Modular Response Analysis of Cellular Regulatory Networks

Frank J. Bruggeman†‡, Hans V. Westerhoff†§, Jan B. Hoek‡ and Boris N. Kholodenko*†

†Department of Molecular Cell Physiology, Biocentrum Amsterdam, de Boelelaan 1087, NL-1081 HV Amsterdam, The Netherlands, ‡Department of Pathology, Anatomy and Cell Biology, Thomas Jefferson University, 1020 Locust St., Philadelphia, PA19107, U.S.A. and §Department of Mathematical Biochemistry, Biocentrum Amsterdam, de Boelelaan 1087, NL-1081 HV Amsterdam, The Netherlands

(Received on 19 December 2001, Accepted in revised form on 28 May 2002)

The sheer complexity of intracellular regulatory networks, which involve signal transducing, metabolic, and genetic circuits, hampers our ability to carry out a quantitative analysis of their functions. Here, we describe an approach that greatly simplifies this type of analysis by capitalizing on the modular organization of such networks. Steady-state responses of the network as a whole are accounted for in terms of intermodular interactions between the modules alone; processes operating solely within modules need not be considered when analysing signal transfer through the entire network. The intermodular interactions are quantified through (local) response coefficients which populate an interaction map (matrix). This matrix can be derived from a biochemical or molecular biological analysis of (macro) molecular interactions that constitute the regulatory network. The approach is illustrated by two examples: (i) mitogenic signalling through the mitogen-activated protein kinase cascade in the epidermal growth factor receptor network and (ii) regulation of ammonium assimilation in Escherichia coli.

© 2002 Elsevier Science Ltd. All rights reserved.

Introduction

The reductionist methods of biochemistry and molecular biology have yielded an impressive amount of data and knowledge of molecular mechanisms and topology of regulatory networks (e.g. Kohn, 1999). However, we are still far from understanding the physiological properties that arise from the interactions that constitute these networks, i.e. from the networking itself. This lack of understanding holds back both the comprehension of molecular cell physiology as a whole and the ability to proceed in experimental analysis thereof. Progress is hampered by the vastness of these networks, which ranges from hundreds to thousands of macromolecular species and their complexes per cell (Kohn, 1999). Although it will become possible to integrate networks of thousands of components numerically, it is not obvious that this by itself will enhance understanding by the human mind. How should analysis proceed in such cases? Here we emphasize the reduction of complexity through modularization.
Modularity appears to be a recurrent phenomenon in the topology of cellular networks (Lauffenburger, 2000; Hartwell et al., 1999). Biochemists and cell biologists treat parts of the cell machinery as semi-autonomous modules. Examples range from Krebs’ cycle and glycolysis in the classical biochemistry of mass-flow networks to kinase–phosphatase couples and mitogen-activating protein kinase (MAPK) pathways in signal transduction networks. A corollary of this modularity is that the number of the interactions between the molecules from within one module to other modules is low as compared to the number of interactions inside the modules. The apparent intractability of large intracellular networks may be overcome by the simplification that results from such a modular treatment.

How can we benefit from the natural modular organization of the regulatory networks in attempting to facilitate understanding? At least two major strategies are followed here. The first strategy is that of further drawing the map of intracellular interactions through directed experimentation and bioinformatics (Marcotte et al., 2001). A second strategy defines what types of modules there are. Accordingly, “levels”, i.e. parts of networks that do not share significant mass flow with other parts (Kahn & Westerhoff, 1991; Kholodenko et al., 1997; Hofmeyr & Westerhoff, 2001) have been identified as modules. The level-type organization is relevant both for transcription and regulation of metabolism (Westerhoff et al., 1989) and for signal transduction (Kahn & Westerhoff, 1991). With respect to intracellular regulation, we now expect that regulatory networks may become more understandable by the implementation of such quantitative methods, which treat segments of regulatory networks as “black-box” modules and relate intrinsic properties of the modules and their interactions to the emergent properties of the entire regulatory network. The latter can be achieved with metabolic control analysis, which provides a suitable theoretical framework.

Metabolic control analysis (MCA) is a type of sensitivity analysis of biochemical systems that makes the explicit connection between global systemic functioning and local kinetic properties of a biochemical network (Kacser & Burns, 1973; Heinrich & Rapoport, 1974). Initially, MCA proved to be successful in determining control exerted by individual enzymes on fluxes and metabolite concentrations and was applied to problems in metabolic regulation and engineering (Fell, 1997). Subsequently, MCA was extended to modular networks composed of signalling, genetic and metabolic subnetworks (Westerhoff & Van Dam, 1987; Brown et al., 1990; Westerhoff et al., 1989; Kahn & Westerhoff, 1991; Schuster et al., 1993; Brand, 1996; Kholodenko et al., 1997, 1998; Hofmeyr & Westerhoff, 2001). The novelty of the method presented here is that it simplifies modular treatment of regulatory networks by focusing only on a subset of intermediates that engage in cross-talk between modules, referred to as communicating intermediates.

In the present paper, we will generalize and extend the method initiated for response analysis of simple networks by Kholodenko et al. (1997) to encompass regulatory networks of any complexity. The response of any modular network to changes in their physiological signals will be explicated in terms of the intramodular responses and intermodular interactions of communicating intermediates that interconnect modules. The resulting expressions are the mathematical equivalent of the (macro)molecular interaction maps that result from molecular–biological studies (Kohn, 1999). Note that the analysis applies to any kind of modular interaction, e.g. protein–protein interactions, etc. Since the number of communicating intermediates is much less than the total number of intermediates, this modular description drastically simplifies the analysis. The approach presented here will be illustrated for two examples of regulatory networks in living cells.

**Mathematical Notation and Theoretical Background**

Scalars are denoted by italic fonts, e.g. \( s \), and matrices by bold capital letters, e.g. \( \mathbf{N} \). Subscripts, as in \( \mathbf{N}_{r,c} \), indicate that \( \mathbf{N} \) has \( r \) rows and \( c \) columns. Similarly, \( \mathbf{s} \), indicates a vector or
(r × 1)-matrix with r elements. The matrix entries are designated as, e.g., \( m_{r,c} \) or \( v_r \), in which \( r \) and \( c \) indicate the row and column index. A partitioned diagonal matrix \( N \) is denoted by \( d g(N) \) with the \( N_i \) referring to the \( i \)-th rectangular matrix entry, whereas all other off-diagonal entries are rectangular null matrices.

The dynamics of a biochemical system with \( m \) intermediates and \( r \) reactions is given by the mass balance equations (Heinrich & Schuster, 1996):

\[
\dot{x}(t, p) = N_m v(x(t, p), p).
\]  

(1)

The concentration vector, stoichiometric matrix and rate vector are denoted by \( x \), \( N \) and \( v \), respectively. An \((i,j)\)-th entry to the matrix \( N \) is the stoichiometric coefficient of the \( i \)-th intermediate of the \( j \)-th reaction. The rates, elements of vector \( v \), depend on the intermediate concentrations, \( x \), and parameters, \( p \), such as \( K_m \), \( V_{max} \), \( K_{eq} \), etc.

On the time-scale characteristic for signal transduction and metabolism, the total amount of interconvertible forms remains constant for each protein. Additionally, certain metabolite pools might be conserved. Such moieties conservation relationships decrease the number of independent intermediates, causing linear dependencies between the rows of the stoichiometric matrix (Reder, 1988). Assuming that the first \( m_0 \) elements of the vector \( x \) are linearly independent, one can decompose \( x \) into \( m_0 \) linearly independent intermediates, \( x \), and \( m - m_0 \) linearly dependent intermediates, \( x^D \), as

\[
x = \begin{bmatrix} x_{m_0} \\ x^D \\ x_{m-m_0} \end{bmatrix}.
\]  

(2)

Accordingly, the first \( m_0 \) linearly independent rows of \( N \) form a reduced stoichiometric matrix, \( N_{m_0}^{0} \), related to \( N \) through the link matrix \( L \) (Reder, 1988):

\[
N = L_{m,m_0} N^0 = \begin{bmatrix} I_{m_0,m_0} \\ L'_{m-m_0,m_0} \end{bmatrix} N^0.
\]  

(3)

Together with eqn (1), the latter two relationships reflect that the dynamics of the total system is fully contained in the dynamics of the independent intermediates, \( x \):

\[
\dot{x}(t, p) = N^0 v(x(t, p), x^D(x(t, p)), p).
\]  

(4)

Under steady-state conditions, the time derivative in eqn (4) equals the null vector.

The steady-state response of the system intermediates to a parameter change is determined by the implicit differentiation of the steady-state equation with respect to \( p \) (Reder, 1988):

\[
\left( \frac{\partial x}{\partial p} \right)_{m_0,r} = - \left( N^0 \left( \frac{\partial v}{\partial s} \right)_{r,m} \right)^{-1} N^0 \left( \frac{\partial v}{\partial p} \right)_{r,r},
\]  

(5)

where

\[
\frac{\partial v}{\partial s} = \begin{bmatrix} \frac{\partial v}{\partial x} & \frac{\partial v}{\partial x^D} \end{bmatrix}
\]

is referred to as the matrix of unscaled elasticities. The matrix \( N^0 (\partial v/\partial s) L \) is the Jacobian matrix of eqn (4).

The operator that transforms parameter sensitivities of individual rates, \( \partial v/\partial p \), into global responses of the concentrations, \( \partial x/\partial p \), is called the matrix of (unscaled) concentration control coefficients, \( G^x_v \):

\[
G^x_v = - \left( N^0 \frac{\partial v}{\partial s} L \right)^{-1} N^0.
\]  

(6)

Post-multiplying \( G^x_v \) with the kernel (null-space), \( K_{r,r-m_0} \), of the stoichiometric matrix \( N \), \( (N^0 K = 0) \) one obtains

\[
G^x_v K = 0_{r,r-m_0}.
\]  

(7)

This relationship is known as the summation theorem (Reder, 1988). The connectivity theorem for concentration control is obtained from eqn (7) after post-multiplication with \( (\partial v/\partial s) L \) (Reder, 1988):

\[
G^x_v \frac{\partial v}{\partial s} L = -I_{m_0,m_0}.
\]  

(8)
In MCA control coefficients are used in a scaled format. The scaled expressions for the various control and response matrices used here can be found in the Appendix.

Results

MODULAR CONTROL ANALYSIS

The modules to be considered here are of the level type, i.e. they do not share net mass flow. Furthermore, we assume that the moiety-conservation relationships either do not involve intermediates that belong to different modules, or the concentrations of such complexes can be neglected when compared to the total amounts of the corresponding moieties. Each module can be assigned an intramodular control concentration matrix which is solely a function of the intrinsic properties of the module. Intramodular control coefficients are defined as the effect of a perturbation of a process within a module on the steady-state value of an intramodular flux or concentration under the condition that all intermodular interactions are kept constant. Operationally, this means that the module is isolated from the system or the fluxes of intermediates that belong to different modules, or the concentrations of such complexes can be neglected when compared to the total amounts of the corresponding moieties. Each module can be assigned an intramodular control concentration matrix which is solely a function of the intrinsic properties of the module. Intramodular control coefficients are defined as the effect of a perturbation of a process within a module on the steady-state value of an intramodular flux or concentration under the condition that all intermodular interactions are kept constant. Operationally, this means that the module is isolated from the system or the remainder of the system is kept constant upon perturbation of a particular parameter of this module.

Here, we analyse a modular regulatory network that contains \( r \) reactions, \( m \) metabolites and \( n \) modules. The structural properties of this network can be expressed in terms of partitioned block-diagonal matrices, i.e. \( \mathbf{N} = dg(\mathbf{N}_i) \), \( \mathbf{K} = dg(\mathbf{K}_i) \) and \( \mathbf{L} = dg(\mathbf{L}_i) \), which, respectively, contain as their \( i \)-th block-diagonal entry the matrices, \( (\mathbf{N}_i)_{m_i, r_i} \), \( (\mathbf{L}_i)_{m_i, m_0} \) and \( (\mathbf{K}_i)_{r_i, r_0} \), describing each module. These structural properties of the modular network guarantee that the intramodular control properties are invariant, i.e. independent of whether the module is isolated or embedded in the entire regulatory network. Modular control analysis expresses the global control properties of the whole network in terms of the intramodular control properties and intermodular interactions. From eqn (6) and the modular properties of the network, it follows that the global concentration control matrix can be written as

\[
\mathbf{G}^r_i = -\left( dg(\mathbf{N}^0_i) \frac{\partial \mathbf{v}}{\partial \mathbf{s}} dg(\mathbf{L}_i) \right)^{-1} \frac{dg(\mathbf{N}^0_i)}{dg(\mathbf{L}_i)}
\]

\[
= - \left( \left( \frac{dg(\mathbf{N}^0_i)}{dg(\mathbf{L}_i)} \right)^{-1} \right) \frac{dg(\mathbf{N}^0_i)}{dg(\mathbf{L}_i)} \right)^{-1}
\]

\[
- \left( \frac{dg(\mathbf{N}^0_i)}{dg(\mathbf{L}_i)} \right)^{-1}
\]

This equation was first derived by Kahn & Westerhoff (1991) and contains the intramodular concentration control matrices of the modules, i.e. \( \Gamma^r_i = -(\mathbf{N}_i^0)^{-1} \).

\[
\mathbf{G}^r_i = -\left( \frac{dg(\Gamma^r_i)}{dg(\mathbf{L}_i)} \right)^{-1} \frac{dg(\Gamma^r_i)}{dg(\mathbf{L}_i)}
\]

\[
= - \mathbf{r}^{-1} \frac{dg(\Gamma^r_i)}{dg(\mathbf{L}_i)}
\]

The matrix \( \mathbf{r} \) quantifies intermodular interactions that form the interaction map of a regulatory network. A non-zero \( (i,j) \)-th entry of matrix \( \mathbf{r} \) indicates that there is a (direct) interaction between modules \( i \) and \( j \) through one or more communicating intermediates of module \( j \) that affect the rate of one or more processes in module \( i \):

\[
(r^r_{ij})_{m_i, m_0} = (\Gamma^r_i)_{m_i, m_0} \left( \frac{\partial \mathbf{v}_j}{\partial \mathbf{s}} \right)_{r_j, m_j} (\mathbf{L}_j)_{m_j, m_0}
\]

It follows from the connectivity theorem [eqn (8)] that if \( i = j \) in eqn (11); \( r^r_{ii} = -I_{m_i, m_0} \). Hence, the interaction map, \( \mathbf{r} \), has the following structure:

\[
\mathbf{r} = \begin{bmatrix}
-\mathbf{I} & \cdots & \mathbf{r}^r_{1m_0} \\
\vdots & \ddots & \vdots \\
\mathbf{r}^r_{nm_0} & \cdots & -\mathbf{I}
\end{bmatrix}_{m_m, m_0}
\]

MODULAR RESPONSE ANALYSIS

This section analyses global responses of intermediates in terms of the structural and
kinetic properties of a modular network. For simplicity, we assume here that the signal, e.g. hormone, growth factor, cytokine or neurotransmitter, only affects a single module directly. In the example section we will consider how to deal with signals that affect multiple modules. The global effect of a signal perturbation on all intermediates can be obtained from eqn (10) through post-multiplication with the vector $\dot{v}/\dot{p}$ that quantifies the effect of the signal perturbation on the individual process rates. Post-multiplication of the global control matrix, $G^r_i$, with $\dot{v}/\dot{p}$ results in a global response vector, $R^r_p$, whereas post-multiplication of the diagonal matrix of intramodular control coefficients results in a vector of intramodular responses, $r^r_p$:

$$R^r_p = -(r)^{-1}r^r_p. \quad (13)$$

The vector $r^r_p$ only contains non-zero entries at the rows that refer to the intramodular responses of the intermediates of module $j$ that are perturbed as a result of a change in one of its signals, i.e. parameters $p_j$. If multiple parameters (signals) are perturbed, i.e. if multiple signals are activated, one obtains response matrices instead of response vectors (see section on MAPK network).

**Reduction of Complexity: Analysis in Terms of Communicating Intermediates**

The preceding analysis considered mechanistically all network intermediates which may hamper its application to large regulatory networks. Here, we simplify the analysis by taking into consideration only those intermediates that mediate interactions between modules, as opposed to the intermediates that merely operate within the modules. The former and the latter intermediates are referred to as communicating and introvert intermediates, respectively. Focusing on the responses of the communicating intermediates, the analysis treats modules as black boxes.

The independent metabolite vector of each module can be decomposed into two subvectors, $x^\text{int}_i (i = 1, \ldots, m^\text{int}_i)$ and $x^\text{com}_i (i = 1, \ldots, m^\text{com}_i)$, for the introvert ($x^\text{int}_i$) and communicating intermediates ($x^\text{com}_i$) of module $i$, respectively. This decomposition results in

$$x_i = \begin{bmatrix} x^\text{com}_i \\ x^\text{int}_i \end{bmatrix}. \quad (14)$$

After an equivalent reordering of the rows and columns of eqn (13) (which involves the decomposition of both the stoichiometric and link matrices of individual modules) one obtains

$$\begin{bmatrix} R^\text{com}_p \\ R^\text{int}_p \end{bmatrix} = - \begin{bmatrix} r^\text{com}_i \\ r^\text{int}_i \end{bmatrix}^{-1} \begin{bmatrix} r^\text{com}_p \\ r^\text{int}_p \end{bmatrix}. \quad (15)$$

A perturbation to introvert intermediates of module $i$ neither affects the steady state of that module nor any process rates outside of module $i$. Hence, $r^\text{com}_i = 0$ and $r^\text{int}_i = -1$. Consequently, the interaction map matrix in the last equation simplifies to

$$r = \begin{bmatrix} r^\text{com}_i \\ r^\text{int}_i \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}. \quad (16)$$

Taking the inverse of this matrix one obtains for eqn (15):

$$\begin{bmatrix} R^\text{com}_p \\ R^\text{int}_p \end{bmatrix} = \begin{bmatrix} (-r^\text{com}_i)^{-1} \\ 0 \end{bmatrix} \begin{bmatrix} r^\text{com}_p \\ r^\text{int}_p \end{bmatrix}. \quad (17)$$

The intermodular response matrix for the communicating intermediates, $r^\text{com}_{ij}$, has as its $(i,j)$-th matrix entry:

$$r^\text{com}_{ij} = (G^r_i)^{m^\text{com}_j \rightarrow m^\text{com}_i} \left( \frac{\partial x^\text{com}_j}{\partial x^\text{com}_i} \right)_{r_i, n^\text{com}_i, n^\text{com}_j}$$

if $i \neq j$; $r^\text{com}_{ii} = L^\text{com}_{ij}, n^\text{com}_i, n^\text{com}_j$. \quad (18)

Importantly, the responses of the communicating species to a parameter change, i.e. $R^\text{com}_p$, can be determined from eqn (15) without any reference to the responses of the introvert intermediates to this parameter change as

$$R^\text{com}_p = -(r^\text{com}_p)^{-1}r^\text{com}_p. \quad (19)$$
The last equation illustrates that our analysis can treat network modules as black boxes that interact through communicating intermediates, whereas the introvert intermediates need not be considered. We refer to the $r_{\text{com}}$-matrix as the reduced interaction map. The reduced interaction map makes it possible to determine the global response of any communicating intermediate to signal changes or perturbations without having to consider all intermediates of the network. According to the definition, the reduced interaction map can be determined by following a reductionist strategy, i.e. by determining the effect of communicating intermediates on each module independently.

By way of eqn (17), modular response analysis (MRA) suggests that the interaction map of a regulatory network can also be determined in a holistic fashion, i.e. by comparing global responses and (“local”) intramodular responses to external perturbations. In the latter strategy multiple parameters, e.g. signals, inhibitors or activators, should be perturbed. If one perturbs as many parameters as the number of chosen communicating intermediates and each individual parameter directly affects the rate(-s) in only one module, one obtains from eqn (17) the following:

$$r_{\text{com}} = -dg(r_p^{\text{com}})(R_p^{\text{com}})^{-1}.$$ 

For each module, vector $p$ should involve $m_{\text{com}}^i$ independent parameters. Here, $m_{\text{com}}^i$ is the number of communicating independent intermediates of module $i$.

**Modular Response Analysis (MRA) for Two Cellular Networks**

The application of MRA to modular networks is illustrated with the following two examples: (i) the eukaryotic MAPK pathway, which transduces signals emanating from growth factor receptors, such as epidermal growth factor (EGF) receptor (Fig. 1), and (ii) the ammonium assimilatory network of *Escherichia coli*, which fine-tunes the ammonium assimilatory flux with respect to the internal nitrogen and carbon statuses (Fig. 2).

MAPK pathways function as central “switchboards” of signal processing in many different eukaryotic species and participate in the regulation of a large number of important physiological processes, such as differentiation, mitosis and apoptosis (Schaeffer & Weber, 1999). MAPK cascades consist of three sequential levels. At each level, phosphorylation and dephosphorylation take place, catalysed by a kinase of a preceding level and a phosphatase of a given level, respectively. The starting point of the cascade is the activation (phosphorylation) of MAPKKK, which subsequently doubly phosphorylates MAPKK. This process is followed by dual phosphorylation of MAPK by doubly phosphorylated MAPKK.

The MAPK-pathway that is activated by EGF consists of RAF (MAPKKK), MEK (MAPKK) and ERK (MAPK) and their phosphatases (Kolch, 2000). Here we will decompose this network conceptually into four modules (Fig. 1). The first module consists of the membrane-associated signalling processes which start with the activation of the EGF receptor (Re) through binding of EGF. This results in the dimerization and auto-phosphorylation of the EGF-receptor and finally, the activation of RAF through protein activation route that involves the following different signalling intermediates: Shc, Gbr2, SOS and Ras (Kholodenko et al., 1999).

Phosphorylated RAF is the communicating intermediate of the first module that interacts with the second module through phosphorylation of MEK. The second and third modules include different phosphorylated species of MEK and ERK, respectively, and their associated phosphatases. Doubly phosphorylated MEK acts as the communicating intermediate of module 2 and interacts with module 3 through phosphorylation of ERK. Doubly phosphorylated ERK is the communicating intermediate of the third module.

*In vivo*, this signal transduction pathway is embedded in a complex network of feedback interactions that act either directly or indirectly via other signalling modules (Kholodenko, 2000; Bhalla & Iyengar, 1999; Kolch, 2000). Here, two regulatory feedbacks that start from activated
ERK will be discussed. The first to be considered is a negative feedback from phosphorylated ERK to SOS (Hu & Bowtell, 1996; Langlois et al., 1995; Cherniak et al., 1995). Activated ERK phosphorylates SOS on multiple serine/threonine residues hereby causing its inactivation, with a concomitant decline in the activity of Ras and subsequently of RAF. The second loop is an indirect positive feedback of phosphorylated ERK to RAF via the additional fourth module. Molecular processes to be considered in this module involve ERK-induced activation of phospholipase A2 (PLA2) and subsequent activation of protein kinase C (PKC) by arachidonic acid (AA) (Bhalla & Iyengar, 1999). Activated PKC acts as the communicating intermediate of the fourth module and interacts with the first module through phosphorylation of RAF, which leads to the activation of the MAPK pathway (Bhalla & Iyengar, 1999).

The individual modules as shown in Fig. 1(A) all contain multiple introvert intermediates, which do not have to be considered in the determination of the global responses of the communicating intermediates [eqn (17)]. The reduced interaction map of the same
network is depicted in Fig. 1(B) and is given by the following matrix:

\[
\begin{pmatrix}
-1 & 0 & r_R^R & r_P^R \\
0 & -1 & 0 & 0 \\
r_M^R & 0 & -1 & 0 \\
0 & 0 & r_E^P & -1
\end{pmatrix},
\]  

(18)

where \( R, M, E \) and \( P \), respectively, denote phosphorylated RAF, doubly phosphorylated MEK, doubly phosphorylated ERK and activated PKC. A comparison of Figs 1(A) and (B) illustrates the reduction of the number of intermediates that have to be considered with the introduction of the reduced interaction map. Additionally, experimental determination of the local responses becomes within reach because only four communicating intermediates have to be monitored experimentally.

A biologically relevant output of the MAPK-pathway is doubly phosphorylated ERK. The
response of activated ERK to a change in the EGF concentration or an effector of module 4 (e.g., phorbol-12-myristate-13-acetate (PMA) (Braz et al., 2002)) can be determined from eqn (17).

\[
\begin{bmatrix}
R^R_{EGF} & R^R_{PMA} \\
R^M_{EGF} & R^M_{PMA} \\
R^E_{EGF} & R^E_{PMA} \\
R^P_{EGF} & R^P_{PMA}
\end{bmatrix} = -
\begin{bmatrix}
-1 & 0 & r^R_E & r^R_P \\
0 & r^M_E & -1 & 0 \\
0 & 0 & r^P_M & -1
\end{bmatrix}^{-1}
\]

(19)

One promise of MRA is that it should facilitate understanding of the response of the network to an external signal in terms of the interactions between the modules only. The latter being given by the \(r\)'s, one can now demonstrate this for the effect of EGF on the phosphorylation state of ERK by multiplying the third row of the inverse matrix of eqn (19) with the vector on the far right in the same equation:

\[
R^E_{EGF} = \frac{\text{MAPK}}{1 - \frac{r^M_E r^R_M r^E_{EGF}}{r^M_E r^R_M r^P_E + r^M_E r^R_P r^E_{EGF}}}.
\]

(20)

The global responses of linear cascades are known to be the product of the responses along the cascade (Small & Fell, 1990; Kholodenko et al., 1997). The last equation shows this for the MAPK cascade without feedbacks where the response of doubly phosphorylated ERK to EGF is a mathematical product of (local) responses of successive levels, \(r^M_E r^R_M r^R_E\). Similarly, the responses along each feedback loop are also represented by products of local responses, e.g., \(r^E_M r^M_P r^P_E\) for the MAPK–SOS loop. The two feedback loops modify the response of the MAPK pathway by subtracting responses along the loops in the denominator of eqn (20). Positive feedbacks tend to amplify the signal by decreasing the denominator of the response coefficient expression whereas negative feedbacks tend to attenuate the response. Although MRA is a powerful tool to analyse steady-state responses, in some cases information about dynamic system properties can also be derived. For instance, an increase in positive feedback can decrease the denominator in eqn (20) to zero, which would formally correspond to infinitely large responses. However, it can be shown that when the denominator in eqn (20) becomes zero, the system passes through a saddle-node bifurcation, which dramatically changes the dynamics of the MAPK pathway resulting in the appearance (or disappearance) of bistability (Kholodenko, 2000). Regulatory networks that portray bistable behavior have been experimentally constructed (Gardner et al., 2000) and observed (Bagowski & Ferrell, 2001; Ferrell & Machleder, 1998). Interestingly, an increase in the strength of a negative feedback, which increases the denominator of eqn (20) and leads to a further signal attenuation, brings about a Hopf bifurcation resulting in sustained MAPK oscillations (Kholodenko, 2000).

In vivo or in silico determination of the local responses, i.e. the interaction map, will show to what extent various regulatory loops contribute to the global response of MAPK. This will be a significant step in the understanding of signal transduction in living cells.

AMMONIUM ASSIMILATION NETWORK

As eukaryotes, prokaryotes comprise complex regulatory networks. One example is the regulatory cascade affecting ammonia assimilation in *E. coli* through the regulation of the activity of glutamine synthetase. This microorganism dedicates two enzyme routes to ammonia assimilation, i.e. glutamate dehydrogenase (GDH) and glutamine synthetase (GS) plus glutamate:2-oxoglutarate aminotransferase (GOGAT) (Rhee et al., 1989). At low concentrations of ammonia, its assimilation is primarily through GS. The
physiological response of E. coli to a decrease in the ammonia concentration is a gradual shift in enzyme activities, i.e., from GDH to GS (Magasanik, 1996). Here, the regulatory network underlying this behavior will be decomposed into four modules [Fig. 2(A)].

The first module involves all small metabolites of E. coli metabolism, where glutamine (G) and α-ketoglutarate (K) are considered as communicating intermediates. In addition, module 1 comprises the enzymes involved in the corresponding metabolic reactions, except for GS which is confined to module 2. The role of the remaining three modules is to regulate the activity of ammonia assimilation through the activity of GS. The second module harbors the expression of the gene encoding GS and translation of its mRNA into the functional protein and, in addition, the enzyme ATase. ATase can progressively inactivate GS by adenylylation of each of the 12 subunits of GS, and it can also catalyse the reverse process. The adenylylation state of GS (GSA) will be considered as the communicating intermediate of module 2. The third module contains the enzyme UTase/UR and all the different species of the trimeric protein PII, including PIIKG1 (P) and PIIUMP3KG3 (PU) [see Fig. 2(A) and (B)]. P and PU serve as communicating intermediates of module 3, which together with glutamine directly affect the activity of ATase (Ninfa et al., 2000; Jaggi et al., 1997). UTase/UR can (de-)uridylylate each subunit of PII as function of the concentrations of glutamine (Ninfa & Atkinson, 2000). Additionally, the various PII species are able to bind α-ketoglutarate. The fourth module is composed of the two-component signalling network NRI/NRII, and it senses the concentration of PIIKG1 (Ninfa et al., 2000). The communicating intermediate of this module is the doubly phosphorylated NRI (T) which is a transcriptional activator of the gene encoding GS in module 2. As a whole, this four-modular network is able to semi-intelligently regulate the ammonium assimilation rate as function of the nitrogen and carbon status (Bruggeman et al., 2000). Importantly, modules 1 and 2 have two communicating intermediates and all modules contain multiple introvert intermediates.

In matrix form, the interaction map of this network is given by

$$
\mathbf{r}^\text{com} = \begin{bmatrix}
-1 & 0 & 0 & r^P_G & r^K_K & 0 \\
0 & -1 & 0 & r^PU_G & r^K_K & 0 \\
0 & 0 & r^G_G & -1 & 0 & 0 \\
0 & 0 & r^K_K & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1
\end{bmatrix} .
$$

A perturbation in the ammonium concentration (N) directly affects module 1 and propagates subsequently through the total network yielding the following global responses of the communicating intermediates:

$$
\begin{bmatrix}
R^P_N \\
R^{PU}_N \\
R^{GA}_N \\
R^G_N \\
R^K_N \\
R^T_N
\end{bmatrix}
= \begin{bmatrix}
-1 & 0 & 0 & r^P_G & r^K_K & 0 \\
0 & -1 & 0 & r^PU_G & r^K_K & 0 \\
0 & 0 & r^G_G & -1 & 0 & 0 \\
0 & 0 & r^K_K & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1
\end{bmatrix}^{-1}
\begin{bmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}
$$

The global response of glutamine to a perturbation in the ammonium concentration can now be expressed as follows:

$$
R^G_N = \frac{1 - \sum K^T_K r^K_N}{1 - \sum r^K_G - \sum r^K_K}.
$$
where

$$\sum \Pi^G = \rho^G + \rho^G \Pi^G + \rho^G \Pi^G \Pi^G \Pi^G$$

$$\sum \Pi^K = \rho^K + \rho^K \Pi^K + \rho^K \Pi^K \Pi^K$$

To reduce complexity further one can zoom-out the network by decreasing the number of modules and communicating intermediates. If modules 2–4 were condensed into one large module, the global response of $G$ to a perturbation in $N$ would be given by eqn (23) with $\sum \Pi^G = \rho^G + \rho^G \Pi^G$ and $\sum \Pi^K = \rho^K + \rho^K \Pi^K$. This illustrates a stepwise approach to modular decomposition of regulatory networks: instead of facing the full complexity of the regulatory network, one can first treat modules 2–4 condensed into one black box module before allowing for additional intermodular interactions and modules. As illustrated by eqns (20) and (23), MRA allows for unravelling the individual contributions of modules and regulatory loops to global network responses. However, the extent of contribution depends strongly on a particular state of the system which is determined by the signal concentration.

In the preceding sections, the analysis has been limited to concentrations of network intermediates. However, in the ammonium assimilatory network a physiologically interesting response is the response of the glutamine synthetase flux, $J_{GS}$, to a change in the ammonium concentration. This response is obtained from the implicit differentiation of $J_{GS} = J_{GS}(G(N), K(N), N)$ with respect to $N$, as follows:

$$R_{N}^{j_{GS}} = r_{N}^{j_{GS}} + r_{G}^{j_{GS}} R_{G}^{j_{GS}} + r_{K}^{j_{GS}} R_{K}^{j_{GS}}.$$

After substitution of the global response of glutamine, $R_{N}^{j_{GS}}$, [eqn (22)] to a change in the ammonium concentration, one obtains

$$R_{N}^{j_{GS}} = r_{N}^{j_{GS}} + r_{G}^{j_{GS}} (1 - \sum \Pi^G) r_{N}^{j_{GS}} + r_{K}^{j_{GS}} (1 - \sum \Pi^K) R_{N}^{j_{GS}}.$$

This equation illustrates that the global response is determined by the intramodular (local) response ($r_{N}^{j_{GS}}$) and the effect of the entire regulatory network through responses of communicating intermediates. Equation (25) demonstrates that $\alpha$-ketoglutarate may eliminate the effect of glutamine on $J_{GS}$, when the term $1 - \sum \Pi^G$ equals zero. Similarly, glutamine may counteract the influence of $\alpha$-ketoglutarate, if $1 - \sum \Pi^K = 0$.

Hence, MRA in combination with a mechanistic model or experimental studies of ammonia assimilation in *E. coli* facilitates the detailed analysis of the cross-talk between the carbon and the nitrogen statuses of the cell and its effect on the regulation of the ammonium-assimilation flux through GS.

**Discussion**

This paper extends earlier work on the application of metabolic control analysis to quantitative studies of control and regulation of modular cellular networks (Westerhoff & Van Dam, 1987; Brown *et al.*, 1990; Brand, 1996). More specifically, modular response analysis is a generalization in matrix form of the method derived by Kholodenko *et al.* (1997, 1998) to encompass modular cellular networks of any complexity. The subsets of intermediates and reactions that form network modules are determined by the non-zero blocks that occur in the block diagonal $L$ and $K$ matrices, respectively (Heinrich & Schuster, 1996). This condition makes it possible to relate global responses of the entire network to intramodular and intermodular responses. A violation of this condition occurs if the total concentration of a communicating intermediate is conserved in module $i$, and its fraction sequestered in a process within module $j$ (affected by that intermediate) cannot be neglected with respect to its total...
concentration (Kholodenko et al., 1994; Fell & Sauro, 1990). This would violate the block-diagonal condition for the L-matrix for module i and therefore, modules i and j should be treated as one module in modular response analysis. This problem is closely related to the violation of the condition of parameter independence of the control coefficient of the parameter choice in MCA (Kholodenko et al., 1995).

Intermodular interactions via communicating intermediates interconnect modules and form the interaction network. The interaction map that results from protein–protein and effector–enzyme interactions is quantitatively expressed in eqn (12) and refers to the responses of all intermediates in the network. The central equation in modular response analysis is the reduced interaction map, i.e. eqn (17), which quantifies the intermodular responses in terms of the communicating species, and offers a drastically reduced description of the global responses of the network. Additionally, MRA offers a method for a rational approach to the description of global responses through the sequential decomposition of modules into submodules as illustrated for the ammonium assimilatory network of E. coli. Thus, in combination with experimental or mechanistic-modelling studies MRA provides a flexible tool to deal with the complexity and modularity of regulatory networks.

At present, modular response analysis also has a number of strong limitations. First of all, it addresses the response of the network to small changes in the signal. In reality, the signal may change quite substantially before an appreciable change in a downstream signalling protein is detected, making MRA a first-order approximation of what actually happens. Such an approximation should be useful for continuous signal transducers. In some systems, signal transduction appears to be switchable, e.g. bistable signal transduction networks, whereas other systems appear to be more continuous, but this may still result from experiments analysing population rather than single cells (Bagowski & Ferrell, 2001; Ferrell & Machleder, 1998). A second limitation is that modular response analysis is only applicable to network that prevail in steady states. Some systems may indeed move between steady states as the signal is altered in strength, but other networks undergo transient activation and subsequent perfect adaptation or down-regulation. For instance, activation of PC12 cells by EGF appears to result in a partial adaptation of the level of activated ERK whereas sustained activation was observed for stimulation with nerve growth factor (NGF) (Marshall, 1995; Brightman & Fell, 2000). Although, MRA may be extended to encompass time-dependent systems in the future, we argue that there remain many unresolved (quasi) steady-state signalling phenomena to be addressed with modular response analysis. The latter does not only hold true for eukaryotic signalling but possibly even more for prokaryotic networks engaged in genetic and metabolic regulation where it is known that steady states represent functional states, e.g. ammonium-assimilatory network. Note that the occurrence of bistability in signalling networks can also be addressed with MRA (a manuscript in preparation, see also the MAPK cascade example earlier).

There is an increasing interest in the analysis of genetic networks. The latter are broadly defined so as to include interactions through proteins and even metabolites (Westerhoff et al., 1989). To the extent that different coupled gene-expression systems do not share mass flow, modular response analysis should be applicable to such systems as well. Application of system analysis tools, e.g. modelling and MRA, to studies of modular biochemical networks, e.g. signal transduction networks or genetic networks, may become important in the current era of quantitative system biology. Modular response analysis offers a stepwise method to increase the complexity of a biochemical system analysis gradually by sequentially adding more modules and, thereby, increasing the number of interactions, before ultimately facing the full complexity of the living cell.

This work was supported by the National Institutes of Health Grant GM59570 and by DARPA Bio-Comp.

REFERENCES


**APPENDIX**

Often the control, response and elasticity coefficients are expressed in scaled form in MCA. Here, the main equations of modular response analysis will be expressed in their scaled form. For the independent scaled control coefficients matrix, one obtains

\[ sc \frac{\Delta G}{x} = \left( \frac{dg(x)}{C_0} \right)^{-1} \frac{\Delta G}{x} dg(J), \]

where \( J \) represents the steady-state flux vector. The scaled global response matrix is given by

\[ sc R_p^x = \left( \frac{dg(x)}{C_0} \right)^{-1} R_p^x dg(p). \]

The scaled interaction map:

\[ sc r = \left( \frac{dg(x)}{C_0} \right)^{-1} r dg(x). \]

The scaled local response matrix:

\[ sc r_p^x = \left( \frac{dg(x)}{C_0} \right)^{-1} r_p^x dg(p). \]

The scaled global response matrix for communicating intermediates:

\[ sc R_p^{x,\text{com}} = \left( \frac{dg(x^{\text{com}})}{C_0} \right)^{-1} R_p^{x,\text{com}} dg(p). \]

The scaled reduced interaction map:

\[ sc r_{x,\text{com}} = \left( \frac{dg(x^{\text{com}})}{C_0} \right)^{-1} r_{x,\text{com}}^{\text{com}} dg(x^{\text{com}}). \]

The scaled local response matrix for communicating intermediates:

\[ sc r_p^{x,\text{com}} = \left( \frac{dg(x^{\text{com}})}{C_0} \right)^{-1} r_p^{x,\text{com}} dg(p). \]