Molecular Systems Biology (2005) doi:10.1038/msb4100031

www.molecularsystemsbiology.com
Article number: 2005.0023

molecular systems biology

REPORT

A theoretical framework for specificity in cell signaling

Natalia L Komarova[1,2,*], Xiufen Zou[1,3], Qing Nie[1] and Lee Bardwell[4,*]

[1] Department of Mathematics, University of California, Irvine, CA, USA, [2] Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA, [3] College of Mathematics and Statistics, Wuhan University, Wuhan, China and [4] Department of Developmental and Cell Biology, University of California, Irvine, CA, USA

* Corresponding authors. NL Komarova, Departments of Mathematics and Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA. Tel.: +1 9498241268; Fax: +1 9498247993; E-mail: komarova@uci.edu or L Bardwell, Department of Developmental and Cell Biology, 5205 McGaugh Hall, University of California, Irvine, CA 92697-2300, USA. Tel.: +1 9498246902; Fax: +1 9498244709; E-mail: bardwell@uci.edu

Received 7.6.05; accepted 23.9.05

Different cellular signal transduction pathways are often interconnected, so that the potential for undesirable crosstalk between pathways exists. Nevertheless, signaling networks have evolved that maintain specificity from signal to cellular response. Here, we develop a framework for the analysis of networks containing two or more interconnected signaling pathways. We define two properties, *specificity* and *fidelity*, that all pathways in a network must possess in order to avoid paradoxical situations where one pathway activates another pathway's output, or responds to another pathway's input, more than its own. In unembellished networks that share components, it is impossible for all pathways to have both mutual specificity and mutual fidelity. However, inclusion of either of two related insulating mechanisms—compartmentalization or the action of a scaffold protein—allows both properties to be achieved, provided deactivation rates are fast compared to exchange rates.

Molecular Systems Biology 18 October 2005; doi:10.1038/msb4100031
Subject Categories: computational methods; signal transduction
Keywords: protein kinase; mathematical model; network; scaffold; signal transduction

Introduction

Cells sense and respond to a wide variety of chemical messages, such as hormones like insulin and adrenalin, which originate from other cells or from the environment. Yet, cells use only a limited number of intracellular signaling proteins to transduce this multitude of signals. For instance, some key intermediate modules, such as the mitogen-activated protein kinase (MAPK) cascade, are activated by an astonishingly high percentage of known stimuli (Lewis *et al*, 1998). Hence, different signals are often transmitted by common components, yet elicit distinct (and appropriate) outcomes. An important unsolved problem in cell biology is to understand how specificity from signal to cellular response is maintained between different signal transduction pathways that share similar (or identical) components, particularly when this occurs in the same cell (Schaeffer and Weber, 1999; Tan and Kim, 1999; Pawson, 2004).

Results and discussion

Simple signaling networks

Figure 1A shows a schematic of a simple signaling network, composed of two pathways, *X* and *Y*. Each pathway in the network has a unique input and a unique output. The input for pathway *X* is designated x_0, and can be taken to represent both the signal itself and its receptor; the input for pathway *Y* is designated y_0. We assume that the network is exposed to only one signal at a time. The activated species of the downstream elements of pathway *X* are designated x_1 and x_2, those of pathway *Y* are y_1 and y_2. When a signal x_0 is present, x_0 'activates' (i.e. causes the production of) component x_1 (which might be a protein kinase or a kinase cascade); x_1 in turn activates x_2 (which might be a terminal kinase or a target transcription factor). Pathway components are also deactivated: proteins that are activated by being phosphorylated by a protein kinase are deactivated when that phosphate group is removed by a protein phosphatase, for instance. Figure 1B shows the typical mound-shaped curve of the time course of activation of a given component in response to a square-pulse input signal. The area under such a curve for the final component of a given pathway can be taken as a measure of that pathway's *total output*. Let us denote the total output of pathway *X* when the cell is exposed to an input signal x_0 as $X_{\text{out}}|X_{\text{in}}$ (read as '*X* output given *X* input', or simply '*X* given *X*').

Interconnections between pathways often serve a useful purpose (Schwartz and Baron, 1999), but here we concern ourselves with undesirable crosstalk, or 'leaking'. In the Figure 1 network, pathway *Y* leaks into pathway *X*, because kinase y_1 is somewhat lacking in substrate selectivity: in addition to phosphorylating its correct target y_2, it also

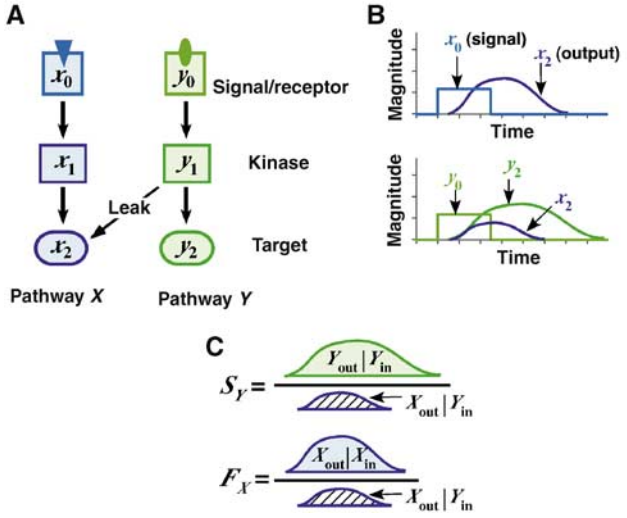

Figure 1 A simple network with crosstalk. (**A**) The network consists of two pathways, *X* and *Y*, that are interconnected because component y_1 activates target x_2. (**B**) Output in response to pulse of signal x_0 (top) or y_0 (bottom). (**C**) Depiction of the ratios equal to the specificity of pathway *Y* and the fidelity of pathway *X*.

phosphorylates the incorrect target x_2. Hence, when the network is stimulated by signal y_0, in addition to the authentic output $Y_{\text{out}}|Y_{\text{in}}$, there is some spurious output $X_{\text{out}}|Y_{\text{in}}$.

Definitions of specificity and fidelity

We define the *specificity* of cascade *X* as the ratio of its authentic output to its spurious output:

$$S_X = \frac{X_{\text{out}}|X_{\text{in}}}{Y_{\text{out}}|X_{\text{in}}} \qquad (1)$$

Thus (as in Figure 1), if pathway *X* is activated by a given signal and this does not affect the output from pathway *Y*, the specificity of *X* with respect to *Y* in response to that signal is infinite, or *complete*.

Similarly, the specificity of cascade *Y* is of the form

$$S_Y = \frac{Y_{\text{out}}|Y_{\text{in}}}{X_{\text{out}}|Y_{\text{in}}} \qquad (2)$$

In Figure 1, since the action of signal y_0 will result in some output from *X*, the specificity of *Y* with respect to *X* is finite (see Figure 1C). Indeed, if S_Y were less than 1, it would mean that the signal for *Y* was actually promoting the output of pathway *X* more than its own output.

The *fidelity* of a pathway is its output when given an authentic signal divided by its output in response to a spurious signal (see Figure 1C):

$$F_X = \frac{X_{\text{out}}|X_{\text{in}}}{X_{\text{out}}|Y_{\text{in}}}, \quad F_Y = \frac{Y_{\text{out}}|Y_{\text{in}}}{Y_{\text{out}}|X_{\text{in}}} \qquad (3)$$

Thus, a pathway that exhibits fidelity (i.e. $F>1$) is activated more by its authentic signal than by others. In contrast, if a pathway has fidelity of less than 1, it is activated more by another pathways' signal than it is by its own. In the Figure 1 network, F_Y is complete, while F_X is finite.

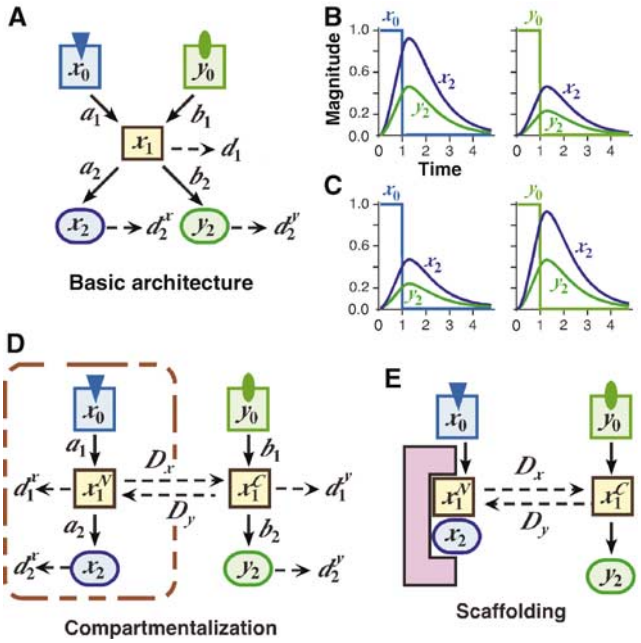

Figure 2 Signaling network with shared components. (**A**) The 'basic architecture'. Component x_1 is common to pathways *X* and *Y*. Although the desired route of signaling is for x_0 to activate x_2 and not y_2, and y_0 to activate y_2 and not x_2, this cannot be achieved with specificity and fidelity for this network. (**B, C**) Numerical simulations of signaling through this network under various sets of parameter values. Values that increase S_X reciprocally decrease S_Y, and values that increase F_X reciprocally decrease F_Y. Shown are the values of outputs x_2 and y_2 in response to inputs x_0 and y_0, applied separately as square pulses of magnitude 1 and duration 1. In panel B, both the specificity and fidelity of cascade *X* are larger than those of cascade *Y*. Parameter values are a_1=2, b_1=1, a_2=2, b_2=1, $d_1 = d_2^x = d_2^y = 1$. We have S_X=2, F_X=2, S_Y=0.5, F_Y=0.5. In panel C, the specificity of cascade *X* is higher than that of cascade *Y* whereas the opposite holds for fidelity values: $F_X < F_Y$. Parameter values are a_1=1, b_1=2, a_2=2, b_2=1, $d_1 = d_2^x = d_2^y = 1$. This yields S_X=2, F_X=0.5, S_Y=0.5, F_Y=2. (**D, E**) Two insulating mechanisms that can augment specificity: (D) compartmentalization; (E) the action of a scaffold protein.

Cascades that share components

In many cases, two signaling pathways share one or more common elements (see Figure 2A). One well-known example is in mammalian PC12 cells, where treatment with epidermal growth factor (EFG) causes the cells to proliferate, whereas treatment with nerve growth factor (NFG) causes the cells to differentiate and sprout neurites, yet both growth factors signal through the same MAPK cascade (Marshall, 1995). Another example is in baker's and brewer's yeast (*Saccharomyces cerevisiae*), where three distinct signaling pathways (mating, invasive growth and osmotic stress response) share elements of the same MAP kinase cascade (van Drogen and Peter, 2002). Experimental data indicate that pathways can be well insulated from one another despite sharing components: treatment of PC12 cells with EGF does not cause them to sprout neurites, and stimulation of yeast with mating pheromone does not activate the stress response, for example (Schaeffer and Weber, 1999; van Drogen and Peter, 2002; Vaudry *et al*, 2002).

This class of networks can be represented by the 'basic architecture' shown in Figure 2A. Here, the parameters a_1 and a_2 are activation rate constants for pathway *X*; a_2 is the rate at which kinase x_1 activates (phosphorylates) target x_2. Similarly,

Table I Equations and solutions for the networks analyzed in this paper[a]

	'Basic architecture' network (Figure 2A)	Compartmentalization/scaffolding network (Figure 2D and E)
Equation(s) for x_1	$dx_1/dt = a_1 x_0(t) + b_1 y_0(t) - d_1 x_1$	$dx_1^N/dt = a_1\, x_0(t) - D_x x_1^N + D_y x_1^C - d_1^x x_1^N$, $dx_1^C/dt = b_1\, y_0(t) - D_y x_1^C + D_x x_1^N - d_1^y x_1^C$
Equations for x_2, y_2	$dx_2/dt = a_2 x_1 - d_2^x x_2$, $dy_2/dt = b_2 x_1 - d_2^y y_2$	$dx_2/dt = a_2 x_1^N - d_2^x x_2$, $dy_2/dt = b_2 x_1^C - d_2^y y_2$
$\bar{x}_1 \vert X_{in}$	$\frac{\bar{x}_0 a_1}{d_1}$	[b]
$X_{out} \vert X_{in}$ $(= \bar{x}_2 \vert X_{in})$	$\frac{\bar{x}_0 a_1 a_2}{d_1 d_2^x}$	[b]
$Y_{out} \vert X_{in}$ $(= \bar{y}_2 \vert X_{in})$	$\frac{\bar{x}_0 a_1 b_2}{d_1 d_2^y}$	[b]
$\bar{x}_1 \vert Y_{in}$	$\frac{\bar{y}_0 b_1}{d_1}$	[b]
$X_{out} \vert Y_{in}$	$\frac{\bar{y}_0 b_1 a_2}{d_1 d_2^x}$	[b]
$Y_{out} \vert Y_{in}$	$\frac{\bar{y}_0 b_1 b_2}{d_1 d_2^y}$	[b]
$S_X = \frac{X_{out} \vert X_{in}}{Y_{out} \vert X_{in}}$	$\frac{a_2 d_2^y}{b_2 d_2^x}$	$\frac{a_2 d_2^y}{b_2 d_2^x} \frac{(d_1^y + D_y)}{D_x}$
$S_Y = \frac{Y_{out} \vert Y_{in}}{X_{out} \vert Y_{in}}$	$\frac{b_2 d_2^x}{a_2 d_2^y}$	$\frac{b_2 d_2^x}{a_2 d_2^y} \frac{(d_1^x + D_x)}{D_y}$
$F_X = \frac{X_{out} \vert X_{in}}{X_{out} \vert Y_{in}}$	$\frac{\bar{x}_0 a_1}{\bar{y}_0 b_1}$	$\frac{\bar{x}_0 a_1}{\bar{y}_0 b_1} \frac{(d_1^y + D_y)}{D_y}$
$F_Y = \frac{Y_{out} \vert Y_{in}}{Y_{out} \vert X_{in}}$	$\frac{\bar{y}_0 b_1}{\bar{x}_0 a_1}$	$\frac{\bar{y}_0 b_1}{\bar{x}_0 a_1} \frac{(d_1^x + D_x)}{D_x}$
$S_{network} = S_X S_Y$	1	$\frac{(d_1^y + D_y)(d_1^x + D_x)}{D_y D_x}$

[a] $\bar{x}_0$, $\bar{y}_0$ = the total amount of signal applied; $\bar{x}_1 \vert X_{in}$ = the total amount of product x_1 in response to signal x_0, and so on; other parameters and expressions are defined in text.
[b] See Supplementary information for full solutions.

b_1 and b_2 are activation rate constants for pathway *Y*. Finally, d_1^x, d_2^x and d_2^y are deactivation rate constants, and can be thought of as representing phosphatase activity, for example. Assuming that both pathways in the network are *weakly activated* (Heinrich *et al*, 2002; Chaves *et al*, 2004), the network can be modeled by a simple system of three linear ordinary differential equations (ODEs) that can be solved to yield precise analytical expressions for pathway outputs, specificities and fidelities in terms of the network parameters (see Table I; also, see Supplementary information for detailed solutions). It can be seen (see Table I) that the specificities of the pathways are independent of the signal strength, and indeed of all parameters that lie upstream or at the level of the shared component. Pathway fidelities, in contrast, depend strongly upon the relative signal strengths and upon the values of upstream parameters. Hence, the two performance metrics, specificity and fidelity, depend on different characteristics of network design. Indeed, it is easy to choose parameters that provide one pathway with specificity but not fidelity.

We define *network specificity* as the product of the pathway specificities:

$$S_{network} = S_X S_Y = \frac{X_{out} \vert X_{in}}{Y_{out} \vert X_{in}} \frac{Y_{out} \vert Y_{in}}{X_{out} \vert Y_{in}} \quad (4)$$

(Note that network fidelity, the product of the pathway fidelities, is always exactly equal to network specificity.) $S_{network}$ provides an indication of the specificity intrinsic in the network architecture. Intuitively, it would seem that the basic architecture does not possess intrinsic specificity. Indeed, it can be seen from Table I that, for the basic architecture, S_X is the reciprocal of S_Y, and F_X is the reciprocal of F_Y, so that $S_{network} = F_{network} = 1$. The specificity of pathway *X* can be increased by changing the magnitude of certain parameters (increasing a_2 or decreasing b_2, for example), but in so doing the specificity of *Y* decreases correspondingly (see Figure 2B and C).

Two other useful network measurements are *mutual specificity* (and *mutual fidelity*), properties that exist if all pathways in the network have specificity (fidelity) greater than 1. The basic architecture never exhibits mutual specificity or mutual fidelity.

Analysis of insulating mechanisms: compartmentalization

Real cellular signaling networks that share components typically contain one or more insulating mechanisms that

are thought to contribute to specificity and fidelity (Tan and Kim, 1999; Schwartz and Madhani, 2004). In *compartmentalization*, different pathways are localized to different cellular compartments, or to different spatial locations within the cell (Figure 2D) (Smith and Scott, 2002; White and Anderson, 2005). The extent of leaking between the two pathways is determined by the efficiency of compartmentalization. For example, assume that the pathway-specific components of pathway X are localized to the nucleus, while those of pathway Y are localized to the cytosol. Although the shared kinase, x_1, is found in both compartments (x_1^{N} is the nuclear pool and x_1^{C} is the cytosolic pool), x_1 activated by x_0 in the nucleus is likely to encounter target x_2, which is also in the nucleus; it will only encounter target y_2 if it diffuses into the cytosol before it is deactivated. Thus, crossover between the two pathways happens when kinase x_1 leaks in or out of the nucleus. D_x is the coefficient for the rate at which x_1 exits the nucleus and enters the cytosol, and D_y is the rate constant for x_1 leaving the cytosol and entering the nucleus. D_x and D_y can be considered as pseudo-diffusion rate constants, or exchange rate constants. The parameters d_1^x and d_1^y are the deactivation constants for x_1 in the nucleus and cytosol, respectively.

Again, assuming weak activation, the network can be modeled with linear ODEs and precise solutions for specificity and fidelity obtained (see Table I and Supplementary information). The specificity of this network is

$$S_{\mathrm{network}} = F_{\mathrm{network}} = \frac{(d_1^x + D_x)(d_1^y + D_y)}{D_x D_y} > 1 \qquad (5)$$

It can be seen that network specificity is greater than in the basic architecture, and is maximized if the exchange rates D_x and D_y are small compared to the deactivation rates d_1^x and d_1^y. Compartmentalization can also provide both mutual specificity and mutual fidelity, as long as the exchange rates balance each other (see Table I). The limiting case where $D_x = D_y = 0$ is equivalent to two noninteracting cascades with complete specificity and fidelity. If the leakage coefficients become very large ($D_x,\ D_y \to \infty$), we again have a fully connected system with a shared element, equivalent to the basic architecture, and $S_{\mathrm{network}} = F_{\mathrm{network}} = 1$.

Role of scaffold proteins

Scaffold proteins, defined here as proteins that bind to two or more consecutively acting components in a pathway, have been shown to enhance the efficiency of signaling, and have also been proposed to augment specificity by several mechanisms (Whitmarsh and Davis, 1998; Levchenko *et al*, 2000; Burack *et al*, 2002; Flatauer *et al*, 2005). In particular, by binding to multiple components of a given pathway, scaffolds may create the equivalent of 'micro-compartments' (Harris *et al*, 2001). That is, if the reactions between those components can only happen on the scaffold (or are much more efficient on the scaffold), then it is as if these scaffolded reactions occur in their own compartment, sequestered away from reactions occurring off-scaffold. In this way, scaffolds may prevent their bound components that are shared with other pathways from straying into those pathways, and protect them from intrusions from those pathways.

To model this sequestration mechanism, we use the equations for compartmentalization, with the meaning of some of the terms interpreted differently (see Figure 2E and Table I). First, x_1^{N} (aNchored x_1) is interpreted to represent kinase x_1 bound to the scaffold and x_1^{C} (Cytosolic x_1) is unbound x_1, free in solution in the cytosol. The equation for $\mathrm{d}x_1^{\mathrm{N}}/\mathrm{d}t$ then indicates that the activation of kinase x_1 by signal x_0 occurs on the scaffold and not in solution, while the equation for $\mathrm{d}x_2/\mathrm{d}t$ indicates that the activation of target x_2 by kinase x_1 also occurs only on the scaffold. In contrast, the corresponding reactions for pathway Y can occur only in solution and not on the scaffold. D_x is the rate constant for the dissociation of x_1 from the scaffold, and D_y is a first-order association constant for the binding of cytosolic x_1 to the scaffolded complex. Leaking between pathways X and Y can occur if x_0-activated x_1 dissociates from the scaffold and encounters y_2, or if x_1 that was activated by y_0 in the cytosol binds to the scaffold (see Figure 2E).

The previous results (see equation (5) and Table I) for specificity and fidelity under compartmentalization also apply to scaffolding: S_X is promoted by a low rate of dissociation of kinase x_1 from the scaffold and a high rate of rebinding; however, these factors reduce S_Y. Obtaining network specificity again requires that deactivation rates be fast relative to exchange rates, so that, for instance, any x_1 that dissociates from the scaffold will be deactivated before it encounters y_2. In this model, d_1^x represents the deactivation of kinase x_1 on the scaffold. Interestingly, one way in which it has been proposed that scaffold proteins might enhance signal transmission is by protecting their bound kinases from the action of phosphatases (Levchenko *et al*, 2000; Heinrich *et al*, 2002), equivalent to lowering d_1^x to close to or equal to zero. Although this might indeed enhance the speed, duration and amplitude of X signaling (Heinrich *et al*, 2002), it would lower both S_Y, F_Y and network specificity.

Conclusion

Here, we presented a framework for the analysis of interconnected biochemical pathways. We defined the *specificity* of a pathway as the ratio of its authentic output to its spurious output, and the *fidelity* of a pathway as its output when given an authentic signal divided by its output in response to a spurious signal. These definitions express commonsense notions that a pathway should stimulate its own output more than another pathway's output, and respond to its own input more than to another's. Moreover, they are simple ratios of pathway output, a property that is readily measurable by modeling or experiment. We also defined the informative metric of *network specificity*, the product of pathway specificities or fidelities. We demonstrated the utility of these metrics by calculating them for simple networks that share components, revealing the limited specificity inherent in simple architectures devoid of specificity-promoting enhancements. Finally, we showed how the insulating mechanisms of compartmentalization and scaffolding are related, and how both require slow exchange rates and fast deactivation rates in order to promote high levels of specificity.

Supplementary information

Supplementary information is available at *Molecular Systems Biology* website (www.nature.com/msb).

Acknowledgements

This work was supported by a National Institutes of Health/National Science Foundation joint initiative on Mathematical Biology through National Institute of General Medical Sciences grant GM75309 (LB, NK, QN), by NIGMS grant P20-GM66051 (LB, QN), by a grant from the National Academies of Sciences Keck Futures Initiative and by NIGMS research grants GM60366 and GM69013 (LB) and by NIGMS grant GM67247 (QN). XZ is supported by Chinese National Natural Science Foundation grant no. 60573168.

References

Burack WR, Cheng AM, Shaw AS (2002) Scaffolds, adaptors and linkers of TCR signaling: theory and practice. *Curr Opin Immunol* **14:** 312–316

Chaves M, Sontag ED, Dinerstein RJ (2004) Optimal length and signal amplification in weakly activated signal transduction cascades. *J Phys Chem B* **108:** 15311–15320

Flatauer LJ, Zadeh SF, Bardwell L (2005) Mitogen-activated protein kinases with distinct requirements for Ste5 scaffolding influence signaling specificity in *Saccharomyces cerevisiae*. *Mol Cell Biol* **25:** 1793–1803

Harris K, Lamson RE, Nelson B, Hughes TR, Marton MJ, Roberts CJ, Boone C, Pryciak PM (2001) Role of scaffolds in MAP kinase pathway specificity revealed by custom design of pathway-dedicated signaling proteins. *Curr Biol* **11:** 1815–1824

Heinrich R, Neel BG, Rapoport TA (2002) Mathematical models of protein kinase signal transduction. *Mol Cell* **9:** 957–970

Levchenko A, Bruck J, Sternberg P (2000) Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. *Proc Natl Acad Sci USA* **97:** 5818–5823

Lewis TS, Shapiro PS, Ahn NG (1998) Signal transduction through MAP kinase cascades. *Adv Cancer Res* **74:** 49–139

Marshall CJ (1995) Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell* **80:** 179–185

Pawson T (2004) Specificity in signal transduction: from phosphotyrosine–SH2 domain interactions to complex cellular systems. *Cell* **116:** 191–203

Schaeffer HJ, Weber MJ (1999) Mitogen-activated protein kinases: specific messages from ubiquitous messengers. *Mol Cell Biol* **19:** 2435–2444

Schwartz MA, Baron V (1999) Interactions between mitogenic stimuli, or, a thousand and one connections. *Curr Opin Cell Biol* **11:** 197–202

Schwartz MA, Madhani HD (2004) Principles of MAP kinase signaling specificity in *Saccharomyces cerevisiae*. *Annu Rev Genet* **38:** 725–748

Smith FD, Scott JD (2002) Signaling complexes: junctions on the intracellular information super highway. *Curr Biol* **12:** R32–R40

Tan PB, Kim SK (1999) Signaling specificity: the RTK/RAS/MAP kinase pathway in metazoans. *Trends Genet* **15:** 145–149

van Drogen F, Peter M (2002) MAP kinase cascades: scaffolding signal specificity. *Curr Biol* **12:** R53–R55

Vaudry D, Stork PJ, Lazarovici P, Eiden LE (2002) Signaling pathways for PC12 cell differentiation: making the right connections. *Science* **296:** 1648–1649

White MA, Anderson RG (2005) Signaling networks in living cells. *Annu Rev Pharmacol Toxicol* **45:** 587–603

Whitmarsh AJ, Davis RJ (1998) Structural organization of MAP-kinase signaling modules by scaffold proteins in yeast and mammals. *Trends Biochem Sci* **23:** 481–485