BMC
Systems Biology

METHODOLOGY ARTICLE

Steady-state global optimization of metabolic non-linear dynamic models through recasting into power-law canonical models

Carlos Pozo[1], Alberto Marín-Sanguino[2], Rui Alves[3], Gonzalo Guillén-Gosálbez[1], Laureano Jiménez[1] and Albert Sorribas[3*]

Abstract

Background: Design of newly engineered microbial strains for biotechnological purposes would greatly benefit from the development of realistic mathematical models for the processes to be optimized. Such models can then be analyzed and, with the development and application of appropriate optimization techniques, one could identify the modifications that need to be made to the organism in order to achieve the desired biotechnological goal. As appropriate models to perform such an analysis are necessarily non-linear and typically non-convex, finding their global optimum is a challenging task. Canonical modeling techniques, such as Generalized Mass Action (GMA) models based on the power-law formalism, offer a possible solution to this problem because they have a mathematical structure that enables the development of specific algorithms for global optimization.

Results: Based on the GMA canonical representation, we have developed in previous works a highly efficient optimization algorithm and a set of related strategies for understanding the evolution of adaptive responses in cellular metabolism. Here, we explore the possibility of recasting kinetic non-linear models into an equivalent GMA model, so that global optimization on the recast GMA model can be performed. With this technique, optimization is greatly facilitated and the results are transposable to the original non-linear problem. This procedure is straightforward for a particular class of non-linear models known as Saturable and Cooperative (SC) models that extend the power-law formalism to deal with saturation and cooperativity.

Conclusions: Our results show that recasting non-linear kinetic models into GMA models is indeed an appropriate strategy that helps overcoming some of the numerical difficulties that arise during the global optimization task.

1 Background

Identifying optimization strategies for increasing strain productivity should be possible by applying optimization methods to detailed kinetic models of the target metabolism. Thus, a rational approach would pinpoint the changes to be done - e.g. by modulating gene expression - in order to achieve the desired biotechnological goals [1-4]. To build such models we can either start from a detailed description of the underlying processes (bottom-up strategy) or we can fit kinetic models to experimental data obtained *in vivo* (top-down strategy).

The bottom-up approach was the original strategy for model building in the biological sciences. Bottom-up kinetic models require information that is seldom available, despite the increasing amount of kinetic data contained in a growing set of databases (for example see [5,6] and the webpage http://kinetics.nist.gov/kinetics/index.jsp). Even in the best case scenarios where kinetic data are available, the data have often been obtained in different labs and under *in vitro* conditions that are hardly ever comparable or representative of the situation *in vivo*. In addition, models built using this strategy often fail to adequately reproduce the known behavior of the target system [7-10]. With the accumulation of time-series data, which were originally generated from the accurate measurement of transient responses, top-down modeling became viable as an alternative to the

* Correspondence: albert.sorribas@cmb.udl.cat
[3]Departament de Ciències Mèdiques Bàsiques, Institut de Recerca Biomèdica de Lleida (IRBLLEIDA), Universitat de Lleida, Montserrat Roig 2, 25008 Lleida, Spain
Full list of author information is available at the end of the article

bottom-up strategy [11]. However, top-down modeling also faces important difficulties. For example, regulatory interactions between metabolites and enzymes are very poorly characterized and most metabolic maps lack such crucial information. Therefore, for a given network structure (i.e. a stoichiometric description) obtained from databases, a large number of alternative regulatory patterns may exist that account for the observed experimental data [12]. Model discrimination among the alternative regulatory patterns requires appropriate experimental design. However, this is seldom considered when performing the time series measurements. Last, but not the least, parameter identifiability in highly non-linear models can be problematic (for a review see [13]).

An additional issue that is common to models built using both strategies is that such detailed kinetic models include non-convexities that lead to the existence of multiple local optima in which standard non-linear optimization algorithms may get trapped during the search. Several stochastic and deterministic global optimization methods have been proposed to overcome this limitation [14]. Deterministic methods, which are the only ones that can rigorously guarantee global optimality, rely on the use of convex envelopes or underestimators to formulate lower-bounding convex problems that are typically combined with spatial branch and bound strategies. Most of these methods are general purpose and assume special structures in the continuous terms of the mathematical model. Because of this, they can encounter numerical difficulties in specific metabolic engineering systems that require the optimization of kinetic models with a large number of non-convexities of different nature.

Given all these issues, it is hardly surprising that linear stoichiometric models have emerged as the most popular tool to analyze genome-wide metabolic networks using optimization techniques. Linear optimization problems can be solved using very fast and efficient algorithms [15,16] that are implemented in almost every kind of computer, ranging from laptops to cloud computing centers. In addition, such models require a relatively small amount of information.

The possibility of condensing information about a very large network in a compact form enabled stoichiometric models to provide interesting insights in many different cases. However, the apparent simplicity in building and analyzing stoichiometric models comes at the cost of neglecting regulatory signals, metabolite levels and dynamic constraints. Accounting for these features in a dynamic way requires using more detailed, non-linear, mathematical models [17,18].

These models go a step further than stoichiometric models by incorporating regulatory influences through a set of ordinary differential equations that can account for the system's dynamics. Building such models is often impossible because the appropriate functional form that needs to be used to describe the dynamical behavior of specific processes is in general unknown. Modeling strategies based on systematic approximated kinetic representations, such as power-laws [19-22], Saturating and Cooperative [23], or convenience kinetics [24], side-step this issue by providing uniform forms that are guaranteed to be accurate over a range of conditions and reduce the amount of information required to build the models. Because of the regularity in the form of the mathematical function, models based on approximate formalisms can be automatically built from the reaction scheme of the target system. The model parameters can subsequently be estimated from experimental data using different procedures [13,25].

Although building and analyzing of comprehensive genome-wide detailed models is still not viable in most cases (see however [26,27]), developing ways to extend large scale optimization analysis to larger and more realistic non-linear kinetic models is an important part of the future of systems biology [18]. In fact, the optimization of certain types of non-linear problems can already be solved very efficiently and geometric programming problems with up to 1,000 variables and 10,000 constraints can be solved in minutes on a personal computer.

Efficient global optimization techniques are available for power-law models [1,28-30], either in S-system form or in Generalized Mass Action (GMA) form (for a review see [31]). In the case of S-system models, a simple logarithmic transformation brings the model to a linear form [1]. In the case of GMA models, the problem can be efficiently solved using branch-and-bound [28,32] and outer-approximation techniques [29,30].

The usefulness of the global optimization techniques developed for GMA models has been shown in the analysis of the adaptive response of yeast to heat shock [29,33]. In essence, starting with a GMA model and considering a set of constraints on flux and metabolite values, we can obtain: (i) The pattern of enzyme activities that maximizes a given objective, (ii) The region of feasible changes in enzyme activities so that the model fulfills a set of constraints on fluxes, metabolites, maximum allowable change in activity, etc., and (iii) A heat map of how the objective function changes within the feasible region. These results share some similarities with those produced with stoichiometric models, but incorporate many additional features.

Based on ideas similar to those that led to the development of the power-law formalism, Sorribas et al. [23] proposed a new Saturable and Cooperative (SC) formalism, that extends the power-law representation to include cooperativity and saturation. Although models

built using this new formalism loses some of the simplicity inherent to the analysis of S-systems and GMA models, they tend to be accurate over a wider range of conditions than both the S-System and GMA representations [23]. Thus, it is important to enlarge the scope of global optimization methods developed for power-law models in order to deal with the SC formalism and analyze under which situations the later models behave better than the former.
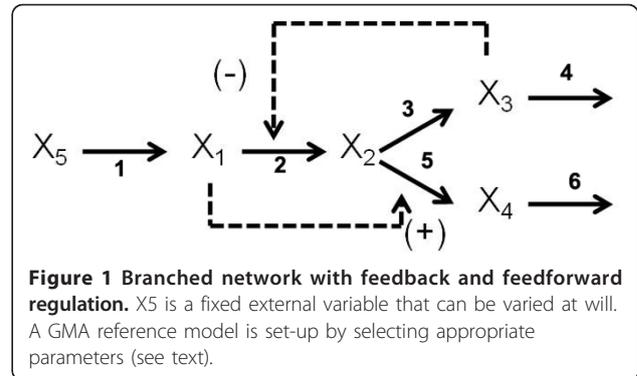
Optimization of SC models faces a number of practical problems common to kinetic non-linear models [34,35]. To sidestep these problems, and in order to be able to use global optimization methods developed for power-law models, we will use a technique called recasting. Recasting permits the exact transformation of a continuous non-linear model with an arbitrary form into a canonical GMA model [36,37]. This transformation is typically performed by increasing the number of variables of the original model. Through this technique, arbitrary non-linear models can be represented using a canonical form such as GMA or S-system that can be used for simulation and optimization purposes, which opens the door for effectively extending the optimization and feasibility analysis originally devised for GMA models to other detailed kinetic models.

In this paper, and as a first step to define a framework for optimization of non-linear models with arbitrary form and extend FBA and related approaches to detailed kinetic models, we shall show the practical utility of recasting SC models into GMA models for optimization purposes. This technique is similar to the symbolic reformulation algorithm proposed by Smith and Pantelides [38]. Our method, however, focuses on obtaining a power-law representation that greatly facilitates global optimization, instead of continuing with the recasting until converting the model to a standard form containing linear constraints and a set of nonlinearities corresponding to bilinear product, linear fractional, simple exponentiation and univariate function terms. After recasting the model to the canonical form, we can apply any of the optimization strategies we have presented for GMA models [29,32] to obtain the global optimum of the original SC problem.

2 Results

2.1 Global optimization of non-linear models through recasting

For a proof of concept of the difficulties of global optimizing non-linear models and of the use of recasting for attaining practical solutions, we shall start by defining a reference biochemical network that corresponds to the reaction scheme in Figure 1. This hypothetical system has a source metabolite X_5 and four internal metabolites. The network includes six reactions and a branch

Figure 1 Branched network with feedback and feedforward regulation. X_5 is a fixed external variable that can be varied at will. A GMA reference model is set-up by selecting appropriate parameters (see text).

point. X_3 acts as a feed-back inhibitor of the synthesis of X_2, while X_1 is an activator of the synthesis of X_4.

The generic model for this system is:

$$
\begin{aligned}
\dot{X}_1 &= v_1 - v_2 \\
\dot{X}_2 &= v_2 - v_3 - v_5 \\
\dot{X}_3 &= v_3 - v_4 \\
\dot{X}_4 &= v_5 - v_6
\end{aligned}
\tag{1}
$$

Each of the velocities is a non-linear function of the involved metabolites. The SC representation, provides a systematic way for defining a functional model of this pathway. As a demonstrative example, let us suppose that the numerical model is:

$$
\begin{aligned}
\frac{dX_1}{dt} &= \frac{20k_1 X_5^1}{X_5^1 + 1} - \frac{40k_2 X_1^1}{\left(X_1^1 + 1\right) X_3^2 \left(1 + \frac{1}{X_3^2}\right)} \\[2ex]
\frac{dX_2}{dt} &= \frac{40k_2 X_1^1}{\left(X_1^1 + 1\right) X_3^2 \left(1 + \frac{1}{X_3^2}\right)} - \frac{7.5k_3 X_2^{2.5}}{X_2^{2.5} + 0.25} \\[2ex]
&\quad - \frac{16k_5 X_1^1 X_2^1}{\left(X_1^1 + 1\right)\left(X_2^1 + 1\right)} \\[2ex]
\frac{dX_3}{dt} &= \frac{7.5k_3 X_2^{2.5}}{X_2^{2.5} + 0.25} - \frac{12k_4 X_3^1}{X_3^1 + 1} \\[2ex]
\frac{dX_4}{dt} &= \frac{16k_5 X_1^1 X_2^1}{\left(X_1^1 + 1\right)\left(X_2^1 + 1\right)} - \frac{8k_6 X_4^1}{X_4^1 + 1}
\end{aligned}
\tag{2}
$$

In these equations, k_r, $r = 1,..,6$ is an auxiliary variable used to model changes in the enzyme activity. At the basal level, $k_r = 1$ for all the reactions. During the optimization tasks, it is possible to limit the maximum change in gene expression by imposing a maximum allowable change in k_r.

We shall now address the following questions:

(i) To what extent can general purpose global optimization methods be applied to SC models?, (ii) Given that a SC model can be recast as a GMA (rGMA), is

this useful for optimization of the original SC model?, (iii) Are the results obtained with the rGMA equivalent to the results of the original SC model?, and (iv) What are the practical advantages of optimizing a rGMA model?.

2.2 Optimization goals

In order to address the questions posed at the end of the previous section we shall define the following optimizations tasks (note that changes in enzyme activities and metabolite concentrations are constrained between $0.2 \leq k_r \leq 5.0$ and $0.1 \leq X_i \leq 10.0$ respectively in all the instances unless otherwise specified):

- O1: What is the optimal pattern of changes in enzyme activities that maximizes the objective function in the new steady-state for a fixed value of X_5?
- O2: What is the optimal pattern of changes in enzyme activities that maximizes the objective function in the new steady-state for a fixed value of X_5 considering a maximum allowable variation of 10% in the steady-state values of the intermediaries?
- O3: What is the optimal pattern of changes in enzyme activities that maximizes the objective function in the new steady-state for a fixed value of X_5 considering changes in the output flux from X_4 of less than 10% with respect to its reference value?
- O4: What is the best set of changes, assuming that we can only manipulate three enzymes, that maximizes the objective function in the new steady-state for a fixed value of X_5 considering a maximum variation of 10% in the steady-state values of the intermediaries?

Two different objective functions (OF), steady-state concentration of X_3 and flux v_4, have been considered for each optimization case, except for O3. This latter case has been optimized in terms only of the first objective (i.e., steady-state concentration of X_3), because limits on v_4 are already included in the formulation of the optimization problem.

2.3 Global optimization of SC models using BARON

We first address the optimization of the aforementioned model in their original SC form using state of the art global optimization techniques. The model was coded in the algebraic modeling system GAMS 23.0.2 and solved with the commercial global optimization package BARON v.8.1.5. on an Intel 1.2 GHz machine. An optimality gap (i.e., tolerance) of 0.2% was set in all the instances. As can be seen in Table 1, BARON produce results with an optimality gap (OG) below the specified tolerance.

Table 1 Results for the maximization of X_3 and v_4 and optimization goals O1-O4 using BARON v.8.1.5. for a tolerance of 0.2%

O	k_1	k_2	k_3	k_4	k_5	k_6	X_3	OG (%)	CPU (s)
1	0.26	5.00	4.97	0.20	0.20	0.54	8.30	0.20	136.17
2	0.20	0.24	0.22	0.20	0.21	0.20	1.10	0.00	0.06
3	0.60	5.00	5.00	0.53	0.20	0.27	5.39	0.20	96.39
4	0.99	1.15	1.00	0.96	1.00	1.00	1.10	0.00	1.42

O	k_1	k_2	k_3	k_4	k_5	k_6	v_4	OG (%)	CPU (s)
1	4.61	5.00	5.00	5.00	0.72	1.20	37.40	0.20	157.83
2	3.22	3.73	5.00	4.99	0.21	0.22	31.33	0.00	1.67
3	0.88	0.94	0.88	0.96	0.23	3.00	6.60	0.00	10.53
4	1.16	1.00	1.34	1.34	1.00	1.00	7.61	0.00	3.61

Table 1 only shows one solution for each particular instance. However, BARON identified in each case a set of equivalent optima (i.e, solutions with the same objective function value) involving different changes in enzyme activities, which indicates that the optimization problem is somehow degenerated. This redundancy is a consequence of the system's structure and has practical implications. As an example, we have calculated some of these equivalent points for case O1-v_4 using the *NumSol* option of BARON (see Figure 2). In particular, a well defined triangular region containing the changes in k_2 and k_5, and k_1 and k_2 that lead to the same objective function value is identified. Within these regions, one can decide which combination of changes should be selected based on additional cost arguments, as they all show the same performance in terms of the predefined objective function. This region could be further reduced by imposing additional constraints to the optimization.

2.4 Recasting SC models into GMA models

Any SC model can be recast into a GMA canonical model by introducing the auxiliary variables $z_{rj} = K_{rj} + X_j^{n_{rj}}$. Substitution and differentiation generates the following recast GMA (rGMA) model:

$$\dot{X}_i = \sum_{r=1}^{p} \mu_{ir} V_r \prod_{j=1}^{n+m} X_j^{n_{rj}} z_{rj}^{-1} \qquad i = 1, .., n \qquad (3a)$$

$$\dot{z}_{rj} = n_{rj} X_j^{n_{rj}-1} \dot{X}_j \qquad \begin{aligned} r &= 1, .., p \\ j &= 1, ..., n+m \end{aligned} \qquad (3b)$$

with appropriate initial conditions $X_j(0) = X_{j_0}$ and $z_{rj}(0) = K_{rj} + X_{j_0}^{n_{rj}}$.

For simulation purposes, model (3) is equivalent to the original SC model. As discussed in [36], a model recast into a GMA model has the same steady-state that the original non-linear model. The steady-state equations of the rGMA model can be expressed as:
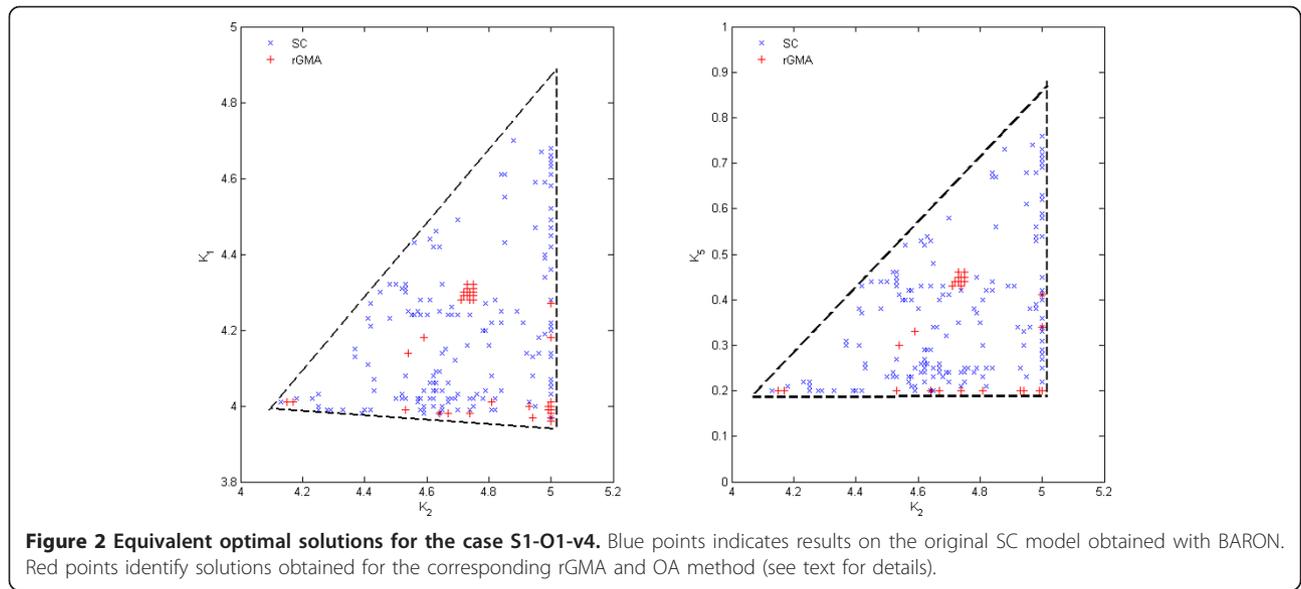
Figure 2 Equivalent optimal solutions for the case S1-O1-v4. Blue points indicates results on the original SC model obtained with BARON. Red points identify solutions obtained for the corresponding rGMA and OA method (see text for details).

$$\sum_{r=1}^{p} \mu_{ir} V_r \prod_{j=1}^{n+m} X_j^{n_{rj}} z_{rj}^{-1} = 0 \quad i = 1, .., n \tag{4a}$$

$$n_{rj} X_j^{n_{rj}-1} \dot{X}_j = 0 \qquad \begin{array}{l} r = 1, .., p \\ j = 1, ..., n+m \end{array} \tag{4b}$$

2.5 Steady-state optimization of SC models through recasting

The steady-state solutions of Eqn. (4b) satisfy also Eqn. (4a). Thus, for optimization purposes, the steady-state constraints of interest are:

$$\sum_{r=1}^{p} \mu_{ir} V_r \prod_{j=1}^{n+m} X_j^{n_{rj}} z_{rj}^{-1} = 0 \quad i = 1, .., n \tag{5a}$$

$$K_{rj} + X_{j_0}^{n_{rj}} = z_{rj_0} \qquad \begin{array}{l} r = 1, .., p \\ j = 1, ..., n+m \end{array} \tag{5b}$$

According to these results, the optimization problem can be stated as:

$$
\begin{aligned}
&min - OF && OF = \{X_i \text{ or } v_r\} \\
&s.t. \\
&\sum_{r=1}^{p} \mu_{ir} k_r v_r \prod_{J=1}^{n+m} X_j^{n_{rj}} z_{rj}^{-1} = 0 && \begin{array}{l} i = 1, .., n \\ j = 1, ..., n+m \end{array} \\
&K_{rj} + X_{j_0}^{n_{rj}} = z_{rj_0} && \begin{array}{l} r = 1, .., p \\ j = 1, ..., n+m \end{array} \\
&X_{iL} \leq X_i \leq X_{iU} && i = 1, .., n \\
&k_{rL} \leq k_r \leq k_{rU} && r = 1, .., p \\
&..... \text{ additional constraints } ....
\end{aligned}
\tag{6}
$$

In our reference model, we shall consider the following constraints:

$$
\begin{aligned}
&min - OF && OF = \{X_3, v_4\} \\
&s.t. \\
&\sum_{r=1}^{p} \mu_{ir} k_r v_r \prod_{J=1}^{n+m} X_j^{n_{rj}} z_{rj}^{-1} = 0 && i = 1, .., n \\
&K_{rj} + X_{j_0}^{n_{rj}} = z_{rj_0} && \begin{array}{l} r = 1, .., p \\ j = 1, ..., n+m \end{array}
\end{aligned}
$$

Specific constraints for each optimization task
(O1, O3 only)

$$0.1 \leq X_i \leq 10 \qquad i = 1, ..., n$$

(O1, O2, O3 only)

$$0.2 \leq k_r \leq 5 \qquad r = 1, ..., p$$

(O2, O4 only)

$$0.9 X_i^{BAS} \leq X_i \leq 1.1 X_i^{BAS} \qquad i = 1, ..., n \tag{7}$$

(O3 only) and (OF : X_3 only)

$$0.9 v_4^{BAS} \leq v_4 \leq 1.1 v_4^{BAS}$$

(O4 only)

$$
\begin{aligned}
&k_r = k_{r1} + k_{r2} + k_{r3} && r = 1, ..., p \\
&k_r^{LB} \gamma_{r1} \leq k_{r1} \leq (1 - \delta) \gamma_{r1} && r = 1, ..., p \\
&(1 - \delta) \gamma_{r2} \leq k_{r2} \leq (1 + \delta) \gamma_{r2} && r = 1, ..., p \\
&(1 + \delta) \gamma_{r3} \leq k_{r3} \leq k_r^{UB} \gamma_{r3} && r = 1, ..., p \\
&\gamma_{r1} + \gamma_{r2} + \gamma_{r3} = 1 && r = 1, ..., p \\
&\sum_{r=1}^{p} \gamma_{r1} + \sum_{r=1}^{p} \gamma_{r3} \leq ME = 3
\end{aligned}
$$

Once the problem has been recast into a rGMA, its mathematical structure can be exploited in order to improve the efficiency of the solution procedure, as demonstrated by the authors in previous works. This problem has a GMA form except for the auxiliary constraint 5b, which is required to recast the SC into the rGMA. This constraint can be easily handled by means of relaxation techniques and exponential transformations similar to those used by the authors in their global optimization algorithms for pure GMA models [32,33]. In particular, two algorithms were developed for the global optimization of GMA models: a customized outer-approximation (OA, [30]) and a tailored spatial branch-

and-bound (sBB, [32]). The authors showed that the numerical performance of these methods depends on the specific problem being solved, and that none of them is clearly better than the other one. Here, we use the OA algorithm to solve 6, as this method proved to be faster than sBB for problems of smaller size ([32]). Again, the main body of the algorithm was coded in GAMS 23.0.2, using CPLEX 11.2.1 as MILP solver for the master subproblems and CONOPT 3.14 s as NLP solver for the slave subproblems of the algorithm. For a fair comparison, we also set a tolerance of 0.2%, the same as when using BARON.

As can be seen in Table 2, the optimization of the rGMA formulation using our customized OA yields similar results to those obtained when BARON is applied to the original SC model. In some cases, significant reductions in computational time are attained with our OA algorithm. While BARON took a total time of 407.68 CPU seconds for solving the 8 instances, the customized OA algorithm solved the same problems in 8.5 CPU seconds.

Note that the objective function values obtained with the SC and rGMA models only differ within the tolerance imposed. In some cases, discrepancies regarding the enzymatic profiles calculated are observed mainly due to the system's structure, that is, to the fact that the problem contains multiple solutions attaining the same performance in terms of objective function value but involving different enzymatic configurations, as discussed in section 2.3.

To further investigate this issue, we apply the multi-solution capability of BARON to the rGMA model (Figure 2). Again, different equivalent optima are obtained, but this time the dispersion of the equivalent solutions generated for a given case tend to concentrate either in the center or in the extremes of the region containing the solutions with the same objective function value calculated with the SC model.

The region illustrated in Figure 2 should not be misunderstood as a feasibility region. In fact, solutions do exist outside this region, but they lead to worse objective function values. To further clarify this issue, we consider a grid of values for k_2 and k_5 in the region defined by constraints $4 \leq k_2 \leq 5$ and $0.2 \leq k_5 \leq 0.8$, and solve the optimization problem within each cell applying BARON to the SC model, and our OA to the rGMA model. Recall that these linear constraints define a region that contains that in Figure 2. The results obtained in this optimization are illustrated in Tables 3 and 4, and are exactly equal for both methods. However, the CPU time is much lower when using our OA algorithm applied to rGMA (11,811 CPU seconds for generating all the points with BARON applied to the SC model vs 17 CPU seconds with the customized OA applied to the rGMA model; as shown in Tables 5 and 6).

2.6 Difficult optimization tasks can be solved via recasting

The reference model can be optimized either by general purpose techniques or by rGMA specific methods such as the customized OA. However, even with this simple

Table 2 Results for the maximization of X_3 and v_4 using the rGMA model and optimization goals O1-O4 using the customized OA for a tolerance of 0.2%

O	k_1	k_2	k_3	k_4	k_5	k_6	X_3	OG (%)	CPU (s)
1	0.26	5.00	5.00	0.20	0.20	0.20	8.30	0.20	2.94
2	0.21	0.22	0.21	0.20	0.20	0.20	1.10	0.00	0.06
3	0.60	5.00	5.00	0.53	0.20	0.24	5.40	0.13	2.35
4	1.00	1.05	0.97	0.92	1.00	1.00	1.10	0.00	0.23
O	k_1	k_2	k_3	k_4	k_5	k_6	v_4	OG (%)	CPU (s)
1	3.96	5.00	5.00	5.00	0.20	2.99	37.47	0.00	0.16
2	3.22	3.55	5.00	4.99	0.20	0.21	31.33	0.17	0.66
3	0.68	1.79	1.12	1.27	0.20	0.21	6.60	0.00	0.12
4	1.16	1.00	1.34	1.34	1.00	1.00	7.61	0.11	1.98

Table 3 Results (objective function) of the optimization of case O1- v_4 for specific regions of k_2 and k_5 obtained with BARON for the SC model

k_5/k_2	1	2	3	4	5	6	7	8
8	36.50	36.71	36.90	37.08	37.24	37.37	37.47	37.47
7	36.62	36.83	37.02	37.19	37.34	37.46	37.47	37.47
6	36.75	36.95	37.14	37.31	37.44	37.47	37.47	37.47
5	36.88	37.08	37.26	37.41	37.47	37.47	37.47	37.47
4	37.02	37.21	37.38	37.47	37.47	37.47	37.47	37.47
3	37.15	37.34	37.47	37.47	37.47	37.47	37.47	37.47
2	37.29	37.46	37.47	37.47	37.47	37.47	37.47	37.47
1	37.43	37.47	37.47	37.47	37.47	37.47	37.47	37.47

Domain of each $k_r (4 \leq k_2 \leq 5; 0.2 \leq k_5 \leq 0.8)$ has been split into 8 intervals with equal width.

Table 4 Results (objective function) of the optimization of case O1-v_4 for specific regions of k_2 and k_5 obtained with the customized OA for the rGMA model

k_5-k_2	1	2	3	4	5	6	7	8
8	36.50	36.71	36.90	37.08	37.24	37.37	37.47	37.47
7	36.62	36.83	37.02	37.19	37.34	37.46	37.47	37.47
6	36.75	36.95	37.14	37.31	37.44	37.47	37.47	37.47
5	36.88	37.08	37.26	37.41	37.47	37.47	37.47	37.47
4	37.02	37.21	37.38	37.47	37.47	37.47	37.47	37.47
3	37.15	37.34	37.47	37.47	37.47	37.47	37.47	37.47
2	37.29	37.46	37.47	37.47	37.47	37.47	37.47	37.47
1	37.43	37.47	37.47	37.47	37.47	37.47	37.47	37.47

Domain of each $k_r (4 \leq k_2 \leq 5; 0.2 \leq k_5 \leq 0.8)$ has been split into 8 intervals with equal width.

example, we may encounter instances that are hard to solve using standard techniques. Consider, for instance, the same reaction scheme as before but this time with the alternative parameters indicated in the following model:

$$\frac{dX_1}{dt} = \frac{11.11 k_1 X_5^{2.86}}{X_5^{2.86} + 0.81}$$
$$- \frac{12.35 k_2 X_1^{1.54}}{\left(X_1^{1.54} + 0.61\right) X_3^{6.81} \left(0.11 + \frac{1}{X_3^{6.81}}\right)}$$

$$\frac{dX_2}{dt} = \frac{12.35 k_2 X_1^{1.54}}{\left(X_1^{1.54} + 0.61\right) X_3^{6.81} \left(0.11 + \frac{1}{X_3^{6.81}}\right)}$$
$$- \frac{4.44 k_3 X_2^{4.14}}{X_2^{4.14} + 0.11}$$
$$- \frac{7.41 k_5 X_1^{0.51} X_2^{26.51}}{\left(X_1^{0.51} + 0.19\right)\left(X_2^{26.51} + 0.11\right)} \qquad (8)$$

$$\frac{dX_3}{dt} = \frac{4.44 k_3 X_2^{4.14}}{X_2^{4.14} + 0.11} - \frac{4.44 k_4 X_3^{4.14}}{X_3^{4.14} + 0.11}$$

$$\frac{dX_4}{dt} = \frac{7.41 k_5 X_1^{0.51} X_2^{26.51}}{\left(X_1^{0.51} + 0.19\right)\left(X_2^{26.51} + 0.11\right)}$$
$$- \frac{6.67 k_6 X_4^{1.57}}{X_4^{1.57} + 1.40}$$

The optimization task of interest being:

- O5: Which is the optimal pattern of changes in enzyme activities that maximize v_6 in the new steady-state for a fixed value of X_5 and considering the following constraints?

$$0.3 \le X_1 \le 30$$
$$0.1 \le X_2 \le 10$$
$$0.1 \le X_3 \le 10 \qquad (9)$$
$$0.6 \le X_4 \le 50$$
$$0.1 \le k_r \le 20 \quad r = 1, ..., p$$

When BARON is employed to solve this case using the native SC form, it cannot reduce the optimality gap

Table 5 Results (CPU time in seconds) of the optimization of case O1- v_4 for specific regions of k_2 and k_5 obtained with BARON for the SC model

k_5/k_2	1	2	3	4	5	6	7	8
8	212.53	308.53	185.64	201.80	222.30	201.53	139.16	178.31
7	194.81	161.16	215.80	196.81	344.73	243.02	0.03	174.81
6	234.30	203.75	147.08	180.69	328.34	254.42	304.11	280.53
5	212.08	282.41	329.33	237.34	208.02	292.27	200.00	154.62
4	288.00	160.14	92.94	235.80	172.69	147.14	56.11	150.28
3	125.56	111.17	150.27	187.52	337.97	158.16	112.66	264.12
2	239.70	190.59	100.03	138.47	106.38	205.14	119.39	246.34
1	140.42	102.12	80.45	21.69	73.12	96.61	89.94	80.03

Domain of each $k_r(4 \le k_2 \le 5; 0.2 \le k_5 \le 0.8)$ has been split into 8 intervals with equal width.

Table 6 Results (CPU time in seconds) of the optimization of case O1-v_4 for specific regions of k_2 and k_5 obtained with the customized OA for the rGMA model

k_5/k_2	1	2	3	4	5	6	7	8
8	0.13	0.27	0.23	0.18	0.17	0.19	0.28	0.28
7	0.26	0.28	0.28	0.26	0.28	0.23	0.32	0.25
6	0.32	0.30	0.28	0.28	0.27	0.23	0.19	0.25
5	0.31	0.21	0.25	0.25	0.26	0.28	0.27	0.29
4	0.25	0.27	0.32	0.30	0.25	0.27	0.26	0.28
3	0.20	0.22	0.28	0.28	0.29	0.30	0.19	0.53
2	0.28	0.25	0.19	0.19	0.22	0.17	0.30	0.25
1	0.23	0.24	0.26	0.27	0.23	0.21	0.24	0.31

Domain of each $k_r(4 \le k_2 \le 5; 0.2 \le k_5 \le 0.8)$ has been split into 8 intervals with equal width.

below the specified tolerance after 1 hour of CPU time. In contrast, when the model is recast into its rGMA form and our OA method is applied, the global optimum can be determined with an optimality gap of 2% in 10.95 seconds (see Table 7). This illustrates both, the utility of using the rGMA as a canonical form for dealing with the optimization of SC models, and the computational efficiency of our global optimization methods specifically designed to take advantage of the mathematical structure of the GMA.

3 Discussion

While experimental tools to manipulate gene expression are already available, there is no established set of guidelines on how these tools can be used to achieve a certain goal. So far, two main difficulties have prevented model driven optimization from becoming a standard in providing such guidelines: (i) the lack of information to build detailed kinetic models and (ii) the computational difficulties that arise upon the optimization of such models. The latter can be exemplified by the application of mixed integer non-linear optimization techniques (MINLP) in the context of kinetic models presented in [34,35]. In such cases, the optimization task showed to be computationally very demanding and global optimality could not be guaranteed in many cases. We propose that using models with a standardized structure may offer a solution to both problems. On one hand, approximate kinetics, such as the SC formalism, can provide very accurate approximations and retain key features of the system like saturation and cooperativity. On the other hand, these formalisms can be automatically recast

Table 7 Results of the optimization of model 8 with BARON (SC model) and the customized OA (rGMA model)

Solver	k_1	k_2	k_3	k_4	k_5	k_6	OF	OG (%)	CPU (s)
BARON (SC)	6.24	5.16	0.46	0.6	8.46	9.09	60.36	45.18	3600
OA (rGMA)	6.25	5.17	0.45	0.6	8.44	9.1	60.46	2.18	10.95

into GMA form and using efficient global optimization methods developed specifically for this canonical representation. Although this technique will certainly have limitations, our previous results indicate that it can be applied to models of moderate complexity without major problems [32]. Optimization of GMA models comprising up to 60 reactions and 40 metabolites offer no limitation to our technique. We have shown how these methods can be easily used to optimize SC via recasting into rGMA models while still being quite efficient.

Our results can be of particular interest for dealing with multicriteria optimization on realistic models. This kind of problems are relevant when exploring the adaptive response to changing conditions, were conflictive goals may be at play [39,40]. Particularly, we should notice that several multi-objective optimization techniques, such as the weighted sum or epsilon constraint methods [41] are based on solving a set of auxiliary single-objective problems. These approaches could directly benefit from the numerical advances presented in this work. This kind of problems are relevant when exploring the adaptive response to changing conditions, were conflictive goals may be on play [39,40]. The highly efficient OA algorithm applied to rGMA models provide a practical way for extending multicriteria optimization methods, for instance as used in [39], to non-linear kinetic models. It is in principle possible to make use of methods such as ours to analyze the optimality of large scale dynamic systems much in the same way that Flux Balance Analysis can be applied to analyze the stoichiometry of an organism on a genomic scale. To make this possible, however, extensive experimental and modeling efforts would be required to characterize the most important properties of the involved processes. In fact, we anticipate that practical limitations to apply the techniques presented here in solving larger problems will be dominated by the lack of information about the component processes and metabolites rather than by the technical capacity of the optimization technique presented here. Although a complete kinetic characterization of the processes in a complete metabolic network may yet be far, information on kinetic orders and saturation fractions is easier to obtain. In this context, the SC formalism provides a sound approximation that results in a mathematical representation useful for simulation and optimization through recasting.

4 Conclusions

We expect that the possibility of building models using non-linear approximate formalisms and of subsequently optimizing these models will trigger interest in the experimental characterization of the components of cellular metabolism. After the genomic explosion, we need to step back and begin to measure enzyme activities, metabolite levels, and regulatory signals on a larger scale than we used to do before, if we want to understand the emergence of the dynamic properties of biological systems and to be able to develop successful biotechnological applications.

5 Methods
5.1 Modelling strategies

The process of model building and optimization can be used to understand how a system should be changed in order to achieve specific biotechnological goals or how the same system has evolved in order to more efficiently execute a given biological function. Different trade-offs are considered during the modeling process. On the one hand, one wants to use models that are as simple as possible to guarantee numerical tractability. Unfortunately simplifications may lead to models whose accuracy is only ensured for a limited range of physiological conditions. On the other hand, models that are very detailed and accurate over a wide range of physiological conditions are typically more difficult to analyze and optimize. Needless to say, the type of modeling strategy and the model one chooses to implement have a large impact on the results of the analysis. The most widely used strategies in the context of optimization are: (1) Stoichiometric models, (2) Kinetic models, and (3) Approximated models.

The three strategies have as a starting point a set of ordinary differential equations, in which the dependent variables or nodes are the chemical species whose dynamical behavior one is interested in studying. For a system with n dependent variables, p processes and m independent variables, the node equations are written as follows:

$$\dot{X}_i = \sum_{r=1}^{p} \mu_{ir} v_r \qquad i = 1, .., n \qquad (10)$$

μ_{ir} stands for the stoichiometry of each metabolite X_i in each reaction r with respect to metabolite i and can be derived from the reaction scheme.

At this stage, the various strategies begin to differ in the way that they implement and analyze the equations. Typically, Flux balance analysis (FBA) and related techniques consider only the steady state behavior of the system, and treat v_r as a variable whose value can be changed in order to optimize specific steady state constraints. To accomplish this, FBA-like methods attempt to find solutions for the following system of linear equations:

$$0 = \sum_{r=1}^{p} \mu_{ir} v_r \quad i = 1, .., n \qquad (11)$$

This system of equations is solved under different assumptions. A typical problem is that of understanding the effect of knocking out different genes from the system. This analysis can be performed by setting $v_r = 0$ for the process(es) that depend on the product of the genes that are knocked out. Once these constraints are set, linear optimization techniques can be used to identify the region of the variable space that satisfies the steady state and optimizes at the same time a set of specific measurable aspects of the systems [42-44]. It must be noted that FBA analysis of Eqn. (11) does not account for the regulatory effects that can result from gene knockout and it cannot be used to predict changes in metabolic concentrations and temporal responses. Thus, optimization constraints are limited to steady-state fluxes [15].

To overcome these limitations, we must use more complex kinetic models where the effect of changing the values of the variables on the fluxes is taken into account. This requires defining a functional form for each v_r in Eqn. (10). Often, this functional form is drawn from a number of classical enzyme kinetic rate-laws. As a result, we use an approximate expression for the kinetic behavior of each elementary process whose form depends on the underlying mechanism of the process. The reason for this is that the classical rate laws are rational functions of the variables and they are built upon different types of simplifying assumptions on the detailed mechanism of the reactions. Such assumptions range from considering that the elementary chemical steps of the catalytic process occur at very different timescales to assuming that the concentration of the catalyst and of the reactants differ in orders of magnitude. Thus, rate laws such as the popular Michaelis-Menten are approximations to the actual mechanism in specific conditions. However, more often than not, one does not have enough information to judge if such conditions meet those one is trying to model. Thus, using rational enzyme kinetics in models lacks a sound theoretical ground. In fact, within the complex architecture of the intracellular milieu, many of the assumptions that justify these classical rate-laws may not hold [45-47]. Even in the best case scenario where a detailed kinetic model using classical enzyme kinetics can be derived and numerically identified, it may be hard to globally optimize that model using general purpose algorithms. As we will show here, available optimization techniques may fail to solve fairly trivial optimization tasks even in simple models. These numerical difficulties can be overcome by defining reformulated models based on canonical representations that are easier to handle using customized global optimization algorithms devised for specific canonical functional forms.

As an alternative, theoretically well supported canonical representations can be derived using approximation theory. One type of such representations are power-law models. In a power-law model, each v_r in Eqn. (10) is approximated as [19,21]:

$$v_r(X_1, .., X_n, ...X_{n+m})$$
$$\approx \gamma_r \prod_{j=1}^{n+m} X_j^{f_{jr}} \quad r = 1, .., p \tag{12}$$

This approximation is derived at a given operating point $(X_{1_0}, X_{2_0}, .., X_{(n+m)_0})$ as a first-order Taylor series representation of the target function in log-log space. This approximation can generate models with different representations. The two that are most commonly used are the S-system representation and the GMA representation. The S-system representation is obtained by lumping the various processes that contribute to the synthesis of a given metabolite into a global process of synthesis V_i^+ and those that contribute to the utilization of a given metabolite into a global degradation process

$$\dot{X}_i = \sum_{r=1}^{p} \mu_{ir} v_r$$
$$= \sum_{r=1}^{p} \mu_{ir}^+ v_r - \sum_{r=1}^{p} \mu_{ir}^- v_r \qquad :$$
$$= V_i^+ - V_i^- \qquad i = 1, .., n$$

$$\dot{X}_i = \sum_{r=1}^{p} \mu_{ir} v_r$$
$$= \sum_{r=1}^{p} \mu_{ir}^+ v_r - \sum_{r=1}^{p} \mu_{ir}^- v_r \tag{13}$$
$$= V_i^+ - V_i^- \qquad i = 1, .., n$$

Then, the aggregated processes are represented by power-law functions:

$$\dot{X}_i = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} -$$
$$\beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}} \quad i = 1, .., n \tag{14}$$

Alternatively, the GMA form is obtained representing each individual v_r as a power-law:

$$\dot{X}_i = \sum_{r=1}^{p} \mu_{ir} v_r$$
$$= \sum_{r=1}^{p} \left(\mu_{ir} \gamma_r \prod_{j=1}^{n+m} X_j^{f_{rj}} \right) \quad i = 1, .., n \tag{15}$$

The parameters in these representations have a clear physical interpretation. Kinetic orders, the exponents in the power-laws, are local sensitivities of the fluxes, either individual (f_{rj} for v_r) or aggregated (g_{ij} for V_i^+ and h_{ij} for V_i^-), with respect to X_j. Rate-constants (α_i, β_i and γ_r) are parameters that are computed so that the flux in the model at steady state is equal to the operating flux at the operating point for the metabolites. Parameter estimation techniques have been developed so that power-law parameters can be calculated from experimental measurements [13]. It should also be noted that the use of estimation procedures (i.e., least-squares), alternate regression or similar procedures to estimate power-law parameters from dynamic curves lead to a power-law representation that is no longer local according to the classical definition [48-50]. Those models may, by definition, slightly improve their accuracy over strictly local models.

To complement the power-law approach, the Saturable and Cooperative (SC) formalism was introduced by Sorribas et al. [23] as an extension of the ideas that led to the power-law formalism. The SC representation of a given velocity is:

$$v_r(X_1, .., X_n, ...X_{n+m}) \approx \frac{V_r \prod\limits_{j=1}^{n+m} X_j^{n_{rj}}}{\prod\limits_{j=1}^{n+m} \left(K_{rj} + X_j^{n_{rj}} \right)} \qquad (16)$$

This representation can be obtained from a power-law model defined at a given operating point $X_0 = (X_{10},.., X_{(n+m)0})$ through the following relationships:

$$n_{rj} = \frac{f_{rj}}{1 - p_{rj}} \qquad \begin{aligned} r &= 1, .., p \\ j &= 1, .., n+m \end{aligned} \qquad (17)$$

$$K_{rj} = \frac{1 - p_{rj}}{p_{rj}} X_{i0}^{n_{rj}} \qquad \begin{aligned} r &= 1, .., p \\ j &= 1, .., n+m \end{aligned} \qquad (18)$$

Thus SC uses the same information as the power-law except for the new parameters p_{rj} (saturation fractions), which are defined as:

$$p_{rj} = v_{r0}/V_{rj} \qquad \begin{aligned} r &= 1, .., p \\ j &= 1, .., n+m \end{aligned} \qquad (19)$$

where $v_{r0} = v_r(X_{10},.., X_{n0},... X_{(n+m)0})$ and V_{rj} is either the limit velocity (saturation) when $X_j \to \infty$ if $n_{rj} > 0$, or the limit velocity when $X_j \to 0$ if $n_{rj} < 0$.

Using SC models for global optimization can raise some numerical issues. These difficulties can be avoided to a large extent by recasting SC models into a canonical GMA model, through the introduction of auxiliary variables, as will be shown in the next section.

5.2 Recasting non-linear models into power-law canonical models by increasing the number of variables

Non-linear models can be *exactly* recast into GMA or S-system models through the use of auxiliary variables [36]. As a result, the final model is an exact representation of the original model, written in a canonical form. In other words, the resulting GMA model is not an approximation to the original model: it is an exact replica of it. To avoid confusion, hereafter, we refer to a GMA model that exactly recasts another as an rGMA model.

As a very simple introductory example, consider a linear pathway with two internal metabolites X_1 and X_2 and a source metabolite X_3 (Figure 3). In this pathway, X_2 is a competitive inhibitor of the synthesis of X_1 from the source metabolite. A generic model using Michaelis-Menten kinetic functions, assuming a competitive inhibition of the first reaction by X_2, can be written as:

$$\begin{aligned} \dot{X}_1 &= \frac{V_1 X_3}{K_1(1 + K_i/X_2) + X_3} \\ &\quad - \frac{V_2 X_1}{K_2 + X_1} \\ \dot{X}_2 &= \frac{V_2 X_1}{K_2 + X_1} - \frac{V_3 X_2^2}{K_3^2 + X_2^2} \end{aligned} \qquad (20)$$

in which X_3 is an externally fixed variable.

Recasting this model as a rGMA can be done as follows. First, let us define three new variables:

$$\begin{aligned} X_4 &= K_1(1 + Ki/X_2) + X_3 \\ X_5 &= K_2 + X_1 \\ X_6 &= K_3^2 + X_2^2 \end{aligned} \qquad (21)$$

We can now write the model in 20 as:

$$\begin{aligned} \dot{X}_1 &= V_1 X_3 X_4^{-1} - V_2 X_1 X_5^{-1} \\ \dot{X}_2 &= V_2 X_1 X_5^{-1} - V_3 X_2^2 X_6^{-1} \end{aligned} \qquad (22)$$

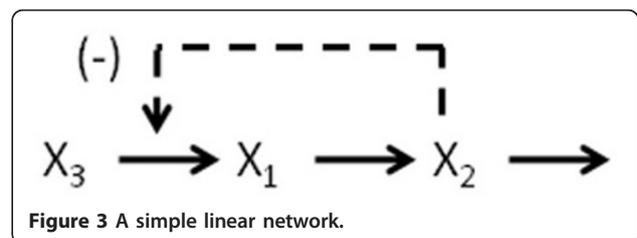with initial conditions $X_1(0) = X_{1_0}$ and $X_2(0) = X_{2_0}$.

Figure 3 A simple linear network.

To complete the recasting we must now provide the equations that follow the change in the new variables over time. These are given by the following equations:

$$\dot{X}_4 = -\frac{K_1 K_i \dot{X}_2}{X_2^2}$$
$$= V_3 K_1 K_i X_6^{-1} - V_2 K_1 K_i X_1 X_5^{-1}$$
$$X_2^{-2} \dot{X}_5 = \dot{X}_1 \tag{23}$$
$$= V_1 X_3 X_4^{-1} - V_2 X_1 X_5^{-1}$$
$$\dot{X}_6 = 2X_2 \dot{X}_2$$
$$= 2V_2 X_1 X_2 X_5^{-1} - 2V_3 X_2^3 X_6^{-1}$$

with initial conditions $X_4(0) = K_1(1 + K_i/X_{2_0}) + X_{3_0}$, $X_5(0) = K_2 + X_{1_0}$, and $X_6(0) = K_3^2 + X_{2_0}^2$.

The resulting rGMA model (22-23) is an exact representation of model in (20). Hence, for a set of appropriate initial conditions, the simulation of the dynamic response using either the model recast as a rGMA or the original model will produce the same trajectory. In principle, any non-linear model can be recast into a rGMA following a similar procedure [36]. This can be extremely useful, because it allows for the application of tailored global optimization procedures originally devised for GMA models [28-30,32,51,52] to generic non-linear models.

Author details
[1]Departament d'Enginyeria Química, Universitat Rovira i Virgili, Avinguda Països Catalans 26, 43007-Tarragona, Spain. [2]Technische Universität München. Fachgebiet für Systembiotechnologie, Boltzmannstr. 15 85748 Garching, Germany. [3]Departament de Ciències Mèdiques Bàsiques, Institut de Recerca Biomèdica de Lleida (IRBLLEIDA), Universitat de Lleida, Montserrat Roig 2, 25008 Lleida, Spain.

Authors' contributions
AM-S suggested the potential utility of recasting for optimizing non-linear kinetic models. AS and AM-S elaborate on the recasting of SC models and planned the work. CP, GG-G and LJ implemented the OA algorithm and worked out the technical solution for applying it to a rGMA model. CP and GG-G performed the optimization tasks. AS and RA defined the reference model and obtained the numerical parameters used in the paper. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

References
1. Voit EO: **Optimization in integrated biochemical systems.** *Biotechnol Bioeng* 1992, **40(5)**:572-82.
2. Alvarez-Vasquez F, Gonzalez-Alcon C, Torres NV: **Metabolism of citric acid production by Aspergillus niger: model definition, steady-state analysis and constrained optimization of citric acid production rate.** *Biotechnol Bioeng* 2000, **70**:82-108.
3. Banga JR: **Optimization in computational systems biology.** *BMC Syst Biol* 2008, **2**:47.
4. Rodriguez-Prados JC, de Atauri P, Maury J, Ortega F, Portais JC, Chassagnole C, Acerenza L, Lindley ND, Cascante M: **In silico strategy to rationally engineer metabolite production: A case study for threonine in Escherichia coli.** *Biotechnology and bioengineering* 2009, **103(3)**:609-620.
5. Scheer M, Grote A, Chang A, Schomburg I, Munaretto C, Rother aSC M, Stelzer M, JT, DS: **BRENDA, the enzyme information system in 2011.** *Nucleic Acids Res* 2011, **39**:670-676.
6. Rojas I, Golebiewski M, Kania R, Krebs O, Mir S, Weidemann A, Wittig U: **SABIO-RK: a database for biochemical reactions and their kinetics.** *BMC Systems Biology* 2007, **1**:S6.
7. Shiraishi F, Savageau MA: **The tricarboxylic acid cycle in Dictyostelium discoideum. I. Formulation of alternative kinetic representations.** *The Journal of biological chemistry* 1992, **267(32)**:22912-22918.
8. Shiraishi F, Savageau MA: **The tricarboxylic acid cycle in Dictyostelium discoideum. II. Evaluation of model consistency and robustness.** *The Journal of biological chemistry* 1992, **267(32)**:22919-22925.
9. Shiraishi F, Savageau MA: **The tricarboxylic acid cycle in Dictyostelium discoideum. III. Analysis of steady state and dynamic behavior.** *The Journal of biological chemistry* 1992, **267(32)**:22926-22933.
10. Shiraishi F, Savageau MA: **The tricarboxylic acid cycle in Dictyostelium discoideum. IV. Resolution of discrepancies between alternative methods of analysis.** *The Journal of biological chemistry* 1992, **267(32)**:22934-22943.
11. Fonseca LL, Sanchez C, Santos H, Voit EO: **Complex coordination of multi-scale cellular responses to environmental stress.** *Molecular bioSystems* 2010.
12. Sorribas A, Cascante M: **Structure identifiability in metabolic pathways: parameter estimation in models based on the power-law formalism.** *The Biochemical journal* 1994, **298(Pt 2)**:303-311.
13. Chou IC, Voit EO: **Recent developments in parameter estimation and structure identification of biochemical and genomic systems.** *Mathematical Biosciences* 2009, **219(2)**:57-83.
14. Grossmann I, Biegler LT: **Part II. Future perspective on optimization.** *Computers and Chemical Engineering* 2004, **28**:1193-1218.
15. Terzer M, Maynard ND, Covert MW, Stelling J: **Genome-scale metabolic networks.** *Wiley interdisciplinary reviews. Systems biology and medicine* 2009, **1(3)**:285-297.
16. Gianchandani EP, Chavali AK, Papin JA: **The application of flux balance analysis in systems biology.** *Wiley interdisciplinary reviews. Systems biology and medicine* 2010, **2(3)**:372-382, [JID: 101516550; ppublish].
17. Voit EO: **Design principles and operating principles: the yin and yang of optimal functioning.** *Math Biosci* 2003, **182**:81-92.
18. Jamshidi N, Palsson BO: **Formulating genome-scale kinetic models in the post-genome era.** *Molecular systems biology* 2008, **4**:171.
19. Savageau MA: **Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions.** *Journal of theoretical biology* 1969, **25(3)**:365-369.
20. Savageau MA: **Biochemical systems analysis. II. The steady-state solutions for an n-pool system using a power-law approximation.** *Journal of theoretical biology* 1969, **25(3)**:370-379.
21. Savageau MA: **Biochemical systems analysis. 3. Dynamic solutions using a power-law approximation.** *Journal of theoretical biology* 1970, **26(2)**:215-226.
22. Savageau MA: *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology* Reading, Mass.: Addison-Wesley; 1976.
23. Sorribas A, Hernandez-Bermejo B, Vilaprinyo E, Alves R: **Cooperativity and saturation in biochemical networks: a saturable formalism using Taylor series approximations.** *Biotechnology and bioengineering* 2007, **97(5)**:1259-1277.

24. Liebermeister W, Klipp E: Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theoretical biology & medical modelling* 2006, **3**:41.

25. Goel G, Chou IC, Voit EO: System Estimation from Metabolic Time Series Data. *Bioinformatics (Oxford, England)* 2008.

26. Ni T, Savageau M: Model assessment and refinement using strategies from biochemical systems theory: Application to metabolism in human red blood cells. *Journal of Theoretical Biology* 1996, **179(4)**:329-368.

27. Arkin A, Ross J, McAdams H: Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. *Genetics* 1998, **149(4)**:1633-1648.

28. Polisetty PK, Gatzke EP, Voit EO: Yield optimization of regulated metabolic systems using deterministic branch-and-reduce methods. *Biotechnol Bioeng* 2008, **99(5)**:1154-69.

29. Guillén-Gosálbez G, Sorribas A: Identifying quantitative operation principles in metabolic pathways: A systematic method for searching feasible enzyme activity patterns leading to cellular adaptive responses. *BMC Bioinformatics* 2009, **10**.

30. Guillén-Gosálbez G, Pozo C, Jiménez L, Sorribas A: A global optimization strategy to identify quantitative design principles for gene expression in yeast adaptation to heat shock. *Computer Aided Chemical Engineering* 2009, **26**:1045-1050.

31. Voit EO: *Computational Analysis of Biochemical Systems. A Practical Guide for Biochemists and Molecular Biologists* Cambridge, U.K.: Cambridge University Press; 2000.

32. Pozo C, Guillén-Gosálbez G, Sorribas A, Jiménez L: A Spatial Branch-and-Bound Framework for the Global Optimization of Kinetic Models of Metabolic Networks. *Industrial and Engineering Chemistry Research* 2010.

33. Sorribas A, Pozo C, Vilaprinyo E, Guillén-Gosálbez G, Jiménez L, Alves R: Optimization and evolution in metabolic pathways: global optimization techniques in Generalized Mass Action models. *Journal of Biotechnology* 2010, **149(3)**:141-153.

34. Chassagnole C, Noisommit-Rizzi N, Schmid J, Mauch K, Reuss M: Dynamic modeling of the central carbon metabolism of Escherichia coli. *Biotechnol Bioeng* 2002, **79**:53-73.

35. Nikolaev E: The elucidation of metabolic pathways and their improvements using stable optimization of large-scale kinetic models of cellular systems. *Metab Eng* 2010, **12**:26-38.

36. Savageau MA, Voit EO: Recasting nonlinear differential equations as S-systems: a canonical nonlinear form. *Mathematical Biosciences* 1987, **87**:83-115.

37. Voit EO: Recasting nonlinear models as S-systems. *Mathematical and Computer Modelling* 1988, **11(C)**:140-145.

38. Smith EMB, Pantelides CC: A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers and Chemical Engineering* 1999, **23(4-5)**:457-478.

39. Vera J, de Atauri P, Cascante M, Torres NV: Multicriteria optimization of biochemical systems by linear programming: application to production of ethanol by Saccharomyces cerevisiae. *Biotechnol Bioeng* 2003, **83(3)**:335-43.

40. Vilaprinyo E, Alves R, Sorribas A: Use of physiological constraints to identify quantitative design principles for gene expression in yeast adaptation to heat shock. *BMC Bioinformatics* 2006, **7**:184.

41. Ehrgott M: *Multicriteria Optimization* Springer; 2005.

42. Famili I, Forster J, Nielsen J, Palsson BO: Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proc Natl Acad Sci USA* 2003, **100(23)**:13134-9.

43. Famili I, Mahadevan R, Palsson BO: k-Cone analysis: determining all candidate values for kinetic parameters on a network scale. *Biophys J* 2005, **88(3)**:1616-25.

44. Price ND, Papin JA, Schilling CH, Palsson BO: Genome-scale microbial in silico models: the constraints-based approach. *Trends Biotechnol* 2003, **21(4)**:162-9.

45. Savageau MA: Influence of fractal kinetics on molecular recognition. *Journal of Molecular Recognition: JMR* 1993, **6(4)**:149-157.

46. Savageau MA: Michaelis-Menten mechanism reconsidered: implications of fractal kinetics. *Journal of theoretical biology* 1995, **176**:115-124.

47. Savageau MA: Development of fractal kinetic theory for enzyme-catalysed reactions and implications for the design of biochemical pathways. *Bio Systems* 1998, **47(1-2)**:9-36.

48. Hernandez-Bermejo B, Fairen V, Sorribas A: Power-law modeling based on least-squares minimization criteria. *Mathematical biosciences* 1999, **161(1-2)**:83-94.

49. Hernandez-Bermejo B, Fairen V, Sorribas A: Power-law modeling based on least-squares criteria: consequences for system analysis and simulation. *Mathematical biosciences* 2000, **167(2)**:87-107.

50. Alves R, Vilaprinyo E, Hernandez-Bermejo B, Sorribas A: Mathematical formalisms based on approximated kinetic representations for modeling genetic and metabolic pathways. *Biotechnology and Genetic Engineering Reviews* 2008, **25**:1-40.

51. Marin-Sanguino A, Torres NV: Optimization of biochemical systems by linear programming and general mass action model representations. *Math Biosci* 2003, **184(2)**:187-200.

52. Marin-Sanguino A, Voit EO, Gonzalez-Alcon C, Torres NV: Optimization of biotechnological systems through geometric programming. *Theor Biol Med Model* 2007, **4**:38.