Network models in the study of metabolism

Arsenio Rodríguez
Centro Nacional de Biotecnologia Agricola
Instituto de Estudios Avanzados and
Postgrado en Biologia, Universidad Simon Bolivar
Caracas, Venezuela

Diogenes Infante*
Centro Nacional de Biotecnología Agrícola
Instituto de Estudios Avanzados
Apdo 17606 Parque Central
Caracas 1015-A, Venezuela
Tel: 58 212 903 5185
Fax: 58 212 903 5093
E-mail: dinfante@idea.gob.ve

Abbreviations: FBA: flux balance analysis
MFA: metabolic flux analysis
SKM: structural kinetics model
TIA: terpenoids indol alkaloid

The systematic study of the genetic fingerprint (genomics) and the biochemistry (metabolites) that goes with a specific cellular process requires the characterization of all the small molecules that form the profile of metabolites and the associated genes. The metabolome represents the collection of all the metabolites during certain process in an organism. The transcriptome represents the gene expression profile, all the messengers RNA in a defined condition. Then to understand the whole process, the studies of metabolites must be accompanied with studies of the gene expression, hence the metabolome must be accompanied by the transcriptome, so we can identify genes and metabolites whose synthesis is induced by a specific process, an infection or stress. Studies of metabolomics generate an enormous amount of data, then they need mathematical and computational tools to establish the correlations between the biochemical and genetic data, and to build up networks that represent the complex metabolic interactions that occur in each case, using tools like Graph and Networks Theory to elucidate the emergent properties inherent to the complex interactions of the metabolic maps. This paper describes the major mathematical tools that can be used for these studies, with emphasis on a semi-qualitative proposal known as the kinetic structural model.

The systematic approach as a tool for the study of metabolism

For a reasonable, objective and complete vision of our external world a connection between our senses, our minds and nature is necessary. In order to make this connection, we generally make use of abstract structures known as "Models". The mathematical diagramation of these models in formal systems or maps allows us to infer causality and make deductions about our environs. A simple representation of the components of a system are input/output block diagrams.

$$f : A \rightarrow B \qquad [1]$$

Where f represents a process, A the input and B the output. Clearly, B could be the input for another process and the system can thus be visualized as a network (Figure 1) with all the possible interactions between the inputs and the outputs (Oltvai and Barabási, 2002). This interrelated system allows us to study a phenomenon from a holistic perspective instead of the separate analysis of each of its constituent parts, where (in the holistic perspective) each element of the process f represents a bi-functional unit whose activity is defined by its context within the model and at the same time contributes to that context and thus the
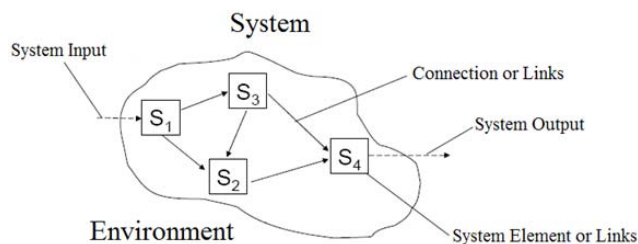
*Corresponding author

Figure 1. Schematic representation of a system.

environs or superstructure as a whole. From a conceptual point of view it is a self-referential system, which means if the context or the system is severely disrupted, the superstructure, as defined by its functional component or components, will cease to exist and the functional component will also disappear. Thus the whole is more than the sum of its parts, each part having its own semantics and its own context (Mikulecky, 2005).

Metabolism may be understood as a series of interrelated chemical-enzymatic reactions that are continually operating to maintain vital functions far from the thermodynamic equilibrium. Metabolism is a natural phenomenon that may be studied from the perspective of f being a metabolic process in an organism and A a group of metabolites that transform into a group of products B. These branched and interconnected reactions (that generate several specific products) represent the metabolic pathways. The presence or absence of an enzyme along one of these pathways varies with the type of cell, tissue or organism and depends on its physiological state (nutritional condition, state or stress). The behavior of the metabolic network after specific variations along some of its pathways and flows is difficult to infer, as the cellular response to perturbations (genetic or environmental) may involve other networks. (For example: cell signaling and transcriptional cascades) (Schmidt and Dandekar, 2002; Qi and Ge, 2006). Thus, for the study of metabolism, its development and applications in biotechnology and metabolic engineering, it is fundamental to elucidate the organization and the biological significance of the metabolic pathways (Yang et al. 1998). This review aims to elucidate the importance of the study of metabolism, based on the concept of the network and to describe the major mathematical approaches that permit modeling the metabolic networks.

Metabolic networks as a source of information

As has already been explained, in order to gain a complete vision of metabolism, it is necessary to go beyond a simple description of its pathways. This global vision could be achieved by the study of metabolism based on the network concept, whereby the characterization of the relationships and interconnections between the elements are taken into account rather than simply looking them as a collection of interconnecting elements (Lazebnik, 2002). The particular

way in which the elements are linked should be described thoroughly in order to precisely depict the pattern of connections, which is known as the network topology.

In an abstract sense, a network may be described by a graph $N = (V, E)$ with a set of elements represented by nodes or vertices V such that each node or vertex v_i is a member of the set V, where $i=1, 2, 3…$ M, and M is the total number of nodes or vertices ($v_i Î V$). Each pair of nodes define a link and these L links also form a set $E = (e_j) \mid j = (1,2,3,…,L)$. Otherwise, a network is the relation of the Cartesian Product $N = V_1$ x V_2 for all pairs of nodes v (v_r,v_s) | r,s=1, 2, 3…, M; where again M is the total number of nodes network. If the trajectory between two nodes is important for its description, each link between two nodes would have a particular orientation (Barabási et al. 1999). This type of graph is known as a directional graph or *digraph* and characterizes the structure or topology of systems such as the metabolic networks (Figure 2). The constitutive relations that describe the network elements and the topology of the network may be formalized independently and then combined to provide a *solution* to the network (Almaas and Barabási, 2006). A network is said to have a solution when everything observed within it can be specified. When the network is directed (digraph) the nature of its solution is a system of pathways that can be formulated as a set coupled kinetic differential equations that describe the relationships within the network (Mikulecky, 2005).

An incidence matrix I = L x M, where L = n° of links and M = n° of nodes, shows the links in the columns and the nodes
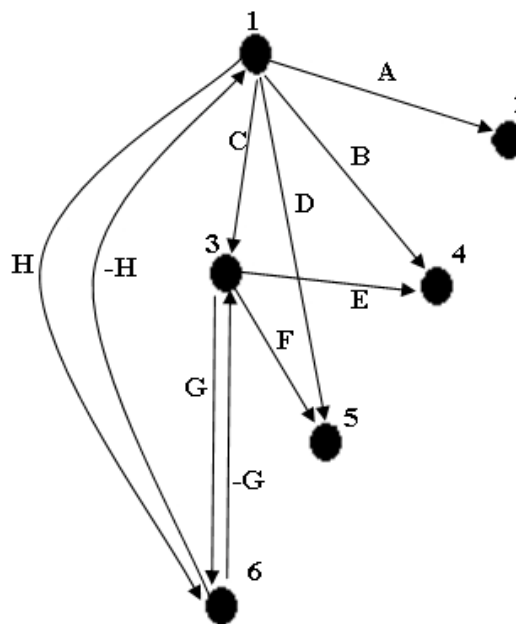

Figure 2. Example of a directed graph (digraph) _N_ representing a network _R_, with _M_ = 6 nodes and _L_ = 10 links.

$$I = \begin{bmatrix} & A & B & C & D & E & F & G & -G & H & -H \\ -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & -1 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \end{bmatrix} \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{matrix}$$

Figure 3. Incidence matrix *I* (6 x 10) associated with the network illustrated in figure 1.

in the rows, where the positive 1's (+1's) represent the incidence of the links that enter a node v_s and the negative 1's (-1's) the incidence of the links that leave a node v_r. The zeros represent the non-incidence (null incidence) between links and nodes (Figure 3). The incidence matrix is a computational tool that allows the development of algorithms that facilitate the calculation of the parameters inherent to the network.

General properties of graphs

Average degree of a graph. The number of *K* vertices being associated with a node are known as the degrees of the network and in the case of a digraph *N*; *K* could be divided into incoming degrees K_i and outgoing degrees K_o, where $K = K_i + K_o$. The average degree of a homogeneous undirected graph is $Z = <K> = 2 L/M$, since by definition each link is joined to two nodes. Nevertheless, in heterogeneous graphs, where the nodes do not have the same number of links, this value is not valid

(Fernández and Solé, 2005).

Nodal distance. The distance between two nodes in a graph *N* is $d(v_i, v_j)$ and is defined as the smallest number of links (shortest path) between two nodes v_i and v_j. In the example illustrated in Figure 2, the pair of nodes (1,6) would have $d(1,6) = 2$ and for the pair of nodes (1,5) $d(1,5) = 1$.

Mean path length or graph radius (l). This is defined as the average distance between any two vertices (v_i, v_j):

$$\ell = \frac{1}{N(N-1)} \sum_{\forall i,j} d(v_i, v_j) \qquad [2]$$

This equation also describes the average number of steps that must be taken from any node in order to pass through the entire graph.

Clustering coefficient. Graphs may show local structures, which may be examined by means of the study of the average clustering coefficient of a network (C) according to the equation:

$$C = \frac{1}{N} \sum_{i=1}^{N} c_i = \frac{1}{N} \sum_{i=1}^{N} \frac{2E_i}{k_i(k_i - 1)} \qquad [3]$$

Where k_i represents the number of vertices or nodes neighboring the *i*-th vertex of the graph, E_i represents the links of the k_i and N represents the total number of vertices or nodes in the graph. This property gives the density of
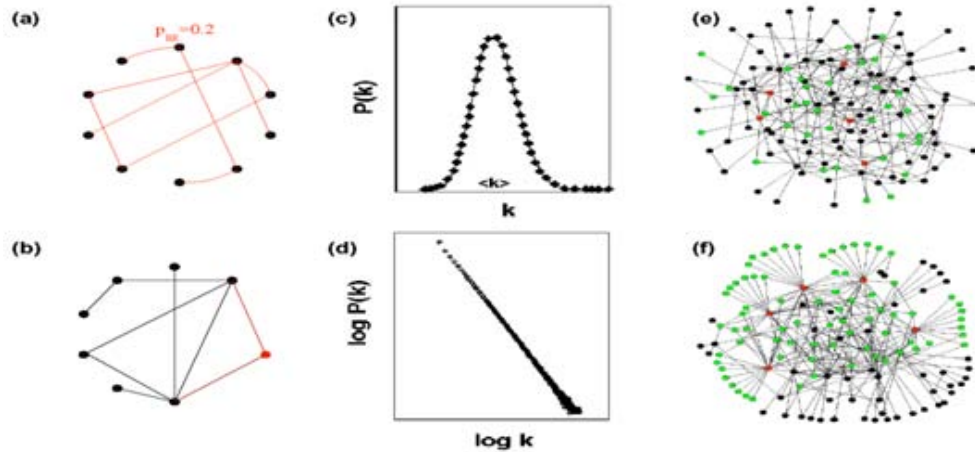
Figure 4. The fundamental differences between a random network (for example an Erdős-Rényi network) and a scale free network. a) shows a random network of $N = 10$ and $P = 0.2$. This type of network grows with a homogeneous degree distribution, *ie.*, $K \approx <K>$ and the degree function grows in the same way as for a Poisson distribution (c); where most of the nodes (e) have more or less the same number of links (red dots) as their neighboring nodes (green dots). (b) In contrast, scale free networks are non-homogeneous, the characteristic degree distribution is not peaked and instead approximates a power law (d); these types of network grow preferentially around a few nodes "Hubs" -the red dots- where most of the links are concentrated: green dots (f) (Barabási, 2003).
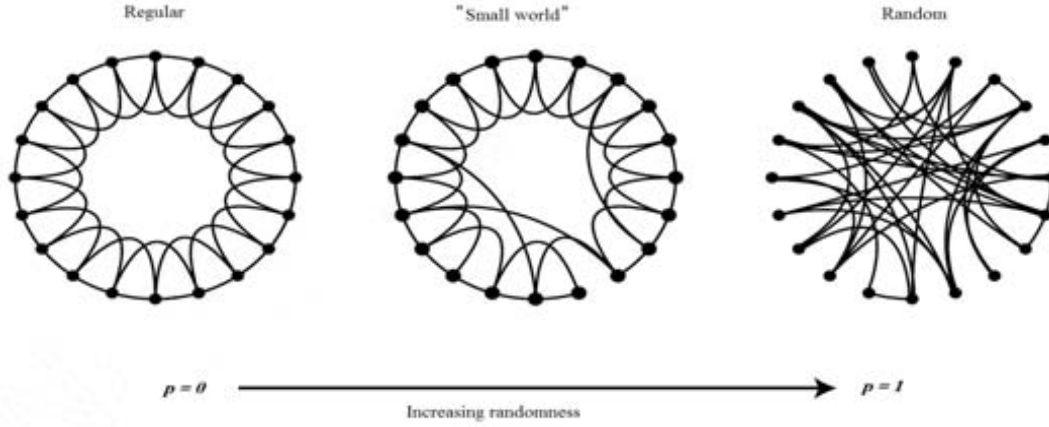
Figure 5. The small world effect is an intermediate between a regular and a random network, as can be seen by simulating a network of *n* nodes with *k* links as a probability function *p* (Watts and Strogatz, 1998).

links in the neighborhood of a particular node (Dorogovtsev et al. 2003).

Degree distribution. One of the principal statistics of a network is its degree distribution $p_{(k)}$ which describes the fraction of nodes that have a certain degree *k* for a graph N, *i.e.* it indicates the probability *p* that a randomly selected node has *k* links. For a random graph (RG) the probability $p_{(k)}$ of linking two nodes is:

$$p_{(k)} = \binom{N-1}{k} p^k \left(1-p\right)^{N-1-k} \qquad [4]$$

When $N \rightarrow \infty$ $p_{(k)}$ approximates a Poisson distribution where $Z = p/N$:

$$p_{(k)} \approx \frac{Z^k e^{-z}}{k!} \qquad [5]$$

Many biological phenomena do not adjust to random graphs (where the links are uniformly distributed with probability p about a mean $< K >$), since natural networks have been observed to show a heterogeneous degree distribution where a small number of nodes has a large number of links (Wodak et al. 2004). Natural networks, such as those described from genetic or metabolic protein-protein interactions show a degree distribution that adjusts to the Bounded Power Law as shown in equation 6:

$$p_{(k)} \sim (k + k_o)^{-\gamma} \, e^{-k/k_o} \qquad [6]$$

Where k_0 is a constant, γ is the degree exponent of the scale and depending on the network has values between 2 and 3, usually $\gamma = 2.5$ and k_c is the bounded degree and indicates the maximum number of links in the network. These networks are known as "scale free" networks and

essentially mean that the network has a small and finite number of highly interconnected nodes (hub nodes) and a large number of scantily connected nodes (Barabási, 2003). In Figure 4, a random graph (for example, Erdős-Rényi) and a scale free graph are compared.

The value of γ determines several network properties and the smaller γ is the more important the hubs nodes are. For values of γ > 3 hubs nodes are not important. If 3 > γ > 2, a hierarchical organization of hubs nodes appears, where the more connected hubs are linked to a small fraction of the total number of nodes. When γ = 2, a network emerges where the most connected hubs are in contact with a large fraction of the total number of nodes. The properties of scale free networks are only observed when γ < 3, explicitly, when the dispersion of the distribution $P_{(k)}$, defined as $\sigma^2 = < k^2 > - < k >^2$ increases with the number of nodes. When γ > 3, scale free network characteristics do not manifest (Barabási and Oltvai, 2004).

An important property associated with scale free networks occurs when new nodes *m* are preferentially attached with probability Π to the most connected nodes *i* of a pre-existing network with m_o nodes, according to the equation:

$$\Pi_i = k_i / \sum_j k_j \qquad [7]$$

Where k_i are the nodes neighboring *i*. This property is known as the Principles of Preferential Addition or Barabási-Albert model (Almaas and Barabási, 2006).

"Small world" effect. Also know as the Watts and Strogatz Model or the "small world" effect (analogous to the "six degrees of separation" phenomenon described for social networks) has been observed in biological networks. This phenomenon described for the first time by Stanley Milgram, was brought to light in an experiment whereby Milgram investigated the number of links that are needed to
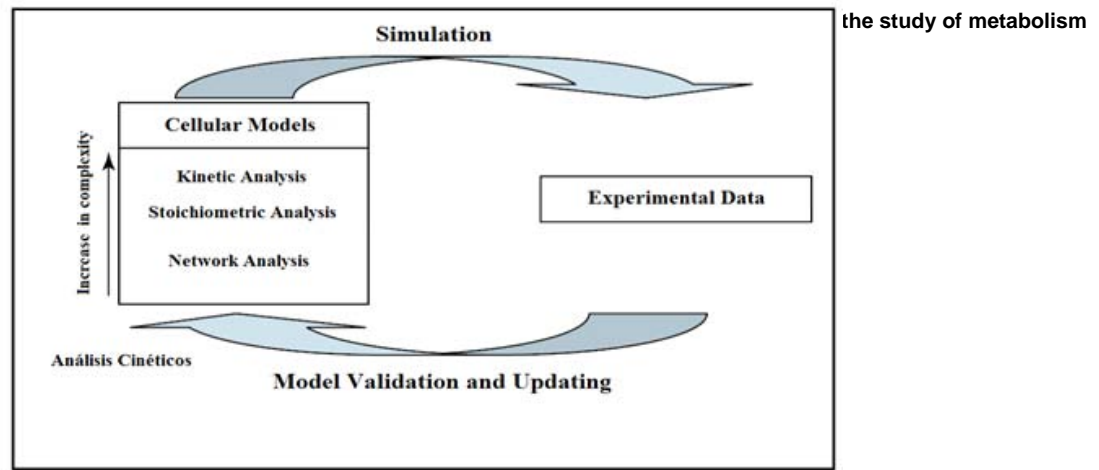
Figure 6. Iterative process in the development of metabolic models. (Adapted from Borodina and Nielsen, 2005).

connect one person with another within a class or social group (writers, actors, etc) and arrived to an estimate of a maximum of six degree of social contacts. Small world graphs, defined semi-quantitatively, are intermediate between random graphs and regular or crystalline ones (Figure 5), because they have a small mean path length (l) as do random networks, but show a high clustering coefficient (C) as do regular networks (Watts and Strogatz, 1998) (for a quantitative definition see Humphries and Gurney (2008)). In these networks the path between any pair of nodes is shorter compared to random networks of a similar size and connectivity and thus the speed of the dissemination of information is higher than in other types of network (Alm and Arkin, 2003).

In general small world networks are characterized by the following properties (Newman, 2001):

1. The local neighborhood of each node is preserved as for regular networks.

2. The network diameter (Shortest average distance between pairs of nodes) increases logarithmically with the number of nodes.

3. Thus, small world networks are those that permit the connection of two nodes with very few links (Amaral et al. 2000).

Emergent properties of scale free networks. The study of scale free and small world networks allows the perception of characteristics that only appear when observed in context, they are not apparent when only individual elements are examined. These characteristics are known as Emergent Properties. The emergent properties associated with scale free and small world networks are: Robustness: the sustainability of network functionality in spite of the loss of some nodes or links; Preferential attachment: nodes new to the network bind with a higher probability to hub nodes; Modular Patterns: the emergence of groups of nodes with discrete functions, separable from

the rest of the network and recurring circuit elements: Metabolic Networks use regulatory circuits such as feedback inhibition in many different pathways. The transcriptional network of *Escherichia coli* has been shown to display a small set of recurring circuit elements termed "network motifs". Each network motif can perform a specific information processing task such as filtering out spurious input fluctuation, generating temporal programs of expression or accelerating the throughput of the network. (Alon, 2003; Takemoto and Oosawa, 2007).

Metabolic networks are examples of scale free and small world models. In these networks the nodes represent the metabolites and the links the enzymatic reactions that transform certain metabolites into others. These networks are bipartite digraphs: a sub group of nodes $M_1 \in N$, the substrate, and a sub group of links $V_2 \in N$, the product.

Studies undertaken of the metabolic networks of 43 organisms (Alm and Arkin, 2003) have found that:

a) They are scale free networks and non random.

b) They are small world networks.

c) They are modular.

The metabolic networks studied showed a fixed diameter that varied very little between organisms. Paradoxically, each hub node was linked with a low probability to other hub nodes (property known as *Dissortativity*) and at the same time were scale free, since most of the metabolic routes were strongly linked to a few nodes with high connectivity (property known as *Network Hierarchization*) (Jeong et al. 2000). These results suggest the presence of new properties associated with these types of network.

Although these findings have helped disentangle general principles in biological design, the risk of simplifying natural processes that are dynamic, evolutionary and subject to selective pressure remains. It is thus necessary to apply tools that allow the study of these networks over time (Schilling et al. 1999; Holme and Huss, 2003).

METABOLIC NETWORKS FROM THE PERSPECTIVE OF MATHEMATICAL MODELS

In recent years a huge mass of data has been obtained from the "omics" technologies, such as genomics, transcriptomics y metabolomics. This information has been stored in extensive data bases available on the Internet. The study and integrative analysis of these data bases is supported by mathematical instruments such as cellular models (Fiehn, 2001). The evolution of these tools is an iterative process whereby the models are compared with the experimental data, which both validate and improve them, leading to the formulation of new *a posteriori* hypotheses (Figure 6) (Borodina and Nielsen, 2005). Essentially, these models seek to translate the present knowledge of cellular components and their interactions into a mathematical representation, to provide a precise, formal and unambiguous language for the study of the behavior of metabolism integrated into a system: Systems Biology (Lazebnik, 2002).

Mathematical models can vary in complexity and focus, nevertheless the model should be judged on the basis of the objective that it pursues, because metabolic networks are dynamic entities with flows of material produced by the biological cellular activity and subject to hierarchical regulatory interactions under cellular direction or canalization (Jørgensen et al. 2005). Thus, the description of metabolic phenotypes should include the composition of the metabolites, the measurement of their flows and the dynamics of the network. The mathematical and computational models most commonly used for the clarification of the structure and dynamics of metabolic networks are discussed in the next section including an

analysis of their most important virtues and limitations (Figure 7).

Topological models and the analysis of qualitative networks

The qualitative approach fundamentally uses graph theory and describes metabolic networks as bipartide digraphs with start nodes (precursor metabolites), end nodes (end products) and connections that link both nodes in a directional manner (chemical-enzymatic reactions) (Alterovitz and Ramoni, 2006). The enzymatic reactions are governed by chemical and thermodynamic laws.

According to Xia et al. (2004) the most important rules of metabolic restriction are:

· Stoichiometry: Refers to the number of molecules that participate in the chemical reaction. This property does not vary for the same reaction in different organisms and does not change with pressure, temperature or other physicochemical parameters. Chemical Stoichiometry appears as a topological property of metabolic networks.

· Relative reaction rate: This property is determined by the thermodynamics of the reaction and is dependent on the physicochemical conditions of the environs.

· Absolute reaction rate: This property is determined by the genetic function and is specific for each enzyme. If changes that modify the enzymes occur in the genome, these will show up in the chemical reactions.
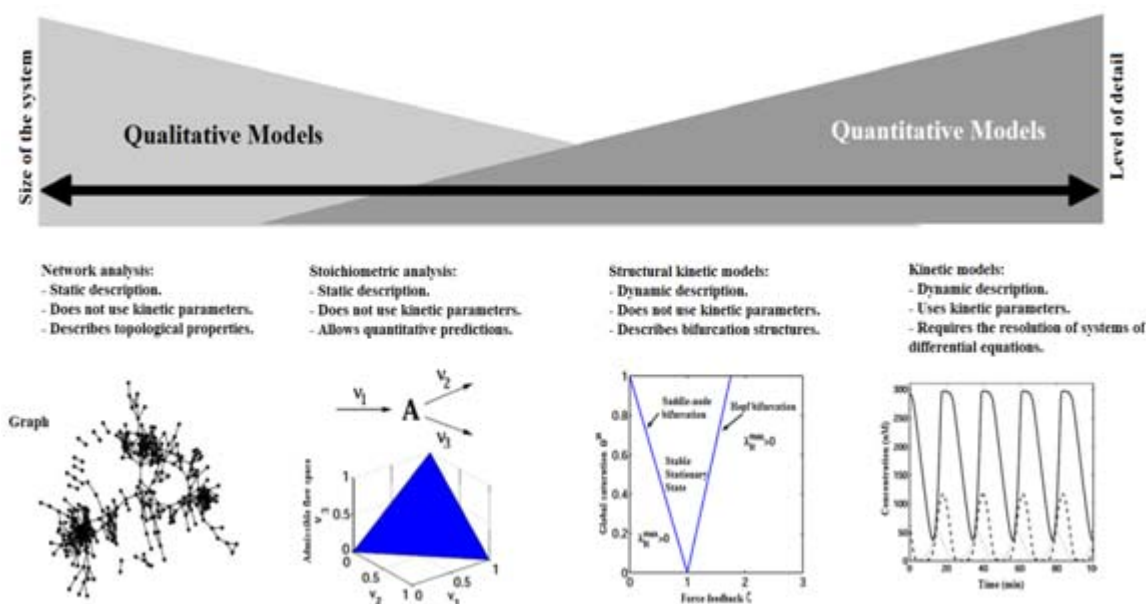
The properties mentioned above provide limited

Figure 7. Comparative representation of the most common mathematical and computational models used for the analysis of metabolic networks (adapted from Steuer, 2007; Kahlem and Birney, 2006).

information on the functional state of any metabolic network restricted to a temporal or spatial interval and to the physicochemical nature of the extra and intra cellular environment (Schilling et al. 1999; Fiehn, 2001). Biological systems are directed by a cellular organization which determines the temporal nature of the interactions. This behavior is linked to properties of a superior order peculiar to cellular networks -self-organization and self-assemble *(Autopoiesis)*- that permit the components of the system to spontaneously create a functional network (Brinkmeier and Schank, 2005; Palsson, 2006). This behavior is associated with scale free and metabolic networks (Dorogovtsev and Méndez, 2002).

Another quality inherent to biological networks is the bi-linearity and high interconnectivity between its components, thus cellular networks behave as hypergraphs (Wagner, 2001). Hypergraphs have topologically non linear

properties, with a higher growth of functional states than is expected from the number of components in the network. In other words, the number of phenotypic functions that derive from a genome is not linearly correlated with the number of genes present in that genome. This fundamental characteristic promotes the expression of different possible functional states (similar phenotypic behaviors). Nevertheless, an organism does not exploit or completely use all of its similar functional states. At a particular moment in time and space, only a limited subgroup of these states will be selected and expressed by the cells (Palsson, 2006). Thus, metabolic networks are under inviolable restrictions (hard restrictions) associated with the thermodynamics of the organism (balance of mass and energy), their evolutionary history (genetic and regulatory restrictions) and restrictions related to the environment, allowing a profile of solutions (phenotypic structure) in a given interval of space and time (Förster et al. 2003).
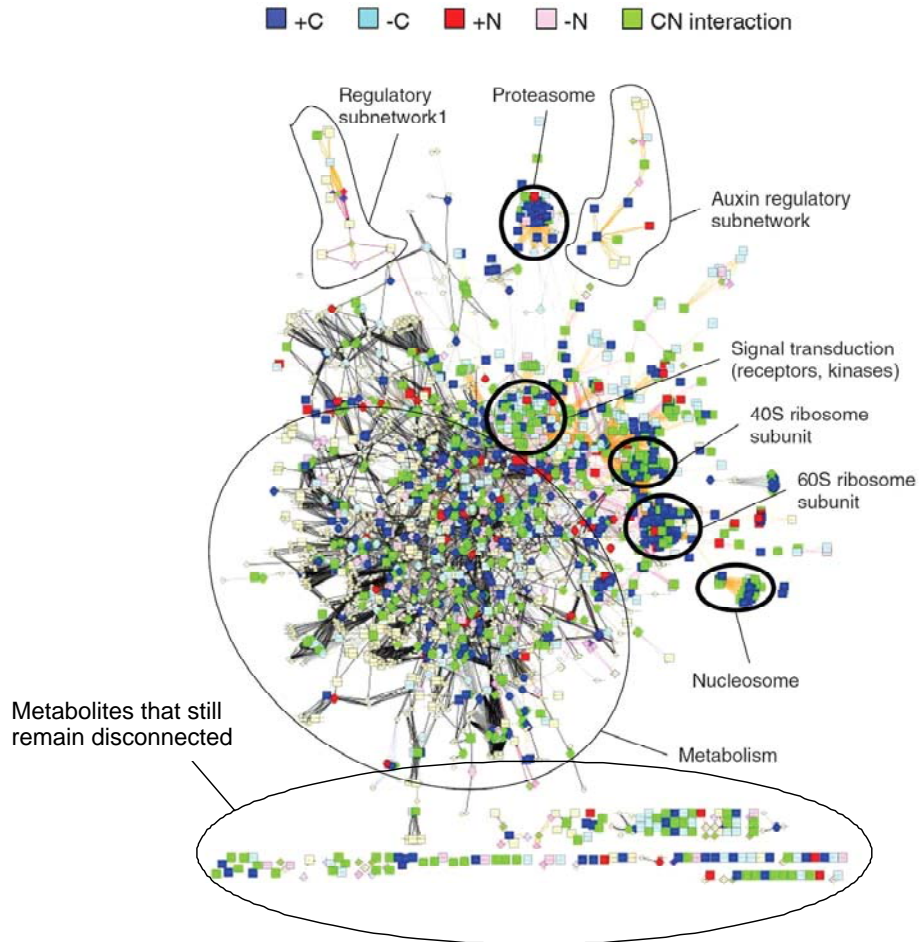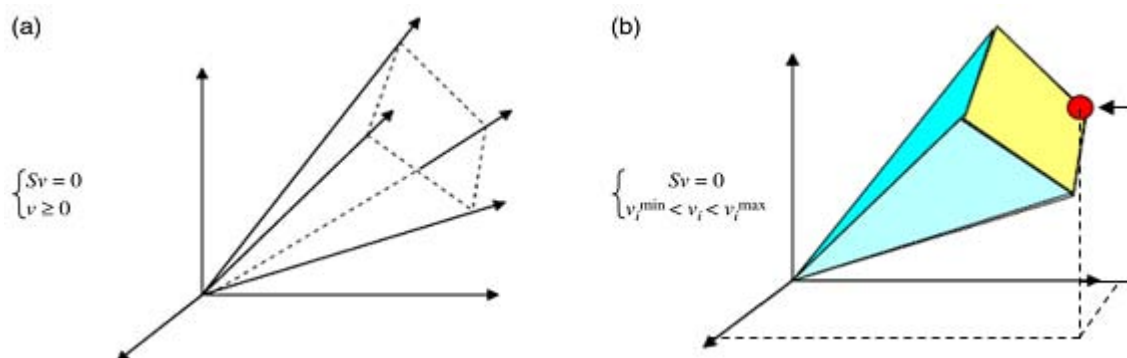

Figure 8. Reconstruction of a qualitative network model of a generic plant cell (adapted from Gutíerrez et al. (2007))

7

Figure 9. The solution space of metabolic networks. In the three-dimensional unconstrained flux space of a hypothetical network, fluxes can assume any value.
a) Solutions obtained by restrictions stoichiometry, enzymatic limitations or the balance of mass, then the feasible solution space is reduced to a cone;
b) restrictions by using an objective function, such as maximizing biomass accumulation, can identify a coordinate (set of flux values) on an of the feasible flux cone (red point). (adapted from (Lanzeni et al. 2008).

Why is it necessary to study the topology of metabolic networks? How can structural information help to clarify the role of individual nodes and their relationships within the network? Which of the functions or solutions of the network will be exploited by the cells, given the existence of equivalent, multiple states of the network or silent phenotypes? How do the responses of the cells change over time as a result of selective pressure (Evolutionary history)? (Vitkup et al. 2006).

These questions must be answered in the future with the support of these tools. An approach to finding the volume and shape of possible phenotypes or solutions for the optimization of biological traits under genetic, environmental or evolutionary restrictions, is performed by the group of Braunstein et al. (2008) where they suggest that the best technique that allows this characterization is based on the method of Monte Carlo sampling (MCS) of the area of metabolic flux of the network under steady state, given the complexity of metabolic networks and the enormous computational cost involved in these calculations. They propose a computational strategy known as message-passing algorithm derived from the field of Statistical Physics and Information Theory. This algorithm based on conjecture or Bethe approximation (Bethe Ansatz) allows calculating the volume of a convex polygon of incomplete higher-order dimension. Was successfully used in the characterization of a metabolic network of 46 reactions and 34 metabolites in blood and red blood cells in predicting the effect on the disruption of some genes of central metabolism (gene knock-out) in *E. coli* (Braunstein et al. 2008). Also this approach was used iteratively in the human metabolic network to identify missing components. There are 356 metabolites that still remain disconnected in the published version that serve to highlight the areas where more detailed work is required (Patil and Nielsen, 2005).

On the other hand, the need for a system-wide approach to network construction in plants has been recognized and used to construct networks in *Arabidopsis*, which to date integrates the regulatory and metabolic networks that contains 7,635 nodes (6,176 genes and 1,459 metabolites) and 230,900 interactions among these nodes (Figure 8). The results of qualitative network analysis of Gutíerrez et al. (2007) provide a starting point for future studies, identifying the regulatory factors - or network hubs - that are likely to be important for the regulation of gene networks in *Arabidopsis* roots in response to Carbon and Nitrogen metabolites. Combining existing knowledge into qualitative network models and using this as a platform to interpret microarray data, which allowed them to identify molecular machines.

In a gene to metabolite study, Rischer et al. (2006) use a holistic approach to characterize the terpenoids indol alkaloid (TIA) biosynthesis in the Madagascar periwinkle *Catharantus roseus* cultured cells. They combine the use of genome wide profiling of transcript using cDNA-amplified fragment length polymorphism (cDNA-AFLP) combined with metabolic profiling, to characterize the synthesis of Ajmalicine, an antihypertensive alkaloid, and Vinblatine and Vincristine, antineoplastic bisindole alkaloids, the later two are produce in very low amounts *in vivo* and are not synthesized *in vitro*. The accumulation profiles of 178 metabolites were combined with the expression profiles of the 417 transcripts for an integrated analysis. They create a correlation network for the complete set of metabolites and genes, which provide insights into the complex regulation of TIA metabolism but also of secondary metabolism in general. They were able to visualize an unbiased subset centered on the Tabersonine node, whose gene-to-metabolite network around consisted of 11 metabolites and 13 genes representing the nearest neighbors. This result increased the practical potential of metabolic engineering of this important medicinal plant.

In principle, the structure affects the function, a fact which can be used to discover or predict new functions or even the evolutive origin of the metabolic network (Solé and Valverde, 2008). The topology dependent approach has contributed to the understanding of general aspects of

$$\downarrow V_1$$
$$X_1$$
$$\downarrow V_2$$
$$X_2$$
$$\downarrow V_3$$
$$X_3$$
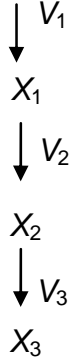
Figure 10. Hypothetical pathway of the transformation of the metabolites X_1 to X_3 by means of the enzymatic reactions V_2 and V_3. For simplicity, the system is open with the input represented by V_1.

metabolic networks (scale free, small world effect, preferential attachment, modularity, network motifs) (Jin et al. 2007). The development of metabolic maps along with the progress of implementation of computer platforms to simulate cellular functions under the tutelage of tools such as graph theory, have been an important first step in understanding the contribution of cellular metabolism (Ishii et al. 2004). Nonetheless, this perspective has not been able to resolve questions inherent to variations in time, is ambiguous in its representation for detailed discussions of structure and is limited as regards of the interpretation and comparative analysis of experimental metabolic data associated with other biological networks (For example: transcriptomic networks or networks of protein interactions) (Gross, 2005), due to the hypergraphic nature of metabolism. Then it is important to incorporate information relevant to flows within the network or changes of the metabolic mass over time, for a complete analysis of metabolic systems (Steuer, 2007; Domijan and Kirkilionis, 2008) due to the static nature of the topological models. We continue in this review with the conceptual models that incorporate the kinetic parameters of metabolism.

Stoichiometric analysis

Although the development of dynamic models that permit a complete simulation of the cellular system have been attempted, this has proved to be a difficult task due to the lack of unambiguous information on the kinetic and regulatory properties of metabolic reactions (Tomita et al. 1999).

Nevertheless, even in the absence of kinetic data, relevant information on the theoretic capacities and operative modes of metabolic systems may be obtained using different Stoichiometric approaches, such as: Flux Balance Analysis (FBA) and Metabolic Flux Analysis (MFA), and others applications: Elemental Mode Analysis (EMA) and Extreme Pathway Analysis (EPA) (Ciliberto et al. 2007;

Rios-Estepa and Lange, 2007). The Flux Analysis is based on the description of the flow reactions of a metabolic network under certain restrictions or boundary conditions of biological networks: a) physicochemical and thermodynamic restrictions, b) topological restrictions c) environmental restrictions and d) regulatory restrictions (Kauffman et al. 2003). FBA translates enzymatic reactions into flow differential equations such as:

$$\frac{dx_{i(t)}}{dt} = Sv_{j(x,k)} \qquad [8]$$

where X, describes the concentration vector $X_i = (x_1, x_2, \ldots, x_m)$, for $I = 1,2,3,\ldots,$ m metabolites, V describes the metabolic flow vector: $V_j = (v_1, v_2, \ldots, v_n)$ for $j = 1,2,3,\ldots,$ n reactions and the Stoichiometric matrix S as the linear transformation of the vector V in X. The general dynamic mass balance equation for a metabolic network (Equation 8) is represented by a set or system of differential equations (Varma and Palsson, 1994):

$$\frac{dx_{i(t)}}{dt} = \sum_{j=1}^{n} s_{ij} \cdot v_{j_{(x,k)}} \qquad [9]$$

Each equation represents the sum of all the individual flows v_k in which the metabolite is either synthesized or degraded x_i. There are m metabolites (x_i) and n reactions (v_i) in the network, thus the dimensions (dim) of the vectors X, V and of the matrix S are:

$$\dim(X) = m, \quad \dim(V) = n, \quad \dim(S) = m \times n \quad [10]$$

The information on the Stoichiometric of the reactions is represented by these equations and by the S matrix (the incidence or Stoichiometric matrix). This matrix correspond to the inter-phase between the biochemical and genetic data (Edwards and Palsson, 2000) and the starting point for the mathematical calculations necessary (using *in silico* tools) to elucidate the properties of biological networks. In other words, the S matrix relates the mathematical properties with the biochemical properties of the network (Morgan and Rhodes, 2002).

Assuming that metabolic networks operate between successive quasi-stationary states (the macroscopic variables -metabolite concentration and flux- are maintained within a tolerance interval defined within a time period) where a 0 net flux balance exists, in accordance with the law of conservation of mass (Roscher et al. 1998).

The distributions of the flows that satisfy this condition belong to the null space of the S matrix ($< \Delta S > = 0$) converting Equation 9 in:

$$\frac{dx_{i(t)}}{dt} = \sum_{j=1}^{n} s_{ij} \cdot v_{j(x,k)} = 0 \qquad [11]$$

Due to the fact that the system is indeterminate, since m < n, which permits infinite solutions *V* of the differential equations, it is necessary the incorporation of experimental information for the imposition of some restrictions, for example: nutrient availability (α) and the maximum flux allowed for an enzymatic route (β), in this case an object function (V_j) is obtained, such as $\alpha \leq V_j \leq \beta$ (Segrè et al. 2002). Thereafter, with the use of linear programming the optimum solutions for the system can be found (Bro et al. 2006; Nielsen, 2007). This reduction in the space of admissible flux provides the basis for FBA (Reed and Palsson, 2003) (Figure 9).

Several metabolic networks have been analyzed in yeast, plants and mammalian cell culture using these tools. Among them the photosynthetic prokaryote *Synechocystis* sp. PCC 6803 is of special interest to plant biologists as a model system (Schwender et al. 2004). FBA has not been applied on a plant model yet, but promises to be a useful tool for developing embryos, in which optimal biomass accumulation and light use-efficiency are also reasonable assumptions or constraints (Schwender, 2008).

MFA is a powerful methodology for the determination metabolic pathway fluxes (*f*). In this approach, the intracellular fluxes are calculated using a stoichiometric model (metabolic pathway map) for the major intracellular reactions and applying mass balances around intracellular metabolites. A set of measured extracellular fluxes are used as input to the calculations, typically uptake rates of substrates and secretion rates of metabolites. The result of flux calculation is a metabolic flux map showing a diagram of the biochemical reactions included in the calculations with an estimate of the rate at which each reaction in the diagram occurs at steady state (Libourel and Shachar-Hill, 2008).

In a study of the metabolic reconfiguration in response to oxidative stress Ralser et al. (2007) showed that there is a dynamic rerouting of the metabolic flux to the pentose phosphate pathway, with the concomitant generation of the reduced electron carrier nicotinamide adenine dinucleotide phosphate (NADPH), which is a conserved post transcriptional response to oxidative stress. After Grant (2008) this study also demonstrates the need to integrate genomic, biochemical and *in silico* modeling approaches to fully understand how cells regulate metabolic fluxes during oxidative stress conditions.

In addition to quantification of pathway fluxes, metabolic flux analysis can provide additional insights about other important cell physiological characteristics. As shown here: (1) Identification of branch point control (nodal rigidity) in cellular pathways, (2) Identification of alternative pathways, (3) Calculation of non measured extracellular fluxes and (4) calculation of maximum theoretical yields (Stephanopoulos et al. 1998).

In MFA, kinetic parameters are not required, all that is needed are the stoichiometric data and the data for the metabolic demand of the network, however, additional available information could be incorporated in the future (for example thermodynamic restrictions) (Liebermeister and Klipp, 2006).

One of the main disadvantages of this method is that it assumes the metabolic network is in steady-state. Computationally, steady-state models are easiest to manage and steady-state systems have been used productively in plant studies over the last dozen years to yield extensive and detailed flux maps of central metabolism (Saito, 2009). An alternative to estimate flows and develop predictive models in non-plants systems is the use of the dynamic labeling methods. Dynamic labeling traces the change in the level of a metabolite in a metabolic pathway after the application of a stable isotope-labeled compound. The time-dependent decrease in the labeling level of the precursor as well as the increase in that of a down-stream metabolite in the pathway is monitored as a dynamic process. The value of metabolic flux is determined by fitting the model describing the labeling dynamics of the pathway to the observed labeling data (Matsuda et al. 2007).

In addition, a number of studies have used batch cultures in which plant cell suspensions are grown for several days on a medium that contains ^{13}C- or ^{14}C-labelled glucose as carbon source. Baxter and colleges establish a method for deriving non steady- state fluxes based on the mass-balance of identifiable forms of molecules. Following the introduction of a ^{13}C-labelled precursor, a time-series of samples are analyzed by gas chromatography-mass spectrometry to give information both on change in metabolite pool size and the percentage labelling of a series of mass isotopomers of each metabolite (Baxter et al. 2007). Using threonine biosynthesis as a model, they demonstrated that for a reaction in which the carbon skeleton is conserved, it is possible to over-determine fluxes using mass balance of mass isotopomers and there was no requirement for metabolic or isotopic steady state. They also demonstrate how linear regression methods can be used to estimate non-steady state fluxes from mass balance equations.

Kinetic or deterministic analysis (quantitative models)

In the same way as for Graph analysis, the Stoichiometric approach does not permit the direct description of the dynamic properties of the systems. Specifically, does not permit conclusions about the dynamic behavior for a future metabolic state, for example, if we wish to use the flow distribution function $V_{(X_0)}$, with (Xo) representing variable

metabolic concentrations over a period of time (Yang et al. 2005).

For this purpose, kinetic models that permit the satisfactory description of the dynamics of metabolic networks have been developed. The metabolic control analysis (MCA) allows the clustering of metabolic pathways as a set of ordinary differential equations assuming spatial homogeneity (Cronwright et al. 2002). To better understand this mathematical model we shall use the following example, suppose a small metabolic route, which for simplicity is assumed that contains no elements or regulatory structures and assume the transformation of the species X_i for I = 1,2,3…m metabolites associated with the V_j for j = 1,2,3…n chemical-enzymatic reactions that in order to manage the system of equations that arise from the model does not take into account the metabolic precursors needed in any system as illustrated in Figure 10; where m = 3 and n = 3 (so the system is determinate). The system is completely described by the set of *i-th* linear differential equations, shown below in both their canonical and matrix form in equation 12, where $X_{(t)}$ represents the *i-th* time dependent metabolic concentration vector and **S** denotes the stoichiometric matrix (incidence matrix) with *ij* elements.

$$\frac{dx_{i(t)}}{dt} = \sum_{j=1}^{m} s_{ij} \cdot v_{j(x,k)} \qquad \forall \quad i=1,2,3…n; \, j=1,2,3…m$$

$$\frac{d}{dt}\begin{bmatrix} x_{1(t)} \\ X_{2(t)} \\ X_{3(t)} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} v_{1(X,k)} \\ v_{2(X,k)} \\ v_{3(X,k)} \end{bmatrix} \qquad [12]$$

The *j-th* kinetic enzymatic vector $V_{j\,(x,k)}$ describes a non linear function dependent on the concentration of the metabolite *X* and kinetic parameters *K* inherent to the enzyme, such as the Michaelis-Menten constant (K_m), the maximum catalytic velocity (V_{max}) and the equilibrium constant of the reaction (k_{eq}). Once the *K* parameters, the initial conditions $X_{(o)}$ and the functional form of the differential equations have been obtained, the system can be numerically resolved and the behavior of the metabolic pathway and its metabolites completely described over time (Fell, 1992).

Several dynamic properties, such as the dynamic response to perturbations, the stability of a metabolic state and the transition of an oscillating behavior can be obtained using a local linear approximation of the system. This local approximation is obtained by means of the expansion of a Taylor Series of the metabolic system in a stationary state (not necessarily unique or stable). The linear term for the expansion is the Jacobian Matrix *J* of the first order partial derivatives of the system of kinetic equations for $X_{(t)}$ (equation 13). This matrix represents the dynamic response of the system in the neighborhood of a stationary state.

$$J = \begin{bmatrix} \dfrac{\partial f_{(1)}}{\partial X_1} & \dfrac{\partial f_{(1)}}{\partial X_2} & \dfrac{\partial f_{(1)}}{\partial X_3} \\ \dfrac{\partial f_{(2)}}{\partial X_1} & \dfrac{\partial f_{(2)}}{\partial X_2} & \dfrac{\partial f_{(2)}}{\partial X_3} \\ \dfrac{\partial f_{(3)}}{\partial X_1} & \dfrac{\partial f_{(3)}}{\partial X_2} & \dfrac{\partial f_{(3)}}{\partial X_3} \end{bmatrix} \qquad [13]$$

In contrast to the mass balance equation (equation 11), which does not provide information on the stability of the metabolic state when disrupted, the real part of the eigenvalue (characteristic root) of the Jacobian describes the possible scenarios of stability and their transitions from the metabolic state, for example: Hopf type bifurcations (Steuer, 2007). If we take the metabolic pathway of Figure 10 and define {K_i / i=1,2,3…n} as the chemical equilibrium constant of each reaction, the metabolic flows will be established by mass action kinetics in their simplest form and the system of explicit kinetic equations will be:

$$\frac{dx_1}{dt} = -k_2 X_1 = f_{(1)}$$

$$\frac{dx_2}{dt} = k_2 X_1 - k_3 X_2 = f_{(2)} \qquad [14]$$

$$\frac{dx_3}{dt} = k_3 X_2 = f_{(3)}$$

The Jacobian matrix of equation 13 becomes:

$$J = \begin{bmatrix} -k_2 & 0 & 0 \\ k_2 & -k_3 & 0 \\ 0 & k_3 & 0 \end{bmatrix} \qquad [15]$$

Now that we know the elements of *J*, we can describe the stability of the system. In general terms, the negative real part of the eigenvalue of the Jacobian *J* indicates a stable solution for the system, which converges to a stationary state. On the other side, the positive real part of the eigenvalue of *J* describes an unstable solution and the system does not converge to a stationary state. In Figure 11 it is shown the different dynamic behaviors of the system depending on the eigenvalues of the Jacobian matrix evaluated around a stationary state X^o.

Let us consider the 3 x 3 *J* matrix of equation 15. The cubic eigenvalue equation λ becomes $-\lambda^3 - (k_2 + k_3)\lambda^2 + k_2 \cdot k_3 \lambda$ and the eigenvalues:

$$\lambda_1 = \frac{-(k_2 + k_3) + \sqrt{(k_2 + k_3)^2 + 4 \cdot k_2 k_3}}{2} \; ;$$

$$\lambda_2 = \frac{-(k_2 + k_3) - \sqrt{(k_2 + k_3)^2 + 4 \cdot k_2 k_3}}{2}$$

If we assume integer numbers for k_1 and $k_2 > 0$, our system will not converge to a stationary state. As can be observed, a limitation of this approach is the necessity to determine the values of the kinetic parameters k.

A defining characteristic of living cells is the ability to respond dynamically to external stimuli while maintaining homeostasis under resting conditions. Capturing both of these features in a single kinetic model is difficult because the model must be able to reproduce both behaviors using the same set of molecular components. A novel approach differs critically from metabolic flux analysis and previous genome-scale metabolic network reconstructions because it accommodates nonlinear terms that describe the dynamic behavior of each reaction in the system. Previous large scale network reconstructions typically use a stoichiometric matrix to represent the gross flux of metabolites in the system. Purvis et al. (2009) using a modified kinetic modeling (models are built in a stepwise fashion, beginning with small ''resting'' networks that are combined to form larger models with complex time-dependent behaviors), not only reduce the computational cost of fitting experimental time-series data but can also provide insight into limitations on system concentrations and architecture. They have preserved the mathematical form of each kinetic rate equations as reported in the literature, allowing models to be built from existing data in a ''bottom-up'' fashion while still allowing calibration to whole-system experimental data. This feature will substantially improve the accuracy of dynamical system simulation and parameter estimation. This method was illustrate showing how 77 reactions from 17 primary data sources were integrated to construct an accurate model of intracellular calcium and phosphoinositide metabolism in the resting and activated human platelet.

Although the explicit kinetic models quantitatively describe the dynamics and complex properties of metabolic systems (oscillations, multi-stability or irreversible commutations), their computational construction becomes complicated when the components of the network are numerous and the kinetic parameters incomplete (Cronwright et al. 2002).

Structural Kinetic Models (SKM): A semi-quantitative approach

The modified kinetic model proposed by Steuer et al. (2006) is an intermediate between the stoichiometric and deterministic models; it describes the stability and robustness of the possible metabolic states of a network using the minimum relevant kinetic information and it identifies the important interactions and the parameters that rule the dynamic properties of metabolic systems, without the knowledge of the explicit functional forms of the differential equations of the enzymatic kinetic rate.

For most natural phenomena is unnecessary an explicit kinetic model and the J matrix, expressed in terms of the derived partial equations (equation 13), generally requires the explicit kinetic knowledge of the enzymatic reactions (values of the kinetic parameters k). Nevertheless, it is still possible to specify the structure of the Jacobian matrix, even without this information, if each element in the matrix is restricted to a definite time interval (Strogatz, 1994;
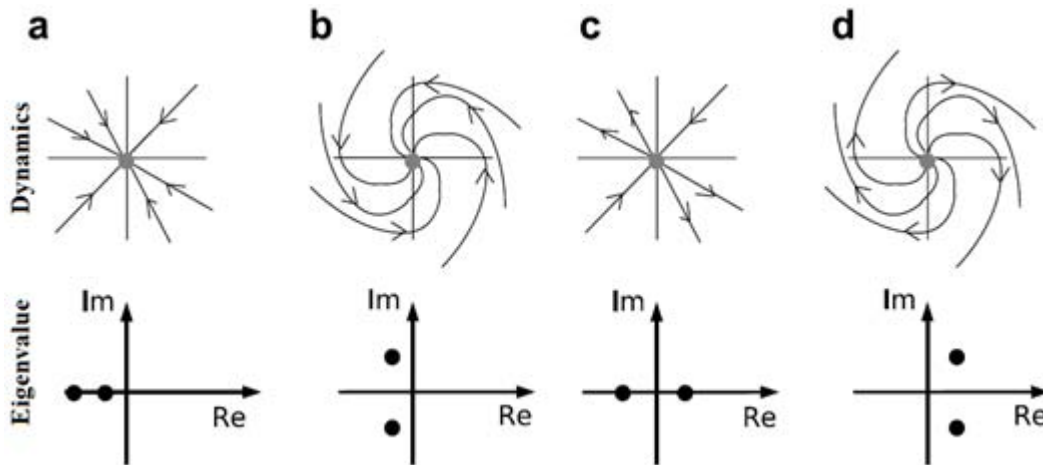
Figure 11. Classification of local dynamics in the neighborhood of a stationary state according to the eigenvalues of the Jacobian after a perturbation.
a) Stable node, the real part of the eigenvalues are negative and the system returns to its original state;
b) Stable focus, the stationary state oscillates back to its original state;
c) Saddle, the metabolic state is unstable and the perturbation is amplified exponentially;
d) Unstable focus, the stationary state is unstable with an oscillating divergence (adapted from Steuer (2007)).

Scott, 2007). Just as the stoichiometric balance imposes restrictions and defines a space of flux solution, the Jacobian matrix imposes a set of possible dynamic behaviors for the metabolic network and defines the dynamic capabilities of any metabolic state.

The structural kinetic models (SKM) are based on the development of a parametric representation of the Jacobian matrix, in a way that each element in the matrix is defined within a time interval without the need of further kinetic information of the metabolic system under study.

The parametric Jacobian matrix (J_x) is defined as the matricial product, $J_x = \Lambda \bullet \theta_X^\mu$ where the matrix Λ reflects the metabolic and stoichiometric state characterized by X^o and V^o in a stationary state (not necessarily unique or stable). The matrix θ_X^μ is made of elements that denote the effective kinetic order o normalized saturation of each reaction with respect to the substrates, products and even possible modulating effectors (for example: allosteric effectors). Both matrices participate in the *parameter space* associated with the metabolic system of interest. Each stationary metabolic state is specified by V^o and X^o and those elements in θ_X^μ different from zero are defined as an interval of dynamic moment, with the kinetic parameters not necessarily known.

The SKM model proposes the rewriting of the system of differential equations described in equation 12 to:

$$\frac{d}{dt}\frac{X_{i(t)}}{X_i^\circ} = \sum_{j=1}^{m}\frac{V_j^\circ}{X_i^\circ}S_{ij}\cdot\frac{V_{j(x,k)}}{V_j^\circ} \quad \forall \ i=1,2,3...n \ \text{[16]}$$

where X_i° indicates the concentrations of the *i*-th metabolite and $V_j^\circ = V(X^o,k)$ represents the distribution of the *j*-th flux, both parameters being associated with a stationary state (not necessarily unique or stable).

By definition the matrices Λ and μ are written as:

$$\Lambda_{ij} = \frac{V_{j(X^\circ,k)}}{X_{i(t)}^\circ}\cdot S_{ij} \ ; \ \mu_{(X)j} := \frac{V_{j(X,k)}}{V_{j^\circ_{(X^\circ,k)}}} \qquad \text{[17]}$$

The Jacobian is re-written with respect to the *y* variable, by means of the transformation $y_i := X_{i(t)} / X^\circ_{i(t)}$:

$$J_x = \Lambda \cdot \theta^\mu_y \quad \text{con} \quad \theta^\mu_y := \left.\frac{\partial\mu(y)}{\partial y}\right|_{y^\circ=1} \qquad \text{[18]}$$

The stationary metabolic state in our example in Figure 10 is then defined by the metabolic concentration $X^\circ_i := [X_1, X_2]$ and metabolic flow by $V^\circ_i := V_1^\circ = V_2^\circ = V^\circ_3$.

With SKM, the system of kinetic equations is replaced by a parametric representation of the metabolic pathway in terms of the Jacobian matrix, becoming in our example:

$$\Lambda = \begin{bmatrix} \dfrac{v^0}{x_1^0} & -\dfrac{v^0}{x_1^0} & 0 \\ 0 & \dfrac{v^0}{x_2^0} & -\dfrac{v^0}{x_2^0} \\ 0 & 0 & \dfrac{v^0}{x_3^0} \end{bmatrix} ; \theta_y^\mu = \begin{bmatrix} 0 & \theta^2_{x_1} & 0 \\ \theta^2_{x_1} & 0 & 0 \\ 0 & \theta^3_{x_2} & 0 \end{bmatrix} \text{[19]}$$

The elements of the matrix Λ represent: a) the structure of the system as defined by the vector V° and the temporal scale as defined by the vector X°, b) the elements of the matrix θ_y^μ represent the normalized degree of saturation of each reaction with respect to its substrate, where $\theta_y^\mu \in [0,1]$ for all substrates and $\theta_y^\mu \in [0,-1]$ for all of the metabolic products (Bapat, 2000). The saturation or effective kinetic order, even with unknown values is defined for an interval: $\theta_y^\mu \in [0,1]$ of the reaction V2 with respect to the substrate X1 and $\theta_y^\mu \in [0,-1]$ of the reaction V3 with respect to the substrate X2. Assuming similar conditions for V2 and V3 as in the explicit kinetic equations and θμy = 1, the Jacobian matrix is specified by the metabolic state [V°; X°1; X°2] is:

$$\mathbf{J}_x = \Lambda\theta_y^\mu = \begin{pmatrix} -\dfrac{v^0}{X_1^0} & -\dfrac{v^0}{X_1^0} \\ \dfrac{v^0}{X_2^0} & -\dfrac{v^0}{X_2^0} \end{pmatrix} \qquad \text{[20]}$$

The Jacobian is evaluated in terms of the four generalized parameters with the purpose of studying the dynamic behavior of a system after a disturbance around a stationary state X°, V° (Steuer et al. 2006; Steuer et al. 2007).

Because lacking or incomplete enzyme-kinetic information that is necessary for classic kinetic models, this proposed method create bridges between topology based approaches and explicit kinetic models of metabolic networks. The SKM models permit to make quantitative conclusions about the dynamics of the system based on a minimum of information (qualitative) compared to the explicit kinetic models (quantitative).
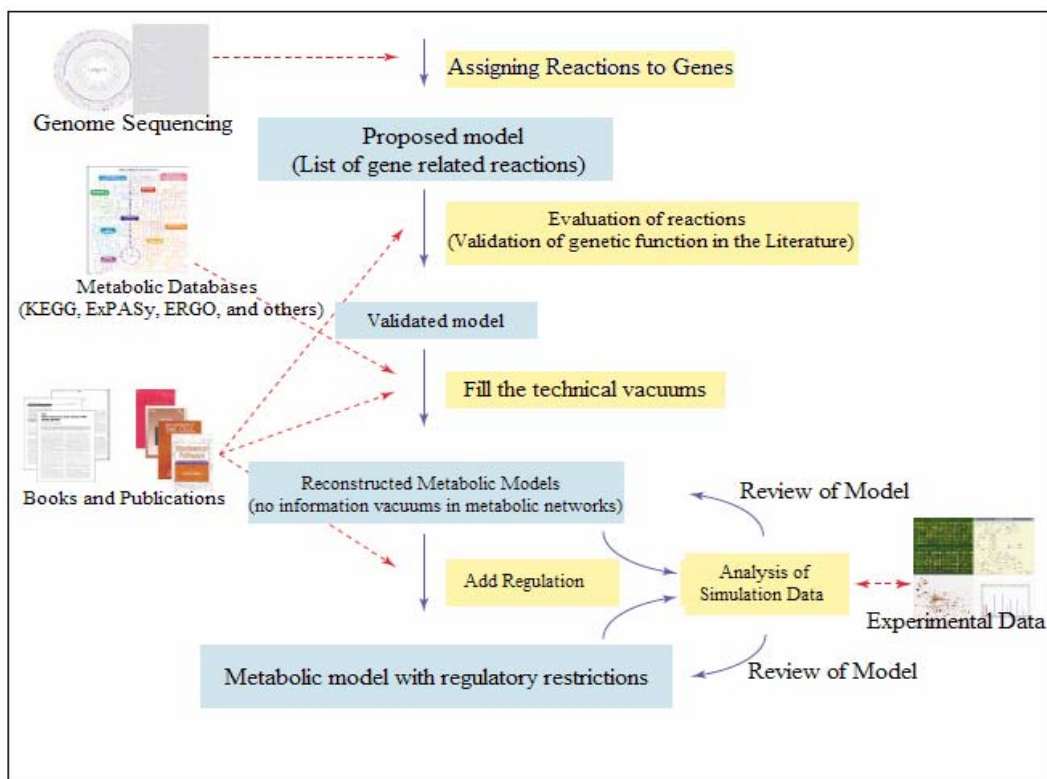
Figure 12. Process of metabolic model reconstruction. (Adapted from Borodina and Nielsen, 2005)

This approach was used by Grimbs et al. (2007), to represent the metabolic network of the human erythrocyte. In particular, transitions to instability, occurring via a loss of a stable steady state, were previously argued to play a crucial role in senescence and metabolic collapse of erythrocytes and may act as a primary signal for cell removal in patients with hemolytic anemia. While usually an investigation of such transitions necessitates the construction of explicit kinetic models, this approach allows the draw of quantitative conclusions about the stability of metabolic states in response to an increased ATP demand, occurring under conditions of osmotic or mechanic stress. The authors demonstrated that different metabolic states, each satisfying the flux balance equation and thermodynamic constraints, can nonetheless show drastic differences in the ability to ensure stability and maintain metabolic homeostasis.

DISCUSSION AND CONCLUDING REMARKS

There is a growing interest in analyzing the topological organization of metabolic networks in order to make inferences about their dynamic and functional properties. Although we must emphasize those mathematical models by themselves do not allow conclusions to be drawn about the dynamic behavior of biological systems. The translation of metabolic network graphs into mathematical expressions does permit the formal specification of the dynamic capabilities of the system and the identification of targets for biotechnological modifications (Steuer, 2007).

Functional behavior emerges from the nonlinear interactions between genes, proteins and metabolites within metabolic and regulatory networks. It has also been argued that the kinetic study of the network of enzymatic reactions (fluxomics) is key for the study of the dynamics of metabolic networks. The quantification of metabolic flows using isotopomers- 13Cbased metabolic flux analysis- (Rousu et al. 2003; Sauer, 2006), the probabilistic graphical models (bayesian, continuous differential kinetic, Markovian, master-slave synchronization) (Weitzke and Ortoleva, 2003; Wodak et al. 2004; Ao, 2005; Pécou, 2005) and random Boolean network models (Kauffman, 1969) have tried to resolve this dilemma, although some contradictions do exist as the minimum amount of information required and the relevance of the modularity of the network (Stumpf and Wiuf, 2005; Ingram et al. 2006; Solé and Valverde, 2006; Hormozdiari et al. 2007).

In this review several analytical approaches have been examined, the first model, although qualitative, characterizes the topological properties of metabolic networks without making any inferences about their dynamic behavior. The second approach explicitly describes the dynamics of the system once the kinetic parameters of each component are known, although this may be a disadvantage in cases where the systems are incomplete or large, due to the high computational costs of the calculations. Lastly, semi-quantitative models that permit inferences about the stability and dynamics of the

system without the need for an exhaustive knowledge of every constituent element were described, although chemical or functional restrictions may be required in order to obtain reasonable solutions. Although these approaches have been successful, they are limited in regard the application of regulatory parameters (feedback loops). In short, these models require iterative processes for the incorporation of new information and the validation of this information using experimental data, thus permitting the evolution and improvement of these mathematical tools (Figure 12).

Although the approaches described here have both advantages and limitations for the study of the functionality and evolution of biological systems, all of the models reviewed agree that robustness is a common property of biological networks. The robustness of biological systems lies in the persistence of their functionality in the face of perturbations such as the removal of network links or nodes. In other words, robustness represents the capacity of the systems to maintain their stability in spite of changes to either their internal or external environment.

A mathematical definition of robustness (R) of a system (s) measured as a function of the property (a) could be, according to Kitano (2007):

$$R_{a,p}^{s} = \int_{p} \Psi(p) D_{a}^{s}(p) dp$$

Where the function $\Psi(p)$ represents the probability of the perturbation of space p and $D_{a}^{s}(p)$ an evaluation function of (p) that determines the extension and degree of the perturbation, defined as:

$$D_{a}^{s}(p) = \begin{cases} 0, \ p \in A \subset P \\ f_{a}(p) / f_{a}(0), \ p \in P \setminus A \end{cases}$$

Both metabolic flow and kinetic analyses permit the quantitative evaluation of the redistribution of the flows when enzymatic reactions are perturbed and to clarify the implications of unstable states and oscillations around a stationary state. The dynamic stability of the stationary state is important for homeostatic sustainability and robustness against disturbances. It would be illuminating if, in the future, efforts were directed towards the elucidation of this type of behavior in biological systems.

The reconstruction of metabolic networks offers a mathematical tool that permits the recuperation of information inherent to their context, within the Systems Biology approach, thus allowing an approximation to the understanding of the cellular behavior, based on the quantity of information about metabolism accumulated from the advances that have been made in the fields of genomics, transcriptomics and metabolomics (Karp et al.

2002). Thus, mathematical models and simulations can be used as a guide to the design and improvement of biotechnological production systems (Sweetlove and Fernie, 2005).

Finally, different approaches have been established in how the network is controlled, in this review we assume the metabolic fluxes, which means that each pathway has a rate limited step controlling the flux by only enzymes *i.e.* if a step is rate limiting it is expected that overexpression of the corresponding enzyme will give an increase in the flux proportional to the increase in expression. In a recent review Morandini (2009) propose a different view, in which post-translational modification of enzymes and feed-back inhibition control the homeostasis of metabolism rather that the flux control and then he proposes more extensive use of control elements, like transcription factor and metabolic shorcuts for plant biotechnology. We do not see any contradiction in both models, because to manipulate the flux control it is necessary to work with the cell circuitry elements, the problem to be solved is where in the network modification need to be applied, as it is stated in the work of Rischer et al. (2006) in *C. roseus*.

REFERENCES

ALM, Eric and ARKIN, Adam P. Biological networks. *Current Opinion in Structural Biology*, April 2003, vol. 13, no. 2, p. 193-202.

ALMAAS, Eivind and BARABÁSI, Albert-László. Power laws in biological networks. In: KOONIN, Eugene V.; WOLF, Yuri I. and KAREV, Georgy P. eds. *Power Laws, Scale-Free Networks and Genome Biology*. Springer US, 2006, p. 1-11.

ALON, U. Biological networks: the tinkerer as an engineer. *Science*, 2003, vol. 301, no. 5641, p. 1866-1867.

ALTEROVITZ, Gil and RAMONI, Marco. Discovering biological guilds through topological abstraction. *AMIA Annual Symposium Proceedings Archive*, 2006, vol. 2006, p. 1-5.

AMARAL, L.A.N.; SCALA, A.; BARTHÉLÉMY, M. and STANLEY, H.E. Classes of small-world networks. *Proceedings of the National Academy of Sciences of the United States of America*, October 2000, vol. 97, no. 21, p. 11149-11152.

AO, P. Metabolic network modeling: including stochastic effects. *Computers and Chemical Engineering*, October 2005, vol. 29, no. 11-12, p. 2297-2303.

BAPAT, R. Linear algebra: eigenanalysis. In: KENNETH, R.; MICHAELS, J.; GROSS, J.; GROSSMAN, J. and SHIER, D. eds. *Handbook of Discrete and Combinatorial Mathematics*. CRC Press, 2000, p. 440-452. ISBN 0-8493-0149-1.

BARABÁSI, Albert-László; ALBERT, Réka and JEONG, Hawoong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, October 1999, vol. 272, no. 1-2, p. 173-187.

BARABÁSI, Albert-László. Emergence of scaling in complex networks. In: BORNHOLDT, S. and SCHUSTER, H.G. eds. *Handbook of Graphs and Networks: From the Genome to the Internet*. H.G. WILEY-VCH GmbH and Co., 2003, p. 69-82.

BARABÁSI, Albert-László and OLTVAI, Zoltán N. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, February 2004, vol. 5, no. 2, p. 101-113.

BAXTER, C.J.; LIU, J.L.; FERNIE, A.R. and SWEETLOVE, L.J. Determination of metabolic fluxes in a non-steady-state system. *Phytochemistry*, August-September 2007, vol. 68, no. 16-18, p. 2313-2319.

BORODINA, Irina and NIELSEN, Jens. From genomes to *in silico* cells via metabolic networks. *Current Opinion in Biotechnology*, June 2005, vol. 16, no. 3, p. 350-355.

BRAUNSTEIN, Alfredo; MULET, Roberto and PAGNANI, Andrea. Estimating the size of the solution space of metabolic networks. *BMC Bioinformatics*, May 2008, vol. 9, p. 240-252.

BRINKMEIER, M. and SCHANK,T. *Network Analysis*. Springer-Verlag; Berlin, 2005. vol. 3418. 471 p. ISBN 3-540-24979-6.

BRO, Christoffer; REGENBERG, Birgitte; FÖRSTER, Jochen and NIELSEN, Jens. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metabolic Engineering*, March 2006, vol. 8, no. 2, p. 102-111.

CILIBERTO, Andrea; CAPUANI, Fabrizio and TYSON, John J. Modeling networks of coupled enzymatic reactions using the total quasi-steady state approximation. *PLoS Computational Biology*, March 2007, vol. 3, no. 3, p. e45-e55.

CRONWRIGHT, Garth R.; ROHWER, Johann M. and PRIOR, Bernard A. Metabolic control analysis of glycerol synthesis in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, September 2002, vol. 68, no. 9, p. 4448-4456.

DOMIJAN, Mirela and KIRKILIONIS, Markus. Graph theory and qualitative analysis of reaction networks. *Networks and Heterogenous Media*, June 2008, vol. 3, no. 2, p. 295-322. DOROGOVTSEV, S. and MENDEZ; J.F.F. Evolution of networks. *Advances in Physics*, June 2002, vol. 51, no. 4, p. 1079-1187.

DOROGOVTSEV, S.; MENDES, J. and SAMUKHIN, A. Principles of statistical mechanics of uncorrelated random networks. *Nuclear Physics B*, September 2003, vol. 666, no. 3, p. 396-416.

EDWARDS, Jeremy S. and PALSSON, Bernard O. Metabolic flux balance analysis and the *in silico* analysis of *Escherichia coli* k-12 gene deletions. *BMC Bioinformatics*, 2000, vol. 1, p. 1.

FELL, D. Metabolic control analysis: a survey of its theoretical and experimental development. *Biochemical Journal*, 1992, vol. 286, p. 313-330.

FERNÁNDEZ, Paul and SOLÉ, Ricard V. Graphs as models of large-scale biochemical organization. In: BONCHEV, Danail and ROUVRAY, Dennis H. eds. *Complexity in Chemistry, Biology and Ecology*. Springer US, 2005, p. 155-189.

FIEHN, O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics*, June 2001, vol. 2, no. 3, p. 155-168.

FÖRSTER, Jochen; FAMILI, Iman; FU, Patrick; PALSSON, Bernard and NIELSEN, Jens. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Research*, February 2003, vol. 13, no. 2, p. 244-253.

GRANT, Chris M. Metabolic reconfiguration is a regulated response to oxidative stress. *Journal of Biology*, January 2008, vol. 7, no. 1, p. 1-4.

GRIMBS, Sergio; SELBIG, Joachim; BULIK, Sascha; HOLZHÜTTER, Hermann-Georg and STEUER, Ralf. The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks. *Molecular Systems Biology*, November 2007, vol. 3, no. 146, p. 1-12.

GROSS, Liza. Charting the interplay between structure and dynamics in complex networks. *PLoS Biology*, October 2005, vol. 3, no. 11, p. e369.

GUTÍERREZ, Rodrigo A.; LEJAY, Laurence V.; DEAN, Alexis; CHIAROMONTE, Francesca; SHASHA, Dennis E. and CORUZZI, Gloria M. Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in *Arabidopsis*. *Genome Biology*, 2007, vol. 8, no. 1, p. R7.

HOLME, Petter and HUSS, Mikael. Discovery and analysis of biochemical subnetwork hierarchies. In: GAUGES, R.; KUMMER, U.; PAHLE, J. and ROST, U. eds. 3rd *Workshop on Computation of Biochemical Pathways and Genetic Networks: European Media Lab Proceedings*. Berlin, Logos, 2003, vol. 12, p. 3-9.

HORMOZDIARI, Fereydoun; BERENBRINK, Petra; PRŽULJ, Nataša and SAHINALP, S. Cenk. Not all scale-free networks are born equal: the role of the seed graph in PPI network evolution. *PLoS Computational Biology*, July 2007, vol. 3, no. 7, p. e118-e130.

HUMPHRIES, Mark D. and GURNEY, Kevin. Network "Small-World-Ness": a quantitative method for determining canonical network equivalence. *PLoS ONE*, April 2008, vol. 3, no. 4, p. e0002051.

INGRAM, Piers J.; STUMPF, Michael P.F. and STARK, Jaroslav. Network motifs: structure does not determine function. *BMC Genomics*, May 2006, vol. 7, p. 108.

ISHII, Nobuyoshi; ROBERT, Martin; NAKAYAMA, Yoichi; KANAI, Akio and TOMITA, Masaru. Toward large-scale modeling of the microbial cell for computer simulation. *Journal of Biotechnology*, September 2004, vol. 113, no. 1-3, p. 281-294.

JEONG, H.; TAMBOR, B.; ALBERT, R.; OLTVAI, Z. and BARABÁSI, A.-L. The large-scale organization of metabolic networks. *Nature*, October 2000, vol. 407, p. 651-654.

JIN, Guangxu; ZHANG, Shihua; ZHANG, Xiang-Sun and CHEN, Luonan. Hubs with network motifs organize modularity dynamically in the protein-protein interaction network of yeast. *PLoS ONE*, November 2007, vol. 2, no. 11, p. e1207-e1216.

JØRGENSEN, Kirsten; RASMUSSEN, Anne Vinther; MORANT, Marc; NIELSEN, Allan Holm; BJARNHOLT, Nanna; ZAGROBELNY, Mika; BAK, Søren and LINDBERG, MØLLER, Birger. Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinion in Plant Biology*, June 2005, vol. 8, no. 3, p. 280-291.

KAHLEM, Pascal and BIRNEY, Ewan. Dry work in a wet world: computation in systems biology. *Molecular Systems Biology*, July 2006, vol. 2, p. 40-44.

KARP, Peter D.; PALEY, Suzanne and ROMERO, Pedro. The pathway tools software. *Bioinformatics*, February 2002, vol. 18, no. 90001, p. 225-232.

KAUFFMAN, Kenneth J.; PRAKASH, Purusharth and EDWARDS, Jeremy S. Advances in flux balance analysis. *Current Opinion in Biotechnology*, October 2003, vol. 14, no. 5, p. 491-496.

KAUFFMAN, S.A. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, March 1969, vol. 22, no. 3, p. 437-467.

KITANO, Hiroaki. Towards a theory of biological robustness. *Molecular Systems Biology*, September 2007, vol. 3, no. 137, p. 1-7.

LANZENI, Stefano; MESSINA, Enza and ARCHETTI, Francesco. Graph models and mathematical programming in biochemical network analysis and metabolic engineering design. *Computers and Mathematics with Applications*, March 2008, vol. 55, no. 5, p. 970-983.

LAZEBNIK, Yuri. Can a biologist fix a radio? Or, what I learned while studying apoptosis. *Cancer Cell*, September 2002, vol. 2, no. 3, p. 179-182.

LIBOUREL, Igor G.L. and SHACHAR-HILL, Yair. Metabolic flux analysis in plants: from intelligent design to rational engineering. *Annual Review of Plant Biology*, June 2008, vol. 59, p. 625-650.

LIEBERMEISTER, Wolfram and KLIPP, Edda. Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theoretical Biology and Medical Modelling*, December 2006, vol. 3, p. 41-54.

MATSUDA, Fumio; WAKASA, Kyo and MIYAGAWA, Hisashi. Metabolic flux analysis in plants using dynamic labeling technique: application to tryptophan biosynthesis in cultured rice cells. *Phytochemistry*, August-September 2007, vol. 68, no. 16-18, p. 2290-230.

MIKULECKY, Donald C. The circle that never ends: can complexity be made simple? In: BONCHEV, Danail and ROUVRAY, Dennis. eds. *Complexity in Chemistry, Biology, and Ecology*. D. H. Springer US, 2005, p. 97-153.

MORANDINI, Piero. Rethinking metabolic control. *Plant Science*, April 2009, vol. 176, no. 4, p. 441-451.

MORGAN, John A. and RHODES, David. Mathematical modeling of plant metabolic pathways. *Metabolic Engineering*, January 2002, vol. 4, no. 1, p. 80-89.

NEWMAN, M.E.J. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, January 2001, vol. 98, no. 2, p. 404-409.

NIELSEN, Jens. Principles of optimal metabolic network operation. *Molecular Systems Biology*, July 2007, vol. 3, p. 126-127.

OLTVAI, Zoltán N. and BARABÁSI, Albert-László. System biology: life's complexity pyramid. *Science*, October 2002, vol. 298, no. 5594, p. 763-764.

PALSSON, B. *Systems Biology: Properties of reconstructed networks*. 1st ed. New york, Cambridge University Press, 2006, 322 p. ISBN 978-0-511-14705-0.

PATIL, Kiran Raosaheb and NIELSEN, Jens. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proceedings of the National Academy of*

Sciences of the United States of America, 2005, vol. 102, no. 8, p. 2685-2689.

PÉCOU, Elizabeth. Splitting the dynamics of large biochemical interaction networks. *Journal of Theoretical Biology*, February 2005, vol. 232, no. 3, p. 375-384.

PURVIS, Jeremy E.; RADHAKRISHNAN, Ravi and DIAMOND, Scott L. Steady-state kinetic modeling constrains cellular resting states and dynamic behavior. *PLoS Computational Biolology*, March 2009, vol. 5, no. 3, p. e1000298.

QI, Yuan and GE, Hui. Modularity and dynamics of cellular networks. *PLoS Computational Biology*, December 2006, vol. 2, no. 12, p. e174-e183.

RALSER, Markus; WAMELINK, Mirjam; KOWALD, Axel; GERISCH, Birgit; HEEREN, Gino; STRUYS, Eduard A.; KLIPP, Edda; JAKOBS, Cornelis; BREITENBACH, Michael; LEHRACH, Hans and KROBITSCH, Sylvia. Dynamic rerouting of the carbohydrate flux is key to counteracting oxidative stress. *Journal of Biology*, December 2007, vol. 6, no. 4, p. 10.1-10.18.

REED, Jennifer L. and PALSSON, Bernhard O. Thirteen years of building constraint-based *in silico* models of *Escherichia coli*. *Journal of Bacteriology*, May 2003, vol. 185, no. 9, p. 2692-2699.

RIOS-ESTEPA, Rigoberto and LANGE, Bernd Markus. Experimental and mathematical approaches to modeling plant metabolic networks. *Phytochemistry*, August-September 2007, vol. 68, no. 16-18, p. 2351-2374.

RISCHER, Heiko; ORESIC, Matej; SEPPANEN-LAAKSO, Tuulikki; KATAJAMAA, Mikko; LAMMERTYN, Freya; ARDILES-DIAZ, Wilson; VAN MONTAGU, Marc C.E.; INZE, Dirk; OKSMAN-CALDENTEY, Kirsi-Marja and GOOSSENS, Alain. Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. *Proceedings of the National Academy of Sciences of the United States of America*, April 2006, vol. 103, no. 14, p. 5614-5619.

ROSCHER, Albrecht; EMSLEY, Lyndon; RAYMOND, Philippe and ROBY, Claude. Unidirectional steady state rates of central metabolism enzymes measured simultaneously in a living plant tissue. *Journal of Biological Chemistry*, September 1998, vol. 273, no. 39, p. 25053-25061.

ROUSU, Juho; RANTANEN, Ari; MAAHEIMO, Hannu; PITKÄNEN, Esa; SAARELA, Katja and UKKONEN, Esko. A method for estimating metabolic fluxes from incomplete isotopomer information. In: PRIAMI, C. ed. *Computational Methods in Systems Biology*, Berlin, Springer, January 2003, p. 88-103.

SAITO, Kazuki. Plant functional genomics based on integration of metabolomics and transcriptomics: toward plant systems biology. In: NAKANISHI, Shigetada; KAGEYAMA, Ryoichiro and WATANABE, Dai. eds. *Systems Biology: The Challenge of Complexity*, Berlin, Springer Science, 2009, p. 135-142.

SAUER, Uwe. Metabolic networks in motion: ^{13}C-based flux analysis. *Molecular Systems Biology*, November 2006, vol. 2, no. 62, p. 1-10.

SCHILLING, Christophe H.; SCHUSTER, Stefan; PALSSON, Bernhard O. and HEINRICH, Reinhart. Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. *Biotechnology Progress*, 1999, vol. 15, no. 3, p. 296-303.

SCHMIDT, S. and DANDEKAR, T. Metabolic pathways. In: COLLADO-VIDES, J. and HOFESTÄDT, R. eds. *Gene Regulation and Metabolism: Postgenomic Computational Approaches*. Cambridge, MIT Press, 2002, p. 251-272.

SCHWENDER, Jörg; OHLROGGE, John and SHACHAR-HILL, Yair. Understanding flux in plant metabolic networks. *Current Opinion in Plant Biology*, June 2004, vol. 7, no. 3, p. 309-317.

SCHWENDER, Jörg. Metabolic flux analysis as a tool in metabolic engineering of plants. *Current Opinion in Biotechnology*, April 2008, vol. 19, no. 2, p. 131-137.

SCOTT, A. *The nonlinear univers*e. Berlin, Springer-Verlag, 2007, 364 p.

SEGRÈ, Daniel; VITKUP, Dennis and CHURCH, George M. Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America*, November 2002, vol. 99, no. 23, p. 15112-15117.

SOLÉ, Ricard V. and VALVERDE, Sergi. Are network motifs the spandrels of cellular complexity? *Trends in Ecology and Evolution*, August 2006, vol. 21, no. 8, p. 419-422.

SOLÉ, Ricard V. and VALVERDE, Sergi. Spontaneous emergence of modularity in cellular networks. *Journal of the Royal Society Interface*, January 2008, vol. 5, no. 18, p. 129-133.

STEUER, Ralf. Computational approaches to the topology, stability and dynamics of metabolic networks. *Phytochemistry*, August-September 2007, vol. 68, no. 16-18, p. 2139-2151.

STEUER, Ralf; GROSS, Thilo; SELBIG, Joachim and BLASIUS, Bernd. Structural kinetic modeling of metabolic networks. *Proceedings of the National Academy of Sciences of the United States of America*, August 2006, vol. 103, no. 32, p. 11868-11873.

STEUER, Ralf; NUNES-NESI, Adriano; FERNIE, Alisdair R.; GROSS, Thilo; BLASIUS, Bernd and SELBIG, Joachim. From structure to dynamics of metabolic pathways: application to the plant mitochondrial TCA cycle. *Bioinformatics*, March 2007, vol. 23, no. 11, p. 1378-1385.

STROGATZ, S. *Nonlinear dynamics and chaos*. New York, Perseus Books Publishing, 1994, 498 p. ISBN 0-201-54344-3.

STEPHANOPOULOS, Gregory N.; ARISTIDOU, Aristos A. and NIELSEN, Jens. *Metabolic engineering: principles and methodologies*. Academic Press, 1998, 725 p. ISBN 0-12-666260-6.

STUMPF, Michael P. and WIUF, Carsten. Sampling properties of random graphs: the degree distribution. *Physical Review*, September 2005, vol. 72, no. 3, p. 036118.

SWEETLOVE, Lee J. and FERNIE, Alisdair R. Regulation of metabolic networks: understanding metabolic complexity in the systems biology era. *New Phytologist*, October 2005, vol. 168, no. 1, p. 9-24.

TAKEMOTO, Kazuhiro and OOSAWA, Chikoo. Modeling for evolving biological networks with scale-free connectivity, hierarchical modularity, and disassortativity. *Mathematical Biosciences*, August 2007, vol. 208, no. 2, p. 454-468.

TOMITA, Masaru; HASHIMOTO, Kenta; TAKAHASHI, Kouichi; SHIMIZU, Thomas Simon; MATSUZAKI, Yuri; MIYOSHI, Fumihiko; SAITO, Kanako; TANIDA, Sakura; YUGI, Katsuyuki; VENTER, J. Craig and HUTCHISON III, Clyde A. E-CELL: software environment for whole-cell simulation. *Bioinformatics*, 1999, vol. 15, no. 1, p. 72-84.

VARMA, Amit and PALSSON, Bernhard O. Metabolic flux balancing: basic concepts, scientific and practical use. *Nature Biotechnology*, 1994, vol. 12, no. 10, p. 994-998.

VITKUP, Dennis; KHARCHENKO, Peter and WAGNER, Andreas. Influence of metabolic network structure and function on enzyme evolution. *Genome Biology*, 2006, vol. 7, no. 5, p. R39-R48.

WAGNER, A. How to reconstruct a large genetic network from n genes perturbations in fewer than n^2 easy steps. *Bioinformatics*, December 2001, vol. 17, no. 12, p. 1183-1197.

WATTS, Duncan J. and STROGATZ, Stephen H. Collective dynamics of "small-world" networks. *Nature*, June 1998, vol. 393, no. 6684, p. 440-442.

WEITZKE, Elizabeth L. and ORTOLEVA, Peter J. Simulating cellular dynamics through a coupled transcription, translation, metabolic model. *Computational Biology and Chemistry*, October 2003, vol. 27, no. 4-5, p. 469-480.

WODAK, Shoshana J.; CASTURA, Jeffrey and ORSI, Chris. Integrative bioinformatics: making sense of the networks. *Drug Discovery Today: Technologies*, October 2004, vol. 1, no. 2, p. 179-187.

XIA, Yu; YU, Haiyuan; JANSEN, Ronald; SERINGHAUS, Michael; BAXTER, Sarah; GREENBAUM, Dov; ZHAO, Hongyu and GERSTEIN, Mark. Analyzing cellular biochemistry in terms of molecular networks. *Annual Review of Biochemistry*, March 2004, vol. 73, p. 1051-1087.

YANG, Chin-Rang; SHAPIRO, Bruce E.; MJOLSNESS, Eric D. and HATFIELD, G. Wesley. An enzyme mechanism language for the mathematical modeling of metabolic pathways. *Bioinformatics*, March 2005, vol. 21, no. 6, p. 774-780.

YANG, Yea-Ting; BENNETT, George N. and SAN, Ka-Yiu. Genetic and metabolic engineering. *Electronic Journal of Biotechnology*, 1998, vol. 1, no. 3, p. 134-141.

Note: Electronic Journal of Biotechnology is not responsible if on-line references cited on manuscripts are not available any more after the date of publication.
Supported by UNESCO / MIRCEN network.

19