# RECONSTRUCTION OF CELLULAR SIGNALLING NETWORKS AND ANALYSIS OF THEIR PROPERTIES

*Jason A. Papin*[*‡], Tony Hunter[§], Bernhard O. Palsson[*] and Shankar Subramaniam[*‖]*

Abstract | The study of cellular signalling over the past 20 years and the advent of high-throughput technologies are enabling the reconstruction of large-scale signalling networks. After careful reconstruction of signalling networks, their properties must be described within an integrative framework that accounts for the complexity of the cellular signalling network and that is amenable to quantitative modelling.

*[*]Department of Bioengineering, 9500 Gilman Drive, Mail Code 0412, University of California, San Diego, La Jolla, California 92093-0412, USA.*
*[‡]Department of Biomedical Engineering, University of Virginia, Box 800759, Health System, Charlottesville, Virginia 22908, USA.*
*[§]The Salk Institute for Biological Studies, Molecular and Cellular Biology Laboratory, 10010 North Torrey Pines Road, La Jolla, California 92037-1099, USA.*
*[‖]San Diego Supercomputing Center, Department of Chemistry & Biochemistry, University of California, 9500 Gilman Drive, San Diego, California 92093-0505, USA.*
*Correspondence to B.P. and S.S.*
*e-mails: palsson@ucsd.edu; shankar@ucsd.edu*
doi:10.1038/nrm1570

Recent genomic technologies have resulted in increasingly more detailed descriptions of signalling mechanisms, which have generated reconstructions of ever-larger signalling networks. Such reconstructions will enable a systemic understanding of signalling network function, which is crucial for studying diseases as diverse as asthma and cancer[1]. Although an increasing number of databases provide invaluable information for biological research[2], the characterization of properties that arise from whole-cell function requires integrated, mathematical descriptions of the relationships between different cellular components[3,4]. A NETWORK RECONSTRUCTION includes a chemically accurate representation of all of the biochemical events that are occurring within a defined signalling network, and incorporates the interconnectivity and functional relationships that are inferred from experimental data. Network reconstructions provide the framework for the application of mathematical methods that can quantitatively describe the properties of signalling networks.

Cellular signalling networks operate over several ORDERS OF MAGNITUDE in spatio-temporal scales. The differences in scale are readily evident in three popular examples of AUTOCRINE, PARACRINE and ENDOCRINE signalling events. In the first example, HELPER T CELLS secrete interleukin-2, which, in turn, binds to receptors on the plasma membrane of these secreting cells to stimulate cell proliferation. In the second example, helper T cells in a lymph node secrete interleukin-2, which binds to receptors on the surface of neighbouring B cells and induces cell differentiation. Third, differentiated B cells secrete antibodies, which bind to the antigens on the surface of specific bacteria throughout the body; this eventually results in PHAGOCYTOSIS of the antibody-bound bacteria.

These three examples involve both extracellular and intracellular signalling mechanisms. The endpoints of signalling events include quick responses ($<10^{-1}$ seconds), such as protein modifications and changes in $Ca^{2+}$ concentrations, as well as slow responses (from minutes to hours), such as transcriptional regulation, cell migration, cell-cycle control, cell proliferation and apoptosis.

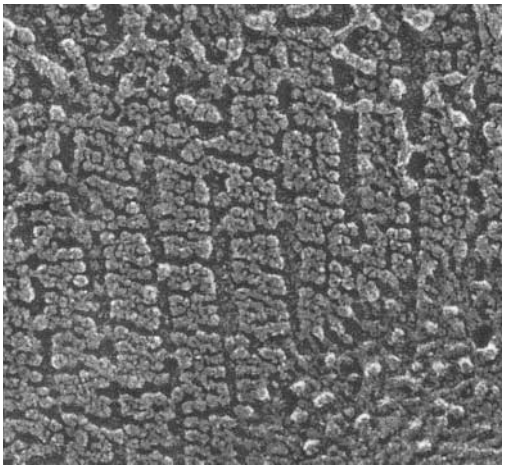
A systematic approach to the reconstruction and mathematical analysis of large-scale signalling networks requires the integration of several events that happen at diverse spatio-temporal scales[4–6]. The large number of components, the degree of interconnectivity, the differences in spatio-temporal scales, and the complex control of signalling networks are becoming evident in the integrated genomic and proteomic analyses that are emerging[7]. Whole-network analyses are necessary to elucidate the global properties amidst the complexity of signalling systems. In this review, estimates of orders of magnitude in human cellular signalling networks are discussed and then issues that are associated with the reconstruction of such networks are presented. A brief survey of existing and developing structural and dynamic analyses of signalling networks is also included. Examples are provided from immune-response cells and their associated

### Box 1 | Orders of magnitude — signalling networks

Simple calculations based on orders of magnitude can provide insights into biochemical systems and answer fundamental questions. For example: how many ligand molecules might a typical cell be exposed to at a given instant in time? What is the maximum number of receptors in a cell membrane? And how many transcription factors are needed to control 25,000 genes? Regarding the second question, three assumptions are made for estimates of the maximum number of signalling receptors on the surface of a typical mammalian cell. First, a maximum of 25% of the cell surface area comprises receptor proteins. Second, the average cell radius is 10 μm. This radius corresponds to a cell surface area of ~1,200 $\mu m^2$ (surface area = $4\pi r^2$) if the cell is spherical. Given the first assumption, 300 $\mu m^2$ of cell surface area can comprise receptor proteins. Third, the average receptor protein has a radius of 5 nm (REF. 106). This radius corresponds to a circle with an area of 75 $nm^2$ for a single receptor protein. (The figure shows acetylcholine receptors on the axon terminal surface of a neuron.) With these three assumptions, we can make an estimate of the maximum number of receptor proteins on a given cell (see equation 1).

$$\frac{300\ \mu m^2}{75\ nm^2} = 4 \times 10^6 \text{ maximum number of receptors} \quad (1)$$

This value seems a reasonable absolute maximum for the total number of receptor proteins considering published *in vivo* data. For example, $10^4$–$10^5$ receptor proteins of one type have been observed for some cell types[37]. This estimated maximum of $4 \times 10^6$ receptor proteins implies that there could be a maximum of 40 different receptor types of $10^5$ receptor proteins each. Certainly, these maximum constraints might be adjusted for a given cell that has a greater surface area, smaller- or larger-sized receptor proteins, or a greater percentage of its plasma membrane devoted to receptor protein. In addition, the plasma membrane might consist of channels, pores and other functional proteins. The figure is provided courtesy of J. Heuser, Washington University School of Medicine, and is reproduced with permission from REF. 130 © (1979) The Rockefeller University Press.



NETWORK RECONSTRUCTION
The process of integrating different data sources to create a representation of the chemical events that underlie a biochemical reaction network.

ORDER OF MAGNITUDE
A simple, quantitative estimate of a parameter.

AUTOCRINE
Describing, or relating to, a cell that produces the ligands by which it is activated.

PARACRINE
Describing, or relating to, a regulatory cell that secretes an agonist into intercellular spaces from which it diffuses to a target cell other than the one that produces it.

ENDOCRINE
Describing, or relating to, a gland or group of cells that makes hormones and secretes them into the blood, lymph or intercellular fluid.

HELPER T CELL
A T cell that functions as an inducer of the effector cells for humoral and cell-mediated immunity. These cells recognize and bind to antigen.

PHAGOCYTOSIS
An actin-dependent process by which cells engulf external particulate material by extension and fusion of pseudopods.

signalling reactions, where available. Although this review focuses on cellular signalling networks that affect transcriptional regulation, many other processes — including mechanotransduction, cytoskeletal organization, organelle assembly, vesicular trafficking, and metabolism — are also tightly coupled with cellular signalling networks, and their relationship to each other is an active area of research[8]. In particular, mathematical formulation of the coupling between mechanical processes and biochemistry poses significant challenges[9,10].

The complexity that is apparent from the examples outlined below argues for a conceptual departure from causal, small-scale descriptions to systemic reconstructions of signalling networks that can directly integrate with metabolic and regulatory network reconstructions. The reconstructed networks can then be modelled in varying degrees of detail to understand their complexity and make quantitative predictions. Although mathematical formulation of these networks is important, by no means can it be argued that a systemic network can be reduced to a set of differential equations. The argument is presented that the level of detail of a cellular signalling network can range from inclusion of every single moiety that is involved to coarse-grained descriptions of key processes, and that the biological context and expected quantitative outcome dictate the type of mathematical formulation of the biochemical network.

## Taking stock of components and interactions

A methodical, quantitative analysis of any complex system should be preceded by a survey of the orders of magnitude of key parameters and variables. First, estimates of the number of components in the human cellular signalling network are discussed (BOX 1). Second, orders of magnitude for the interactions of these components are discussed in the context of the degree of combinatorial control that these interactions afford. Third, extracellular interactions between networks are presented in the context of a cell that emits and receives a signal.

*Intracellular components.* The human genome contains ~25,000 genes[11]. A human being develops from a single cell into $10^{14}$ cells that are comprised of more than 200 different cell types[12,13] (TABLE 1). This developmental process and the subsequent homeostatic mechanisms of the organism require precise orchestration of cellular functions, which is achieved through signalling and other networks. The human cellular signalling network includes genes for 1,543 signalling receptors[14], 518 protein kinases[15] and ~150 protein phosphatases[16–18]. The activity of these components of the human signalling network can result in the activation (or inhibition) of transcription factors (of which there are estimated to be more than 1,850 in the human genome[19]), which then direct cellular regulatory processes.

A primary mechanism to control the number of intracellular components occurs at the level of post-transcriptional mRNA processing[20]. At present, estimates indicate that 40–60% of the genes in the human genome are subject to alternative-splicing events[21,22]. With estimates of an average of 8 exons per gene (although at the

Table 1 | The scope of the human cellular signalling network

| Network component | Number | References |
|---|---|---|
| Cells | $10^{14}$ | 12 |
| Cell types | 200 | 13 |
| Genes | 25,000 | 14,19 |
| Percentage of genes with splice variants | 40–60 | 21 |
| Average number of exons per alternatively spliced gene | 8 | 14 |
| Maximum number of exons per alternatively spliced gene (taken from *TTID* that encodes titin) | 234 | 14 |
| Average number of splice variants per gene across the genome | 2.5 | 19 |
| Percentage of alternatively spliced genes with signalling function | 75 | 24 |
| Average number of post-translational modifications per protein (current estimates) | 2.5 | (see below*) |
| Genes for transcription factors | 1,850 | 19 |
| Genes for protein kinases | 518 | 15 |
| Genes for protein phosphatases | 150 | 16–18 |
| Genes for receptors | 1,543 | 14 |
| Genes for GPCRs (for endogenous ligands) | 367 | 33 |

*See the Human Proteomics Initiative in the online links box.

upper extreme, the gene that encodes titin (*TTID*) might have up to 234 exons[14]), there are potentially a large number of splice variants for each gene. There is evidence that cellular signalling networks use alternative splicing to control the number and types of network component. For example, there are several splice variants of the gene that encodes the β isoform of inhibitor of nuclear factor (NF)-κB (IκBβ), which is a protein that inhibits the activity of the NF-κB transcription factor by preventing its translocation from the cytoplasm to the nucleus. The IκBβ1 and IκBβ2 proteins, which arise from differential splicing of the *IκBβ* gene (which is also known as *NFKBIB*), have different proteolytic-degradation properties and this confers differential control on the signalling network[23]. There are also at least 10 alternatively spliced exons in the gene that encodes CD44, the hyaluronan receptor, which is crucial for the immune response. CD44 is a cell-surface glycoprotein that is expressed in lymphocytes and other cell types. Splicing of alternative exons of *CD44* has the potential to generate 1,024 isoforms ($2^{10}$), of which ~30 have already been identified[20]. Initial estimates indicate that genes corresponding to proteins with signalling functions comprise more than 75% of all alternatively spliced genes[24]. This remarkable bias in alternative-splicing events implies that cells use this mechanism extensively to achieve the extraordinary specificity that is required in signalling systems. Assuming an average of 2.5 splice variants per gene across the entire genome (although this is probably a conservative estimate[19]), it is estimated that there are mRNA transcripts for 3,858, 1,295 and 375 distinct receptors, kinases and phosphatases, respectively, in the human cellular signalling network.

The next level of combinatorics in the number of signalling network components involves post-translational modifications (PTMs) — for example, phosphorylation, acetylation, methylation, and so on. At present, databases that catalogue PTMs indicate that there are, on average, at least 2.5 modifications per protein (see the Human Proteomics Initiative in the online links box, and REF. 25). This is likely to be a significant underestimate considering the difficulty in comprehensively identifying PTMs, especially when proteins are present at low levels. Proteins are also subject to proteolytic events that further regulate their activity. For example, as a result of splicing and cleavage events, many growth factor receptors can exist in soluble or membrane-bound forms, each of which might have a distinct function in cellular signalling[26]. The PTMs might not function independently of each other. However, three independent PTMs would correspond to eight distinct states of a given protein (each of the 3 PTMs could be present or absent, so $2^3 = 8$), and each state could have different binding and activating properties. Assuming that there are 3 PTMs per protein product of each mRNA transcript, and consequently 8 distinct protein states, it can be estimated that there would be 30,864, 10,360 and 3,000 distinct receptor, kinase and phosphatase states, respectively, each with potentially different properties in the network.

These numbers constitute nearly a 20-fold increase in the number of protein states over the respective number of genes, which approximates the estimated 10–100-fold increase in the number of components of the human proteome over the human genome[27]. There is also increasing evidence that the modification of proteins is an important mechanism for controlling the function of the cellular signalling network. For example, histone H3 (referred to as H3), which regulates gene expression by organizing chromatin structure, can be differentially methylated, phosphorylated and acetylated by various signalling proteins[28]. The methylated and acetylated forms of H3 show distinct properties that affect gene expression. Extensive and varied modifications of histone proteins might be a general method of regulating gene expression[29].

Alternative splicing and PTMs can therefore generate many more components of signalling networks than is indicated by the number of annotated genes on the human genome (TABLE 1). Tissue-specific gene expression and alternative splicing limit the number of different protein states that are potentially present in a single cell type, so the scope of this combinatorial control is speculative. However, there is emerging evidence that the many possible combinations of cellular components contribute to the specificity of responses of cellular signalling networks to various stimuli.

*Links and connectivity.* In addition to the increasing number of signalling network components that arises from alternative splicing and PTM, the interactions between the network components allow for an even greater degree of combinatorial control. A simple hypothetical example shows how the interactions between even small numbers of elements give rise to a broad spectrum of distinct functional states. The homo- and heterodimerization of only 224 proteins would provide sufficient specificity (for example, as activating protein complexes) to control the expression of all 25,000 of the genes in the human genome (in general, 'x' proteins can form [x + (x–1) + (x–2) + … +1] different homo- and heterodimers). If a given regulatory protein was associated with several genes, then the number of required homo- and heterodimers for such specificity would be even less than 224. There is emerging experimental evidence that the combinatorics of interactions among cellular signalling network components is indeed a primary mechanism for generating larger (and more specialized) signalling networks. A recent study showed that the interactions between proteins in yeast could be rearranged by the synthesis of chimeric protein constructs of components of the mitogen-activated protein kinase (MAPK) pathways so as to redirect the cellular responses between mating and high-osmolarity signalling stimuli[30]. Initial estimates of the number of interactions in the yeast proteome indicated that there are an average of five interacting partners per protein[31]. The role of combinatorics of protein complexes in cellular signalling networks has been discussed in a recent review[32].

Furthermore, a relatively small number of receptors can allow for the discrimination of a large number of environmental stimuli. A recent study analysed the repertoire of G-PROTEIN-COUPLED RECEPTORS (GPCRs) in the human genome and identified 367 GPCRs for endogenous ligands[33]. The expression profiles of 100 GPCRs in the mouse genome in 26 different tissues also indicated that most of the receptors were expressed in various tissues, but that each tissue had a unique profile of expressed receptors. Although there are families of receptors within which members bind to the same ligand (for example, GPCRs for adrenaline), assuming that a mere 1% of the estimated 1,543 receptors in the human genome[14] (15 receptor proteins) can be independently expressed in any given cell type, then the cell could potentially respond to 32,768 different ligand combinations. (2 independent states (bound or unbound ligand) of 15 receptor proteins corresponds to $2^{15}$ = 32,678 different bound-ligand states; in general, x objects each with y different states can generate $y^x$ unique combinations.) Although several receptors might trigger the same intracellular signalling responses, these 'back-of-the-envelope' calculations emphasize that a small number of intracellular signalling proteins and receptor proteins operating in a combinatorial manner can allow for a large diversity of function in signalling networks. This is a concept for which experimental evidence is surfacing.

Indeed, mathematical studies of metabolic networks have indicated that the number of functional states in a biochemical network grows much faster than the number of components. For example, in metabolic networks, the addition of a single reaction to a network could increase the number of functional pathways by several-fold[34]. This property is also likely to be found in signalling networks, particularly in consideration of the number of PTMs and splice variants, and the degree of interconnectivity of signalling network components. It is apparent that signalling networks will similarly require mathematical analyses to describe their network properties[34,35].

*Signal reception.* The maximum number of receptor proteins per typical somatic cell is estimated to be a few million (BOX 1). As the calculation of this maximum number depends on the percentage of plasma membrane proteins that is represented by other functional proteins (for example, channels and pumps) as well as the size of the receptor protein, variations in these properties could reduce or augment this maximum number of receptor proteins. Furthermore, the cell might regulate its size and the proportion of plasma membrane that consists of receptor proteins to increase or decrease the number of available receptor proteins. However, this maximum value is reasonable considering that some cells are estimated to contain $10^4$–$10^5$ of each type of receptor[36,37] and that A431 cells, which have a 10–30-fold greater density of epidermal growth factor (EGF) receptors than other cell types, contain up to ~2 x $10^6$ EGF receptors[38,39]. With 4 x $10^6$ receptor proteins per cell, for example (BOX 1), there could be 40 receptor types with $10^5$ of each type of receptor protein. A recent survey indicated that there are more than 30 different receptor types[1] present in the human MAST CELL. A cell of twice the radius (for example, a MEGAKARYOCYTE[40], which can have a radius of ~20 μm), which is subject to the other assumptions listed above, would have an estimated maximum of 2 x $10^7$ receptor proteins per cell (BOX 1). In addition to receptors on their surfaces, cells also have large pools of internal, membrane-enclosed endosomal vesicles that can cycle to and from the plasma membrane and supply receptor proteins to the plasma membrane, either constitutively or inducibly[41]. Endosomal receptor proteins can also signal[42], and this would further increase the effective number of receptor proteins.

The number of receptors of any one type on a cell can vary greatly. For example, in the 70Z/3 murine pre-B-lymphocyte cell line, there are as few as ~100 high-affinity

G-PROTEIN-COUPLED RECEPTOR
(GPCR). A seven-helix membrane-spanning cell-surface receptor that signals through heterotrimeric GTP-binding and -hydrolysing G proteins to stimulate or inhibit the activity of a downstream enzyme.

MAST CELL
A type of leukocyte with large secretory granules that contain histamine and various protein mediators.

MEGAKARYOCYTES
Bone-marrow precursor cells that give rise to blood platelets. During differentiation, megakaryocytes become polyploid by endoreplication.

and ~900 low-affinity interleukin-1 receptors per cell[43]. Assuming that there are ~30,000 distinct receptor-protein types (from estimates of the number of receptor genes, splice variants, and PTMs as described above), the maximum of $4 \times 10^6$ receptor proteins per cell corresponds to ~130 receptors of each receptor type ($4 \times 10^6$ receptors divided by $3 \times 10^4$ receptor types), the same order of magnitude that is seen for the smallest observed number of receptors as shown by the number of interleukin-1 receptors described above. As some splice variants of receptor proteins and their corresponding PTMs will not affect ligand specificity, the maximum of $4 \times 10^6$ receptor proteins per cell could effectively correspond to 2,600 receptors per receptor type ($4 \times 10^6$ receptors divided by 1,543 genes for receptor proteins).

So, cells can express a few receptor types (~10–40) in high numbers (~$10^5$ per cell), or many receptor types (~ 2,000–30,000) in small numbers (~$10^2$ per cell) on their surface. The former might be more common in highly differentiated and specialized cells, whereas the latter might be more typical of stem cells or undifferentiated cells.

*Signal transmission.* Constraints on the transmission of a signal from one cell to another are two-fold. First, the flux is constrained by the maximum production/secretion rate of a signalling protein. Immunoglobin genes have strong promoters, and transcription and translation rates that correspond to the production of 2,000–8,000 antibody molecules (each with a molecular mass of ~150,000) per effector B cell, per second have been measured[44,45]. This rate corresponds to the secretion of 1 picogram of protein per cell per hour, which is the maximum mass of a signalling molecule that can be secreted by the cell per hour (this maximum rate is estimated by multiplying the number of antibody molecules, the molecular weight, the inverse of Avogadro's constant, and a time-scaling factor). Second, owing to diffusion limitations that are imparted by the extracellular environment, there are also spatial constraints on the propagation of a signal. Effective paracrine signalling distances have been estimated to be within the range of 25 cell diameters, with time constants of 10–30 minutes for smaller proteins such as soluble cyto/chemokines[46]. Although lipid modifications or binding of the protein to the EXTRACELLULAR MATRIX might significantly limit the number of cell diameters that a protein can travel, the existence of cytonemes (long cellular processes or filopodia that can extend for several cell diameters) might even allow this limit to be exceeded[47].

The brief considerations outlined above help define the orders of magnitude of the content and connectivity of signalling networks. As kinetic values become available the dynamic properties of signalling networks can also begin to be estimated — for example, network reaction flux rates. An appreciation of the orders of magnitude in signalling networks is the first step in analysing their properties because it allows for reasonable assumptions regarding limitations on possible network function. The next step involves network reconstruction in a biochemically and genetically consistent fashion within a framework that is amenable to mathematical analysis.

## Reconstructing a signalling network

The most laborious step in the reconstruction of a cellular signalling network involves the delineation of all of the known individual biochemical processes that comprise the network. Once the proteins that are expressed in a given cell type are identified, there are two main considerations for the network reconstruction process: first, the scope of the reconstruction, or the number of reactions and components to include; and second, the level of detail that is accounted for in the underlying biochemical processes. Ultimately, there will be genome-scale networks, which include information on the biochemical reactions that comprise them — as there are now for metabolism[48].

*Scope of reconstructed signalling networks.* Signalling network reconstruction can be approached in three different ways (see BOX 2, panel a). The first approach consists of reconstructions of highly connected 'NODES' in networks. Such reconstructions involve comprehensively listing the compounds and reactions that are associated with a given protein, ion or metabolite. The second approach to network reconstruction involves forming linear 'pathways' that connect signalling inputs to signalling outputs. For example, such a pathway might be the delineation of all of the steps from the binding of a growth factor to its receptor through to the subsequent activation of a transcription factor that induces the expression of target genes. The reconstruction and analysis of the well-studied, pheromone-activated MAPK pathway in yeast has shown the utility of such an approach[49,50]; for example, the mechanism by which the MAPK Fus3 is dephosphorylated and localized at particular steps in the SIGNALLING PATHWAY has been hypothesized[49]. The third approach consists of identifying SIGNALLING 'MODULES'. Such modules historically consist of groups of compounds and proteins that function together under certain conditions on the basis of phenomenological reasoning — for example, the Jun N-terminal kinase (JNK), p38 and extracellular signal-regulated kinase (ERK)/MAPK modules. These modules have led to detailed kinetic analyses that traced the concentrations of various effector proteins and helped to understand processes such as feedback mechanisms. For example, the EGF-receptor system has been extensively analysed, and the effects of receptor internalization and autocrine signalling loops have been described in detail[51].

Significant interest in modular descriptions of signalling networks has developed recently[52–54]. There are three characteristics that can be considered for such modular descriptions. First, modular networks can be described with precisely defined inputs and outputs[55]. Second, the relative timescales within a modular network should be comparable[56]. Without comparable timescales, one network function could

EXTRACELLULAR MATRIX (ECM). The complex, multi-molecular material that surrounds cells. The ECM comprises a scaffold on which tissues are organized, it provides cellular microenvironments and it regulates various cellular functions.

SIGNALLING NODE
A highly connected compound in an intracellular signalling network.

SIGNALLING PATHWAY
A linear set of reactions that connects an input to an output in an intracellular signalling network.

SIGNALLING MODULE
An intuitive grouping of reactions from an intracellular signalling network that have a related function.

## Box 2 | Reconstructing a signalling network

Network reconstruction involves the integration of several sources of data to describe the biochemical transformations that occur in a given network. Contextual specificity is a crucial consideration in answering five questions for signalling network reconstruction:

- What proteins and other network components participate?
- What are the ligand–receptor interactions?
- What are the receptor–intracellular-component interactions?
- What are the intracellular-component–intracellular-component interactions?
- What are the intracellular-component–DNA interactions?

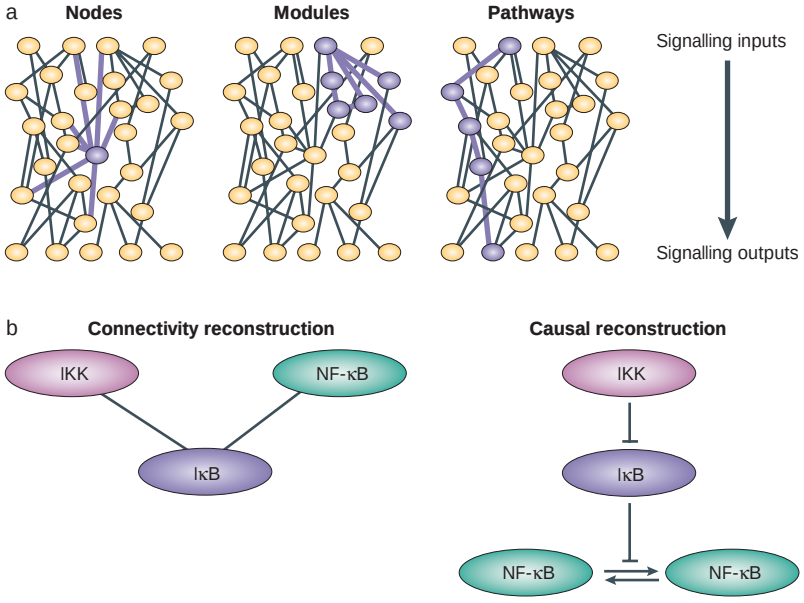
Genome annotation, biochemical experimentation, cell-physiology characterizations, expression arrays, and other such data sources each provide different types of datum that answer these questions and contribute to the reconstruction of a given cellular signalling network.

### Scope

Owing to a lack of comprehensive data regarding the interactions in a network, most signalling network reconstructions focus on particular nodes, modules or pathways in a given network (see figure, panel **a**). These reconstructions usually consist of a list of associations between network components. Network 'nodes' describe the many interactions that a given compound (for example, $Ca^{2+}$) participates in. Network 'modules' consist of a group of related reactions that often incorporate feedback mechanisms. Network 'pathways' connect a signalling input to a signalling output. Each of these types of reconstruction has distinct advantages for analytical purposes. However, some properties emerge from the interconnectivity of the nodes, modules and pathways with other network components.

### Resolution

Reactions amongst components in signalling networks are chemical transformations. There are three levels of resolution in reconstructions (see figure, panel **b**). A connectivity reconstruction lists the associations between network components (for example, nuclear factor (NF)-κB is functionally connected to IκB kinase (IKK) through the inhibitor of NF-κB (IκB)). A more detailed causal reconstruction describes cause-and-effect relationships and is often analysed with differential equations (for example, IKK interacts with the IκB–NF-κB complex such that NF-κB is activated). As signalling reactions are chemical transformations, they can also be represented by a more mechanistic description — for example, a stoichiometric matrix. This representation accounts for all chemical events that occur in a given network. For example, one IKK complex binds to and phosphorylates one IκB–NF-κB complex with two ATP molecules, which leads to the degradation of IκB and the nuclear localization of NF-κB ($R_1$–$R_4$). This relationship can be written out as a series of stoichiometric equations and its accompanying matrix, as indicated in panel **c** of the figure. pp, phosphate groups.




| Compounds | Reactions | | | |
|---|---|---|---|---|
| | $R_1$ | $R_2$ | $R_3$ | $R_4$ |
| IKK | –1 | +1 | 0 | 0 |
| IKK complex | +1 | –1 | 0 | 0 |
| IκBpp–NF-κB | 0 | +1 | –1 | 0 |
| IκBpp (degraded) | 0 | 0 | +1 | 0 |
| NF-κB (cytoplasm) | 0 | 0 | +1 | –1 |
| IκB–NF-κB | –1 | 0 | 0 | 0 |
| ATP | –2 | 0 | 0 | 0 |
| ADP | 0 | +2 | 0 | 0 |
| NF-κB (nucleus) | 0 | 0 | 0 | +1 |

Part **a** is reproduced from REF. 81 © (2003) Elsevier.

be sufficiently separated in time from another function such that there is effectively no interaction between the two. Third, the components of the modular network should be spatially colocalized[56]. Once these three criteria have been met, the CONTEXTUAL SPECIFICITY of the network can be maintained. However, even with such precise definitions, modular descriptions of signalling networks require intuitive, subjective, human-based decisions (for example, whether or not to include a scaffold protein or the activating receptor tyrosine kinase in the description of the module that comprises MEKK (MAPK and ERK kinase (MEK) kinase), MEK and MAPK). These considerations compel seeking a definition of modules from a mathematical, network-based perspective rather than subjective reasoning.

With increasing genomic and proteomic efforts, the scope of the reconstructed networks can be extended from these piecemeal reconstructions of nodes, modules and pathways to include a more comprehensive account of the signalling processes that occur in an entire cellular system. Recent studies have illustrated that actual properties of cellular signalling networks emerge from the interconnectivity of the entire system. For example, a

CONTEXTUAL SPECIFICITY
This takes into account the context in which a given signalling network property is observed — for example, splice variants of a particular protein might only exist in a cell when it is in a particular differentiated state.

recent model of β-adrenergic signalling in the rat ventricular myocyte accounted for several signalling interactions, and integrated the signalling network with excitation–contraction coupling that is fundamental to myocyte function[57]. This larger network model predicted that changes in myocyte contractility could primarily be accounted for by the activity of L-TYPE $Ca^{2+}$ CHANNELS and the phosphorylation of phospholamban — an example of a systems property that arises from the analysis of an integrated (and not piecemeal) network.

Importantly, signalling networks do not function in isolation from metabolic and regulatory processes. For example, recent experimental work has partly elucidated the signalling reactions that connect an isotype of protein kinase C (PKCθ) to the activation of NF-κB[58]. PKC is also intimately connected with lipid metabolism, as diacylglycerol (DAG) is one requirement for its activation. Indeed, lipids are frequent, fundamental components of signalling networks. Furthermore, ATP and GTP (which are principal cofactors in cellular signalling reactions) are primary products of metabolic networks. Transcriptional responses to stimuli can also lead to the synthesis of proteins that inhibit or activate corresponding signalling reactions. Indeed genes (or rather the corresponding sections of DNA) could be considered as components in a signalling network. The association of a transcription factor with the corresponding regulatory region of a gene might then be considered as another reaction in the network. Because signalling, metabolism and regulation are so tightly coupled with each other, the divisions between them are arbitrary. There is therefore a need to consider integrative approaches that can account for several 'types' of component and interaction. Recent work has involved the integration of a model of cardiac mitochondrial energy metabolism with what is known about $Ca^{2+}$ dynamics[59], which is a step towards this goal of integrating cellular signalling, metabolic and regulatory reconstructions.

*Level of detail in a reconstruction.* A concurrent consideration in the reconstruction of a signalling network is the desired level of detail, which is a primary function of the amount of data that are available. A reconstruction of associations involves a description of simple connectivity (for example, NF-κB is functionally connected to IκB kinase (IKK)) or a more involved set of relationships that shows more intermediates between a signalling input and a signalling output (for example, NF-κB is functionally connected to IKK through IκB; see BOX 2 and REF. 60). Network-connectivity reconstructions that consist of associations between components are amenable to several types of structural analysis (as discussed below). The next level of detail in a reconstruction enables the creation of cause-and-effect relationships (for example, IKK interacts with the IκB–NF-κB complex such that NF-κB is activated; see REF. 36 for an example of a causal reconstruction). Kinetic relationships build on these causal relationships, assigning scaling factors and time constants between different properties of interest. Stoichiometric reconstructions are based on chemically accurate representations of the biochemical transformations that occur between components in a signalling network — for example, one IKK molecule binds to and phosphorylates one IκB–NF-κB complex with two ATP molecules, which leads to the degradation of IκB and the nuclear localization of NF-κB (BOX 2). This level of detail requires an account of all components of the network (such as ATP) and contextual specificity (for example, NF-κB in the cytoplasm is distinct from NF-κB in the nucleus) that are necessary to drive a signal from stimulus to response. Much effort is needed to provide the data that are necessary for stoichiometric reconstructions of signalling networks. Stoichiometric models of metabolic networks are well developed and have been found to be quite informative[61,62].

*Data collection for the network reconstruction process.* Experimental techniques are continually being developed to identify the components and decipher the interactions in cellular signalling networks (FIG. 1). At the most crude level, the interactions between signalling networks (such as endocrine and paracrine signalling mechanisms) are readily deciphered through spatio-temporal measurements of the concentrations of extracellular molecules and the subsequent phenotypic changes that they induce. However, the connection between many extracellular signalling events and intracellular responses has remained a black box. High-throughput genome sequencing has provided the extensive 'parts list' of these intracellular signalling networks. The control of the expression of these parts can be deciphered, to an extent, by genome-wide location analysis — a recently developed technique that uses chromatin immunoprecipitation (ChIP)-chip assays to identify the binding sites of transcription factors throughout the genome[63] (see REF. 64 for a recent ChIP-chip analysis). However, work to characterize the mechanisms that connect extracellular signal inputs to the control of transcription factors was, until recently, restricted to painstaking experimental efforts with biochemical, genetic and pharmacological-intervention techniques.

High-throughput techniques to elucidate the mechanisms that connect an extracellular signalling stimulus to the control of transcription factors are still in their infancy (see REFS 65,66 for descriptions of various technologies that are important for reconstructing these intracellular signalling networks). Despite their shortcomings, such technologies are leading to the characterization of intracellular signalling mechanisms on a large scale. These developing technologies can be grouped into two categories: first, biochemical techniques and expression systems for characterizing protein–protein interactions; and second, assays for piecing together functional properties.

Perhaps the most widely used technique for deciphering protein–protein interactions involves YEAST TWO-HYBRID ASSAYS. However, so far there is only a small degree of congruence between different data sets that are generated using this approach[67]. In addition, yeast two-hybrid experiments lack contextual specificity — false-positive results occur partly because spatially or temporally segregated proteins that would

L-TYPE $Ca^{2+}$ CHANNELS
A form of voltage-operated $Ca^{2+}$ channel in cardiac muscle that has a high electrical threshold.

YEAST TWO-HYBRID ASSAY
A technique that is used to test whether two proteins physically interact with each other. One protein is fused to the GAL4 activation domain and the other to the GAL4 DNA-binding domain, and both fusion proteins are introduced into yeast. The expression of a GAL4-regulated reporter gene indicates that the two proteins physically interact.

GUANINE NUCLEOTIDE-EXCHANGE FACTOR
(GEF). A protein that facilitates the exchange of GDP (guanine diphosphate) for GTP (guanine triphosphate) in the nucleotide-binding pocket of a GTP-binding protein.

ICAT
(isotope coded affinity tag). ICAT probes have different masses, but are chemically identical. They incorporate a reactive cysteine, a biotin moiety, and eight deuteriums in place of eight hydrogens, and they are used to specifically label, by mass-difference, identical proteins in two separate samples for the identification and semiquantitative comparison of abundance.

SILAC
(stable isotope labelling by amino acids in culture). An experimental technique used to study hormone-activated protein complexes.

SH2 PROFILING
A technique based on the Far-Western assay that is used to identify SH2-binding domains in protein extracts.

TAIS
(target-assisted iterative screening). A method for screening protein products of a cDNA library that bind to a target protein.

RNAi
(RNA interference). A form of post-transcriptional gene silencing in which expression or transfection of dsRNA induces degradation — by nucleases — of the homologous endogenous transcripts. This mimics the effect of the reduction, or loss, of gene activity.

GREEN FLUORESCENT PROTEIN
(GFP). An autofluorescent protein that was originally identified in the jellyfish *Aequorea victoria*.

FRET
(fluorescence resonance energy transfer ). The non-radiative transfer of energy from a donor fluorophore to an acceptor fluorophore that is typically <80 Å away. FRET will only occur between fluorophores in which the emission spectrum of the donor has a significant overlap with the excitation of the acceptor.

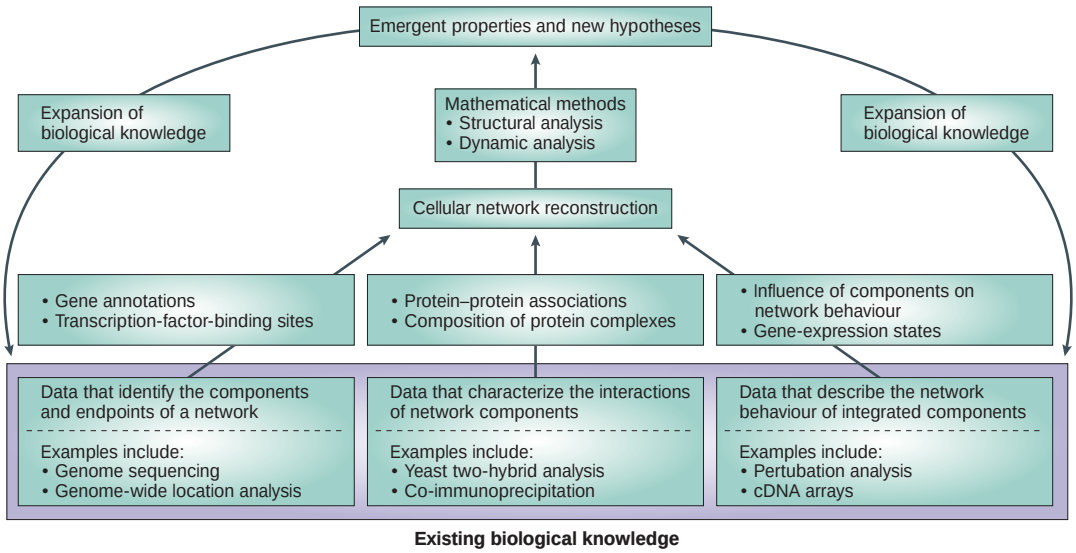



Figure 1 | **Integrative and iterative process of cellular signalling network reconstruction.** The existing biological knowledge for a given system is composed of several types of datum. Each datum type provides unique information that can be incorporated into a chemically accurate reconstruction (for example, a stoichiometric matrix). The first type of datum includes the identification of components and endpoints of a network (for example, genome sequencing or genome-wide location analysis). The second type of datum characterizes the interactions between network components (for example, yeast two-hybrid and immunoprecipitation data identify protein–protein interactions and protein complexes). The third datum type describes the network behaviour of the integrated components (for example, perturbation analysis and cDNA arrays delineate how entire networks function under various conditions). Each of these results provides unique types of datum that can be used to generate a cellular network reconstruction (for example, genome sequencing enables the annotation of the genes that are present in a given organism). With a network reconstruction, dynamic and structural analysis techniques can be used to describe emergent properties of the network, and generate new hypotheses. These characterizations then expand and revise the foundation of biological knowledge for the given system. This process can be iterated to offer increasingly more accurate descriptions of a given biochemical network.

not interact *in vivo* are 'thrown together' in this assay. A modified yeast two-hybrid approach that is known as the 'split-ubiquitin membrane yeast two-hybrid system' identifies interactions between membrane-bound proteins — such interactions cannot be studied with earlier yeast two-hybrid systems[68]. The son-of-sevenless (SOS; a GUANINE NUCLEOTIDE-EXCHANGE FACTOR for RAS) recruitment system (SRS) involves the interaction of proteins that are fused to SOS with membrane-bound or transmembrane[69] proteins that activate RAS. Interacting proteins induce cell growth. It has also been possible to characterize multiprotein complexes in *Saccharomyces cerevisiae* using mass-spectrometry approaches[70,71]. Additional biochemical techniques for investigating intracellular signalling networks are also being developed. These include isotope-coded affinity tags (ICATS), stable isotope labelling by amino acids in cell culture (SILAC), Src-homology-2 (SH2) PROFILING, and target-assisted iterative screening (TAIS; see REF. 72 for a recent review). Although these approaches are only beginning to be systematically applied on a large scale, the initial results are promising[73].

Recent advances in experimental techniques for investigating functional relationships in signalling networks show a strong potential for future analyses. Four approaches are highlighted here. First, perturbation analysis monitors genome-wide changes in expression (which are representative of transcriptional control) after disrupting specific components of a network — it has been used to refine models of galactose use in yeast[74]. Second, knockdown strategies (RNA interference (RNAi)) were recently used to elucidate components of the Hedgehog signalling pathway in *Drosophila melanogaster*[75], and the interactions between de-ubiquitylating enzymes and the IKK complex that is involved in NF-κB signalling[76]. Such knockdown strategies will certainly develop into a powerful tool for deciphering larger cellular signalling networks on a genome-wide scale[77,78]. Third, protein arrays are being developed to generate high-throughput data on the activity of proteins[79]. Fourth, fluorescence imaging technologies are generating data regarding protein localization and the dynamics of signalling processes[80,81]. For example, transformation of *Schizosaccharomyces pombe* cells with a fusion genomic library that contained fragments of the *S. pombe* genome fused to the gene that encodes GREEN FLUORESCENT PROTEIN (GFP) enabled the evaluation and localization of proteins to 11 distinct cellular compartments[82]. Recently, imaging technology also facilitated a global analysis of protein localization in *S. cerevisiae* and discriminated between 22 different cellular locales. It was used to assign a cellular compartment location for 70% of proteins for which the location was not previously known[83]. Furthermore, fluorescence resonance energy transfer (FRET) is a technique that can be used to decipher specific signalling mechanisms because it can indicate molecular proximity. For example, FRET has been used to study signalling events at the membrane,

and showed that 14% of EGF receptors in A431 cells were oligomerized before growth factor binding[84]. It has also been used to study membrane-associated signalling mechanisms (for example, activation of heterotrimeric G-protein complexes might involve rearrangement of the subunits rather than their dissociation[85]) and intracellular signalling events (such as the phosphorylation states of insulin-receptor substrates[86]).

As each of these techniques has distinct advantages and disadvantages, there is a growing need to integrate various data sources to represent most accurately the biochemical processes in a network.

*Integrative and iterative approaches.* Initial studies that have integrated disparate data sources have recently been published. For example, a recent study integrated three distinct data sets (from RNAi phenotyping, expression arrays and yeast two-hybrid analysis) to generate observations at a systems level for the *Caenorhabditis elegans* germline. Such observations included the tendency of essential proteins to interact with each other and similar expression profiles of pairs of genes that interact at the protein level[87]. Gene-expression array data for *Escherichia coli* and *S. cerevisiae* have also been integrated with reconstructed regulatory networks, which has yielded interesting results[88]. These include the result that the regulatory patterns observed from expression arrays and calculated from literature-based reconstructions are generally less consistent with each other for genes controlled by repressor proteins than for genes controlled by activator proteins.

Integrative approaches have also been used recently to characterize several signalling reactions in the RAS–ERK/MAPK[89], transforming growth factor β (TGFβ)[90], and tumour-necrosis factor α (TNFα)–NF-κB[91] signalling pathways. These approaches require the integration of several types of datum, as no single experimental protocol can accurately characterize all of the parameters that are necessary for a systems-level biological description. Large-scale-reconstruction efforts have already begun to integrate various experimental technologies and resources to reconstruct large-scale signalling networks. The Alliance for Cellular Signaling (see the online links box)[6] has focused resources on elucidating signalling mechanisms in the murine B cell[92] and cardiac myocyte[93] and, more recently, in a mouse macrophage cell line RAW 264.7. The Cell Migration Consortium and LIPID MAPS consortium (see the online links box) have also begun work on elucidating components of signalling networks in RAW 264.7 cells.

Underlying such large-scale efforts is the iterative process of reconstructing cellular signalling networks (FIG. 1). Reconstruction incorporates diverse data sets to represent biochemical processes in a given system. Each source of data contributes unique information, which needs to be integrated into a network reconstruction that is based on a mathematical framework (for example, a stoichiometric matrix). Mathematical methods can then be used to generate new hypotheses and describe the properties that emerge from network analyses. These computed data can be compared with the existing literature to revise and expand existing data and reconstructions for the given biological system. Some mathematical methods that are being developed for the analyses of biochemical networks will now be discussed.

CLUSTERING ANALYSIS
An approach for identifying and grouping similar data points.

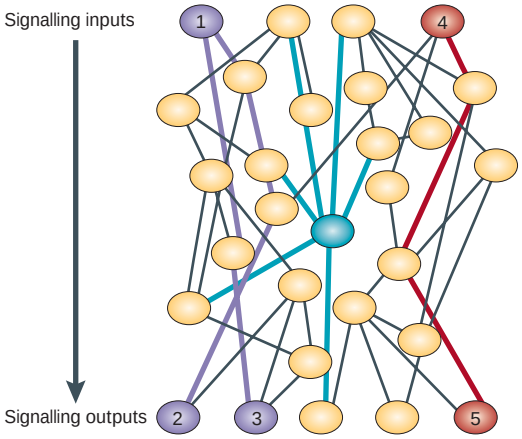



Figure 2 | **Structural analyses of signalling networks.** Structural analyses can identify components that are well or poorly connected (and therefore of potential interest for drug targeting). For example, the green component is the most highly connected in this schematic of a signalling network, and so drugs that inhibit the activity of this hypothetical component could have the broadest effect on the functions of the network. Structural analyses can also characterize which signalling inputs generate which signalling outputs. For example, in this schematic of a signalling network, the signalling inputs 1 and 4 can generate signalling outputs 2, 3 and 5.

## Mathematical analysis of network properties

Large-scale signalling networks are complex. Their complexity necessitates the use of methods from systems sciences, which are quite mathematical. Structural and dynamic analyses — that is, mathematical analyses that measure the time-invariant/topological and the time-variant properties of a network, respectively — can provide different results that can be integrated to characterize the properties of reconstructed signalling networks.

*Structural network analysis.* Large-scale networks can undergo structural analysis in their entirety, as this does not require an extensive knowledge of the parameters that have been determined from detailed experimentation. Structural analyses of connectivity reconstructions (which detail the existence of functional relationships between network components, see BOX 2) can generate hypotheses regarding the structure of the global network as well as the function of individual proteins (for example, see FIG. 2). Recently published examples illustrate the analyses that have led to hypotheses concerning global, modular and individual protein function. At a global level, a scale-free nature was observed in the yeast protein–protein interaction network (that is, there are a small number of network components with a high degree of connectivity)[94]. At a modular level, CLUSTERING ANALYSIS of the yeast signalling protein

Table 2 | Orders of magnitude for timescales in signalling networks

| Cellular signalling process | Time (in seconds) | References |
|---|---|---|
| ***Activities*** | | |
| Kinase/phosphatase reactions | $10^{-3}$ | 98,99 |
| Protein conformational changes | $10^{-3}$ | 106,127 |
| Cell-scale protein diffusion (passive) | $10^{0}$–$10^{1}$ | 101–103 |
| Cell-scale protein diffusion (active) | $<10^{0}$ | 104,105,128 |
| ***Responses*** | | |
| Cell migration | $10^{0}$–$10^{2}$ | 44,129 |
| Receptor internalization | $10^{2}$ | 43,111–113 |
| Transcriptional control | $10^{2}$ | 108–110 |
| Cellular growth | $10^{4}$ | 44 |

interaction network successfully partitioned groups of proteins that are associated with known signalling families such as the RAS–ERK/MAPK pathway[53]. At the level of individual protein function, SPECTRAL ANALYSIS defined groups of protein–protein interactions that had functional significance in the yeast proteome[95]. This spectral analysis provided hypotheses regarding the potential functions of proteins on the basis of their position in the network topology.

Initial structural analyses of causal reconstructions of signalling networks (BOX 2) have also highlighted the value of these analyses in describing network properties (for example, see REF. 96). So far, stoichiometric analyses of signalling networks are limited, owing to a lack of corresponding reconstructions. Recent work by our group involved the analysis of network features of the JAK (Janus-activated kinase)–STAT (signal transducer and activator of transcription) signalling network[97]. This analysis of a stoichiometric reconstruction has led to descriptions of protein synthesis requirements and energy demands of signalling networks, as well as mathematical definitions of network properties such as crosstalk and pathway redundancy.

*Dynamic network analysis.* A dynamic analysis of a reconstructed signalling network can be carried out once the associated kinetic parameters are known. The timescales that are associated with signalling processes can be estimated (TABLE 2), and can be crudely divided into two groups: signalling activities and signalling responses. Signalling activities typically occur rapidly. For example, most protein conformational changes, kinase/phosphatase reactions[98,99], and the physical movement of signalling compounds by diffusion or cytoskeleton-dependent mechanisms[100–105] occur over a timeframe that ranges from fractions of a second to seconds. However, signalling responses can occur over a wider range of timescales. Signalling responses that are coupled with metabolic processes or 'intermediate phenotypes' (for example, the conversion of glycogen to glucose, or increases in cyclic-AMP concentrations in response to stimuli) can occur over a timeframe of fractions of a second[106,107], as can elements of chemotactic and mechanotransduction behaviour[106]. However, other signalling responses occur over a timescale that is an order of magnitude slower. For example, transcriptional events[108–110], cellular growth[44] and receptor internalization[43,111–113] require several minutes, or longer, in response to a signal. This timescale separation is a crucial consideration for dynamic network analyses and can lead to simplifications that enable more thorough analyses which would otherwise be difficult. Further work is needed to account for the many timescales that occur in signalling systems. One promising approach involves MONTE CARLO SAMPLING for efficiently simulating a system of kinetic equations over several timescales[114].

As numerical values for kinetic parameters are typically difficult to obtain[115], dynamic analyses are usually only carried out for causal and stoichiometric reconstructions of smaller cellular signalling network reconstructions (FIG. 3). These studies have analysed complex network properties. For example, kinetic descriptions of complex reaction networks with feedback loops have been analysed to describe BISTABLE BEHAVIOUR[116].

The coupling of experimental data with these mathematical analyses can enable the identification of previously unknown signalling mechanisms. An elegant study that shows the benefit of integrative experimental and mathematical analyses deciphered the importance of particular IκB isoforms in feedback loops that involved the NF-κB SIGNALLING MODULE. Predictions were experimentally verified in knockout mouse models[117]. The WNT signalling module, which is important for development as well as oncogenesis, was recently represented using an extensive set of kinetic reactions[118]. Predictions were made for dynamic profiles of concentrations of β-catenin and other signalling mediators, and these matched experimental results. These dynamic analyses of signalling modules show the complex properties that can be studied once the reconstruction of only a limited number of reactions has been completed and experimental data are integrated with the model predictions.

## Conclusions

The main objective of this review is to emphasize the role and importance of the careful reconstruction of signalling networks and the advent of mathematical methods for their analysis. A brief inventory of the experimental approaches that are being developed to help unravel signalling mechanisms and an appreciation of the orders of magnitude in the human cellular signalling network have been provided. Furthermore, a survey of the existing structural and dynamic analysis methods was presented, with a sampling of the results and hypotheses that they have provided. Although lipids, proteins and metabolites are the principal components of signalling networks as they are known at present, further research in molecular biology will uncover additional signalling components and their concomitant mechanisms — as exemplified by the recent discovery of the regulatory functions of microRNAs[119]. As the different biochemical reactions that occur in entire cellular signalling networks are delineated, reconstructions will contain sufficient detail to simultaneously account for all of the necessary

SPECTRAL ANALYSIS
A method derived from graph theory that describes high-level structures in complicated networks of relationships.

MONTE CARLO SAMPLING
An approach for choosing pseudo-random data points that represent the characteristics of a larger population or function.

BISTABLE BEHAVIOUR
A property in which there are two stable points of a dynamic system, which provides a sense of 'memory'.

WNT PROTEINS
A family of highly conserved secreted signalling molecules that regulate cell–cell interactions during embryogenesis.

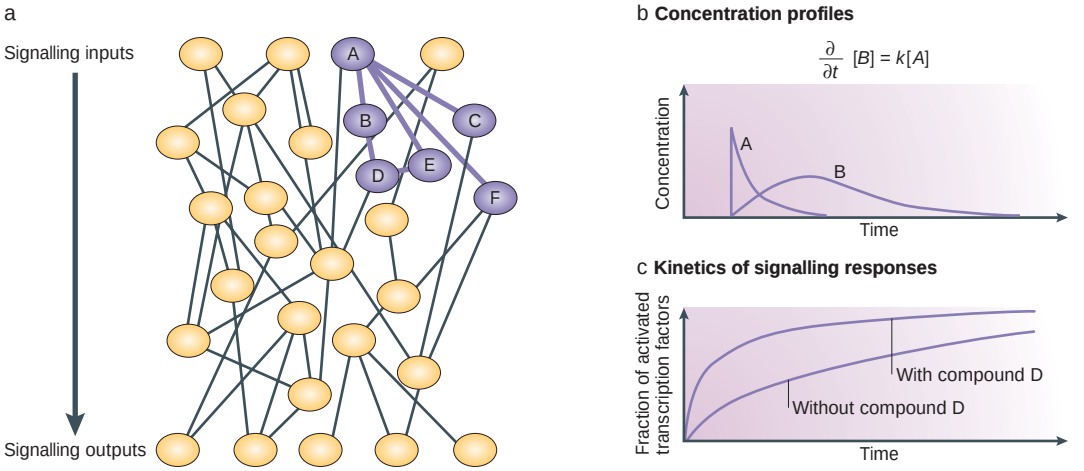



Figure 3 | **Dynamic analyses of cellular signalling networks.** When there is sufficient knowledge of parameters for a subset of a cellular signalling network (**a**; subset components A–F), complex dynamic profiles of various concentrations can be modelled and studied with dynamic analyses of signalling networks — as illustrated in the representative plot (**b**). For example, in this hypothetical system, an increase in the concentration of A results in an increase in the concentration of B. After A is depleted, B is slowly degraded. This relationship is indicated with the differential equation that relates the time derivative of B with the concentration of A and the constant *k*. Complex kinetics of network responses to signalling inputs can also be predicted from models that consist of kinetic descriptions of individual reactions (**c**). For example, in this hypothetical network, the fraction of activated transcription factors might be significantly reduced when compound D is absent from the network.

cellular network components (such as the demands for ATP and protein synthesis), and provide the level of detail that is necessary to integrate signalling networks with reconstructions of regulatory and metabolic networks[120,121]. These steps will lead to the reconstruction of whole-cell signalling networks. Recent efforts to automate network reconstruction processes[122] and extract molecular-interaction data from the literature[123] are important computational developments that will allow these more extensive network reconstructions.

Improvements in analysis methods will enable the characterization of previously incomplete sections of network reconstructions, which will further direct experimental programmes. Furthermore, network analysis might reveal strategies that are used in signalling cascades, and thereby generate theoretical arguments for network function. Stoichiometric signalling network reconstructions are likely to drive structural and dynamic network analyses and therefore lead to predictions of systems-level behaviour that will be important for future biochemical and medical research[124,125]. Network analysis will therefore meet the growing need to move beyond causal descriptions to mechanistic, bioengineering analysis of cellular signalling networks at various levels of detail[126].


1. Finkel, T. & Gutkind, J. S. *Signal transduction and human disease* (Wiley–Liss, Hoboken, New Jersey, USA, 2003).
2. Li, J. *et al*. The Molecule Pages database. *Nature* **420**, 716–717 (2002).
3. Levchenko, A. Dynamical and integrative cell signaling: challenges for the new biology. *Biotechnol. Bioeng.* **84**, 773–782 (2003).
4. Weng, G., Bhalla, U. S. & Iyengar, R. Complexity in biological signaling systems. *Science* **284**, 92–96 (1999).
5. Sivakumaran, S., Hariharaputran, S., Mishra, J. & Bhalla, U. S. The database of quantitative cellular signaling: management and analysis of chemical kinetic models of signaling networks. *Bioinformatics* **19**, 408–415 (2003).
6. Gilman, A. G. *et al*. Overview of the Alliance for Cellular Signaling. *Nature* **420**, 703–706 (2002).
**Provides an overview of the first coordinated, multi-institutional effort to systematically unravel the signalling mechanisms of model systems.**
7. Ge, H., Walhout, A. J. & Vidal, M. Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet.* **19**, 551–560 (2003).
**Emphasizes the need for integrating and reconciling data from several experimental sources to create the most accurate representation of biochemical networks.**
8. Helmke, B. P. & Schwartz, M. A. Putting the squeeze on mechanotransduction. *Dev. Cell* **6**, 745–746 (2004).
9. Ingber, D. E. Tensegrity I. Cell structure and hierarchical systems biology. *J. Cell Sci.* **116**, 1157–1173 (2003).
10. Ingber, D. E. Tensegrity II. How structural networks influence cellular information processing networks. *J. Cell Sci.* **116**, 1397–1408 (2003).
11. Pennisi, E. Human genome. A low number wins the GeneSweep Pool. *Science* **300**, 1484 (2003).
12. Hood, L. & Galas, D. The digital code of DNA. *Nature* **421**, 444–448 (2003).
13. Vander, A. J., Sherman, J. H. & Luciano, D. S. *Human physiology: the mechanisms of body function* (WCB McGraw–Hill, Boston, Massachusetts, USA 1998).
14. Venter, J. C. *et al*. The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
15. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912–1934 (2002).
**The authors present the first systematic inventory of a complete set of signalling network components (protein kinases) in the human cell.**
16. Wang, W. Q., Sun, J. P. & Zhang, Z. Y. An overview of the protein tyrosine phosphatase superfamily. *Curr. Top. Med. Chem.* **3**, 739–748 (2003).
17. Forrest, A. R. *et al*. Phosphoregulators: protein kinases and protein phosphatases of mouse. *Genome Res.* **13**, 1443–1454 (2003).
18. Alonso, A. *et al*. Protein tyrosine phosphatases in the human genome. *Cell* **117**, 699–711 (2004).
19. Lander, E. S. *et al*. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
**This, and reference 14, reported the initial sequence of the human genome. They provide a glimpse into the 'parts list' of signalling networks from which further reconstructions and analyses will be developed.**
20. Roberts, G. C. & Smith, C. W. Alternative splicing: combinatorial output from the genome. *Curr. Opin. Chem. Biol.* **6**, 375–383 (2002).
21. Modrek, B. & Lee, C. A genomic view of alternative splicing. *Nature Genet.* **30**, 13–19 (2002).
22. Thanaraj, T. A. *et al*. ASD: the alternative splicing database. *Nucleic Acids Res.* **32**, D64–D69 (2004).
23. Hirano, F. *et al*. Alternative splicing variants of IκBβ establish differential NF-κB signal responsiveness in human cells. *Mol. Cell. Biol.* **18**, 2596–2607 (1998).
24. Modrek, B., Resch, A., Grasso, C. & Lee, C. Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucleic Acids Res.* **29**, 2850–2859 (2001).
25. O'Donovan, C., Apweiler, R. & Bairoch, A. The human proteomics initiative (HPI). *Trends Biotechnol.* **19**, 178–181 (2001).
26. Heaney, M. L. & Golde, D. W. Soluble receptors in human disease. *J. Leukoc. Biol.* **64**, 135–146 (1998).
27. Service, R. F. Proteomics. High-speed biologists search for gold in proteins. *Science* **294**, 2074–2077 (2001).

28. Kiekhaefer, C. M., Grass, J. A., Johnson, K. D., Boyer, M. E. & Bresnick, E. H. Hematopoietic-specific activators establish an overlapping pattern of histone acetylation and methylation within a mammalian chromatin domain. *Proc. Natl Acad. Sci. USA* **99**, 14309–14314 (2002).
29. Fischle, W., Wang, Y. & Allis, C. D. Binary switches and modification cassettes in histone biology and beyond. *Nature* **425**, 475–479 (2003).
30. Park, S. H., Zarrinpar, A. & Lim, W. A. Rewiring MAP kinase pathways using alternative scaffold assembly mechanisms. *Science* **299**, 1061–1064 (2003).
31. Grigoriev, A. On the number of protein–protein interactions in the yeast proteome. *Nucleic Acids Res.* **31**, 4157–4161 (2003).
32. Hlavacek, W. S., Faeder, J. R., Blinov, M. L., Perelson, A. S. & Goldstein, B. The complexity of complexes in signal transduction. *Biotechnol. Bioeng.* **84**, 783–794 (2003).
33. Vassilatis, D. K. *et al*. The G protein-coupled receptor repertoires of human and mouse. *Proc. Natl Acad. Sci. USA* **100**, 4903–4908 (2003).
34. Papin, J. A., Price, N. D., Wiback, S. J., Fell, D. & Palsson, B. O. Metabolic pathways in the post-genome era. *Trends Biochem. Sci.* **28**, 250–258 (2003).
35. Bornheimer, S. J., Maurya, M. R., Farquhar, M. G. & Subramaniam, S. Computational modeling reveals how interplay between components of the GTPase-cycle module regulates signal transduction. *Proc. Natl Acad. Sci. USA* **101**, 15899–15904 (2004).
36. Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D. & Muller, G. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotechnol.* **20**, 370–375 (2002).
37. Kholodenko, B. N., Demin, O. V., Moehren, G. & Hoek, J. B. Quantification of short term signaling by the epidermal growth factor receptor. *J. Biol. Chem.* **274**, 30169–30181 (1999).
38. Masui, H., Castro, L. & Mendelsohn, J. Consumption of EGF by A431 cells: evidence for receptor recycling. *J. Cell Biol.* **120**, 85–93 (1993).
39. Buss, J. E., Kudlow, J. E., Lazar, C. S. & Gill, G. N. Altered epidermal growth factor (EGF)-stimulated protein kinase activity in variant A431 cells with altered growth responses to EGF. *Proc. Natl Acad. Sci. USA* **79**, 2574–2578 (1982).
40. Handin, R. I., Lux, S. E. & Stossel, T. P. *Blood: principles and practice of hematology* (Lippincott Williams & Wilkins, Philadelphia, USA, 2003).
41. Maxfield, F. R. & McGraw, T. E. Endocytic recycling. *Nature Rev. Mol. Cell Biol.* **5**, 121–132 (2004).
42. Wang, Y., Pennock, S. D., Chen, X., Kazlauskas, A. & Wang, Z. Platelet-derived growth factor receptor-mediated signal transduction from endosomes. *J. Biol. Chem.* **279**, 8038–8046 (2004).
43. Bomsztyk, K., Stanton, T. H., Smith, L. L., Rachie, N. A. & Dower, S. K. Properties of interleukin-1 and interferon-γ receptors in B lymphoid cell line. *J. Biol. Chem.* **264**, 6052–6057 (1989).
44. Alberts, B. *Molecular biology of the cell* 4th edn (Garland Science, New York, 2002).
45. Savinell, J. M., Lee, G. M. & Palsson, B. O. On the orders of magnitude of epigenic dynamics and monoclonal-antibody production. *Bioproc. Eng.* **4**, 231–234 (1989).
46. Francis, K. & Palsson, B. O. Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion. *Proc. Natl Acad. Sci. USA* **94**, 12258–12262 (1997).
47. Ramirez-Weber, F. A. & Kornberg, T. B. Cytonemes: cellular processes that project to the principal signaling center in *Drosophila* imaginal discs. *Cell* **97**, 599–607 (1999).
48. Reed, J. L., Vo, T. D., Schilling, C. H. & Palsson, B. O. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* **4**, R54 (2003).
49. van Drogen, F., Stucke, V. M., Jorritsma, G. & Peter, M. MAP kinase dynamics in response to pheromones in budding yeast. *Nature Cell Biol.* **3**, 1051–1059 (2001).
50. Kusari, A. B., Molina, D. M., Sabbagh, W. Jr., Lau, C. S. & Bardwell, L. A conserved protein interaction network involving the yeast MAP kinases Fus3 and Kss1. *J. Cell Biol.* **164**, 267–277 (2004).
51. Wiley, H. S., Shvartsman, S. Y. & Lauffenburger, D. A. Computational modeling of the EGF-receptor system: a paradigm for systems biology. *Trends Cell Biol.* **13**, 43–50 (2003).
**This review recounts the success of models of the EGF-receptor signalling system and argues for integrative computational and experimental approaches for dissecting signalling mechanisms.**
52. Milo, R. *et al*. Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
53. Rives, A. W. & Galitski, T. Modular organization of cellular networks. *Proc. Natl Acad. Sci. USA* **100**, 1128–1133 (2003).
54. Hartwell, L. H., Hopfield, J. J., Leibler, S. & Murray, A. W. From molecular to modular cell biology. *Nature* **402**, C47–C52 (1999).
55. Yi, T. M., Huang, Y., Simon, M. I. & Doyle, J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc. Natl Acad. Sci. USA* **97**, 4649–4653 (2000).
56. McAdams, H. H. & Shapiro, L. A bacterial cell-cycle regulatory network operating in time and space. *Science* **301**, 1874–1877 (2003).
57. Saucerman, J. J., Brunton, L. L., Michailova, A. P. & McCulloch, A. D. Modeling β-adrenergic control of cardiac myocyte contractility *in silico*. *J. Biol. Chem.* **278**, 47997–48003 (2003).
58. Lucas, P. C., McAllister-Lucas, L. M. & Nunez, G. NF-κB signaling in lymphocytes: a new cast of characters. *J. Cell Sci.* **117**, 31–39 (2004).
59. Cortassa, S., Aon, M. A., Marban, E., Winslow, R. L. & O'Rourke, B. An integrated model of cardiac mitochondrial energy metabolism and calcium dynamics. *Biophys. J.* **84**, 2734–2755 (2003).
60. Spirin, V. & Mirny, L. A. Protein complexes and functional modules in molecular networks. *Proc. Natl Acad. Sci. USA* **100**, 12123–12128 (2003).
61. Kauffman, K. J., Prakash, P. & Edwards, J. S. Advances in flux balance analysis. *Curr. Opin. Biotechnol.* **14**, 491–496 (2003).
62. Price, N. D., Papin, J. A., Schilling, C. H. & Palsson, B. O. Genome-scale microbial *in silico* models: the constraints-based approach. *Trends Biotechnol.* **21**, 162–169 (2003).
63. Ren, B. *et al*. Genome-wide location and function of DNA binding proteins. *Science* **290**, 2306–2309 (2000).
64. Odom, D. T. *et al*. Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**, 1378–1381 (2004).
65. Zhu, H. & Snyder, M. 'Omic' approaches for unraveling signaling networks. *Curr. Opin. Cell Biol.* **14**, 173–179 (2002).
66. Graves, P. R. & Haystead, T. A. A functional proteomics approach to signal transduction. *Recent Prog. Horm. Res.* **58**, 1–24 (2003).
67. von Mering, C. *et al*. Comparative assessment of large-scale data sets of protein–protein interactions. *Nature* **417**, 399–403 (2002).
68. Stagljar, I., Korostensky, C., Johnsson, N. & te Heesen, S. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins *in vivo*. *Proc. Natl Acad. Sci. USA* **95**, 5187–5192 (1998).
69. Aronheim, A., Zandi, E., Hennemann, H., Elledge, S. J. & Karin, M. Isolation of an AP-1 repressor by a novel method for detecting protein–protein interactions. *Mol. Cell. Biol.* **17**, 3094–3102 (1997).
70. Gavin, A. C. *et al*. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
71. Ho, Y. *et al*. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**, 180–183 (2002).
72. Stagljar, I. Finding partners: emerging protein interaction technologies applied to signaling networks. *Sci. STKE* pe56 (2003).
73. Blagoev, B., Ong, S. E., Kratchmarova, I. & Mann, M. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nature Biotechnol.* **22**, 1139–1145 (2004).
74. Ideker, T. *et al*. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**, 929–934 (2001).
**The authors present an approach for systematically interrogating a biochemical network.**
75. Lum, L. *et al*. Identification of Hedgehog pathway components by RNAi in *Drosophila* cultured cells. *Science* **299**, 2039–2045 (2003).
76. Brummelkamp, T. R., Nijman, S. M., Dirac, A. M. & Bernards, R. Loss of the cylindromatosis tumour suppressor inhibits apoptosis by activating NF-κB. *Nature* **424**, 797–801 (2003).
77. Zheng, L. *et al*. An approach to genomewide screens of expressed small interfering RNAs in mammalian cells. *Proc. Natl Acad. Sci. USA* **101**, 135–140 (2004).
78. Boutros, M. *et al*. Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303**, 832–835 (2004).
79. Nielsen, U. B., Cardone, M. H., Sinskey, A. J., MacBeath, G. & Sorger, P. K. Profiling receptor tyrosine kinase activation by using Ab microarrays. *Proc. Natl Acad. Sci. USA* **100**, 9330–9335 (2003).
80. Phair, R. D. & Misteli, T. Kinetic modelling approaches to *in vivo* imaging. *Nature Rev. Mol. Cell Biol.* **2**, 898–907 (2001).
81. Meyer, T. & Teruel, M. N. Fluorescence imaging of signaling networks. *Trends Cell Biol.* **13**, 101–106 (2003).
82. Ding, D. Q. *et al*. Large-scale screening of intracellular protein localization in living fission yeast cells by the use of a GFP-fusion genomic DNA library. *Genes Cells* **5**, 169–190 (2000).
83. Huh, W. K. *et al*. Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
84. Martin-Fernandez, M., Clarke, D. T., Tobin, M. J., Jones, S. V. & Jones, G. R. Preformed oligomeric epidermal growth factor receptors undergo an ectodomain structure change during signaling. *Biophys. J.* **82**, 2415–2427 (2002).
85. Bunemann, M., Frank, M. & Lohse, M. J. Gi protein activation in intact cells involves subunit rearrangement rather than dissociation. *Proc. Natl Acad. Sci. USA* **100**, 16077–16082 (2003).
86. Sato, M., Ozawa, T., Inukai, K., Asano, T. & Umezawa, Y. Fluorescent indicators for imaging protein phosphorylation in single living cells. *Nature Biotechnol.* **20**, 287–294 (2002).
87. Walhout, A. J. *et al*. Integrating interactome, phenome, and transcriptome mapping data for the *C. elegans* germline. *Curr. Biol.* **12**, 1952–1958 (2002).
88. Herrgard, M. J., Covert, M. W. & Palsson, B. O. Reconciling gene expression data with known genome-scale regulatory network structures. *Genome Res.* **13**, 2423–2434 (2003).
89. Plavec, I. *et al*. Method for analyzing signaling networks in complex cellular systems. *Proc. Natl Acad. Sci. USA* **101**, 1223–1228 (2004).
90. Tewari, M. *et al*. Systematic interactome mapping and genetic perturbation analysis of a *C. elegans* TGF-β signaling network. *Mol. Cell* **13**, 469–482 (2004).
91. Bouwmeester, T. *et al*. A physical and functional map of the human TNF-α/NF-κB signal transduction pathway. *Nature Cell Biol.* **6**, 97–105 (2004).
**High-throughput experimental technologies are beginning to be used to interrogate signalling networks at a cellular level, as shown in this reference and in reference 87.**
92. Sambrano, G. R. *et al*. Unravelling the signal-transduction network in B lymphocytes. *Nature* **420**, 708–710 (2002).
93. Sambrano, G. R. *et al*. Navigating the signalling network in mouse cardiac myocytes. *Nature* **420**, 712–714 (2002).
94. Jeong, H., Mason, S. P., Barabasi, A. L. & Oltvai, Z. N. Lethality and centrality in protein networks. *Nature* **411**, 41–42 (2001).
95. Bu, D. *et al*. Topological structure analysis of the protein–protein interaction network in budding yeast. *Nucleic Acids Res.* **31**, 2443–2450 (2003).
96. Schuster, S., Kholodenko, B. N. & Westerhoff, H. V. Cellular information transfer regarded from a stoichiometry and control analysis perspective. *Biosystems* **55**, 73–81 (2000).
97. Papin, J. A. & Palsson, B. O. The JAK–STAT signaling network in the human B-cell: an extreme signaling pathway analysis. *Biophys. J.* **87**, 37–46 (2004).
98. Goodman, O. B. Jr. *et al*. Role of arrestins in G-protein-coupled receptor endocytosis. *Adv. Pharmacol.* **42**, 429–433 (1998).
99. Vuong, T. M. & Chabre, M. Deactivation kinetics of the transduction cascade of vision. *Proc. Natl Acad. Sci. USA* **88**, 9813–9817 (1991).
100. Teruel, M. N. & Meyer, T. Translocation and reversible localization of signaling proteins: a dynamic future for signal transduction. *Cell* **103**, 181–184 (2000).
101. Lillemeier, B. F., Koster, M. & Kerr, I. M. STAT1 from the cell membrane to the DNA. *EMBO J.* **20**, 2508–2517 (2001).
102. Elowitz, M. B., Surette, M. G., Wolf, P. E., Stock, J. B. & Leibler, S. Protein mobility in the cytoplasm of *Escherichia coli*. *J. Bacteriol.* **181**, 197–203 (1999).
103. Swaminathan, R., Hoang, C. P. & Verkman, A. S. Photobleaching recovery and anisotropy decay of green fluorescent protein GFP–S65T in solution and cells: cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. *Biophys. J.* **72**, 1900–1907 (1997).
104. Allan, V. Membrane traffic motors. *FEBS Lett.* **369**, 101–106 (1995).
105. Theurkauf, W. E. Premature microtubule-dependent cytoplasmic streaming in *cappuccino* and *spire* mutant oocytes. *Science* **265**, 2093–2096 (1994).
106. Stryer, L. *Biochemistry* (W. H. Freeman, New York, 1995).
107. Neves, S. R., Ram, P. T. & Iyengar, R. G protein pathways. *Science* **296**, 1636–1639 (2002).
108. Zubay, G. *In vitro* synthesis of protein in microbial systems. *Annu. Rev. Genet.* **7**, 267–287 (1973).
109. Rivett, A. J. Regulation of intracellular protein turnover: covalent modification as a mechanism of marking proteins for degradation. *Curr. Top. Cell Regul.* **28**, 291–337 (1986).

110. McAdams, H. H. & Arkin, A. Simulation of prokaryotic genetic circuits. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 199–224 (1998).
111. Chang, D. Z., Wu, Z. & Ciardelli, T. L. A point mutation in interleukin-2 that alters ligand internalization. *J. Biol. Chem.* **271**, 13349–13355 (1996).
112. Ferguson, S. S. Evolving concepts in G protein-coupled receptor endocytosis: the role in receptor desensitization and signaling. *Pharmacol. Rev.* **53**, 1–24 (2001).
113. Jullien, J., Guili, V., Reichardt, L. F. & Rudkin, B. B. Molecular kinetics of nerve growth factor receptor trafficking and activation. *J. Biol. Chem.* **277**, 38700–38708 (2002).
114. Resat, H., Wiley, H. S. & Dixon, D. A. Probability-weighted dynamic Monte Carlo method for reaction kinetics simulations. *J. Phys. Chem. B* **105**, 11026–11034 (2001).
115. Bailey, J. E. Complex biology with no parameters. *Nature Biotechnol.* **19**, 503–504 (2001).
116. Bhalla, U. S. & Iyengar, R. Emergent properties of networks of biological signaling pathways. *Science* **283**, 381–387 (1999).
**Describes some of the first large-scale analyses of signalling reactions.**
117. Hoffmann, A., Levchenko, A., Scott, M. L. & Baltimore, D. The IκB–NF-κB signaling module: temporal control and selective gene activation. *Science* **298**, 1241–1245 (2002).
**Shows the powerful integration of mathematical modelling with experimental investigation.**
118. Lee, E., Salic, A., Kruger, R., Heinrich, R. & Kirschner, M. W. The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol.* **1**, 116–132 (2003).
119. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
120. Forster, J., Famili, I., Fu, P., Palsson, B. B. & Nielsen, J. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* **13**, 244–253 (2003).
121. Davidson, E. H. *et al*. A genomic regulatory network for development. *Science* **295**, 1669–1678 (2002).
122. Steffen, M., Petti, A., Aach, J., D'Haeseleer, P. & Church, G. Automated modelling of signal transduction networks. *BMC Bioinformatics* **3**, 34 (2002).
123. Krauthammer, M. *et al*. Of truth and pathways: chasing bits of information through myriads of articles. *Bioinformatics* **18** (Suppl. 1), S249–S257 (2002).
124. Cohen, P. Protein kinases — the major drug targets of the twenty-first century? *Nature Rev. Drug Discov.* **1**, 309–315 (2002).
125. Dancey, J. & Sausville, E. A. Issues and progress with protein kinase inhibitors for cancer treatment. *Nature Rev. Drug Discov.* **2**, 296–313 (2003).
126. Lazebnik, Y. Can a biologist fix a radio? — Or, what I learned while studying apoptosis. *Cancer Cell* **2**, 179–182 (2002).
127. Branden, C. & Tooze, J. *Introduction to protein structure* (Garland Pub., New York, USA, 1999).
128. Giannakakou, P. *et al*. p53 is associated with cellular microtubules and is transported to the nucleus by dynein. *Nature Cell Biol.* **2**, 709–717 (2000).
129. Devreotes, P. & Janetopoulos, C. Eukaryotic chemotaxis: distinctions between directional sensing and polarization. *J. Biol. Chem.* **278**, 20445–20448 (2003).
130. Heuser, J. E. & Salpeter, S. R. Organization of acetylcholine receptors in quick-frozen, deep-etched, and rotary-replicated Torpedo postsynaptic membrane. *J. Cell Biol.* **82**, 150–173 (1979).



## Acknowledgments

We would like to thank the Whitaker Foundation for a Graduate Fellowship to J.P., and acknowledge grants from the National Institutes of Health to S.S. T.H. is a Frank and Else Schilling American Cancer Society Research Professor. B.O.P. is a member of the Scientific Advisory Board of Genomatica Inc. We would also like to acknowledge valuable input from A. Hoffmann, and thank T. Allen, S. Becker, N. Price and J. Reed for detailed feedback on the manuscript.


## Competing interests statement

The authors declare no competing financial interests.

## Online links

DATABASES

The following terms in this article are linked online to:

Entrez: http://www.ncbi.nlm.nih.gov/entrez
CD44 | *NFKBIB* | *TTID*
*Saccharomyces* genome database: http://www.yeastgenome.org/
Fus3
Swiss-Prot: http://us.expasy.org/sprot/
GFP | IκBβ | JNK | RAS | SOS | TNFα

FURTHER INFORMATION

Alliance for Cellular Signaling: http://www.signaling-gateway.org
Bioinformatics and Computational Biology: http://genome.ucsd.edu
Cell Migration Consortium: http://www.cellmigration.org
Genomatica: http://www.genomatica.com
Human Proteomics Initiative: http://us.expasy.org/sprot/hpi/
LIPID MAPS consortium: http://ww w.lipidmaps.org
Molecular and Cellular Biology laboratory: http://pingu.salk.edu/~hunter/
Systems Biology Research Group: http://systemsbiology.ucsd.edu

Access to this links box is available online.