

Characterizing Criticality of Proteins by System Dynamics

Using *Escherichia coli* central carbon metabolism as a working example

Ru-Dong Li¹ and Lei Liu^{1,2*}

¹ Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

² Shanghai Center for Bioinformatics Technology, Shanghai, China

* Corresponding author

Abstract—Systems biology calls for studying system-level properties of genes and proteins rather than their individual chemical/biological properties. Up to date, most studies aiming at this goal are confined to topology-based approach. However, proteins have tertiary structures and specific functional roles, especially in metabolic systems. Thus topological properties such as connectivity, path length, etc., are not good surrogates for protein properties. In the present work, we developed a method to directly assess protein system-level properties based on system dynamics and *in silico* knockout tests. Applying the method to *E. coli* central carbon metabolic system, we found that transaldolase and transketolase-b had great impact on the system in terms of both system states and dynamical stability, while glucose-6-phosphate isomerase exerted very little influence. This finding is highly consistent with experimental characterization of metabolic essentiality. We also found that enzymes could affect a distant metabolite or enzyme even greater than a close neighbor. Our work may create a new angle for evaluating protein criticality in a system.

Keywords—Criticality; System-level property; System dynamics; Metabolic system

I. INTRODUCTION

Systems biology focuses on studying properties of molecules like genes and proteins at the system level, especially their constitutive and functional role as system components. By exploring their interplay with the overall system, we can evaluate how critical a building block is and how different parts vary in properties [1, 2]. Based on such knowledge, we can understand how a system is formed, how the system-level function is achieved and whether it can be modified according to our needs, thus enhancing researches in drug target selection, synthetic bio-system engineering, complex diseases, etc [3, 4]. *E. coli* is the best-studied organism, for which knowledge has accumulated in each of its biological hierarchies, e.g. genetic regulation, genomic information, metabolism, etc [5-7]. The central carbon metabolism contains glycolysis and pentose phosphate pathways as principal parts (supplementary file, Figure S1). It is the most common and conservative pathway among prokaryotes as well as having close resemblance in eukaryotes [8, 9]. Up to date, many genome-scale networks have been built on the organism and pathway to reveal essentiality of genes and proteins [10, 11]. However, most of such studies are based on network topology. They define system-level properties as connectivity of a molecule or shortest path

lengths, etc [12, 13]. Such properties usually have poor consistency with experimental characterization, especially on the protein level. For example, multiple studies suggest that proteins with large connectivity in a protein-protein interaction network do not correspond to essentiality. Also, many enzymes associated with large number of companions exert very little influence on cell growth rate [6, 14, 15]. We think the possible reason is that these modeling and analysis methods do not account for specific biochemical/biological functions of proteins, only purely regarding them as vertices in an abstract graph.

Under such consideration, we developed a method based on kinetic system, which can accurately reflect system dynamics and has explicit context on the biophysical/biochemical basis [5, 16]. Because *E. coli* central carbon metabolic system is the only one with comprehensive kinetic data, we used it as our model. By *in silico* knocking out an enzyme, we explored how it influenced the system state, i.e. whether state fluctuations were restricted in a limited area or spread throughout a broader range; and how large their amplitudes were. Moreover, we investigated the dynamical stability of the residual system to see whether the system maintained or lost metabolic robustness after removing the enzyme (Figure 1). From these computations, we gave characterization of protein criticality and our discoveries were consistent with published experiments. Furthermore, our methods may create a new viewpoint for protein criticality characterization.

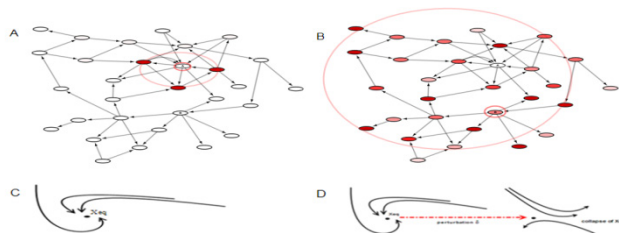


Figure 1. Schematic illustration of the protein criticality characterization method. Subfigure A and B show vertices exerting very different impacts on the system states. When a particular vertex is perturbed, the impact is within a very limited area near the epicenter (red circle in A). When another vertex is perturbed, the impact spreads throughout the network, with many distant vertices severely affected (red circle in B). Here color gradient corresponds to influence amplitude, with empty color representing zero effect. Subfigure C and D show the dynamical property of stable and unstable equilibriums. A stable equilibrium (X_{eq}) attracts its neighboring trajectories to it (C); and an unstable one repels its neighboring trajectories (D). A stable equilibrium may

collapse due to perturbations on critical system components. Vertices exerting severe impacts on system states, or cause qualitative changes in system dynamics upon perturbation are attributed with critical system-level properties.

II. RESULTS

A. System State Fluctuation

We first simulated the kinetic system to obtain metabolite kinetics and flux distributions under normal conditions. Next, we carried out *in silico* knockout of enzymes by modifying the corresponding parameters and re-simulating the system. Following the definition in previous researches, we regarded concentrations as the primary state of metabolic system [6, 17]. We calculated state deviations of the modified system from the original one and we got the fluctuation amplitude of each metabolite upon the enzyme's removal. Here we encoded them with a vector f . Second, we assessed the area of the influence by calculating the distances of metabolites from the removed enzyme and encoded them with a vector d . This allowed us to see whether the influence was within a limited radius or propagated to distant parts of the system. In short, we used a vector pair $U = (d, f)$ to represent system state fluctuation (see section METHODS for details).

We computed 3 enzymes associating to a common metabolite fructose-6-phosphate: Transketolase-b (TKb), transaldolase (TA), and glucose-6-phosphate isomerase (PGI). Since these enzymes mutually formed alternative pathways, we computed their system-level properties to examine if they shared similar characteristics or had distinct constitutive/functional roles in the system. Interestingly, we found that they had very asymmetric properties as revealed by our method. TKb had relatively large influences on many metabolites in the system compared with TA and PGI, especially for glucose, sedoheptulose-7-phosphate, and erythrose-4-phosphate (Figure 2A). TA also had influences on many metabolites, with ribulose-5-phosphate, ribose-5-phosphate, xylulose-5-phosphate, and sedoheptulose-7-phosphate having the greatest amplitudes (Figure 2B). TKb's overall influences were especially superior to TA at the distance ≥ 3 (Figure 2A and 2B), indicating that TKb could impact distant areas more strongly than TA and exert a more persistent impact with respect to distance. In contrast to them, PGI had a very little influence with none of the amplitudes being large. Its deletion did not exert as much fluctuation as TKb and TA at the distances of 1 and 2, and the influences were even negligible when distance ≥ 3 (Figure 2C). Hence, for system-level properties so far as system state fluctuation was considered, TKb and TA were much superior to PGI.

More interestingly, we found that the most severely influenced metabolites did not always concentrate in the very neighborhood of the perturbed enzyme. For example, the largest 2 impacts of TKb were at the distances of 2 and 3 but not at the distance of 1 (Figure 2A and 2D). Similar patterns were also seen from the results of TA and PGI (Figure 2B and 2C). This suggested that in contrast to the intuition that perturbation would cause largest changes to its neighborhood, enzymes could exert distant effects due to the leverage of system dynamics. We also examined how enzyme knockout influenced the system from the enzyme-centric view with our method. With each enzyme representing a reaction, we

computed flux change amplitudes and impact radius on the enzyme-centric metabolic network in the same way as stated previously. The results showed a similar pattern (supplementary Figure S2).

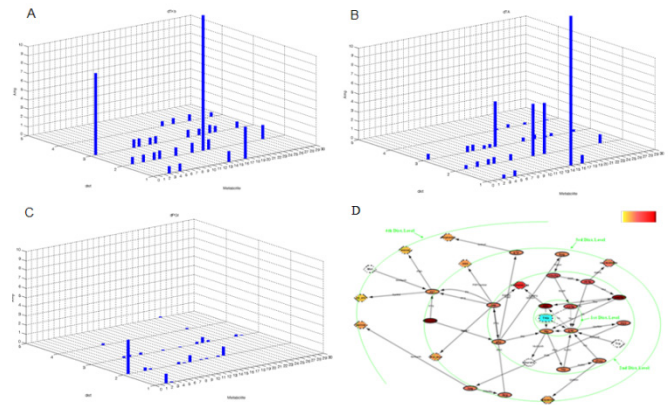


Figure 2. System state fluctuations caused by TKb, TA and PGI deletions. (A) and (B) show that TKb and TA exert relatively large influence on system states, where the amplitudes are much stronger than those of PGI. TKb has larger impact than TA at distances >3 . On the contrary, PGI exerts obviously small impact upon its deletion. The overall amplitudes are much smaller and they become especially negligible at distances >3 , indicating PGI's influence is within a limited area (Subfigure C). The results show that TKb and TA cause larger state fluctuation to the system than PGI, indicating they have different system-level properties. Here subfigure titles “dTKb, dTA and dPGI” mean “deletion of TKb, TA and PGI respectively”; x-axis is metabolite indexes, see supplementary Text S1 for the mapping of indexes to full names; y-axis is the distance of metabolite from the deleted enzyme; z-axis is the impact amplitude. Subfigure (D) is a direct demonstration of (A) in biological context. It is organized as a metabolite network with TKb highlighted in blue and metabolites arranged along green circles representing distance levels. The amplitude is proportional to color gradient (upper right corner, “red-yellow” corresponds to “strong-slight”). (A) and (D) show that 2 of the 3 most affected metabolites locate on the 2nd and 3rd distance levels and the overall impact on these levels are not inferior to that on the 1st level, exemplifying that an enzyme can affect distant metabolites even greater than its closest neighbors and TKb deletion mediates a persistent impact with respect to distances.

B. System Dynamical Stability

We found that the original system had an asymptotically stable equilibrium point X_{eq} in a large range of the parameter/state space, which made all trajectories in a wide neighborhood tending to it. This gives rise to metabolic robustness, as slight perturbations in initial values do not cause large changes in system states [18, 19]. Since this property is well in accord with Lyapunov stability, we could characterize an enzyme's criticality by examining the bifurcations of X_{eq} exerted by removal of the enzyme in mathematics. Such bifurcations include: (1) whether knocking out this enzyme made the residual system have no equilibrium; (2) if the residual system still had equilibrium, how far its location deviated from X_{eq} ; and (3) was its stability property changed (the orbit structure in its neighborhood changed). Equilibrium was computed by dynamical simulation and optimization methods. When it was located, its location deviation from X_{eq} was calculated and its neighborhood orbit structures were described by the rules of topological conjugate (see section METHODS for details).

After *in silico* knockout of TKb, the residual system had large qualitative change in dynamics. It had an equilibrium far away from X_{eq} (Figure 3A) and had very different stability property. It was an unstable one with the trajectory

representing sedoheptulose-7-phosphate kinetics being divergent and the 2 dimensions representing ribulose-5-phosphate and xylulose-5-phosphate forming a limit cycle when certain initial value held. By setting different initial values on the 2-dimensional plane of the limit cycle and investigating trajectory dynamics, it was seen that the limit cycle was an unstable one. Trajectories on the plane inside its range converged to the equilibrium's projection on the plane; and trajectories outside its range spread quickly through both dimensions (Figure 4A). Likewise, deleting TA also caused the system equilibrium to relocate to a similar distance (Figure 3B). Also, this equilibrium had similar neighborhood orbit structures to those of TKb (Figure 4B). It was also an unstable one with one dimension being divergent and another two dimensions forming an unstable limit cycle. For TA, the divergent dimension was 6-phosphogluconate and the two cycling dimensions were xylulose-5-phosphate and sedoheptulose-7-phosphate. In contrast to TA and TKb, PGI again showed a very different property. After deleting PGI, the residual system still had an equilibrium locating very near to X_{eq} (Figure 3C). Moreover, this equilibrium was also asymptotically stable, with all dimensions converging to it (Figure 4C). Therefore, the stability property of the residual system of PGI knockout was just like that of the original system, while the knockout of TKb or TA qualitatively changed the system dynamics. Hence, for system-level properties so far as dynamical stability was considered, TKb and TA were more critical than PGI.

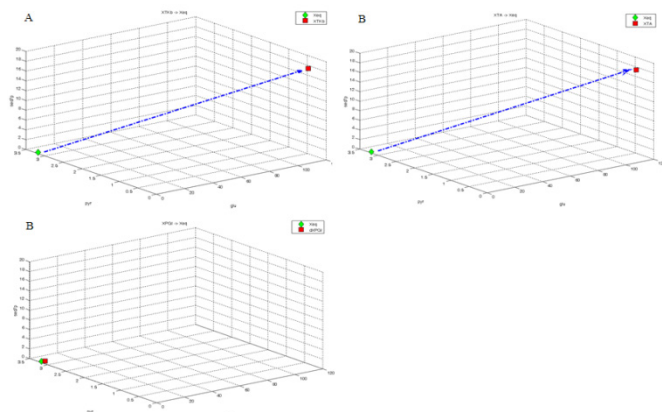


Figure 3. Equilibrium deviations caused by TKb, TA and PGI deletions. (A): The equilibrium deviation caused by TKb deletion, the 3 dimensions shown are metabolite *glu*, *pyr*, and *sed7p* respectively. They are chosen because deviations are the most significant along the 3 dimensions. The green diamond denotes X_{eq} , which is the equilibrium of the original system. The red solid represents X_{TKb} , the equilibrium of the residual system upon TKb deletion. The blue arrow marks the distance of shift. (B and C): The equilibrium deviation caused by TA, PGI deletion, respectively. Equilibriums are denoted as X_{TA} and X_{PGI} . Legends are the same as (A). The results show that TKb and TA cause obviously larger equilibrium shifts than PGI. PGI hardly shifts the original equilibrium, thus the residual system upon its deletion still equilibrates in the neighborhood of the original equilibrium. Abbreviations used here: *glu* – glucose, *pyr* – pyruvate, *sed7p* – sedoheptulose-7-phosphate.

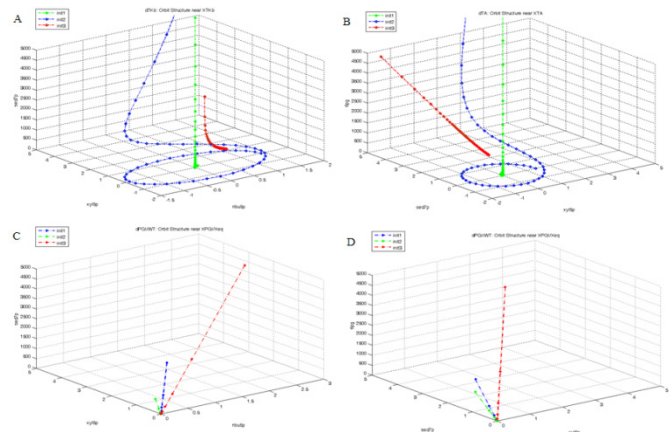


Figure 4. Orbit Structures in the neighborhood of equilibriums of the TKb-deletion, TA-deletion and PGI-deletion systems. The figures are drawn according to topology conjugate. Thus point (0,0,0) in the coordinates just represents the equilibrium, not the quantities. The figures only manifest that the curves are topologically equivalent to the orbits in the neighborhood of the equilibrium, not to be interpreted as real-value trajectories. The figures describe the qualitative system dynamics. (A): The orbit structure near X_{TKb} upon TKb deletion. The 3 dimensions are *ribu5p*, *xyl5p* and *sed7p*, which differ in dynamics from the original system. All the other dimensions exhibit similar dynamics as the original system. The dimension *sed7p* is divergent, while *ribu5p* and *xyl5p* can form an unstable limit cycle. The 3 colors represent 3 orbits with different initial values, corresponding to convergence, limit cycle and divergence on the 2D plane (*ribu5p*, *xyl5p*). (B): The orbit structure near X_{TA} upon TA deletion. The orbit structure is similar to that of TKb deletion, but the 3 dimensions differing in dynamics from the original system are *xyl5p*, *sed7p* and *6pg*. (C and D): The orbit structure near X_{PGI} upon PGI deletion is exactly the same as the orbits near X_{eq} in the original system (WT). No matter which initial value is set, the orbits eventually converge to the equilibrium. Orbits in dimensions of (*ribu5p*, *xyl5p*, *sed7p*) and (*xyl5p*, *sed7p*, *6pg*) are shown in (C) and (D) respectively. The results show that X_{TKb} and X_{TA} are unstable, while X_{PGI} is stable. Deleting TKb or TA causes qualitative changes in system dynamics while deleting PGI does not have such impact, implying TKb and TA have more critical system-level properties than PGI. Abbreviations used here: *ribu5p* – ribulose-5-phosphate, *xyl5p* – xylulose-5-phosphate, *6pg* – 6-phosphogluconate.

C. Comparison with Experimental Characterization

We compared our characterization of system-level properties with characterization of essentiality from experimental basis. Kim *et al.*'s work on *E. coli* metabolism defined a set of essential metabolites and demonstrated experimentally that if the flux-sum of an essential metabolite reduced by more than 50%, the cell growth rate would decrease by more than 50% correspondingly [6]. There were 12 metabolites in central carbon metabolism overlapping with the set of essential metabolites. We examined the flux-sums of the 12 metabolites by utilizing the simulation power of the kinetic model on perturbations, including enzyme deletion. A naive method was modifying the correspondent enzymatic parameter to zero and leaving the rest of the system as they originally were. However, the theory of Minimization of Metabolic Adjustment (MOMA) suggested that when a perturbation occurred, the system adjusted itself to some extent towards a state that was relatively close to normal [17]. Since MOMA was accepted as a rationale, we adopted it in flux simulation upon enzyme deletions, formulating the computation as an optimization problem and solving it numerically (see section METHOD for details). We found that the flux-sums of the essential metabolites were reduced much more than 50% by deleting TKb or TA (Figure 5A and 5B), thus TKb or TA deletion would result in more than 50%

reduction in cell growth rate according to Kim *et al.* On the contrary, deleting PGI did not cause any of the essential metabolite flux-sum to drop by 50% (Figure 5C), thus had relatively slight effect on cell growth. This indicated that TKb and TA had much more weight in functional essentiality than PGI, which supported our characterization of criticality.

Noteworthy, enzyme TKb is encoded by gene *tkbB*, TA is encoded by *talA*, and PGI is encoded by *pgi*. According to Baba *et al.*'s work on *E. coli* single-gene knockout mutants that measured cell growth rates upon various gene deletions, the cell growth rates of the *tkbB* and *talA* knockout mutants eventually dropped to a low level that was only about 30% (or below) of the wild type growth; while the growth rate of the *pgi* knockout mutant stayed at a level well above 50% [20]. This result was also in accord with our prediction.

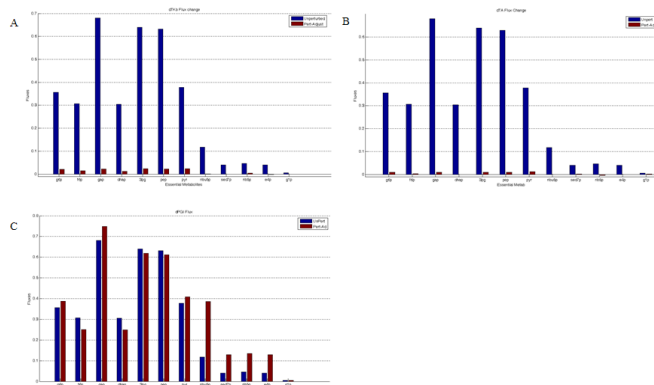


Figure 5. Sum of fluxes of essential metabolites. (A): Flux-sum values of 12 essential metabolites after TKb is deleted. The x-axis represents metabolites and y-axis represents the flux-sum values. The blue bars represent the flux-sum values in the original system; the red bars represent flux-sum values in the system with enzyme knockout. All metabolites suffer devastating flux reductions. (B): Flux-sum values of 12 essential metabolites after TA is deleted. Legends are identical to (A). All metabolites also suffer devastating flux reductions upon TA deletion. (C): Flux-sum values of 12 essential metabolites after PGI is deleted. The flux-sums of all essential metabolites are sustained above 81% of the original value. Some metabolites even have increased fluxes due to metabolic compensation effect of alternative pathways. The observations support our conclusion that TKb and TA have more critical properties over PGI in the system. Abbreviations used here: *g6p* – glucose-6-phosphate, *f6p* – fructose-6-phosphate, *gap* – glyceraldehydes-3-phosphate, *dhap* – dihydroxyacetonephosphate, *3pg* – 3-phosphoglycerate, *pep* – phosphoenolpyruvate, *rib5p* – ribose-5-phosphate, *e4p* – erythrose-4-phosphate, *glp* – glucose-1-phosphate.

III. DISCUSSIONS

Studying system-level properties of biomolecules is essential to systems biology [2, 21]. But most studies are based on network topology and not working very well at the protein level [6, 14, 15]. To overcome the drawbacks of these approaches, we propose a method based on kinetic modeling. In a kinetic system, how a component influences the system is determined by both its position and kinetic parameters. Position is equivalent to a topological property, while kinetic parameters reflect specific biochemical/biological functions. Both kinds of information are integrated in kinetic modeling and revealed by dynamical simulation [16, 22].

By using our method in *E. coli* central carbon metabolism, we have found two categories of enzymes with very distinct properties. We find that deleting TKb or TA mediates large impact on the system in terms of system state fluctuation,

while PGI knockout exerts very small influence. This indicates TKb and TA might be more critical building blocks in the system than PGI. Besides simulating state fluctuations from the metabolite-centric view, we have also evaluated how the enzymes impact flux distributions in the enzyme-centric network, where a similar pattern is observed. We also find enzymes can mediate large influence on distant metabolites or enzymes. For example, TKb can exert a large impact on glucose, whose distance from TKb is 3; similarly, TA can impact on a remote compound ribulose-5-phosphate. Although PGI has small impact amplitudes, its strongest actions are not at the distance of 1 but at the distance of 2. This is because bio-systems have complex structures consisting of branches, alternative pathways and loops, as well as various kinetic parameters differing in orders of magnitudes [1, 6, 23]. Such structure acts as a special leverage, determining special ways of interaction and effect propagation. Only kinetic modeling can reveal such behaviors. Hence, our kinetic model based method is useful in knowledge discovery. By computation, we can detect remote interacting spots as well as assessing their properties at the system level. Such analyses can give us more clues on selecting potential regulatory targets for use in drug development, metabolic engineering, etc.

Also, TKb or TA knockout mediates qualitative change in system stability while system dynamics remains almost unchanged upon PGI knockout. Bio-systems in normal conditions are subjected to biological robustness as they structurally consist of abundant branches, alternative pathways and loops [6, 18, 23]. Thus valid formula of a bio-system usually has a stable equilibrium, which can attract neighborhood trajectories to it, allowing slight changes to be tolerated without disturbing normality. Intuitively, if perturbing a component destroys a stable equilibrium and qualitatively changes system dynamics, the component should be regarded as critical. By examining equilibriums and their properties, we can assess whether the system has or losses stability. The results indicate that TKb and TA are more critical building blocks in the system than PGI.

By utilizing the power of kinetic model for approaching real-time events, we simulate fluxes after enzyme deletions. We relate the results to a previous study of functional essentiality [6]. The comparison shows that our characterization of criticality is well supported by functional essentiality. In addition, it is noteworthy that multiple metabolites (e.g. ribulose-5-phosphate, sedoheptulose-7-phosphate, etc.) in the pentose phosphate pathway have increased flux-sums (Figure 5C). This is because the cutoff of PGI gives rise to alternative pathway activation of TKb and TA, 2 backup paths for generating the essential metabolite fructose-6-phosphate. Thus the fluxes through relative reactions are compensated, resulting in local amplified flux-sums for these metabolites. This is a likely result in accordance to the MOMA theory [17]. Although MOMA can compensate system state and fluxes to some degrees, our results show that perturbations on critical building blocks such as TA and TKb cannot be smoothed by such compensation (Figure 5A and 5B). This is because the compensation effect of MOMA is mainly mediated by alternative pathways [6]. When a critical component is deleted, leaving inferior

components as backup to rely on, the system state cannot be compensated efficiently. This is the case of TA or TKb deletion. On the contrary, deleting PGI leaves its two superior alternative pathways at the “ON” state and system state can be efficiently compensated through MOMA. This gives a hint on how system-level property evaluation can help in bio-system modifications as drug development and metabolic engineering. We can delete some system components with inferior properties, and this action will result in the alternative compensation effect of their alternative pathways with superior properties. And the phenotypes of targets at some local areas will have automatic compensation due to the leverage of system structure and MOMA effect. Therefore, comprehensive methods of exploring system-level properties can help us make use of bio-complexity in engineering, as well as in knowledge discovery.

We develop our method of characterizing protein criticality based on kinetic models. But noteworthy, the application of our method is not confined to metabolic systems. For instance, we can derive gene transcription rate by the power-law formalism, the Hill equation of functional influence, or equations of chemical kinetic actions to form a dynamical model of genes and transcription factors. Or we can describe the ligand-receptor and protein-protein binding effects by the mass action law and build models for signaling networks [22]. We even do not have to obtain a set of exact parameters that can fit the modeled solutions to experimental observations. In such cases, kinetic models can still be used to analyze the generic behavioral potential of the system, e.g. in what parameter ranges will the system exhibit certain dynamics and how they change with parameters. Such qualitative predictions can also be useful in revealing general principles governing complex biological systems. And naturally, complicated bifurcation dynamics will be harder to analyze. But the idea of our method can be well applied once the coexisting dynamical characteristics in bifurcation are associated with specific biological implications, just similar to the cases in theoretical biology [24]. By integrating knowledge from different levels and using theoretical generic forms of models [16, 25], kinetic modeling will be eventually feasible for more organisms. Hence instead of the conventional topology-based approaches, we propose that complex system be studied by casting the network into kinetic equations and computing the system dynamics; and meanwhile, properties of individual system components can also be studied on this basis. Overall, our method may provide a new viewpoint in revealing constitutive/functional properties of building blocks in a biological system.

IV. METHODS

A. Kinetic Modeling

We utilized existing kinetic data in *E. coli* metabolism and adopted a published modeling framework as our working platform [5]. The model included all enzymes in glycolysis and pentose phosphate pathways. There were 24 enzymes and 6 transport/biosynthetic reactions relating to external processes. There were 18 internal metabolites and 12 external metabolites and biosynthesis products. If we considered all these reactions and metabolites, there were 30 metabolites and 30 reactions in our model (supplementary Figure S1). The

model could also be recasted into an enzyme-centric network, by adding a directed connection from enzyme A to B if any of A’s products was B’s substrate. We could explicitly see the interactions among enzymes (through metabolites) from the enzyme-centric view. A schematic representation of it was shown in supplementary Figure S3.

All kinetic rate equations were formulated according to biochemical mechanisms [5]. Most of them were casted in the uni-/bi-substrate Michaelis-Menten formalism. The kinetics for each metabolite was expressed by an ordinary differential equation (ODE, Eqn (1)).

$$\frac{dX}{dt} = A \cdot R(X, P) + B(X, P) \quad (1)$$

Here vector X denoted system state and P denoted kinetic parameters. R was a function vector collocated by all rate equations, and A was the stoichiometric matrix. B was the term standing for extra reactions (e.g. transport, metabolite utilization for cellular growth, etc). Most parameters could be found in published studies and the rest could be estimated using the experimental conditions, as well as steady-state reaction rates and concentrations reported in previous studies [5, 26, 27]. For complete descriptions of metabolites, reactions, forms of kinetic rate equations and ODEs, see supplementary Text S1.

B. Dynamical Simulation and State Fluctuation

By substituting in an initial value, a typical Cauchy problem was formed and numerical integration curves could be computed for Eqn (1). We used the Gear method in computation so as to alleviate the stiffness problem of the ODE system [28]. With an initial value for normal conditions, we obtained the dynamical states of the system X^0 , i.e. concentration kinetics of metabolites under normal conditions. After deleting an enzyme, we computed the kinetics of the residual system X^e to see how it deviated from the original state. Thus the influence of the deleted enzyme could be assessed. Assuming solution X was organized as a matrix and each column represented the kinetics of a metabolite, we could calculate the amplitude of metabolite k ’s state fluctuation as

$$f_k = \left\| X_{\cdot k}^e - X_{\cdot k}^0 \right\|_2 \quad (2)$$

We could calculate the distances of metabolites from the deleted enzyme by the metabolite-centric network. Metabolites directly associating with the enzyme were assigned a distance of 1; metabolites not directly associating with the enzyme but associating with the 1st distance level metabolites within a direct single reaction were assigned a distance of 2, and so on. We combined the distance and amplitude results to see in which ranges influences occurred and how strong they were. We also computed the flux distributions of the residual system based on the metabolites concentrations and rate equations. Thus we could observe how the flux distributions deviated from the original system and assessed them in the same way as Eqn (2). The distances of effects could be directly counted from the enzyme-centric network. Furthermore, we could combine amplitude and distance data into a single measurement for assessing the

overall impact, both for metabolite-centric network and enzyme-centric network (Eqn (3)).

$$M(d, f) = \sqrt{\sum_k f_k d_k^n}, n \in N_+ \quad (3)$$

C. Dynamical Stability

By observing the trends in system trajectories within adequately long time intervals and large parameter ranges, we could find the trajectories tending to some area if the system had equilibrium. And if it did not have equilibrium, the trajectories spread out along some dimensions running beyond several orders of magnitudes if simulation intervals were adequately long. In order to locate the equilibrium, we used the system state at the end time point of simulation as an initial guess, and utilized the trust-region method to solve the problem [29]. By carefully refining the numerical tests, equilibriums could be computed and distances from the original X_{eq} could be calculated by the Euclid norm.

We defined the dynamical stability property following Lyapunov stability, which has explicit context in physics and is suitable for describing metabolic robustness [30, 31]. The stability of an equilibrium is determined by the eigenvalues of the Jacobian matrix evaluated at the equilibrium (Eqn (4)). If all eigenvalues have negative real parts, the equilibrium is asymptotically stable; if any of them has a positive real part, the equilibrium is unstable; and if the Jacobian matrix has a pair of purely imaginary conjugate eigenvalues, a limit cycle is likely to bifurcate out of the equilibrium.

$$J_{X_{eq}} = \left[\frac{\partial(A \cdot R(X, P))}{\partial X} \right]_{X=X_{eq}} \quad (4)$$

The Hartman-Grobman Theorem and Center Manifold Theorem prove that if the Jacobian matrix evaluated at an equilibrium has 2 conjugate purely imaginary eigenvalues, N_s eigenvalues with negative real parts and N_u eigenvalues with positive real parts, the trajectories of Eqn (1) near the equilibrium are topologically equivalent to those of Eqn (5). Here β was a part of the kinetic parameters and σ was +1 according to our system. In other words, trajectories (near the equilibrium) of Eqn (5) are topological conjugates with those of Eqn (1). And because Eqn (5) is much simpler, we could investigate it instead of studying the complex Eqn (1). In this way, we explicitly drew the orbit structures of Eqn (5) near the equilibrium and could know the qualitative system dynamics of Eqn (1) accordingly.

$$\left\{ \begin{array}{l} \frac{dy_1}{dt} = \beta y_1 - y_2 + \sigma y_1 (y_1^2 + y_2^2) \\ \frac{dy_2}{dt} = y_1 + \beta y_2 + \sigma y_2 (y_1^2 + y_2^2) \\ \frac{dy^{N_s}}{dt} = -y^{N_s} \\ \frac{dy^{N_u}}{dt} = y^{N_u} \end{array} \right. \quad (5)$$

D. MOMA and Flux-sum

MOMA suggested that metabolic systems were subjected to biological robustness. When perturbed, it was able to adjust itself to a state that was relatively close to the original state. We could formulate this process as an optimization problem as in Eqn (6).

$$\begin{array}{l} \min S(P_\mu) = \|X(P_\mu) - X_0\|_2 \\ \text{s.t.} \left\{ \begin{array}{l} \frac{dX}{dt} = A \cdot R(X, P_\mu) \\ \bar{0} \leq X \\ \bar{0} \leq P_\mu \\ X(t_0 = 0) = C_0 \end{array} \right. \end{array} \quad (6)$$

Here P_μ was the parameter set with corresponding enzymatic parameter deleted, X_0 was the original state and C_0 was the initial value. A state that was closest to X_0 in the feasible space could be solved with the genetic algorithm, a heuristic numerical approach that can alleviate computation difficulty in large variable space to some extent.

We adopted the definition of essential metabolite and flux-sum in Kim *et al.*'s work on *E. coli* metabolism [6]. The 12 essential metabolites occurred in central carbon metabolism were shown in Figure 5. Here the flux-sum of metabolite k was defined as Eqn (7),

$$\Phi_k = \sum_{i \in \Omega_k} A_{ki} \cdot R_i(X, P) \quad (7)$$

where Ω_k was the index set of reactions producing metabolite k .

After MOMA computation, we obtained one (or more) set of parameters and system states. Using rate equations, we simulated the fluxes and calculate flux-sums according to Eqn (7).

ACKNOWLEDGMENT

We thank Yi-Xue Li for helpful advices about system dynamics analysis.

SUPPLEMENTARY FILES

The supplementary file includes full-sized figures 1-5, 3 additional figures S1-S3 and Text S1, which encloses descriptions of metabolites, enzymes, kinetic rate equations, ODEs and legends of Figure S1-S3. It can be found at link: http://www.biosino.org/download/rkli/additional_files2.

REFERENCES

- [1] Senger RS, Papoutsakis ET: Genome-scale model for *Clostridium acetobutylicum*: Part I. Metabolic network resolution and analysis. *Biotechnol Bioeng* 2008, 101:1036-1052.
- [2] Hood L: Systems biology: integrating technology, biology and computation. *Mech Ageing Dev* 2003, 124:9-16.
- [3] Bailey JE: Toward a science of metabolic engineering. *Science* 1991, 252:1668-1675.
- [4] Li H, Zhan M: Systematic intervention of transcription for identifying network response to disease and cellular phenotypes. *Bioinformatics* 2006, 22:96-102.

- [5] Chassagnole C, Noisommit-Rizzi N, Schmid JW, Mauch K, Reuss M: Dynamic modeling of the central carbon metabolism of *Escherichia coli*. *Biotechnol Bioeng* 2002, 79:53-73.
- [6] Kim PJ, Lee DY, Kim TY, Lee KH, Jeong H, Lee SY, Park S: Metabolite essentiality elucidates robustness of *Escherichia coli* metabolism. *Proc Natl Acad Sci U S A* 2007, 104:13638-13642.
- [7] Neidhardt FC, Curtiss R, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff W, Riley M, Umberger HE: *Escherichia coli* and *Salmonella*: Cellular and molecular biology. Washington, DC: ASM Press; 1996.
- [8] Shen-Orr S, Milo R, Mangan S, Alon U: Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 2002, 31:64-68.
- [9] Milo R, Itzkovitz S, Kashtan N, Levitt R, Shen-Orr S, Ayzenshtat I, Sheffer M, Alon U: Superfamilies of Evolved and Designed Networks. *Science* 2004, 303:1538-1542.
- [10] Ma H, Zeng A-P: Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics* 2003, 19:270-277.
- [11] Covert MW, Palsson BO: Transcriptional Regulation in Constraints-based Metabolic Models of *Escherichia coli*. *J Biol Chem* 2002, 277:28058-28064.
- [12] Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL: The large-scale organization of metabolic networks. *Nature* 2000, 407:651-654.
- [13] Vitkup D, Kharchenko P, Wagner A: Influence of metabolic network structure and function on enzyme evolution. *Genome Biology* 2006, 7:R39.
- [14] Gerdes SY, Scholle MD, Campbell JW, Balazsi G, Ravasz E, et al: Experimental Determination and System Level Analysis of Essential Genes in *Escherichia coli* MG1655. *J Bacteriol* 2003, 185:5673-5684.
- [15] Ghim C-M, Goh K-I, Kahng B: Lethality and synthetic lethality in the genome-wide metabolic network of *Escherichia coli*. *Journal of Theoretical Biology* 2005, 237:401-411.
- [16] Li RD, Li YY, Lu LY, Ren C, Li YX, Liu L: An improved kinetic model for the acetone-butanol-ethanol pathway of *Clostridium acetobutylicum* and model-based perturbation analysis. *BMC Syst Biol* 2011, 5:S12.
- [17] Segrè D, Vitkup D, Church GM: Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci U S A* 2002, 99:15112-15117.
- [18] Kitano H: Biological robustness. *Nat Rev Genet* 2004, 5:826-837.
- [19] Wagner A: Robustness and evolvability in living systems. Princeton: Princeton University Press; 2005.
- [20] Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H: Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2006, 2.
- [21] Ideker T, Galitski T, Hood L: A new approach to decoding life: Systems biology. *Annu Rev Genom Human Genet* 2001, 2:343-372.
- [22] Lee JM, Gianchandani EP, Eddy JA, Papin JA: Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Comput Biol* 2008, 4:e1000086.
- [23] Kim D, Kwon YK, Cho KH: Coupled positive and negative feedback circuits form an essential building block of cellular signaling pathways. *BioEssays* 2006, 29:85-90.
- [24] Huang S, Guo Y-P, May G, Enver T: Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. *Developmental Biology* 2007, 305:695-713.
- [25] Lee LW, Yin L, Zhu X, Ao P: Generic enzymatic rate equation under living conditions. *J Biol Syst* 2007, 15:495-514.
- [26] Bhattacharya M, Fuhrman L, Ingram A, Nickerson KW, Conway T: Single-Run Separation and Detection of Multiple Metabolic Intermediates by Anion-Exchange High-Performance Liquid Chromatography and Application to Cell Pool Extracts Prepared from *Escherichia coli*. *Analytical Biochemistry* 1995, 232:98-106.
- [27] Buziol S, Bashir I, Baumeister A, Claßen W, Noisommit-Rizzi N, Mailinger W, Reuss M: New bioreactor-coupled rapid stopped-flow sampling technique for measurements of metabolite dynamics on a subsecond time scale. *Biotechnology and Bioengineering* 2002, 80:632-636.
- [28] Hairer E, Wanner G: Solving ordinary differential equations II: Stiff and differential-algebraic problems, 2nd edn. Berlin: Springer-Verlag; 1996.
- [29] Conn AR, Gould NIM, Toint PL: Trust-region methods: Society for Industrial Mathematics; 1987.
- [30] Kuznetsov YA: Elements of applied bifurcation theory 3rd edn: Springer; 2004.
- [31] Strogatz SH: Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering: Westview Press; 2001.