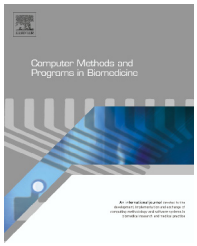
COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE XXX (2013) XXX–XXX

journal homepage: www.intl.elsevierhealth.com/journals/cmpb



Structural identifiability analyses of candidate models for *in vitro* Pitavastatin hepatic uptake☆,☆☆

Thomas R.B. Grandjean[a,*], *Michael J. Chappell*[a,1], *James W.T. Yates*[b,2], *Neil D. Evans*[a]

[a] *School of Engineering, University of Warwick, Coventry CV4 7AL, UK*
[b] *AstraZeneca R&D, Alderley Park, Macclesfield SK10 4TG, UK*

ARTICLE INFO

Keywords:
Pitavastatin
Compartmental models
Structural identifiability

ABSTRACT

In this paper a review of the application of four different techniques (a version of the similarity transformation approach for autonomous uncontrolled systems, a non-differential input/output observable normal form approach, the characteristic set differential algebra and a recent algebraic input/output relationship approach) to determine the structural identifiability of certain *in vitro* nonlinear pharmacokinetic models is provided. The Organic Anion Transporting Polypeptide (OATP) substrate, Pitavastatin, is used as a probe on freshly isolated animal and human hepatocytes. Candidate pharmacokinetic non-linear compartmental models have been derived to characterise the uptake process of Pitavastatin. As a prerequisite to parameter estimation, structural identifiability analyses are performed to establish that all unknown parameters can be identified from the experimental observations available.

1. Introduction

Pitavastatin is a drug used to treat hypercholesterolaemia. It shows active uptake into hepatocytes, mediated mainly by Organic Anion Transporting Polypeptide (OATP) 1B1 [1,2]. It is desired to investigate the nonlinear kinetics of in vitro hepatic uptake of the OATP substrate, Pitavastatin, and quantify the mechanisms present both structurally and numerically. Experiments utilising the 'oil spin' methodology described by [3] have been designed at AstraZeneca to investigate the nonlinear kinetics of *in vitro* hepatic uptake. Six candidate models are proposed to characterise the uptake process. In order to perform parameter estimation and compare the models to ascertain those most suitable for predictive purposes, it is a necessary requirement to first ask "do the observations uniquely determine the unknown model parameters?". In the models derived, as in most biomedical systems modelling, the model parameters have biological meaning and it is desired to establish whether it is at all possible to estimate their values from experimental data. Techniques for structural identifiability analysis look to determine whether unknown model parameters have unique values given the available observation(s) [4].

☆☆ This work is supported by the UK MRC Capacity Building Initiative and AstraZeneca.
* *Corresponding author*. Tel.: +44 02476572904.
E-mail addresses: t.grandjean@warwick.ac.uk, t.grandjean@hotmail.com (T.R.B. Grandjean), M.J.Chappell@Warwick.ac.uk (M.J. Chappell), James.Yates@AstraZeneca.com (J.W.T. Yates), Neil.Evans@warwick.ac.uk (N.D. Evans).
[1] Tel.: +44 024765 24309.
[2] Tel.: +44 01625 512996.
0169-2607/$ – see front matter © 2013 The Authors. Published by Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cmpb.2013.06.013

2. Structural identifiability

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 [7,8]. It is important to note that an *a priori* structurally identifiable model does not necessarily guarantee a posteriori numerical parameter identifiability, for example [9], however it does greatly increase the confidence in the parameter estimation process for the given system observation.

Numerous techniques for performing a structural identifiability analysis on linear parametric models exist and this is a well-understood topic [5,10]. In comparison, there are relatively few techniques available for nonlinear systems (the Taylor series approach [11], similarity transformation based approaches [12,13], and differential algebra techniques) [14,15] and significant computational problems can arise for these, even for relatively simple models [16,17].

In this paper, four methods are reviewed: a version of the similarity transformation approach for autonomous uncontrolled systems [18], a non-differential input/output observable normal form approach [19], the characteristic set differential algebra approach [14,15], and a recently introduced algebraic input/output relationship approach [19]. Each approach is performed on all of the Pitavastatin pharmacokinetic models developed in order to ascertain whether the unknown system parameters can be identified uniquely or otherwise for the observation available and to compare their performance.

For a given output, an *unidentifiable* parameter can take an (uncountably) infinite set of values, whereas a *nonuniquely (locally) identifiable* parameter can take any of a distinct (countable) set of values. A parameter is *globally identifiable* if for a given output, it can only take one value.

If all of the unknown parameters are globally identifiable, the system model is structurally globally identifiable (SGI). In the case that all parameters are locally identifiable and at least one is non-uniquely identifiable then the model is structurally locally identifiable (SLI). In the case where at least one parameter is unidentifiable then the model is structurally unidentifiable (SU).

Due to the complex nature of the analytical approaches, a symbolic computational package, namely Maple 2010 (Maplesoft) [20], was used to perform the analyses.

3. Models

As described previously, Pitavastatin is a substrate of OATP, which actively mediates the transport of the drug across the hepatocyte membrane. Diffusion also takes place at the cell membrane, where the drug flows in and out of the cell according to the concentration gradient. These two mechanisms can be represented by the compartmental model shown in Fig. 1. In the model, each compartment represents a different component of the hepatocyte cell.

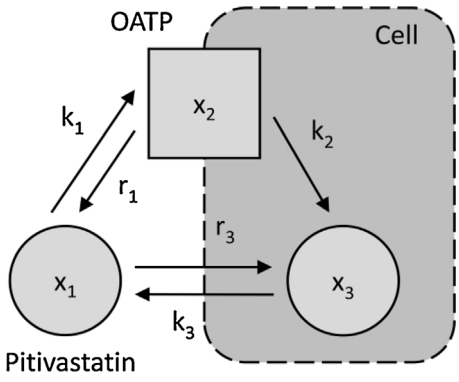


Fig. 1 – Basic conceptual model representation.

A known concentration of Pitavastatin can be added to the medium in which the hepatocytes sit at the beginning of the experiment (x_1). The substrate actively binds to OATP (x_2) and is mediated into the cell (x_3). Extracellular Pitavastatin (x_1) also flows into the cell (x_3) by diffusion with rate constants k_3 and r_3.

3.1. System equations

The system of ordinary differential equations describing the models is derived using classical mass-balance principles as per [4] for example. The corresponding model equations are given by:

$$\dot{x}_1 = k_3x_3 - r_3x_1 - k_1x_1(T_0 - x_2) + r_1x_2 \quad (1)$$

$$\dot{x}_2 = k_1x_1(T_0 - x_2) - (r_1 + k_2)x_2 \quad (2)$$

$$\dot{x}_3 = r_3x_1 - k_3x_3 + k_2x_2 \quad (3)$$

where T_0 is the total number of transporter binding sites on OATP. Here the unknown parameter set, p, is given by:

$$p = \{k_1, r_1, k_2, k_3, r_3, T_0\}. \quad (4)$$

The initial conditions are given by:

$$x_1(0) = D, \quad x_2(0) = 0, \quad x_3(0) = 0, \quad (5)$$

where D is the initial dose in μmol (1 million cells). The initial concentration of the medium in which the hepatocytes sit at the beginning of the experiment is known and given in μmol/L. The initial volume is 1 mL and 1 million cells are used thus the initial concentration is multiplied by a factor of 10^{-3} to convert it from μmol/L to μmol (1 million cells). x_1, x_2, x_3 therefore denote quantities in μmol (1 million cells).

Please cite this article in press as: T.R.B. Grandjean, et al., Structural identifiability analyses of candidate models for *in vitro* Pitavastatin hepatic uptake, Comput. Methods Programs Biomed. (2013), http://dx.doi.org/10.1016/j.cmpb.2013.06.013

Finally, the observation of the system is given by:

$$y = k \cdot (x_2 + x_3) \tag{6}$$

where k is the observation gain, i.e. the observation is the sum of the quantity of Pitavastatin in compartments 2 and 3. The observed concentration is given in nmol/L (0.1 million cells–100 μL is taken from 1 mL) and the measured volume is 900 μL. The observation is therefore multiplied by 9×10^{-6} to convert from nmol/L (0.1 million cells) to μmol (1 million cells). As the units for the initial dose D and the observation are equivalent, the observation gain k is effectively 1.

It is possible to simplify the model of the form (1)–(3), by noting that adding (1)–(3) equals zero, as below

$$\dot{x}_1 + \dot{x}_2 + \dot{x}_3 = 0. \tag{7}$$

Integrating with respect to time and solving for the initial conditions yields

$$x_1 + x_2 + x_3 = D \tag{8}$$

Re-arranging (8)

$$x_1 = D - x_2 - x_3 \tag{9}$$

and substituting (9) into the model of the form (1)–(3) reduces it to a system of two ordinary differential equations;

$$\dot{x}_2 = k_1(D - x_2 - x_3)(T_0 - x_2) - (r_1 + k_2)x_2 \tag{10}$$

$$\dot{x}_3 = r_3(D - x_2 - x_3) - k_3x_3 + k_2x_2 \tag{11}$$

with initial conditions

$$x_2(0) = 0, \quad x_3(0) = 0 \tag{12}$$

and observation

$$y = k \cdot (x_2 + x_3). \tag{13}$$

Similarly x_2 can be eliminated by re-arranging (8) to

$$x_2 = D - x_1 - x_3, \tag{14}$$

or alternatively, x_3 can be eliminated by re-arranging (8) to give

$$x_3 = D - x_2 - x_3. \tag{15}$$

The resulting system equations and corresponding initial conditions and observation are shown in Table 1.

3.2. *Pseudo steady state assumption*

It is possible to reduce the above model using a commonly applied approximation in the biological/pharmacokinetic modelling literature. The necessary assumption is that the binding to the transporter occurs very rapidly over the time scale of the rate of appearance of intracellular compound [4].

Taking the proposed model of the form (1)–(3), this is equivalent to assuming that the OATP association and dissociation rate constants, k_1 and r_1 are considerably faster than the other rates, namely the flow into the cell, k_2 and the diffusion into and out of the cell, k_3 and r_3, i.e. there is rapid equilibration of OATP. If this assumption is true then instantaneously after the experiment has begun, the amount of substrate bound to transporter (x_2) is effectively constant, the rate of change of OATP (2) can be set to zero and the right hand side of (2) can be re-arranged to give

$$x_2 = \frac{T_0x_1}{K_M + x_1} \tag{26}$$

where

$$K_M = \frac{r_1 + k_2}{k_1} \tag{27}$$

is the relevant Michaelis–Menten constant. Substituting (26) back into the original system equations, (1) and (3) become

$$\dot{x}_1 = k_3x_3 - r_3x_1 - \frac{V_Mx_1}{K_M + x_1} \tag{28}$$

$$\dot{x}_3 = r_3x_1 - k_3x_3 + \frac{V_Mx_1}{K_M + x_1} \tag{29}$$

respectively, where

$$V_M = k_2T_0 \tag{30}$$

is the maximum velocity of the reaction. The unknown parameter set, p, is now given by

$$p = \{V_M, K_M, k_3, r_3, T_0\}. \tag{31}$$

The initial conditions are

$$x_1(0) = D_1, \quad x_3(0) = 0, \tag{32}$$

where D_1 is a constant. It is possible to simplify the model of the form (28) and (29), by noting that adding (28) and (29) equals zero, as below

$$\dot{x}_1 + \dot{x}_3 = 0. \tag{33}$$

Integrating (33) with respect to time and solving for the initial condition yields

$$x_1 + x_3 = D_1. \tag{34}$$

Re-arranging (34)

$$x_1 = D_1 - x_3 \tag{35}$$

and substituting (35) into the model of form (28) and (29) reduces it to one ordinary differential equation. Alternatively x_3 can be eliminated by re-arranging (34) as

$$x_3 = D_1 - x_1. \tag{36}$$

Please cite this article in press as: T.R.B. Grandjean, et al., Structural identifiability analyses of candidate models for *in vitro* Pitavastatin hepatic uptake, Comput. Methods Programs Biomed. (2013), http://dx.doi.org/10.1016/j.cmpb.2013.06.013

Table 1 – Alternate system equations for the model of the form (1)–(3).

System equations	Initial conditions	Observation
Eliminating x_2:		
$\dot{x}_1 = k_3x_3 - r_3x_1 - k_1x_1(T_0 - D + x_1 + x_3) + r_1(D - x_1 - x_3)$ (16)	$x_1(0) = D$ (18)	$y = k \cdot (D - x_1)$ (20)
$\dot{x}_3 = r_3x_1 - k_3x_3 + k_2(D - x_1 - x_3)$ (17)	$x_3(0) = 0$ (19)	
Eliminating x_3:		
$\dot{x}_1 = k_3(D - x_1 - x_2) - r_3x_1 - k_1x_1(T_0 - x_2) + r_1x_2$ (21)	$x_1(0) = D$ (23)	$y = k \cdot (D - x_1)$ (25)
$\dot{x}_2 = k_1x_1(T_0 - x_2) - (r_1 + k_2)x_2$ (22)	$x_2(0) = 0$ (24)	

Table 2 – Alternate system equations for the model of the form (28)–(29).

System equation	Initial condition	Observation
Eliminating x_1:		
$\dot{x}_3 = r_3(D_1 - x_3) - k_3x_3 + \frac{V_M(D_1 - x_3)}{K_M + D_1 - x_3}$ (37)	$x_3(0) = 0$ (38)	$y = k\left(\frac{T_0(D_1 - x_3)}{K_M + D_1 - x_3} + x_3\right)$ (39)
Eliminating x_3:		
$\dot{x}_1 = r_3x_1 - k_3(D_1 - x_1) + \frac{V_Mx_1}{K_M + x_1}$ (40)	$x_1(0) = D_1$ (41)	$y = k\left(\frac{T_0x_1}{K_M + x_1} + D_1 - x_1\right)$ (42)

Both variations are shown in Table 2.

3.3. *Non specific binding*

Another candidate model developed allows for non specific binding [21] of Pitavastatin at the cell wall. This is where the medium in which the cells sit in at the beginning of the experiment is not homogenous and a concentration gradient occurs at the hepatic cell wall. This is described by adding a fourth compartment (x_4) to represent a concentration present at the cell wall, as shown in Fig. 2

The corresponding set of nonlinear ordinary differential equations characterising the proposed model are given by,

$$\dot{x}_1 = k_4x_4 - r_4x_1 - k_1x_1(T_0 - x_2) + r_1x_2 \quad (43)$$

$$\dot{x}_2 = k_1x_1(T_0 - x_2) - (r_1 + k_2)x_2 \quad (44)$$

$$\dot{x}_3 = r_3x_4 - k_3x_3 + k_2x_2 \quad (45)$$

$$\dot{x}_4 = r_4x_1 - (k_4 + r_3)x_4 + k_3x_3 \quad (46)$$

with initial conditions

$$x_1(0) = D, \quad x_2(0) = 0, \quad x_3(0) = 0, \quad x_4(0) = 0 \quad (47)$$

and observation

$$y = k \cdot (x_2 + x_3 + x_4). \quad (48)$$

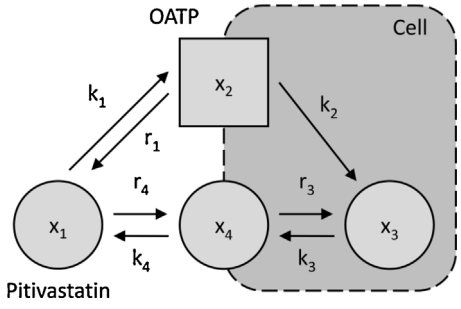


Fig. 2 – Non specific binding model representation.

Here the unknown parameter set, p, is given by

$$p = \{k_1, r_1, k_2, k_3, r_3, k_4, r_4, T_0\}. \quad (49)$$

Again it is possible to simplify the model of the form (43)–(46), by noting that adding (43)–(46) gives

$$\dot{x}_1 + \dot{x}_2 + \dot{x}_3 + \dot{x}_4 = 0. \quad (50)$$

Integrating (50) with respect to time and solving for the initial condition yields

$$x_1 + x_2 + x_3 + x_4 = D. \quad (51)$$

Re-arranging (51) and substituting into the model of the form (43)–(46) reduces it to three ordinary differential equations. The four alternative structures eliminating x_1, x_2, x_3, and x_4 respectively are shown in Table 3.

The same pseudo steady state assumption as per 3.2, can also be made to obtain another representation of this model of the form

$$\dot{x}_1 = k_4x_4 - r_4x_1 - \frac{V_Mx_1}{K_M + x_1} \quad (80)$$

$$\dot{x}_3 = r_3x_4 - k_3x_3 + \frac{V_Mx_1}{K_M + x_1} \quad (81)$$

$$\dot{x}_4 = r_4x_1 - (k_4 + r_3)x_4 + k_3x_3. \quad (82)$$

with initial conditions

$$x_1(0) = D_1, \quad x_3(0) = 0, \quad x_4(0) = 0 \quad (83)$$

and observation

$$y = k \cdot \left(\frac{T_0x_1}{K_M + x_1} + x_3 + x_4\right). \quad (84)$$

The unknown parameter set, p, is now given by

$$p = \{V_M, K_M, k_3, r_3, k_4, r_4, T_0\}. \quad (85)$$

Please cite this article in press as: T.R.B. Grandjean, et al., Structural identifiability analyses of candidate models for *in vitro* Pitavastatin hepatic uptake, Comput. Methods Programs Biomed. (2013), http://dx.doi.org/10.1016/j.cmpb.2013.06.013

Table 3 – Alternate system equations for the model of the form (43)–(46).

System equations	Initial conditions	Observation
Eliminating x_1:		
$\dot{x}_2 = k_1(D - x_2 - x_3 - x_4)(T_0 - x_2) - (r_1 + k_2)x_2$ (52)	$x_2(0) = 0$ (55)	
$\dot{x}_3 = r_3x_4 - k_3x_3 + k_2x_2$ (53)	$x_3(0) = 0$ (56)	$y = k \cdot (x_2 + x_3 + x_4)$ (58)
$\dot{x}_4 = r_4(D - x_2 - x_3 - x_4) - (k_4 + r_3)x_4 + k_3x_3$ (54)	$x_4(0) = 0$ (57)	
Eliminating x_2:		
$\dot{x}_1 = k_4x_4 - r_4x_1 - k_1x_1(T_0 - D + x_1 + x_3 + x_4) + r_1(D - x_1 - x_3 - x_4)$ (59)	$x_1(0) = D$ (62)	
$\dot{x}_3 = r_3x_4 - k_3x_3 + k_2(D - x_1 - x_3 - x_4)$ (60)	$x_3(0) = 0$ (63)	$y = k \cdot (D - x_1)$ (65)
$\dot{x}_4 = r_4x_1 - (k_4 + r_3)x_4 + k_3x_3$ (61)	$x_4(0) = 0$ (64)	
Eliminating x_3:		
$\dot{x}_1 = k_4x_4 - r_4x_1 - k_1x_1(T_0 - x_2) + r_1x_2$ (66)	$x_1(0) = D$ (69)	
$\dot{x}_2 = k_1x_1(T_0 - x_2) - (r_1 + k_2)x_2$ (67)	$x_2(0) = 0$ (70)	$y = k \cdot (D - x_1)$ (72)
$\dot{x}_4 = r_4x_1 - (k_4 + r_3)x_4 + k_3(D - x_1 - x_2 - x_4)$ (68)	$x_4(0) = 0$ (71)	
Eliminating x_4:		
$\dot{x}_1 = k_4(D - x_1 - x_2 - x_3) - r_4x_1 - k_1x_1(T_0 - x_2) + r_1x_2$ (73)	$x_1(0) = D$ (76)	
$\dot{x}_2 = k_1x_1(T_0 - x_2) - (r_1 + k_2)x_2$ (74)	$x_2(0) = 0$ (77)	$y = k \cdot (D - x_1)$ (79)
$\dot{x}_3 = r_3(D - x_1 - x_2 - x_3) - k_3x_3 + k_2x_2$ (75)	$x_3(0) = 0$ (78)	

It is possible to simplify the model of the form (80)–(82), by noting that adding (80)–(82) gives

$$\dot{x}_1 + \dot{x}_3 + \dot{x}_4 = 0 \tag{86}$$

Integrating (50) with respect to time on both sides yields

$$x_1 + x_3 + x_4 = D_1, \tag{87}$$

where D_1 is a constant. Re-arranging (87) and substituting into the model of the form (80)–(82) reduces it to two ordinary differential equations. The three variations, eliminating x_1, x_3, and x_4 respectively are shown in Table 4.

3.4. *Drug metabolism models*

Although it is suspected that there is minimal *in vitro* metabolism [22] the final candidate models developed also account for the drug metabolising within the cell. This involves adding an elimination term, k_4, to the third compartment, as shown in Fig. 3.

The corresponding set of nonlinear ordinary differential equations characterising the proposed model are given by,

$$\dot{x}_1 = k_3x_3 - r_3x_1 - k_1x_1(T_0 - x_2) + r_1x_2 \tag{103}$$

$$\dot{x}_2 = k_1x_1(T_0 - x_2) - (r_1 + k_2)x_2 \tag{104}$$

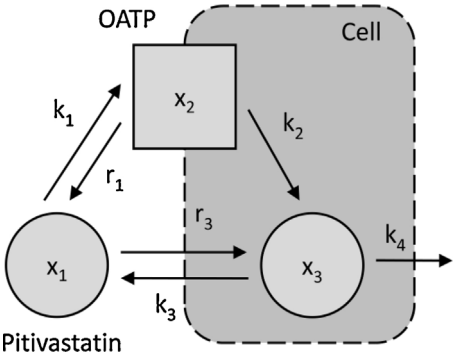


Fig. 3 – Metabolite model representation.

$$\dot{x}_3 = r_3x_1 - (k_3 + k_4)x_3 + k_2x_2 \tag{105}$$

with the initial conditions and observation are now given by

$$x_1(0) = D, \quad x_2(0) = 0, \quad x_3(0) = 0 \tag{106}$$

and observation

$$y = k \cdot (x_2 + x_3). \tag{107}$$

The unknown parameter set, p, is given by

$$p = \{k_1, r_1, k_2, k_3, r_3, k_4, T_0\}. \tag{108}$$

The same pseudo steady state assumption can also be made to obtain another representation of this model of the form

$$\dot{x}_1 = k_3x_3 - r_3x_1 - \frac{V_Mx_1}{K_M + x_1} \tag{109}$$

$$\dot{x}_3 = r_3x_1 - (k_3 + k_4)x_3 + \frac{V_Mx_1}{K_M + x_1} \tag{110}$$

where the initial conditions are

$$x_1(0) = D_1, \quad x_3(0) = 0, \tag{111}$$

and the observation is now given by

$$y = k \cdot \left(\frac{T_0x_1}{K_M + x_1} + x_3\right). \tag{112}$$

The unknown parameter set, p, is given by

$$p = \{V_M, K_M, k_3, r_3, k_4, r_4, T_0\}. \tag{113}$$

Table 4 – Alternate system equations for the model of the form (80)–(82).

System equation	Initial condition	Observation
Eliminating x_1:		
$\dot{x}_3 = r_3x_4 - k_3x_3 + \frac{V_M(D_1-x_3-x_4)}{K_M+D_1-x_3-x_4}$ (88)	$x_3(0)=0$ (90)	$y = k\cdot\left(\frac{T_0(D_1-x_3-x_4)}{K_M+D_1-x_3-x_4} + x_3 + x_4\right)$ (92)
$\dot{x}_4 = r_4(D_1-x_3-x_4) - (k_4+r_3)x_4 + k_3x_3$ (89)	$x_4(0)=0$ (91)	
Eliminating x_3:		
$\dot{x}_1 = k_4x_4 - r_4x_1 - \frac{V_Mx_1}{K_M+x_1}$ (93)	$x_1(0)=D_1$ (95)	$y = k\cdot\left(\frac{T_0x_1}{K_M+x_1} + D_1 - x_1\right)$ (97)
$\dot{x}_4 = r_4x_1 - (k_4+r_3)x_4 + k_3(D_1-x_1-x_4)$ (94)	$x_4(0)=0$ (96)	
Eliminating x_4:		
$\dot{x}_1 = k_4(D_1-x_1-x_3) - r_4x_1 - \frac{V_Mx_1}{K_M+x_1}$ (98)	$x_1(0)=D_1$ (100)	$y = k\cdot\left(\frac{T_0x_1}{K_M+x_1} + D_1 - x_1\right)$ (102)
$\dot{x}_3 = r_3(D_1-x_1-x_3) - k_3x_3 + \frac{V_Mx_1}{K_M+x_1}$ (99)	$x_3(0)=0$ (101)	

4. Techniques

4.1. A similarity transformation approach for uncontrolled systems (STAUS)

Given a linear model structure, this approach generates all the linear models that have the same input/output behaviour. It has also been successfully applied to non-linear models by mapping the state equations to a linear set [23]. Given a non-linear mathematical model of the following general form:

$$\dot{\boldsymbol{x}}(t,\boldsymbol{p}) = \boldsymbol{f}(\boldsymbol{x}(t,\boldsymbol{p}),\boldsymbol{p}) \tag{114}$$

$$\boldsymbol{x}(0,\boldsymbol{p}) = \boldsymbol{x}_0(\boldsymbol{p}) \tag{115}$$

$$\boldsymbol{y}(t,\boldsymbol{p}) = \boldsymbol{h}(\boldsymbol{x}(t,\boldsymbol{p}),\boldsymbol{p}) \tag{116}$$

where $\boldsymbol{p}$ is the r dimensional vector of unknown parameters. The n dimensional vector $\boldsymbol{x}(t,\boldsymbol{p})$ is the state vector, such that $\boldsymbol{x}_0(\boldsymbol{p})$ is the initial state and $\boldsymbol{y}(t,\boldsymbol{p})$ is the observation vector. For an autonomous system with no input, this approach initially entails establishing an Observability Rank Criterion (ORC). This is performed by defining a function $\boldsymbol{H}$ given by

$$\boldsymbol{H}(x,\boldsymbol{p}) = (\mu_1(x,\boldsymbol{p}),\ldots,\mu_n(t,\boldsymbol{p}))^T \tag{117}$$

where $\mu_1(\boldsymbol{x},\boldsymbol{p})$ is the observation function $\boldsymbol{h}$, and $\mu_n(\boldsymbol{x},\boldsymbol{p})$ is the *Lie derivative* of the previous term, given by

$$\mu_n(x,\boldsymbol{p}) = L_f\mu_{n-1}(x) = \frac{\partial\mu_{n-1}}{\partial x}(x)\cdot f(x) \tag{118}$$

where $\boldsymbol{h}$ is the observation from (114) and $\boldsymbol{f}$ the vector of the system coordinate functions given by (115). If the Jacobian matrix with respect to $\boldsymbol{x}$, evaluated at $\boldsymbol{x}_0(\boldsymbol{p})$, of the resultant function $\boldsymbol{H}(\cdot,\boldsymbol{p})$ is non singular, then the system (114)–(116) is said to satisfy the ORC and it is possible to construct a smooth mapping from the state corresponding to a parameter vector $\bar{\boldsymbol{p}}$, indistinguishable from $\boldsymbol{p}$, to the state corresponding to $\boldsymbol{p}$. For a particular $\boldsymbol{p}$, let $\boldsymbol{H}_p$ denote the vector field $\boldsymbol{H}(\cdot,\boldsymbol{p})$. According to Theorem 4 from [23], a smooth map λ is calculated using

$$\boldsymbol{H}_p(\lambda(x)) = \boldsymbol{H}_{\bar{p}}(x) \tag{119}$$

Equations can then be derived from the initial conditions $\boldsymbol{x}_0$, the model structure $\boldsymbol{f}$ and the observation function $\boldsymbol{h}$ by using:

$$\boldsymbol{\lambda}(\boldsymbol{x}_0(\bar{\boldsymbol{p}})) = \boldsymbol{x}_0(\boldsymbol{p}) \tag{120}$$

$$\boldsymbol{f}(\boldsymbol{\lambda}(\boldsymbol{x}(t,\bar{\boldsymbol{p}})),\boldsymbol{p}) = \frac{\partial\boldsymbol{\lambda}}{\partial\boldsymbol{x}}(\boldsymbol{x}(t,\bar{\boldsymbol{p}}))\boldsymbol{f}(\boldsymbol{x}(t,\bar{\boldsymbol{p}}),\bar{\boldsymbol{p}}) \tag{121}$$

$$\boldsymbol{h}(\boldsymbol{\lambda}(\boldsymbol{x}(t,\bar{\boldsymbol{p}})),\boldsymbol{p}) = \boldsymbol{h}(\boldsymbol{x}(t,\bar{\boldsymbol{p}}),\bar{\boldsymbol{p}}) \tag{122}$$

These are solved for $\boldsymbol{p}$ and the model is structurally globally identifiable (SGI) if $\boldsymbol{p} = \bar{\boldsymbol{p}}$ is the only solution, that is to say all the parameters are globally uniquely identifiable. A set of distinct solutions gives rise to a structurally locally identifiable (SLI) model. Otherwise the model is structurally unidentifiable (SU).

4.2. A sufficient condition for unidentifiability

As a straightforward consequence of the similarity transformation approach for uncontrolled systems described above the approach gives rise to a sufficient condition for unidentifiability [23]. In this instance, instead of calculating the smooth map λ from solving (119) using the observation vector field $\boldsymbol{H}$, the smooth map λ is assumed to be of the form

$$\lambda(x) = (t_1x_1,\ldots,t_nx_n)^T \tag{123}$$

where $t_i(\neq 0)\in\mathbb{R}$ and n is the number of states. Again the identities (120)–(122) are used to generate the relevant equations which are solved for $\bar{\boldsymbol{p}}$ and $t_1,\ldots,t_n$. If there are an (uncountably) infinite number of solutions, then the model is structurally unidentifiable (SU).

4.3. Differential algebra approach using characteristic sets (DAACS)

This approach consists of generating the input/output structure of the given model of the general form (114)–(116) solely in terms of the observation function y and its derivatives using characteristic sets [14,15]. This method requires the model to satisfy the ORC and is implemented using the Rosenfeld–Groebner algorithm in Maple 2010, which calculates a characteristic set for the model with a particular ranking of variables, where one member of the characteristic set gives the input/output map. A second input/output map is generated by substituting $\boldsymbol{p}$ for $\bar{\boldsymbol{p}}$ in the original map. If

equating the monomials of these two functions produces only one solution for the unknown parameters, then the system is SGI.

4.4. *Algebraic input/output relationship approach (Ai/oRA)*

This is the most recent approach, developed by [19]. Given a model of the general form (114)–(116) that satisfies the ORC, this approach generates the corresponding input/output map for the system. This approach requires calculating the *Lie derivatives* of the observation function, defined in (118). These are used as inputs into the Univariate Polynomial or Groebner Bases algorithms in Maple, producing the input/output relationship for the model. Again, a second input/output map is generated by substituting $\boldsymbol{p}$ for $\bar{\boldsymbol{p}}$ in the original input/output relationship. If equating the monomials of these functions produces only one solution for the unknown parameters, then the system is SGI.

4.5. *Non-differential input/output observable normal form approach (NDi/oONF)*

This approach also generates the input/output structure for the given model, solely in terms of *Lie derivatives* of the observation function, defined in (118). This is achieved using a co-ordinate transformation into the Observable Normal Form. Given a model of the general form (114)–(116) which satisfies the ORC, the *Lie derivatives* are calculated and solved simultaneously as a system of equations to obtain expressions for all the states in terms of these *Lie derivatives*. These are subsequently substituted into the derivative with respect to time of the highest *Lie derivative* to give an input/output map of the model. If equating the monomials of this function produces only one solution for the unknown parameters, then the system is SGI (See Appendix for example Maple code).

4.6. *Taylor series expansion*

This general method, introduced in [11], is commonly used for systems with a single input and can be applied to both linear and non-linear systems. The observation function is expanded as a Taylor series around the known initial condition. The Taylor series coefficients are measurable and unique for a particular output. Equating the Taylor series coefficients obtained from $y(t, \boldsymbol{p})$ with those derived from $y(t, \bar{\boldsymbol{p}})$ produces a system of equations. If there is only one solution for the unknown parameters, then the model is SGI. The total number of unknown model parameters determines the minimum number of Taylor series coefficients required to establish structurally identifiability and this causes significant computational problems in models with numerous unknown model parameters.

5. Results

All the candidate models derived satisfy the ORC and none can be shown to be unidentifiable using the sufficient condition for unidentifiability described above. The Taylor series expansion approach was also applied to all of the models to ascertain structural identifiability. However this method did not converge for any of the models proposed. The approach is intractable if it is not possible to obtain sufficient Taylor series coefficients or if the Taylor series coefficients prove too complex in structure to yield solutions for the unknown parameters using symbolic tools such as Maple 2010 or Mathematica 9 (Windows XP Pro 2002 SP3, Intel Quad CPU 1.98 GHz, 2.99GB of RAM). The results for the remaining approaches are summarised in Table 5.

The models of the form (1)–(3), (10) and (11), (16) and (17), and (21) and (22) can all be shown to be structurally globally identifiable (SGI) via all four techniques. Both the differential algebra approach using characteristic sets (DAACS) and the algebraic input/output relationship approach (Ai/oRA) produce the same input/output map for the four models and confirm that the models of the form (1)–(3), (10) and (11), (16) and (17), and (21) and (22) are structurally equivalent.

The similarity transformation approach for uncontrolled systems (STAUS), the non-differential input/output observable normal form approach (NDi/oONF), and the differential algebra approach using characteristic sets (DAACS) provide no further information about the remaining models. In all instances the complexity, in particular the non-linear terms and the fact that the observation is a sum of compartments, means that solutions prove intractable. It is suspected that there is not enough memory available for Maple 2010 to perform the required symbolic calculations.

The algebraic input/output relationship approach (Ai/oRA) proves to be the most successful on the hepatic uptake models derived. It can be shown that both models of the form (37) and of the form (40) are SGI and have the same input/output map, confirming that both models are indeed structurally equivalent. Although it is suspected the model of the form (28) and (29) has the same structure, Maple 2010 is unable to perform the required symbolic manipulations. The algebraic input/output relationship approach (Ai/oRA) requires as many *Lie derivatives* as there are states, hence reducing the model of the form (28) and (29) to a one state model significantly simplifies the computation.

The algebraic input/output relationship approach (Ai/oRA) also shows that the non specific binding models of the form (52)–(54), (59)–(61), (66)–(68), and (73)–(75) are all structurally locally identifiable. Again, all four models produce the same input/output map and the approach confirms that the models are structurally equivalent. There are two solutions for certain parameters, however T_0 and k_1 are globally identifiable. Again, although it is suspected the non specific binding model of the form (43)–(46) has the same structure, Maple 2010 is unable to perform the required symbolic manipulations. This provides another example where it proves very advantageous to reduce the system to its minimal form.

The algebraic input/output relationship approach (Ai/oRA) demonstrates that the models of the form (88)–(89) and (93)–(94) are structurally globally identifiable. Both models produce the same input/output map and the approach confirms that the models are structurally equivalent. Although it is suspected the model of the form (80)–(82) and (98)–(99) have the same structure, Maple 2010 is unable to perform the required symbolic manipulations. It is interesting that the approach works for models of the form (88)–(89) and (93)–(94) but not for the models of the from (98) and (99) as they each have two

Please cite this article in press as: T.R.B. Grandjean, et al., Structural identifiability analyses of candidate models for *in vitro* Pitavastatin hepatic uptake, Comput. Methods Programs Biomed. (2013), http://dx.doi.org/10.1016/j.cmpb.2013.06.013

Table 5 – Summary of the structural identifiability of all six candidate models considered using all four approaches (SGI: structurally globally identifiable; SLI: structurally locally identifiable; DNC: does not converge).

Approach	Basic						
	Original					Pseudo steady state assumption	
	(1)–(3)	(10)–(11)	(16)–(17)	(21)–(22)	(28)–(29)	(37)	(40)
4.1 STAUS	*SGI*	*SGI*	*SGI*	*SGI*	DNC	DNC	DNC
4.3 DAACS	*SGI*	*SGI*	*SGI*	*SGI*	DNC	DNC	DNC
4.5 Ai/oRA	*SGI*	*SGI*	*SGI*	*SGI*	DNC	*SGI*	*SGI*
4.6 NDi/oONF	*SGI*	*SGI*	*SGI*	*SGI*	DNC	DNC	DNC

Approach	Non specific binding				
	Original				
	(43)–(46)	(52)–(54)	(59)–(61)	(66)–(68)	(73)–(75)
4.1 STAUS	DNC	DNC	DNC	DNC	DNC
4.3 DAACS	DNC	DNC	DNC	DNC	DNC
4.5 Ai/oRA	DNC	*SLI*	*SLI*	*SLI*	*SLI*
4.6 NDi/oONF	DNC	DNC	DNC	DNC	DNC

Approach	Non specific binding				Metabolite	
	Pseudo steady state assumption				Original	Pseudo
	(80)–(82)	(88)–(89)	(93)–(94)	(98)–(99)	(103)–(105)	(109)–(110)
4.1 STAUS	DNC	DNC	DNC	DNC	DNC	DNC
4.3 DAACS	DNC	DNC	DNC	DNC	DNC	DNC
4.5 Ai/oRA	DNC	*SGI*	*SGI*	DNC	DNC	DNC
4.6 NDi/oONF	DNC	DNC	DNC	DNC	DNC	DNC

states and very similar model equations, suggesting that the success of the approach sometimes depends on which substitution is made. The authors suspect that the number of rational functions in the system equations is the reason for the varied success in this instance. Both models of the form (88)–(89) and (93)–(94) only have one equation that includes a rational function, (88) and (93) respectively, whereas the model of the from (98)–(99) has rational functions in both equations. This is re-enforced by noting that the models of the form (37) and (40) only include a single rational function each, whereas the model of the form (28) and (29) has one in each of the state space ordinary differential equations. Unfortunately this method does not converge for models of the form (103)–(105) and of the form (109)–(110) as Maple 2010 cannot calculate the required input/output relationship. Again, it is suspected that there is insufficient memory to perform the necessary symbolic calculations.

Note: For the pseudo steady models of the form (37) and of the form (40), which are both SGI, the unknown parameters given in (31), i.e. K_M, V_M, k_3, r_3, T_0 are structurally identifiable. Similarly for the pseudo steady models of the form (88)–(89) and of the form (93)–(94), which are both SGI, the unknown parameters given in (85), i.e. K_M, V_M, k_3, r_3, k_4, r_4, T_0 are structurally identifiable. In both instances, we can infer from (30) that k_2 is also structurally identifiable as both V_M and T_0 are structurally identifiable, whereas k_1 and r_1 are not identifiable. However in these four models, structural identifiability of the micro parameters k_1, k_2, and r_1 is not required for the models to be SGI as they are replaced by the macro parameters K_M and V_M, and do not appear in the model structure.

All of the analyses above were also duplicated in Mathematica 9 which yielded the same results as Maple 2010.

6. Conclusions

All of the candidate models derived satisfy the ORC and none satisfy the condition for unidentifiability described above. The traditional approach, namely the Taylor Series expansion approach proves intractable for all of the models. All of the candidate models have been shown to be SGI, except the models of the form (52)–(54), (59)–(61), (66)–(68) and (73)–(75) which are SLI and the models of the form (28)–(29), (43)–(46), (80)–(82), (98)–(99), (103)–(105), and (109)–(110) where none of the techniques used converged to a solution. The algebraic input/output relationship approach was the most successful technique of those considered, leading to conclusive results for 12 out of the 18 candidate models. The approach was found to be most successful when systems were reduced to their minimal form and the number of equations containing rational functions was reduced as far as possible. The similarity transformation approach for uncontrolled systems, the non-differential input/output observable normal form approach, and the differential algebra approach using characteristic sets were only successfully applied to four of the models considered. The successful application of the approaches may vary for different models. Other approaches were also used, but failed to produce solutions. The direct test [25] is not applicable for the models described here as it requires measuring at least one compartment directly. DAISY [24] is another differential algebra software tool, which was unable to solve the

Please cite this article in press as: T.R.B. Grandjean, et al., Structural identifiability analyses of candidate models for *in vitro* Pitavastatin hepatic uptake, Comput. Methods Programs Biomed. (2013), http://dx.doi.org/10.1016/j.cmpb.2013.06.013

algebraic equations generated. GenSSI [26] produces identifiability tableaus and reduces them a number of times but fails to solve the remaining equations.

Parameter estimation using real rat, dog, and human data from AstraZeneca has also been performed [27] and reveals that the models of the form (37) and of the form (40) fit the data very well. For some of the other models (i.e. the model of the form (1)–(3)) some of the parameters are not numerically identifiable as the data are quite sparse.

Further models incorporating both non-specific binding and *in vitro* metabolism were considered, however none of the approaches have yet to yield any conclusive results with regards to structural identifiability and this analysis is ongoing for these models.

Appendix A. Appendix

A.1. *Performing structural identifiability analysis in Maple 2010*

The Maple code used to perform the structural identifiability analysis is very similar to the code proposed by Evans *et al.* [19]. It has been adjusted to be able to cope with rational functions. The procedures use short command names and so the following packages are loaded:

with(LinearAlgebra): with(Groebner):

The following procedure, *lieDer*, determines the *Lie derivative* of *H* along *F*:

```
lieDer:= proc (H, F, vars)
local N, V;
N:= Dimension(F);
V:= map((a, b) -> diff(b, a), vars, H):
DotProduct(F, V, conjugate = false)
end:
```

while the next procedure, *listLieDer*, determines the list of the first *N Lie derivatives* of *H* along *F*:

```
listLieDer:= proc (H, F, N, vars)
local L, i, tmp;
L:= [denom(H)·y[0]-numer(H)];
tmp:= H;
for i to N do
tmp:= lieDer(tmp, F, vars);
L:= [op(L), denom(tmp)·y[i]-numer(tmp)]
od;
end:
```

Evans *et al.* [19] approach to determining the output equation using Grobner Bases is implemented by the following procedure.

```
outptEqn:= proc (F, H, vars)
local N, L, vars2;
N:= Dimension(F);
L:= listLieDer(H, F, N, vars);
L:= map(expand, L);
Vars2:= [op(vars),y[N]];
UnivariatePolynomial(y[N], L, vars2)
end:
```

Taking the model of the form (40) as an example, the identifiability analysis is performed by first defining the model equation(s) (*F*), the observation (*H*), and the state(s) (*vars*) then determining the output equation (*out*):

```
F:= Vector(r[3]·x[1] − k[3]·(D[1] − x[1]) + V[M]·x[1]/(K[M] + x[1])):
H:= T[0]·x[1]/(K[M] + x[1]) + D[1] − x[3]:
vars:= [x[1]]:
out:= outptEqn(F,H,vars);
```

Next the coefficients of the output equation (*u*) are extracted:

```
yout:= {seq(y[t],t = 0.Dimension(F))}:
p:= collect(out,yout,'distributed'):
u:= {coeffs(p,yout)}:
```

And a second set of coefficients (*ub*) is generated:

```
ub:= eval(u,[K[M] = Kb[M], V[M] = Vb[M], k[3] = kb[3], r[3] = rb[3],
  T[0] = Tb[0], D[1] = Db[1]]):
```

Finally, equating the two sets of coefficients and solving for the unknown parameters:

```
eqns:= convert(u,list) − convert(ub,list):
sol:= solve(eqns, {Kb[M],Vb[M],kb[3],rb[3],Tb[0],Db[1]});
```

The only solution is $\{Kb[M] = K[M], Vb[M] = V[M], kb[3] = k[3], rb[3] = r[3], Tb[0] = T[0], Db[1] = D[1]\}$, which implies that the system is Structurally Globally Identifiable.

REFERENCES

[1] S. Shimada, H. Fujino, T. Morikawa, M. Moriyasu, J. Jojima, Uptake mechanism of pitavastatin, a new inhibitior of HMG-CoA reductase, in rat hepatocytes, Drug Metabolism & Pharmacokinetics 18 (4) (2003) 245–251.
[2] M. Hirano, K. Maeda, Y. Shitara, Y. Sugiyama, Contribution of OATP2 (OATP1B1) and OATP8 (OATP1B3) to the hepatic uptake of of pitavastatin in humans, Journal of Pharmacology and Experimental Therapeutics 311 (2004) 139–146.
[3] A.M. Hassen, D. Lam, M. Chiba, E. Tan, M. Geng, K. Sandy Pang, Uptake of sulfate conjugates by isolated rat hepatocytes, Drug Metabolism and Disposition 24 (1996) 792–798.
[4] J.A. Jacquez, Compartmental Analysis in Biology and Medicine, BioMedware, Ann Arbor, MI, 1996.
[5] R. Bellman, K.J. Åström, On structural identifiability, Mathematical Biosciences 7 (1970) 329–339.
[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.
[7] Miao Hongyu, Xia Xiaohua, S. Perelson Alan, Wu Hulin, On identifiability of nonlinear ODE models and applications in viral dynamics, SIAM Review archive 53 (February (1)) (2011) 3–39.
[8] O.-T. Chis, J.R. Banga, E. Balsa-Canto, Structural identifiability of systems biology models: a critical comparison of methods, PLoS ONE 6 (11) (2011) e27755.
[9] T.R.B. Grandjean, M.J. Chappell, J.W.T. Yates, K. Jones, G. Wood, T. Coleman, Compartmental modelling of the pharmacokinetics of a breast cancer resistance protein, Computer Methods and Programs in Biomedicine 104 (2) (2011) 81–92, 0169-2607.
[10] E. Walter, Identifiability of Parametric Models, Pergamon, Oxford, 1987.
[11] H. Pohjanpalo, System identifiability based on the power series expansion of the solution, Mathematical Biosciences 41 (1978) 21–33.
[12] E.T. Tunali, T.J. Tarn, New results for identifiability of nonlinear systems, IEEE Transactions on Automatic Control 32 (1987) 146–154.

[13] S. Vajda, K.R. Godfrey, H. Rabitz, Similarity transformation approach to identifiability analysis of nonlinear compartmental models, Mathematical Biosciences 93 (1989) 217–248.
[14] 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, Progress in Systems and Control Theory, vol. 14, Boston, Birkhaüser, 1993.
[15] L. Ljung, T. Glad, On global identifiability for arbitrary model parametrizations, Automatica 30 (1994) p265–p276.
[16] J.E. Jiménez-Hornero, I.M. Santos-Dueñas, A.I. Garci A-Garci, Structural identifiability of a model for the acetic acid fermentation process, Mathematical Biosciences 216 (December) (2008) 154–162 (Epub 2008 September 27).
[17] N. Meshkat, M. Eisenberg, J.J. Distefano 3rd, An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner Bases, Mathematical Biosciences 222 (2009) p61–p72.
[18] N.D. Evans, M.J. Chapman, M.J. Chappell, K.R. Godfrey, Identifiability of uncontrolled nonlinear rational systems, Automatica 38 (2002) 1799–1805.
[19] N.D. Evans, H. Moyse, D. Lowe, D. Briggs, R. Higgins, D. Mitchell, D. Zehnder, M.J. Chappell, Structural identifiability of surface binding reactions involving heterogeneous analyte: application to surface plasmon resonance experiments, Automatica 49 (2013) 48–57.
[20] Maple (2010 – version 14) Technical reference, Maplesoft, Waterloo, ON, Canada, 2010.
[21] B. Obrink, B. Waemegaard, H. Pertoft, Specific binding of rat liver plasma membranes by rat liver cells, Biochemical and Biophysical Research Communications 77 (1977) 665–670.
[22] H. Fujino, I. Yamada, J. Kojima, M. Hirano, H. Matsumoto, M. Yoneda, Studies on the metabolic fate of NK-104, a new inhibitor HMG-CoA reductase (5): in vitro metabolism and plasma protein binding in animals and human, Xenobiotics and Metabolic Disposition, 14 (1999) 415–424.
[23] N.D. Evans, L.J. White, M.J. Chapman, K.R. Godfrey, M.J. Chappell, The structural identifiability of the susceptible infected recovered model with seasonal forcing, Mathematical Biosciences 194 (2005) 175–197.
[24] G. Bellu, M.P. Saccomani, S. Audoly, L. D'Angio, DAISY: a new software tool to test global identifiability of biological and physiological systems, Computer Methods and Programs in Biomedicine Vol88 (2007) 52–61.
[25] L. Denis-Vidal, G. Joly-Blanchard, An easy to check criterion for (un)identifiability of uncontrolled systems and its applications, IEEE Transactions on Automatic Control 45 (2000) 768–771.
[26] E. Walter, Y. Lecourtier, Global approaches to identifiability testing for linear and nonlinear state space models, Mathematics and Computers in Simulation 24 (1982) 472–482.
[27] T.R.B. Grandjean, M.J. Chappell, J.W.T. Yates, A.L. Lench, C.J. O'Donnell, Experimental and mathematical analysis of *in vitro* pitavastatin hepatic uptake across species, Drug Metabolism & Disposition (2013) (submitted for publication).

Please cite this article in press as: T.R.B. Grandjean, et al., Structural identifiability analyses of candidate models for *in vitro* Pitavastatin hepatic uptake, Comput. Methods Programs Biomed. (2013), http://dx.doi.org/10.1016/j.cmpb.2013.06.013