Direct Response Analysis in cellular signalling networks

N. Shayeghi a, T. Ng b, A.C.C. Coolen a,b,*

a Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK
b Randall Division of Cell and Molecular Biophysics, King’s College London, New Hunts House, London SE1 1UL, UK

A R T I C L E   I N F O

Article history:
Received 9 December 2011
Received in revised form
17 March 2012
Accepted 2 April 2012
Available online 9 April 2012

Keywords:
Systems biology
Signal transduction
Direct Response Analysis

A B S T R A C T

Direct Response Analysis is a general computational tool for quantifying direct functional interactions between components in cellular signalling systems from experimental perturbations and measurements alone. This paper aims to reveal the biological meaning of the direct response coefficients obtained upon applying DRA to simple Michaelis–Menten type proteomic and gene regulatory systems. These systems describe dimer formation and dissociation, protein production and decay, and transcription. We derive explicit formulae for the direct response coefficients in terms of biochemical reaction rates, and clarify the potential and limitations of the DRA method. We find that response coefficients are strongly asymmetric, and that they balance persistent characteristics of reactions (e.g., the ratios of on- and off rates) against the time-scales over which these reactions act; fast reactions give stronger response coefficients. The direct interactions between protein species, caused by dimer formation, are effectively negative. We illustrate our results with numerical simulations.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Macroscopic biological complexity often results from non-linear interaction among microscopic components, the properties of which appear quite different in context and in isolation. Over the years, many studies have aimed at understanding how components of large and complex biological systems are integrated and coordinated into a functioning macroscopic unit. Nowadays such work is called Systems Biology (Murray, 2001; Citri and Yarden, 2006; Sorkin and Goh, 2008; Ciaccio et al., 2010). The systems approach is particularly fruitful in cellular biology, where the different responses induced in cells (such as proliferation, survival, and motility) are thought to be regulated by signalling via complicated cascades of protein reactions, initiated by ligand binding. Network-level analysis is needed to infer the signal transduction through changes in the activities and concentrations of multiple proteins and their complexes; this information then needs to be mapped to biochemical reaction equations, in order to understand the signalling dynamics quantitatively. Such research typically involves two areas of mathematical analysis (Chen et al., 2009). The first focuses on statistical and topological properties of the interaction network, and inference of biologically relevant modules (Casey et al., 2007). The second on the dynamical analysis of biochemical protein reactions (Murray, 2001; Cornish-Bowden, 2004), usually represented by coupled ordinary differential equations.

If we seek to model mathematically the various cellular biochemical reactions we face serious obstacles. Even if we know the relevant signalling pathways in full, such models typically have many parameters that must be measured or estimated, such as initial protein concentrations and reaction rate constants (e.g. complex formation and enzymatic reactions Cornish-Bowden, 2004). Moreover, often we do not know the pathways in full, and need mathematical analysis in conjunction with biological experiments in order to elucidate the pathway structure of the system. We then require effective means to incorporate experimental data into the model in a manner that reproduces cell specific responses while taking into account the uncertainty in the precise biochemistry of the signalling and the inability of experiments to measure all relevant processes (Chen et al., 2009; Aldridge et al., 2006). The purpose of Direct Response Analysis (DRA) (Kholodenko et al., 1997, 2002; Bruggeman et al., 2002; Sontag et al., 2004; Cho et al., 2005; Andreù et al., 2005; Kholodenko, 2007), proposed by Kholodenko and coworkers, is to achieve this. It is a general protocol with which to extract information on the direct pairwise influence which components in a complex dynamical system exert upon each other (as opposed to influence mediated via other components), from only the measurement of responses of these variables to controlled system perturbations. DRA has generally been referred to in literature as Modular Response Analysis, but we feel that this does not capture sufficiently the essence of the method, which is its ability to separate direct interactions between components from those that are mediated via third parties. However, it is not yet clear what Kholodenko's response coefficients represent exactly in terms of reaction rates and other biophysical parameters of the cellular signalling system from which they are calculated.
In this paper we aim to clarify the meaning of the direct response coefficients that emerge from DRA, by applying DRA to simple models of synthetic proteomic and gene regulation signalling systems. We express the response coefficients explicitly in terms of biochemical parameters (reaction rates), and thereby achieve a better understanding of what these coefficients generally tell us about the underlying biology.

2. Direct Response Analysis

Let us first review briefly the method of Direct Response Analysis (Kholodenko et al., 1997, 2002). We consider a general dynamical system, modelled by a closed set of ordinary differential equations for N dynamical variables \( x = (x_1, \ldots, x_N) \) (such as log-concentrations of proteins and protein complexes, or gene expression levels), and in which \( K \) parameters \( \theta = (\theta_1, \ldots, \theta_K) \) represent time-independent external perturbations (such as RNA interference):

\[
\frac{dx_i}{dt} = f_i(x, \theta), \quad i = 1 \ldots N
\]  

(1)

The functions \( f_i \) are assumed to be smooth, so that their partial derivatives exist. In DRA one assumes also that the system (1) relaxes to a unique equilibrium state\(^1\); the fixed point is written as \( x = x_0(\theta) \), and is to be solved from the fixed-point equations \( f_i(x) = 0 \) for all \( i \). We next define direct response coefficients \( r_{ij} \) for the unperturbed system, and show how these can be calculated upon measuring the system's response to perturbations:

- We first consider the unperturbed system, i.e. \( \theta = 0 \), and write its stationary state simply as \( \mathbf{x}(\mathbf{0}) = \mathbf{x}_0 \). We now imagine a constrained version of (1), where only \( x_i \) evolves via \( dx_i/dt = f_i(x, \mathbf{0}), \) \( x_i \) is kept at \( x_i = x_i^0 + \epsilon \) (with \( \epsilon \) small) and all other components are held at their equilibrium values \( x_k = x_k^0 \). The new stationary value for \( x_i \) will be of the form \( x_i^* = x_i + r_{ij}\epsilon + C(\epsilon^2) \); the factor \( r_{ij} \) measures the direct linear effect of small perturbations in \( x_j \) on \( x_i \) close to the equilibrium state, and is called the direct response coefficient of the interaction \( j \to i \). Its value is determined by the fixed-point equation:

\[
f_i(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, x_i + r_{ij}\epsilon + C(\epsilon^2), \mathbf{x}_{i+1}, \ldots, \mathbf{x}_{N}) = 0
\]  

(2)

Working out this identity, using \( f_i(\mathbf{x}, \mathbf{0}) = 0 \) and the short-hands \( J_{ij} = (\partial f_i(\mathbf{x}, \mathbf{0})/\partial x_j)\big|_{x=x^0} \) for the elements of the Jacobian at the fixed-point, gives \( J_{ij} r_{ij} + J_{ji} k + C(\epsilon^2) = 0 \), and hence

\[
r_{ij} = -J_{ij}/J_{ii}
\]  

(3)

- Calculation of (3) requires knowing \( J_{ij} \) and \( J_{ii} \), i.e. partial derivatives of the functions \( f_i(\mathbf{x}, \theta) \) in (1). However, we often do not know these functions. The next step in DRA is therefore to construct a protocol for measuring all derivatives \( (J_{ij}) \) via perturbations. Differentiation of the general fixed-point \( f_i(\mathbf{x}(\theta), \theta) = 0 \) of (1) with respect to the perturbations gives

\[
\frac{\partial f_i}{\partial x_k} |_{x=x_0(\theta)} = \sum_{j=1}^{N} J_{ik} \delta_{kj} = \left[ \frac{\partial f_i}{\partial \theta_j} \right]_{x=x_0(\theta)}
\]  

(4)

The entries \( M_{ij} \) give the changes in fixed-point values of all components, resulting from small system perturbations; they can in principle be measured or estimated experimentally. The \( L_{ij} \) are generally not known; one could limit oneself to combinations \( (i,j) \) for which one expects to have \( L_{ij} = 0 \), but this brings in dangerous and uncontrolled assumptions. However, if we limit ourselves to perturbations of production rates\(^2\), i.e. \( f_i(\mathbf{x}, \theta) = f_i(\mathbf{x}) + \theta_i \), then \( L_{ij} = \delta_{ij} \), and we can find the \( (J_{ij}) \) by solving

\[
\text{for all } (i,j) : \quad \sum_{k=1}^{N} J_{ik} M_{kj} = -\delta_{ij}
\]  

(5)

These are \( N^2 \) equations for \( N^2 \) unknowns, the solution of which amounts simply to inverting a matrix.

The strength of DRA is that it requires no prior knowledge of the forces \( f_i(\mathbf{x}, \theta) \) in the unperturbed system, which allows it to be used for uncovering these forces. But DRA also has weaknesses. First, it assumes that our system obeys closed equations, and (unless we turn to general perturbations \( \theta \), where we need ad hoc assumptions on which of the \( L_{ij} \) are zero) it demands that we are able to perturb all production rates of all components independently. In a proteomic system we would then have to perturb production rates of individual complexes independently, for assuming closed equations demands that complexes are included in our component list. This cannot be done without violating chemical conservation laws. Second, DRA requires perturbations that are infinitesimally small; in cellular signalling systems this is experimentally very hard. In most applications of DRA so far the perturbations have been gene knockdowns, so the above formulae do not apply. One could try to remedy the situation somewhat by extending the perturbation theory to higher orders, but this has not yet been done and would still rule out knockdown perturbations. In addition, some of the applicability limitations of DRA can be lifted if it is applied in a modular framework, with dynamical variables representing the activity of a module as a whole, as in e.g. Santos et al. (2004), Cho et al. (2005), and Yalamanchili et al. (2006).

Here we are concerned with a different question: even if we succeed in measuring the response coefficients (3), how do we interpret the numbers we find? What do these coefficients mean in terms of biophysical reaction rates, when we apply DRA to cellular signalling systems? Answering these questions will aid our translation of experimental results into practical biological knowledge, which is ultimately the objective of studies that use DRA, such as Santos et al. (2007). To achieve this we calculate the response coefficients (3) for simple mathematical models of proteomic reaction equations and gene regulation.

3. Application to simple proteomic reaction networks

3.1. Derivation of explicit formulae for direct response coefficients

Suppose our cellular signalling system is a simple protein interaction network, described by standard kinetic reaction equations (see e.g. Murray, 2001). If the reactions taking place are hetero-dimer formation and dissociation, production, and decay, and if for simplicity we leave out higher order protein complexes and post-translational modifications, then such a system with n
protein species would be described by equations of the form

\[
\frac{d}{dt} x_i = \sum_{j=1}^n \left( k_{ij}^+ x_j - k_{ij}^- x_i \right) + \gamma_i x_i,
\]

\[
\frac{d}{dt} y_j = k_{ij}^+ x_i - k_{ij}^- x_j
\]

(6)

Here \( x_i \) and \( y_j \) denote the concentrations of unbound protein \( i \), and the hetero-dimer formed by \( i \) and \( j \). The \( k_{ij}^\pm \) are the on/off rates for the formation and dissociation of the hetero-dimers (if \( k_{ij}^\pm = 0 \) such dimers are not forming); \( \gamma_i \) and \( \gamma_j \) represent protein production and decay rates. We consider only hetero-dimers, so \( k_{ij}^\pm = k_{ji}^\pm = 0 \) for all \( i \). Our Eq. (6) obey the conservation laws \( d/dt [x_j + \sum_{i:j} x_{ij}] = \gamma_j - \gamma_i \) for all \( i \). We define \( x_j = x_i \) and \( k_{ij}^\pm = k_{ji}^\pm \) for all \((ij)\). It is trivial to determine the stationary solution of (6), giving

\[
x_i = \frac{\gamma_j}{\gamma_i} x_j = \frac{k_{ji}^+}{k_{ij}^-} \frac{\gamma_j}{\gamma_i} x_j
\]

(7)

It is also trivial to calculate the entries of the Jacobian matrix.

Since our components include two distinct groups of dynamical variables, viz. free protein concentrations labelled by a single index \( i \), and concentrations of complexes labelled by index pairs \((ij)\), it is natural to write the associated forces as

\[
f_i(x) = \sum_{j}^n \left( k_{ij}^+ x_j - k_{ij}^- x_i \right) + \gamma_i x_i = k_{ij}^+ x_j - k_{ij}^- x_i + \gamma_i x_i
\]

(8)

and gives for the nonzero direct response coefficients:

\[
r_{ij} = \frac{\partial f_i}{\partial x_j} = k_{ij}^+ - k_{ij}^- + \gamma_i
\]

(9)

From this, via (3), follow our direct response coefficients of DRA.

They are only defined for distinct pairs of interacting components, so we have \( r_{ij} \) with \( i \neq j \), \( r_{ijk} \), and \( r_{ijkl} \) with \( (ij) \neq (k,l) \).

\[
r_{ij} = -\frac{k_{ij}^+}{\sum k_{ij}^+ x_j + \gamma_i}, \quad r_{ijk} = k_{ij}^- (\delta_{jk} + \delta_{ji}) - k_{ij}^+ (\delta_{ji} + \delta_{jk}), \quad r_{ijkl} = 0
\]

(10)

Equivalently, with (7) we can eliminate the equilibrium concentrations and express the response coefficients fully in terms of biochemical system parameters:

\[
r_{ij} = -\frac{k_{ij}^+}{\sum k_{ij}^+ x_j / \gamma_i}, \quad r_{ijk} = k_{ij}^- (\delta_{jk} + \delta_{ji}) - k_{ij}^+ (\delta_{ji} + \delta_{jk}), \quad r_{ijkl} = 0
\]

(11)

### 3.2. Interpretation and illustration

One understands the signs of the above coefficients. Increasing the amount of an unbound protein \( \ell \), at equilibrium, will increase the formation rate of all dimers involving \( \ell \), and hence reduce the amount of unbound protein of all its binding partners (so \( r_{ij} < 0 \) for all \( i \) with \( k_{ij}^+ > 0 \)), while increasing the dimer concentrations (so \( r_{ij} > 0 \) for all complexes \((ij)\) that involve protein \( \ell \)). Similarly, increasing the amount of the \((ij)\) dimer, in equilibrium, increases the number of \((ij)\) dissociations and hence the concentrations of its constituent proteins; since it has no direct effect on other dimers we should indeed expect \( r_{ijkl} = 0 \). For further interpretation it is useful to write our equations and results in terms of the following quantities: \( x_i = \gamma_i / \gamma_j \) (stationary concentrations of unbound protein), \( y_j = k_{ij}^+ / k_{ij}^- \) (which control the equilibrium balance of bound versus unbound \((ij)\) dimers), \( \tau_{ij} = 1/k_{ij}^- \)

(12)

(13)

and gives for the nonzero direct response coefficients:

\[
r_{ij} = \frac{-\gamma_i}{\sum \tau_{ij} x_j + \gamma_i}
\]

(14)

We now see more clearly how direct response coefficients balance the persistent impact of interactions against the timescales over which the interactions act. For instance, if protein decay is much slower than dimer formation and dissociation, i.e. \( \tau_{ij} = 1/k_{ij}^- \) for all \((ij)\)

(15)

### Fig. 1

The DRA map for a simple synthetic protein interaction system with \( N = 5 \) proteins, described by Eq. (6). All rate constants were drawn randomly from a truncated Gaussian distribution. Complexes are represented by grey squares and unbound proteins by red circles. Arrows represent the direct response coefficients (in direction and magnitude) given in (11); interactions between unbound proteins are always negative and shown in red, and interactions between unbound proteins and dimers are always positive and shown in black (thicker and darker arrows indicate stronger direct responses).
e.g. forward and reverse direct response coefficients $r_{ij}$ and $r_{ji}$, or between direct response coefficients and the type of reaction rate functions one might expect to dominate their values, and how the coefficients are distributed. The results are shown in Figs. 2 and 3. We observe no significant correlation between $r_{ij}$ and $r_{ji}$ (Pearson coefficient < 0.05), see e.g. Fig. 2A. Fig. 2B shows a histogram estimator of the distribution $\rho(r) = N^{-2} \sum_{i<j} \delta[r - r_{ij}]$, which suggests that for the chosen reaction rate statistics the direct response coefficients are, to good approximation, distributed according to a power law. We next plot the coefficients $r_{ij}$ against the associated on-rates $k_{ij}^+$ in Fig. 2C (showing only a modest positive correlation, as expected from (11)), and the ratios $r_{ij}/r_{ji}$ against the corresponding binding ratios $k_{ij}^-/k_{ji}^-$ in Fig. 2D (where no relation is observed). In Fig. 2E and F we observe a strong correlation between the $r_{ij}$ and the protein synthesis/degradation ratios $\theta_i/\gamma_i$ (Pearson coefficient 0.7), but almost no correlation between the reverse response coefficient $r_{ji}$ and $\theta_i/\gamma_i$, i.e. the direct response of protein $i$ to perturbation of protein $j$ depends solely on typical concentration of unbound protein $i$. In Fig. 3 Left and Right we detect a weak inverse linear relation between $r_{ijk}$ and $r_{jki}$, (correlation −0.1). Furthermore, we do observe a roughly linear relation between $r_{ijk} + r_{jki}$ and $k_{ij}^-/k_{ji}^-$. Overall we can deduce that the most important parameters impacting upon direct response coefficients between unbound proteins are the on-rates for complex formation and their synthesis/degradation ratios. The direct responses of free protein concentrations to perturbations in concentrations of complexes (and vice versa) are influenced strongly by the binding ratios of the complexes.

The above regularities are preserved if we use different random initialisations or different distributions of rate constants.

4. Application to a combined proteomic reaction and gene regulation network

4.1. Construction of closed equations for equilibrating gene regulation dynamics

If we wish to apply DRA as a tool for uncovering gene regulation pathways (Yalamanchili et al., 2006; Mettetal et al., 2006; Raingeaud et al., 1996; Williams et al., 2004; Workman et al., 2006), the variables to be perturbed and measured are gene expression levels. Experimental timescales will then have to be sufficiently large for the transcription system to equilibrate; gene regulation is about three orders of magnitude slower than protein reactions. In Eq. (6) the variable $\theta_i$ can be interpreted as representing the expression level of gene $i$, so we now extend (6) with further simple equations that describe the evolution of the $\theta_i$, giving

$$\frac{d}{dt} \theta_i = \sum_{j=1}^{n} \left[ k_{ij}^+ x_j - k_{ij}^- x_i \right] + \theta_i - \gamma_i \theta_i$$

$$\tau \frac{d}{dt} \gamma_i = F(A_i(x), R_i(x)) - \beta_i$$

$$A_i(x) = \sum_j x_j^+ x_{ji} + \sum_{jk} x_j x_{jk}$$

$$R_i(x) = \sum_j x_j x_i + \sum_{jk} x_j x_{jk}$$
Here \( A_i(x) \) and \( R_i(x) \) represent formulae for the cumulative activation and repression of gene \( i \), in which nonnegative parameters \( U_{ji}^+, J_{ji}^+ \) specify the efficacies of the various proteins \( j \) and dimers \( (i, j) \) as activators (+) or repressors (−) of gene \( i \), and with \( J_{ji}^-, J_{ji}^- \) for all \( (i, j) \). If protein \( j \) does not transcribe gene \( i \) then \( J_{ji}^+ = J_{ji}^- = 0 \), etc. \( F(A, R) \) is a nonnegative nonlinear Hill-type function, with \( \partial F / \partial A \geq 0 \), \( \partial F / \partial R \leq 0 \), such as \( F(A, R) = \text{Erf}(A - R) \) or \( F(A) = (A - R)^{3/2} \). The parameter \( \tau \) gives the characteristic timescale for gene regulation dynamics.

Since the proteomic dynamics is much faster than gene regulation, we may assume that (16) reaches equilibrium very fast, and replace the instantaneous values \( x \) in (18) by their equilibria \( x_{\text{eq}} \) (given in (7)). We also choose the simple form \( F(A, R) = F(A - R) \) for our transcription function. All this reduces our equations to the following simple set that involves expression levels only:

\[
\frac{d}{dt} x = G(x) - \bar{y},
\]

\[
U_{ij} = \sum_{jk} U_{ij}^+ + \sum_{jk} U_{ij}^+ \bar{y}_j \bar{y}_k, \quad \bar{y}_j = \frac{J_{ji}^+ - J_{ji}^-}{\partial \bar{y}_j \bar{y}_k}
\]

(19)

(20)

The DRA formalism demands that (19) evolves to a unique fixed-point. Numerical examination, however, reveals that this is not always the case; the system (19) can evolve into oscillatory trajectories. To overcome this problem and get meaningful results from DRA, we have to postulate that, whenever the fixed-points of (19) are locally unstable, asymptotically the system (19) will exhibit persistent oscillations close to these fixed-points. This would allow us to apply DRA after all, in reasonable approximation. In order to find the unstable fixed points of (19) numerically, we modify (19) into

\[
\frac{d}{dt} \bar{y}_j = G_j(x) - \bar{y}_j - \frac{1}{2} \sum_{\ell} \left( F(U_{\ell}(\bar{y})) - \bar{y}_j \right) \frac{\partial^2 F(U_{\ell}(\bar{y}))}{\partial \bar{y}_{\ell}^2}
\]

(21)

All stationary solutions of (19) are also stationary solutions of (21). However, in contrast to (19), the modified process (21) is guaranteed to evolve to a fixed point, i.e. a solution of

\[
\sum_{\ell} \left( F(U_{\ell}(\bar{y})) - \bar{y}_j \right) \frac{\partial^2 F(U_{\ell}(\bar{y}))}{\partial \bar{y}_{\ell}^2} = 0
\]

(22)

since it defines gradient descent on the Lyapunov function

\[
\mathcal{L}(x) = \frac{1}{2} \sum_{\ell} \left( \bar{y}_j - F(U_{\ell}(\bar{y})) \right)^2
\]

(23)

Moreover, the global minima of (23), i.e. the dominant solutions of (22), are indeed the fixed-points of (19).

4.2. Derivation of explicit formulae for direct response coefficients

Since stationary can now be relied upon, we can calculate from (21) after equilibration the direct response coefficient \( R_i \) that quantifies the impact of expression level \( x_j \) on expression level \( x_i \):

\[
R_i = \frac{\partial G_i(x)}{\partial x_j} = \frac{\partial^2 F(U_j(x))}{\partial x_i \partial x_j}
\]

(24)

with

\[
\frac{\partial}{\partial x_j} G_i(x) = \frac{\partial}{\partial x_j} \sum_{\ell} \left( F(U_{\ell}(x)) - \bar{y}_i \right) \left( \delta_{ij} - \frac{\partial}{\partial x_j} F(U_{\ell}(x)) \right)
\]

\[
= \sum_{\ell} \left( \left( \frac{\partial F(U_j(x))}{\partial x_j} - \frac{\partial F(U_j(x))}{\partial x_i} \right) \frac{\partial^2 F(U_{\ell}(x))}{\partial x_i \partial x_j} \right)
\]

(25)

Upon choosing the simple sigmoidal function \( F(U) = \frac{1}{2} + \frac{1}{2} \tanh(U) \), with \( F(U) = \frac{1}{2} - \tanh(U) \), the remaining partial derivatives come out to be

\[
\frac{\partial^2 F(U_j(x))}{\partial x_i \partial x_j} = \frac{1}{2} \left[ 1 - \tanh^2(U_j(x)) \right] \frac{\partial F(U_j(x))}{\partial x_i}
\]

\[
\frac{\partial^2 F(U_j(x))}{\partial x_i \partial x_j} = \frac{1}{2} \left[ 1 - \tanh^2(U_j(x)) \right] \frac{\partial F(U_j(x))}{\partial x_i} - \tanh(U_j(x)) \frac{\partial F(U_j(x))}{\partial x_i} \frac{\partial F(U_j(x))}{\partial x_j}
\]

(26)

(27)

in which we have

\[
\frac{\partial F(U_j(x))}{\partial x_i} = U_{ij} + 2 \sum_k U_{ijk} \delta_{jk}, \quad \frac{\partial^2 F(U_j(x))}{\partial x_i \partial x_j} = 2 U_{ij}
\]

(28)

We should bear in mind that Eq. (21) will generally have more than one fixed-point solution; for each such solution we can calculate corresponding direct response coefficients \( R_i \), and these will generally differ between fixed-points. Fig. 4 shows an example of the resulting DRA map for the (modified) gene regulation system (21), in which genes correspond to circular nodes, and where nonzero direct response coefficients are calculated from (24) and drawn as directed arrows (in shades of red when negative, and shades of black when positive). Transcription rates in this example were generated randomly. Again we observe clearly the general asymmetry of the direct response coefficients.
Direct Response Analysis (DRA) is a versatile computational tool proposed by Kholodenko (2007, and see references therein) for quantifying functional interactions between components in cellular signalling systems from experimental perturbations and measurements alone. In this paper we sought to understand in greater depth the potential and limitations of this method when applied to cellular signalling systems, and the biological meaning of the direct response coefficients that would result from application of DRA to simple Michaelis–Menten type proteomic and gene regulatory systems by deriving explicit formulae for the direct response coefficients in terms of biochemical reaction rates.

As with any mathematical formalism, the validity of DRA rests on various assumptions and conditions. It is hard to see how these can be satisfied in the context of proteomic signalling systems. For instance, DRA demands that one perturbs and measures independently all dynamical variables of a closed dynamical system. In proteomic pathways this implies that one has to include the concentrations of all complexes that can be formed by the proteins under study. But independent perturbation of concentrations of individual complexes without perturbing the concentrations of the constituent proteins is impossible in practice, as it would violate biochemical conservation laws. Second, DRA is a linear perturbation theory, and thus demands that all perturbations are infinitesimally small; yet in practice it is often applied in a context where perturbations take the form of gene knockdowns—violating the basis of the method. It is not clear to what extent the outcome of DRA can be trusted when applied to a situation that fails to meet its conditions for validity. If DRA is applied to gene regulation pathways a further complication arises: the typical equations describing the dynamics of expression levels in gene regulation networks do not always evolve to equilibrium, in conflict with another assumption of DRA. One can circumvent the problem via a slightly modified dynamical process that does go to stationary states, but this requires further assumption that the gene regulation dynamics stabilises into oscillatory trajectories close to fixed-points.

We have expressed the direct response coefficients of DRA explicitly in terms of biochemical rate constants, for simple models of proteomic and gene regulation systems. This led to valuable insight into what exactly is measured by these coefficients. For instance:

1. Response coefficients are strongly asymmetric (i.e. directional). The direct effect of perturbing the concentration of a protein A on that of another protein B with which it interacts is not invariant under exchanging the roles of A and B. The same is true for perturbations of expression levels.
2. Response coefficients balance persistent characteristics of reactions (e.g. the ratios of on- and off rates) against the time-scales over which these reactions act. The direct response mediated by a biochemical reaction increases with the impact of that reaction on steady-state concentrations, but decreases with the characteristic time scale of this reaction (fast reactions give stronger response coefficients).
3. The direct interactions between protein species, caused by dimer formation, are negative. The intuition is that increasing the concentration of a protein leads to increased formation of dimers with its reaction partners; this reduces the amount of free protein of its partners. This is why oscillatory protein concentration dynamics can be linked to short loops of odd length in protein interaction networks (Carlin et al., 2011).

The main reasons for using DRA in the study of real biological systems have traditionally been to understand the system responses that are observed experimentally following specific biological perturbations such as gene knockdown and cell adhesion, and to have a systematic tool for mapping unknown cellular signalling pathways. Our present work might add to these the potential for calculating estimates of actual kinetic parameters from experimental perturbations, either by explicit inversion of identities such as (11) and (24), or by using these identities in a Bayesian estimation framework. The main limitation of our analysis is that it is based on a simplified mathematical representation of protein–protein interaction and gene regulation, as will be any quantitative study. Our results are thus valid to the extent that this representation captures the dominant signalling events in the cell. More specifically: our model does not include molecular complexes of order higher than two, post-translational modifications, stochasticity, or spatial variation of molecular concentrations, and involves simplified equations for transcription.

However, we hope that this paper may aid experimentalists who use the DRA method for uncovering cellular signalling pathways, in exposing the dangers caused by violation of the assumptions underlying DRA, and in providing a better understanding of the biochemical meaning of the direct response coefficients which they would be measuring.

Acknowledgments

This study was supported by the Biotechnology and Biological Sciences Research Council (BBSRC, grant BB/G007160/1) of the United Kingdom.

References


