

## FEDMA – a simple algorithm for theoretical modeling of linear metabolic pathways: from fuzzy data sets to prediction and experiment

Piotr H. Pawlowski<sup>1</sup>✉ and Piotr Zielenkiewicz<sup>1,2</sup>

<sup>1</sup>*Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warszawa, Poland;*

<sup>2</sup>*Plant Molecular Biology Laboratory, Warsaw University, Warszawa, Poland*

Received: 07 January, 2005; accepted: 02 March, 2005

**Key words:** chemical kinetics, theoretical modeling, algorithm, linear metabolic pathway

**A theoretical model of a chain of irreversible Michaelis-Menten reactions proceeding inside a living cell, taking cell growth, division and subcellular compartmentation into account, was proposed. It became a basis for the construction of a “fuzzy” enzymatic data-modeling algorithm (FEDMA) – a procedure allowing the estimation of missing parameter values for the modeled system, in accordance both with the derived theoretical rules and the available experimental data. The obtained tool was tested to model the heme biosynthesis pathway in *Saccharomyces cerevisiae*, where about 40% of parameters remain unknown. The missing parameters estimated by means of FEDMA fall in the range of expected values.**

Computer modeling is a powerful tool for the investigation of cellular metabolism. It allows us to analyze the mechanisms underlying various aspects of the cell's behavior. There are several existing conceptual tools designed for this purpose, for example the “Metabolic Control Analysis” approach (Schuster *et al.*, 2000), or the “power-law” ap-

proach for the modeling of biological systems (Savageau & Voit, 1982; Voit & Savageau, 1982a; 1982b). Appropriate computer software is also available: Gepasi (Mendes, 1993), Dynafit (Kuzmic, 1996), Plas (Voit & Ferreira, 1998), Stochsim (Morton-Firth & Bray, 1998), E-Cell (Tomita *et al.*, 1999), and others. These utilities, of high mathematical

---

✉Correspondence to: Piotr H. Pawlowski, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, A. Pawińskiego 5A, 02-106 Warszawa, Poland; fax: (48) 3912 1623; e-mail: [piotr@ibb.waw.pl](mailto:piotr@ibb.waw.pl)

**Abbreviations:** FEDMA, “fuzzy” enzymatic data-modelling algorithm; M-M, Michaelis-Menten; others abbreviations are given on page 76.

complexity, are capable of fitting a wide class of models to dynamic experimental data. They can determine the steady-state values of required variables, or analyze the direction of changes resulting from small disturbances of the system. Unfortunately, they are not well designed to deal with models of underdefined systems, for which some of the parameter values are lacking. They cannot use all the information derived from data concerning the system's behavior under modified conditions. They do not take into account the temporal changes of some physical parameters, such as compartment volume, which is important in models considering the metabolism of living, growing, and dividing cells over a large time scale. Furthermore, the incorporated iterative algorithms are more time-consuming than direct analytical calculations, which reduces the net time remaining for the modification of chosen parameters.

We therefore propose a new simple algorithm for the estimation of missing parameter values (the rate of the substrate external supply, Michaelis constants, turnover numbers, and concentrations of the total substrate and enzyme) in a simple case of a linear enzymatic pathway. To do this, we first created a general theoretical model of a chain of irreversible Michaelis-Menten (M-M) reactions proceeding inside a living cell. In its general formulation, the model includes a set of kinetic equations which describe a system of  $N$  non-degrading enzymes operating sequentially, according to M-M kinetics in a varying-volume compartment. Detailed mathematical analysis considers only the steady state (strictly the average state of periodized process – see Theoretical model, Scheme 1) and relates the mean concentrations of all included metabolites to fluxes and other intrinsic parameters of the setup.

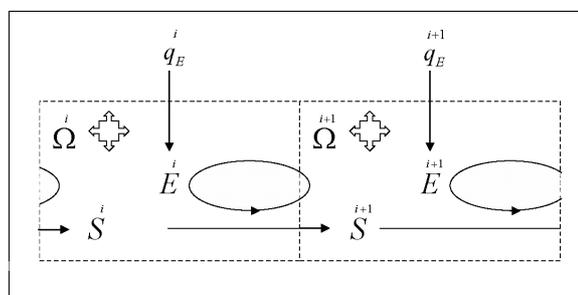
The formula derived lets us propose a “fuzzy” enzymatic data modeling algorithm (FEDMA). It allows fast estimation of unknown parameters, temporal means of concentrations or flux of the postulated meta-

bolic model. It makes it also possible to consider simultaneously several varying, modified copies of the system, and to treat them not as predictions, but as a source of additional information, as part of the initial setup. The output of the algorithm depends on which parameter values we consider reliable, and which are labeled as “uncertain” and subject to additional estimation. We decided to test this new approach by applying it to the yeast heme biosynthesis pathway. In this case, the following formulation of the special model was based on our previous knowledge of the subject and the data available in the literature concerning enzyme kinetics and mean intracellular metabolite concentrations.

#### THEORETICAL MODEL OF A CHAIN OF MICHAELIS-MENTEN REACTIONS PROCEEDING INSIDE A LIVING CELL

##### General model formulation

Let us consider a chain of  $N$  enzymatic M-M reactions proceeding inside a living (growing and dividing) cell, purveyed by adequate substrate and enzyme inflows, with each reaction taking place in a separate compartment of varying volume (Scheme 1):



**Scheme 1.** Theoretical model of a chain of M-M (Michaelis-Menten) reactions operating inside a living cell.

$$\dot{S}^i + \dot{E}^i \frac{k_{1+}^i}{k_{1-}^i} \{SE\}^i \frac{k_{2+}^i}{k_{2-}^i} S^i + \dot{E}^i \quad (i = 1, 2 \dots N) \quad (1)$$

The temporal behavior of such a system is generally described by the following set of indexed equations:

$$\frac{d}{dt} \dot{S}^i = -k_{1+}^i \dot{S}^i \dot{E}^i + k_{1-}^i \{SE\}^i + q_S^i - \dot{S}^i \frac{d}{dt} \ln(\Omega^i) \quad (2)$$

$$\frac{d}{dt} \dot{E}^i = -k_{1+}^i \dot{S}^i \dot{E}^i + k_{1-}^i \{SE\}^i + k_{2+}^i \{SE\}^i + q_E^i - \dot{E}^i \frac{d}{dt} \ln(\Omega^i) \quad (3)$$

$$\frac{d}{dt} \{SE\}^i = k_{1+}^i \dot{S}^i \dot{E}^i - k_{1-}^i \{SE\}^i - k_{2+}^i \{SE\}^i - \{SE\}^i \frac{d}{dt} \ln(\Omega^i) \quad (4)$$

where  $t$  means time,  $\Omega^i$  is the volume of the  $i$ -th compartment,  $S^i, E^i, \{SE\}^i$  are the local concentrations of the free substrate, free enzyme and substrate–enzyme complex for the  $i$ -th reaction, respectively, and  $k_{1+}^i, k_{1-}^i, k_{2+}^i$  denote the reaction rate coefficients for reaction  $i$ . Since the considered reactions form a chain, the values are not independent, and the product of reaction  $i-1$  is the substrate of reaction  $i$ . Therefore, the substrate influxes  $q_S^i$  are related to the rates of catalysis  $v^i$  (where  $v^i \stackrel{\text{def}}{=} k_{2+}^i \{SE\}^i$ ) as the following recurrent for-

$$q_S^i = \begin{cases} q_S^1 & i=1 \\ v^{i-1} \frac{\Omega^{i-1}}{\Omega^i} & i>1 \end{cases} \quad (5)$$

mula describes:

where  $q_S^1$  is a predefined initial substrate influx, which may be time-dependent. The predefined enzyme influxes  $q_E^i$  may also depend on time.

In eqns. 2–4, the components proportional to the logarithmic rate  $\frac{d}{dt} \ln(\Omega^i)$  describe changes in metabolite dilution, resulting from changes in compartment volume. It is assumed that all compartments increase their

volume during a period  $\tau$  and then return to their initial size. This volume increase-decrease scheme is repeated cyclically.

For the influxes  $q_S^i$  and  $q_E^i$  either constant patterns or also periodic ones with the interval  $\tau$ , are presupposed.

### Model development

As it was shown (Pawlowski & Zielonkiewicz, 2004) for a M-M reaction operating under the conditions of periodic changes in the process volume and constant or periodic (synchronous) influx rates, concentrations of total enzyme and also those of free substrate, free enzyme, substrate–enzyme complex and product can reach a state of stable oscillations around a certain constant level – an effect we termed “periodization”. In the case of the model presented above, by analogy, a system of periodically oscillating chemical phenomena with the period  $\tau$  should be achieved.

In general, the average state of a chain of “periodized” M-M reactions in a living cell can be quantitatively discussed by detailing the mean values  $\langle \underline{X}^i \rangle$  – mean concentrations of chosen metabolites, related to the total cell volume  $\Omega_c$  and the time interval  $\tau$ . Here, the symbol  $\langle \underline{X}^i \rangle$  denotes the weighted time-averaged

value  $\frac{1}{\tau} \int_t^{t+\tau} \dot{X}^i \frac{\Omega^i}{\Omega_c} d\xi$ , where  $d\xi$  represents infinitesimal time increase.

Mathematical analysis of the formulated model shows that for a chain of M-M reactions mean metabolite concentrations are related to the averaged substrate and enzyme influxes (as well as other parameters) by approximate flux balances, supplemented with the modified M-M equation:

$$\langle \underline{q}_S^i \rangle = \kappa \langle \underline{S}^{*i} \rangle + \langle \underline{v}^i \rangle \quad (6)$$

$$\langle \underline{q}_E^i \rangle = \kappa \langle \underline{E}^{*i} \rangle \quad (7)$$

$$\langle \underline{v}^i \rangle = k_{2+}^i \frac{\langle \underline{E}^{*i} \rangle \langle \underline{S}^i \rangle}{K^i + \langle \underline{S}^i \rangle} \quad (8)$$

where  $\langle \underline{S}^* \rangle$  and  $\langle \underline{E}^* \rangle$  are mean total substrate  $\left( \underline{S}^* = \underline{S} + \{SE\} \right)$  and enzyme  $\left( \underline{E}^* = \underline{E} + \{SE\} \right)$  concentrations, respectively,

$$\kappa = \frac{\ln \left( \frac{\Omega_c^{fin}}{\Omega_c^{ini}} \right)}{\tau} \quad (9)$$

is the mean logarithmic rate of cell volume change, with  $\Omega_c^{ini}$  and  $\Omega_c^{fin}$  representing the initial and final volume of a growing cell,

$$\overset{i}{K} = \frac{\left( \overset{i}{K}_m + \frac{\kappa}{\overset{i}{k}_{1+}} \right)}{\overset{i}{\Gamma}} \quad (10)$$

denotes the apparent equilibrium constant, with the well known Michaelis constant

$$\overset{i}{K}_m = \frac{\overset{i}{k}_{1-} + \overset{i}{k}_{2+}}{\overset{i}{k}_{1+}} \quad (11)$$

and  $\overset{i}{\Gamma}$  – a parameter describing the mean ratio of the cell and compartment volume, calculated as:

$$\overset{i}{\Gamma} = \frac{1}{\tau} \int_t^{t+\tau} \frac{\Omega_c}{\Omega} d\xi \quad (12)$$

The analyzed model allows us also to determine the mean substrate influxes  $\langle \underline{q}_S \rangle$ . This can be achieved by averaging of the substrate influx formula (Eqn. 5), which leads to the following relationship:

$$\langle \underline{q}_S \rangle = \begin{cases} \langle \underline{q}_S^1 \rangle & i = 1 \\ \langle \underline{v}^{i-1} \rangle & i > 1 \end{cases} \quad (13)$$

The method for  $\langle \underline{q}_E \rangle$  determination is out of the scope of this paper.

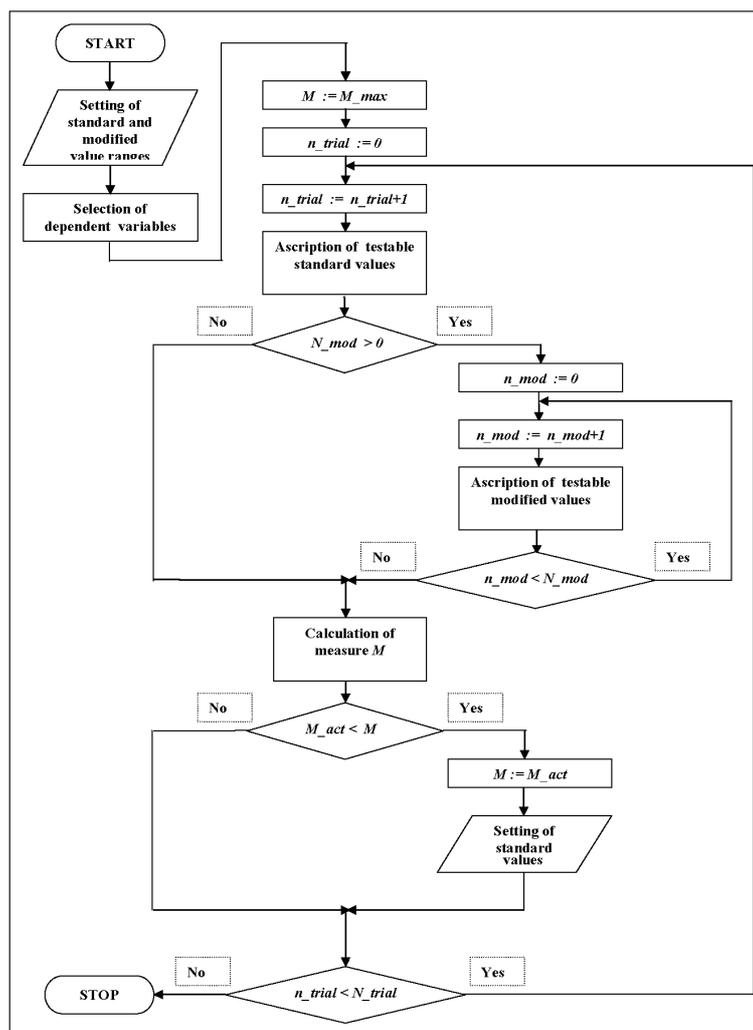
Equations 6–8 (supplemented with 9–12) can be easily derived by averaging of Eqns. 2–4, with the simplifying approximation:  $\langle \underline{ab} \rangle = \langle \underline{a} \rangle \langle \underline{b} \rangle$ . Along with Eqn. 13, they provide us with a set of general rules which have to be fulfilled by the averaged concentrations, fluxes and other parameters of any special model, if it is to remain consistent with the proposed general scheme.

### FEDMA – “FUZZY” ENZYMATICAL DATA MODELING ALGORITHM

The theoretical model presented above generally describes a non-branched, linear pathway of irreversible M-M reactions, and incorporates some features of living cells: growth, division and cellular compartmentation. Following it the FEDMA algorithm (Scheme 2) was proposed which allows estimation of the lacking  $\langle \underline{q}_S \rangle$ ,  $\langle \underline{S}^* \rangle$ ,  $\langle \underline{E}^* \rangle$ ,  $\overset{i}{K}$  and  $\overset{i}{k}_{2+}$  values for any special implementation in such a way that they fulfill the general rules (Eqns. 6–8, 13) and are in agreement with the available experimental data.

The values are sought within set boundaries that define acceptable ranges for each variable, according to the investigator’s knowledge. It is possible to introduce more than one set of ranges, representing a cell under different conditions (environmental change, mutants, etc.). By doing so we facilitate proper parameter estimation: since more information is fed into the program, the resulting values are more reliable. The value of the parameter  $\kappa$  is treated as a known constant. An arbitrarily defined measure  $M$  is introduced to evaluate the discrepancy between the current model values and the experimental data. In default, it describes the sum of relative distances of parameter proposed values from the central points of its assumed acceptable ranges, but it can be freely redefined. Minimization of  $M$  is the goal of FEDMA.

FEDMA consists of six main blocks. These are:



Scheme 2. Block diagram of FEDMA („fuzzy” enzymatic data-modeling algorithm).

- ◆ setting of standard and modified value ranges,
- ◆ selection of dependent variables,
- ◆ ascription of testable standard values,
- ◆ ascription of testable modified values,
- ◆ calculation of the measure  $M$ ,
- ◆ setting of standard values.

First, sets of standard and (optionally) modified ranges of the  $\langle \underline{q}_S^i \rangle$ ,  $\langle \underline{S}^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $K^i$  and  $k_{2+}^i$  values are set. Next, one “dependent” variable for each reaction is selected from the group:  $\langle \underline{S}^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $K^i$ , and  $k_{2+}^i$ , according to certain rules which take into consideration the width of the ranges and whether they are susceptible to modification. This value will be

calculated in a way that makes it consistent with the others, so that generally the least constrained one (acceptable in the widest range) is chosen. Then a course of repetitive trials is performed. In each trial, values are ascribed to all “testable” (independent) variables of the standard and modified sets, then the dependent values and finally the measure  $M$  are calculated.  $\langle \underline{q}_S^i \rangle$  and the values for the testable variables are chosen randomly from within the set ranges, and the dependent values are calculated according to the rules of the assumed theoretical model (Eqns. 6–8, 13). If modified conditions are also considered, a similar procedure is followed for the ascription of testable modified values, for each additional set. Here modified ranges are

considered, and those values which are not affected by the modification, are replaced by single values obtained during the fitting of standard. Finally, the measure  $M$  is calculated: as the sum of relative distances between the current values assigned for the testable (standard and modified) variables and the middle of the introduced acceptable ranges. In this manner, each whole trial is assigned a score (measure  $M$ ) according to how well all values fit the inserted data. This procedure is run iteratively, until the score is thought satisfactory. The obtained values are then saved as final standard model values. A detailed description of the algorithm, particular blocks and applied procedures is included in the Appendix.

The FEDMA can also be used as a simple substrate calculator. Running one trial for a set of single-point ranges of  $\langle \underline{q}_S^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $\overline{K}$  and  $k_{2+}^i$  and wide ranges of  $\langle \underline{S}^i \rangle$  yields new values for the  $\langle \underline{S}^i \rangle$  concentrations.

The computer program was implemented in Borland Delphi 2.0. This software was run on a PC-compatible computer based on a Pentium III/533 MHz processor with 128 Mbytes of RAM.

## AN EXEMPLARY TEST: MODELING OF HEME BIOSYNTHESIS

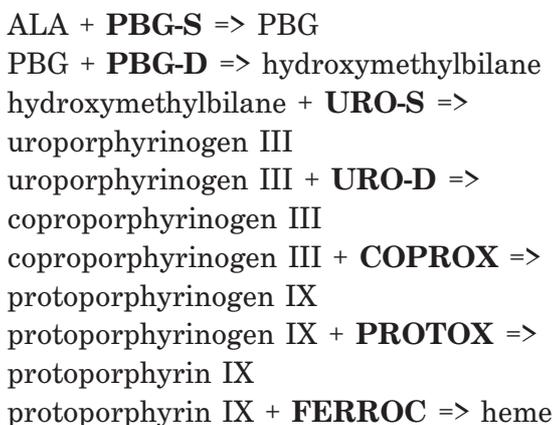
### *Setup for the special model*

To test the obtained numeric tool – FEDMA – we constructed a special model of a fragment of the heme biosynthesis pathway in *Saccharomyces cerevisiae* (Hoffman *et al.*, 2003).

Abbreviations used in the below text: ALA, 5-aminolevulinic acid; PBG, porphobilinogen; PBG-S, PBG synthase; PBG-D, PBG deaminase; URO-S, uroporphyrinogen III synthase; URO-D, uroporphyrinogen III decarboxylase; COPROX, coproporphyrino-

gen III oxidase; PROTOX, protoporphyrinogen IX oxidase; FERROC, ferrochelatase.

The modeled fragment encompasses a linear chain of seven reactions, leading from ALA to heme:



Due to procedural reasons, an eighth process (not listed), catalyzed by a “dead” enzyme, is additionally incorporated in order to complete the equations setup.

From the point of view of FEDMA full determination of the state of the considered system requires introducing ranges of experimental values for total substrate and enzyme concentrations ( $\langle \underline{S}^i \rangle$  and  $\langle \underline{E}^i \rangle$ ), as well as for the kinetic parameters ( $\overline{K}$  and  $k_{2+}^i$ ) of the included enzymes and the influx of the initial substrate  $\langle \underline{q}_S^i \rangle$ . Additionally, the single value  $\kappa$  is needed.

Approximate values of mean total metabolite concentrations  $\langle \underline{S}^i \rangle$  were derived from the literature (Table 1). In two cases of neighbor substrates (hydroxymethylbilane – uroporphyrin III and protoporphyrinogen IX – protoporphyrin IX) it was only possible to estimate the ranges for the summarized intracellular metabolite content, so we decided to attribute a range from zero to the maximum for the sum – to each of the related substrates. The additional contribution from these uncertainty to the score  $M$  was therefore added as the relative distance of the summarized concentrations from the centers of the ranges for the summarized values.

**Table 1. Content of some metabolites of the yeast heme biosynthesis pathway and their calculated intracellular concentrations.**

All data concern wild-type yeast cells cultured aerobically in medium containing glucose. Intracellular concentrations were calculated using the following values: dry weight of one cell =  $0.33 \cdot 10^{-10}$  [g] (unpublished results), cell volume =  $36.6 \cdot 10^{-15}$  [dm<sup>3</sup>] (Tyson *et al.*, 1979).

Metabolite	Content in the cell [nmol/g dry weight]	References	Intracellular concentration [M]
ALA	100–10 000	Rytka <i>et al.</i> , 1984; Carvajal <i>et al.</i> , 1990; Labbe-Bois & Volland, 1997; Labbe-Bois <i>et al.</i> , 1977	$9.0 \cdot 10^{-5}$ – $9.0 \cdot 10^{-3}$
PBG	50–60	Rytka <i>et al.</i> , 1984	$4.5 \cdot 10^{-5}$ – $5.4 \cdot 10^{-5}$
Uroporphyrins <sup>a</sup>	1.5–3.5	Labbe-Bois & Labbe, 1990; Rytka <i>et al.</i> , 1984; Kurlandzka <i>et al.</i> , 1988	$1.4 \cdot 10^{-6}$ – $3.2 \cdot 10^{-6}$
Coproporphyrin	1.5–7.5	Labbe-Bois & Labbe, 1990; Rytka <i>et al.</i> , 1984; Kurlandzka <i>et al.</i> , 1988; Pretlow & Sherman, 1967	$1.4 \cdot 10^{-6}$ – $6.8 \cdot 10^{-6}$
Protoporphyrins <sup>b</sup>	0–1.0	Labbe-Bois & Labbe, 1990; Rytka <i>et al.</i> , 1984; Kurlandzka <i>et al.</i> , 1988; Pretlow & Sherman, 1967	$0$ – $9 \cdot 10^{-7}$
Heme	70	Rytka <i>et al.</i> , 1984	$6.0 \cdot 10^{-5}$

<sup>a</sup>Products of spontaneous oxidation of both uroporphyrinogen I (the product of non-enzymatic cyclization of hydroxymethylbilane) and III. <sup>b</sup>Both protoporphyrin IX and the product of spontaneous oxidation of protoporphyrinogen IX.

Single-point (zero-length) ranges for total intracellular enzyme concentrations  $\langle \underline{E}^i \rangle$  could be estimated only for three of the seven enzymes (Table 2): URO-D, PROTOX, and FERROC. For the rest no data on enzyme content in yeast cells was to be found, so wide but reasonable ranges of possible values were established by preliminary simulations (for PBG-S and PBG-D the range  $0$ – $10^{-7}$  [M] was accepted, for URO-S:  $0$ – $10^{-9}$  [M], and for

COPROX:  $0$ – $10^{-6}$  [M]). These were not included during calculation of the measure  $M$ .

Literature data on kinetic parameters was available (Table 3), yielding single-point ranges for  $k_{2+}^i$ . Along with the assumptions:  $\kappa \ll k_{2-}^i + k_{2+}^i$  and  $\Gamma = 1$ , it also allowed us to estimate the ranges for  $K^i$ , basing on the ranges for  $K_m^i$ . In practice only one such range had a non-zero length (for the enzyme URO-S).

**Table 2. Total enzyme content for three enzymes of the yeast heme biosynthesis pathway and their calculated intracellular concentrations.**

Intracellular concentrations were calculated using the following values: dry weight of one cell =  $0.33 \cdot 10^{-10}$  [g] (unpublished results), cell volume =  $36.6 \cdot 10^{-15}$  [dm<sup>3</sup>] (Tyson *et al.*, 1979), total cellular proteins = 55% of the cell's dry weight (Alberts *et al.*, 1983).

Reaction number ( <i>i</i> )	Enzyme	Content in the cell	References	Intracellular concentration [M]
4	URO-D	0.007% of total cellular proteins	Felix & Brouillet, 1990	$8.7 \cdot 10^{-7}$
6	PROTOX	3 pmol/mg of total cellular proteins	Camadro <i>et al.</i> , 1986	$1.5 \cdot 10^{-6}$
7	FERROC	0.005% of total cellular proteins	Camadro <i>et al.</i> , 1994	$6.3 \cdot 10^{-7}$

**Table 3. Experimental values for kinetic parameters of the yeast heme biosynthesis enzymes.**

Reaction number ( <i>i</i> )	Enzyme	$K_m^i$ [M]	$k_{2+}^i$ [1/s]	References
1	PBG-S	$3.6 \cdot 10^{-4}$	1.2	Borrvalho <i>et al.</i> , 1990
2	PBG-D	$1.9 \cdot 10^{-5}$	$0.5^a$	Correa Garcia <i>et al.</i> , 1991; Shoolingin-Jordan, 1995
3	URO-S	$5.0 \cdot 10^{-6}$ – $2.6 \cdot 10^{-5b}$	$250^a$	Shoolingin-Jordan, 1995; Alwan <i>et al.</i> , 1989; Hart & Battersby, 1985; Tsai <i>et al.</i> , 1987
4	URO-D	$6.0 \cdot 10^{-9}$	0.021	Felix & Brouillet, 1990
5	COPROX	$5.0 \cdot 10^{-8}$	0.07	Camadro <i>et al.</i> , 1986
6	PROTOX	$1.0 \cdot 10^{-7}$	0.62	Camadro <i>et al.</i> , 1994

<sup>a</sup>No data on *S. cerevisiae* enzymes available; mean values of parameters for enzymes from other organisms are given. <sup>b</sup>No data on *S. cerevisiae* enzymes available; range of values occurring in other organisms is given.

By definition, for the last – “dead” – enzyme operating with heme  $k_{2+1}^8 = 0$  was fixed, at arbitrary chosen  $K_m^8$  and  $\langle \underline{E}^* \rangle$ .

A logarithmic rate of cell volume change was established as  $\kappa = 1.28 \cdot 10^{-4} [\text{s}^{-1}]$ , according to eqn. 9, where:  $\Omega_c^{fin} / \Omega_c^{ini} = 2$  and  $\tau = 1.5$  [h] (average lifetime of *S. cerevisiae* cells under standard laboratory conditions) were assumed.

The unknown hypothetical range for  $\langle \underline{q}_S^1 \rangle$  was also established in preliminary experiments ( $8 \cdot 10^{-8}$ – $9 \cdot 10^{-8}$  [Ms<sup>-1</sup>]).

We decided to include some information concerning the pathway’s behavior in mutant yeast strains as well, in order to facilitate ap-

propriate parameter estimation, so data on porphyrin accumulation in *hem12* and *hem15* mutant strains blocked in heme biosynthesis (lacking in the URO-D or ferrochelatase enzymes, respectively) was also introduced into the program, giving two additional setups of modified  $\langle \underline{S}^i \rangle$  ranges (Table 4). For simplicity, modified ranges were attributed only to Uroporphyrinogen III and Protoporphyrin IX. The appropriate enzyme concentrations were reduced to zero, and for the concentrations of heme and the intermediates downstream of the block wide ranges (0–100 [M]) were set, in order to monitor “numeric leakage”; other parameters were left unchanged.

**Table 4. Data on *hem12* and *hem15* yeast mutants, incorporated into the model.**

All data concern yeast cultured aerobically in medium containing glucose. Intracellular concentrations were calculated using the following values: dry weight of one cell =  $0.33 \cdot 10^{-10}$  [g] (unpublished results), cell volume =  $36.6 \cdot 10^{-15}$  [dm<sup>3</sup>] (Tyson *et al.*, 1979)

Relevant genotype	Metabolite accumulated	Content in the cell [nmol/g of dry weight]	References	Intracellular concentration [M]
<i>hem12-3</i>	Uroporphyrins <sup>a</sup>	70–84	Żołądek <i>et al.</i> , 1996	$6.1 \cdot 10^{-5}$ – $7.2 \cdot 10^{-5}$
<i>hem15-3</i>	Protoporphyrins <sup>b</sup>	65–70	Żołądek <i>et al.</i> , 1996	$6.0 \cdot 10^{-5}$ – $6.1 \cdot 10^{-5}$

<sup>a</sup>Products of spontaneous oxidation of both uroporphyrinogen I (the product of non-enzymatic cyclization of hydroxymethylbilane) and III. <sup>b</sup>Of this: about 70% corresponds to protoporphyrin IX and 30% to the product of spontaneous oxidation of protoporphyrinogen IX.

In this way, by running  $10^5$  FEDMA trials, we managed to determine a set of standard values for all quantities required in the model (including all  $\langle \underline{S}^* \rangle$  values,  $\langle \underline{E}^* \rangle$  values for four enzymes, the one lacking  $K$  value for URO-S, and the influx rate  $\langle \underline{q}_S \rangle$  of ALA (Table 5)), giving proper results for the wild-type and *hem12Δ* and *hem15Δ* yeast strains (Fig. 1).

**Table 5. Standard (wild-type) values determined by FEDMA.**

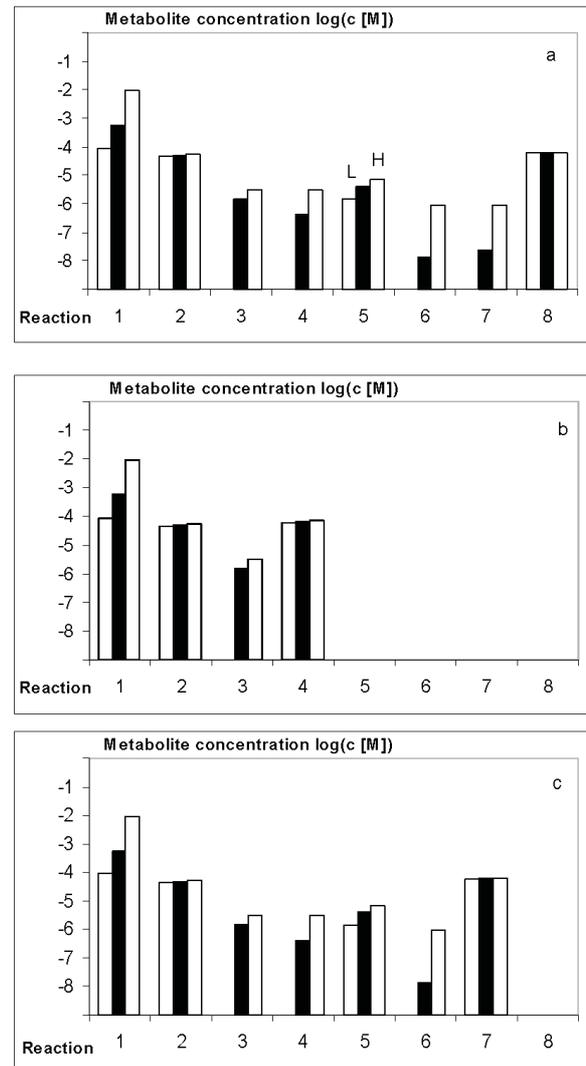
Reaction number ( <i>i</i> )	Mean total substrate concentration [M]	Mean total enzyme concentration [M]
1	$5.8 \cdot 10^{-4}$	$2.0 \cdot 10^{-8}$
2	$4.9 \cdot 10^{-5}$	$2.4 \cdot 10^{-8}$
3	$1.5 \cdot 10^{-6}$	$3.4 \cdot 10^{-10}$
4	$4.1 \cdot 10^{-7}$	$8.7 \cdot 10^{-7}{}^a$
5	$4.1 \cdot 10^{-6}$	$1.2 \cdot 10^{-7}$
6	$1.4 \cdot 10^{-8}$	$1.5 \cdot 10^{-6}{}^a$
7	$2.4 \cdot 10^{-8}$	$6.3 \cdot 10^{-7}{}^a$
8	$6.2 \cdot 10^{-5}$	“dead” enzyme arbitrary value

Additional parameters:

Influx of ALA	$8.9 \cdot 10^{-8}$ [M/s]
$K$ for URO-S	$1.3 \cdot 10^{-5}$ [M]

<sup>a</sup>Experimental values.

We did some modifications of the standard project to consider the effects of compartmentation and budding. Three analogous investigations with FEDMA were done, but replacing  $\Gamma = 1$ , by:  $\Gamma = 10$ , or  $\Gamma = 100$  – for all mitochondrial enzymes (PBG-S, PROTOX, FERROC), or  $\Omega_c^{fin}/\Omega_c^{ini} = 2$  by  $\Omega_c^{fin}/\Omega_c^{ini} = 1.25$  for the estimation of  $\kappa$ , leaving other quantities unchanged. Relative differences usually less than 10% (not shown) between the newly obtained and standard set of values, when compared with the hypothetical ranges, were observed.



**Figure 1. FEDMA estimations of mean metabolite (substrate) concentrations for the wild-type (a) and mutant strains (b, c).**

For each metabolite the minimal (L) and maximal (H) experimental value is shown and the single-point value assigned by the algorithm is depicted between them. Logarithmic scale was used. Values smaller than  $10^{-9}$  are not shown.

To check the predictive value of the finally constituted model we applied FEDMA as a substrate calculator and conducted a computer simulation of yeast cells being supplied with additional ALA (5-fold increase in the  $\langle \underline{q}_S \rangle$  value relatively to standard, and  $\langle \underline{E}^* \rangle$ ,  $K$ ,  $k_{2+}$  values unchanged). As a result, with ALA concentration 5 times higher than

under normal conditions, heme production remained unchanged. Some of the intermediate concentrations did increase: PBG 2-fold and coproporphyrinogen 3-fold, and the two substrates between them very slightly (less than 1.4 times).

## DISCUSSION

The proposed general theoretical model of a chain of M-M reactions inside a living cell, when compared to the classic one (Goldbeter, 1991) includes important novel features: cell growth, cell division and cellular compartmentation. These new elements allow for the description of the effect of metabolite diluting during cell growth, as the physical process having, with the processes of metabolite inflow (synthesis) and outflow (conversion), a simultaneous impact on the resultant picture of the observed kinetics. Along with the “periodization” idea this leads to formulation of rules obeyed by the mean levels of fluxes and metabolite and enzyme concentrations during repetitive cycles of cell division.

The developed FEDMA algorithm is a practical application of the conclusions of the general model. As shown for the heme biosynthesis, it may be a useful tool in looking for the most reliable values of parameters when only “fuzzy” experimental data is available. The possibility of introducing additional sets of value ranges, corresponding to modifications of the investigated system, allows all the metabolic information available to be effectively used in the process of estimating the missing parameters.

It should be stressed that the presented algorithm comprises two important simplifications regarding the considered processes. These are: the irreversibility of enzymatic reaction and the hyperbolic type of kinetics. As such, it covers a broad range of approaches, among which the most obvious is the chain of enzymatic processes with non-reversible catalysis (or with negligible backward transport

of a product) preceded by the step of reversible substrate–enzyme binding, where the cooperativity with an effector ligand can be neglected. For the case of more complicated metabolic phenomena it needs further development.

The question is – are the above limitations acceptable in the analysis of the special case of heme biosynthesis pathway? According to KEGG (Ogata *et al.*, 1999) the map of porphyrin and chlorophyll metabolism (<http://www.genome.jp/kegg/pathway/map/map00860.html>) it is suggested that the considered pathway may be effectively modeled as a set of one-way phenomena, where the rate of a given step does not depend on its product. There is also an evidence for simple Michaelis-Menten kinetics in the cases of PBG-S (Senior *et al.*, 1996), PBG-D (Fumagalli *et al.*, 1991), COPROX (Mahlitz, 2002), and FERROC (Shi & Ferreira, 2004). On the other hand, there is no clear evidence for sigmoidal kinetics in the case of the rest of the enzymes. Although there are some reports about product inhibition of URO-S from *Euglena gracilis* (Hart & Battersby, 1985), susceptibility to substrate or protoporphyrin IX inhibition of URO-D from *Rhodobacter sphaeroides* (Jones & Jordan, 1993), inhibition of PROTOX from *Bos taurus* by the substrate at a high concentration ( $> 0.05$  [mM]) (Siepkner *et al.*, 1987), and physiologically insignificant inhibition of PROTOX from *Saccharomyces cerevisiae* by heme (Poulson & Polglase, 1975), the poor knowledge of relevant parameters describing metabolite cooperation in yeast, and the sake of simplicity, justify the accepted M-M approximation.

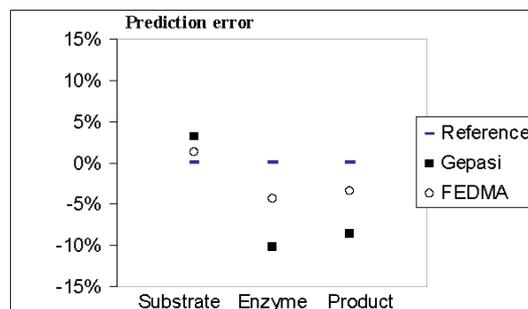
In our previous paper (Pawlowski & Zielenkiewicz, 2004) in Fig. 2 we showed that compartment volume change could essentially influence the kinetics and steady state behavior of an M-M process. According to our knowledge there is no kinetics software other than FEDMA where this effect is algorithmically included. The Gepasi simulator is one of those programs where it is possible to take

into account the consequences of cell volume expansion – by appropriate definition of some metabolite outflows. Gepasi can also optimize one set of unknown parameters, at one time. For comparison, our FEDMA – many sets of parameters with modified ranges, at one time, with a common accuracy measure.

To confront the most important properties of FEDMA and Gepasi we proposed a special comparative test. There, single M-M predefined reference processes in a varying volume compartment were considered, with a constant substrate inflow, a constant concentration of the total enzyme, accumulation of product, and metabolite diluting due to the processing volume increase. Details are included in the description of the attached presentation (Fig. 2). For precisely defined parameters, steady state metabolite concentrations were estimated as the limit values of the results of Gepasi “time course” simulation. Next, the enzyme concentration  $\langle \underline{E}^* \rangle_R$  of the reference process was  $\pm 100\%$  “fuzzed” in the range  $0-2 \langle \underline{E}^* \rangle_R$ . Then, its correct value and the proper values of the steady state metabolite concentration were reconstructed in the  $\pm 100\%$  “fuzzed” ranges, by the FEDMA and Gepasi optimization genetic algorithm. The test’s results show a better quality of the modeling by FEDMA. Moreover, the processing times adequately equal to 5 and 310 [s], respectively, also favor our algorithm.

In practice, the quality of predictions possible with FEDMA, depends, of course, on the amount and quality of information included in the investigation. As could be seen generally for the second reaction (Figs. 1a, b, c), the eighth reaction in wild-type strain (Fig. 1a), the fourth (Fig. 1b) and the seventh reaction (Fig. 1c) in mutants – FEDMA fit can be pretty good. For other reactions the relatively wide ranges of uncertainty, rising to two orders of magnitude, do not justify such an optimistic view. We knew from the literature (Labbe-Bois & Labbe, 1990) that supplying wild-type yeast cells with additional ALA in

the medium does not result in an increase of the pathway’s productivity: the amount of intracellular PBG remains unchanged, no porphyrin accumulation is detected and no additional heme synthesized, although the ALA uptake from the medium is effective and



**Figure 2. FEDMA and Gepasi comparison test.**

Single M-M predefined reference process in a varying volume compartment was considered, with constant rate of substrate inflow and constant concentration of the total enzyme. Accumulation of the product was assumed. For arbitrary but precisely defined parameters:  $\langle \underline{q}_S \rangle_R = 10^{-7}$  [M/s],  $\frac{1}{K_R} = 5 \cdot 10^{-4}$  [M],  $k_{2+R}^1 = 1$  [1/s] and  $\langle \underline{E}^* \rangle_R = 5 \cdot 10^{-8}$  [M] – steady state values,  $\langle \underline{S}^* \rangle_R$  and  $\langle \underline{S}^* \rangle_R$ , of the total substrate and the product concentration were obtained as limit values of the results obtained by the means of Gepasi “time course” algorithm within the 54 000 [s] range, where the metabolite diluting due to the compartment volume increase was modeled by the appropriate addition of the outflow equal to  $-\ln(2) \cdot [\text{metabolite concentration}] / \tau$ , with  $\tau = 5400$  [s]. For simplicity the free substrate concentration was assumed equal to that of the total substrate. Next, simulated investigations of the parameters of the above process were performed, but under condition of modeled uncertainty. To reach the goal, it was assumed that total enzyme concentration of the investigated process falls in the range  $0-2 \langle \underline{E}^* \rangle_R$ . Similarly it was assumed that the steady state values of total metabolite concentrations fall in the range:  $0-2 \langle \underline{S}^* \rangle_R$  and  $0-2 \langle \underline{S}^* \rangle_R$ . Then two attempts of investigations were performed: by Gepasi optimization genetic algorithm (generations = 200, population = 20) and FEDMA (no. of simulations,  $N_{\text{trial}} = 10^5$ ). Prediction error, i.e., the relative difference between prediction of a given algorithm and the real value of the investigated parameter, for desired concentrations of total substrate, total enzyme and the product, of the reference process is shown.

the intracellular ALA concentration rises significantly (depending on the concentration of ALA in the medium) – reaching a new steady state level, stabilized by the balance between inflow, outflow and diluting. Our simulation of a 5-fold increase in the  $\langle \underline{q}_S \rangle$  value predicted ALA accumulation and unchanged heme level. On the other hand, the accompanying predictions of an increase in PBG and coproporphyrinogen content are not consistent with the experimental data. This discrepancy shows that more information about the behavior of the modified system, especially for the conditions of a change in ALA influx,

The use of the term “fuzzy” data warrants a comment. Here the adjective “fuzzy” does not mean that the theory of fuzzy sets or fuzzy logics was applied *explicite*. It rather means that the simultaneously analyzed data are uncertain and additionally spread among many classes of physical quantities.

The presented algorithm can easily be implemented in any programming environment. The DELPHI implementation (Fig. 3) can be obtained from the authors (piotrp@ibb.waw.pl).

We are grateful to Joanna Rytka, with whom this project was initiated, and to her co-workers Monika Góra and Marta Hoffman for

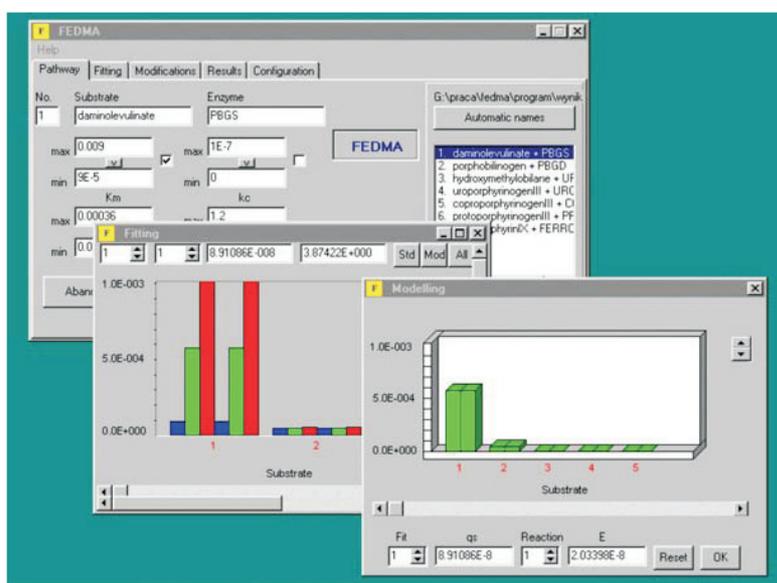


Figure 3. FEDMA. The DELPHI implementation.

ought to be incorporated into the data setup analyzed by FEDMA. In the case of other “*in silico*” experiments, further development of the model, for example addition of the omitted first enzyme in the pathway – ALA synthase – may be fruitful.

We believe that the proposed general model with the FEDMA algorithm can be useful for identifying the rate-controlling steps of the metabolic pathways inside the cell, analytical description of the pathways without feedback signaling, and other metabolic control investigations. Moreover, it could be also adopted “locally” for branched systems.

their invaluable help in the investigation of real parameter values and critical discussion of the proposed model of heme biosynthesis.

## REFERENCES

- Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. (1983) Macromolecules: structure, shape, and information. In *Molecular Biology of the Cell*. Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD, eds, pp 89–138. Garland Publishing Inc., New York.

- Alwan AF, Mgbeje BIA, Jordan PM. (1989) Purification and properties of uroporphyrinogen III synthase (co-synthase) from an overproducing recombinant strain of *Escherichia coli* K-12. *Biochem J.*; **264**: 397–402.
- Borrvalho LM, Ortiz CHD, Panek AD, Mattoon JR. (1990) Purification of  $\delta$ -aminolevulinatase dehydratase from genetically engineered yeast. *Yeast.*; **6**: 319–30.
- Camadro JM, Chambon H, Jolles J, Labbe P. (1986) Purification and properties of coproporphyrinogen oxidase from the yeast *Saccharomyces cerevisiae*. *Eur J Biochem.*; **156**: 579–87.
- Camadro JM, Thome F, Brouillet N, Labbe P. (1994) Purification and properties of protoporphyrinogen oxidase from the yeast *Saccharomyces cerevisiae*. *J Biol Chem.*; **269**: 32085–91.
- Carvajal E, Panek AD, Mattoon JR. (1990) Isolation and characterization of a new mutant of *Saccharomyces cerevisiae* with altered synthesis of 5-aminolevulinic acid. *J Bacteriol.*; **172**: 2855–61.
- Correa Garcia SR, Rossett MV, Batlle AM. (1991) Studies on porphobilinogen deaminase from *Saccharomyces cerevisiae*. *Z Naturforsch.*; **46**: 1017–23.
- Felix F, Brouillet N. (1990) Purification and properties of uroporphyrinogen decarboxylase from *Saccharomyces cerevisiae*. *Eur J Biochem.*; **188**: 393–403.
- Fumagalli SA, Kotler ML, Rossetti MV, Batlle AM. (1991) Studies on uroporphyrinogen biosynthesis in pig liver. *Z Naturforsch.*; **46**: 1101–8.
- Goldbeter A. (1991) A minimal cascade model for the mitotic oscillator involving cyclin and cdc2 kinase. *Proc Natl Acad Sci USA.*; **88**: 9107–11.
- Hart GJ, Battersby AR. (1985) Purification and properties of uroporphyrinogen III synthase (co-synthetase) from *Euglena gracilis*. *Biochem J.*; **232**: 151–60.
- Hoffman M, Góra M, Rytka J. (2003) Identification of rate-limiting steps in yeast heme biosynthesis. *Biochem Biophys Res Commun.*; **310**: 1247–53.
- Jones RM, Jordan PM. (1993) Purification and properties of the uroporphyrinogen decarboxylase from *Rhodobacter sphaeroides*. *Biochem J.*; **293**: 703–12.
- Kurlandzka A, Żołądek T, Rytka J, Labbe-Bois R, Urban-Grimal D. (1988) The effects *in vivo* of mutationally modified uroporphyrinogen decarboxylase in different *hem12* mutants of baker's yeast (*Saccharomyces cerevisiae*). *Biochem J.*; **253**: 109–16.
- Kuzmic P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data. Application to HIV proteinase. *Anal Biochem.*; **237**: 260–73.
- Labbe-Bois R, Labbe P. (1990) Tetrapyrrole and heme biosynthesis in the yeast *Saccharomyces cerevisiae*. In *Biosynthesis of Heme and Chlorophylls*. Dailey HA, ed, pp 235–86. McGraw-Hill Book Co., New York.
- Labbe-Bois R, Volland C. (1977) Changes in the activities of the protoheme-synthesizing system during the growth of yeast under different conditions. *Arch Biochem Biophys.*; **179**: 565–77.
- Labbe-Bois R, Rytka J, Litwińska J, Biliński T. (1977) Analysis of heme biosynthesis in catalase and cytochrome deficient yeast mutants. *Mol Gen Genet.*; **156**: 177–83.
- Mahlitz E. (2002) Die sauerstoffabhängige Coproporphyrinogen III Oxidase (HemF) aus *Escherichia coli*. Rekombinante Herstellung, biochemische und biophysikalische Charakterisierung. Der Gemeinsamen Naturwissenschaftlichen Fakultät der Technischen Universität Carolo-Wilhelmina zu Braunschweig. Dissertation.
- Mendes P. (1993) GEPASI: A software package for modelling the dynamics, steady states and control of biochemical and other systems. *Comput Applic Biosci.*; **9**: 563–71.
- Morton-Firth CJ, Bray D. (1998) Predicting temporal fluctuations in an intracellular signalling pathway. *J Theor Biol.*; **192**: 117–28.
- Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. (1999) KEGG: Kyoto encyclope-

- dia of genes and genomes. *Nucleic Acids Res.*; **27**: 29–34.
- Pawłowski PH, Zielenkiewicz P. (2004) Biochemical kinetics in changing volumes. *Acta Biochim Polon.*; **51**: 231–43.
- Poulson R, Polglase WJ. (1975) The enzymic conversion of protoporphyrinogen IX to protoporphyrin IX. Protoporphyrinogen oxidase activity in mitochondrial extracts of *Saccharomyces cerevisiae*. *J Biol Chem.*; **250**: 1269–74.
- Pretlow TP, Sherman F. (1967) Porphyrins and zinc porphyrins in normal and mutant strains of yeast. *Biochim Biophys Acta.*; **148**: 629–44.
- Rytka J, Biliński T, Labbe-Bois R. (1984) Modified uroporphyrinogen decarboxylase activity in a yeast mutant which mimics porphyria cutanea tarda. *Biochem J.*; **218**: 405–13.
- Savageau MA, Voit EO. (1982) Power-law approach to modeling biological systems; I. Theory. *J Ferment Technol.*; **60**: 221–8.
- Schuster S, Dandekar T, Mauch K, Reuss M, Fell DA. (2000) Recent developments in metabolic pathway analysis and their potential implications for biotechnology and medicine. In *Technological and Medical Implications of Metabolic Control Analysis*. Cornish-Bowden A, Cardenas ML, eds, pp 57–66. Kluwer Academic Publishers, Dordrecht.
- Senior NM, Brocklehurst K, Cooper JB, Wood SP, Erskine P, Shoolingin-Jordan PM, Thomas PG, Warren MJ. (1996) Comparative studies on the 5-aminolaevulinic acid dehydratases from *Pisum sativum*, *Escherichia coli* and *Saccharomyces cerevisiae*. *Biochem J.*; **320**: 401–12.
- Shi Z, Ferreira GC. (2004) Probing the active site loop motif of murine ferrochelatase by random mutagenesis. *Biol Chem.*; **279**: 19977–86.
- Shoolingin-Jordan PM. (1995) Porphobilinogen deaminase and uroporphyrinogen III synthase: structure, molecular biology, and mechanism. *J Bioenerg Biomemb.*; **27**: 181–96.
- Siepkner LJ, Ford M, de Kock R, Kramer S. (1987) Purification of bovine protoporphyrinogen oxidase: immunological cross-reactivity and structural relationship to ferrochelatase. *Biochim Biophys Acta.*; **913**: 349–58.
- Tomita M, Hashimoto K, Takahashi K, Shimizu T, Matsuzak Y, Miyosh F, Saito K, Tanida S, Yugi K, Venter JC, Hutchison C. (1999) E-CELL: Software environment for whole cell simulation. *Bioinformatics.*; **15**: 72–84.
- Tsai SF, Bishop DF, Desnick RJ. (1987) Purification and properties of uroporphyrinogen III synthase from human erythrocytes. *J Biol Chem.*; **262**: 1268–73.
- Tyson CB, Lord PG, Wheals AE. (1979) Dependency of size of *Saccharomyces cerevisiae* cells on growth rate. *J Bacteriol.*; **138**: 92–8.
- Voit EO, Savageau MA. (1982a) Power-law approach to modeling biological systems; II. Application to ethanol production. *J Ferment Technol.*; **60**: 229–32.
- Voit EO, Savageau MA. (1982b) Power-law approach to modeling biological systems; III. Methods of analysis. *J Ferment Technol.*; **60**: 233–41.
- Voit EO, Ferreira AEN. (1998) Buffering in models of integrated biochemical systems. *J Theor Biol.*; **191**: 429–38.
- Żołądek T, Nhi NB, Rytka J. (1996) *Saccharomyces cerevisiae* mutants defective in heme biosynthesis as a tool for studying the mechanism of phototoxicity of porphyrins. *Photochem Photobiol.*; **64**: 957–62.

## APPENDIX

The FEDMA algorithm (Scheme 2) consists of six main blocks. The initial two set up the variables for future processing, and the next four operate in the course of  $1-2 \cdot 10^9$  repetitive fitting trials. These are:

### The setting of standard and modified value ranges

Here, the sets of ranges for the standard values:  $[q_{s\_min}, q_{s\_max}]$ ,  $[Stot\_min[i], Stot\_max[i]]$ ,  $[Etot\_min[i], Etot\_max[i]]$ ,  $K\_min[i], K\_max[i]$ , and  $[kc\_min[i], kc\_max[i]]$ , for seeking the  $\langle \underline{q}_S^1 \rangle$ ,  $\langle \underline{S}^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $K$ , and  $k_{2+}^i$  values ( $i = 1, 2 \dots N$ ), respectively, are introduced into the program. There can also be up to 30 modified sets of ranges for any type of cell modification considered. If a certain value is known, a single-point range is set, where the lower and upper limits are equal.

### The selection of dependent variables

In this block, one "dependent" variable for each path unit (reaction) from each set is selected from the group:  $\langle \underline{S}^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $K$ , and  $k_{2+}^i$ , according to the weights attributed. The weights are calculated as the ratio of the width of the corresponding range to its central value. For values from the unmodified set, the weights are assigned the value -1 (minus one). The variable with the highest weight is chosen from each group. If two variables are of equal weight, the selection is done randomly. The results of the selection are saved.

### The ascription of testable standard values

Here, in each trial the value  $\langle \underline{q}_S^1 \rangle$ , is randomly chosen from the corresponding standard range. Then, the path unit index  $i$  is set as  $i = 1$ . Next:

- A. The mean substrate influx of the unit under consideration is established by appropriate procedure (P.1).
- B. The values of the three testable ("independent") variables from the group:  $\langle \underline{S}^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $K$ , and  $k_{2+}^i$  are chosen from the corresponding ranges.

C. The value of the fourth, dependent variable is calculated, according to the appropriate procedure (P. 2-5).

D. The value of the index  $i$  is increased by one:  $i = i + 1$ .

Steps A-D are conducted for all path units.

### The ascription of testable modified values

If any modified value sets are considered, the procedure for the ascription of values for the testable variables is performed for each of these sets. Those values which are supposed to remain unmodified are replaced by single-point values, obtained in the preceding fitting of the standard set.

### The calculation of measure $M$

Here the actual value of  $M$  for each trial is calculated. Initially, a large number  $M\_max$  is assigned to the measure  $M$ . In practice,  $M\_max = 1 \cdot 10^{+4932}$ . For the calculation of the current value of  $M$ , any arbitrarily defined formula using the fitted (standard and/or modified) values for  $\langle \underline{q}_S^1 \rangle$ ,  $\langle \underline{S}^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $K$ , and  $k_{2+}^i$  can be applied.

### The setting of standard values

If the measure  $M$  assigned to the current trial is smaller than the previous value of  $M$ , then the current values for the standard set are saved as the final standard model. Up to nine next-best trials are also saved.

### Procedures

For given  $\langle \underline{q}_S^i \rangle$  and  $\kappa$ , for every three arbitrarily chosen variables from the group:  $\langle \underline{S}^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $K$ , and  $k_{2+}^i$  the value of the fourth variable, and the metabolite influx of the next path unit, can always be established in such a way that the group, the inflow and

the outflow for the  $i$ -th unit will remain in agreement with the rules of the assumed theoretical model (eqns. 6–8, 13). This is achieved by the procedures applied in FEDMA. They permit the determination of:

P.1:  $\langle \underline{q}_S^i \rangle$  – for assumed  $\langle \underline{q}_S^{i-1} \rangle$ ,  $\langle \underline{S}^{i-1} \rangle$ , and  $\kappa$  values, as:

$$\langle \underline{q}_S^i \rangle = \begin{cases} \langle \underline{q}_S^1 \rangle & i=1 \\ \langle \underline{q}_S^{i-1} \rangle - \kappa \langle \underline{S}^{i-1} \rangle & i>1 \end{cases} \quad (\text{A.1})$$

P.2:  $\langle \underline{S}^{i*} \rangle$  – for assumed  $\langle \underline{q}_S^i \rangle$ ,  $\langle \underline{E}^i \rangle$ ,  $K$ ,  $k_{2+}^i$ , and  $\kappa$  values, as:

$$\langle \underline{S}^i \rangle = \langle \underline{S}^i \rangle_1 + \frac{\langle \underline{E}^i \rangle - \langle \underline{S}^i \rangle_1}{K + \langle \underline{S}^i \rangle_1} \quad (\text{A.2})$$

P.3:  $\langle \underline{E}^{i*} \rangle$  – for assumed  $\langle \underline{q}_S^i \rangle$ ,  $\langle \underline{S}^{i*} \rangle$ ,  $K$ ,  $k_{2+}^i$ , and  $\kappa$  values, as:

$$\langle \underline{E}^{i*} \rangle = \frac{(\langle \underline{S}^{i*} \rangle - \langle \underline{S}^i \rangle_2)(K + \langle \underline{S}^i \rangle_2)}{\langle \underline{S}^i \rangle_2} \quad (\text{A.3})$$

P.4:  $K$  – for assumed  $\langle \underline{q}_S^i \rangle$ ,  $\langle \underline{S}^{i*} \rangle$ ,  $\langle \underline{E}^{i*} \rangle$ ,  $k_{2+}^i$ , and  $\kappa$  values, as:

$$K = \frac{\langle \underline{E}^{i*} \rangle - \langle \underline{S}^i \rangle_2}{(\langle \underline{S}^{i*} \rangle - \langle \underline{S}^i \rangle_2)} \quad (\text{A.4})$$

P.5:  $k_{2+}^i$  – for assumed  $\langle \underline{q}_S^i \rangle$ ,  $\langle \underline{S}^{i*} \rangle$ ,  $\langle \underline{E}^{i*} \rangle$ ,  $K$ , and  $\kappa$  values, as:

$$k_{2+}^i = \frac{\langle \underline{q}_S^i \rangle - \kappa \langle \underline{S}^{i*} \rangle}{(\langle \underline{S}^{i*} \rangle - \langle \underline{S}^i \rangle_3)} \quad (\text{A.5})$$

where  $\langle \underline{S}^i \rangle_1$ ,  $\langle \underline{S}^i \rangle_2$ ,  $\langle \underline{S}^i \rangle_3$  represent three different ways of determining  $\langle \underline{S}^i \rangle$ , as a function of certain quantities, including special chosen variables, and excluding others.

Strictly speaking:

$$\langle \underline{S}^i \rangle_1 = \frac{1}{2} \left[ \frac{\langle \underline{q}_S^i \rangle}{\kappa} - \left( 1 + \frac{k_{2+}^i}{\kappa} \right) \langle \underline{E}^{i*} \rangle - K + \sqrt{\left( \frac{\langle \underline{q}_S^i \rangle}{\kappa} - \left( 1 + \frac{k_{2+}^i}{\kappa} \right) \langle \underline{E}^{i*} \rangle - K \right)^2 + 4 \frac{\langle \underline{q}_S^i \rangle}{\kappa} K} \right] \quad (\text{A.6})$$

$$\langle \underline{S}^i \rangle_2 = \left( 1 + \frac{\kappa}{k_{2+}^i} \right) \langle \underline{S}^{i*} \rangle - \frac{\langle \underline{q}_S^i \rangle}{k_{2+}^i} \quad (\text{A.7})$$

$$\langle \underline{S}^i \rangle_3 = \frac{1}{2} \left[ \langle \underline{S}^{i*} \rangle - \langle \underline{E}^{i*} \rangle - K + \sqrt{\left( \langle \underline{S}^{i*} \rangle - \langle \underline{E}^{i*} \rangle - K \right)^2 + 4 \langle \underline{S}^{i*} \rangle K} \right] \quad (\text{A.8})$$