Integrative Biology

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/ibiology

Integrative Biology

RSCPublishing

ARTICLE

An entropy-like index of bifurcational robustness for metabolic systems

Cite this: DOI: 10.1039/xoxxooooox

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxooooox

www.rsc.org/

Jimmy G. Lafontaine Rivera[a#], Yun Lee[a#] & James C. Liao[a,b*]

Natural and synthetic metabolic pathways need to retain stability when faced against random changes in gene expression levels and kinetic parameters. In the presence of large parameter changes, a robust system should specifically avoid moving to an unstable region, an event that would dramatically change system behavior. Here we present an entropy-like index, denoted as *S*, for quantifying the bifurcational robustness of metabolic systems against loss of stability. We show that *S* enables the optimization of a metabolic model with respect to both bifurcational robustness and experimental data. We then demonstrate how the coupling of Ensemble Modeling and *S* enables us to discriminate alternative designs of a synthetic pathway according to bifurcational robustness. Finally, we show that *S* enables the identification of a key enzyme contributing to the bifurcational robustness of yeast glycolysis. The different applications of *S* demonstrated illustrate the versatile role it can play in constructing better metabolic models and designing functional non-native pathways.

Introduction

The major role of most metabolic systems is to support cellular growth, maintenance, or adaptation without losing stability despite perturbations in the environment[1,2]. When the environmental or physiological conditions change, gene expression levels or kinetic parameters may drift outside their typical working ranges and lose stability. The stochastic nature of transcriptional and translational mechanisms[3] is one such source of noise and a robust system would maintain homeostasis despite the perturbations. Failure to retain a stability may lead to accumulation of toxic metabolites or depletion of essential intermediates. This detrimental state could cause growth arrest and has been linked to many diseases such as diabetes and cancer[4]. In metabolic engineering, this situation leads to loss of production or cell death.

Non-linear system behavior changes qualitatively and dramatically when parameters cross a bifurcation point and exhibits instability or multiplicity of steady states. Thus, a necessary but insufficient requirement for stable or damped-oscillatory metabolic pathway design is to avoid crossing a bifurcation point in the presence of random perturbations in gene expression levels and environmental conditions that change kinetic parameters. For metabolic systems, a fixed-point bifurcation may cause the stable steady state (or fixed-point) to become unstable[5-7], or mark the emergence of undamped oscillations[8,9] or multiple steady states[10] .

The distance away from an unstable region is defined as the bifurcational robustness[5], which measures the ability to return to a fixed point upon perturbation. Thus, building a theoretical foundation of robustness, and in particular defining a simple way to quantify it, represents a key challenge in systems biology[11]. For small, local perturbations, stability criteria are well defined using linear stability analysis[12]. For large perturbations, one must explore global properties of the system. It is important to make the distinction between bifurcational robustness, which quantifies the tendency to avoid sudden change in dynamic regime due to parameter changes, and local sensitivity, which quantifies the changes in performance (flux, period of oscillation) as a function of changes in parameters within the same dynamic regime. The latter does not quantify the distance away from bifurcation.

Natural metabolic pathways are presumed to be at least bifurcationally robust against stochastic changes in protein expression levels. Thus models of natural metabolic pathways

need to be similarly robust. However, there is no quantitative way to characterize bifurcational robustness in the presence of random parameter changes. Without a quantitative index, optimization of models for bifurcational robustness becomes difficult, if not impossible. Therefore, our goal here is to develop a quantitative index for bifurcational robustness, and show that such an index enables the optimization of the bifurcational robustness of metabolic models. The developed index is easy to compute and applies to metabolic systems of various scale and complexity. Interestingly, the mathematical form of our robustness index resembles the definition of entropy in thermodynamics and information theory[13]. We show that this entropy-like index, denoted as *S*, negatively correlates with bifurcational robustness. Metabolic systems with a small *S* are highly robust against bifurcation, and are more likely to retain a steady state under random perturbations affecting every enzyme than systems with a large *S*.

The utility of *S* was demonstrated through three examples. First, we show that the bifurcational robustness of a native pathway model can be significantly improved by applying a multi-objective optimization with *S* as an objective. Second, we show that the integration of *S* with EMRA[5] is able to discriminate, without any prior knowledge of kinetic parameters, the difference in bifurcational robustness between two configurations of a non-native metabolic pathway[14]. Possible sources of pathway failure were also identified. Finally, by quantifying *S* in a series of yeast glycolysis models incorporating different features, we identified pyruvate decarboxylase as a key enzyme determining the robustness of yeast glycolysis, a finding consistent with earlier studies[15,16]. Together, our results demonstrate that *S* may serve as an unbiased standard by which the bifurcational robustness is judged.

Results

Lack of robustness in existing metabolic models

In a survey of the robustness of existing metabolic models, we simulated natural perturbations and recorded the response of thirteen kinetic models of metabolic pathways with fitted parameters (Supp. Fig. 1). This *in silico* experiment was designed to mimic the real biological situation where protein expression levels, which affect kinetic parameters, vary randomly and non-specifically[3]. Since natural metabolic pathways are presumed to be bifurcationally robust in such situations, the computational models of these metabolic systems need to be similarly robust. To our surprise, the selected models displayed varying degrees of robustness against bifurcation. Some models are very robust and almost always retain stability after perturbation, even though their steady-state flux and metabolite concentrations are changed. Others respond poorly even to moderate perturbations (Supp. Fig. 1) where the system becomes unstable with some metabolites accumulating or vanishing, leading to system failure. If one accepts the assumption that natural systems are robust against enzyme expression perturbations, these models do not reflect this assertion. This unexpected result highlighted the need for the optimization of metabolic models with respect to bifurcational robustness, which in turn calls for the development of a quantitative robustness index.

Searching for a robustness index

The natural perturbation of environmental or physiological conditions often affects the expression levels of many genes, which in turn affect the kinetic parameters of all enzymes. Thus, an appropriate description of bifurcational robustness should focus on the probability that a metabolic system will not cross a bifurcation point and retain a stable steady state when every enzyme is randomly perturbed. Although this probability of retaining stability, denoted as P_{SS}, is a function of all enzymes, a simple approximation can be obtained under the assumption that there is no "crosstalk" between enzymes. That is, the probability of stability retention under the perturbation of a single enzyme is independent of other enzymes. In this simplest case, the expression for P_{SS} reduces to

$$P_{SS} \approx \prod_{i=1}^{n} p_i, \quad (1)$$

where p_i denotes the probability that the system will retain stability if only enzyme i is subject to variation.

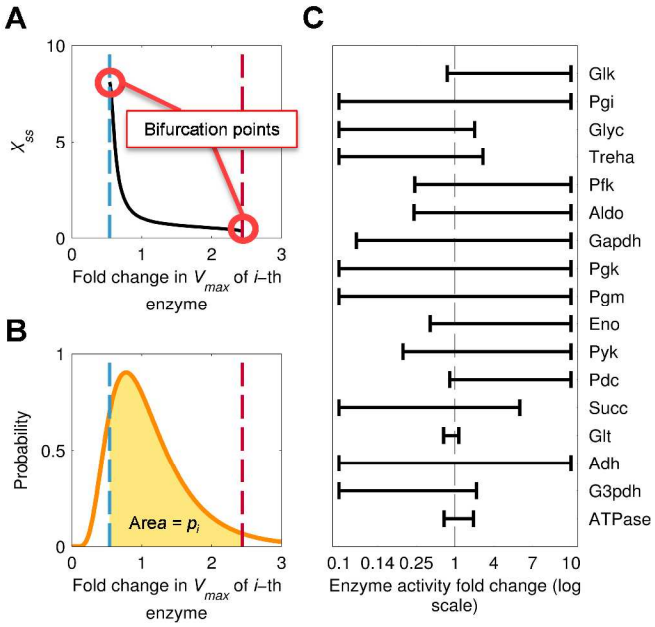

Figure 1. The continuation method enables the detection of bifurcation points. (**A**) Starting from a default steady state, the continuation method traces the trajectory of X_{SS} (the steady-state value of some metabolite concentration *X*) as it varies according to enzyme activity levels. In this example, the system loses stability (no stable fixed point exisits) when the enzyme activity is increased by over 2.5-fold or decreased by over 50%. (**B**) Given the probability density function (pdf; orange curve), we can calculate p_i (the probability of retaining a steady state when *i*-th enzyme is subject to a random perturbation) as the area under the pdf and between the bifurcation points (red and blue dashed lines). (**C**) The bifurcation points of V_{max} define the boundaries of single-enzyme perturbations. Abbreviations are defined in Supplementary Table 1.

When the environmental perturbation affects only one enzyme, the probability that the system will cross a bifurcation point, denoted as p_i, is determined by the area bounded by the bifurcation points and the probability density function (Fig. 1A & B; Methods). Fortunately, the bifurcation points can be readily determined based on the continuation method described previously[5]. As an example, Figure 1C shows the bifurcation points with respect to each enzyme in Teusink *et al*'s yeast glycolysis model[17]. Clearly, the model can tolerate wide variation in some enzymes, such as glucose 6-phosphate isomerase (PGI) and phosphoglycerate kinase (PGK), whereas a moderate perturbation in other enzymes, such as ATPase, can lead to a bifurcation and the loss of stability.

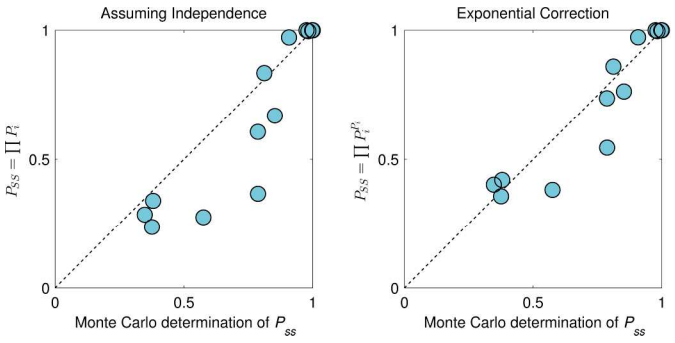

Figure 2. Comparison of the robustness (Probability of retaining stability, P_{ss}) determined by *in silico* experiment and two approximations. (A) The first theoretical approximation was $P_{SS} \approx \prod p_i$. **(B)** The second theoretical approximation was $P_{SS} \approx \prod p_i^{p_i}$. In both cases the Monte Carlo approximation of P_{SS} was calculated by randomly perturbing every enzyme.

For any metabolic model with all parameters specified or fitted, we can calculate p_i for every enzyme and use the formula in Eq. 1 to approximate P_{SS}. Figure 2 shows the performance of Eq. 1 in approximating the P_{SS}'s of the 13 metabolic models discussed above when compared to the Monte Carlo simulation (Methods). Clearly, the approximation yields a similar trend as the simulation results, but it tends to underestimate the true probability and may not be an appropriate index.

In an attempt to remedy the observed underestimation from the independence assumption, we applied an exponential correction factor to each individual probability:

$$P_{SS} \approx \prod_i p_i^{p_i} . \quad (2)$$

With an exponential correction, our rationale is to increase the value of p_i since the effect of crosstalk is more likely to be strong when p_i becomes smaller. Figure 2B shows the performance of this new approximation against the Monte Carlo simulation. Compared to the approximation under an independence assumption, Eq. 2 yields an improved correlation with the simulation results. Since determination of crosstalk between every enzyme is technically challenging, if not impossible, we believe that Eq. 2 provides a reasonable approximation without increasing computational costs. Interestingly, the mathematical form of Eq. 2 resembles the definition of entropy in thermodynamics and information theory[13], except that the exponent does not have a negative sign.

Here we propose an entropy-like robustness index, denoted as S, which corresponds to the negative of the logarithm of Eq. 2:

$$S = -\sum_i p_i \log(p_i). \quad (3)$$

Like entropy, S also enjoys the additive property as do thermodynamic and information theoretic entropy. That is, the system-level robustness S can be regarded as a simple sum of the enzyme-level robustness, $S_i = -p_i \cdot \log(p_i)$, which is determined solely by p_i. One difference between S and the thermodynamic entropy is that each pi is not mutually exclusive. Other thermodynamic properties or information theoretic properties of entropy do not easily carry over to the robustness index S. In robustness, the lower S the more robust the system is.

To test the performance of S as an index of bifurcational robustness, we calculated S for 13 metabolic models and compared those values to the Monte Carlo simulation of random perturbations (Figure 3). As expected, for the 13

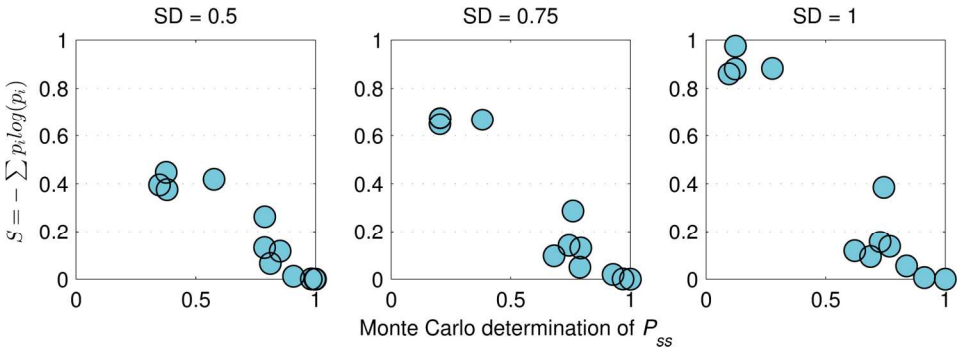

Figure 3. *S* is a proper index of bifurcational robustness for metabolic systems. For the 13 BioModels database models (blue filled circles; Methods), *S* decreases with increasing probability of stability retention (P_{ss}), which is determined by Monte Carlo simulation.

Integrative Biology

RSCPublishing

ARTICLE

models considered, *S* negatively correlates with bifurcational robustness. That is, models with a small *S* are highly robust against bifurcation, and are more likely to retain a steady state under random perturbations than models with a large *S*. In fact, the difference in *S* between robust and non-robust systems becomes more apparent when a higher perturbation level is tested. These results suggest that *S*, which can be calculated efficiently using the continuation method[5], may serve as an unbiased standard of bifurcational robustness. In the following sections, we will demonstrate how *S* can be used to: (i) optimize the bifurcational robustness of metabolic models with fitted parameters; (ii) compare the bifurcational robustness of alternative synthetic pathway designs when parameters are unknown; and (iii) identify key features determining the bifurcational robustness of a metabolic system.

Parameter optimization using S

Given its scalar nature, *S* can be readily incorporated into commonly used optimization algorithms for parameter fitting. To demonstrate this functionality, we re-fit the parameters of Teusink *et al*'s yeast glycolysis model[17] by using *S* as an optimization objective. In this particular case, we applied a multi-objective optimization so as to simultaneously (i) minimize the discrepancy between available data (metabolite concentrations and fluxes) and model predictions, and (ii) minimize the model's *S* value (See Methods). By incorporating *S* in the objective function, this algorithm was indeed able to significantly improve the bifurcational robustness of an otherwise non-robust model (Figure 4A). More importantly, neither metabolite concentrations nor fluxes required anything larger than a 2-fold change to accomplish this (Figure 4B). These results demonstrate that *S* enables the optimization of bifurcational robustness of a metabolic model and that such robustness optimization can be readily integrated into any parameter-fitting routine.

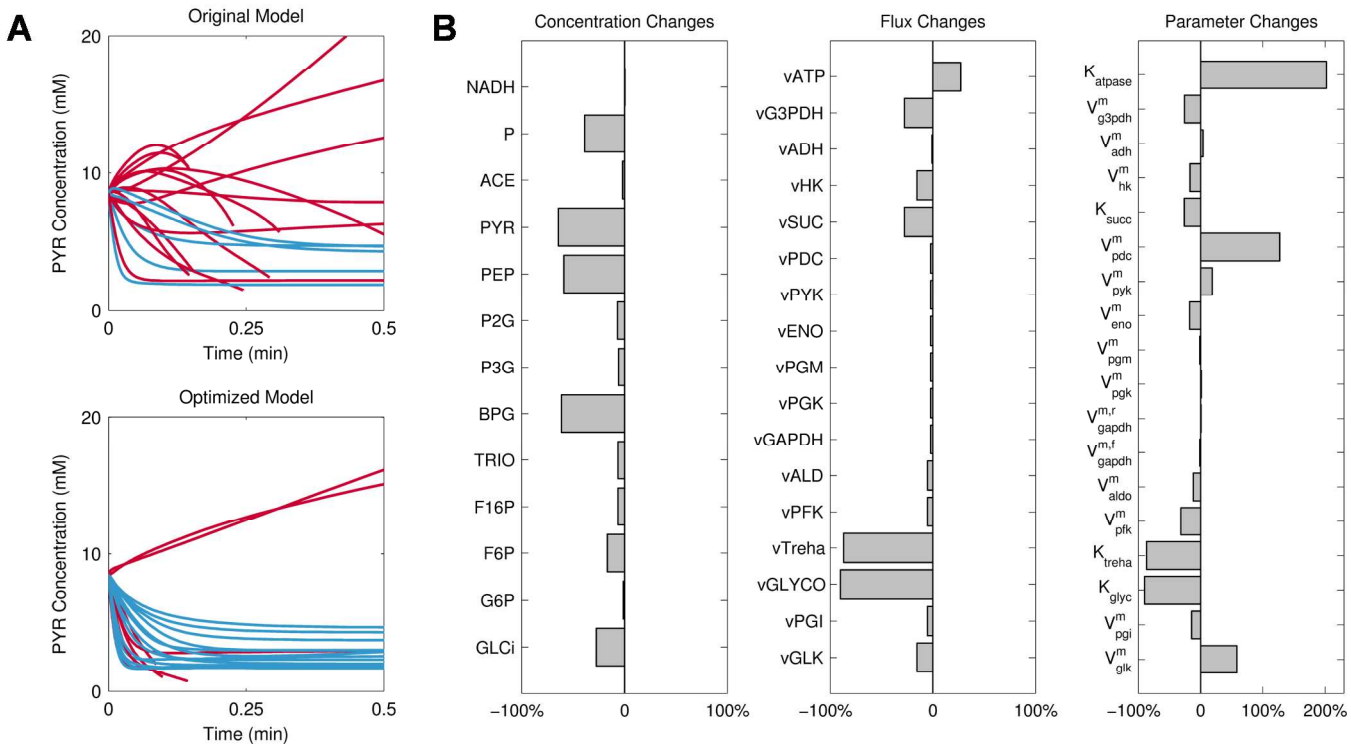

Figure 4. Optimization incorporating *S* returns a model with significantly improved robustness. (A) Comparison of the bifurcational robustness in the original and the optimized model by random perturbation. (—)Stability retained; (—) Stability lost **(B)** Percentage change in steady-state metabolite concentrations, steady-state fluxes and kinetic parameters between the best model returned by optimization and Teusink *et al.*'s model. For abbreviations, see Supplementary Tables 1 and 2.

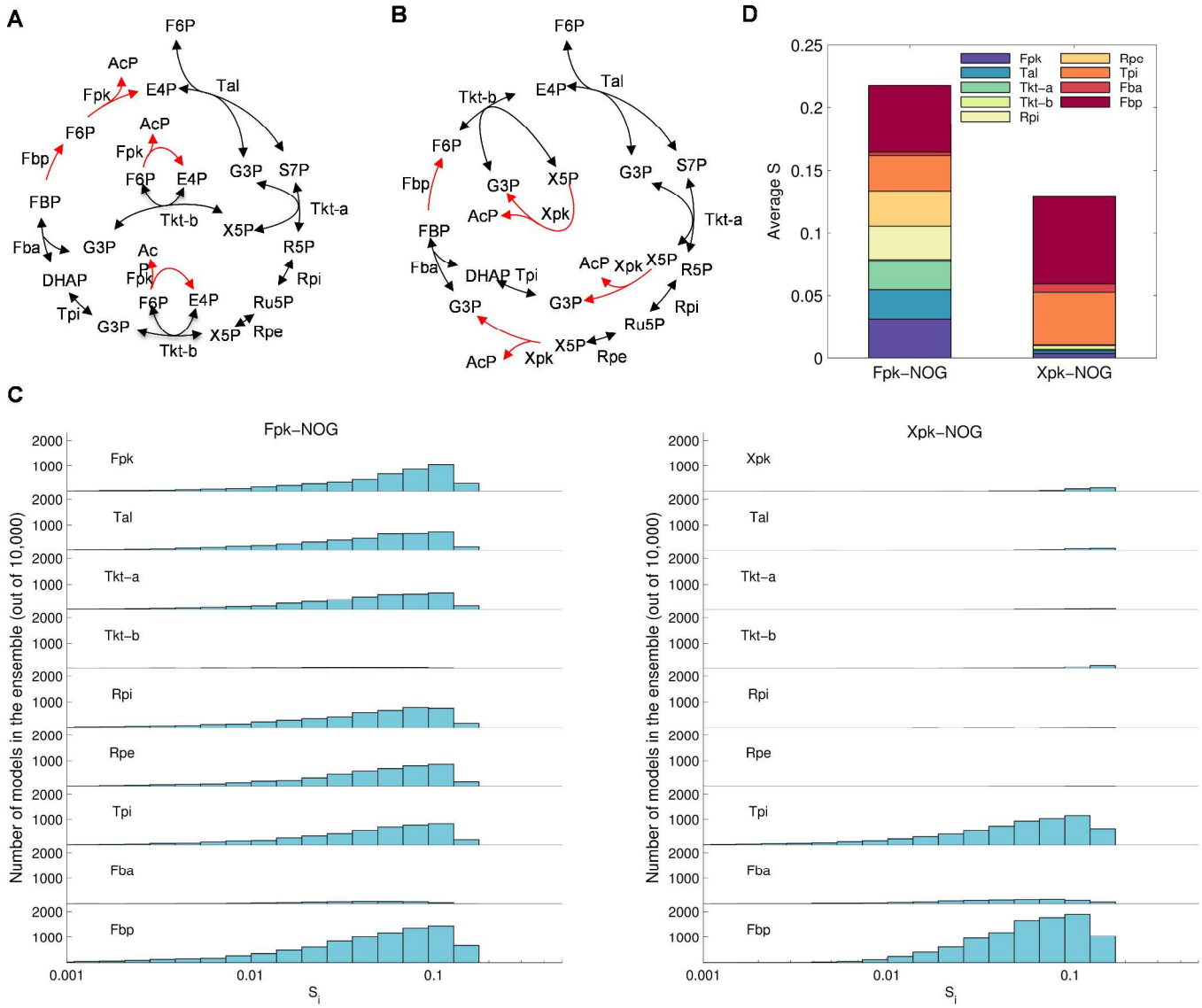

Figure 5. Robustness analysis of two configurations of a non-oxidative glycolysis (NOG) using *S*. (A-B) Two configurations of NOG where black double arrows denote reversible reactions and red arrows denote irreversible reactions. **(A)** In Fpk-NOG, phosphoketolase exhibits enzymatic activity only towards F6P and is denoted Fpk **(B)** In Xpk-NOG, phosphoketolase exhibits enzymatic activity only towards X5P and is denoted Xpk **(C)** Histogram of Si throughout the ensemble for every enzyme in each configuration. **(D)** The average *S* for each configuration is calculated as the sum of the average S_i of every enzyme.

Robustness index in the design of non-native pathways

Besides the optimization of the bifurcational robustness of existing models, *S* can be useful in non-native pathway design even when the kinetic parameters are unavailable. To address the uncertainty of kinetic parameters, we applied the calculation of *S* to an ensemble of models representing the feasible kinetic space[18-20]. Such an ensemble approach has recently been adopted to evaluate the bifurcational robustness of non-native pathways and to identify configurations that are more likely to be functional[5].

Here we demonstrate the utility of *S* in non-native pathway design using two configurations of a synthetic non-oxidative glycolysis (NOG)[14]: Fpk-NOG (Fig. 5A) and Xpk-NOG (Fig. 5B). Fpk-NOG contains a specific homolog of phosphoketolase (termed Fpk) that only reacts with fructose 6-phosphate (F6P), whereas the phosphoketolase in Xpk-NOG (termed Xpk) only reacts with xylulose 5-phosphate (X5P). For each configuration, we constructed an ensemble of 10,000 models by random sampling (Methods) and calculated the distributions of S_i. As shown in Fig. 5C, both Fpk-NOG and Xpk-NOG are quite sensitive to the changes in the activity of triose phosphate isomerase (Tpi) and fructose 1,6-bisphosphatase (Fbp) (Fig. 5C). Nevertheless, the Fpk-NOG is considerably less robust than Xpk-NOG as it is also sensitive to many other enzymes (Fig. 5C; left column). This conclusion is also confirmed by Fig. 5D, where the average *S* of each configuration is visualized as the stacked contributions of average S_i. Although Xpk-NOG as a whole has a lower average *S* than Fpk-NOG, the high

average S_i of Tpi and Fbp might indicate a potential problem during strain construction. As this example illustrates, the coupling of Ensemble Modeling with the calculation of S allows us to assess the robustness of a pathway design and identify possible causes of failure.

Determining the cause of non-robustness

Another possible use of *S* is in the identification of key features contributing to the bifurcational robustness of a metabolic system. To demonstrate this utility, we used van Heerden *et al.'s* model of yeast glycolysis[21] as an example. This model is adapted from the model developed by Teusink *et al.*[17] with five major changes:

1. Hexokinase (HK) inhibition by glucose 6-phosphate (G6P)
2. Consideration of phosphate as a free variable
3. Activation of pyruvate kinase (PYK) by fructose 1,6-bisphosphate (FBP)
4. A 6.1 fold rise in the V_{max} of pyruvate decarboxylase (PDC)
5. Trehalose and glycogen fluxes were considered as functions of G6P

Although these modifications are seemingly minor, the new model has a significantly lower *S* than the original model of Teusink *et al.* (Figure 6, inset). This is particularly interesting because other adaptations of Teusink *et al.*'s model, such as Pritchard *et al.*'s[22] and Conant *et al.*'s[23] glycolysis models, do not show a similar improvement in bifurcational robustness (Supplementary Fig. 1). This result suggests that some modifications made by van Heerden *et al.* are particularly important for robustness improvement.

To identify the key determinant of robustness in van Heerden's model, we constructed 16 alternative models by reversing all possible combinations of the first four major changes and calculated their respective *S* (Figure 6). As expected, the case where all four major changes were reversed (Figure 6, blue bar) is lost the robustness and exhibits a 43-fold increase in *S* when compared to van Heerden's model (Figure 6, orange bar). Interestingly, we found that high V_{PDC}^{m} reduced *S* (increased robustness) 25- to 100-fold when compared to models with a lower V_{PDC}^{m}. Thus, the activity of PDC appears to be the most important factor in the robustness of the model. In addition, the inhibition of HK by glucose 6-phosphate also contributes to the reduction of S and thus increase in robustness.

The finding that a high V_{PDC}^{m} is critical for the overall robustness of yeast glycolysis is consistent with several observations. First, a pyruvate-decarboxylase-negative (Pdc-) *Saccharomyces cerevisiae* mutant lacking all three PDC genes (PDC1, PDC5 and PDC6) not only exhibited a three-fold lower growth rate in rich medium containing glucose than the isogenic wild-type strain, but was also unable to grow in minimal medium with glucose as the sole carbon source[15]. Second, the PDC6 gene, whose expression is either very low or absent in wild-type *S. cerevisiae*, was highly induced in the presence of excess sugars[16], suggesting that extra Pdc activity is

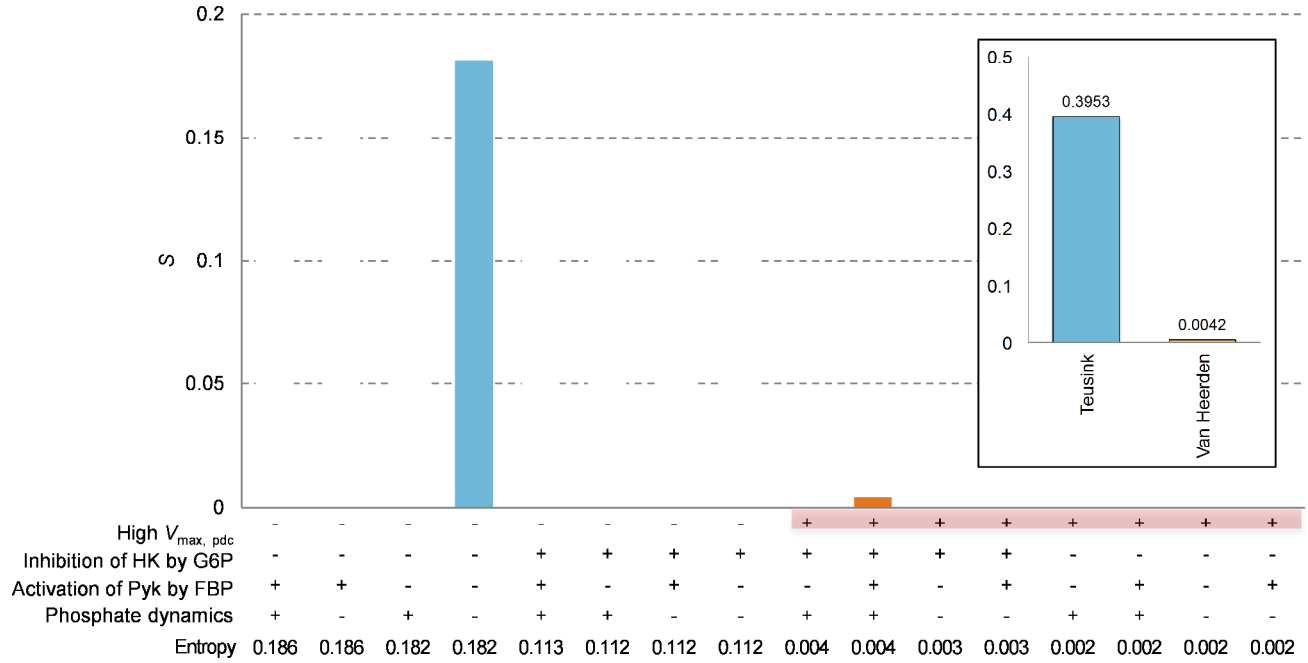

Figure 6. *S* helps identify pyruvate decarboxylase activity as a key parameter for the bifurcational robustness of yeast glycolysis. The inset shows the large difference in *S* between Teusink *et al.*'s [18] (blue bar) and van Heerden *et al.*'s[21] glycolysis models. To identify which of the first four major changes incorporated in van Heerden *et al.*'s model accounts for the robustness improvement, we constructed 16 alternative models by reversing every combination of the four changes and calculated each model's *S*. The orange bar corresponds to the original model, whereas the blue bar corresponds to the extreme case where all four changes were removed.

Integrative Biology RSCPublishing

ARTICLE

beneficial for growth under high-sugar stress conditions. In these tests, the changes to trehalose and glycogen production kinetics were not reverted because the alternative systems would seldom reach a default steady state. These results demonstrate the utility of *S* in identifying key features that are essential for the global robustness of a metabolic system.

Discussion

Robustness is an inherent property of biological systems to maintain desired function when faced with perturbations in environmental or physiological conditions. However, the intrinsic nonlinearity of metabolic systems suggests that a system can suddenly lose a stable steady state in the presence of random perturbations if a bifurcation point is crossed. Therefore, a bifurcationally robust system is needed to tolerate large changes in gene expression levels or kinetic parameters without crossing a bifurcation point. It is necessary to note that the robustness against bifurcation is different from local sensitivity[24-29], which concerns the quantitative change of system properties against small perturbations, but is equally important.

Here we develop an index for bifurcational robustness, denoted as *S*, and show that it negatively correlates with a system's robustness against bifurcation in the presence of random parameter (or enzyme) changes. Interestingly, the definition of *S* is mathematically similar to the entropy in thermodynamics and information theory[13]. As a result, *S* is also an extensive property as is the thermodynamic or information theoretic entropy. That is, the robustness of a metabolic system against random enzyme changes (*S*) is a sum of the robustness with respect to random changes in individual enzymes (S_i). This additive property gives us a tool for identifying the possible obstacles to the in vivo observability of a non-native pathway. However, other entropy properties in thermodynamics and information theory do not readily apply.

We demonstrated the utility of *S* using three examples. First, we show that *S* enables the optimization of bifurcational robustness of a yeast glycolysis model[17]. Given that experimental data used for fitting is generally sparse, adding *S* as an additional optimization objective can further reduce the uncertainty of parameter values that are otherwise loosely constrained[30,31]. Indeed, the parameter changes returned by the optimization not only improve the bifurcational robustness dramatically, but also correspond with the upgrades made to the model in a subsequent modelling effort (Supplementary Information). This first demonstration sheds light on the important role that parameter optimization considering robustness can play in building better metabolic models.

The calculation of *S* calls for a model with all parameters specified or fitted, which is not always possible. For example, non-native pathway design normally starts with a list of enzymatic reactions that constitute the pathway, but the key parameters (e.g. V_{max}'s) are strain and condition dependent and usually unknown. To address the uncertainty of kinetic parameters, we show that *S* can be readily integrated with Ensemble Modeling for Robustness Analysis[5] to enable the design of robust non-native pathways. Depending on the goal of a metabolic pathway designer, one can either investigate the distributions of S_i to identify potential sources of pathway failure, or simply use *S* to differentiate robust designs from non-robust designs (Fig. 5D). Given the versatility of *S*, we expect its combination with EMRA will bring unique value to the design of viable non-native pathways.

Finally, we show that *S* enables the comparison of a series of related models based on their bifurcational robustness. Since each model incorporates distinct features, a comparison of these models should elucidate the feature(s) that improve robustness the most. Indeed, by comparing the 16 different models of yeast glycolysis[21], we found that a high pyruvate decarboxylase activity is necessary for the robustness of this pathway. This interesting finding not only is supported by physiological data[15,16], but it may also offer an indirect explanation to the large number of PDC genes in the *S. cerevisiae* genome[15]. Even though the models we used consider only enzyme kinetics and regulations at the kinetic level, other regulatory mechanisms, such as transcriptional and post-transcriptional regulations, can also be included in the model, and the computation can proceed identically.

Methods

Models of metabolic systems with known parameters

To investigate the bifurcational robustness of metabolic systems, we examined thirteen ordinary differential equation (ODE)-based kinetic models of various metabolic systems from the BioModels Database[32]. We selected these models because: (i) they are based on rate expressions exhibiting saturation kinetics, which is essential for discussing large perturbations; (ii) the model descriptions are sufficient for simulation; and (iii) each model reaches a non-trivial steady state (also called the default steady state) when simulating with the default parameters. Lumped mass action and power-law models were excluded because they represent local behavior and do not exhibit saturation kinetics in large perturbations.

Bifurcation detection with the continuation method

Here we use a continuation method[6,7,33] as a computationally cheap and scalable alternative to study the steady-state response to parameter perturbations. In general, this method aims to find a connected path of steady-state solutions ($\mathbf{x}_{SS}$) as follows:

$$\frac{d\mathbf{x}}{\mathrm{dt}} = \mathbf{F}(\boldsymbol{x}_{SS}, \mathbf{p}) = \mathbf{0} \qquad (4)$$

Since $\mathbf{F}(\mathbf{x}_{SS},\mathbf{p})$ is always equal to zero, it follows that the total derivative of $\mathrm{F}(\mathbf{x}_{SS},\mathbf{p})$ with respect to p is also zero:

$$\frac{d\mathbf{F}(\boldsymbol{x}_{SS}, \mathbf{p})}{\mathrm{d}\mathbf{p}} = \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}_{SS}}\frac{d\boldsymbol{x}_{SS}}{d\boldsymbol{p}} + \frac{\partial \boldsymbol{F}}{\partial \boldsymbol{p}} = \mathbf{0} \qquad (5)$$

Rearranging Eq. 5 then yields Eq. 6:

$$\frac{d\boldsymbol{x}_{SS}}{d\boldsymbol{p}} = -\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{x}_{SS}}\right)^{-1}\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{p}} = \mathbf{0} \qquad (6)$$

which defines the derivatives of steady-state concentrations with respect to kinetic parameters and sets the ground for parameter continuation. Starting from the set of parameters that characterize the reference steady state, the integration of Eq. 6 can proceed in the direction where a specific parameter (*e.g.*, $V_{\max}$) is increased or decreased (Fig. 1A). The corresponding solution, which traces a trajectory in the $\mathbf{x}_{SS}$-$\mathbf{p}$ space, will then characterize how the steady state changes according to the parameter of interest. It should be noted that Eq. 6 is mathematically equivalent to the steady-state first-order sensitivity equations. Therefore, as the algorithm (*i.e.*, the differential equation solver) proceeds in the parametric domain, the (local) sensitivity profile of metabolite concentrations with respect to parameters will be updated simultaneously as the steady state moves along the $\mathbf{x}_{SS}$-$\mathbf{p}$ trajectory.

Given that Eq. 6 is ill-defined when the Jacobian matrix ($\partial\mathbf{F}/\partial\mathbf{x}_{SS}$) is not invertible, it is important to detect the point where the Jacobian matrix becomes singular. This boundary point is also known as the bifurcation point[6,7]. Therefore, the bifurcation point defines the parameter range where a system is functional with a stable steady state. In practice, the Jacobian almost always becomes badly conditioned when the system is near a bifurcation point. When this happens, we stop the solver and declare such an edge case a bifurcation point. Additionally, due to the nature of numerical integration, it is possible to "jump" over the exact point of singularity. To account for this, we always check if any of the negative eigenvalues of the Jacobian matrix becomes positive in case a bifurcation point has escaped the detection.

Calculation of S

Unless otherwise specified, parameters were assumed to vary according to a log-normal distribution with the median equal to the parameter's value at the reference steady state and a standard deviation of $\log(\mathrm{X}/\mathrm{X}_{\mathrm{original}})$ equal to .5. Such distribution has a standard deviation of .36, which is conservative compared to the variations reported in Newman et al.[3] where the average ratio of standard deviation in protein level to protein level was found to be around 1.

Parameter optimization utilizing *S*

Given that Teusink *et al.* model of yeast glycolysis[17] has a known robustness issue (Supplementary Information), we seek to modify the parameters so that the tuned model can satisfy experimental data and exhibit high robustness. To accomplish this, maximal rate constants ($\boldsymbol{V}^{\mathrm{m}}$) were subject to optimization with the following objective function:

$$\min_{\boldsymbol{V}^m}\left(\sum_i \frac{\left|V_i - V_i^{org}\right|}{V_i^{org}} + \sum_j \frac{\left|X_j - X_j^{org}\right|}{X_j^{org}} + 100\cdot S\right) \qquad (6)$$

Here, V_i represents the steady-state flux of reaction i and X_j represents the concentration of metabolite j. The superscript *org* refers to the particular values in the Teusink *et al.* model. In this study, a weighting of 100 was used to over-emphasize *S*. Optimization was performed using a simulated annealing algorithm in MATLAB (The MathWorks, Inc., USA). We allowed parameters to vary within 10-fold of their original values, and reported the results with the lowest objective value among 20 independent runs.

Calculation of *S* for models with unknown parameters

To demonstrate the utility of *S* in quantifying the robustness of non-native pathways, we constructed an ensemble of 10,000 models for each of the two configurations of a non-oxidative glycolysis (NOG) pathway as described previously[5]. Given that this is a heterologously expressed pathway, expression levels are assumed to have much higher uncertainty. Therefore, the probabilities of steady state retention under single-enzyme perturbation (p_i) were calculated assuming a log-uniform distribution between 0.1-fold and 10-fold of the default activity.

Reverting Changes to van Heerden model

The yeast glycolysis model by van Heerden *et al.*[21] is derived from the model by Teusink *et al.*[17] with five major modifications (*cf.* Results). (These have been mentioned in the text)

1. Inhibition of hexokinase (HK) by glucose 6-phosphate (G6P)
2. Consideration of phosphate as a free variable
3. Activation of pyruvate kinase (PYK) by fructose 1,6-bisphosphate (FBP)
4. A 6.1 fold increase in the $V_{\max}$ of pyruvate decarboxylase (PDC)
5. Trehalose and glycogen fluxes became functions of G6P

To study the degree by which each of these modifications improves the bifurcational robustness, we created new models by reverting the main traits of these changes. In general, if the modifications required changes to the rate laws (such as the first three modifications), we adjusted the corresponding maximal rate so as to maintain the original steady-state flux.

However, such adjustments are not always possible. The fourth modification, for example, required an increase of the maximal rate of Pdc, meaning that reverting this modification will lead to a change in the steady state, in particular the pyruvate concentration.

In this study, only the first four modifications were reverted. We did not revert the fifth modification because it causes very erratic behavior where many models don't even reach a default steady state. For the four modifications considered, we reverted all possible combinations of these changes.

Notes and references

[a] Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, 5531 Boelter Hall, Los Angeles, California 90095, USA.

[b] UCLA-DOE Institute for Genomics and Proteomics, University of California, Los Angeles, 611 Young Drive East, Los Angeles, California 90095, USA.

[#] These authors contributed equally to the work.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

1 *Kitano, H. Biological robustness. Nature reviews. Genetics **5**, 826-837 (2004).*

2 *Stelling, J., Sauer, U., Szallasi, Z., Doyle, F. J., 3rd & Doyle, J. Robustness of cellular functions. Cell **118**, 675-685, doi:10.1016/j.cell.2004.09.008 (2004).*

3 *Newman, J. R. et al. Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature **441**, 840-846, doi:10.1038/nature04785 (2006).*

4 *DeBerardinis, R. J. & Thompson, C. B. Cellular metabolism and disease: what do metabolic outliers teach us? Cell **148**, 1132-1144 (2012).*

5 *Lee, Y., Lafontaine Rivera, J. G. & Liao, J. C. Ensemble Modeling for Robustness Analysis in engineering non-native metabolic pathways. Metabolic engineering **25**, 63-71 (2014).*

6 *Chow, S. N., Hale, J. K. & Malletparet, J. Applications of Generic Bifurcation .1. Arch Ration Mech An **59**, 159-188 (1975).*

7 *Guckenheimer, J. & Holmes, P. in Nonlinear oscillations, dynamical systems, and bifurcations of vector fields 117-165 (Springer, 1983).*

8 *Chandra, F. A., Buzi, G. & Doyle, J. C. Glycolytic oscillations and limits on robust efficiency. Science (New York, N.Y.) **333**, 187-192 (2011).*

9 *Danø, S., Sørensen, P. G. & Hynne, F. Sustained oscillations in living cells. Nature **402**, 320-322 (1999).*

10 *Chung, J. D. & Stephanopoulos, G. On physiological multiplicity and population heterogeneity of biological systems. Chemical Engineering Science **51**, 1509-1521, doi:http://dx.doi.org/10.1016/0009-2509(95)00306-1 (1996).*

11 *Kitano, H. Towards a theory of biological robustness. Molecular systems biology **3**, 137, doi:10.1038/msb4100179 (2007).*

12 *Lyapunov, A. M. The general problem of motion stability. Annals of Mathematics Studies **17** (1892).*

13 *Shannon, C. E. A Mathematical Theory of Communication. The Bell System Technical Journal **27**, 379-423, 623-656 (1948).*

14 *Bogorad, I. W., Lin, T. S. & Liao, J. C. Synthetic non-oxidative glycolysis enables complete carbon conservation. Nature **502**, 693-697, doi:10.1038/nature12575 (2013).*

15 *Flikweert, M. T. et al. Pyruvate Decarboxylase: An Indispensable Enzyme for Growth of Saccharomyces cerevisiae on Glucose. Yeast **12**, 247-257 (1996).*

16 *Erasmus, D. J., van der Merwe, G. K. & van Vuuren, H. J. Genome-wide expression analyses: Metabolic adaptation of Saccharomyces cerevisiae to high sugar stress. FEMS Yeast Res **3**, 375-399 (2003).*

17 *Teusink, B. et al. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. European journal of biochemistry / FEBS **267**, 5313-5329 (2000).*

18 *Tan, Y., Lafontaine Rivera, J. G., Contador, C. A., Asenjo, J. A. & Liao, J. C. Reducing the allowable kinetic space by constructing ensemble of dynamic models with the same steady-state flux. Metabolic engineering **13**, 60-75, doi:10.1016/j.ymben.2010.11.001 (2011).*

19 *Tan, Y. & Liao, J. C. Metabolic ensemble modeling for strain engineers. Biotechnology journal **7**, 343-353, doi:10.1002/biot.201100186 (2012).*

20 *Tran, L. M., Rizk, M. L. & Liao, J. C. Ensemble modeling of metabolic networks. Biophysical journal **95**, 5606-5617, doi:10.1529/biophysj.108.135442 (2008).*

21 van Heerden, J. H. et al. Lost in Transition: Start-Up of Glycolysis Yields Subpopulations of Nongrowing Cells. Science (New York, N.Y.) **343**, 1245114 (2014).
22 Pritchard, L. & Kell, D. B. Schemes of flux control in a model of Saccharomyces cerevisiae glycolysis. European journal of biochemistry / FEBS **269**, 3894-3904 (2002).
23 Conant, G. C. & Wolfe, K. H. Increased glycolytic flux as an outcome of whole-genome duplication in yeast. Molecular systems biology **3**, 129, doi:10.1038/msb4100170 (2007).
24 Savageau, M. A. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. Nature **229**, 542-544 (1971).
25 Wang, L., Birol, I. & Hatzimanikatis, V. Metabolic control analysis under uncertainty: framework development and case studies. Biophysical journal **87**, 3750-3763, doi:10.1529/biophysj.104.048090 (2004).
26 Heinrich, R. & Rapoport, T. A. A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. European journal of biochemistry / FEBS **42**, 89-95 (1974).
27 Wang, L. & Hatzimanikatis, V. Metabolic engineering under uncertainty. I: framework development. Metabolic engineering **8**, 133-141, doi:10.1016/j.ymben.2005.11.003 (2006).
28 Wang, L. & Hatzimanikatis, V. Metabolic engineering under uncertainty--II: analysis of yeast metabolism. Metabolic engineering **8**, 142-159, doi:10.1016/j.ymben.2005.11.002 (2006).
29 Voit, E. O. Computational analysis of biochemical systems: a practical guide for biochemists and molecular biologists. (Cambridge University Press, 2000).
30 Gutenkunst, R. N. et al. Universally sloppy parameter sensitivities in systems biology models. PLoS computational biology **3**, e189 (2007).
31 Jaqaman, K. & Danuser, G. Linking data to models: data regression. Nature reviews. Molecular cell biology **7**, 813-819, doi:10.1038/nrm2030 (2006).
32 Le Novere, N. et al. BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic acids research **34**, D689-691, doi:10.1093/nar/gkj092 (2006).
33 Allgower, E. E. L. & Georg, K. Introduction to numerical continuation methods. Vol. 45 (Society for Industrial and Applied Mathematics, 2003).

