Mathematical models of metabolic pathways
J Varner* and D Ramkrishna†

There have been recent advances in metabolic flux analysis. In particular, the marriage of traditional flux balancing with NMR isotopomer distribution analysis holds great promise for the detailed quantification of physiology. Nevertheless, flux analysis yields only static snapshots of metabolism. To robustly predict the time evolution of metabolic networks, dynamic mathematical models, especially those that contain a description of both gene expression as well as enzyme activity, must be utilized. When mechanistic and regulatory information is not available, heuristic-based methods, such as the cybernetic framework, can be employed to describe the action of these control mechanisms. In the ‘high-information’ future, as more biological information becomes available, such heuristic-based approaches can be replaced by mechanistic mass-action representations of physiology that stem directly from genetic sequence.

Introduction

Metabolic networks are highly complex nonlinear reaction systems whose functions are tightly coordinated to meet the physiological demands of living organisms. Thus, directed attempts to reengineer physiology often result, for a number of reasons, in unexpected and, more seriously, undesirable results. The complexity of this special class of models for mathematical analysis and the rigorous, quantitative design guidance that it brings.

In what follows, we review two major classes of metabolic design tools, namely, stoichiometric models and flux analysis, and dynamic mathematical models of metabolic networks. The former is a purely descriptive tool that yields detailed quantitative snap-shots of metabolism that could be employed to gain insight into physiology. The latter, stochastic models and flux analysis, are based on linear algebraic equations relating experimental measurements of substrate inputs and product outputs to the steady-state intracellular flux distribution via a stoichiometric model. The origin of FBA follows from the work of Aiba and Matsuoka [1] who calculated fluxes for citric acid cycle in Canadula hypocotyls for the purpose of identifying the most experimentally consistent metabolic model. This approach was further refined by Papoutsakis et al. [2–4] to describe fermentations of butyric acid, butanol, mixed-acid, and propionic acid bacteria. Similar concepts have been applied to analyze growth on different carbon sources in Saccharomyces cerevisiae [5], hydrome production by Corynebacterium glutamicum [6–10], the analysis of physiological ramifications of mutations in C. glutamimc [11•], as well as investigation of the metabolism of peptide amino acids in Chinese Hamster Ovary cells grown in complex medium [12]. Although widely applicable to a diverse range of systems, strictly speaking FBA is restricted to steady state situations (although Jorgensen has applied it to analyze fed-batch fermentation’s of Pseudomallium chrysogenum by assuming a pseudo-steady state [13]) and may require restrictive assumptions regarding metabolically active pathways or cofactor dependence. Furthermore, FBA only allows the determination of net fluxes between metabolic pools, and as such is unable to assess the degree of reversibility of reversible reactions. Restrictive assumptions regarding the stoichiometric network, which are a mathematical consequence of determining generalized inverses, can be overcome by linear programming. Recently, such a procedure has been successfully employed by Pramunuk and Kesling [14••] to investigate Escherichia coli metabolism.

Stoichiometric models and flux analysis

The various types of metabolic flux analysis offer at differing levels of complexity a descriptive snapshot of physiology. This technique has been gaining popularity, especially amongst biologists because of its relative ease of application and, furthermore, its unobtrusive conceptual nature (i.e. easily understood by non-mathematicians). Metabolic flux analysis can be broadly subdivided into three components: metabolic flux balancing analysis (FBA), NMR, GC/MS mass spectrometry (MS) flux analysis methods; and finally, an approach which integrates the previous two, termed the combined approach.

FBA, the simplest of the flux analysis techniques, involves the solution of a set of linear algebraic equations relating experimental measurements of substrate inputs and product outputs to the steady-state intracellular flux distribution via a stoichiometric model. The origin of FBA follows from the work of Aiba and Matsuoka [1] who calculated fluxes for citrate production in Canadula hypocotyls for the purpose of identifying the most experimentally consistent metabolic model. This approach was further refined by Papoutsakis et al. [2–4] to describe fermentations of butyric acid, butanol, mixed-acid, and propionic acid bacteria. Similar concepts have been applied to analyze growth on different carbon sources in Saccharomyces cerevisiae [5], hydrome production by Corynebacterium glutamicum [6–10], the analysis of physiological ramifications of mutations in C. glutamimc [11•], as well as investigation of the metabolism of peptide amino acids in Chinese Hamster Ovary cells grown in complex medium [12]. Although widely applicable to a diverse range of systems, strictly speaking FBA is restricted to steady state situations (although Jorgensen has applied it to analyze fed-batch fermentation’s of Pseudomallium chrysogenum by assuming a pseudo-steady state [13]) and may require restrictive assumptions regarding metabolically active pathways or cofactor dependence. Furthermore, FBA only allows the determination of net fluxes between metabolic pools, and as such is unable to assess the degree of reversibility of reversible reactions. Restrictive assumptions regarding the stoichiometric network, which are a mathematical consequence of determining generalized inverses, can be overcome by linear programming. Recently, such a procedure has been successfully employed by Pramunuk and Kesling [14••] to investigate Escherichia coli metabolism.

A complementary approach to the determination of intracellular flux distribution follows from the application of NMR and/or GC/MS techniques in conjunction with 13C substrate labeling and traditional FBA. This approach augments traditional FBA by providing key intracellular information, such as flux ratios at metabolic branch points, that can function as constraints that, in combination with the stoichiometric model, uniquely identify intracellular...
fluctuations. The combined application of NMR and/or GC/MS techniques has been employed to calculate flux ratios at key metabolic branch points [16,17••] and to determine degrees of reversibility in complex networks [18–20]. Additionally, recently Schmidt et al. [22] have used labeling measurements to verify the predictions of traditional FBA. The combined application of labeling measurements and FBA alleviates, somewhat, the need to make restrictive biological assumptions; however, this technique still remains strictly valid only in steady state situations.

To determine intracellular flux distribution in transient situations, for example, in fed-batch cultivation or under shifting process conditions, NMR techniques can be employed in the absence of a stoichiometric model. Using mixtures of unlabeled and uniformly labeled substrates, Szyperski [23] has shown that it is possible to follow the cleavage of covalent bonds in the carbon backbone of biomolecules by measuring the degree of coupling between adjacent carbon atoms in the amino acids. This labeling pattern, termed the isotope distribution, can be utilized to not only determine intracellular flux ratios and active metabolic pathways [16,24] but also to yield a comprehensive characterization of network topology and an approximate assessment of exchange fluxes [25••,26–28]. Thus, this class of NMR measurement, in addition to determining intracellular fluxes in transient situations, can also be utilized to define the reactions and refine assumptions about active pathways. As a complementary theoretical counterpart to the experimental determination of intracellular fluxes via isotope distributions can be simulated using isotope models, thereby providing a prediction of the ramifications of assumptions on active pathways or network structure. This concept which has been applied to simple networks [29,30] has been greatly simplified by Schmidt et al. [31,32•], and hence, is directly applicable to complex networks.

**Dynamic mathematical models of metabolic networks**

While flux correlation analysis provides a descriptive snapshot of physiology, clearly, it’s role in the rational design or reengineering of metabolism is limited by its inability to address the dynamic evolution of the system. Obviously, this is true because stoichiometric models contain no kinetic or regulatory information. Thus, to predict system-time evolution (e.g. how metabolite or enzyme levels change with time) the stoichiometric foundation of metabolic flux analysis must be augmented by a description of the system kinetics. Kinetic mathematical modeling has long been employed, in the context of complex chemical reaction systems, to predict the effect of changes in process parameters. In this spirit, a movement within the biochemical engineering community has grown around the use of mechanistic kinetic mathematical models to predict the effects of environmental or genetic change. The single-cell simulations of the response of E. coli to mixed nitrogen sources [33–36] and perturbations in metabolism resulting from introduction of recombinant plasmids [37–40] are early examples of this movement. When viewed broadly, however, the results of such attempts, in a biotechnological context, have been disappointing when compared with experiment and the rigorous promise of kinetic mathematical modeling has gone unrealized [41].

The key distinction between conventional chemical reaction systems and metabolic networks, which is often missing in kinetic metabolic network models, is the influence of regulation and control. In conventional chemical reaction systems, knowledge of the kinetics completes the treatment of the system. In biological systems, however, all levels of metabolic function (i.e. transcription, translation, and catalytic activity) are tightly integrated and coordinated with the global environment of the organism, hence yielding adaptability in the face of changing conditions. Thus, the “conventional” mathematical treatment of a metabolic network, encompassing only the kinetics and stoichiometry, is often hard pressed to correctly predict system adaptation because it lacks a description of the forces driving the adaptation. This capability, however, is exactly what is required for the reengineering of physiology.

In cases where detailed enzymatic regulatory information is available (e.g. the mass action understanding of genetic or post-translational regulatory/control mechanisms), mechanistic kinetic models, describing regulation of enzyme activity via augmented kinetic expressions, can be constructed. For example, Reza et al. [42] has formulated a detailed model of glycolysis in S. cerevisiae describing the short-time scale system response to a glucose pulse added to continuous culture. Dae et al. [43] constructed a mechanistic kinetic model of polyhydroxyalkanoate production in plant plastids and simulated product production rates in light and dark conditions. Pissara et al. [44] formulated a kinetic model describing the formation of penicillin V and other key by-products in fed-batch cultures of P. chrysosporium that was employed, along with metabolic control analysis (MCA), to describe metabolic control properties. Hatzimanikatis et al. [45,46] took this concept a step further and formulated a kinetic model of glycolysis in E. coli that was combined with a mixed-integer linear programming optimization routine to predict optimum regulatory properties of network enzymes. In all of these models, however, no description of transcription or translation was included, only post-translation control. As such, they would not be able to describe significant physiological adaptation (shifts in expression profiles).

To capture significant adaptation it is vital to include the regulatory/control mechanics driving gene expression, in addition to those governing enzyme activity. In cases where detailed information is known for small networks or individual operons, mechanistic models of transcription, such as the genetically structured model of lac operon function in E. coli constructed by Lee and Bailey [47,48] or
more recently the structured lac operon model of Keasling and co-workers [49], can be formulated. This class of model, when married with a description of metabolite levels, could possibly predict shifts in gene expression. For large networks or cases where mechanistic information is not available, however, a genetically structured description may become intractable or impossible. In these situations, heuristic-based methodologies can be employed. Recently, Liao and co-workers [50••] have developed a fuzzy-logic approach, based upon qualitative knowledge of enzyme properties, that compensates for incorrectly or inadequately modeled kinetics. Although presented in a different context, this algorithm could easily be adapted to predict gene expression profiles based on qualitative information. Van Riel et al. [51•] have formulated an approach where gene expression is controlled via the action of a key transcription regulator. They have utilized this concept to predict the response of central nitrogen metabolism to ammonia and glutamine pulses in S. cerevisiae.

We approached the problem of representing metabolic regulation and control in the absence of specific mechanisms from a different perspective. We postulated that control of metabolic networks, specifically gene expression, can be described using cybernetic principles, that is, physiology operates to optimally satisfy nutritional objectives [52••,53•]. Employing such a principle, in combination with an approximate knowledge of system kinetics, affords a mathematical modeling framework that is potentially capable of predicting modification of enzyme expression and activity profiles in the face of genetic or environmental perturbation. This tool has been successfully employed to predict system response to flux diversion away from lysine production toward threonine in batch cultivations of Corynebacterium lactofermentum [54••], as well as for penicillin V production in fed-batch cultivations of P. chrysogenum [55]. More recently, Varner et al. [56•] have formulated a cybernetic model of glycolysis in E. coli that was able to predict significant system response to overexpression of phosphofructokinase and pyruvate kinase under a number of different experimental conditions. Cybernetic principles have also been employed by van Riel et al. [57] to predict the response of central nitrogen metabolism to ammonia pulses in continuous cultures of S. cerevisiae under glutamine-limited conditions. In this case, however, a somewhat different mathematical formalism is applied.

As adept as cybernetic models are at predicting local physiological responses, they do not yield information regarding specific genetic alterations other than overexpression and/or deletion of network enzymes. To investigate more specific alterations, for example, control-protein binding disruption or promoter sequence reengineering, Varner and Bailey have married a cybernetic treatment of translational and post-translational regulation with a genetically structured treatment of transcription. This hybrid framework has been employed to predict strategies for the alleviation of glucose control of maltoose metabolism in S. cerevisiae (J Varner and JE Bailey, unpublished data).

Models in a post-sequencing, highly parallel era

Current paradigms (e.g. the fundamental processes by which we analyze biological systems) are in a state of flux following from massive parallel data generation and the availability of genome sequences. The latter yields a particularly interesting challenge, that is, the rational utilization of the fundamental blueprint of an organism. Furthermore, the former, with the increasing prevalence of automated high-throughput screening/selection technologies and the growing trend toward ‘combinatorial minimization’ (e.g. large numbers of microscale physiological experiments), necessitates the immediate development of large-scale data handling, storage and analysis machinery. In the coming ‘high-information’ era, mathematical models will most certainly play a central role. For example, Bhalla and Iyengar [58] have recently formulated and employed a very detailed mechanistic mathematical model of general signaling pathways to simulate molecular-level interactions present within signaling networks. This model, in conjunction with the material presented throughout this review, is indicative of the class of tools required to organize and dynamically analyze large quantities of biological information.

In terms of sequence utilization, one conceptually interesting approach developed by Palsson and co-workers [59] couples the formulation of the stoichiometric model of a target organism with its genetic sequence. This linkage, which bases the formulation of the stoichiometric matrix upon the functional assignment of open reading frames, can then be employed, via solution of a linear programming problem, to predict phenotype-genotype relationships [59]. Yin and co-workers [60•] took the concept of simulating the phenotype-genotype relationship a step further by marrying a detailed kinetic model of phage T7 function with a genetic algorithm to simulate the phenotype resulting from shuffled genomes. These techniques, when combined with sequence and genome-wide expression level phylogenetic algorithms, such as those outlined by Ukkonen and co-workers [61] (among many others), are key to the rational utilization of sequence data.

Conclusions

Stoichiometric modeling and metabolic flux analysis can provide very detailed and valuable snapshots of physiology. In a physiological reengineering context, however, stoichiometric models alone are of limited value because they are incapable of addressing dynamic system evolution and, in particular, adaptation resulting from genetic or environmental change. To consider network time evolution and physiological adaptation, stoichiometry must be augmented with kinetics as well as metabolic regulation and control. In the ‘high-information’ future, the marriage between kinetics, stoichiometry and metabolic regulation
may be a mechanistic one. Today, however, in the absence of such a luxury, heuristic methods such as fuzzy-logic or cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

References and recommended reading

Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

2. Papoutsakis ET: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

3. Papoutsakis ET, Meyer CL: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

4. Papoutsakis ET, Meyer CL: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

5. Cortassa S, Aon JC, Aon MA: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

6. Vallino JJ, Stephanopoulos G: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

7. Vallino JJ, Stephanopoulos G: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

8. Vallino JJ, Stephanopoulos G: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

9. Vallino JJ, Stephanopoulos G: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

10. Vallino JJ, Stephanopoulos G: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

11. Vallino JJ, Stephanopoulos G: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

12. Nyberg GB, Balcarcel RR, Follstad BD, Stephanopoulos G, E. coli: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

13. Nyberg GB, Balcarcel RR, Follstad BD, Stephanopoulos G, E. coli: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

14. Pramanik J, Keasling JD: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

15. Pramanik J, Keasling JD: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

16. Pramanik J, Keasling JD: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

17. Sauer U, Hatamian-Majd V, Bailey JE, Hochul M, Sogzewski T: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

19. Wochert W, de Graaf AA: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

20. Wochert W, de Graaf AA: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

21. Zupke C, Stephanopoulos G: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

22. Zupke C, Stephanopoulos G: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

23. Szyperski T: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

24. Hochul M, Patzch H, Osthell D, Wochert K, Sogzewski T: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

25. Bacsor TB, Stephanopoulos G, Wang DIC: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

27. Szyperski T: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.

33. Domach MM, Leung SK, Cahn RE, Cocks GC, Shuler ML: 
Papers of particular interest, published within the annual period of review, References and recommended reading

cybernetics must be employed. It is only through this combination that we can arrive at a predictive methodology for the rational design of metabolic networks.


44. Pissara PN, Nielsen J, Bazin MJ: Analytical technique. This work is a sizeable leap forward both in terms of the mathematical representation and manipulation of sequence information as well as computational technique.


