

# The ins and outs of signalling

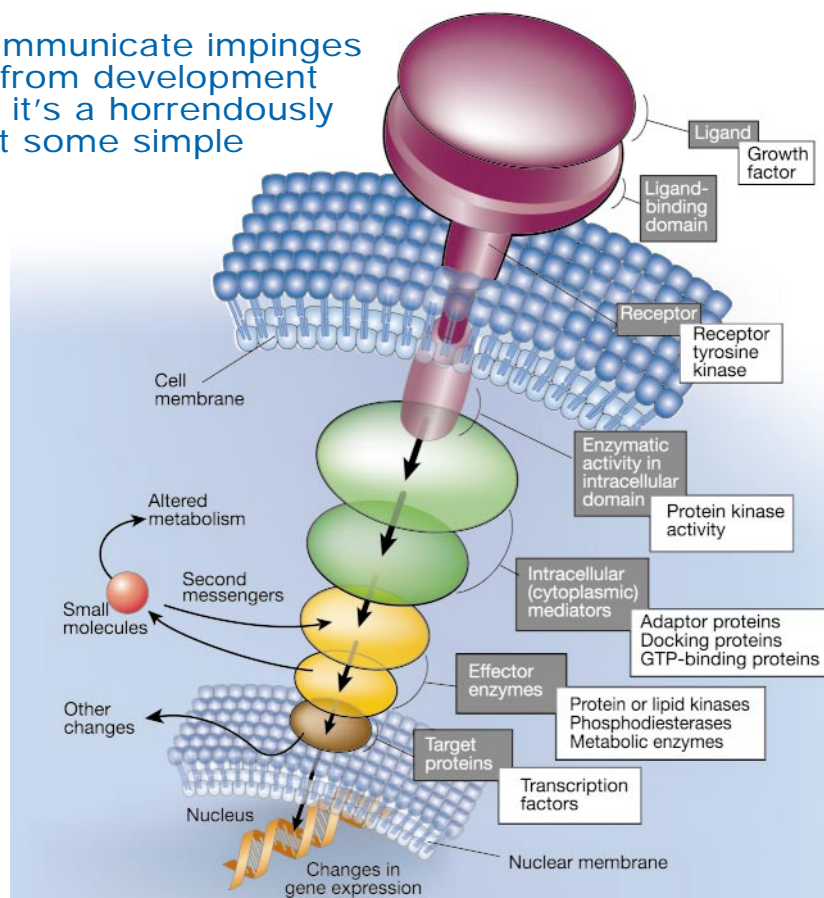
Julian Downward

The study of how cells communicate impinges on all aspects of biology, from development to disease. At first glance it's a horrendously complicated business, but some simple themes are emerging.

All cells must continually sense their surrounding environment and make decisions on the basis of that information. Single-celled organisms must be able to tell which nutrients are nearby and regulate their metabolic processes accordingly. Cells in multicellular organisms such as ourselves must sense the presence of neighbouring cells and hormones when making decisions such as to whether to proliferate, move or die. These processes all require the transfer of information (Fig. 1) from detection systems referred to as receptors through intermediate molecules within the cell, to cause changes in the expression of genes and the activity of enzymes — the specialized protein machinery that carries out all of a cell's functions. 'Signal transduction', 'cell signalling' or simply 'signalling' is the study of the mechanisms by which this transfer of biological information comes about.

Signalling can be studied at the level of the individual cell or the whole organism. For individual cells, signalling is crucial to decisions about division, specialization, death and metabolic control. In more specialized cells it is central to immunity and the transmission of nerve impulses, to name but two examples. At the level of whole multicellular organisms, signalling controls growth and development, as well as aspects of metabolism and behaviour. Not surprisingly, then, signalling malfunctions underlie many human diseases (Box 1, overleaf).

The history of the study of biological signalling can be read in the Nobel Prizes for Physiology or Medicine awarded over the past 30 years, starting with Sutherland in 1971 and culminating with the 2000 prize to Carlsson, Greengard and Kandel for their studies of signal transduction in the nervous system. The work of these and many other scientists has established the outlines of numerous signalling pathways. These vary enormously in their details. Yet despite this diversity, a much simpler underlying logic is beginning to emerge. Here I aim to explore key themes in biological signalling, and to illustrate them with some of the better-understood signalling systems. I hope that the reader will be convinced that signal transduction is not the impenetrable soup of acronyms that it might at first appear to be.



**Figure 1** A generic signalling pathway. The grey boxes indicate general components of signalling pathways; the white boxes show specific examples. On the outside of the cell membrane, receptors bind to biologically active ligands such as growth factors. As a result, enzymatic activity associated with the intracellular part of the receptor is altered. This can affect the association of the receptor with intracellular mediators, or the localization or function of those mediators. These in turn alter the activity of 'effector' enzymes. Some effectors can move to the nucleus and control gene expression, or they can induce other proteins to do so. Others target small molecules, either generating further signalling mediators (second messengers) or controlling the metabolic state of the cell. Real signalling pathways may bypass entire classes of these molecules, or may have several components in one or more class, working either in series or in parallel.

## Signalling nuts and bolts

**Receptors.** Signalling pathways start with receptor proteins that are able to sense a change in the environment outside the cell. Receptors are most commonly found at the cell surface, where they bind to extracellular molecules that cannot penetrate the plasma membrane — the lipid boundary between the cell and the outside world. Receptors on the cell surface can bind to water-soluble signalling proteins such as growth factors and peptide hormones, which may be produced at a distant site in the body (and delivered in the bloodstream) or by neighbouring cells. Receptors can also bind to small water-soluble molecules such as nutrients.

As a result of engagement of these receptor

proteins by their binding partners ('ligands'), a signal is transferred across the plasma membrane<sup>1</sup>. In most cases, the receptor itself spans the membrane. Ligand binding causes a change in the shape of the protein; this change is transmitted from the extracellular part of the receptor to the part inside the cell. Sometimes this involves the formation of dimers of receptor molecules — two receptors bonded together. This tends to occur for receptor proteins with amino-acid chains that cross the membrane only once. For receptors that fold up so that they span the membrane several times, ligand binding may cause different parts of the molecule to reorientate themselves with respect to each other.

Inside the cell, the change in the recep-

tor's conformation can result in the stimulation of an activity that is integral to the intracellular part of the receptor. One of the most common, and perhaps the best studied, activities is that of the kinases — enzymes that catalyse the transfer of a phosphate group from adenosine triphosphate (ATP, the cell's soluble energy-storing molecule) to protein substrates. Two principal classes of kinases exist: those, such as members of the activin/transforming growth factor (TGF)- $\beta$  receptor family, that transfer phosphate to serine and threonine amino-acid residues in target proteins; and those, including the growth-factor-receptor tyrosine kinase family, that target tyrosine residues<sup>2,3</sup>.

Other receptors have different activities. For example, some act as ion channels and pumps, allowing ions into or out of the cell to control cellular functions. Alternatively, the part of the receptor within the cell may interact with other signalling proteins and pass on the information to them by inducing changes in conformation and activity. A common example is provided by the abundant family of 'G-protein-coupled' receptors.

Cell-surface receptors can bind to insoluble structures as well as to soluble molecules. For example, integrins are receptors that bind to the extracellular matrix<sup>4</sup> — the glue that holds the body together. Cell-adhesion molecules are responsible for binding cells to their neighbours<sup>5</sup>. These types of receptor are important not only for structural reasons — in establishing tissue architecture — but also in informing cells of the presence of other cells and of the matrix, and enabling them to behave accordingly, for example by ceasing to grow when they are surrounded by other cells. Finally, receptors inside the cell bind to signalling molecules that can pene-

trate the plasma membrane, such as steroid hormones and nitric oxide.

**Intracellular mediators.** Inside the cell, receptors pass on the signal to other molecules. As mentioned above, receptor kinases catalyse the transfer of a phosphate group onto substrate proteins, which may be enzymes with activities that are directly affected by this phosphorylation. For example, receptors for TGF- $\beta$  phosphorylate proteins of the SMAD family (which are transcriptional regulators that work to control gene expression), resulting in their activation<sup>6</sup>. Receptors themselves are often modified by self-phosphorylation; this occurs for growth-factor receptors that stimulate phosphorylation on tyrosine residues within themselves.

The phosphorylation of substrate proteins can also affect their interactions with other molecules. For example, phosphorylated residues in a protein can act as binding sites for specific recognition domains in other proteins<sup>7</sup>. A domain in a protein is a self-folding unit with a particular sequence and conformation, and certain domains allow proteins to recognize each other. So, as a result of phosphorylation, for example, protein complexes can assemble, resulting in changes in the localization or activity of enzymes.

Some proteins in these complexes, referred to as 'adaptor' or 'docking' proteins, may work only to bring together other signalling molecules. One example is provided by the Ras pathway (Fig. 2). This pathway has been particularly well studied, not least because mutation of the *ras* gene occurs in many cancers. In this pathway, binding of a growth factor to its cell-surface receptor leads to the self-phosphorylation of the receptor. The SH2 domain of an adaptor protein, Grb2, grabs on to the phosphoryl-

ated tyrosine residue; Grb2 thereby links the activated receptor to an enzyme, Sos, that regulates Ras. As both the receptor and Ras are found at the plasma membrane, the outcome of the assembly of this complex is that Sos has increased access to Ras<sup>8</sup>. Ras then activates a cascade of intracellular protein kinases, culminating in activation of a kinase called ERK (a so-called MAP kinase), which directly influences gene expression.

**Nuclear events.** Like the Ras pathway, many signalling pathways end with a change in the gene-expression programme of the cell<sup>9</sup>. This usually requires the movement of a protein from the body of the cell to the nucleus in response to activation of the signalling pathway. Specific recognition systems control the import and export of proteins to and from the nucleus. These systems recognize sequence motifs in the proteins, and the accessibility of the motifs may be altered as a result of phosphorylation or complex formation<sup>10</sup>. In the growth-factor-activated Ras pathway, ERK moves to the nucleus as a result of its phosphorylation and activation by the previous kinase in the cascade. Once inside the nucleus, ERK phosphorylates and activates proteins that are involved in the transcription of genes into messenger RNAs. These mRNAs may then be translated into protein, so altering the protein composition of the cell and leading to changes in cell function.

### Signalling concepts

**Specificity.** The route that information takes from outside a cell to the endpoint of a signalling pathway is defined by the molecular interactions of the proteins in the pathway. These interactions are therefore critical. They are frequently induced only upon activation of the pathway, and they must be sufficiently

## Box 1 Signalling and disease

Many diseases involve malfunction of signalling pathways. In particular, much of the basic work on the regulation of cell proliferation has been carried out to obtain a better understanding of cancer. Cancer is a disease of signal malfunctioning that is driven by microevolutionary processes. Once a cell suffers an inactivating mutation in a growth-inhibitory pathway (that is, in a tumour-suppressor gene), or an activating mutation in a growth-promoting pathway (in other words, in an oncogene), it will have a competitive proliferative advantage over its neighbouring cells<sup>23</sup>. This can lead eventually to malignant growth of a mutant clone of cells at the expense of the whole organism. To avoid cancer, animals have

developed several signalling pathways that prevent uncontrolled cell proliferation, meaning that only after several mutations have occurred in the same cell does cancer result.

The study of oncogenes has led to the identification of many signalling proteins that are involved in cell growth, including growth factors, their receptors, intracellular mediators and transcription factors. Likewise, the study of tumour-suppressor genes has led to the identification of proteins that negatively regulate cell-cycle progression, cell survival and the ability to invade surrounding tissue.

Many other non-infectious diseases are caused by defects in signalling pathways. Diabetes, for

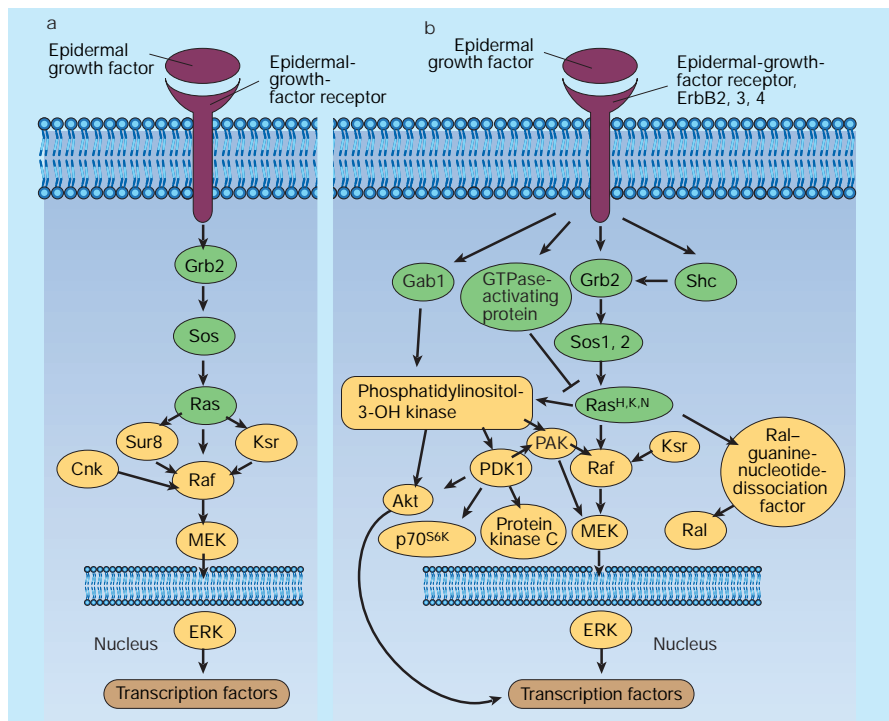
example, results from defects in the insulin-signalling pathway used to control blood glucose levels<sup>24</sup>. Insulin works through a receptor-tyrosine-kinase signalling pathway that results in activation of phosphatidylinositol-3-OH kinase, Akt and other intermediate enzymes, particularly in fat and muscle cells. Ultimately, these control the amounts of the glucose receptor at the cell surface, as well as the levels of various metabolic enzymes.

Moreover, many developmental diseases result from signalling defects, one example being forms of achondroplasia (dwarfism) in which the receptor tyrosine kinase for fibroblast growth factor is mutated. Immunological disorders can also have a basis in signalling

malfunction. For example, in agammaglobulinaemia (which is characterized by a lack of immunoglobulin antibodies in the blood), mutation in the B-cell tyrosine kinase Btk often occurs, resulting in a failure of this enzyme to respond to activation of phosphatidylinositol-3-OH kinase.

We can hope that in-depth characterization of signalling pathways will lead eventually to an ability to intervene in diseases in which those pathways are defective. For example, a drug called STI571, which inhibits the tyrosine kinase Abl, is now being used successfully to treat certain leukaemias in which this enzyme has been erroneously activated<sup>25</sup>.

J.D.



**Figure 2** Two ways of looking at one signalling pathway — the Ras pathway — from growth factor to changes in gene expression. **a.** The genetic view is based largely on studies of the development of flies, worms and, to some extent, mice. **b.** The molecular cell-biological view is derived mainly from work on mammalian cells in tissue culture. The genetic approach gives a simpler picture, but principally considers only components that are essential for normal development of the organism under laboratory conditions. Being based largely on simple model organisms, genetics has been less useful in studying the function of the several closely related members of gene families that are typically found in mammals (such as the epidermal-growth-factor receptor and Erb2, 3 and 4). The molecular cell-biological view is more complex: it is pieced together from experiments done under different conditions and in different cell types, often relying on overexpression of proteins in the pathway. It reveals many possible subtleties of the pathway, but these are unlikely to occur together to a significant extent at the same time in the same cell. The most effective modelling of signalling pathways draws on experimental evidence of both types to reach a mutually consistent picture.

specific to ensure that signals do not become misrouted. How is specificity achieved?

A strong theme that has emerged over the past decade is that proteins are made up of modular units, which belong to families of related structures<sup>11</sup>. Modules may be either domains or smaller motifs, such as some phosphorylation sites. Modules can direct protein–protein interactions through their ability to interact with other modules (Fig. 3, overleaf). They may also have enzymatic activity, which in signalling proteins is often used to regulate other molecules. A large protein may contain several different modules, each of which behaves similarly to related modules in other proteins. A protein is thus made up of generic functional building blocks, which have been shuffled around during evolution to yield different combinations of interactions and activities.

To ensure the accuracy of information flow, irrelevant interactions between proteins need to be kept at a manageable level. The use of interacting modules narrows down the number of connections that can occur. But there is only a limited number of module families, and many proteins will

contain the same modules, so how can they distinguish between related targets? Several of the design features of signalling pathways may aid specificity, including the use of combinations of modules to achieve higher selectivity for a given protein–protein interaction, and of scaffolds of adaptor proteins to help to bring partners together<sup>11,12</sup>. The localization of proteins to specific cellular compartments can also limit the range of potential interactions.

Signalling pathways presumably evolved to transfer information with acceptable selectivity under normal conditions. But a common experimental technique for analysing signalling pathways involves overexpressing either normal or mutant proteins to see the effects on the cell. Unfortunately, as many proteins contain the same modules, this technique may well result in perturbation of pathways other than the one that contains the protein of interest. It is not clear how big a problem this is. But there is a real danger that many data from overexpression experiments will have to be re-evaluated if we are to obtain a clear overall picture of cell signalling.

**Complexity.** The size of even the simplest

genome of a unicellular organism allows for enormous complexity of signalling. A sizeable proportion of the genome of a multicellular organism encodes signalling molecules; for instance, over 2% of the genome of the worm *Caenorhabditis elegans* encodes its 400 kinases<sup>13</sup>. Evolution drives the generation of robust, adaptable systems, and the modular nature of signalling proteins makes it easy to generate new variants to meet new environmental challenges. There is probably little benefit to be had from simplicity in itself, as the energy required to make so many genes and their products is amply repaid by the improved performance of the organism. Often, organisms have developed several ways to do the same thing, and this redundancy perhaps reflects the importance of some signalling pathways. Nonetheless, it is likely that the apparent complexity of many published signalling diagrams may be at least partly due to the experimental techniques used (Fig. 2).

**Signal integration.** Cells receive inputs from many signalling pathways at the same time and must interpret them together, in the context of each other, before making decisions. There are several known ways in which cells do this, although this is an area where much work remains to be done.

For example, individual enzymes can receive input from several pathways, which often feed into different modules of the same enzyme. This type of integration process is known as crosstalk, whereby an important signalling pathway is influenced by the activity of another, even if the second pathway cannot fully control the first. Examples of this include the ability of phosphatidylinositol-3-OH kinase and another kinase, PAK, to feed into the MAP kinase pathway that is triggered by activated Ras.

As well as crosstalk between pathways, feedback loops can occur, in which a component further on in a pathway either positively or negatively influences the activity of an earlier component. Negative-feedback systems are common in signalling. In contrast, positive-feedback systems are rarer, as they can be inherently unstable. However, it is common to find a degree of amplification built into signalling pathways, such that each step in the pathway results in an increase in signal strength, turning a weak signal at the cell surface into a strong one in the nucleus.

Signal integration also occurs at the level of gene expression. Different signals lead to activation of different sets of transcription factors, so affecting the expression of different genes. The cell's overall response will depend on how the induced proteins interact with and influence each other.

**Localization and translocation.** The location of a signalling protein within a cell affects the molecules with which that protein can interact. Movement of a signalling protein to a different cellular location — translocation — may result in a change in its interacting part-

ners and hence in its activity. Regulation of the localization of signalling proteins is a key aspect of many signal-transduction pathways. Thus the activation of phosphatidylinositol-3-OH kinase by receptor tyrosine kinases leads to the generation of a specific lipid, phosphatidylinositol-3,4,5-trisphosphate, at the plasma membrane. The lipid binds to so-called pleckstrin-homology domains in proteins such as Akt, a serine kinase<sup>14,15</sup>. This causes Akt to translocate to the plasma membrane, where it comes into contact with another protein kinase that phosphorylates and activates it. Phosphorylated Akt can then move to other sites in the cell, including the nucleus, while maintaining its activity.

Control of localization is also important in the control of many transcription factors, which can be kept out of the nucleus, and hence away from their target genes, by regulated interactions with other proteins. For example, the Forkhead transcription factors are inactivated when phosphorylated by Akt. Phosphorylation creates a binding site for 14-3-3 proteins, impairing the import of these transcription factors into the nucleus<sup>14</sup>.

**Signal amplitude and duration.** The response of cells to activation of a particular signalling pathway depends on the strength of that activation. Pathways can show graded responses, like a rheostat — the stronger the activation of the intermediate proteins in the pathway, the stronger the final activity. In these cases, because different cells may show different sensitivities to a signal, low signal strengths might activate a subset of the responses that are activated by high signal strengths. In addition, some pathways work as on/off switches — once signal strength rises above a certain level, positive feedback results in full activation of downstream targets.

The time course of a signalling pathway can also be critical. Transient activation of a pathway may have quite different effects to long-term activation. For example, in cultured PC12 cells, epidermal growth factor induces transient activation of the MAP kinase pathway, causing the cells to multiply slowly. Nerve growth factor, in contrast, induces sustained activation of the MAP kinase cascade, and the cells become specialized to resemble nerve cells<sup>16</sup>. During the stimulation of cell proliferation by growth factors, signals may be needed at specific periods of the cell-division cycle but not at others. What's more, signals may be interpreted differently at different points in the cell cycle, because of changes in other combinations of signals.

**The future**

Our ability to understand signalling pathways will be revolutionized by the availability of full genome sequences. Complete annotated genome sequences are available for a yeast, a nematode worm, a mustard plant and a fruit-fly, and the sequencing, assembly and annotation of the human genome are approaching

Partner 1	Partner 2	Interaction used by
Phosphorylated tyrosine residues	SH2 domain PTB domain	Growth-factor receptor tyrosine kinase
Phosphorylated serine/threonine residues	FHA domain 14-3-3 domain	Enzymes within signalling pathways, such as Raf Transcription factors
Motifs rich in proline residues	SH3 domain WW domain	Adaptor proteins regulated by growth factors Also used in regulating cytoskeletal, cytoplasmic and nuclear proteins
Phosphoinositide lipids	PH domains	Growth-factor-regulated enzymes, such as Akt
DD, DED, CARD domains	DD, DED, CARD domain	Used in regulation of programmed cell death
G(S/T)XVI sequence (at carboxy terminus)	PDZ domain	Ion channels Also used for control of cell polarity

**Figure 3 Some common protein-interaction modules used in signalling. Modules can direct protein-protein interactions through their ability to specifically recognize other modules. Abbreviations: SH2 and SH3, Src-homology domains 2 and 3; PTB, phosphotyrosine-binding domain; FHA, Forkhead-associated domain; WW, tryptophan-tryptophan domain; PH, pleckstrin-homology domain; DD, death domain; DED, death-effector domain; CARD, caspase-recruitment domain; G(S/T)XVI, amino-acid sequence glutamate, serine/threonine, any amino acid, valine, isoleucine; PDZ, a domain present in so-called PSD-95, Dlg and ZO1/2 proteins<sup>7</sup>.**

completion<sup>17,18</sup>. These innovations have fundamentally redefined the study of all aspects of biology, including signalling. In yeast, worms and flies, we can now find every member of every family of signalling proteins, and this is rapidly becoming the case for humans as well. As more is learnt, it will be possible to fill in more and more detail about the function of as-yet-uncharacterized proteins, simply from the sequence of their genes.

Of course, knowledge of gene sequences will not in itself fully define protein function. But such knowledge will allow the creation of comprehensive genome-wide databases of signalling information. For example, protein-interaction maps are being drawn for all the signalling components encoded by yeast and worm genomes; this approach will generate a matrix of all possible protein interactions in these organisms<sup>19</sup>. And the function of each gene in these organisms is being systematically knocked out, one at a time, in the effort to learn more<sup>20</sup>. Another approach is to use DNA microarrays to study the level of gene expression in response to activation of different pathways or deletion of different genes<sup>21</sup>. For example, by searching for similarities in gene-expression profiles between mutant yeast lacking a specific gene and normal yeast treated with a drug of unknown function, it can be possible to determine how and in what pathway the drug works.

Ultimately, genomic approaches like this, combined with conventional techniques, may lead to the generation of comprehensive diagrams of cellular signalling pathways. But we will have to learn a great deal more about handling complexity before we will be able to make complete sense of such a mass of data. It is likely that information in these quantities will never be easily understandable without the assistance of computer modelling. At

present, modelling of signalling networks is in its infancy, although relatively simple parts of signal-transduction pathways have been modelled successfully<sup>22</sup>. Comprehensive web-based databases of information about signalling pathways are being built, and may in the future progress from being mere descriptive tools to predictive ones. ■

*Julian Downward is in the Signal Transduction Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK. e-mail: downward@icrf.icnet.uk*

- Schlessinger, J. *Cell* **103**, 211–225 (2000).
- Hunter, T. *Phil. Trans. R. Soc. Lond. B* **353**, 583–605 (1998).
- Cohen, P. *Trends Biochem. Sci.* **25**, 596–601 (2000).
- Schlaepfer, D. D. & Hunter, T. *Trends Cell Biol.* **8**, 151–157 (1998).
- Daniel, J. M. & Reynolds, A. B. *Bioessays* **19**, 883–891 (1997).
- Massague, J. & Chen, Y. G. *Genes Dev.* **14**, 627–644 (2000).
- Pawson, T. & Scott, J. D. *Science* **278**, 2075–2080 (1997).
- Shields, J. M. *et al. Trends Cell Biol.* **10**, 147–154 (2000).
- Kornberg, R. D. *Trends Cell Biol.* **9**, M46–M49 (1999).
- Christophe, D. *et al. Cell Signal.* **12**, 337–341 (2000).
- Pawson, T. *Nature* **373**, 573–580 (1995).
- Whitmarsh, A. J. & Davis, R. J. *Trends Biochem. Sci.* **23**, 481–485 (1998).
- Plowman, G. D. *et al. Proc. Natl Acad. Sci. USA* **96**, 13603–13610 (1999).
- Datta, S. R. *et al. Genes Dev.* **13**, 2905–2927 (1999).
- Fruman, D. A. *et al. Annu. Rev. Biochem.* **67**, 481–507 (1998).
- Marshall, C. J. *Cell* **80**, 179–185 (1995).
- International Human Genome Sequencing Consortium *Nature* **409**, 860–921 (2001).
- Venter, J. C. *et al. Science* **291**, 1304–1351 (2001).
- Walhout, A. J. M. & Vidal, M. *Nature Rev. Mol. Cell Biol.* **2**, 55–62 (2001).
- Oliver, S. G. *et al. Trends Biotechnol.* **16**, 373–378 (1998).
- Schulze, A. & Downward, J. *J. Cell Sci.* **113**, 4151–4156 (2000).
- Endy, D. & Brent, R. *Nature* **409**, 391–395 (2001).
- Hanahan, D. & Weinberg, R. A. *Cell* **100**, 57–70 (2000).
- Withers, D. J. & White, M. *Endocrinology* **141**, 1917–1921 (2000).
- Mauro, M. J. & Druker, B. J. *Curr. Opin. Oncol.* **13**, 3–7 (2001).

**Related websites**

- ♦ [genome-www.stanford.edu/saccharomyces](http://genome-www.stanford.edu/saccharomyces)
- ♦ [www.wormbase.org](http://www.wormbase.org)
- ♦ [www.flybase.org](http://www.flybase.org)
- ♦ [www.ncbi.nlm.nih.gov/genome/guide/human](http://www.ncbi.nlm.nih.gov/genome/guide/human)