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; \( V_{i0} \)) 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)

<table>
<thead>
<tr>
<th>ORF</th>
<th>Gene</th>
<th>Enzyme name</th>
<th>Fold induction</th>
<th>mRNA copies/cell</th>
<th>Transcription (copies/hour)</th>
</tr>
</thead>
<tbody>
<tr>
<td>YPR026W</td>
<td>ATH1</td>
<td>Vacuolar α,α'-trehalase</td>
<td>1.99</td>
<td>0.2</td>
<td>0</td>
</tr>
<tr>
<td>YGR254W</td>
<td>ENO1</td>
<td>Enolase I</td>
<td>1.91</td>
<td>17.10</td>
<td>48.1</td>
</tr>
<tr>
<td>YHR174W</td>
<td>ENO2</td>
<td>Enolase 2</td>
<td>2.07</td>
<td>61.10</td>
<td>134.7</td>
</tr>
<tr>
<td>YKL060C</td>
<td>FBA1</td>
<td>Fructose-bisphosphate aldolase II</td>
<td>1.09</td>
<td>74.10</td>
<td>167.9</td>
</tr>
<tr>
<td>YCL040W</td>
<td>GLK1</td>
<td>Glucokinase</td>
<td>18.77</td>
<td>3.70</td>
<td>3.8</td>
</tr>
<tr>
<td>YKL152C</td>
<td>GPM1</td>
<td>Phosphoglycerate mutase</td>
<td>1.17</td>
<td>32.30</td>
<td>64.0</td>
</tr>
<tr>
<td>YDL021W</td>
<td>GPM2</td>
<td>Phosphoglycerate mutase</td>
<td>3.76</td>
<td>0.10</td>
<td>0</td>
</tr>
<tr>
<td>YOL056W</td>
<td>GPM3</td>
<td>Phosphoglycerate mutase</td>
<td>0.51</td>
<td>0.90</td>
<td>0.3</td>
</tr>
<tr>
<td>YFR065C</td>
<td>HXK1</td>
<td>Hexokinase I</td>
<td>14.03</td>
<td>3.70</td>
<td>5.8</td>
</tr>
<tr>
<td>YGL253W</td>
<td>HXK2</td>
<td>Hexokinase II</td>
<td>1.91</td>
<td>19.30</td>
<td>71.9</td>
</tr>
<tr>
<td>YHR094C</td>
<td>HXT1</td>
<td>Hexose transporter (low affinity)</td>
<td>0.79</td>
<td>11.2</td>
<td>41</td>
</tr>
<tr>
<td>YMR011W</td>
<td>HXT2</td>
<td>Hexose transporter (high affinity)</td>
<td>7.11</td>
<td>7.0</td>
<td>12.4</td>
</tr>
<tr>
<td>YDR345C</td>
<td>HXT3</td>
<td>Hexose transporter (low affinity)</td>
<td>1.85</td>
<td>12.9</td>
<td>37.5</td>
</tr>
<tr>
<td>YHR092C</td>
<td>HXT4</td>
<td>Hexose transporter (low affinity)</td>
<td>3.18</td>
<td>3.1</td>
<td>10.6</td>
</tr>
<tr>
<td>YDR343C</td>
<td>HXT6</td>
<td>Hexose transporter (high affinity)</td>
<td>21.56</td>
<td>11.10</td>
<td>18.9</td>
</tr>
<tr>
<td>YDR342C</td>
<td>HXT7</td>
<td>Hexose transporter (high affinity)</td>
<td>19.43</td>
<td>9.90</td>
<td>15.1</td>
</tr>
<tr>
<td>YIL214W</td>
<td>HXT8</td>
<td>Hexose transporter (putative)</td>
<td>1.27</td>
<td>0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>YIL170W</td>
<td>HXT9</td>
<td>Hexose transporter (putative)</td>
<td>1.37</td>
<td>0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>YDR001C</td>
<td>NTH1</td>
<td>Neutral α,α'-trehalase</td>
<td>1.75</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>YBR001C</td>
<td>NTH2</td>
<td>Neutral α,α'-trehalase</td>
<td>4.41</td>
<td>0.4</td>
<td>1.3</td>
</tr>
<tr>
<td>YGR240C</td>
<td>PFK1</td>
<td>Phosphofructokinase alpha subunit</td>
<td>0.80</td>
<td>8.80</td>
<td>23</td>
</tr>
<tr>
<td>YMR205C</td>
<td>PFK2</td>
<td>Phosphofructokinase beta subunit</td>
<td>1.17</td>
<td>6.80</td>
<td>20</td>
</tr>
<tr>
<td>YBR196C</td>
<td>PGI1</td>
<td>Glucose 6-phosphate isomerase</td>
<td>2.13</td>
<td>15.40</td>
<td>30.7</td>
</tr>
<tr>
<td>YCR012W</td>
<td>PGK1</td>
<td>Phosphoglycerate kinase</td>
<td>2.23</td>
<td>60.40</td>
<td>110.5</td>
</tr>
<tr>
<td>YKL127W</td>
<td>PGM1</td>
<td>Phosphoglucomutase, minor isom</td>
<td>1.10</td>
<td>1.80</td>
<td>4.9</td>
</tr>
<tr>
<td>YMR105C</td>
<td>PGM2</td>
<td>Phosphoglucomutase, major isom</td>
<td>1.14</td>
<td>0.30</td>
<td>0</td>
</tr>
<tr>
<td>YAL038W</td>
<td>PYK1</td>
<td>Pyruvate kinase</td>
<td>1.13</td>
<td>50.40</td>
<td>101.3</td>
</tr>
<tr>
<td>YOR347C</td>
<td>PYK2</td>
<td>Pyruvate kinase</td>
<td>3.81</td>
<td>0.30</td>
<td>1</td>
</tr>
<tr>
<td>YIL052W</td>
<td>TDH1</td>
<td>Glycereraldehyde 3-phosphate dehydrogenase 1</td>
<td>2.20</td>
<td>3.60</td>
<td>12.7</td>
</tr>
<tr>
<td>YJR090C</td>
<td>TDH2</td>
<td>Glycereraldehyde 3-phosphate dehydrogenase 2</td>
<td>1.75</td>
<td>34.10</td>
<td>98.9</td>
</tr>
<tr>
<td>YGR192C</td>
<td>TDH3</td>
<td>Glycereraldehyde 3-phosphate dehydrogenase 2</td>
<td>1.85</td>
<td>89.00</td>
<td>188.1</td>
</tr>
<tr>
<td>YBR126C</td>
<td>TPS1</td>
<td>α,β-Trehalose 6-Phosphate Synthase</td>
<td>7.06</td>
<td>2.8</td>
<td>2.8</td>
</tr>
<tr>
<td>YDR074W</td>
<td>TPS2</td>
<td>α,β-Trehalose 6-Phosphate Phosphatase</td>
<td>4.69</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>YMR261C</td>
<td>TPS3</td>
<td>α,β-Trehalose 6-Phosphate Synthase</td>
<td>3.03</td>
<td>1</td>
<td>1.9</td>
</tr>
<tr>
<td>YNL241C</td>
<td>ZWF1</td>
<td>Glucose 6-phosphate dehydrogenase</td>
<td>3.76</td>
<td>1.60</td>
<td>3.8</td>
</tr>
</tbody>
</table>

*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 $\alpha_i$ and $\beta_i$ and kinetic orders $g_{ij}$ and $h_{ij}$ reads

$$\dot{X}_i = \alpha_i X_1^{g_{i1}} \cdots X_n^{g_{i,n+m}} - \beta_i X_1^{h_{i1}} \cdots X_n^{h_{i,n+m}}.$$  (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
S. M. Radosavljevic and T. Radivojevitch

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; Rodriguez-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 Martin and Ursini, 1996). The reason to include the G6PDH branch here is that the reduced coenzyme 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; Wiederlick 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 and uptake (Curto et al., 1995), were later used by others for purposes of yield optimization (Torres et al., 1997; Rodriguez-Acosta 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

\[
\dot{X}_1 = 0.9018X_2^{-0.2344}X_6 - 3.1833X_1^{0.7464}X_5^{0.0243}X_7
\]
\[
\dot{X}_2 = 3.1833X_1^{0.7464}X_5^{0.0243}X_7 - 0.8187X_2^{0.664}X_5^{-0.354}
\]
\[
\dot{X}_3 = 0.5234X_2^{0.7318}X_5^{0.3941}X_8 - 0.0148X_3^{0.584}X_4^{0.03}
\]
\[
\dot{X}_4 = 0.022X_3^{0.6159}X_5^{1.308}X_9X_1^{-0.6088} - 0.0945X_4^{0.05}
\]
\[
\dot{X}_5 = 0.0880X_3^{0.3505}X_4^{0.2665}X_5^{0.0243}X_9^{0.5}X_{10}X_{14}^{-0.3044}
\]
\[
- 3.3270X_1^{0.2123}X_2^{0.190}X_5^{0.362}X_7^{0.286}X_8^{0.257}
\]
\[
\times X_1^{-0.002}X_3^{-0.457}
\]
The model contains dependent variables \( X_1, X_2, X_3, X_4, X_5 \) and independent variables \( X_6, X_7, X_8, X_9, X_{10} \). The independent variables are glucose transport, 19.7 mM min\(^{-1}\), hexokinase/glucokinase, 68.5 mM min\(^{-1}\), phosphofructokinase, 31.7 mM min\(^{-1}\), triokinase, 49.9 mM min\(^{-1}\), and trehalase, 3440 mM min\(^{-1}\). The dependent variables are glycogen and trehalose production, the total carbohydrate flux, and the total ATP flux, which are given as the first product of power-law functions on the right-hand side of the fifth equation in (2):

\[
V_{\text{ATP}} = 0.0880X_3^{0.3505}X_4^{0.2665}X_5^{0.0243}X_9^{0.5}X_{10}^{0.5}X_{14}^{0.3044}.
\]

(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 \textit{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 \( \beta_{2,2}X_2^{0.222}X_{11} \) and has a baseline value of 0.014. The kinetic order associated with G6P (\(= X_2 \)) is \( h_{2,2} \). If the parameter \( h_{2,2} = 0.8 \), the flux reads

\[
V_{\text{Tre}} = 0.00098X_2X_{11}.
\]

(4)

If we instead set \( h_{2,2} = 0.8 \), the flux is

\[
V_{\text{Tre}} = 0.0008925X_2^{0.8}X_{11}.
\]

(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 \( h_{2,2} \), the model allows us to evaluate its effect. Setting \( h_{2,2} = 0 \) corresponds to a process saturated with respect to G6P, and setting \( h_{2,2} = 1 \) 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 \textit{et al.}, 1983, p. 1632). The \( K_M \) of G6PDH with respect to G6P for a variety of microorganisms is about 50 \(\mu\)M (McKerns, 1975; Olive and Levy, 1975; Scott, 1975; Shepherd, 1975; Stanbury \textit{et al.}, 1983). The kinetic order of the G6PDH flux with respect to G6P at steady state (\( ([\text{G6P}] = 1.011 \text{ mM}) \) is thus approximately \( K_M/\tau \) (G6P) = \( (50/1,001) \approx 0.05 \) (for this type of computation, see Voit \textit{et al.}, 1991, p. 55). The flux of NADPH production is therefore modeled as

\[
V_{\text{NADPH}} = 1.769X_2^{0.05}X_{15}
\]

(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 \textit{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 over-expressed, 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 phosphofructophosphate 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 × 1.13 + 0.30 × 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 isofrom 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 UDPG. 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 Neurospora 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\textsubscript{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; 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).

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

   \[
   \ln \left( \frac{\text{ATP \ flux}_{\text{Hypo}}}{\text{ATP \ flux}_{\text{HS}}} \right).
   \]

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

<table>
<thead>
<tr>
<th>Scenario</th>
<th>HXT</th>
<th>GLK</th>
<th>PFK</th>
<th>TDH (X_9)</th>
<th>PYK (X_10)</th>
<th>TPS</th>
<th>G6PDH</th>
<th>Responses</th>
<th>Perf.</th>
<th>Cost</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>VA, VN, VT</td>
<td>−11.0</td>
<td>0</td>
<td>−11.0</td>
</tr>
<tr>
<td>HS</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>4.5</td>
<td>3.8</td>
<td>✔</td>
<td>0</td>
<td>9.64</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>VA, VN, VT</td>
<td>−8.58</td>
<td>4.85</td>
<td>−13.43</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>1.5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>PEP, VT</td>
<td>−8.64</td>
<td>13.28</td>
<td>−21.92</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Glu †;</td>
<td>−9.92</td>
<td>2.30</td>
<td>−12.22</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.5</td>
<td>5</td>
<td>4.5</td>
<td>3.8</td>
<td>−6.46</td>
<td>4.32</td>
<td>−10.78</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>1.5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>Glu, FDP †;</td>
<td>−7.03</td>
<td>5.14</td>
<td>−12.17</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>FDP, PEP †;</td>
<td>−7.86</td>
<td>3.91</td>
<td>−11.78</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>FDP †; PEP †;</td>
<td>−4.88</td>
<td>6.75</td>
<td>−11.63</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>FDP, PEP †; VT †;</td>
<td>−8.30</td>
<td>8.14</td>
<td>−16.44</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>5</td>
<td>0.25</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>FDP †; PEP †;</td>
<td>−4.51</td>
<td>5.36</td>
<td>−9.88</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Glu †;</td>
<td>−3.53</td>
<td>7.85</td>
<td>−11.38</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1.2</td>
<td>5</td>
<td>4.5</td>
<td>3.8</td>
<td>−2.01</td>
<td>8.54</td>
<td>−10.55</td>
</tr>
<tr>
<td>12</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1.5</td>
<td>4</td>
<td>5</td>
<td>4.5, 3.8</td>
<td>−3.40</td>
<td>11.76</td>
<td>−15.16</td>
</tr>
</tbody>
</table>

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 NADPH 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, HKX/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, HKX/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, HKX/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, HKX/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
αi - S-system parameter: rate constant of production term for variable Xi
ATH - trehalase gene
βi - S-system parameter: rate constant of degradation term for variable Xi
BST - Biochemical Systems Theory
FBA - aldolase gene
ENO - enolase gene
FDP - fructose 1,6-diphosphate
F6P - fructose 6-phosphate
GAPD - glyceraldehyde 3-phosphate dehydrogenase

8ij - S-system parameter: kinetic order, representing the effect of variable Xj on the production of variable Xi

GLK - glucokinase (gene/enzyme)
GLY-3-P - glyceraldehyde 3-phosphate
GPM - phosphoglycerate mutase gene
G6P - glucose 6-phosphate
G6PDH - glucose 6-phosphate dehydrogenase

hij - S-system parameter: kinetic order, representing the effect of variable Xj on the degradation of variable Xi

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
Vin - 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
Xi - S-system variable; typically Xi is a dependent (state) variable if i ≤ n, or an independent (control)
variable if \( n + 1 \leq i \leq n + m \).

\[
\text{ZWF} = \text{glucose 6-phosphate dehydrogenase gene}
\]

**References**


