BIOINFORMATICS

Vol. 16 no. 11 2000
Pages 1023–1037

Biochemical systems analysis of genome-wide expression data

Eberhard O. Voit and Tomas Radivoyevitch*

Department of Biometry and Epidemiology, 135 Rutledge Avenue, Medical University of South Carolina, Charleston, SC 29425, USA

Received on November 17, 1999; revised on April 6, 2000; accepted on May 23, 2000

Abstract

Motivation: *Modern methods of genomics have produced an unprecedented amount of raw data. The interpretation and explanation of these data constitute a major, well-recognized challenge.*

Results: *Biochemical Systems Theory (BST) is the mathematical basis of a well-established methodological framework for analyzing networks of biochemical reactions. An existing BST model of yeast glycolysis is used here to explain and interpret the glycolytic gene expression pattern of heat shocked yeast. Our analysis demonstrates that the observed gene expression profile satisfies the primary goals of increased ATP, trehalose, and NADPH production, while maintaining intermediate metabolites at reasonable levels. Based on a systematic exploration of alternative, hypothetical expression profiles, we show that the observed profile outperforms other profiles.*

Conclusion: *BST is a useful framework for combining DNA microarray data with enzymatic process information to yield new insights into metabolic pathway regulation.*

Availability: *All analyses were executed with the software PLAS©, which is freely available at http://correio.cc.fc.ul.pt/~aenf/plas.html for academic use.*

Contact: *VoitEO@MUSC.edu*

Introduction

The complete sequencing of several prokaryote and eukaryote genomes, combined with the invention of microarrays, has produced an unprecedented amount of raw data. The interpretation of these data constitutes a well-recognized challenge (DeRisi *et al.*, 1997). The difficulties in explaining the data derive from the underlying genetic and biochemical networks, which are generally very complex and large in size. A true understanding of their function requires computational approaches (Brownstein *et al.*, 1998; Bailey, 1999), which, in turn, are critically dependent on effective mathematical models (Savageau, 1991a,b; Thornton, 1998).

We propose that *Biochemical Systems Theory* (BST) offers such models and therefore provides a tool for the analysis and functional interpretation of genomic data. We show that BST can yield explanations for why some genes respond to an environmental stimulus with slight or strong over-expression or under-expression, whereas other genes apparently do not respond at all. Sometimes the explanations are rather evident, but at other times the observations are puzzling.

As a prototype, our analysis deals with glycolysis in *Saccharomyces cerevisiae* (see Figure 1). This system is well known and small enough to afford a simple illustration of the power and potential of a strong biomathematical approach. Nonetheless, the principles presented here are the same for moderate or large systems and can be applied toward effective analyses of highly regulated metabolic pathways of biologically relevant size. In order to keep the discussion as simple and lucid as possible, the analysis here focuses on data at only one time point after a stimulus. An extended analysis (Voit, in preparation) will take account of dynamical aspects of gene expression patterns as well as enzyme availability.

Data

We use a small portion of two extensive sets. The primary set consists of spotted array (Schena *et al.*, 1995) heat shock data in yeast (http://rana.Stanford.EDU/clustering/Figure2.txt; Eisen *et al.*, 1998). This data set contains relative expression levels of over 2000 yeast genes at 0, 10, 20, 40, 80, and 160 min after transition from 25 to 37 °C. The overall trend in gene expression is a more or less simultaneous deviation from baseline, which reaches its maximum in most cases at about 10 min. Afterwards, essentially all expression slowly returns to baseline toward the end of the observation period. To make our analysis as lucid as possible, we consider only genes coding for glycolytic enzymes and restrict the analysis to measurements 10 min after heat shock (Table 1). Other time points will be considered elsewhere.

*To whom correspondence should be addressed.

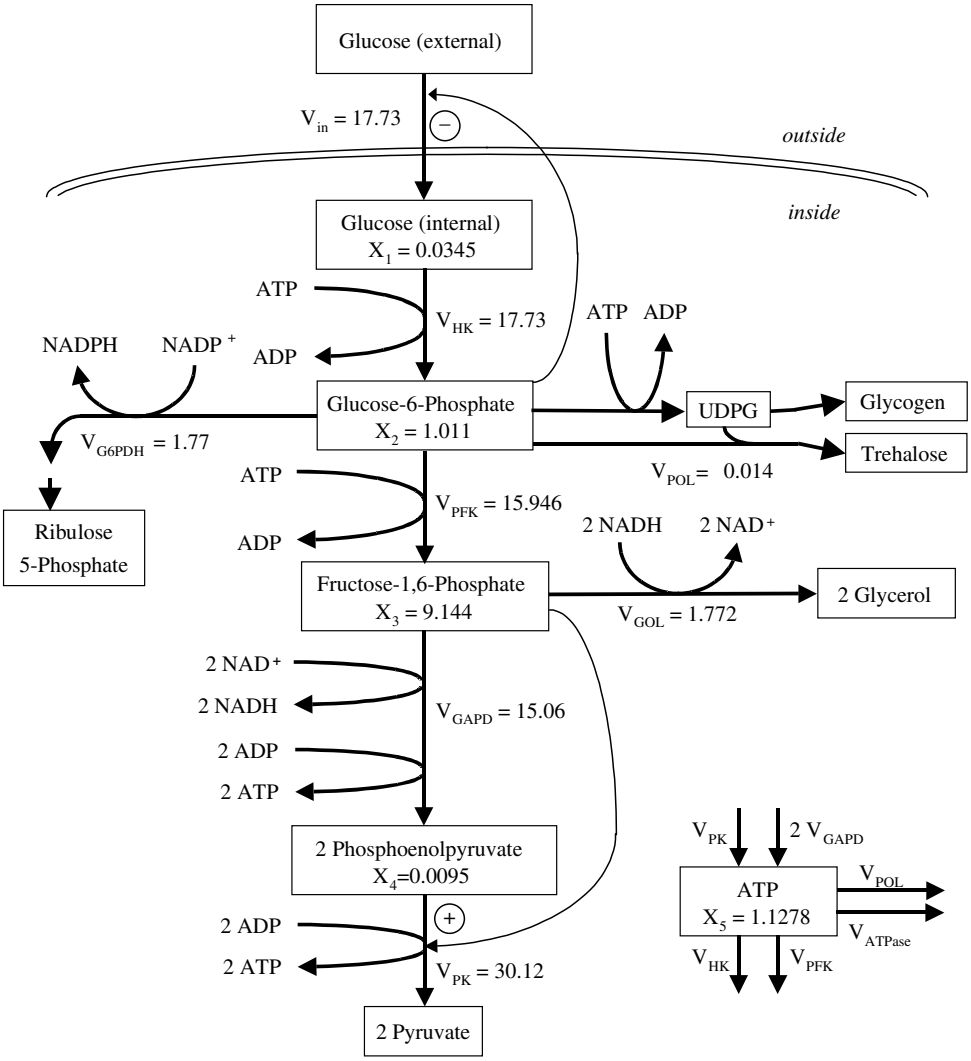

Fig. 1. Fermentation pathway in *Saccharomyces cerevisiae*(adapted from (Galazzo and Bailey, 1990, 1991; Curto *et al.*, 1995). Glucose enters the cell and is immediately phosphorylated to glucose-6 phosphate, G6P. G6P can either proceed toward the pentose shunt to the left, glycogen and trehalose production to the right, or glycolysis straight below. For the glycolytic path, there is an initial investment of two ATP because, when the six-carbon sugar breaks into two three-carbon molecules, each of the products needs to have a negative charge; negative charges produce hydration spheres that help keep small molecules in the cell. Once NAD+ oxidizes the three-carbon aldehyde to form a negatively charged carboxylate, while producing two ATPs, the negative charge provided by the phosphate group is no longer needed, and the initial investment of two ATP is returned to form pyruvate.

The second data set contains baseline mRNA expression levels and transcription rates (Wodicka *et al.*, 1997; Holstege *et al.*, 1998), available from (http://gaiberg.wi.mit.edu/cgi-bin/young_public/lists.cgi?type=H). These expression data (Table 1) are used to estimate overall increases in activities when several enzymes or isozymes are involved in the same step (see below). The reported data are taken at face value; their accuracy is discussed in Alizadeh *et al.* (1998) and Holstege *et al.* (1998).

The observed heat shock profile is not intuitive. Studying the glycolytic pathway in Figure 1, one readily rationalizes increases in hexose transporters (e.g. HXT6, HXT7; V_{in}) and, possibly, in the first step of glycolysis (hexokinase/glucokinase; HXK/GLK; V_{HK}). By contrast, the following steps, catalyzed by glucose 6-phosphate isomerase (PGI) and phosphofructokinase (PFK; V_{PFK}), are essentially unchanged. Why? Later processes, namely glyceraldehyde-3-phosphate dehydrogenase (TDH1-TDH3; V_{GAPD}), phosphoglycerate kinase (PGK1), and pyruvate kinase (PYK1, PYK2; V_{PK}) are slightly increased in activity. Simple inspection cannot explain this pattern.

Table 1. Increases in the expression of genes associated with glycolytic enzymes, 10 min after heat shock, along with numbers of mRNA copies per cell and turnover rates; from (http://rana.Stanford.EDU/clustering/Figure2.txt; http://gaiberg.wi.mit.edu/cgi-bin/young_public/lists.cgi?type=H)

ORF	Gene	Enzyme name	Fold induction	mRNA copies/cell	Transcription (copies/hour)
YPR026W	ATH1	Vacuolar acid α,α-trehalase	1.99	0.2	0
YGR254W	ENO1	Enolase 1	1.91	17.10	48.1
YHR174W	ENO2	Enolase 2	2.07	61.10	134.7
YKL060C	FBA1	Fructose-bisphosphate aldolase II	1.09	74.10	167.9
YCL040W	GLK1	Glucokinase	18.77	3.70	3.8
YKL152C	GPM1	Phosphoglycerate mutase	1.17	32.30	64.0
YDL021W	GPM2	Phosphoglycerate mutase	3.76	0.10	0
YOL056W	GPM3	Phosphoglycerate mutase	0.51	0.90	0.3
YFR053C	HXK1	Hexokinase I	14.03	3.70	5.8
YGL253W	HXK2	Hexokinase II	1.91	19.30	71.9
YHR094C	HXT1*	Hexose transporter (low affinity)	0.79	11.2	41
YMR011W	HXT2	Hexose transporter (high affinity)	7.11	7.0	12.4
YDR345C	HXT3	Hexose transporter (low affinity)	1.85	12.9	37.5
YHR092C	HXT4	Hexose transporter (moderate to low affinity)	3.18	3.1	10.6
YDR343C	HXT6	Hexose transporter (high affinity)	21.56	11.10	18.9
YDR342C	HXT7	Hexose transporter (high affinity)	19.43	9.90	15.1
YJL214W	HXT8	Hexose transporter (putative)	1.27	0.2	0.8
YIL170W	HXT12	Hexose transporter (putative)	1.37	0.4	0.8
YDR001C	NTH1	Neutral α,α-trehalase	1.75	1.5	1.4
YBR001C	NTH2	Neutral α,α-trehalase	4.41	0.4	1.3
YGR240C	PFK1	Phosphofructokinase alpha subunit	0.80	8.80	23
YMR205C	PFK2	Phosphofructokinase beta subunit	1.17	6.80	20
YBR196C	PGI1	Glucose 6-phosphate isomerase	1.21	15.40	30.7
YCR012W	PGK1	Phosphoglycerate kinase	2.23	60.40	110.5
YKL127W	PGM1	Phosphoglucomutase, minor isoform	1.10	1.80	4.9
YMR105C	PGM2	Phosphoglucomutase, major isozyme	14.83	0.30	0
YAL038W	PYK1	Pyruvate kinase	1.13	50.40	101.3
YOR347C	PYK2	Pyruvate kinase	3.81	0.30	1
YJL052W	TDH1	Glyceraldehyde 3-phosphate dehydrogenase 1	2.20	3.60	12.7
YJR009C	TDH2	Glyceraldehyde 3-phosphate dehydrogenase 2	1.75	34.10	98.9
YGR192C	TDH3	Glyceraldehyde 3-phosphate dehydrogenase 2	1.85	89.00	188.1
YBR126C	TPS1	α,α-Trehalose 6-Phosphate Synthase	7.06	2.8	2.8
YDR074W	TPS2	α,α-Trehalose 6-Phosphate Phosphatase	4.69	0.7	0.8
YMR261C	TPS3	α,α-Trehalose 6-Phosphate Synthase	3.03	1	1.9
YNL241C	ZWF1	Glucose 6-phosphate dehydrogenase	3.76	1.60	3.8

*Hexose transporters listed only when more than 0 mRNAs per cell were reported.

Biochemical Systems Theory

Recognizing that rational functions are mathematically inconvenient for the analysis of large biochemical systems, Savageau (1969a,b, 1970, 1972, 1976) proposed approximating kinetic rate laws with multivariate power-law functions. This representation results from Taylor's theorem in logarithmic coordinates. It offers striking procedural, mathematical, and computational benefits, and is the basis of a conceptual framework known as *Biochemical Systems Theory* (BST). BST has proven very useful in a variety of applications (for reviews see Savageau, 1976, 1992, 1995, 1996; Voit, 1991, 2000). A particularly effective variant within BST is the formulation of complex metabolic pathways as *S-systems*.

An S-system representing a pathway with n dependent variables (typically metabolites) and m independent variables (such as enzymes or modulators) consists of n nonlinear ordinary differential equations. The change in each dependent variable is composed of two terms, one capturing all influences that enlarge this variable, and the other one capturing all influences that diminish the variable. Each term contains some or all variables as power functions. A generic S-system equation with *rate constants* α_i and β_i and *kinetic orders* g_{ij} and h_{ij} reads

$$\dot{X}_i = \alpha_i X_1^{g_{i1}} \cdots X_{n+m}^{g_{i,n+m}} - \beta_i X_1^{h_{i1}} \cdots X_{n+m}^{h_{i,n+m}}. \quad (1)$$

BST has matured methodologically to a point where moderately large systems can be analyzed with great efficiency. S-system models of the tricarboxylic acid cycle in *Dictyostelium* (Shiraishi and Savageau, 1992a,b,c,d, 1993, 1996), and of citric acid (Torres, 1994), red blood

cell (Ni and Savageau, 1996a,b), and purine metabolism (Curto *et al.*, 1997, 1998a,b) consist of several dozen variables. Customized software for S-system analysis (PLAS©: http://correio.cc.fc.ul.pt/∼aenf/plas.html; see also Voit, 2000) can handle arbitrarily large systems and is limited only by the capacity of the accessible computer memory. We employ S-systems here to interpret the genomic heat shock data described above.

An S-system model of glycolysis

Galazzo and Bailey (1990, 1991) formulated a model of glycolysis in yeast under different experimental conditions. They used Michaelis–Menten rate laws with standard generalizations for modulated reactions and implemented the numerical model with kinetic information from the literature and with measurements from their own lab. Curto *et al.* (1995) reformulated this model in the structure of S-system equations. This reformulation allowed them to analyze steady-state and dynamic features of the model in a streamlined fashion (see Cascante *et al.*, 1995; Sorribas *et al.*, 1995). In particular, S-systems provided the authors with an objective assessment of the model's sensitivity and robustness, and the detection and diagnosis of problems with the model under various experimental conditions. A robust formulation of the model was later used by others for purposes of yield optimization (Torres *et al.*, 1997; Rodríguez-Acosta *et al.*, 1999)

The model structure and its parameter values are adopted here directly from scenario 1 in Curto *et al.* (1995). The only amendment is the inclusion of the glucose 6-phosphate dehydrogenase (G6PDH) branch. G6PDH oxidizes G6P to 6-phosphogluconate, which is converted in two further steps into ribulose-5-phosphate, thus initiating the pentose shunt. Concomitant with the oxidation on G6P, G6PDH reduces $NADP^+$ to NADPH (for a review, see Martini and Ursini, 1996). The reason to include the G6PDH branch here is that the reduced coezyme NADPH is important in the defense against oxidative stress, where $NADP^+$ is formed in relatively large amounts (e.g. Jollow and McMillan, 1998). Heat shock presumably causes similar stress (Davidson *et al.*, 1996; Godon *et al.*, 1998; Boy-Marcotte *et al.*, 1999). NADPH is furthermore needed in sphingolipid metabolism, which is important in heat shock response (Skrzypek *et al.*, 1999). Because the flux through the G6PDH branch is relatively small, it has almost no effect on the glycolytic flux. It is assumed that $NADP^+$ is available in sufficient quantities; its generation is not modeled.

The flux of trehalose production is even smaller in comparison to the glycolytic flux (Curto *et al.*, 1995), but it seems to be important under heat shock conditions (e.g. Wiemken, 1990; Jenkins *et al.*, 1997). Trehalose is typically present in very low concentrations, but rapidly accumulates under stress conditions, including heat shock (e.g. Neves and François, 1992; Dickson *et al.*, 1997). In one study, where the temperature was shifted to 45 °C, the yeast cells increased their trehalose content from 0.01 to 1 g for each gram of protein (Hottiger *et al.*, 1987). Originally, trehalose was thought to serve as a storage for carbohydrates, but it is now suspected to play the role of a stress protectant (e.g. Wiemken, 1990; Winderickx *et al.*, 1996) that enhances the thermal stability of proteins and thereby improves the cell's resistance to high temperatures (Hottiger *et al.*, 1994). The large quantity of trehalose may also provide energy for mitochondrial biogenesis (Panek and Mattoon, 1977). Genes for both, trehalose production (TPS 1-3) and degradation (NTH 1-2, ATH 1), are overexpressed under heat shock conditions (see Table 1). Trehalose 6-phosphate, the intermediate in the synthesis of trehalose from UDPG, seems to have a regulatory role in heat shock response. It was suggested that it shifts the balance in activity between the two hexokinases and glucokinase (Blázquez *et al.*, 1993; Boy-Marcotte *et al.*, 1999).

Curto *et al.* (1995) represented the effect of G6P on the production of trehalose and glycogen with kinetic orders between 0.37 and 8.6, depending on conditions. This large range was due to the strong sigmoidal dependence of the carbohydrate flux on the concentration of G6P (see Galazzo and Bailey, 1990). For the relatively high concentrations considered here, kinetic orders in the range of 0.8–1 appear to be most reasonable. The particular value turned out to be rather irrelevant for the glycolytic flux itself, but it noticeably affected the flux of carbohydrate production. We will revisit this point in the context of trehalose production under heat shock conditions and in the **Discussion**.

The adjustment of the kinetic order and the inclusion of the G6DPH branch followed without exception the rules of BST (e.g. Savageau, 1976; Voit, 2000). The resulting baseline model (before heat shock) reads

$$\begin{aligned}
\dot{X}_1 &= 0.9018 X_2^{-0.2344} X_6 - 3.1833 X_1^{0.7464} X_5^{0.0243} X_7 \\
\dot{X}_2 &= 3.1833 X_1^{0.7464} X_5^{0.0243} X_7 - 0.8187 X_2^{0.664} X_5^{-0.354} \\
&\quad \times X_8^{0.899} X_{11}^{0.0008} X_{15}^{0.0963} \\
\dot{X}_3 &= 0.5234 X_2^{0.7318} X_5^{-0.3941} X_8 - 0.0148 X_3^{0.584} X_4^{0.03} \\
&\quad \times X_5^{0.119} X_9^{0.944} X_{12}^{0.056} X_{14}^{-0.575} \qquad (2) \\
\dot{X}_4 &= 0.022 X_3^{0.6159} X_5^{0.1308} X_9 X_{14}^{-0.6088} - 0.0945 X_3^{0.05} \\
&\quad \times X_4^{0.533} X_5^{-0.0822} X_{10} \\
\dot{X}_5 &= 0.0880 X_3^{0.3505} X_4^{0.2665} X_5^{0.0243} X_9^{0.5} X_{10}^{0.5} X_{14}^{-0.3044} \\
&\quad - 3.3270 X_1^{0.2134} X_2^{0.190} X_5^{0.362} X_7^{0.286} X_8^{0.257} \\
&\quad \times X_{11}^{0.0002} X_{13}^{0.457}.
\end{aligned}$$

The model contains *dependent* variables (X_1 = internal glucose, X_2 = G6P, X_3 = FDP, X_4 = PEP, X_5 = ATP) and *independent* variables (X_6 = glucose transport, 19.7 mM min^{-1}; X_7 = HXK/GLK, 68.5 mM min^{-1}; X_8 = PFK, 31.7 mM min^{-1}; X_9 = TDH (GAPD in (Curto *et al.*, 1995)), 49.9 mM min^{-1}; X_{10} = PK, 3440 mM min^{-1}; X_{11} = trehalose/glycogen production, 14.31 mM min^{-1}; X_{12} = glycerol production, 203 mM min^{-1}; X_{13} = ATPase, 25.1 mM min^{-1}; X_{14} = NADH/NAD$^+$ ratio, 0.042; X_{15} = G6PDH). The dependent variables are defined by the differential equations above, whereas the independent variables remain constant during each analysis and, therefore, do not require representation by differential equations. In addition to the equations in Equation (2), we consider the fluxes of ATP, trehalose and NADPH at steady state, which provide a measure of heat-shock relevant productivity.

Heat shock leads to an increased demand in ATP, which is met by increased glycolysis (Nickells and Browder, 1988). Under suitable conditions, a shift from about 30 to 40C leads to a 2- to 8-fold increase in ATP concentration (Hottiger *et al.*, 1987; Winkler *et al.*, 1991; Neves and François, 1992). However, without external carbon sources, ATP concentrations may rather rapidly drop during heat shock (e.g. Nickells and Browder, 1988; Browder *et al.*, 1989). Bonini *et al.* (1995) observed a drop in ATP within 15 min after heat shock, if no external carbon sources were available, along with decreases in trehalose, glycogen, and glycolytic metabolites. *Tetrahymena* lost 50% of cellular ATP within the first 2 min of heat shock (Findly *et al.*, 1983). Given the relatively short half-life of ATP, which is on the order of seconds (see Hottiger *et al.*, 1987), the rapid decrease may not be surprising. Because the ATP concentration during heat shock apparently differs widely among organisms, we focus instead on ATP production.

The flux of ATP production (which equals ATP usage at steady state) is given as the first product of power-law functions on the right-hand side of the fifth equation in (2):

$$V_{\mathrm{ATP}} = 0.0880 X_3^{0.3505} X_4^{0.2665} X_5^{0.0243} X_9^{0.5} X_{10}^{0.5} X_{14}^{-0.3044}. \tag{3}$$

At the baseline steady state, its value is 60.17 mM min^{-1}.

The trehalose flux is derived from the models of Galazzo and Bailey (1990, 1991) and Curto *et al.* (1995). According to these authors, this flux constitutes 10% of the glycogen production flux. The total carbohydrate flux (glycogen plus trehalose) is formulated in Curto's notation as $\beta_{2,2} X_2^{h_{22,2}} X_{11}$ and has a baseline value of 0.014. The kinetic order associated with G6P($= X_2$) is $h_{22,2}$. If the parameter $h_{22,2}$ equals 1 (see above), and if 1 part of 11 is used for trehalose production, the flux reads

$$V_{\mathrm{Tre}} = 0.000\,98 X_2 X_{11}. \tag{4}$$

If we instead set $h_{22,2} = 0.8$, the flux is

$$V_{\mathrm{Tre}} = 0.000\,8925 X_2^{0.8} X_{11}. \tag{5}$$

Both fluxes have the same value of 0.001 286 at the baseline steady state, which corresponds to 1/11 of G6P degradation. Even without exact information about the kinetic order $h_{22,2}$, the model allows us to evaluate its effect. Setting $h_{22,2} = 0$ corresponds to a process saturated with respect to G6P, and setting $h_{22,2} = 1$ simulates direct proportionality. The actual value is likely to be bounded by these two settings.

About 10% of G6P is channeled toward the pentose shunt, while 90% is used for glycolysis and other pathways (Stanbury *et al.*, 1983, p. 1632). The K_M of G6PDH with respect to G6P for a variety of microorganisms is about 50 μM (McKerns, 1975; Olive and Levy, 1975; Scott, 1975; Shepherd, 1975; Stanbury *et al.*, 1983). The kinetic order of the G6PDH flux with respect to G6P at steady state ([G6P] = 1.011 mM) is thus approximately $K_M/K_M + [\mathrm{G6P}]) = (50/1,061) \approx 0.05$ (for this type of computation, see Voit *et al.*, 1991, p. 55). The flux of NADPH production is therefore modeled as

$$V_{\mathrm{NADPH}} = 1.769 X_2^{0.05} X_{15} \tag{6}$$

and has the steady-state value 1.77.

Methods of model analysis

Types of analyses

We performed three types of analyses. The first consisted of confirming the steady-state concentrations and fluxes published by Galazzo and Bailey (1990, 1991) and Curto *et al.* (1995). The model results are shown in Figure 1 and are indeed the same as those published.

Second, to simulate the observed heat shock response, the enzyme activities in the S-system model were specified in accordance with the observed gene expression profile (Table 1). We call the S-system model with these numerical settings the *heat-shock model* or, simply, the HS *model*.

Third, to explain why particular genes are overexpressed, while others are not, the HS model was compared with *hypothetical* model implementations in which some of the catalytic steps were altered in some plausible fashion. For instance, representing rate limitation at the first step, we analyzed a scenario in which only the first step was activated, whereas all other activities were kept constant.

Alterations in enzyme activities were implemented in the following fashion.

1. If a single enzyme catalyzes a reaction, the observed change in gene expression, obtained from the microarray data (http://rana.Stanford.EDU/clustering/

Figure2.txt; Eisen *et al.*, 1998), was interpreted as a corresponding change in enzyme activity. The time delay between gene expression and change in enzyme activity was assumed to be the same for all enzymes considered. The one-to-one correspondence between gene expression and enzyme activity, and the common time delays, may not be entirely accurate, but seemed to provide the best default estimate (however, see point 3 below regarding secondary changes in enzyme activities).

2. Several glycolytic steps are catalyzed by two or more enzymes or isozymes. To assess their relative importance, we weighted their altered activities by the corresponding numbers of mRNA copies per cell, which we obtained from http://gaiberg.wi.mit.edu/cgi-bin/young_public/lists.cgi?type=H. As an example, consider the conversion of phosphoenolpyruvate to pyruvate, which is catalyzed by pyruvate kinases 1 and 2. Their increases in gene expression are 1.13 and 3.81, respectively (cf. Table 1). The corresponding numbers of mRNA copies per cell are 50.40 and 0.30. Using these numbers as weights, one obtains as overall increase in expression $(50.40 \times 1.13 + 0.30 \times 3.81)/(50.40 + 0.30) = 1.15$.

   As a second example, the conversion of G6P and UDPG to trehalose 6-phosphate is catalyzed by trehalose 6-phosphate synthase. The two corresponding enzymes, TPS1 and TPS3, are over-expressed 7.06 and 1.6 times, respectively. 2.8 and 1 mRNA copies were measured per cell, respectively (see Table 1). The overall increase in activity, weighted by mRNA copies, is 6-fold. Trehalose production involves a subsequent step, catalyzed by trehalose 6-phosphate phosphatase (TPS2). TPS2 is over-expressed 4.69-fold, which is in the same range as the over-expression of the trehalose 6-phosphate synthase step. In general, an assessment of the overall change in activity in a sequence consisting of several steps with widely differing alterations in activation is not a trivial matter and would require more detailed analysis.

   It must be emphasized that different types of weighting, inaccuracies in measurements, and decisions about the inclusion or exclusion of processes create uncertainties that need to be considered during the interpretation of results. For instance, there are over a dozen hexose transporters, which potentially contribute to the import of glucose into the cell. According to actual measurements (http://gaiberg.wi.mit.edu/cgi-bin/young_public/lists.cgi?type=H), several of these show no mRNA copies per cell and no transcription rates, and our weighting method effectively excludes them from the computation of the overall enzyme activity. Furthermore, some transporters have low and some high affinity. Depending on which transporters are included, the transport activity is increased 7-fold (all transporters considered) or 16-fold (low-affinity transporters excluded). For our analysis, we assume a 10-fold increase as a starting point and analyze the effects of higher or lower values through model simulation. Some of the uncertainties are unavoidable in any method of analysis, and it is a particular strength of the modeling approach that it allows us to test the effects of using higher or lower activities.

3. The response to heat shock is generally regulated at the level of transcription (Lindquist, 1986). In *E.coli*, transcription is activated in less than 1 min, and although a few mRNAs accumulate, most of the heat shock mRNAs begin to decline within 5 min (Lindquist, 1986). In addition to transcriptional regulation, the literature provides some information about increases in enzyme activities under heat shock conditions. Directly pertaining to glycolysis, even though not always investigated in yeast, it appears that phosphofructokinase is somewhat inhibited (Bonini *et al.*, 1995), whereas glyceraldehyde 3-phosphate-dehydrogenase (GAPD, TDH 1) and pyruvate kinase (PYK) are more active after shifts to higher temperatures (Nickells and Browder, 1988; Browder *et al.*, 1989; Marsden *et al.*, 1993; Boucherie *et al.*, 1995). In *Xenopus* embryos, for instance, the specific activity of GAPDH was found to increase up to over 200% (Nickells and Browder, 1988). This was accompanied by a significant accumulation of a GAPDH isozyme with a narrow temperature range of optimal activity that was centered about the temperature during heat shock. The authors suspected heat-induced *de novo* synthesis, but considered increased stability of the enzyme as an alternative. Browder *et al.* (1989) suspected GAPD to be a heat shock protein (HSP35). Entian *et al.* (1984) did not observe changes in GAPDH.

   Marsden *et al.* (1993) found increases in the synthesis of an isoform of pyruvate kinase in *Xenopus* embryos and demonstrated that it enhanced thermal stability, thus classifying pyruvate kinase as a heat shock protein. Rosenzweig (1992) found no change in heat shock sensitivity of yeast over-producing pyruvate kinase.

   The synthesis of pyruvate kinase is markedly induced by increased levels of glucose (Entian *et al.*, 1984), while the key gluconeogenic enzymes are

inactivated. This might be of relevance, given the strong induction of hexose transporters following the shift in temperature.

It seems to be generally accepted that trehalose metabolism is affected by heat shock. Winkler *et al.* (1991) argue that the observed increases in trehalose concentration are solely due to an increase in its precursors, G6P and UPDG. However, many other authors have presented compelling results and explanations for why the induction of trehalose metabolism is a good strategy for responding to heat shock (e.g. Wiemken, 1990; Neves and François, 1992; Boucherie *et al.*, 1995; Winderickx *et al.*, 1996). The modeling approach allows us to explore these observations by increasing the activities of these enzymes to various degrees (see also **Discussion**).

Implementation of the heat shock model

The structure of the earlier model of Curto *et al.* (1995) was used with the slight alterations mentioned above. The implementation of the observed heat shock scenario consisted of resetting all enzyme activities in accordance with the observed profile.

According to the expression profile, the phosphofructokinase step (PFK1-2) is essentially unchanged in activity. In *Neurospsora crassa*, phosphofructokinase is inhibited, while fructose 1,6-bisphosphatase is active, which seems to indicate a channeling of glucose or glycerol toward trehalose (Bonini *et al.*, 1995). Indeed, if PFK is inhibited in the HS model, G6P increases, FDP decreases, and the trehalose flux is elevated. The increase in G6P, by means of its inhibitory effect on hexose transport, leads to a decrease in free glucose (see discussion of scenarios in the **Results** section).

The conversion of FDP into PEP was represented in previous models (Galazzo and Bailey, 1990; Curto *et al.*, 1995) by glyceraldehyde 3-phosphate dehydrogenase (V_{GAPD}). The corresponding three genes, TDH1-3, are over-expressed by a factor of about 1.8. It also appears that these dehydrogenases are somewhat induced at increased temperatures (Nickells and Browder, 1988; Boucherie *et al.*, 1995). Simulations with different induction factors yielded similar results in almost all aspects, except for the concentrations of F6P and PEP. A 1.5-fold induction seemed reasonable.

Several other enzymes are involved in the conversion of F6P into PEP. Like TDH1-3, the genes coding for these enzymes (FBA1, aldolase; PGK1, phosphoglycerate kinase; GPM1-3, phosphoglycerate mutase; and ENO1-2, enolase) are over-expressed by weighted factors between 1 and 2 (see Table 1). Corroborating these measurements, Entian *et al.* (1984) found a marked increase in enolase II, and Iida and Yahara (1985) suspected enolase to be the isoform of a heat shock protein. The condensed model of Curto *et al.* (1995) does not allow us to use this detail information.

The pyruvate kinase genes, PYK1-2, are over-expressed 1.13- and 3.81-fold, which gives a weighted average of about 1.2. Pyruvate kinase is induced under increased temperature (Browder *et al.*, 1989; Marsden *et al.*, 1993) and activated by glucose (Entian *et al.*, 1984). Simulations with several induction factors yielded similar results in all components, except for a noticeable effect on the concentration of PEP and a very slight effect on ATP production. An induction factor of 5 yielded a PEP concentration close to the baseline value that existed before heat shock. This factor was used as default.

Assessment of results

The primary criterion for evaluating different enzyme profiles is the sufficient production of ATP, trehalose, and NADPH. Secondary criteria are the unneeded accumulation of intermediates, which would strain the cell's solubility capacity, and the cost of over-expression.

The criteria are of a different nature, and any simple comparison or scoring of the quality of a hypothetical scenario ('Hypo') requires the merging of incomparable aspects into some metric. Nonetheless, a metric is desirable for assessing the overall performance of a given scenario.

To develop such a metric, it appears reasonable to use ratios of fluxes, concentrations, and costs of over-expression, rather than some absolute values. Furthermore, in order not to assign too much weight to a single high concentration or flux, we decided to use logarithms of ratios. The cost associated with over-expressing a gene and increasing the activity of an enzyme is unknown, but one may assume that the cost for a small degree of over-expression is relatively higher than for further increases. In order to balance costs with benefits in flux production, we therefore decided to use the logarithmic function also for cost ratios. Specifically, the following assumptions underlie the scoring metric.

1. While the baseline profile may be optimal under normal conditions, the observed heat shock (HS) profile is considered to be optimal under heat shock conditions. The heat shock profile is therefore used for comparison.
2. The production of ATP in a hypothetical scenario (Hypo) enters the scoring function as

$$\ln\left[\frac{\mathrm{ATP}\quad \mathrm{flux}_{\mathrm{Hypo}}}{\mathrm{ATP}\quad \mathrm{flux}_{\mathrm{HS}}}\right].$$

Trehalose, and NADPH production fluxes enter the scoring function in the same fashion.

Table 2. Metabolite levels and fluxes of the HS model in comparison to model results before heat shock

Metabolite or flux	Steady-state value of HS model	Relation to value before heat shock
Internal glucose	0.031 mM	89%
G6P	23.18 mM	23.0-fold
FDP	15.98 mM	1.8-fold
PEP	0.096 mM	102%
ATP	6.89 mM	6.1-fold
Glucose input flux	85.0 mM min^{-1}	4.8-fold
ATP production	325.9 mM min^{-1}	5.4-fold
Trehalose production	0.071 mM min^{-1}	55.1-fold
NADPH production	7.87 mM min^{-1}	4.4-fold

3. The deviations of the intermediary metabolites (glucose, G6P, F6P and PEP) from the heat shock profile enter the scoring function as

$$-\text{abs}\left(\ln\left[\frac{\text{Concentration}_{\text{Hypo}}}{\text{Concentration}_{\text{HS}}}\right]\right).$$

4. The sum of the flux and concentration terms constitute a *performance score* that has a value of 0 for the heat shock profile and negative values for all scenarios that produce less ATP, trehalose, and/or NADPH, and/or significantly deviate in metabolite concentrations.
5. The cost of over-expressing a gene or inducing an enzyme enters the scoring function as

$$-\ln\left[\frac{\text{Level of Expression}_{\text{Hypo}}}{\text{Level of expression}_{\text{Baseline}}}\right].$$

   It seemed more reasonable to compare the expression to the normal, baseline conditions rather than to heat shock conditions. This implies that the heat shock profile has a cost, while the baseline scenario does not.
6. The overall score is computed as the sum of the above components.

Results

Observed expression profile upon heat shock and settings for the HS model

Overall, the responses of the HS model appear to be very reasonable. The primary goals are met. ATP is produced at over 5 times the original rate, trehalose is produced at a 50-fold rate, and the production of NADPH is increased almost 5-fold (Table 2).

It is known that ATP demand is increased under heat shock conditions, but no directly applicable, quantitative data are available for comparison. If ATP usage (modeled as the degradation term in the fifth equation of the S-system (Equation 2)) is unchanged, the steady-state ATP concentration increases about 6-fold. This seems to be in the right range (Grba *et al.*, 1979; Hottiger *et al.*, 1987; Winkler *et al.*, 1991; Neves and François, 1992).

The trehalose flux rises 50 times above the original magnitude. Again, no quantitative information is available for direct comparisons. Although not modeled here, the main pathway of trehalose degradation is catalyzed by trehalase, whose genes (NTH1-2, ATH1) are over-expressed about 2.3 times (see Table 1). This is of the right magnitude, because Hottiger *et al.* (1987) measured a 3-fold increase in neutral trehalase activity, along with an increase in trehalose 6-phosphate synthase and a very rapid trehalose turnover. Specifically, trehalose 6-phosphate synthase activity rose in 20 min to 3–4 times the baseline level, and it rose further to 6 times the baseline level in 40 min. Taking account of the time delay between gene expression and altered enzyme activity, these observations are consistent with the array data used here. Accounting for increases in both production and degradation, the steady-state concentration of trehalose is increased 24-fold over baseline, which is consistent with observations (Winkler *et al.*, 1991; Hottiger *et al.*, 1987; Bonini *et al.*, 1995). Contrary to these results and those of others, Blázquez *et al.* (1993) found no increase in trehalose 6-phosphate concentration under heat shock conditions. The model can explain this observation if the excess trehalose 6-phosphate is converted into trehalose or if the trehalase activity is more strongly induced, which might be the case for the authors' conditions.

Neves and François (1992) found significant increases in free glucose and trehalose, which were accompanied by 3-fold decreases in hexose 6-phosphates. Our model also shows an increase in trehalose and, depending on the degree of hexokinase over-expression, can explain the increase in glucose. However, none of our simulations showed a decrease in G6P (however, see **Discussion**).

The secondary goals of the heat shock implementation are also met. With the exception of G6P, the intermediates have concentrations within the normal range (Table 2). G6P is increased 23-fold, which may seem high. However, Winkler *et al.* (1991) found the concentrations of both G6P and UDPG to be increased 5–10-fold. Galazzo and Bailey (1990) and Curto *et al.* (1995) did not explicitly model UDPG, but if a variable for UDPG were included, one would expect it to rise, at the cost of G6P. An explanation for the elevation in G6P may be that this metabolite is the starting point of carbohydrate metabolism; many important pathways stem from this ester, rather than from glucose itself (White *et al.*, 1968, p. 391). The additional supply of 'the essential biosynthetic precursor' of carbohydrate metabolism (DeRisi *et al.*, 1997), G6P,

might therefore be advantageous (see **Discussion** for further comments).

The HS model is robust. In particular, most logarithmic gains, representing the effect of independent variables on the steady state, are smaller than 1, and none exceeds 2. This indicates that perturbations in independent variables are usually attenuated or, at most, mildly amplified. This (local) result is confirmed by simulations, in which we varied uncertain enzyme activities. For instance, the weighted increase in activity for hexose transporters was set to 10 in the HS model, but 6–16-fold increases yielded similar results (details not shown).

The glucokinase/hexokinase activity was increased in the HS model by a factor of 5. This factor essentially yields the baseline level in glucose concentration. A 4-fold increase leads to a steady-state glucose level that is about 20% elevated. Some authors found indeed an accumulation of glucose during heat shock (Hottiger *et al.*, 1987; Neves and François, 1992) Except for the free glucose pool, the steady state is essentially unaffected by different rates of increase. Blázquez *et al.* (1993) executed a series of experiments showing that trehalose 6-phosphate inhibits hexokinase II, and to a lesser degree hexokinase I. If such an inhibition is modeled, the free glucose pool increases correspondingly, but all other metabolites, including G6P and trehalose, are essentially unaffected.

Although not of primary interest, dynamic responses of the HS model are rapid. For instance, an external bolus of glucose, modeled as a 10-fold increase in internal glucose, is absorbed within seconds, with all metabolites returned to within 1% of their steady-state concentrations. Even after a drastic step-switch from baseline to heat shock conditions, the system converges to the new steady state in less than 2 min. This short response time of the model indicates that the system is able to react to environmental changes in due time. A more detailed investigation of the dynamic aspects of the HS model will be discussed elsewhere.

The cost of over-expression associated with the observed HS profile is 9.64. Its performance score is 0, by definition. By comparison, the baseline scenario has no over-expression cost, but shows inferior performance (-11.0) under heat shock conditions.

Exploration of alternate expression strategies

The observed expression profile (Table 1) is by no means intuitive. The transport of hexoses into the cell is increased 10-fold, the following kinase step only 5-fold. The subsequent step is not affected at all. Later steps are increased again. What is the rationale for such an arrangement? One can find some answers by studying alternative, hypothetical profiles that, at first glance, seem plausible. A summary of model responses is presented in Table 3.

1. Equal scaling. If all activities are doubled, the intermediate pools have concentrations between 80 and 132% of the baseline. However, all three production fluxes (ATP, trehalose, NADPH) are much too low for an adequate response. The cost is moderate, but the performance is not good. One could argue that this strategy should be accompanied by the alleged heat-induced increases in the activities of GAPD and pyruvate kinase. These additional increases do not improve the production fluxes, yet lead to a significant decrease in PEP.

2. Stronger equal scaling. An across-the-board 5-fold increase is more expensive (cost $= 11.27$), but produces more ATP, trehalose and NADPH (performance score $= -5.68$). Though increased, trehalose and ATP production are still significantly below that in the HS model (by 15 and 85%, respectively). Overall, the performance is rather low, but the cost high. If this strategy of over-expression is accompanied by increases in the activities of GAPD and pyruvate kinase, the performance in ATP and NADPH fluxes improves, but cost increases and the concentration of PEP is much reduced. Trehalose production is still much too low.

3. Increase input (HXT) only. Glycolysis is essentially a linear pathway, and one could argue that hexose transport into the cell might be the limiting step and that an increase in this initial step alone would provide sufficient material for a heat shock response. Model analysis of this simple and inexpensive strategy reveals undesirable consequences. Several metabolites assume high steady-state values. Internal glucose, FDP, and PEP are increased 7-, 10-, and 25-fold, respectively. The ATP flux is increased 5-fold, as desired, but trehalose and NADPH production are significantly lower than needed. The cost of this scenario is low (2.3), but the performance (-9.92) is low too.

4. Increase input (HXT) only and induce GAPD and PYK. The previous strategy may be accompanied by the alleged heat-induced increases in the activities of GAPD and pyruvate kinase. Indeed, the metabolite concentrations respond more favorably, but the production of trehalose and NADPH is still insufficient. The improvement in performance requires increased cost.

5. Increase input (HXT) only, induce GAPD and PYK, and increase TPS and G6PDH. If the enzymes responsible for trehalose and NADPH production are activated, in addition to increased input and heat induction of GAPD and PYK, the result is much improved. Still, glucose and FDP are unnecessarily elevated.

6. Increase the transport and HXK/GLK steps. One could surmise that the first 'internal' step is important for regulation and that increased activity at this step should increase the overall flow. Implementation shows that FDP

Table 3. Summary of responses to heat shock, given different observed and hypothesized enzyme activities

Scenario	HXT	GLK	PFK	TDH (X_9)		PYK (X_{10})		TPS	G6PDH	Responses	Perf.	Cost	Score
	X_6	X_7	X_8	Expr.	Indct.	Expr.	Indct.	X_{11}	X_{15}				
Baseline	1	1	1	1	1	1	1	1	1	VA, VN, VT↓	−11.0	0	−11.0
HS	10	5	1	2	1.5	1.2	5	4.5	3.8	✔	0	9.64	−9.64
1	2	2	2	2	1	2	1	2	2	VA, VN, VT↓	−8.58	4.85	−13.43
2	5	5	5	5	1.5	5	5	5	5	PEP, VT↓	−8.64	13.28	−21.92
3	10	1	1	1	1	1	1	1	1	Glu ↑; FDP, PEP↑; VN, VT↓	−9.92	2.30	−12.22
4	10	1	1	1	1.5	1	5	1	1	Glu, FDP ↑; VN, VT↓	−6.46	4.32	−10.78
5	10	1	1	1	1.5	1	5	4.5	3.8	Glu, FDP↑; PEP↑	−7.03	5.14	−12.17
6	10	5	1	1	1	1	1	1	1	FDP, PEP↑; VN, VT↓	−7.86	3.91	−11.78
7	10	5	1	1	1	1	1	4.5	3.8	FDP ↑; PEP↑	−4.88	6.75	−11.63
8	10	5	4	1	1	1	1	4.5	3.8	FDP, PEP↑;VT↓	−8.30	8.14	−16.44
9	10	5	0.25	1	1	1	1	4.5	3.8	FDP↑; PEP↑; VA ↓	−4.51	5.36	−9.88
10	10	5	1	2	1.5	1	1	4.5	3.8	PEP↑	−3.53	7.85	−11.38
11	10	5	1	1	1	1.2	5	4.5	3.8	FDP↑	−2.01	8.54	−10.55
12	10	5	1	5	1.5	4	5	4.5	3.8	PEP ↓	−3.40	11.76	−15.16

Responses not specified are close to those observed under heat shock conditions. Symbols: ✔ : Overall good response; ↑: very strong increase (more than 10-fold); ↑: strong increase (between 5 and 10-fold); ↓: very strong decrease (to less than 10%); ↓: strong decrease (to between 10 and 20%). Production of ATP (VA), trehalose (VT) and NADPH (VN) given in comparison to fluxes under the observed heat shock conditions. 'Perf.' denotes the performance score (see Text).

is 10 times increased and PEP 25 times. ATP production is adequate, but production of trehalose and NADPH is not.

7. Additionally increase trehalose and NAPDH production. One might expect that FDP and PEP accumulate because the flux toward trehalose and the pentose pathway has not been increased. Increasing trehalose and NADPH production yields sufficient quantities of these metabolites, but FDP and PEP still have high concentrations.

8. Increase HXT, HXK/GLK, G6PDH, TPS, and PFK. PFK activity is unchanged in the observed profile. Nonetheless, the model allows us to explore what an alteration would accomplish. Raising the activity of the enzyme 4-fold, the system still has severe problems. For instance, FDP and PEP rise to 15-fold and 39-fold, respectively. Furthermore, the G6P concentration is much lower than in the HS model, and this results in lower trehalose production.

9. Increase HXT, HXK/GLK, G6PDH, TPS, but decrease PFK. Because increases in PFK activity have negative consequences, one might want to explore repression. If the activity is lowered to one quarter, the FDP concentration is more reasonable, but PEP is still elevated more than 10-fold. G6P rises to 74-times the original concentration, which leads to increased trehalose and NADPH production. By contrast, ATP production is much lower than in the HS model.

10. Increase HXT, HXK/GLK, G6PDH, TPS, and TDH. This strategy, with a 3-fold increase in TDH, leads to a much-improved solution that is close to the observed scenario, but PEP still exceeds the baseline value by a factor of 25. Stronger increases in TDH lead to a further rise in PEP and decreased ATP production. A weaker increase resembles *Scenario 7*.

11. Increase HXT, HXK/GLK, G6PDH, TPS, and PYK, but not TDH. If TDH is not increased, but PYK is elevated as in the HS model, the response is quite reasonable and cost-effective, with the exception of an FDP concentration that is elevated 10-fold. Further increases in PYK lead to similar responses and higher cost. Because there is no other benefit, these increases are wasteful.

12. Increase TDH and PYK beyond heat shock profile. If TDH and PYK are more strongly activated than observed, PEP falls to about 14% of the baseline level. Otherwise, there are no particular benefits or disadvantages, except for increased cost.

Discussion

The analysis presented here demonstrates the potential of an effective systems approach. With a series of what-if scenarios we provided an explanation for a complex gene expression pattern during heat shock. The analysis provided evidence for the importance of each increased enzyme activity, the lack of benefit in increasing enzyme activities that are indeed not increased in the cell, and the lack of benefit in increasing enzyme activities further than observed. Taken together, these results paint a picture of highly effective coordination, not just with respect to primary goals, such as the production of ATP, trehalose, and NADPH, but also in terms of secondary goals, including the control of intermediates and the minimization of cost. The coordination is in some sense optimized since any simple change in the activity pattern is apparently disadvantageous.

An interesting phenomenon is the strong increase in the activity of hexokinase and glucokinase. All sensitivities of the model with respect to changes in this step are very low, except for a strong negative effect on the internal glucose concentration (see also Cascante *et al.*, 1995) Indeed, altering this enzyme activity alone does not have much effect on any other flux or concentration, except for internal glucose. If hexokinase and glucokinase operate at the baseline level (before heat shock), while all other enzyme activities are increased as in the HS model, the internal glucose concentration rises 8-fold, but all other results are essentially unaltered. The only consequence of the observed increase in hexokinase and glucokinase thus seems to be the maintenance of the low baseline glucose concentration.

Definite reasons for maintaining a 'normal' glucose level are not available, but our conclusion is consistent with experimental evidence indicating that intracellular free sugars do not accumulate during fermentation (Romano, 1986, see also Gancedo, 1986, Neves and François, 1992). In fact, glucose is phosphorylated immediately as it enters the cell (Gancedo, 1986; Neves and François, 1992). One could argue that glucose represses the expression of a number of genes, some of which are associated with heat shock response (DeRisi *et al.*, 1997). For instance, glucose affects the transporters HXT1 and HXT2 (http://www.mips.biochem.mpg.de; Özcan and Johnston, 1999), hexokinase 1 (http://gaiberg.wi.mit.edu/cgi-bin/young_public/lists.cgi?type=H.), and pyruvate kinase (Entian *et al.*, 1984). It may also be advantageous to have instant access to an increased level of G6P instead of glucose. G6P is the critical branchpoint at which the cell decides whether to use carbohydrates for energy production, for the synthesis of pentoses with concomitant generation of NADPH, or for increased production of trehalose or glycogen.

Although there may be reasons for maintaining a high concentration of G6P, one could also suspect that the underlying model needs to be revisited. It is not our purpose to discredit the earlier models, but to show how the combined analysis of biochemical and gene expression data can lead to new insights and to a diagnosis of potential problems. Galazzo and Bailey (1990, 1991) developed the model to reflect normal physiology. Under typical conditions, trehalose is only found in traces (e.g. Hottiger *et al.*, 1987), and it was probably reasonable for Galazzo and Bailey to model trehalose production in a simplified form. Under heat shock conditions, the simplifications may no longer be valid.

Galazzo and Bailey (1990) suggested sigmoid dependence of glycogen production on G6P and assumed trehalose production to be proportional to this process. If true, the process is saturated even under moderately elevated G6P conditions (e.g. [G6P] = 4 mM) and the cell cannot further enhance trehalose production by increasing the concentration of G6P. In the current HS implementation, over-expression of the TPS1/TPS3 complex leads to a 50-fold increase in trehalose production. This may seem high, but the absolute magnitude of this flux is still very small in comparison to the glycolytic flux. It is far from being able to control the G6P concentration or to channel appreciable amounts of glucose into trehalose (Hottiger *et al.*, 1987). To accomplish effective channeling, the trehalose flux would have to be much larger.

The high G6P concentration and the inability of the model to channel more glucose toward trehalose instead of FDP suggest a more detailed model representation of trehalose metabolism under heat shock conditions. Such a model might include the dynamics of ceramide and other sphingolipids, which apparently serve as heat stress signals and control the transcription of TPS2 and the consequent accumulation of trehalose (Dickson *et al.*, 1997; Jenkins *et al.*, 1997; Perry and Hannun, 1998; Wells *et al.*, 1998; Skrzypek *et al.*, 1999)

The present pilot study makes a number of simplifying assumptions, whose validity can presently not be tested. First, there are obvious delays between changes in gene expression and corresponding changes in enzyme activity. We assume that these delays are approximately the same for all enzymes. If so, the metabolic events have the same time sequence as the events at the genomic level. Second, we assume that a particular change in gene expression translates directly into a corresponding change in enzyme activity. Thus, translation and post-translational alterations are ignored. Obviously, the enzyme activities are affected

by the dynamics of their mRNAs. For the present analysis, we treated enzyme activities as changing parallel to heat shocked values and subsequently held them constant. This seemed reasonable because of the rapid response time of the model. A detailed dynamical analysis (Voit, in preparation) will account for the dynamical aspects of the heat shock response. It will address how over- or under-expression for each gene rises from the baseline to some extreme values, before converging back toward the baseline. The analysis will also account for the dynamic production and degradation of enzymes. This type of dynamic investigation will allow us to study whether any of the transients between new states might lead to potentially harmful, temporary excesses or shortages in key metabolites. These extensions of the present model will require a larger number of variables and processes, but this should not pose a problem for the S-system modeling framework.

It is a particular strength of the biomathematical approach that the model analysis identifies what kinds of problems would occur under hypothesized alternate expression strategies. Intuition is not sufficient to execute such assessments, and it seems that no experimental approach is available to interpret the observed data in this fashion or to contrast them with alternate, seemingly viable strategies. In the present analysis, data of two types were available, microarray data and biochemical data. Our combined analysis allowed us to test for consistency between these two types of data and to explain one set of data with a model based on the other set. Once the details of combining diverse pieces of information are better understood, predictions in one domain will arise based on observations in the other.

Acknowledgement

We are grateful to Dr Néstor Torres for providing positive criticism.

Acronyms

ADH - alcohol dehydrogenase gene

α_i - S-system parameter: rate constant of production term for variable X_i

ATH - trehalase gene

β_i - S-system parameter: rate constant of degradation term for variable X_i

BST - Biochemical Systems Theory

FBA - aldolase gene

ENO - enolase gene

FDP - fructose 1,6-diphosphate

F6P - fructose 6-phosphate

GAPD - glyceraldehyde 3-phosphate dehydrogenase

g_{ij} - S-system parameter: kinetic order, representing the effect of variable X_j on the production of variable X_i

GLK - glucokinase (gene/enzyme)

GLY-3-P - glyceraldehyde 3-phosphate

GPM - phosphoglycerate mutase gene

G6P - glucose 6-phosphate

G6PDH - glucose 6-phosphate dehydrogenase

h_{ij} - S-system parameter: kinetic order, representing the effect of variable X_j on the degradation of variable X_i

HS model - S-system model implemented with enzyme activities under heat shock conditions

HXK - hexokinase (gene/enzyme)

HXT - hexose transporter

Hypo - hypothetical scenario

$NADP^+$ - nicotinamide-adenine dinucleotide phosphate

NADPH - reduced nicotinamide-adenine dinucleotide phosphate

NTH - neutral trehalase gene

PEP - phosphoenolpyruvate

PFK - phosphofructokinase (gene/enzyme)

PGI - glucose 6-phosphate isomerase (gene/enzyme)

PGK - phosphoglycerate kinase (gene/enzyme)

PLAS© - Power-Law Analysis and Simulation software

PYK - pyruvate kinase (gene/enzyme)

TDH - glyceraldehyde 3-phosphate dehydrogenase gene

TSP1 - trehalose 6-phosphate synthase gene

TSP2 - trehalose 6-phosphate phosphatase gene

UDPG - uridine diphosphate glucose

VA - ATP production flux

V_{GAPD} - glyceraldehyde-3-phosphate dehydrogenase flux

V_{in} - flux of glucose into the cell

V_{HK} - hexokinase/glucokinase flux

VN, V_{NADPH} - NADPH production flux

V_{PFK} - phosphofructokinase flux

V_{PK} - pyruvate kinase flux

VT - trehalose production flux

X_i - S-system variable; typically X_i is a dependent (state) variable if $i \leq n$, or an independent (control)

variable if $n + 1 \leq i \leq n + m$.
ZWF - glucose 6-phosphate dehydrogenase gene

References

Alizadeh,A., Eisen,M., Botstein,D., Brown,P.O. and Staudt,L.M. (1998) Probing lymphocyte biology by genomic-scale gene expression analysis. *J. Clin. Immunol.*, **18**, 373–379.

Bailey,J.E. (1999) Lessons from metabolic engineering for functional genomics and drug discovery. *Nat. Biotech.*, **17**, 616–618.

Blázquez,M.A., Lagunas,R., Gancedo,C. and Gancedo,J.M. (1993) Trehalose-6-phosphate, a new regulator of yeast glycolysis that inhibits hexokinases. *FEBS Lett.*, **329**, 51–54.

Bonini,B.M., Neves,M.J., Jorge,J.A. and Terenzi,H.F. (1995) Effects of temperature shifts on the metabolism of trehalose in *Neurospora crassa* wild type and a trehalase-deficient (tre) mutant. Evidence against the participation of periplasmic trehalase in the catabolism of intracellular trehalose. *Biochim. Biophys. Acta*, **1245**, 339–347.

Boucherie,H., Bataille,N., Fitch,I.T., Perrot,M. and Tuite,M.F. (1995) Differential synthesis of glyceraldehyde-3-phosphate dehydrogenase polypeptides in stressed yeast cells. *FEMS Microbiol. Lett.*, **25**, 1127–1133.

Boy-Marcotte,E., Lagniel,G., Perrot,M., Bussereau,F., Boudsocq,A., Jacquet,M. and Labarre,J. (1999) The heat shock response in yeast: differential regulations and contributions of the Msn2p/Msn4p and Hsf1p regulons. *Mol. Microbiol.*, **33**, 274–283.

Browder,L.W., Pollock,M., Nickells,R.W., Heikkila,J.J. and Winning,R.S. (1989) Developmental regulation of the heat-shock response. *Dev. Biol.*, **6**, 97–147.

Brownstein,M.J., Trent,J.M. and Boguski,M.S. (1998) Functional genomics. *Trends Guide to Bioinformatics, Elsevier Trend Journals*, Suppl., 27–29.

Cascante,M., Curto,R. and Sorribas,A. (1995) Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using biochemical systems theory and metabolic control analysis. Steady-state analysis. *Math. Biosci.*, **130**, 51–69.

Curto,R., Sorribas,A. and Cascante,M. (1995) Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using biochemical systems theory and metabolic control analysis. Model definition and nomenclature. *Math. Biosci.*, **130**, 25–50.

Curto,R., Voit,E.O., Sorribas,A. and Cascante,M. (1997) Validation and steady-state analysis of a power-law model of purine metabolism. *Biochem. J.*, **324**, 761–775.

Curto,R., Voit,E.O., Sorribas,A. and Cascante,M. (1998a) Mathematical models of purine metabolism in man. *Math. Biosci.*, **151**, 1–49.

Curto,R., Voit,E.O., Sorribas,A. and Cascante,M. (1998b) Analysis of abnormalities in purine metabolism leading to gout and to neurological dysfunctions in man. *Biochem. J.*, **329**, 477–487.

Davidson,J.F., Whyte,B., Bissinger,P.H. and Schiestl,R.H. (1996) Oxidative stress is involved in heat-induced cell death in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **93**, 5116–5121.

Dickson,,R.C., Nagiec,E.E., Skrzypek,M., Tillman,P., Wells,G.B. and Lester,R.L. (1997) Sphingolipids are potential heat stress signals in *Saccharomyces*. *J. Biol. Chem.*, **272**, 30 196–30 200.

DeRisi,J.L., Iyer,V.R. and Brown,P.O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, **278**, 680–686.

Eisen,M.B., Spellman,P.T., Brown,P.O. and Botstein,D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA*, **95**, 14 863–14 868.

Entian,K.D., Frohlich,K.U. and Mecke,D. (1984) Regulation of enzymes and isoenzymes of carbohydrate metabolism in the yeast *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta*, **799**, 181–186.

Findly,R.C., Gillies,R.J. and Shulman,R.G. (1983) *In vivo* phosphorus-31 nuclear magnetic resonance reveals lowered ATP during heat shock of *Tetrahymena*. *Science*, **219**, 1223–1225.

Galazzo,J.L. and Bailey,J.E. (1990) Fermentation pathway kinetics and metabolic flux control in suspended and immobilized *Saccharomyces cerevisiae*. *Enzyme Microb. Technol.*, **12**, 162–172.

Galazzo,J.L. and Bailey,J.E. (1991) Errata. *Enzyme Microb. Technol.*, **13**, 363.

Gancedo,J. (1986) Carbohydrate metabolism in yeast. In Morgan,M.J. (ed.), *Carbohydrate Metabolism in Cultured Cells* Plenum Press, New York, pp. 245–287.

Godon,C., Lagniel,G., Lee,J., Buhler,J.M., Kieffer,S. Perrot,M. *et al.* (1998) The H_2O_2 stimulon in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **273**, 22 480–22 489.

Grba,S., Oura,E. and Suomalainen,H. (1979) Formation of trehalose and glycogen in growing baker's yeast. *Finn. Chem. Lett.*, **1979**, 61–64.

Holstege,F.C. P., Jennings,E.G., Wyrick,J.J., Lee,T.I., Hengartner,C.J., Green,M.R., Golub,T.R., Lander,E.S. and Young,R.A. (1998) Dissecting the regulatory circuitry of a eukaryotic genome. *Cell*, **95**, 717–728.

Hottiger,T., Schmutz,P. and Wiemken,A. (1987) Heat-induced accumulation and futile cycling of trehalose in *Saccharomyces cerevisiae*. *J. Bacteriol.*, **169**, 5518–5522.

Hottiger,T., De Virgilio,C.D., Hall,M.N., Boller,T. and Wiemken,A. (1994) Physiological concentrations of trehalose increase in thermal stability of proteins *in vitro*. *Eur. J. Biochem.*, **219**, 187–193.

Iida,H. and Yahara,I. (1985) Yeast heat shock protein of M_r 48 000 is an isoprotein of enolase. *Nature*, **315**, 688–690.

Jenkins,G.M., Richards,A., Wahl,T., Mao,C., Obeid,L. and Hannun,Y. (1997) Involvement of yeast sphingolipids in the heat stress response of *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **272**, 32 566–32 572.

Jollow,D. and McMillan,D. (1998) Ethnic variation and genetic susceptibility: glucose-6-phosphate dehydrogenase deficiency. In Mendelsohn,M.L., Mohr,L. and Peeters,J.P. (eds), *Biomarkers: Medical and Workplace Applications* Joseph Henry Press, Washington, DC, pp. 227–239.

Lindquist,S. (1986) The heat-shock response. *Annu. Rev. Biochem.*, **55**, 1151–1191.

Marsden,M., Nickells,R.W., Kapoor,M. and Browder,L.W. (1993) The induction of pyruvate kinase synthesis by heat shock in Xenopus laevis embryos. *Dev. Genet.*, **14**, 51–57.

Martini,G. and Ursini,M.V. (1996) A new lease of life for an old enzyme. *BioEssays*, **18**, 631–637.

McKerns,K.W. (1975) Glucose-6-phosphate dehydrogenase from cow adrenal cortex. *Meth. Enzymol.*, **41**, 188–196.

Neves,M.-J. and François,J. (1992) On the mechanism by which a heat shock induces trehalose accumulation in *Saccharomyces cerevisiae*. *Biochem. J.*, **288**, 859–864.

Ni,T.-C. and Savageau,M.A. (1996a) Application of biochemical systems theory to metabolism in human red blood cells. *J. Biol. Chem.*, **271**, 7927–7941.

Ni,T.-C. and Savageau,M.A. (1996b) Model assessment and refinement using strategies from biochemical systems theory: application to metabolism in human red blood cells. *J. Theor. Biol.*, **179**, 329–368.

Nickells,R.W. and Browder,L.W. (1988) A role for glyceraldehyde-3-phosphate dehydrogenase in the development of thermotolerance in Xenopus laevis embryos. *J. Cell Biol.*, **107**, 1901–1909.

Olive,C. and Levy,H.R. (1975) Glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides*. *Meth. Enzymol.*, **41**, 196–201.

Özcan,S. and Johnston,M. (1999) Function and regulation of yeast hexose transporters. *Microbiol. Mol. Biol. Rev.*, **63**, 554–569.

Panek,A.D. and Mattoon,Jr (1977) Regulation of energy metabolism in *Saccharomyces cerevisiae*. *Arch. Biochem. Biophys.*, **183**, 306–316.

Perry,D.K. and Hannun,Y.A. (1998) The role of ceramide in cell signalling. *Biochim. Biophys. Acta*, **1436**, 233–243.

PLAS: http://correio.cc.fc.ul.pt/∼aenf/plas.html.

Rodríguez-Acosta,F., Regelado,C.M. and Torres,N.V. (1999) Nonlinear optimization of biotechnological processes by stochastic algorithms application to the maximization of the production rate of ethanol, glycerol and carbohydrates by *Saccharomyces cerevisiae*. *J. Biotechnol.*, **68**, 15–28.

Romano,A. (1986) Sugar transport systems of baker's yeast and filamentous fungi. In Morgan,M.J. (ed.), *Carbohydrate Metabolism in Cultured Cells* Plenum Press, New York, pp. 225–244.

Rosenzweig,R.F. (1992) Regulation of fitness in yeast overexpressing glycolytic enzymes: responses to heat shock and nitrogen starvation. *Genet. Res.*, **59**, 167–177.

Savageau,M.A. (1969a) Biochemical systems analysis, I. Some mathematical properties of the rate law for the component enzymatic reactions. *J. Theor. Biol.*, **25**, 365–369.

Savageau,M.A. (1969b) Biochemical systems analysis, II. The steady-state solutions for an n-pool system using a power-law approximation. *J. Theor. Biol.*, **25**, 370–379.

Savageau,M.A. (1970) Biochemical systems analysis, III. Dynamic solutions using a power-law approximation. *J. Theor. Biol.*, **26**, 215–226.

Savageau,M.A. (1972) The behavior of intact biochemical control systems. *Curr. Top. Cell. Regl.*, **6**, 63–129.

Savageau,M.A. (1976) *Biochemical Systems Analysis. A Study of Function and Design in Molecular Biology*. Addison-Wesley, Reading, MA.

Savageau,M.A. (1991a) The challenge of reconstruction. *New Biol.*, **3**, 101–102.

Savageau,M.A. (1991b) Reconstructionist molecular biology. *New Biol.*, **3**, 190–197.

Savageau,M.A. (1992) Critique of the enzymologist's test tube. In Bittar,E.E. (ed.), *Fundamentals of Medical Cell Biology* vol. 3A, JAI Press, Greenwich, CT, pp. 45–108.

Savageau,M.A. (1995) Enzyme kinetics *in vitro* and *in vivo*: Michaelis–Menten revisited. In Bittar,E.E. (ed.), *Principles of Medical Biology* vol. 4, JAI Press, Greenwich, CT, pp. 93–146.

Savageau,M.A. (1996) A kinetic formalism for integrative molecular biology: manifestation in biochemical systems theory and use in elucidating design principles for gene circuits. In Collado-Vides,J., Magasanik,B. and Smith,T.F. (eds), *Integrative Approaches to Molecular Biology* MIT Press, Cambridge, MA, pp. 115–146.

Schena,M., Shalon,D., Davis,R.W. and Brown,P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467–470.

Scott,W.A. (1975) Glucose-6-phosphate dehydrogenase from *Neurospora crassa*. *Meth. Enzymol.*, **41**, 177–182.

Shepherd,M.G. (1975) Glucose-6-phosphate dehydrogenase from *Penicillium duponti*. *Meth. Enzymol.*, **41**, 201–205.

Shiraishi,F. and Savageau,M.A. (1992a) The tricarboxylic acid cycle in *Dictyostelium discoideum*. I. Formulation of alternative kinetic representations. *J. Biol. Chem.*, **267**, 22 912–22 918.

Shiraishi,F. and Savageau,M.A. The tricarboxylic acid cycle in *Dictyostelium discoideum*. II. Evaluation of model consistency and robustness. *J. Biol. Chem.*, **267**, 22 919–22 925.

Shiraishi,F. and Savageau,M.A. (1992c) The tricarboxylic acid cycle in *Dictyostelium discoideum*. III. Analysis of steady state and dynamic behaviour. *J. Biol. Chem.*, **267**, 22 926–22 933.

Shiraishi,F. and Savageau,M.A. (1992d) The tricarboxylic acid cycle in *Dictyostelium discoideum*. IV. Resolution of discrepancies between alterntative methods of analysis. *J. Biol. Chem.*, **267**, 22 934–22 943.

Shiraishi,F. and Savageau,M.A. (1993) The tricarboxylic acid cycle in *Dictyostelium discoideum*. V. Systemic effects of including protein turnover in the current model. *J. Biol. Chem.*, **268**, 16 917–16 928.

Shiraishi,F. and Savageau,M.A. (1996) The tricarboxylic acid cycle in *Dictyostelium discoideum*—Two methods of analysis applied to the same model. *J. Theor. Biol.*, **178**, 219–222.

Skrzypek,M.S., Nagiec,M.M., Lester,R.L. and Dickson,R.C. (1999) Analysis of phosphorylated sphingolipid long-chain bases reveals potential roles in heat stress and growth control in *Saccharomyces*. *J. Bacteriol.*, **181**, 1134–1140.

Sorribas,A., Curto,R. and Cascante,M. (1995) Comparative characterization of the fermentation pathway of *Saccharomyces cerevisiae* using biochemical systems theory and metabolic control analysis. Model validation and dynamic behavior. *Math. Biosci.*, **130**, 71–84.

Stanbury,J.B., Wyngaarden,J.B., Fredrickson,D.S., Goldstein,J.L. and Brown,M.S. (eds) (1983) *The Metabolic Basis of Inherited Disease*. 5th edn, McGraw-Hill, Auckland.

Teusink,B., Walsh,M.C., Dam,K.V. and Westerhoff,H.V. (1998) The danger of metabolic pathways with turbo design. *Trends Biochem. Sci.*, **23**, 162–168.

Thornton,J.M. (1998) The future of bioinformatics. *Trends Guide to Bioinformatics, Elsevier Trend Journals* Suppl., pp. 30–31.

Torres,N.V. (1994) Modelling approach to control of carbohydrate metabolism during citric acid accumulation by *Aspergillus niger*. *Biotechnol. Bioeng.*, **44**, 104–118.

Torres,N.V., Voit,E.O., Glez-Alcón,C. and Rodriguez,F. (1997) An indirect optimization method for biochemical systems:

description of method and application to the maximization of the rate of ethanol, glycerol, and carbohydrate production in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.*, **55**, 758–772.

Voit,E.O. (ed.) (1991) *Canonical Nonlinear Modeling. S-System Approach to Understanding Complexity*. Van Nostrand Reinhold, New York, xi+365.

Voit,E.O. (2000) *Computational Analysis of Biochemical Systems. A Practical Guide for Biochemists and Molecular Biologists*. Cambridge University Press, UK, xii+530.

Voit,E.O., Savageau,M.A. and Irvine,D.H. (1991) Introduction to S-systems. In Voit,E.O. (ed.), *Canonical Nonlinear Modeling. S-System Approach to Understanding Complexity* Van Nostrand Reinhold, New York, xi+365.

Wells,G.B., Dickson,R.C. and Lester,R.L. (1998) Heat-induced elevation of ceramide in *Saccharomyces cerevisiae* via *de novo* synthesis. *J. Biol. Chem.*, **273**, 7235–7243.

White,A., Handler,P. and Smith,E.L. (1968) *Principles of Biochemistry*. 4th edn, McGraw-Hill, New York.

Wiemken,A. (1990) Trehalose in yeast, stress protectant rather than reserve carbohydrate. *Antonie van Leeuwenhoek*, **58**, 209–217.

Winderickx,J., de Winde,J.H., Crauwels,M., Hino,A., Hohmann,S., Van Dijck,P. and Thevelein,J.M. (1996) Regulation of genes encoding subunits of the trehalose synthase complex in *Saccharomyces cerevisiae*: novel variations of STRE-mediated transcription control? *Mol. Gen. Genet.*, **252**, 470–482.

Winkler,K., Kienle,I., Burgert,M., Wagner,J. -C. and Holzer,H. (1991) Metabolic regulation of the trehalose content of vegetative yeast. *FEBS Lett.*, **291**, 269–272.

Wodicka,L., Dong,H., Mittman,M., Ho,M.H. and Lockhart,D.J. (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat. Biotech.*, **15**, 1359–1367.