SYSTEMS BIOLOGY

Kinetic modules are sources of concentration robustness in biochemical networks

Damoun Langary[1,2], Anika Küken[1,2], Zoran Nikoloski[1,2]*

Modules represent fundamental building blocks of cellular networks and are thought to facilitate robustness of phenotypes against perturbations. While reaction kinetic shapes the concentration of components and reaction rates, its use in identification of modules entails knowledge of parameter values. Here, we demonstrate that kinetic modules can be efficiently identified on the basis of steady-state reaction rate couplings in large-scale biochemical networks endowed with mass action kinetics without knowledge of parameter values. We then link the kinetic modules of metabolic networks with robustness of metabolite concentrations to perturbations. Analyzing 34 metabolic network models of 26 organisms, we demonstrate that the ordered binding enzyme mechanism leads to increased concentration robustness compared to random binding. Our findings pave the way for usage of modules in synthetic biology and biotechnological applications.

INTRODUCTION

Biochemical networks are composed of reactions through which cellular species (e.g., metabolites, proteins, and transcripts) are transformed into building blocks that support all cellular functions. The cellular functions are driven by the rates of the underlying biochemical reactions. Reaction rates are governed by specific kinetics that depends on the concentrations of substrates and effectors. A kinetic module can then be captured by couplings of reaction rates, whereby any two reactions in a kinetic module exhibit a fixed ratio of rates in a given set of states that the network obtains. This definition agrees with the intuitive notion of a module, whereby all reactions in the module are expected to respond in unison to an environmental or internal perturbation; it is also general enough to allow the consideration of spatiotemporal or steady-state behavior of reaction rates. We also note that the concept of a kinetic module differs from the concept of a cluster or community in a network (*1*), which can be determined only from the structure of the network with no relation to the dynamics of processes involving the network components.

Metabolic modeling has already contributed efficient approaches to determining steady-state coupling of reaction rates (*2*–*4*); however, these approaches uncover only a subset of steady-state couplings since they neglect the kinetics of reaction rates. Further, reaction kinetics involves multiple parameters that are challenging to obtain at a genome-scale level (*5*). Therefore, we ask whether couplings of reaction rates and the corresponding kinetic modules can be determined on the basis of the structure of the underlying networks for any values of the kinetic parameters of commonly used reaction kinetics. Despite some pioneering work (*6*), with applications to specific network types, we continue to lack general answers to these questions for arbitrary large-scale biochemical networks endowed with plausible reaction kinetics. This gap in knowledge has precluded the understanding of the extent to which kinetic modules in the biochemical networks determine robustness of concentrations of the network components, which have important implications to the functionality of these networks.

[1]Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam, Germany. [2]Bioinformatics Department, Institute of Biochemistry and Biology, University of Potsdam, 14476 Potsdam, Germany. *Corresponding author. Email: nikoloski@mpimp-golm.mpg.de

RESULTS

Identification of kinetic modules based on concordance modules

Biochemical networks are fully described by the structure and kinetics of the underlying reactions. The structure of a biochemical network, composed of reactions that transform one set of species into another, can be uniquely described in terms of two matrices: (i) a species-complex ($m \times n$) matrix $\mathbf{Y}$, the so-called stoichiometric map, that captures the molarity with which species enter the substrate or product side of a reaction; the substrate and product side of reactions are referred to as complexes; and (ii) a complex-reaction ($n \times r$) matrix $\mathbf{A}$ that specifies the incidence matrix of the directed graph describing the topological structure of the network, with nodes corresponding to complexes and directed edges to irreversible reactions. For instance, the network in Fig. 1A is composed of nine species that participate in 13 complexes interconnected by 14 reactions (see fig. S1 for the corresponding matrices $\mathbf{Y}$ and $\mathbf{A}$); here, B is a substrate complex to reactions R_2 and R_3 and a product complex to reactions R_1 and R_4.

The rates of biochemical reactions are described by different types of kinetics (*7*). We can write the rate of a reaction as $v(\mathbf{c}; k, \mathbf{p}) = k\psi(\mathbf{c}; \mathbf{p})$, where k denotes the (apparent) rate constant and $\psi(\mathbf{c}; \mathbf{p})$ represents, generally, a nonlinear activation function, parameterized by vector $\mathbf{p}$, that describes how the concentrations of species, $\mathbf{c}$, affect the reaction rate (section S1.1). This form allows us to consider mass action kinetics (*8*, *9*), used in modeling elementary biochemical reactions and derivation of other types of enzyme kinetics (e.g., Michaelis-Menten). With $\mathbf{K}$ denoting a reaction-complex ($r \times n$) matrix with entries corresponding to the reaction rate constants, the change in the concentration of the network species can be written as $\frac{d\mathbf{c}}{dt} = \mathbf{YA}v(\mathbf{c}; \mathbf{p}) = \mathbf{YAK}\boldsymbol{\psi}(\mathbf{c}; \mathbf{p})$ or succinctly $\frac{d\mathbf{c}}{dt} = \mathbf{YA_k}\psi(\mathbf{c}; \mathbf{p})$ (*10*) (see fig. S1). At steady state, it then holds that $\mathbf{YA_k}\psi(\mathbf{c}; \mathbf{p}) = \mathbf{YA}v(\mathbf{c}; \mathbf{p}) = \mathbf{Nv} = \mathbf{0}$, with $\mathbf{N}$, $\mathbf{N} = \mathbf{YA}$, denoting the stoichiometric matrix.

Under mass action kinetics, any two reactions sharing the same substrate complex have coupled rates, since their ratio corresponds to the ratio of the rate constants (Fig. 1B). For instance, reactions R_2 and R_3 in the network in Fig. 1A share the substrate complex B and, as a result, are coupled. The coupling of the rates of two reactions with distinct substrate complexes C_i and C_j is then equivalent to the

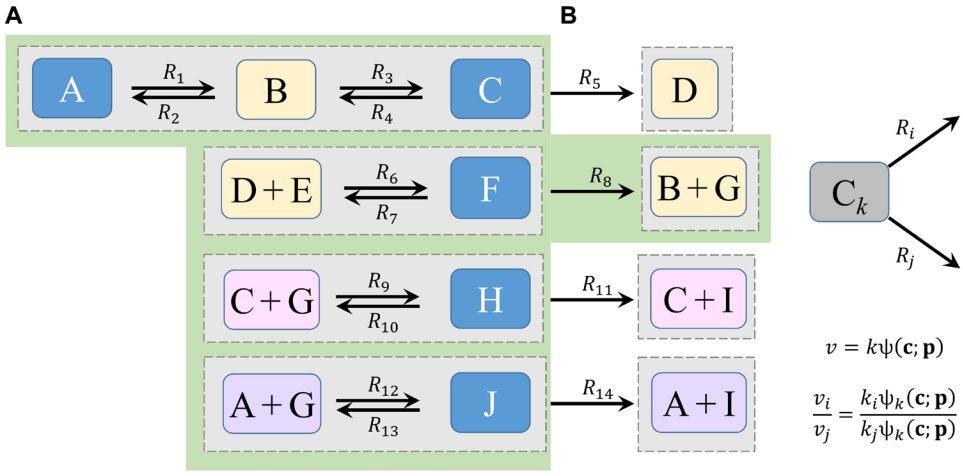

Fig. 1. Illustration of concordance and kinetic modules. (**A**) The toy reaction network is composed of nine species, A to I, and 13 complexes (marked with rounded squares) connected by 14 reactions, R_1 to R_{14}. The four concordance modules are marked with the four colors of the complexes. The balanced complexes are marked in blue. The strong linkage classes are delineated by gray boxes with dashed lines. The complexes that are included in the largest kinetic module are marked in green. The network represents a model of the EnvZ-OmpR system (*25*). (**B**) Any two reactions, R_i and R_j, that share the same complex, here C_k, have coupled rates.

coupling of their corresponding activation functions $\psi_i(\mathbf{c};\mathbf{p})$ and $\psi_j(\mathbf{c};\mathbf{p})$. We say that two complexes in a given network are coupled if their activation functions have the same ratio in all steady states the network obtains. From this definition, it follows that the complex coupling is a reflexive, symmetric, and transitive relation (sections S1.2 and S3). Therefore, complex coupling is an equivalence relation that partitions the set of complexes in a biochemical network into kinetic modules.

Identifying kinetic modules in a given biochemical network is nontrivial, even at steady state with specified parameter values, since it entails specifying the relationship between reaction rates over a set of steady states the network obtains. Here, we rely on the concept of concordance modules (*11*) to develop a general approach for identifying kinetic modules at steady state for any values the kinetic parameters of the used reaction kinetics may take. We say that two complexes C_i and $C_j$belong to the same concordance module if their activities, given by $\mathbf{A}^{i\cdot}\mathbf{v}$ and $\mathbf{A}^{j\cdot}\mathbf{v}$, are proportional, i.e., there exists a constant γ_{ij}, such that $\mathbf{A}^{i\cdot}\mathbf{v} - \gamma_{ij}\,\mathbf{A}^{j\cdot}\mathbf{v} = \mathbf{0}$ for any steady-state flux distribution $\mathbf{v}$ that the network obtains irrespective of reaction kinetics, i.e., when $\mathbf{v}$ satisfies $\mathrm{Y}\mathrm{A}\mathrm{v} = \mathrm{N}\mathrm{v} = \mathbf{0}$. Note that the activity of complex C_i, $\mathbf{A}^{i\cdot}\mathbf{v}$, differs from the activation function of complex C_i, given by $\psi_i(\mathbf{c};\mathbf{p})$. Complexes of zero activity at any steady states—irrespective of kinetics—form the concordance module of balanced complexes (*11*–*13*). For instance, the network on Fig. 1A comprises four concordance modules, depicted in different colors, of which one is composed of balanced complexes, shown in blue. The concordance modules in a given biochemical network can be efficiently determined by solving a set of linear programs (*11*). These modules hold for any reaction kinetics and emerge because of the interplay between the network structure and steady-state functionality of the network.

In addition, on the basis of the directed graph specified by $\mathbf{A}$, the complexes in a biochemical network can be partitioned into strongly connected components, corresponding to the maximal subgraphs in which there is a path from every node (i.e., complex) to any other node in the subgraph. For instance, each strongly connected component in Fig. 1A is depicted by a gray rectangle surrounding the respective complexes. The resulting strongly connected components can be further divided into nonterminal and terminal: While a nonterminal connected component contains at least one complex connected via an outgoing reaction to another component, a terminal connected component does not contain these complexes (Fig. 1A). If a complex belongs to a (non)terminal component, then it will be referred to as (non)terminal. For instance, the complexes D, B + G, C + E, and A + E in the network in Fig. 1A are terminal, while all remaining are nonterminal. Further, a subset of complexes is called autonomous if all incoming reactions are from complexes inside the set. Hence, all nonterminal complexes in Fig. 1 form an autonomous set.

Equipped with these concepts, we define a kinetic module in a biochemical network as a maximal set of mutually coupled complexes with respect to the coupling relation derived from the concordance structure in the network and the stoichiometric flux couplings. We proved that complexes in any autonomous set of nonterminal complexes that are either balanced or belong to any single concordance module form a kinetic module for any values of the kinetic parameters (section S2.3). Given a partition of a network endowed with mass action kinetics into kinetic modules, we also established sufficient conditions for identifying additional couplings between complexes. These additional couplings facilitate merging of kinetic modules and coarsening of the partition of the network into kinetic modules (sections S3.2 and S4.1). Applying these tools to the network in Fig. 1A reveals that all nonterminal complexes along with one terminal complex form a kinetic module, marked in green (more details in section S5.2). We note that our methodology does not guarantee the identification of the coarsest partition of kinetic modules.

Since all these results rely on a combination of efficient graph-theoretic and linear programming approaches, kinetic modules and consequences thereof can be identified and examined in large-scale biochemical networks without knowledge of values for the kinetic parameters. We note that the developed tools facilitate the identification of previously unknown coupling patterns arising from kinetic modules and the potential merging of existing coupling sets, both of which enable the coarsening of identified modules. As

illustrated below, our results demonstrate the presence of large kinetic modules in metabolic networks operating under steady-state constraints, whose reactions have been decomposed into elementary steps using different enzyme kinetic mechanisms allowing the use of mass action kinetics. In addition, we identify network structural conditions under which a kinetic module ensures concentration robustness of specific species (section S3.3), thus demonstrating crucial implications that kinetic modules have for network functionality.

Kinetic modules in large-scale metabolic networks

We note that the developed tools hold for networks endowed with mass action kinetics, precluding their direct applicability to metabolic networks, whose reactions follow different kinetics (e.g., Michaelis-Menten) that depend on the underlying enzyme mechanism. To render the tools applicable to metabolic networks, we considered two types of enzyme kinetic mechanisms, used in metabolic network modeling—ordered kinetic mechanism (*7*, *14*) and random kinetic mechanism (*7*, *15*). To this end, we decomposed every metabolic reaction to its elementary reaction steps, according to the used mechanism, facilitating the use of mass action kinetics (see Materials and Methods). We refer to the resulting models as ordered and random binding, respectively. As a result, we developed ordered and random binding model variants using 34 large-scale metabolic models for 26 species. This recasting substantially increased the size of the investigated models, with random binding models being, on average, twice the size of the ordered binding models. For instance, the ordered binding model of *Trypanosoma cruzi* (ilS312 Trypomastigote) consists of 1046 components, representing metabolites, enzymes, and metabolite-enzyme complexes, and 1339 elementary reactions. For comparison, the random binding model consists of 2185 complexes and 3017 elementary reactions (table S1). The largest number of components was found for the models of *Escherichia coli* (iAF1260b) including 14,973 components and 19,249 elementary reactions in the ordered binding model and 34,801 components and 51,969 elementary reactions in the case of random binding.

On the basis of the sizes of commonly studied metabolic pathways, we expected that the metabolic networks can be decomposed into uniform kinetic modules in terms of size. In contrast, we found that most of the analyzed genome-scale metabolic networks included a giant kinetic module, defined as the module of largest size (table S1). The giant kinetic module comprised a substantially larger fraction of complexes in the ordered binding models in comparison to the random binding models, and the remaining kinetic modules were either composed of three, two, or singleton complexes. For instance, the giant kinetic module in the ordered binding model of *Arabidopsis thaliana* (AraCore) and *Synechocystis* sp. (iSynCJ816) contained the minimum (0.77%) and maximum (29.37%) fractions of all complexes, respectively. In contrast, the giant kinetic module in the random binding models was the smallest for *Chlorella variabilis* (iAJ526) and the largest for *Bacillus subtilis* (iYO844), containing 0.13 and 1.4% of all complexes, respectively (Fig. 2). The size of the giant kinetic module scales with the number of components and reactions in both the ordered and the random binding models (fig. S2).

Next, we determined the percentage of reactions that have their substrate complexes in the giant module, containing at least 10 complexes. For the ordered binding models, on average, 14.15% of the reactions have their substrate complexes in the giant module. The model of *A. thaliana* (AraCore) contained the smallest fraction (0.9%) of these reactions, while the largest fraction (30%) was found in the model of *Synechocystis* sp. (iSynCJ816) (fig. S3). The results for the ordered binding models indicate their large capacity for coordinated steady-state response due to environmental changes. In contrast, for the random binding models, we found that this fraction of reactions with substrate complexes in the giant module was, expectedly, smaller, with the minimum and maximum values attained for the models of *C. variabilis* (iAJ526; 0.08%) and *B. subtilis* (iYO844; 1.32%).

To compare these findings with the modules based on the steady-state reaction couplings irrespective of their kinetics (*2*–*4*), we next determined the substrate complexes of the stoichiometrically coupled reactions in the ordered binding models. These substrate complexes are mutually coupled regardless of the assumed kinetics of the elementary reactions. We found that the giant kinetic module was, on average, 58-fold larger than the biggest group of coupled complexes found from stoichiometric reaction coupling in the ordered and random binding models, respectively (fig. S4). As a result, the proposed tools reveal an unexpectedly larger capacity for coordinating reaction fluxes in genome-scale metabolic networks across different organisms.

Further, it has been observed that different organisms exhibit remarkable robustness of growth against gene knockouts, explained by the existence of alternative pathways (*16*–*18*). In line with the observation of this remarkable robustness, it is expected that essential genes, whose deletion abolishes growth, are depleted in the identified giant modules in the metabolic models. To test the hypothesis, we identified the essential reactions with substrate complexes inside the giant module. Essential reactions were determined by the capacity to abolish growth upon their blocking using flux balance analysis (*19*, *20*). We found that the giant module was depleted for essential reactions in the ordered binding kinetic model of *E. coli*'s metabolism (*14*) ($P = 0.002$, Fisher's exact test). Therefore, we concluded that the presence of a giant module is in line with observations about robustness of growth in *E. coli* and motivates analysis of how this robustness extends on the level of metabolite concentrations.

Concentration robustness from kinetic modules in metabolic networks

Next, we investigated the implications of kinetic modules on the concentration of metabolites. In a biochemical network endowed with mass action kinetics, if two coupled complexes C_i and C_j differ only in a single molecule of a metabolite *S*, then the concentration of *S* exhibits absolute concentration robustness (*6*). For instance, in the network shown in Fig. 1A, complexes C and C + G, which belong to the same kinetic module, differ in a single molecule of G, which then exhibits absolute concentration robustness. Similarly, when two coupled complexes share all but two metabolites, each of which is present in one of the complexes, these metabolites exhibit absolute concentration ratio robustness (*21*). For instance, in the network in Fig. 1A, complexes C + G and A + G share the metabolite G and differ in metabolites C and A, respectively; since these complexes belong to the same kinetic module, C and A show absolute concentration ratio robustness. Therefore, studying the structural network property given by the differences between complexes within a kinetic module allows us to draw important conclusions about concentration robustness in arbitrary networks endowed with mass action kinetics. Identifying metabolites endowed with these concentration robustness properties was previously possible only for classes of networks with specific properties (*6*) or required symbolic computation (*22*)—both not applicable for large-scale networks.

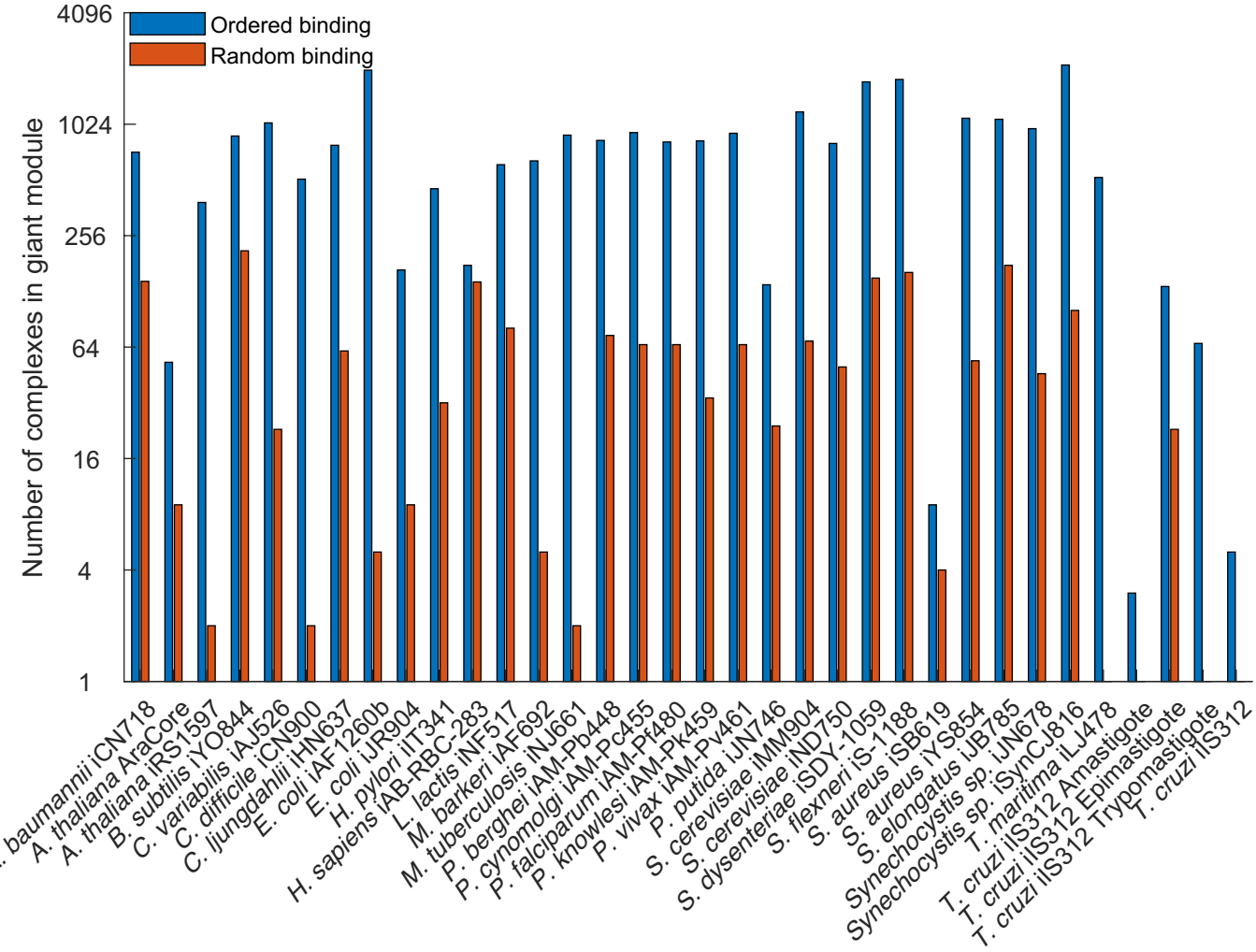

Fig. 2. Giant modules in metabolic networks of different organisms. Comparison of the number of complexes that belong to the giant module identified from kinetic coupling for models of ordered binding (blue bars) or random binding (red bars) in 34 metabolic networks of 26 organisms. See also table S1 column K.

Applying our tools to a large-scale ordered binding kinetic model of *E. coli* (*14*) with mass action kinetics revealed 26 metabolites with absolute concentration robustness, including crucial cofactors, such as adenosine triphosphate (table S2). This is in line with the experimental evidence, demonstrating that adenosine triphosphate exhibits absolute concentration robustness across different growth rates in *E. coli* (*23*). This finding also suggests that maintaining robustness of concentrations of metabolites involved in energy metabolism and used as cofactors ensures functionality of the multitude of reactions that depend on these molecules.

With the identified kinetic modules based on the developed tools in the metabolic models, we asked whether and to what extent enzyme kinetic mechanisms embedded in large-scale networks affect the capacity for absolute concentration (ratio) robustness in selected metabolites. In the case of ordered binding, we found metabolites with absolute concentration robustness in 62% of the analyzed models. The number of metabolites with absolute concentration robustness ranges from one for models of *A. thaliana* (iRS1597), *Homo sapiens* (iAB_RBC_283), *Clostridium difficile* (iCN900), *E. coli* (iAF1260b), *T. cruzi* (iIS312), and *Saccharomyces cerevisiae* (iND750), respectively, to 12 in the model of *C. variabilis* (iAJ526), 18 in the models of *Shigella dysenteriae* (iSDY_1059) and *Synechococcus elongatus* (iJB785), 19 in the model of *Shigella flexneri* (iS_1188), and up to 148 in the model of *Synechocystis* sp. (iSynCJ816) (Fig. 3A). For instance, in the genome-scale model of *C. variabilis*, *Pseudomonas putida*, *S. dysenteriae*, *Lactococcus lactis*, *S. elongatus*, *Synechocystis* sp., and *Thermotoga maritima*, metabolites, such as protons, adenosine diphosphate, adenosine triphosphate, and nicotinamide adenine dinucleotide (phosphate), were found to exhibit absolute concentration robustness (table S3). These metabolites have been suspected to exhibit buffered concentrations against perturbations based on their participation in multiple reactions (*24*). In addition, models of *Synechocystis* sp., *S. dysenteriae* and *S. flexneri*, *Plasmodium knowlesi*, *A. thaliana*, and *T. maritima* showed absolute concentration robustness for the acetyl carrier protein or coenzyme A, while the metabolic models of *Plasmodium cynomolgi*, *S. dysenteriae*, *S. flexneri*, *Synechocystis* sp., and *T. maritima* exhibited absolute concentration robustness for carbon dioxide. Metabolite pairs with absolute concentration ratio robustness were identified in all models of ordered binding that included a giant module. The number of metabolite pairs ranged from

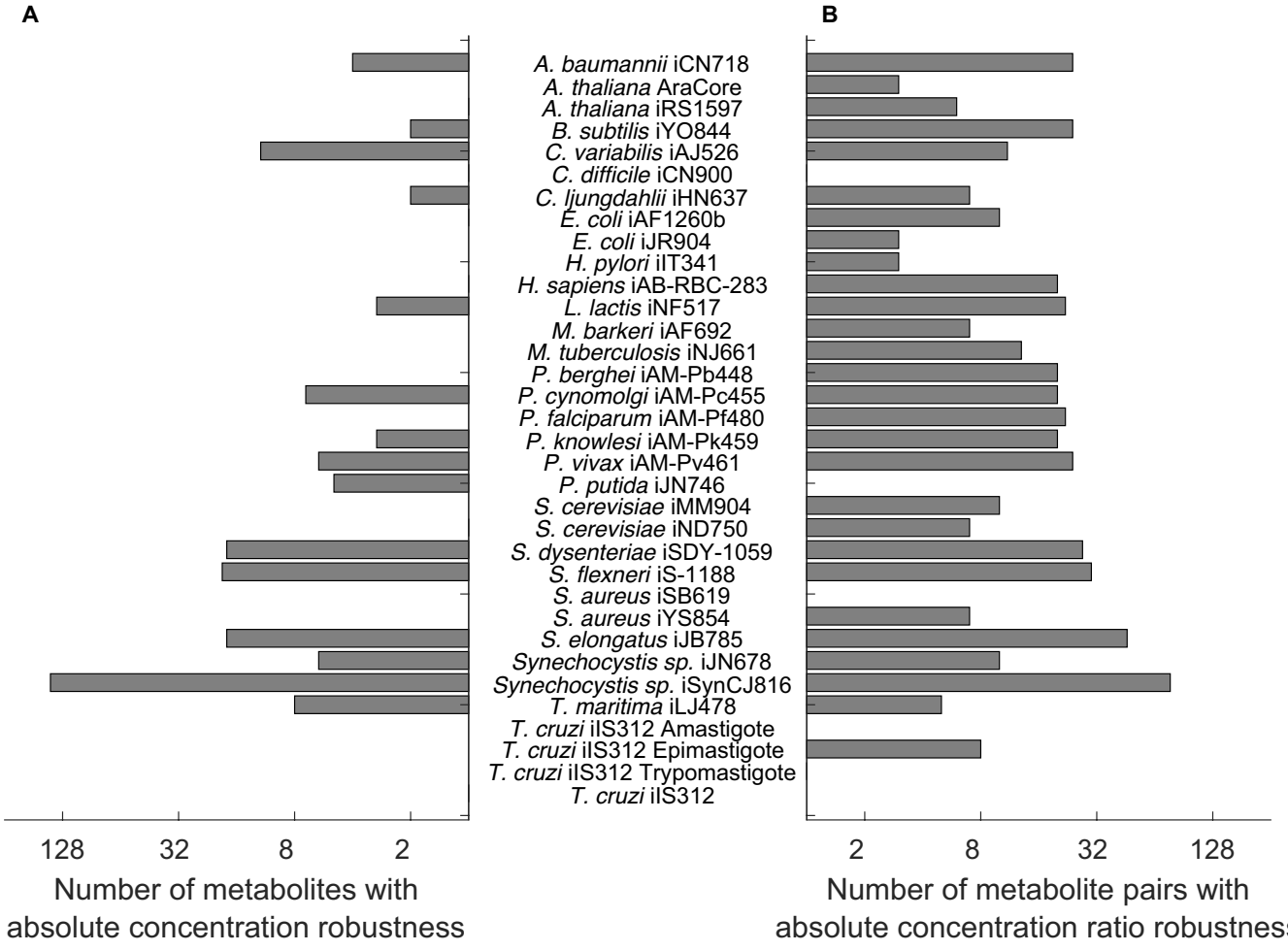

Fig. 3. Metabolites with absolute concentration (ratio) robustness. (**A**) Number of metabolites with absolute concentration robustness. (**B**) Number of metabolite pairs with absolute concentration ratio robustness in the 34 analyzed ordered binding metabolic models.

one in the models of *C. difficile* (iCN900) and *T. cruzi* (iIS312_Trypomastigote) to 46 and 77 in the metabolic models of *S. elongatus* (iJB785) and *Synechocystis* sp. (iSynCJ816), respectively (Fig. 3B).

Concentration robustness and kinetic modules in other biochemical networks

The developed tools, when applied to biochemical networks endowed with mass action kinetics, enable the detection of single and pairs of components exhibiting concentration robustness properties. For instance, the network in Fig. 1A corresponds to a model of the EnvZ-OmpR system (*25*) for which previous tools stay silent concerning the ability to show the concentration robustness of the free (phosphorylated) response regulator OmpR (*6*). The ability to identify kinetic modules in signaling networks broadens the applicability of the developed tools.

DISCUSSION

Modules serve as fundamental building blocks of real-world biochemical networks. However, we are lacking the means to decompose these networks into modules while accounting for functional and operational constraints for plausible reaction kinetics. In addition, the implications that network modules have on the robustness and plasticity of concentration of species have not been addressed, although modules provide a stepping stone in understanding how biochemical networks evolve on the basis of duplication (*26*, *27*). The existing approaches to tackling these problems are computationally prohibitive.

Here, we have developed the tools to address the identification of kinetic modules in biochemical networks of arbitrary structure and size. We note that in contrast to communities and clusters, which can be determined solely on the basis of network structure, kinetic modules consider both the structure and the operational constraints of the network (e.g., steady state and reaction reversibility). Therefore, kinetic modules provide an innovative means to analyze large-scale metabolic networks. Our key finding is that kinetic modules can be efficiently identified from the network structure while considering realistic physicochemical constraints, such as stationarity. Therefore, our results provide a general framework to explore the impact of network modifications on concentration robustness and plasticity of network species in different steady states the network obtains. Robustness of concentration of network components can ensure a nominal level of functionality in case of environmental perturbations, promoting survival, growth, and fitness (*28*). The predictions about robustness of concentrations can serve as key hypotheses

to be tested by quantitative metabolomics technologies in future dedicated experiments.

Given that our methodology does not guarantee the identification of the coarsest partition of kinetic modules, our findings also raise a number of questions, the answers to which could further deepen our understanding of how different biochemical regulation types, modifying the network structure or reaction kinetics, may affect the relationships between concentrations of network species and affect cellular functions. Together, our results indicate that the proposed modular decompositions of biochemical networks open new avenues for biotechnological and synthetic biology applications.

MATERIALS AND METHODS

The analyzed metabolic network models were obtained from the BiGG database of stoichiometric genome-scale metabolic models (*29*) or their respective original publication (for a full list of model references see the overview table available at https://github.com/ankueken/Upstream_Algorithm. The names of the models used can be found in the table S1. Balanced complexes and concordance modules were determined for each model by solving linear programming problems as described in previous publications (*11*, *12*). The identification of concordant complexes is the time-limiting step in identifying kinetic modules. This step scales quadratically with the number of complexes in the model and with the complexity of solving the linear programming problem underlying the determination of the concordance relation. To reduce the runtime, we investigate ratios of activities of complexes in sampled steady-state flux distributions. Complex pairs that vary in their ratio of activities are, in turn, excluded from the set of candidates checked for concordance relation.

Models of ordered and random binding

Models were preprocessed by removal of blocked reactions. Further, all enzyme-catalyzed reactions were split into their elementary reaction steps. Hereby, we distinguished two types of reaction mechanism that consider either (i) a fixed order of substrates binding the enzyme (*7*, *14*) and (ii) random binding of substrates, where all possible orders of substrates binding the enzyme are considered (*7*, *15*).

For instance, given a reaction $A + B \rightarrow C$ that is catalyzed by enzyme E, in the case of ordered binding, with metabolite A binding first, the reaction mechanisms is represented by the following elementary reactions: $A + E \leftrightarrow AE$, $AE + B \leftrightarrow ABE$, and $ABE \rightarrow C + E$; further, in the case of random binding, reactions $A + E \leftrightarrow AE$ and $B + E \leftrightarrow BE$, as well as $AE + B \leftrightarrow ABE$, $BE + A \leftrightarrow ABE$, and $ABE \rightarrow C + E$, are also considered to provide a representation of this mechanism. We note that reversible reactions were split into two irreversible reactions.

Data access

The implementation of the Upstream Algorithm, as well as functions to find balanced and concordant complexes, along with the analyzed models is available at https://doi.org/10.5061/dryad.7pvmcvf4v or https://github.com/ankueken/Upstream_Algorithm.

Supplementary Materials

The PDF file includes:
Figs. S1 to S11
Legends for tables S1 to S3
Supplementary Text
References

Other Supplementary Material for this manuscript includes the following:
Tables S1 to S3

REFERENCES AND NOTES

1. M. E. J. Newman, M. Girvan, Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).
2. A. Larhlimi, L. David, J. Selbig, A. Bockmayr, F2C2: A fast tool for the computation of flux coupling in genome-scale metabolic networks. *BMC Bioinformatics* **13**, 57 (2012).
3. M. Tefagh, S. P. Boyd, Quantitative flux coupling analysis. *J. Math. Biol.* **78**, 1459–1484 (2019).
4. A. P. Burgard, E. V. Nikolaev, C. H. Schilling, C. D. Maranas, Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Res.* **14**, 301–312 (2004).
5. B. O. Palsson, J. T. Yurkovich, Is the kinetome conserved? *Mol. Syst. Biol.* **18**, e10782 (2022).
6. G. Shinar, M. Feinberg, Structural sources of robustness in biochemical reaction networks. *Science* **327**, 1389–1391 (2010).
7. I. H. Segel, *Enzyme Kinetics: Behaviour and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems* (Wiley, ed. 1, 1993).
8. C. M. Guldberg, P. Waage, Ueber die chemische Affinität. § 1. Einleitung. *J. Prakt. Chem.* **19**, 69–114 (1879).
9. P. Waage, C. M. Guldberg, "Studier over Affiniteten [Studies about affinities]," in *Forhandlinger I Videnskabs-Selskabet I Christiania [Transaction of the Scientific Society in Christiania]* (Trykt hos Brøgger & Christie, 1864), pp. 35–45; www.biodiversitylibrary.org/bibliography/62508.
10. M. Feinberg, Foundations of chemical reaction network theory. *Appl. Math. Sci.* **202**, 1–475 (2019).
11. A. Küken, D. Langary, Z. Nikoloski, The hidden simplicity of metabolic networks is revealed by multireaction dependencies. *Sci. Adv.* **8**, eabl6962 (2022).
12. A. Küken, P. Wendering, D. Langary, Z. Nikoloski, A structural property for reduction of biochemical networks. *Sci. Rep.* **11**, 17415 (2021).
13. D. Langary, A. Küken, Z. Nikoloski, The unraveling of balanced complexes in metabolic networks. *Sci. Rep.* **13**, 5712 (2023).
14. A. Khodayari, C. D. Maranas, A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nat. Commun.* **7**, 13806 (2016).
15. W. Liebermeister, E. Klipp, Bringing metabolic networks to life: Convenience rate law and thermodynamic constraints. *Theor. Biol. Med. Model.* **3**, 41 (2006).
16. N. Ishii, K. Nakahigashi, T. Baba, M. Robert, T. Soga, A. Kanai, T. Hirasawa, M. Naba, K. Hirai, A. Hoque, P. Y. Ho, Y. Kakazu, K. Sugawara, S. Igarashi, S. Harada, T. Masuda, N. Sugiyama, T. Togashi, M. Hasegawa, Y. Takai, K. Yugi, K. Arakawa, N. Iwata, Y. Toya, Y. Nakayama, T. Nishioka, K. Shimizu, H. Mori, M. Tomita, Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* **316**, 593–597 (2007).
17. L. M. Blank, L. Kuepfer, U. Sauer, Large-scale ^{13}C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol.* **6**, R49 (2005).
18. W. Weckwerth, M. E. Loureiro, K. Wenzel, O. Fiehn, Differential metabolic networks unravel the effects of silent plant phenotypes. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7809–7814 (2004).
19. A. Varma, B. O. Palsson, Metabolic capabilities of *Escherichia coli*: I. Synthesis of biosynthetic precursors and cofactors. *J. Theor. Biol.* **165**, 477–502 (1993).
20. J. S. Edwards, M. Covert, B. Palsson, Metabolic modelling of microbes: The flux-balance approach. *Environ. Microbiol.* **4**, 133–140 (2002).
21. J. P. Dexter, J. Gunawardena, Dimerization and bifunctionality confer robustness to the isocitrate dehydrogenase regulatory system in *Escherichia coli*. *J. Biol. Chem.* **288**, 5770–5778 (2013).
22. R. L. Karp, M. Pérez Millán, T. Dasgupta, A. Dickenstein, J. Gunawardena, Complex-linear invariants of biochemical networks. *J. Theor. Biol.* **311**, 130–138 (2012).
23. D. A. Schneider, R. L. Gourse, Relationship between growth rate and ATP concentration in *Escherichia coli*: A bioassay for available cellular ATP. *J. Biol. Chem.* **279**, 8262–8268 (2004).
24. H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A. L. Barabási, The large-scale organization of metabolic networks. *Nature* **407**, 651–654 (2000).
25. E. Batchelor, M. Goulian, Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 691–696 (2003).
26. J. B. Pereira-Leal, S. A. Teichmann, Novel specificities emerge by stepwise duplication of functional modules. *Genome Res.* **15**, 552–559 (2005).
27. S. A. Teichmann, M. M. Babu, Gene regulatory network growth by duplication. *Nat. Genet.* **36**, 492–496 (2004).
28. G. Shinar, J. D. Rabinowitz, U. Alon, Robustness in glyoxylate bypass regulation. *PLOS Comput. Biol.* **5**, e1000297 (2009).
29. Z. A. King, J. Lu, A. Dräger, P. Miller, S. Federowicz, J. A. Lerman, A. Ebrahim, B. O. Palsson, N. E. Lewis, BiGG models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **44**, D515–D522 (2016).
30. F. Horn, R. Jackson, General mass action kinetics. *Arch. Ration Mech. Anal.* **47**, 81–116 (1972).

31. M. Feinberg, "Chemical oscillations, multiple equilibria, and reaction network structure," in *Dynamics and Modelling of Reactive Systems* (Elsevier, 1980), pp. 59–130; https://linkinghub.elsevier.com/retrieve/pii/B9780126695502500087.
32. M. Feinberg, Chemical reaction network structure and the stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems. *Chem. Eng. Sci.* **42**, 2229–2268 (1987).
33. M. Feinberg, Chemical reaction network structure and the stability of complex isothermal reactors—II. Multiple steady states for networks of deficiency one. *Chem. Eng. Sci.* **43**, 1–25 (1988).
34. S. Boyd, L. Vandenberghe, *Convex Optimization* (Cambridge Univ. Press, 2013).
35. D. Davidi, E. Noor, W. Liebermeister, A. Bar-Even, A. Flamholz, K. Tummler, U. Barenholz, M. Goldenfeld, T. Shlomi, R. Milo, Global characterization of in vivo enzyme catalytic rates and their correspondence to in vitro k_{cat} measurements. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3401–3406 (2016).
36. G. E. Briggs, A further note on the kinetics of enzyme action. *Biochem. J.* **19**, 1037–1038 (1925).
37. W. Liebermeister, J. Uhlendorf, E. Klipp, Modular rate laws for enzymatic reactions: Thermodynamics, elasticities and implementation. *Bioinformatics* **26**, 1528–1534 (2010).
38. M. D. Johnston, Translated chemical reaction networks. *Bull. Math. Biol.* **76**, 1081–1116 (2014).
39. G. Strang, The fundamental theorem of linear algebra. *Am. Math. Mon.* **100**, 848–855 (1993).
40. J. Neigenfind, S. Grimbs, Z. Nikoloski, On the relation between reactions and complexes of (bio)chemical reaction networks. *J. Theor. Biol.* **317**, 359–365 (2013).
41. J. P. Dexter, T. Dasgupta, J. Gunawardena, Invariants reveal multiple forms of robustness in bifunctional enzyme systems. *Integr. Biol.* **7**, 883–894 (2015).
42. G. H. Golub, C. F. Van Loan, *Matrix Computations* (The John Hopkins Univ. Press, Ed. 4, 2013).
43. R. A. Horn, C. R. Johnson, *Matrix Analysis* (Cambridge Univ. Press, 2013); www.cambridge.org/highereducation/product/9781139020411/book.
44. P. N. Shivakumar, K. H. Chew, A sufficient condition for nonvanishing of determinants. *Proc. Am. Math. Soc.* **43**, 63 (1974).
45. J. Gunawardena, *Chemical Reaction Network Theory for In-Silico Biologists* (Harvard Univ., 2003); http://vcp.med.harvard.edu/papers/crnt.pdf.
46. S. Müller, G. Regensburger, "Generalized mass-action systems and positive solutions of polynomial equations with real and symbolic exponents (invited talk)," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Springer, 2014), vol. 8660 LNCS, pp. 302–323; http://link.springer.com/10.1007/978-3-319-10515-4_22.
47. J. Neigenfind, S. Grimbs, Z. Nikoloski, Biochemical network decomposition reveals absolute concentration robustness. arXiv:1105.0624 [q-bio.MN] (2011).

Acknowledgments

Funding: Z.N., D.L., and A.K. would like to acknowledge funding from the University of Potsdam's Research Focus Group "Evolutionary Systems Biology." **Author contributions:** Writing—original draft: D.L. and Z.N. Writing—review and editing: A.K., D.L., and Z.N. Conceptualization: D.L. and Z.N. Investigation: A.K., D.L., and Z.N. Methodology: D.L. and Z.N. Funding acquisition: Z.N. Validation: A.K. and Z.N. Supervision: Z.N. Formal analysis: A.K., D.L., and Z.N. Project administration: Z.N. Visualization A.K., D.L., and Z.N. Resources: A.K. Data curation: A.K. Software: A.K. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data and code needed to evaluate the conclusions in the paper are present in the paper, the Supplementary Materials, and at https://doi.org/10.5061/dryad.7pvmcvf4v or https://github.com/ankueken/Upstream_Algorithm.

Submitted 26 August 2024
Accepted 24 April 2025
Published 28 May 2025
10.1126/sciadv.ads7269