

A new strategy for assessing sensitivities in biochemical models

BY SVEN SAHLE¹, PEDRO MENDES^{2,3}, STEFAN HOOPS³
AND URSULA KUMMER^{1,*}

¹*Department Modeling of Biological Processes, Institute for Zoology/BIOQUANT,
Im Neuenheimer Feld 267, 69120 Heidelberg, Germany*

²*School of Computer Science and Manchester Centre for Integrative Systems
Biology, University of Manchester, Oxford Road, Manchester M13 9PL, UK*

³*Virginia Bioinformatics Institute, Virginia Tech, Washington Street 0477,
Blacksburg, VA 24061, USA*

An integral part of any systems biology approach is the modelling and simulation of the respective system under investigation. However, the values of many parameters of the system have often not been determined or are not identifiable due to technical experimental difficulties or other constraints. Sensitivity analysis is often employed to quantify the importance of each of the model's parameters in the behaviour of the system. This approach can also be useful in identifying those parts of the system that are most sensitive with the potential of becoming drug targets. A problem of the commonly used methods of sensitivity analysis is that they constitute local methods meaning that they depend directly on the exact parameter space, which in turn is not known exactly. One way to circumvent this problem is to carry out sensitivity analysis over a wide range of values for all parameters, but this is handicapped by expensive computations when the systems are high dimensional. Another approach is to employ global sensitivity analysis, which in this context is mostly based on random sampling methods. In this paper we present an efficient approach that involves using numerical optimizing methods that search a wide region of parameter space for a given model to determine the maximum and minimum values of its metabolic control coefficients. A relevant example for drug development is presented to demonstrate the strategy using the software COPASI.

Keywords: sensitivity analysis; robustness; optimization

1. Introduction

Systems biology raises hopes for more effective drug discovery (Butcher 2005). Research in this area involves an integrated experimental and computational approach comprising quantitative data generation on one hand and modelling

* Author for correspondence (ursula.kummer@bioquant.uni-heidelberg.de).

One contribution of 12 to a Theme Issue 'Biomedical applications of systems biology and biological physics'.

and simulation on the other. This is a promising way to understand the complex mechanisms that underlie normal physiological regulations as well as the development of diseases. In particular, it is useful to identify effective drug targets and to select drugs that have minimal toxicity.

The majority of computational models that represent a kind of virtual biochemistry consist of sets of ordinary differential equations that represent the concentrations of different chemical species as they occur in cells, in tissues or in whole organisms (see [Heinrich & Schuster 1996](#)). These models are brought to life by numeric integration of the differential equations. Obviously, whether on the cellular or physiological level, these models are never complete representations of all processes occurring in the living system. Moreover, in order to generate reliable predictions with such a model, detailed knowledge of kinetic and other parameters seems to be a prerequisite. However, most of the time the parameters for such models are derived from *in vitro* experimental data, which do not necessarily coincide with the *in vivo* conditions and most parameter sets are incomplete as well. It is often the case that the amount of experimental data is not sufficient for an unambiguous parameter estimation, the method of choice if sufficient data are available (for a recent review see [Van Riel 2006](#)), resulting in an underdetermined problem with many possible solutions (sets of parameter values). This raises a number of questions about the validity and practical value of models in these conditions.

One feature that allows one to assess the validity of a model beyond a specific parameter set is the robustness of the system's behaviour with respect to parameter changes. This robustness is a property of the system that makes a lot of sense in the real world. Living systems demonstrate a robust functioning and response of their processes to perturbations. Thus, it is pertinent to assume that many cellular processes should be fairly rigid ([Stephanopoulos & Vallino 1991](#)) and insensitive to parameter changes; on the other hand, some processes must be sensitive towards certain external parameters such that the system may adapt to environmental changes. Therefore, the robustness of biological systems is an interesting question for the understanding of their function *in vivo*. It is also often invoked to justify that some model parameters do not need to be determined accurately but yet represent and predict realistic behaviour.

In order to determine which parameters are important to know in detail and which have a lesser impact on the system's behaviour, a sensitivity analysis is commonly performed. Sensitivity analysis computes how much a certain system property, for example a steady-state concentration, depends on a specific parameter, for example a kinetic constant. A specialized form of sensitivity analysis for biochemical systems is the so-called metabolic control analysis (MCA; [Kacser & Burns 1973](#); [Heinrich & Rapoport 1974](#)), which offers a scaled sensitivity analysis through 'control coefficients'. MCA is very useful as a theoretical framework because it provides a set of summation theorems that explain many system-level properties of biochemical systems. Control coefficients can be defined for any state variable or quantities derived thereof, for example steady-state concentrations and fluxes. Control coefficients measure the response of the system variable in question to infinitesimal changes in the rate of one reaction of the system. There are also response coefficients, which are more general and measure the response of a system variable to infinitesimal changes in any parameter (control coefficients are response coefficients defined for a

parameter that affects the rate of a reaction linearly). MCA has often been employed to validate models in the context of drug discovery or to predict potential targets for drugs (for a review see [Cascante *et al.* 2002](#)).

A problem of MCA and similar sensitivity analysis is the fact that they are only valid for small perturbations of one parameter around a specific operating point in parameter space. Therefore, it is quite possible that a parameter that has been determined by MCA to play only a minor role may actually have a high impact when the values of other parameters are significantly different from those studied. Since usually there is a lot of uncertainty not only for one parameter, but indeed for many, there is a problem with respect to the validity of the results of this type of sensitivity analysis that are valid only around an operating point (hereafter called *local sensitivities*). These considerations are also true for experimentally determined sensitivities since a change in environmental conditions leading to a change of some of the parameters might lead to different sensitivities. Whereas parameters like Michaelis constants are much less prone to differences between the *in vitro* and *in vivo* situations, the above discussion is especially important for parameters such as the limiting rates (V_{\max}) that depend on enzyme concentrations and in some cases on their regulation. Since the *in vitro* enzyme concentration is hardly ever similar to the *in vivo* situation and there is often very limited knowledge about the actual enzyme concentrations in the cell, the values for these parameters are often just educated guesses.

Second-order sensitivities (sensitivities of sensitivities) offer some but only limited help ([Bentele *et al.* 2004](#)). They allow one to determine the robustness of the sensitivities with respect to parameter changes in the immediate vicinity of the studied parameter space. However, despite being predictive for a slightly larger range than first-order sensitivities, they are still essentially local sensitivities.

Another approach to solve the above problem is to scan the parameter space of the model with respect to first- or second-order sensitivities. This is a good way to do a more global analysis; however, it is computationally extremely expensive for any realistic model, since the number of calculations scales exponentially with the number of parameters in the model. Moreover, the results of scanning large numbers of parameters are hard to visualize. Therefore parameter scanning is only practical for very small models and not a workable solution when the parameter space is high dimensional.

In addition to the above, there exist several approaches to global sensitivity analysis (GSA; [Campolongo *et al.* 2000](#)). These have been rarely applied to biological systems so far. Obviously, this situation should be changed to improve the validity of model assessment. Most of the existing GSA methods rely on random sampling with or without probability distributions in parameter space, followed by a statistical (reviewed in [Campolongo *et al.* 2000](#)) or information-theoretic analysis ([Ludtke *et al.* 2008](#)) of the results. Again, random sampling is a very expensive step, except if the density of samples is to be made extremely small (in which case its value is also reduced). Sampling methods are arguably also not feasible if the parameter space is large.

The purpose of this paper is to present a novel approach that combines the use of optimization methods with sensitivity analysis, particularly in the convenient MCA formulation. We use different optimization strategies to find the largest

possible control coefficient in a given parameter space for each given parameter. A model of skeletal muscle glycolysis by Lambeth & Kushmerick (2002) is used to demonstrate the approach.

2. Method

All algorithms used here are available in COPASI, a software package for modelling, simulation and analysis of biochemical systems (Hoops *et al.* 2006). COPASI is available at <http://www.copasi.org>. It is free for academic users. COPASI offers optimization routines that can be applied to any state variable of a model or derived quantities, for example the coefficients of MCA. This means that for a given parameter range, the software tries to find the set of parameters for which the target function has a minimum or maximum value. In this application, we will use flux control coefficients as the target functions, therefore effectively discovering their maximum value over a wide range of parameter values. Rather than random sampling, this approach is then based on established search strategies.

A wide range of different optimization strategies are available, in general and in COPASI. It is then useful to experiment with different optimization algorithms (also known as search algorithms) and compare their performance. This is very important as the problem is essentially one of global optimization, and it is well known that the optimal strategy for global optimization depends very much on the specific problem (Wolpert & Macready 1997). Thus, it is crucial to apply several algorithms to these problems, to gain confidence that the solution obtained is indeed the global optimum (or at least a good local one; Mendes & Kell 1998). The algorithms that were used for this study are described below.

In this study, we derive the sensitivities of the steady-state fluxes of reactions with respect to the reaction rates in the system. These sensitivities are normalized and represent flux control coefficients according to MCA (Kacser & Burns 1973; Heinrich & Rapoport 1974). Thus,

$$C_i^{J_k} = \frac{v_i \partial J_k}{J_k \partial v_i}, \quad (2.1)$$

where $C_i^{J_k}$ represents the flux control coefficient of reaction k with respect to the reaction rate i ; v_i the rate of the controlling reaction; and J_k the flux of the reaction that responds.

The flux control coefficients of a model of interest are optimized in parameter space to find the maximal and minimal coefficients for each metabolite with respect to any parameter of interest (such as kinetic or rate constants). In order to make this query more efficient, it is advisable to select boundaries for the parameter space. First of all, these should be sensible boundaries from the physiological point of view; then they can be further restricted to values around observed intervals. The tighter the boundaries of the search, the smaller will be the computational costs. Second, the result becomes more relevant if the boundaries are chosen with some intelligence. This is due to the fact that it is quite likely that somewhere in parameter space there will be extreme values of control coefficients, but this may well be for parameter values that never occur in a realistic system; therefore there is not much usefulness in such a finding.

The following optimization methods are employed here.

(a) *Particle swarm*

The particle swarm optimization method suggested by Eberhart & Kennedy (1995) is inspired by how social insects seem to carry out efficient food searches as a team. Essentially, the algorithm is based on a group of agents carrying out random walks biased by the results obtained by their neighbours (which requires communication between the agents, or candidate solutions). Thus, each particle has a position X_i and a velocity v_i in the parameter space. The position of each particle corresponds to a complete candidate solution. At each iteration, a new velocity for each particle is calculated depending on its own position and velocity and the position of its best neighbour in the swarm. Once the new velocity is calculated, it is then used to calculate a new position in parameter space, thus updating the candidate solution. It is possible that a good solution may be achieved that is lost later, so the algorithm implementation remembers the best achieved objective function value O (control coefficient) and position M_i (parameter values).

(b) *Genetic algorithm*

The genetic algorithm (GA; Mitchell 1995) is a computational technique that mimics evolution of populations through variation and natural selection. A GA is composed of individuals that reproduce and compete, each one being a potential solution to the (optimization) problem and is represented by a ‘genome’ where each gene corresponds to one adjustable parameter. At each generation of the GA, each individual is paired with one other at random for reproduction. Two offspring are produced by combining their genomes and allowing for ‘cross-over’, i.e. the two new individuals have genomes that are formed from a combination of the genomes of their parents. Also each new gene might have mutated, i.e. the parameter value might have changed slightly. At the end of the generation, the algorithm has doubled the number of individuals. At that point, the algorithm selects half of them to be eliminated through a probabilistic approach, meaning that a good solution is highly likely to remain in the population but also has a small probability of being eliminated; a bad solution, while likely to be eliminated, has a small chance of remaining in the population. After a number of iterations (generations), the population of candidate solutions has improved. The literature on GAs is vast and there are many implementations; the implementation in COPASI is of a real-valued GA and thus is more like an evolution strategy (Beyer 2001).

(c) *Simulated annealing*

Simulated annealing is an optimization algorithm first proposed by Kirkpatrick *et al.* (1983) and is inspired by how perfect crystals are formed, using concepts from statistical mechanics. Perfect crystals grow from a molten substance that is cooled very slowly. At high temperatures, the particles vibrate with wide amplitude and this allows a search for global optima. As the temperature decreases so do the vibrations until the system settles to the global optimum (the perfect crystal). Much like creating perfect crystals, simulated

annealing is only effective if the cooling is slow, otherwise it settles on a local minimum (imperfect crystal). In the simulated annealing algorithm, the objective function is considered a measure of the energy of the system; for each fixed temperature (a temperature cycle), the Metropolis algorithm (Metropolis *et al.* 1953) is used to calculate the minimum energy for that temperature. In each iteration of the Metropolis algorithm, the parameters are changed to a nearby location in parameter space and a new value of the objective function is calculated: if the value is decreased, then the new state is accepted; if it increased then the new state is accepted with a probability that follows a Boltzmann distribution (higher temperature means higher probability of accepting the new state). After a fixed number of iterations of the Metropolis algorithm, the stopping criterion is checked; if it is not time to stop yet, then the system's temperature is reduced and a new round of the Metropolis algorithm is run at lower temperature. Simulated annealing is a stochastic algorithm that is guaranteed to converge if run for an infinite number of iterations. It is one of the most robust global optimization algorithms, although it is fairly slow. The implementation of simulated annealing in COPASI is based on the code of Corana *et al.* (1987).

(d) *Random search*

Random search is an optimization method that attempts to find the optimum by testing the objective function's value on a series of combinations of random values of the adjustable parameters. The random values are generated complying with any boundaries selected by the user. For an infinite number of iterations, this method is guaranteed to find the global optimum of the objective function. In general one is interested in processing a very large number of iterations. We employ this method to compare to other approaches to GSA using random sampling, which is essentially what this algorithm performs.

3. Test case: glycolysis in skeletal muscle

We chose a model published by Lambeth & Kushmerick (2002) for the glycolysis in skeletal muscle. It models an important process of potential significance for drug discovery, describing central metabolism in a human cell type. The model contains 16 reactions and 22 chemical species, and can therefore be considered to be of a moderate size (i.e. it is not trivially small).

The reaction network consists of the usual muscle glycolytic reactions plus creatine kinase, adenylate kinase and an ATPase. Figure 1 shows a schematic of the model. All concentration changes of metabolites are described by ordinary differential equations. These equations represent the sum of reaction rates corresponding to those reactions that directly increase or decrease the respective metabolite concentration. The flux control coefficients were studied in the original publication and it was found that the ATPase exerts the large control on the glycolytic flux.

The model is available in the BioModels database (Le Novère *et al.* 2006). We used the model provided in SBML format (a standard file format for the exchange of biochemical models; Hucka *et al.* 2003). The SBML file was imported into COPASI and used without further modification.

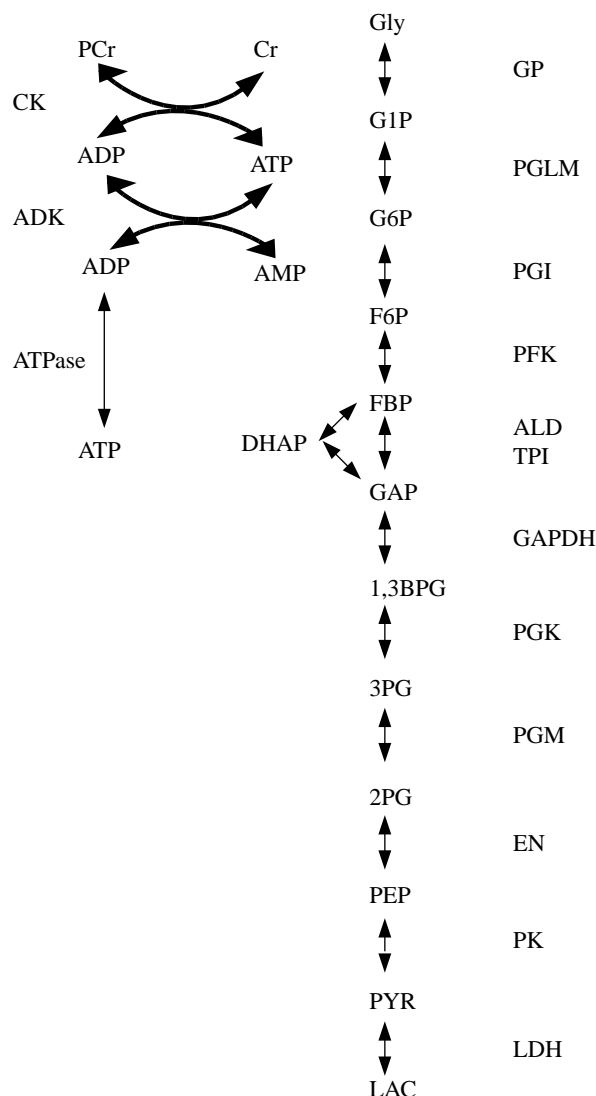


Figure 1. Schematic of the model of glycolysis in skeletal muscle (Lambeth & Kushmerick 2002) used in our study. For the sake of simplicity, ATP, ADP, NADH, etc. are not represented in the course of the glycolysis and only shown for the additional reactions used in the model, which are creatine kinase (CK), adenylate kinase (AK) and ATPase.

4. Results

We started with the original parameter set in Lambeth & Kushmerick (2002) and calculated the flux control coefficients (FCCs) that quantify the control exerted by the different reactions of the model towards the main branch flux (i.e. all reactions except those catalysed by adenylate kinase or creatine kinase). As described in Lambeth & Kushmerick (2002), flux is almost completely controlled by the ATPase reaction (table 1, column 2).

Table 1. Flux control coefficients (FCCs) exerted by each reaction (see also figure 1 for a representation of the reaction) of the model on the glycolytic flux. (The second column gives the FCCs for the original parameter values given in Lambeth & Kushmerick (2002). The third and fourth columns show an estimation of the minimal and maximal values of the respective FCCs, when the V_{\max} parameters are allowed to vary between half and double their original value. Correspondingly, columns five and six show the minimal and maximal values of the FCCs when the V_{\max} parameters can vary between one-tenth and 10 times the original value. All minimum and maximum values were estimated using the ‘particle swarm’ optimization method in COPASI with settings ‘iteration limit’=150 and ‘swarm size’=30.)

	original	ranges of V_{\max} values			
		0.5–2 times		0.1–10 times	
		minimum	maximum	minimum	maximum
GP	−0.00027	−0.00060	−0.00005	−1.10392	0.04807
PGLM	0.0	0.0	0.0	−0.79746	0.00160
PGI	0.0	0.0	0.0	−0.11764	0.00017
PFK	−0.00001	−0.00007	0.00001	−0.76544	0.00284
ALD	0.0	0.0	0.0	0.0	0.14336
TPI	0.0	0.0	0.0	0.0	0.01432
GAPDH	0.0	0.0	0.0	0.0	1.45508
PGK	0.0	0.0	0.00001	0.0	0.76405
PGM	0.0	0.0	0.00001	0.0	0.61275
EN	0.0	0.0	0.00006	0.0	6.36451
PK	0.00026	0.00005	0.00139	0.0	1.36505
LDH	0.0	0.0	0.00001	0.0	0.12333
CK	0.0	0.0	0.0	0.0	0.0
ADK	0.0	0.0	0.0	0.0	0.0
ATPase	0.99971	0.99856	0.99994	−4.60501	1.00000
outflux	0.00029	0.00010	0.00062	0.0	1.25773

To illustrate our approach, we then assume that the V_{\max} values of the reactions are not known very well, which in fact is rather realistic since they depend on *in vivo* enzyme concentrations and these are not easy to determine. Therefore, we want to determine the variation of FCC by each reaction when the V_{\max} values of all reactions can vary within a defined range. We performed optimizations of the FCCs, determining both the minima and maxima that each one of the 16 FCCs can take in that parameter space. The maxima and minima of each FCC were calculated with a separate optimization run. Any parameter sets that did not result in a steady-state solution were discarded. Steady states were computed using the Newton method as well as by forward integration.

It is perhaps useful to consider the biological meaning of a certain reaction having a minimum and a maximum value of FCC. The minimum value means that in every possible condition within the allowed variation of *all* parameters, that particular reaction has always that amount of control or higher. The maximum value of FCC means that in any possible condition within the allowed range of variation, that reaction has at best that amount of control. It is important to realize that the search is global and all parameters can change.

Thus any reaction with a considerably large *minimum* value of FCC is always important in determining the flux. This would have interesting physiological and pharmacological implications: such a reaction is indeed very effective at decreasing or increasing the flux and therefore is a good drug target. This also has important consequences for modellers, who then need to be able to estimate rather accurately the values of the kinetic constants of that reaction (since any small changes would have large effects on the model). Conversely, if the *maximum* value of the FCC by a certain reaction is low, then in any possible condition within the range of variation allowed that reaction has almost no effect on the flux. Such a reaction would be a really bad drug target (even if very good drugs would be available for it), since inhibiting or activating it would have only negligible effect on the flux. For a modeller, this is a reaction that one can leave with rather crude estimates for its parameters, since more or less accuracy will have no effect on the model in general.

Results of optimizations using the particle swarm algorithm are shown in [table 1](#). We performed this exercise twice, the first time within a fairly tight range of values (1/2 and 2 times their nominal value in [Lambeth & Kushmerick 2002](#)), and with a much wider range the second time (1/10 and 10 times their nominal value in [Lambeth & Kushmerick 2002](#)). When the search space was tightly around the nominal values in the original publication, the maxima and minima of the FCCs did not differ much from those of the original parameter set. This means that all the V_{\max} values could be changed arbitrarily within that range without changing the conclusion that the flux is essentially controlled by the ATPase reaction.

When the search space under investigation was enlarged (columns 5 and 6 of [table 1](#)), dramatic changes were observed. Now several other reactions can gain high control of the flux in some regions of parameter space. As an example, the maximum control that phosphoglycerate kinase (PGK) can have over the flux is now 0.76405, suggesting that there are conditions (within the parameter range) where this reaction has significant control over the glycolytic flux. For other reactions, however, the maximal FCC is still 0.0 (creatine kinase and adenylate kinase) or small (e.g. phosphofructokinase or triose phosphate isomerase), meaning that even on a global scale and within a wide range of parameters these reactions do not control glycolytic flux. Several of the reactions—ATPase and the first reactions of glycolysis—are observed to be able to exert negative control, whereas none of the others are able to do so. (Negative control means that if the rate of those reactions is increased the flux is *decreased*.)

In order to test the reliability and performance of the method, we repeated the optimization runs several times using different optimization methods. As an example, [figure 2](#) shows the progress of the optimization of the maximum of the FCC exerted by pyruvate kinase on the flux of interest for several runs of different algorithms. The abscissa corresponds to the number of different sets of parameters for which a steady state has been calculated, and the ordinate corresponds to the largest FCC that was found after this number of calculations was performed. Each curve corresponds to one optimization run. The stochasticity of the algorithms used results in a different trajectory each time. By comparing the curves for the repeated runs of the same algorithm, we can estimate how long the algorithm typically takes to produce a useful value for the maximum of the FCC.

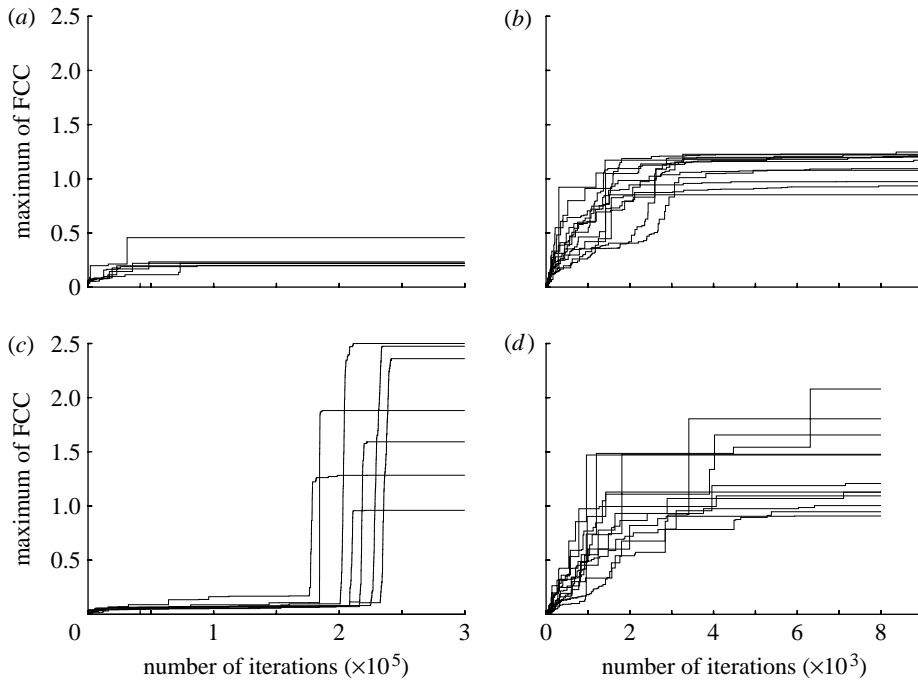


Figure 2. Progress of optimization runs using different algorithms. The ordinate is the candidate value of the maximal FCC (of pyruvate kinase towards the glycolytic flux) while the abscissa is the number of steady-state calculations performed. These runs correspond to the optimizations with a parameter range of V_{\max} within 0.1–10 times their original value. The number of steady-state calculations is roughly proportional to the computational time. On a current CPU (2.2 GHz), 1000 steady-state calculations for this model using numerical integration combined with the Newton method take approximately 6 min. (a) Random search (number of iterations=300 000) and (c) simulated annealing (using default parameters in COPASI). (b) Particle swarm (iteration limit=300, swarm size=30) and (d) genetic algorithm (number of generations=400, population size=20).

Obviously, the larger the high-dimensional parameter space, the more difficult it is to find the global optimum of the sensitivities. In this case, we have a 16-dimensional parameter space, where each parameter may span a range of two orders of magnitude. We have no indication that any of the algorithms we used was able to find the absolute global maximum, i.e. the absolute largest or smallest FCC within the given range of parameters. Indeed, every single optimization run resulted in a different optimum, corresponding to a different parameter set. This difficulty is expected due to the high dimensionality of the parameter space and the existence of multiple minima/maxima of comparable values. It is also partly because we did not allow the optimization algorithms to run for long enough time. However, we note that the particle swarm and GAs were in all cases able to estimate the order of magnitude of the maximum or minimum reliably within less than 5000 steady-state calculations. We never observed extreme situations where one run of the optimization would find a control coefficient to be insignificant, while another run would find it to be potentially significant. This means that the technique as presented here is sufficient to quickly

evaluate the importance of the different reaction steps for the flux control. It also means that one does not need to carry out excruciatingly long optimization runs (which would eventually get very close to the true global optima).

We observed that in some cases, singularities appeared in some FCCs, where these coefficients exhibit very large spikes (with values larger than 1000) in a very narrow region of parameter space. These singularities are not reliably found in each run.

Comparing the computational efficiency of the different methods and checking the number of function evaluations, it becomes apparent that both particle swarm and GA were quite efficient for our purpose and the given model. Simulated annealing was approximately one order of magnitude slower in convergence. Random search was even slower, and failed to produce comparable optima within 100 000 function evaluations, which is 20 times the computational effort required by the best optimizers to find the optima.

5. Discussion

We presented a new strategy to assess sensitivities of a biochemical model, specifically control coefficients, exerted by the model's parameters. Motivated by the large degree of uncertainty in such models, we argue that it is necessary to use methods beyond the conventional local sensitivity analysis, for example MCA, to assess which parameters need to be determined with higher accuracy.

Our approach features the optimization of control coefficients in a given parameter space. The knowledge of the minimal and maximal values for these coefficients allows a more confident estimate of the importance of the respective reactions. We applied our method to a model of central metabolism in muscle by Lambeth & Kushmerick (2002) as a case study. We find for this model that the original results remain valid—namely the high control by ATPase—if the confidence interval for the parameters (V_{\max} in this case) is just to 0.5–2 times the original values. However, when this interval was increased, i.e. when we had less confidence in the values of these parameters, the situation changed dramatically. Several other reactions are able to exert high levels of control over the glycolytic flux. However, we also observed that even in this wider range, there are still reactions that are not able to have significant flux control. These findings are important information that should be obtained in the study of a biochemical model in general. They are able to answer the following questions.

- Which reactions are potentially sensitive for partial inhibitors (drugs) and which are more robust?
- Which parameters (in this case V_{\max}) should be determined with more accuracy to increase confidence in the sensitivity analysis?

Owing to the high dimensionality and complexity of the problem of finding global optima, there are some limitations in this approach. Thus, if the quest is to be finished within a short time scale, it is sometimes hard to reproduce the exact value of the optimum. However, the trends (robust or sensitive) are highly reproducible when using the more efficient optimization methods (particle swarm and GA). In addition, the comparison with random search shows that this

limitation is even stronger with sampling approaches. The same applies to the problem of the presence of singularities with very high values of the control coefficient. These were generally not found by random search.

We believe that the approach presented is an efficient means to find the general trends of control and sensitivity in complex biochemical systems. It is clearly more efficient than random sampling. More models have to be checked to study the scalability of the approach; however, we note that the model tested here was already not that small.

We are excited by the ease with which this method can be applied, which places it within the reach of modellers. With COPASI, at least, one can very quickly program the steps we described even for very large models. This should make it easy to obtain a quick but fairly accurate idea of which reactions are important in a model at a global scale, without having to carry out complex Monte Carlo samplings that require specialized software and very long computational runs.

We thank the NIGMS (grant GM080219), the Klaus Tschira Foundation, the BMBF and the EPSRC for generous funding of this project. Matthias Reuss is acknowledged for fruitful discussions.

References

- Bentele, M., Lavrik, I., Ulrich, M., Stösser, S., Heermann, D. W., Kalthoff, H., Krammer, P. H. & Eils, R. 2004 Mathematical modeling reveals threshold mechanism in CD95-induced apoptosis. *J. Cell Biol.* **166**, 839–851. (doi:10.1083/jcb.200404158)
- Beyer, H.-G. 2001 *The theory of evolution strategies*. Berlin, Germany: Springer.
- Butcher, E. C. 2005 Can cell systems biology rescue drug discovery? *Nat. Rev. Drug Discov.* **4**, 461–467. (doi:10.1038/nrd1754)
- Campolongo, F., Saltelli, A. & Tarantola, S. 2000 Sensitivity analysis as an ingredient of modeling. *Stat. Sci.* **15**, 377–395. (doi:10.1214/ss/1009213004)
- Cascante, M., Boros, L. G., Comin-Anduix, B., de Atauri, P., Centelle, J. J. & Lee, P. W. N. 2002 Metabolic control analysis in drug discovery and disease. *Nat. Biotechnol.* **20**, 243–249. (doi:10.1038/nbt0302-243)
- Corana, A., Marchesi, M., Martini, C. & Ridella, S. 1987 Minimizing multimodal functions of continuous variables with the simulated annealing algorithm. *ACM Trans. Math. Software* **13**, 262–280. (doi:10.1145/29380.29864)
- Eberhart, R. C. & Kennedy, J. 1995 A new optimizer using particle swarm theory. In *Proc. of the 6th Int. Symp. on Micromachine and Human Science, Nagoya, Japan*, pp. 39–43.
- Heinrich, R. & Rapoport, T. A. 1974 A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *Eur. J. Biochem.* **42**, 89–95. (doi:10.1111/j.1432-1033.1974.tb03318.x)
- Heinrich, R. & Schuster, S. 1996 *The regulation of cellular systems*. Berlin, Germany: Springer.
- Hoops, S. et al. 2006 COPASI—a COMplex PATHway Simulator. *Bioinformatics* **22**, 3067–3074. (doi:10.1093/bioinformatics/btl485)
- Hucka, M. et al. 2003 The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531. (doi:10.1093/bioinformatics/btg015)
- Kacser, H. & Burns, J. A. 1973 The control of flux. *Symp. Soc. Exp. Biol.* **27**, 65–104.
- Kirkpatrick, S., Gelatt Jr, C. D. & Vecchi, M. P. 1983 Optimization by simulated annealing. *Science* **220**, 671–680. (doi:10.1126/science.220.4598.671)
- Lambeth, M. J. & Kushmerick, M. J. 2002 A computational model for glycogenolysis in skeletal muscle. *Ann. Biomed. Eng.* **30**, 808–827. (doi:10.1114/1.1492813)

- Le Novère, N. *et al.* 2006 BioModels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res.* **34**, D689–D691. (doi:10.1093/nar/gkj092)
- Ludtke, N., Panzeri, S., Brown, M., Broomhead, D. S., Knowles, J., Montemurro, M. A. & Kell, D. B. 2008 Information-theoretic sensitivity analysis: a general method for credit assignment in complex networks. *J. R. Soc. Interface* **5**, 223–235. (doi:10.1098/rsif.2007.1079)
- Mendes, P. & Kell, D. B. 1998 Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics* **14**, 869–883. (doi:10.1093/bioinformatics/14.10.869)
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N. & Teller, A. H. 1953 Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092. (doi:10.1063/1.1699114)
- Mitchell, M. 1995 *An introduction to genetic algorithms*. Boston, MA: MIT Press.
- Stephanopoulos, G. & Vallino, J. J. 1991 Network rigidity and metabolic engineering in metabolite overproduction. *Science* **252**, 1675–1681. (doi:10.1126/science.1904627)
- Van Riel, N. A. 2006 Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. *Brief Bioinform.* **7**, 364–374. (doi:10.1093/bib/bbl040)
- Wolpert, D. H. & Macready, W. G. 1997 No free lunch theorems for optimization. *IEEE Trans. Evol. Comp.* **1**, 67–82. (doi:10.1109/4235.585893)