Review TRENDS in Biochemical Sciences Vol.31 No.5 May 2006

Systems analyses characterize integrated functions of biochemical networks

Erwin P. Gianchandani[1], David L. Brautigan[2] and Jason A. Papin[1]

[1]Department of Biomedical Engineering, University of Virginia, Box 800759, Health System, Charlottesville, VA 22908, USA
[2]Center for Cell Signaling and Department of Microbiology, University of Virginia, Box 800577, Health System, Charlottesville, VA 22908, USA

Metabolic, regulatory and signaling pathways have been characterized in detail over the past century. As the amount of genomic, proteomic and metabolic data has increased, and the mathematical and analytical capabilities of interrogating these data have advanced, the overlapping roles of pathway constituents have been described. These developments reflect the truly integrated nature of subcellular biochemical networks. Systems analyses, including the reconstruction of stoichiometric networks, provide a key set of tools for quantifying overlap among the metabolic, regulatory and signaling functions of network components. Accounting for this integration is crucial for accurately describing the function of biochemical networks.

Introduction

Our ability to generate high-throughput experimental data has led to the detailed characterization of biochemical networks. For example, mass spectrometry, a technique for identifying and quantifying known compounds in chemically complex mixtures, has been used as a high-throughput means to study metabolomics [1]. Similarly, DNA microarrays and yeast two-hybrid systems have generated detailed lists of biological components and their associated interactions in regulatory and signaling networks. Alongside this explosion in experimental techniques, much progress has been made in quantitative analyses of networks [2]. For example, the reconstruction of a stoichiometric matrix involves assimilating knowledge of what genes and associated proteins are present in a system, in addition to the stoichiometry of associated chemical transformations. Thus, this reconstruction integrates the individual components and associated interactions elucidated by screening technologies into a large-scale description of biochemical networks. It forms the basis for subsequent analysis of the fundamental characteristics of the network. Consequently, networks can be reconstructed and modeled *in silico*, and the biological function to which they give rise can be analyzed, interpreted and predicted [3]. Ultimately, genome-scale experimentation iteratively coupled with stoichiometric network reconstruction and *in silico* systems analyses can generate new hypotheses and can discover emergent properties of biological systems.

Applying systems analysis techniques to high-throughput data sets in this manner can identify the multifunctionality of networks and their components. Several multifunctional components have been characterized so far (Table 1). As summarized in Ref. [4], many glycolytic enzymes are multifunctional proteins and not simply the single-function components of the glycolytic pathway that they were originally thought to be. For example, hexokinase, which catalyzes the phosphorylation of glucose during glycolysis, is also a regulator of apoptosis [5–9]. Association of hexokinase with the mitochondrial membrane is crucial to determining its functional specificity (Figure 1). Many of these multifunctional components have crucial roles in fundamental disease processes.

Furthermore, a recent study investigating genetic pleiotropy (i.e. genes that encode proteins involved in multiple cellular processes) in the budding yeast *Saccharomyces cerevisiae* found a much higher degree of pleiotropy than would be expected simply from random variation [10]. This finding, obtained through a high-throughput experimental assay coupled with a biclustering algorithm for data analysis, offers empirical evidence in support of the presence, and importance, of pleiotropy in biological systems. Specifically, researchers produced strains containing single-gene deletions of 4700 nonessential genes, analyzed them for growth under 21 different conditions, and observed 216 strains showing significant growth defects across more than three of the environments – a number much higher than expected on the basis of random variation.

Here, we describe specifically how stoichiometric network reconstructions and associated mathematical analyses can be used to interrogate biochemical networks to identify new properties, including previously unknown overlaps of network components and functions. We present an overview of metabolic, regulatory and signaling processes, and of how these types of networks have been treated separately until recently. (Note that here we use 'regulation' to refer specifically to transcriptional regulatory networks and 'signaling' to represent intracellular reactions that drive responses to the extracellular environment.) We provide the motivation for considering

Corresponding author: Papin, J.A. (papin@virginia.edu).
Available online 17 April 2006

 0968-0004/$ - see front matter doi:10.1016/j.tibs.2006.03.007

Table 1. Examples of network overlap: multifunctional proteins and metabolites[a]

Compound	Metabolic	Regulatory	Signaling	Refs
Aconitase	Mitochondrial TCA cycle enzyme[b]	Translation regulator	Unknown	[41,42]
ATP	Principal product of energy metabolism[b]	Nucleotide required for RNA and DNA synthesis; global regulator of DNA coiling	Phosphate is ubiquitous in signaling reactions	[43]
Arg5,6 protein	Arginine biosynthesis[b]	Binds to specific nuclear and mitochondrial loci of DNA	Unknown	[44]
BID	Unknown	Regulator of apoptosis[b]; mediator of cell-cycle arrest (and possibly DNA repair)	Unknown	[45,46]
Cyclin D3	Unknown	Unknown	Mediator of cell proliferation[b]; potentiates apoptosis in the presence of Caspase 2	[47]
Enolase 1	Glycolytic enzyme[b]	Transcriptional regulator	Unknown	[4,48–50]
Glucose-6-phosphate isomerase	Glycolytic enzyme[b]	Unknown	Involved in cell motility and invasion	[51–55]
Glyceraldehyde-3-phosphate dehydrogenase	Glycolytic enzyme[b]	Transcriptional coactivator; mediator of apoptosis	Unknown	[56]
Insulin	Glucose uptake	Unknown	PI3K signaling	[57,58]
Interleukin-1	Inhibition of fatty acid synthesis	Unknown	Involved in the immune system[b]	[57]
Hexokinase	Glycolytic enzyme[b]	Global regulator of apoptosis	Unknown	[4,6]
Lactate dehydrogenase	Glycolytic enzyme[b]	Transcriptional coactivator	Unknown	[56]
Nicotinamide adenine dinucleotide	Metabolic cofactor[b]	Alters transcription factor DNA-binding properties	Ca^{2+} signaling; poly(ADP)-ribose polymerases	[56,59]
Phosphoglucose isomerase	Glycolytic enzyme[b]	Unknown	Acts as neuroleukin, ACF and DMM	[12–14]
Phosphoinositides	Lipid biosynthesis[b]	Modulate chromatin structure	PI3K signaling	[60]
PutA protein	Proline dehydrogenase[b]; pyrroline-5-carboxylate dehydrogenase	Transcriptional repressor	Unknown	[12,61,62]
Rae1 protein	Unknown	mRNA export protein[b]; involved in spindle assembly	Unknown	[17]
Riboswitches and ribozymes	Metabolite binding can regulate activity	Cleave RNA transcripts (ribozymes)[b]; control gene expression[b]	Unknown	[63]
Sialic acid	Oligosaccharide synthesis[b]	Unknown	Apoptotic signaling	[64]
Ubiquitin and ubiquitin-like proteins	Protein binding can regulate activity[b]	Transcriptional regulators; regulators of apoptosis; facilitate DNA repair	Unknown	[65]

[a]Abbreviations: ACF, autocrine motility factor; DMM, differentiation and maturation mediator; TCA, trichloroacetic acid.
[b]Initial characterizations of compounds.

the interconnectedness of these biochemical networks by highlighting examples of how multifunctional components are involved in fundamental processes linked to disease. Finally, we describe how coupling high-throughput data with stoichiometric network reconstructions and systems analysis techniques is yielding insights into biological systems, including the previously unrealized diversity of functions among biochemical network components. Here, we illustrate the utility of these emerging techniques in characterizing the pathways involved in human disease.

The need to understand integrated, overlapping functions

The interconnectivity of metabolic, regulatory and signaling processes is illustrated by the growing number of network components that show several roles across multiple pathways (Table 1). By systemically identifying and characterizing the overlapping actions of these multifunctional components, it becomes possible to capture a complete functional profile of cellular response and, ultimately, to gain a much better appreciation of fundamental disease processes [11].

Examples of overlapping actions of multifunctional components

The integrated nature of biochemical networks is exemplified by multifunctional proteins and metabolites (Table 1). Protein function can vary as a consequence of changes in cellular localization, in oligomeric state, in the cellular concentration of a ligand, substrate, cofactor or product, or in any combination of these [12]. For example, the protein PutA in *Escherichia coli* has pyrroline-5-carboxylate dehydrogenase activity when it is associated with the plasma membrane, but lacks enzymatic activity when it binds DNA as a transcriptional repressor [12–14]. Likewise, neuropilin, a surface receptor on both endothelial cells and nerve axons, detects different ligands depending on the cell type on which it is expressed [12,15]: on endothelial cells, it detects vascular endothelial growth factor and indicates when new blood cells are needed; on nerve axons, by contrast, it detects semaphorin III and helps to steer axons to their proper destinations.

Similarly, novel functions of several glycolytic enzymes have been identified on the basis of cellular localization [4]. For example, glyceraldehyde-3-phosphate

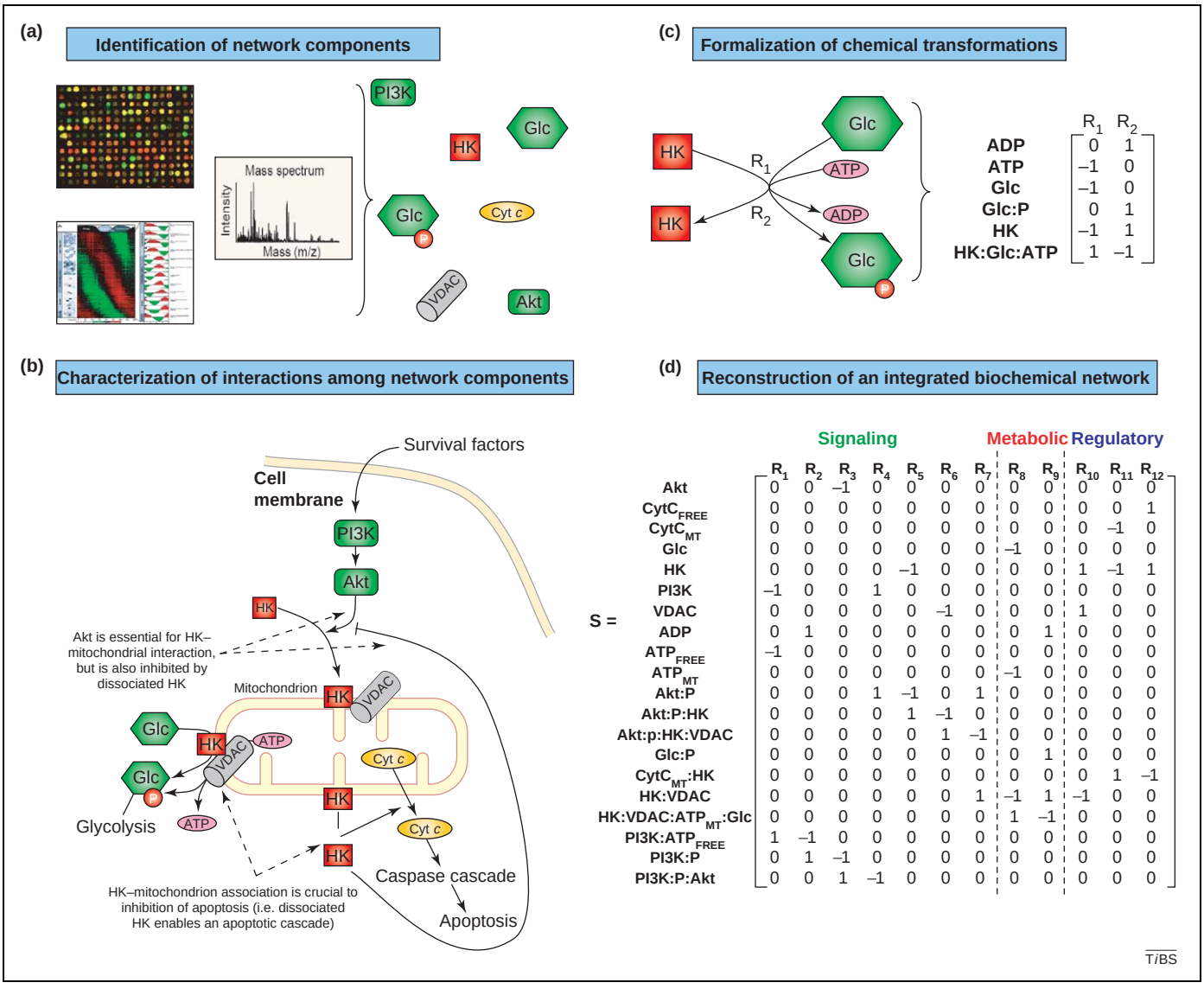

Figure 1. An example of multifunctional roles and associated systems analyses. The steps involved in the stoichiometric reconstruction of a biochemical network are illustrated by the hexokinase-driven apoptotic cascade [4–9]. Specifically, experimental data identify network components **(a)**, for example phosphatidylinositol 3-kinase (PI3K), hexokinase (HK) and glucose (Glc), and component interactions **(b)**, such as the conversion of glucose to glucose-6-phosphate by hexokinase. These components are subsequently formalized stoichiometrically and synthesized in matrix form **(c)**. The rows of the matrix correspond to the network components and the columns correspond to the chemical transformations or interactions between the components. The elements of the matrix correspond to the stoichiometric coefficients of the associated chemical transformations. A negative number indicates an 'input' to the reaction, and a positive number indicates an 'output' from the reaction. For example, the conversion of glucose to glucose-6-phosphate is represented by two reactions in the stoichiometric matrix. The first column corresponds to the association of Glc and ATP with HK. The second column corresponds to the dissociation of ADP and glucose-6-phosphate from HK. Ultimately, the stoichiometric matrix provides a mathematical formalism to account for interactions of components across various cellular processes, including metabolic, regulatory and signaling pathways, in terms of chemical transformations **(d)**.

dehydrogenase (GAPDH), the pivotal glycolytic enzyme that catalyzes the conversion of glyceraldehyde 3-phosphate (GAP) into 1,3-bisphosphoglycerate, has also been shown to initiate cell death [16]. In the presence of nitric oxide GAPDH can be *S*-nitrosylated, which augments its binding to Siah1 (an E3 ubiquitin ligase with a nuclear localization signal). Together, the bound proteins translocate to the nucleus where GAPDH stabilizes Siah1, thereby facilitating the ability of Siah1 to degrade nuclear proteins.

Association of the mRNA export protein Rae1 during different phases of the cell cycle leads to two distinct functions – namely, translational regulation and cell-cycle control. Originally identified as essential to the export of mRNA transcripts out of the nucleus, Rae1 has recently been shown to localize to mitotic spindles and around chromosomes in *Xenopus* egg extracts [17]. Furthermore, depleting the protein from these egg extracts markedly inhibits mitotic spindle assembly and leads to chromosome misalignment. Consequently, the proposed hypothesis is that Rae1 functions as a scaffold during mitotic spindle assembly, recruiting additional proteins that are necessary for proper spindle formation [17].

These examples are merely a fraction of the ever-growing set of multifunctional network components that are important in individually diverse yet systemically integrated pathways. Indeed, they are demonstrative of the breadth of function of, and degree of overlap among, network components. Moreover, because any component that functions in multiple cellular pathways effectively 'constrains' each pathway in which it is involved, a holistic approach that considers this 'network overlap' must be

used when studying biochemical networks and assessing network properties. When developing a therapeutic intervention for a specific target protein, for example, it would help to take into account component interactions and associated phenotypic effects in cellular processes to minimize or to prevent potentially adverse consequences. This 'universal accounting' is possible only with a structured framework for reconstructing and analyzing the role of a component in a network context.

Examples of functional implications of multifunctional components

When overlapping actions of network components are identified and characterized, they provide a basis for the further study of the resultant cellular response, which eventually leads to a much better understanding of fundamental disease processes [11]. Consider autophagy, for example, an intracellular process that is mechanistically linked with growth factor signaling and that supports survival in some cells but leads to death in others [18] (Figure 2). In general, intracellular metabolite supply is regulated by, first, the acquisition of extracellular nutrients; and second, the activation of catabolic metabolism, which degrades intracellular macromolecules in the absence of appropriate growth stimuli and nutrient uptake.

Specifically, growth factor signaling activates the phosphatidylinositol 3-kinase and Akt (PI3K/Akt) pathway, in addition to its downstream effectors including mammalian target of rapamycin (mTOR), and ultimately maintains the expression of nutrient transporters such as glucose transporter 1 (GLUT1). In turn, these nutrient transporters ensure proper nutrient uptake, which enables cellular metabolism and cell growth to proceed normally. Diminished growth factor signaling, however, inactivates the PI3K/Akt pathway and correspondingly triggers a loss of expression of cell-surface nutrient receptors, a reduction in nutrient uptake, and a decrease in cellular metabolism and cell growth. In some cases of diminished growth factor signaling, an autophagic response (i.e. the activation of catabolic metabolism) attempts to compensate for this loss of nutrient supply and enables organisms to survive long periods of starvation, although this compensation might well contribute to tumor resistance to chemotherapy, wherein cells that are intentionally starved (e.g. through therapeutic intervention) keep themselves alive [18]. In other cases, however, inactivation of the PI3K/Akt pathway leads to

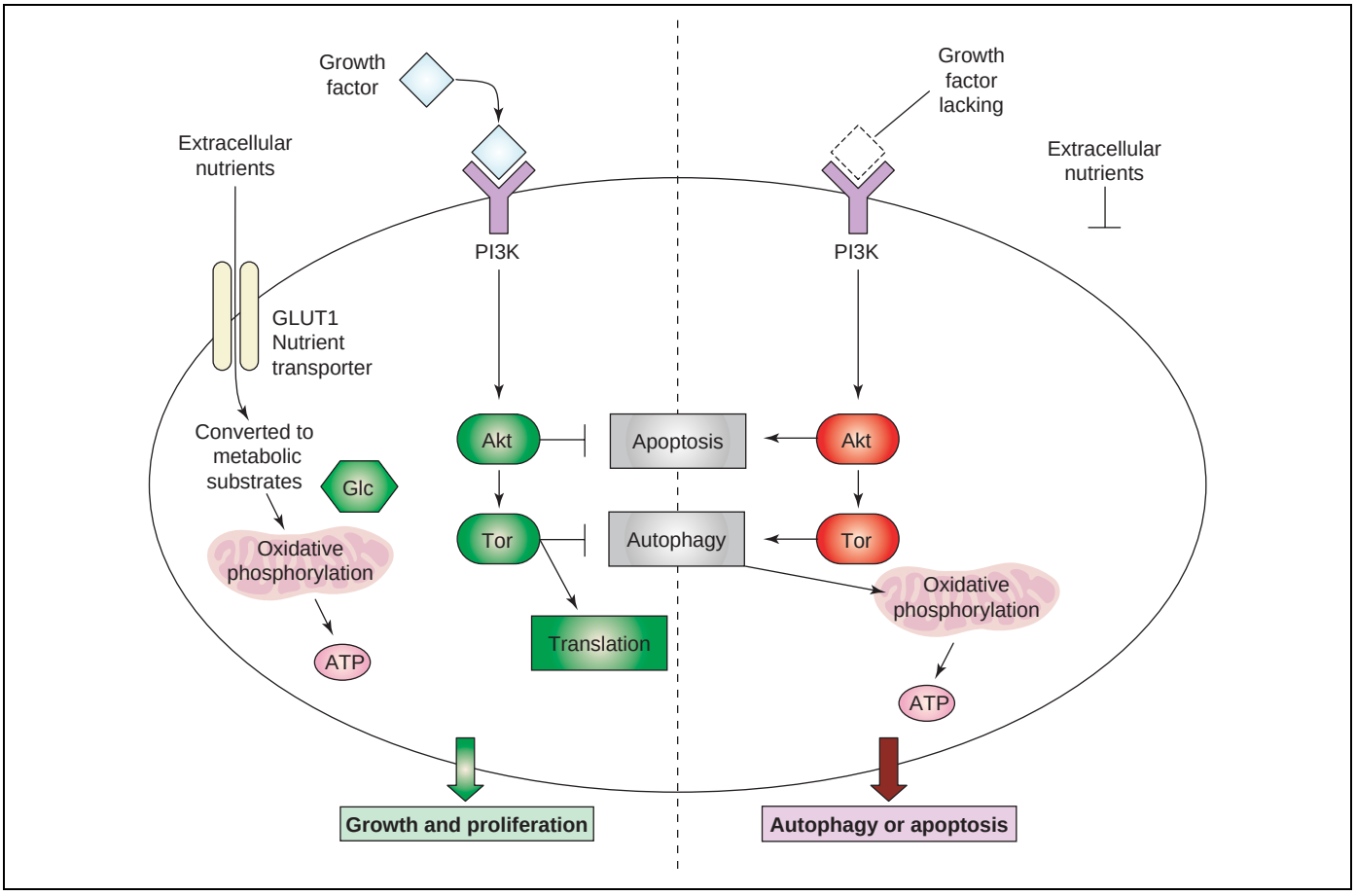

Figure 2. Autophagy as an impetus for understanding 'network overlap'. Growth factor signaling is known to activate the phosphatidylinositol 3-kinase and Akt (PI3K/Akt) pathway and its downstream effectors, including target of rapamycin (Tor), and thereby to maintain nutrient uptake, cellular metabolism and cell growth. Activation of the PI3K/Akt pathway in the presence of growth factors potentiates anabolic processes such as translation and cell growth, while suppressing catabolic processes including autophagy and apoptosis (left). Conversely, when growth factors are lacking, the inactivated PI3K/Akt pathway results in a loss of expression of cell-surface nutrient transporters, reduced nutrient uptake and decreased metabolic substrates for oxidation by the mitochondria (right). In some situations, autophagy is induced and cytoplasmic components are degraded to provide a compensatory source of metabolic substrates to support ATP production and cell survival; in others, however, apoptotic cell death occurs. Green shading denotes activation of network components, and red shading denotes inhibition. Reproduced, with permission, from Ref. [18]. © (2005) Macmillan Magazines (www.nature.com/reviews).

apoptotic cell death [18]. Understanding how the same set of autophagic proteins leads to both cell survival and cell death is therefore an active area of research [18].

Experimental evidence underscores the multifunctionality of many components of the autophagy process. For example, mTOR is a component of both growth factor signaling and nutrient-sensing pathways that supports protein synthesis when metabolic conditions are favorable and/or when it is stimulated by growth factors. Thus, as in other types of integrated biochemical pathways, elucidating the integrated functions of proteins involved in autophagy will help to characterize better this intrinsic cellular process and, eventually, will help to explain the potential therapeutic benefit of manipulating autophagy [18].

Another recent study has shown that breast cancer tumors subvert the normal wound-healing process to build a network of blood vessels [19], demonstrating how important it is to understand fully the multifunctional roles of components of a biochemical pathway. Tumors were shown to produce stromal cell derived factor 1 (SDF-1), which recruits cells that are required for growing blood vessels, similar to what happens at sites of tissue injury [19]. In particular, carcinoma-associated fibroblasts (CAFs) extracted from human breast cancers were found to show traits similar to myofibroblasts and to promote the growth of breast carcinomas significantly more than were normal mammary fibroblasts. CAFs secreted SDF-1, thereby recruiting endothelial progenitor cells into the carcinoma to enhance angiogenesis and thus tumor growth. Indeed, the tumor stroma was found to resemble the stroma present at sites of tissue injury [20]. This recent work demonstrates that carcinoma cells can use the normal host stromal response by converting recruited cells into CAFs, which in turn recruit endothelial progenitor cells to the tumor via SDF-1. As in autophagy, understanding the multifunctional roles of the components and pathways involved in angiogenesis will help to refine our understanding of how the wound-healing process is exploited by tumors and will potentially shed light on cancer treatment.

Stoichiometric network reconstructions and associated analyses

Coupled with systems analysis techniques, high-throughput experimental data provide a mechanism for satisfying our ever-growing need to characterize and to quantify network overlap. Specifically, a stoichiometric matrix representation provides a mechanism for capturing a detailed description of a biochemical network and forms the basis for further analyses of fundamental network characteristics, such as cross-talk, redundancy and modularity (see Ref. [21] for an example of network analyses of the JAK–STAT signaling network in the human B cell).

Reconstructing the stoichiometric matrix of a biochemical network

A stoichiometric reconstruction is the end product of the careful integration of data on the chemical transformations in a system with defined boundaries. The process of generating a stoichiometric reconstruction can be illustrated by an initial reconstruction of the hexokinase-driven apoptotic cascade discussed (Figure 1). Essentially, data on network components and the interactions between components are synthesized into a matrix form. The rows of the matrix correspond to the network components, whereas the columns represent the chemical transformations or interactions between the components. The elements of the matrix correspond to the stoichiometric coefficients of the associated chemical transformations. Typically, a negative number indicates an 'input' and a positive number indicates an 'output'. For example, a hypothetical reaction wherein 1 mol of A and 1 mol of B are combined to produce 1 mol of C is represented in a single column of the matrix, with the consumption of 1 mol of A denoted by -1, the consumption of 1 mol of B denoted by -1, and the production of 1 mol of C denoted by $+1$ in the corresponding rows for A, B and C, respectively.

This stoichiometric matrix provides a mathematical formalism to account for interactions of components across various cellular processes. For example, the chemical transformations involving a protein that catalyzes the conversion of a metabolic substrate into a product would be represented in columns of the stoichiometric matrix. If the same protein has a signaling function (e.g. forming a heterodimer with another protein in which the dimer now acts as a transcription factor), the same row of the matrix would have entries in multiple columns (i.e. reactions). Because the stoichiometric framework attempts to capture the flow of information in accordance with the principle of conservation of mass, it is applicable to metabolic, regulatory and signaling reactions alike, because these processes can be described as mass-balanced reactions [2,22].

Systems analyses that use the stoichiometric matrix

Several mathematical approaches are being developed that can analyze properties of these network reconstructions. For example, the 'constraint-based' approach is a set of mathematical techniques that characterize properties of a biochemical network. It is based on the principle that all possible phenotypes of a given network satisfy fundamental constraints (e.g. physical laws such as the conservation of mass and environmental factors including nutrient availability, pH and osmolarity) that are inherently imposed on the molecular functions of all cells [23]. These constraints are described mathematically through 'balances' (constraints that are associated with conserved quantities) and 'bounds' (constraints that limit the numerical ranges of individual variables) [23]. By identifying these constraints and stating them mathematically, the spectrum of possible phenotypes is narrowed, and thus understanding of the cellular network function is enhanced. These approaches have been successfully used to analyze metabolism, regulation and cellular signaling processes [21,23]. Although the mathematical details of such analysis methods are beyond the scope of this review, two general approaches to constraint-based analysis provide good examples of these techniques.

The first, called 'network-based pathway analysis', is a class of mathematical techniques that involves the

computation of pathways in a biochemical network that satisfy a set of properties [24]. Two examples of these pathways are elementary flux modes and extreme pathways. An elementary flux mode constitutes the minimal set of enzymes in a biochemical network that can operate at steady state, assuming that all irreversible reactions proceed in the appropriate direction [25]. Extreme pathways, a closely related approach [26], are mathematically derived vectors that can be used to characterize the phenotypic potential of a defined biochemical network [27]. Elementary modes and extreme pathways have broad applications in biochemical networks [28]. For example, elementary mode analysis has been used to predict the growth behavior of wild-type and mutant strains of *E. coli* [29]. Another application of elementary mode analysis to the metabolic network of *E. coli* has led to the identification of the most efficient pathway for the production of green fluorescent protein [30]. Extreme pathway analysis has been carried out on a genomic scale for amino acid production in *Haemophilus influenzae* [31] and protein production in *Helicobacter pylori* [32] to quantify network redundancy, and has been recently applied to the JAK–STAT signaling network of the human B cell [21].

The second, termed 'flux–balance analysis', is a set of techniques that uses the stoichiometric matrix to obtain a quantitative description of network properties. It involves deriving a feasible set of steady-state fluxes that optimizes a stated cellular objective (e.g. maximizing biomass production in a metabolic network) that is subject to a set of constraints such as conservation of mass (i.e. the underlying network stoichiometry) [23]. Flux–balance analysis can be used to evaluate the performance of a system under various perturbations such as different cellular objectives or varying constraints, and the resultant sets of fluxes can be compared with each other and with experimental data to yield predictive models of large-scale biochemical networks exposed to different conditions [33]. For example, the robustness of a network can be evaluated by varying the upper bound constraint on the flux through a particular reaction or pathway and by observing the resultant growth rate [34]. This approach has been used to demonstrate the robustness of *E. coli* to changes in the activity of individual enzymes or pathways [35]. The effect of various gene knockouts and gene additions can also be explored by constraining the associated reaction fluxes or including the additional reaction functionality and by evaluating the resultant phenotype *in silico* [34]. Gene knockouts have been evaluated extensively in *E. coli* [33].

Successful characterization of multifunctionality in networks

Mathematical modeling and analysis techniques can provide considerable insight into integrated biological systems. A systems analysis approach combining experimental and computational efforts has been used to interrogate the overall system behavior of nucleocytoplasmic transport, a highly integrated function that overlaps with many 'traditional' mechanisms of cellular signaling and transport. This work has resulted in the first quantitative description of Ran flux between the nuclear and cytoplasmic compartments in eukaryotic cells [36]. The model predicts that the Ran exchange factor RCC1, and not the flux capacity of the nuclear pore complex, is the crucial regulator of steady-state flux across the nuclear pore complex. Furthermore, by indicating that the free RanGTP concentration is orders of magnitude lower than the dissociation constant for Ran binding to importins, the model suggests that our understanding of the importin complex is incomplete and that additional factors might be essential. In the context of integrated biological systems, this model of nucleocytoplasmic transport is an ideal system because it captures, in a single structured framework, multiple cellular components acting across distinct cellular processes.

Other integrative models of biochemical pathways that have emerged recently have met with equal success. For example, a comprehensive model, comprising the dynamics of receptor stimulation and the mitogen-activated protein kinase cascade, activation of gene expression and adaptation of cellular metabolism, has recently described the response of *S. cerevisiae* to osmotic shock [37]. To generate the model, the steady-state and dynamic behavior of the components of osmotic shock were studied by using stoichiometric relationships defined through a series of differential equations. The model has yielded information about the dynamic operation of the processes underlying osmotic adaptation in yeast, including the time courses of many steps in the signaling cascade for which experimental data have not yet been generated.

Two additional integrative analyses have used constraint-based approaches. In one, the integrated regulatory and metabolic networks of *E. coli* were extensively interrogated and additional regulatory rules were systematically predicted and validated [38]. This analysis was the first to account for regulatory processes on a genomic scale with a stoichiometric formalism of a metabolic network. In the other, the stoichiometric matrix of the JAK–STAT signaling network was reconstructed and analyzed by using constraint-based methods, as mentioned, to quantify overlap between metabolic demands and signaling pathways [21].

Furthermore, quantitative modeling has been successfully applied to the study of integrated systems in other disciplines. For example, the utility of the real-time optimization (RTO) strategy has been demonstrated in the field of chemical engineering, where it has been used to maintain at economic optimum an integrated plant with many different types of process units, including, for example, a monitoring system, a feed system and a control apparatus [39]. In addition, a dynamic RTO strategy has been recently developed to maintain an integrated plant setting at economic optimum [40]. Rather than using a conventional RTO strategy that assumes merely a steady-state model of the plant, the dynamic RTO formulation attempts to consider the long transient dynamics, often lasting several days, that are characteristic of an integrated plant owing to its recycle loops, transportation delays and large intermediate storage capacities [40]. A dynamic, integrated plant is similar to a biological system: just as different plant components have overlapping

functions that operate on different timescales, the components of a biological system are overlapping and range in time from slow (regulatory networks) to fast (signaling networks). In addition, the dynamic RTO formulation requires individual plant components to be synthesized in a single structured framework, much as biochemical network components and their interconnections are represented in a stoichiometric formalism.

These and similar success stories underscore the potential that stoichiometric network reconstructions coupled with the mathematical modeling and analysis tools emerging in systems biology offer to the study of integrated biochemical networks, wherein functional components of individual networks are overlapping across spatial and temporal scales. Indeed, ongoing efforts across several laboratories involve both further developing systems analysis tools for merging metabolic, regulatory and signaling network reconstructions, and extending the kind of dynamic analysis applied to plant settings in chemical engineering to these integrated stoichiometric network reconstructions.

Concluding remarks

With the rapid increase in genomic data over the past decade has come realization of the extensive interconnectivity of metabolic, regulatory and signaling mechanisms. Indeed, numerous examples of network overlap have been published in the past few years (Table 1). Here, we have described how incorporating and accounting for the multifunctionality of network components is imperative to fully understanding the biochemical pathways that are involved in fundamental disease processes. We have further illustrated how systems analyses, including stoichiometric network reconstructions and associated mathematical analyses, have successfully investigated genome-scale network properties of independent biochemical networks and how, by coupling these computational approaches with experimental studies, novel hypotheses about biochemical networks can be generated, evaluated and validated. Most importantly, we have shown how systems analysis techniques can now integrate metabolic, regulatory and signaling processes, thereby revealing new examples of network overlap.

Stoichiometric network reconstruction involves synthesizing network components and their interactions into a single structured framework. Simply assimilating this knowledge of metabolic, regulatory and signaling processes into a single stoichiometric matrix can yield insight about individual network components, including any experimentally observed integrative functions. Furthermore, stoichiometric network reconstruction serves as the basis for mathematical analyses of network properties, such as cross-talk, redundancy and modularity, and the results of these analyses can generate hypotheses about potentially overlapping functions of network components. We have shown how systems analyses can be used to elucidate integrative functions about potentially therapeutically relevant attributes of biological systems.

Indeed, the integrative models of metabolic, regulatory and signaling pathways that result from this systems analysis approach might yield 'digital cells' that can be interrogated *in silico* to assess component and network properties in a holistic fashion. These models have tremendous potential to generate new avenues of biochemical and medical research at a time when assimilation and analysis of the growing wealth of experimental data are becoming increasingly necessary. Many challenges remain to meet this goal, including careful delineation of the timescales associated with metabolic, regulatory and signaling processes, and continued exhaustive characterization of the protein–protein interactions and chemical transformations that connect a given genotype to a cellular phenotype.

Acknowledgements

We thank Jong Min Lee and Bernhard Palsson for valuable comments on the manuscript and integrated analysis of biochemical networks. We also thank the Whitaker Foundation and National Institutes of Health (NIH) for financial support (E.P.G. is supported by a NIH Biotechnology Training Grant and D.L.B. is supported by NIH Grants GM65362 and CA77584).

References

1 Buchholz, A. *et al*. (2001) Quantification of intracellular metabolites in *Escherichia coli* K12 using liquid chromatographic–electrospray ionization tandem mass spectrometric techniques. *Anal. Biochem.* 295, 129–137
2 Papin, J.A. *et al*. (2005) Reconstruction of cellular signalling networks and analysis of their properties. *Nat. Rev. Mol. Cell Biol.* 6, 99–111
3 Palsson, B. (2004) Two-dimensional annotation of genomes. *Nat. Biotechnol.* 22, 1218–1219
4 Kim, J.W. and Dang, C.V. (2005) Multifaceted roles of glycolytic enzymes. *Trends Biochem. Sci.* 30, 142–150
5 Majewski, N. *et al*. (2004) Hexokinase–mitochondria interaction mediated by Akt is required to inhibit apoptosis in the presence or absence of Bax and Bak. *Mol. Cell* 16, 819–830
6 Majewski, N. *et al*. (2004) Akt inhibits apoptosis downstream of BID cleavage via a glucose-dependent mechanism involving mitochondrial hexokinases. *Mol. Cell. Biol.* 24, 730–740
7 Gottlob, K. *et al*. (2001) Inhibition of early apoptotic events by Akt/PKB is dependent on the first committed step of glycolysis and mitochondrial hexokinase. *Genes Dev.* 15, 1406–1418
8 Pastorino, J.G. *et al*. (2002) Mitochondrial binding of hexokinase II inhibits Bax-induced cytochrome *c* release and apoptosis. *J. Biol. Chem.* 277, 7610–7618
9 Rathmell, J.C. *et al*. (2003) Akt-directed glucose metabolism can prevent Bax conformation change and promote growth factor-independent survival. *Mol. Cell. Biol.* 23, 7315–7328
10 Dudley, A.M. *et al*. (2005) A global view of pleiotropy and phenotypically derived gene function in yeast. *Mol. Syst. Biol.* DOI:10.1038/msb4100004
11 Palsson, B.O. *et al*. (2003) Development of network-based pathway definitions: the need to analyze real metabolic networks. *Trends Biotechnol.* 21, 195–198
12 Jeffery, C.J. (2003) Moonlighting proteins: old proteins learning new tricks. *Trends Genet.* 19, 415–417
13 Ostrovsky de Spicer, P. and Maloy, S. (1993) PutA protein, a membrane-associated flavin dehydrogenase, acts as a redox-dependent transcriptional regulator. *Proc. Natl. Acad. Sci. U. S. A.* 90, 4295–4298
14 Muro-Pastor, A.M. *et al*. (1997) Regulation of gene expression by repressor localization: biochemical evidence that membrane and DNA binding by the PutA protein are mutually exclusive. *J. Bacteriol.* 179, 2788–2791
15 Soker, S. *et al*. (1998) Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor. *Cell* 92, 735–745
16 Hara, M.R. *et al*. (2005) *S*-nitrosylated GAPDH initiates apoptotic cell death by nuclear translocation following Siah1 binding. *Nat. Cell Biol.* 7, 665–674

17 Blower, M.D. *et al*. (2005) A Rae1-containing ribonucleoprotein complex is required for mitotic spindle assembly. *Cell* 121, 223–234
18 Lum, J.J. *et al*. (2005) Autophagy in metazoans: cell survival in the land of plenty. *Nat. Rev. Mol. Cell Biol.* 6, 439–448
19 Orimo, A. *et al*. (2005) Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* 121, 335–348
20 Dvorak, H.F. (1986) Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. *N. Engl. J. Med.* 315, 1650–1659
21 Papin, J.A. and Palsson, B.O. (2004) The JAK–STAT signaling network in the human B-cell: an extreme signaling pathway analysis. *Biophys. J.* 87, 37–46
22 Reed, J.L. *et al*. (2006) Towards multidimensional genome annotation. *Nat. Rev. Genet.* 7, 130–141
23 Price, N.D. *et al*. (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* 2, 886–897
24 Papin, J.A. *et al*. (2003) Metabolic pathways in the post-genome era. *Trends Biochem. Sci.* 28, 250–258
25 Schuster, S. *et al*. (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol.* 17, 53–60
26 Papin, J.A. *et al*. (2004) Comparison of network-based pathway analysis methods. *Trends Biotechnol.* 22, 400–405
27 Schilling, C.H. *et al*. (2000) Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J. Theor. Biol.* 203, 229–248
28 Wagner, C. and Urbanczik, R. (2005) The geometry of the flux cone of a metabolic network. *Biophys. J.* 89, 3837–3845
29 Stelling, J. *et al*. (2002) Metabolic network structure determines key aspects of functionality and regulation. *Nature* 420, 190–193
30 Vijayasankaran, N. *et al*. (2005) Metabolic pathway structures for recombinant protein synthesis in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 68, 737–746
31 Papin, J.A. *et al*. (2002) The genome-scale metabolic extreme pathway structure in *Haemophilus influenzae* shows significant network redundancy. *J. Theor. Biol.* 215, 67–82
32 Price, N.D. *et al*. (2002) Determination of redundancy and systems properties of the metabolic network of *Helicobacter pylori* using genome-scale extreme pathway analysis. *Genome Res.* 12, 760–769
33 Segre, D. *et al*. (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15112–15117
34 Kauffman, K.J. *et al*. (2003) Advances in flux balance analysis. *Curr. Opin. Biotechnol.* 14, 491–496
35 Edwards, J.S. and Palsson, B.O. (2000) Robustness analysis of the *Escherichia coli* metabolic network. *Biotechnol. Prog.* 16, 927–939
36 Smith, A.E. *et al*. (2002) Systems analysis of Ran transport. *Science* 295, 488–491
37 Klipp, E. *et al*. (2005) Integrative model of the response of yeast to osmotic shock. *Nat. Biotechnol.* 23, 975–982
38 Covert, M.W. *et al*. (2004) Integrating high-throughput and computational data elucidates bacterial networks. *Nature* 429, 92–96
39 Cutler, C.R. and Perry, R.T. (1983) Real-time optimization with multivariate control is required to maximize profits. *Comput. Chem. Eng.* 7, 663–667
40 Tosukhowong, T. *et al*. (2004) An introduction to a dynamic plant-wide optimization strategy for an integrated plant. *Comput. Chem. Eng.* 29, 199–208
41 Klausner, R.D. and Rouault, T.A. (1993) A double life: cytosolic aconitase as a regulatory RNA binding protein. *Mol. Biol. Cell* 4, 1–5
42 Rouault, T.A. *et al*. (1991) Structural relationship between an iron-regulated RNA-binding protein (IRE-BP) and aconitase: functional implications. *Cell* 64, 881–883
43 Hatfield, G.W. and Benham, C.J. (2002) DNA topology-mediated control of global gene expression in *Escherichia coli*. *Annu. Rev. Genet.* 36, 175–203
44 Hall, D.A. *et al*. (2004) Regulation of gene expression by a metabolic enzyme. *Science* 306, 482–484
45 Zinkel, S.S. *et al*. (2005) A role for proapoptotic BID in the DNA-damage response. *Cell* 122, 579–591
46 Kamer, I. *et al*. (2005) Proapoptotic BID is an ATM effector in the DNA-damage response. *Cell* 122, 593–603
47 Mendelsohn, A.R. *et al*. (2002) Cyclin D3 activates Caspase 2, connecting cell proliferation with cell death. *Proc. Natl. Acad. Sci. U. S. A.* 99, 6871–6876
48 Ghosh, A.K. *et al*. (1999) Functional domains of *c-myc* promoter binding protein 1 involved in transcriptional repression and cell growth regulation. *Mol. Cell. Biol.* 19, 2880–2886
49 Subramanian, A. and Miller, D.M. (2000) Structural analysis of α-enolase. Mapping the functional domains involved in down-regulation of the c-myc protooncogene. *J. Biol. Chem.* 275, 5958–5965
50 Feo, S. *et al*. (2000) ENO1 gene product binds to the c-*myc* promoter and acts as a transcriptional repressor: relationship with Myc promoter-binding protein 1 (MBP-1). *FEBS Lett.* 473, 47–52
51 Sun, Y.J. *et al*. (1999) The crystal structure of a multifunctional protein: phosphoglucose isomerase/autocrine motility factor/neuroleukin. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5412–5417
52 Liotta, L.A. *et al*. (1986) Tumor cell autocrine motility factor. *Proc. Natl. Acad. Sci. U. S. A.* 83, 3302–3306
53 Niinaka, Y. *et al*. (1998) Expression and secretion of neuroleukin/phosphohexose isomerase/maturation factor as autocrine motility factor by tumor cells. *Cancer Res.* 58, 2667–2674
54 Nabi, I.R. *et al*. (1990) Identification of B16-F1 melanoma autocrine motility-like factor receptor. *Cancer Res.* 50, 409–414
55 Watanabe, H. *et al*. (1996) Tumor cell autocrine motility factor is the neuroleukin/phosphohexose isomerase polypeptide. *Cancer Res.* 56, 2960–2963
56 Shi, Y. and Shi, Y. (2004) Metabolic enzymes and coenzymes in transcription – a direct link between metabolism and transcription? *Trends Genet.* 20, 445–452
57 Matarese, G. and La Cava, A. (2004) The intricate interface between immune system and metabolism. *Trends Immunol.* 25, 193–200
58 Shamji, A.F. *et al*. (2003) Integration of growth factor and nutrient signaling: implications for cancer biology. *Mol. Cell* 12, 271–280
59 Berger, F. *et al*. (2004) The new life of a centenarian: signalling functions of NAD(P). *Trends Biochem. Sci.* 29, 111–118
60 Cantley, L.C. (2002) The phosphoinositide 3-kinase pathway. *Science* 296, 1655–1657
61 Chaput, M. *et al*. (1988) The neurotrophic factor neuroleukin is 90% homologous with phosphohexose isomerase. *Nature* 332, 454–455
62 Faik, P. *et al*. (1988) Mouse glucose-6-phosphate isomerase and neuroleukin have identical 3′ sequences. *Nature* 332, 455–457
63 Winkler, W.C. *et al*. (2004) Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* 428, 281–286
64 Murrell, M.P. *et al*. (2004) The systems biology of glycosylation. *ChemBioChem* 5, 1334–1347
65 Welchman, R.L. *et al*. (2005) Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat. Rev. Mol. Cell Biol.* 6, 599–609