substrate concentration – *Dose*, the initial intracellular inhibitor concentration – $Initial_{in}$, and the initial extracellular inhibitor concentration – $Initial_{out}$. *Table 4* provides the parameter estimates for best fits obtained for the full model using all 32 data sets where inhibitor is present.

Insert Table 4

There are some discrepancies between the reduced and full model parameter estimates. It is suspected these are mainly due to the different order of the nucleus binding reaction – n , which affects the curvature of the fit (a higher value producing a more sigmoidal shaped curve). For the reduced model of form (9) – (13), FACSIMILE converged towards a value of $n = 4.35$, producing the little step at the beginning of the times series, whereas for the full model a value of $n = 1$ proved more stable. The higher RSS value for the full model relative to the reduced model is believed to be due to the higher number of degrees of freedom and larger number of data sets used.

Fig. 5 illustrates the best fit associated with *Table 4*, obtained for the full model using 32 data sets where inhibitor is present.

Insert Figure 5

Fig. 5 is split into eight separate charts (a-h), one for each initial inhibitor concentration. It can be seen from each one that the initial step is less prominent.

8. DISCUSSION AND CONCLUSIONS

The model adequately reproduces the observed time series. Though introduced into the model to improve the model fits the estimated values of T_0 and N_0 do not differ greatly between experimental conditions, which suggests that the conditions within each well of the multi-well plate are effectively constant. Parameters were estimated with a reasonable level of confidence, which can be judged by the SDLN values for each parameter.

Given these estimates, the binding affinity of FTC for the BCRP transporter can be calculated to be

$$K_D = \frac{k_i^+}{k_i^-} = 1.67 \times 10^{-5} \mu M \quad (35)$$

This is understandable given the effect small concentrations of FTC have on the observed fluorescence-time profiles.

The models adequately describe the data observed. It can be further seen that BCRP mediated cellular kinetics can be indirectly measured in this way. A structural identifiability analysis was performed successfully with two methods; the similarity transformation/exhaustive modelling approach and the differential algebra approach. These demonstrate that all models derived are uniquely identifiable for the experiments/observations available, adding greater confidence to the numerical parameter estimation carried out.

By modelling the kinetics of the system the binding kinetics for FTC can be inferred. A more simplistic approach that used the change in the steady state fluorescence would only yield a concentration of FTC *outside of the cell* that would alter the observed fluorescence by a given amount and such information would be difficult to relate to *in vivo* data. The binding affinity can be compared to blood concentrations observed to assess the impact of BCRP on a drug's ability to penetrate cancer cells and this aspect is currently under investigation.

REFERENCES

- [1] Doyle LA, Yang W, Abruzzo LV, Krogmann T, Gao Y, Rishi AK and Ross DD (1998) A multidrug resistance transporter from human MCF-7 breast cancer cells. *Proc Natl Acad Sci USA* 95:15665-15670.
- [2] Lalande ME, Ling V and Miller RG (1981) Hoechst 33342 dye uptake as a probe of membrane permeability changes in mammalian cells. *Proc Natl Acad Sci USA* 78:363-367.
- [3] Paine SW, Parker AJ, Gardiner P, Webborn PJH, and Bailey RJ. (2008). Prediction of the pharmacokinetics of atorvastation, cerivastatin and indomethacin using kinetic models applied to isolated rat hepatocytes. *Drug Metabolism and Distribution*. 36: 1365-1374
- [4] Baker M, and Parton T (2007). Kinetic determinants of hepatic clearance: Plasma protein binding and hepatic uptake. *Xenobiotica*. 37: 1110-1134.
- [5] R. Bellman, K.J. Åström, On structural identifiability, *Math. Biosci.* 7 (1970) 329.
- [6] K.R. Godfrey, J.J. DiStefano III, Identifiability of model parameters, in: E. Walter (Ed.), *Identifiability of Parametric Models*, Pergamon, Oxford, 1987, p. 1 (Chapter 1).
- [7] E. Walter (Ed.), *Identifiability of Parametric Models*, Pergamon, Oxford, 1987.
- [8] H. Pohjanpalo, System identifiability based on the power series expansion of the solution, *Math. Biosci.* 41 (1978) 21.
- [9] E.T. Tunali, T.J. Tarn, New results for identifiability of nonlinear systems, *IEEE Trans. Automat. Contr.* 32 (1987) 146.
- [10] S. Vajda, K.R. Godfrey, H. Rabitz, Similarity transformation approach to identifiability analysis of nonlinear compartmental models, *Math. Biosci.* 93 (1989) 217.
- [11] M. Fliess, S.T. Glad, An algebraic approach to linear and nonlinear control, in: H.L. Trentelman, J.C. Willems (Eds.), *Essay on Control: Perspectives in the Theory and its Applications*, vol. 14, *Progress in Systems and Control Theory*, Birkhäuser, Boston, 1993.
- [12] L. Ljung, T. Glad, On global identifiability for arbitrary model parametrizations, *Automatica* 30 (1994) 265.
- [13] S.Y.A. Cheung, N.D. Evans, M.J. Chappell, K.R. Godfrey, P.J. Smith, R.J. Errington (2008) 'Exploration of the intercellular heterogeneity of topotecan uptake into human breast cancer cells through compartmental modelling' *Mathematical Biosciences* 213 (2), 119 - 134 (0025-5564)
- [14] N.D. Evans, M.J. Chapman, M.J. Chappell, K.R. Godfrey, Identifiability of uncontrolled nonlinear rational

systems, *Automatica* 38 (2002) 1799.

[15] Evans N D, White L J, Chapman M J, Chappell M J, Godfrey K R (2005) 'The structural identifiability of the Susceptible Infected Recovered model with seasonal forcing', *Mathematical Biosciences*, 194 (2), 175 - 197 (0025-5564).

[16] *Facsimile (Version 4.0) Technical Reference*. (1995), Laboratory, A.T.H., Didcot, Oxon, uUK.

Table Legends:

Table 1: Description of the inter-compartmental rate transfer and compartments

Table 2: Matrix showing number of time series for each experimental set up

Table 3: Well-determined parameters for reduced model of form (9) – (13) with no inhibitor present

Table 4: Well- determined parameters for full model of form (1) – (8) with inhibitor present

Inter-compartmental rate transfers		Compartments	
$k_{S_{in}}$	Marker compound cellular influx	[SO]	Extracellular concentration of marker compound
$k_{S_{out}}$	Marker compound cellular efflux		
$k_{I_{in}}$	Inhibitor cellular influx	[SI]	Intracellular concentration of marker compound
$k_{I_{out}}$	Inhibitor cellular efflux		
k_N^+	Marker compound nuclear binding	[IO]	Extracellular concentration of inhibitor compound
k_N^-	Marker compound nuclear dissociation		
k_S^+	Marker compound transporter binding	[II]	Intracellular concentration of inhibitor compound
k_S^-	Marker compound transporter dissociation		
k_I^+	Inhibitor transporter binding	[TS]	Marker bound to transporter
k_I^-	Inhibitor transporter dissociation		
$k_{S_{tran}}$	Transporter flow back to marker compound	[TI]	Inhibitor bound to transporter
$k_{I_{tran}}$	Transporter flow back to inhibitor		
		[NS]	Marker bound to the nucleus (DNA)

Table 1: Description of the inter-compartmental rate transfer and compartments

		Hoeschst concentration [μM]					
		0.5	1	2	5	10	Total
Inhibitor Concentration [μM]	0	2	2	4	2	2	12
	0.1	2	2	4	2	2	12
	0.32	2	2	4	2	2	12
	1	2	2	4	2	2	12
	3.16	2	2	4	2	2	12
	10	2	2	4	2	2	12
	31.6	2	2	4	2	2	12
	100	2	2	4	2	2	12
	Total	16	16	32	16	16	96

Table 2: Matrix showing number of time series for each experimental set up

Parameters	Value	SDLN	Parameters	Value	SDLN
kS_{in}	$2.45 \times 10^{-7} \text{ s}^{-1}$	42.7%	T_{03}	6.83×10^5	8.6%
kS_{out}	3.23 s^{-1}	6.5%	T_{04}	2.27×10^5	6.3%
k_N^+	$6.68 \times 10^{-4} \text{ s}^{-1}$	39.3%	T_{05}	4.41×10^6	5.8%
k_N^-	$4.43 \times 10^{-2} \text{ s}^{-1}$	42.5%	N_{01}	2.49×10^4	29.6%
k_S^+	$1.30 \times 10^{-5} \text{ s}^{-1}$	7.3%	N_{02}	3.11×10^3	12.6%
k_S^-	$1.05 \times 10^{-3} \text{ s}^{-1}$	10.9%	N_{03}	9.91×10^2	3.5%
kS_{tran}	$1.26 \times 10^{-4} \text{ s}^{-1}$	51.5%	N_{04}	1.33×10^3	3.2%
n	4.35	7.8%	N_{05}	1.84×10^3	2.1%
T_{01}	2.91×10^5	7.2%	$Dose$	1.55×10^7	36.3%
T_{02}	4.48×10^5	9.4%	RSS	7.44	-

Table 3: Well-determined parameters for reduced model of form (9) – (13) with no inhibitor present

Parameters	Value	SDLN	Parameters	Value	SDLN
kS_{in}	$4.69 \times 10^{-7} \text{ s}^{-1}$	3.5%	T_{05}	5.22×10^5	fixed
kS_{out}	$4.43 \times 10^{-2} \text{ s}^{-1}$	157%	T_{06}	1.34×10^6	fixed
kI_{in}	$2.48 \times 10^{-4} \text{ s}^{-1}$	27.6%	T_{07}	2.91×10^6	fixed
kI_{out}	$1.36 \times 10^{-3} \text{ s}^{-1}$	71.2%	T_{08}	2.06×10^6	fixed
k_N^+	$1.75 \times 10^{-4} \text{ s}^{-1}$	137%	N_{01}	1.55×10^3	fixed
k_N^-	$7.92 \times 10^{-4} \text{ s}^{-1}$	59.2%	N_{02}	3.71×10^3	fixed
k_S^+	$2.36 \times 10^{-6} \text{ s}^{-1}$	136%	N_{03}	4.79×10^3	fixed
k_S^-	$1.81 \times 10^{-3} \text{ s}^{-1}$	46.6%	N_{04}	8.97×10^3	fixed
k_I^+	$5.27 \times 10^{-9} \text{ s}^{-1}$	53.8%	N_{05}	6.16×10^3	fixed
k_I^-	$3.15 \times 10^{-4} \text{ s}^{-1}$	138%	N_{06}	6.75×10^3	fixed
kS_{tran}	$2.56 \times 10^{-4} \text{ s}^{-1}$	79.8%	N_{07}	7.25×10^3	fixed
kI_{tran}	$2.04 \times 10^{-4} \text{ s}^{-1}$	168%	N_{08}	5.70×10^3	fixed
T_{01}	6.55×10^5	fixed	n	1	fixed
T_{02}	1.43×10^6	fixed	$Dose$	3.20×10^6	fixed
T_{03}	7.53×10^5	fixed	$Initial_{in}$	5.75×10^4	fixed
T_{04}	7.57×10^5	fixed	$Initial_{out}$	4.71×10^3	fixed
RSS	48.627	-	-	-	-

Table 4: Well- determined parameters for full model of form (1) – (8) with inhibitor present

Figure Legends:

Fig 1: Model representation

Fig 2: Sample data plot for different initial Hoechst concentrations with no inhibitor present Fig. 3: Reduced model representation

Fig. 4: FACSIMILE fits for reduced model of form (9) – (13) without inhibitor

Fig. 5: FACSIMILE fits for full model of form (1) – (8) with inhibitor

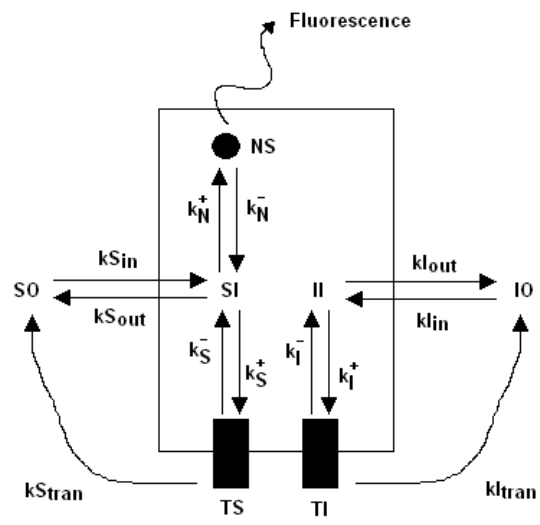


Fig 1: Model representation

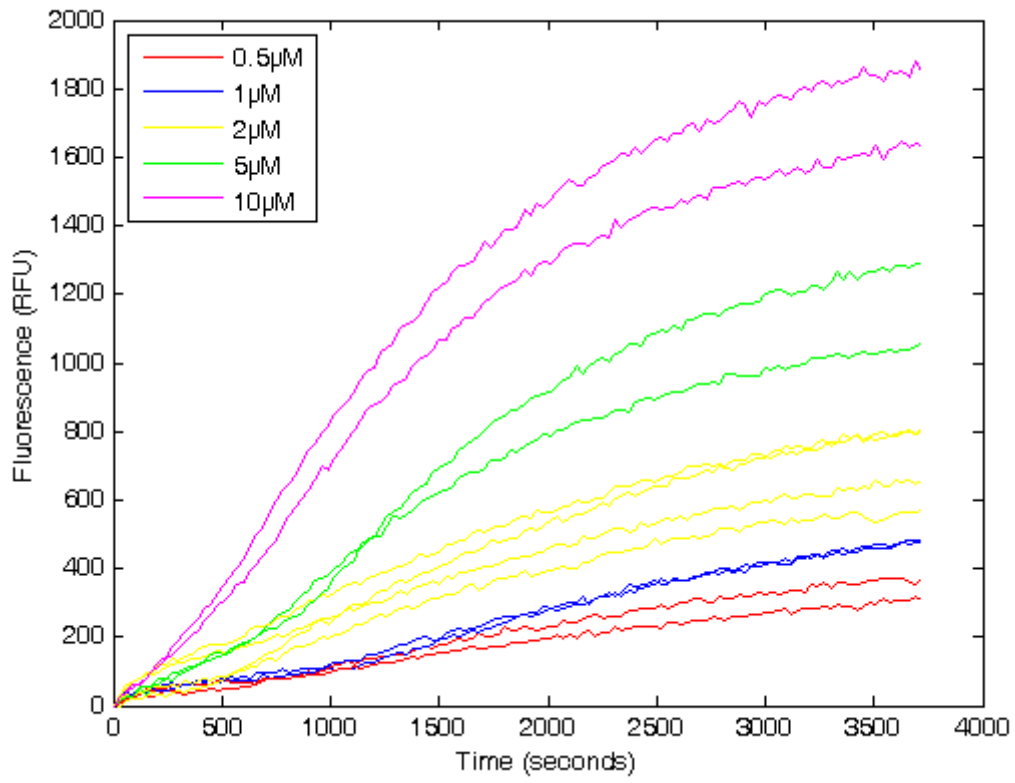


Fig 2: Sample data plot for different initial Hoechst concentrations with no inhibitor present

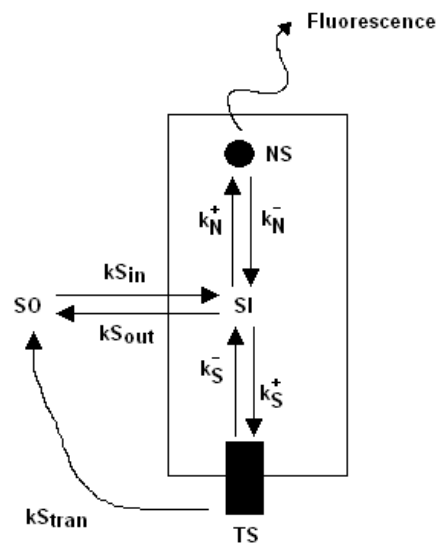


Fig. 3: Reduced model representation

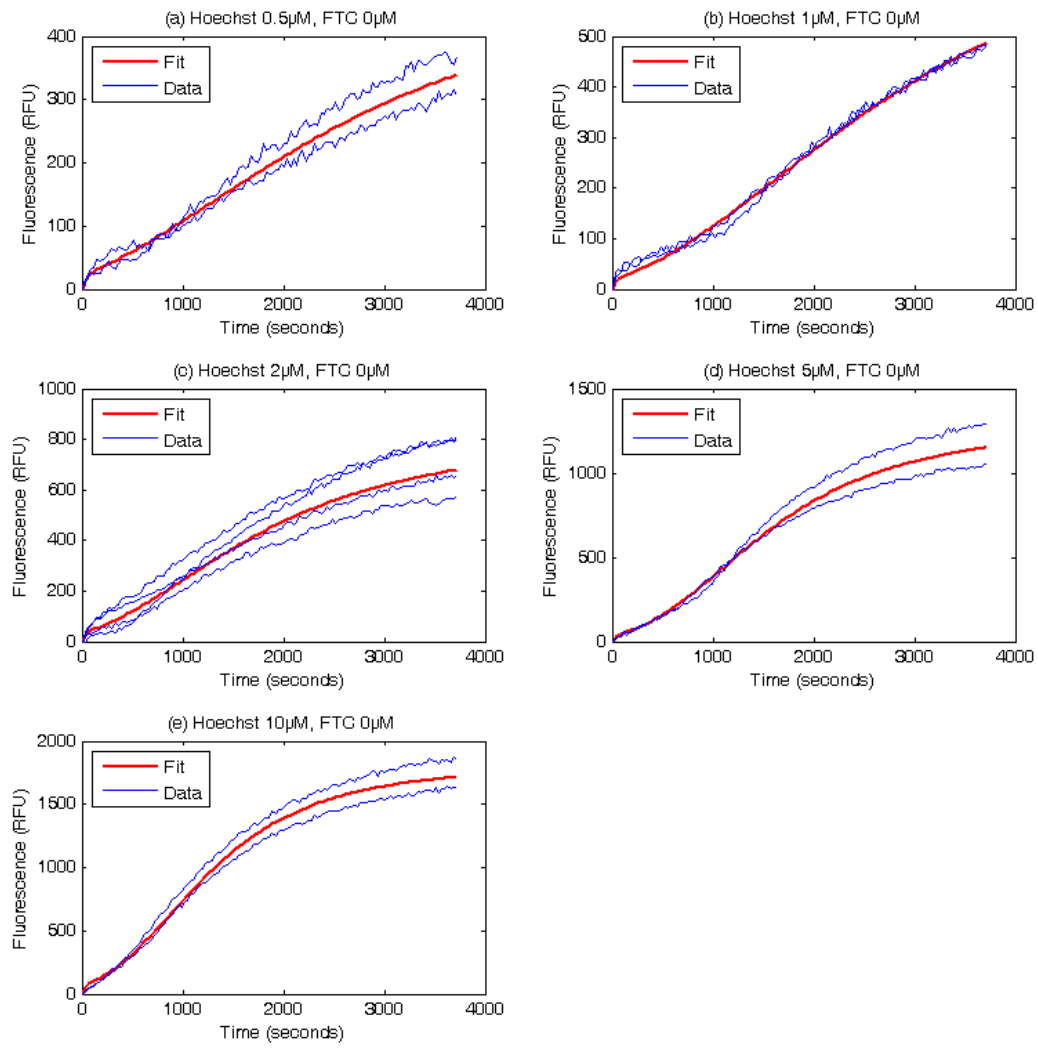


Fig. 4: FACSIMILE fits for reduced model of form (9) – (13) without inhibitor

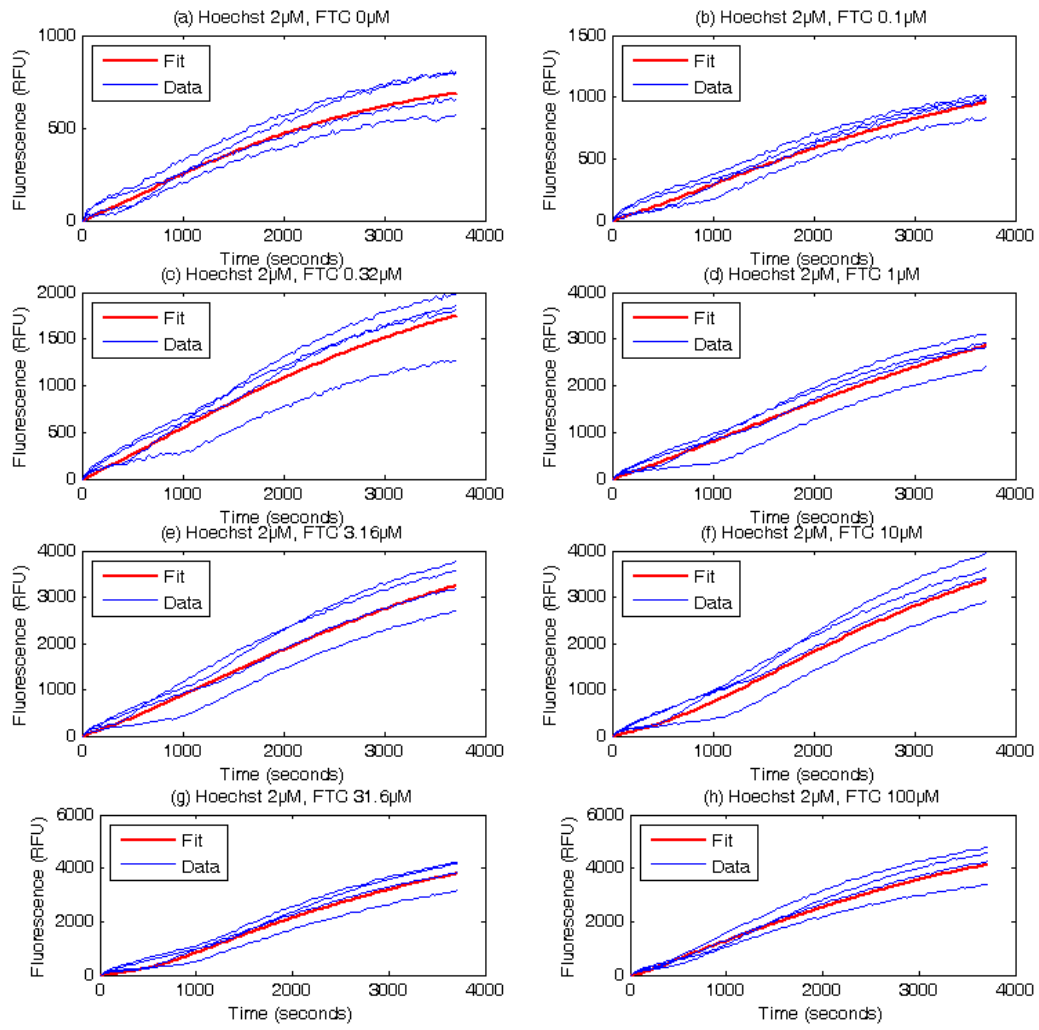


Fig. 5: FACSIMILE fits for full model of form (1) – (8) with inhibitor