

NEWS AND VIEWS

Circuit diagrams for biological networks

Kurt W Kohn and Mirit I Aladjem

Laboratory of Molecular Pharmacology, National Cancer Institute, NIH, Bethesda, MD, USA

Molecular Systems Biology 17 January 2006; doi:10.1038/msb4100044

Metabolic maps have long been a staple of biochemistry students, providing clear and concise charts depicting the flow of metabolites and energy in cells. However, depicting the molecular networks involved in signaling pathways that regulate cell function have proven challenging, due to the enormous amount of information that needs to be conveyed for each participant in the network and the cross-connections between pathways. This challenge must nevertheless be addressed in order to understand the underlying design of such networks, and to utilize the findings of modern biology most effectively to combat diseases, such as cancers, that arise from defects in cell regulation. Another difficulty is that bioregulatory networks are replete with interconnections and loops that make intuition about network function unreliable; therefore, computer simulations may be needed. In a recent issue of *Nature Biotechnology*, Kitano *et al* (2005) describe a notation for biological network diagrams, 'process diagrams', the formalism of which allows a straightforward conversion of human-readable diagrams into machine-readable documents.

Hiroaki Kitano and his Symbiotic Systems Project have in recent years focused their attention on bioregulatory networks, how they convey functionality and robustness on biological organisms and how they can be diagrammed and simulated. Kitano (2004) recently reviewed the fascinating field of biological robustness, and insightfully discussed the features that convey robustness to a network. Aiming for a deep understanding of biological networks, he and his colleagues devised useful tools. One tool is the systems biology markup language (SBML; www.sbml.org) that allows scientists to integrate and exchange data in a clear, unambiguous and standardized format (Hucka *et al*, 2003). A second is CellDesigner, a tool for creating 'process diagrams' that can be linked to computer-readable SBML files (Funahashi *et al*, 2003). Kitano and co-workers envision that process diagrams will form the basis of a Systems Biology Graphical Notation (SBGN), which will be developed as a web-based community effort (www.sbgm.org). Importantly, these tools are freely available. The essential feature of process diagrams is that they show the event sequences or pathways in a network.

Kitano *et al* (2005) now describe the process diagram notation formally as consisting of 'state nodes' and 'transition nodes'. State nodes represent entities in the biological process, such as proteins, RNA or genes, and transition nodes represent modulations of the reactions, such as association, dissociation, activation or inhibition. 'Edges' are defined as going between state nodes and transition nodes, or vice versa. This leads to a connectivity matrix that serves to define networks in a manner suitable for computer simulation.

Visually, process diagrams show each occurrence of a molecular species fully, including the name of each monomolecular component and its modification states (Figure 1A). This makes the individual reactions easy to interpret. The trade-off is that molecular species that engage in multiple reactions have to be represented multiple times, which can make it difficult to survey the full set of interactions of a given molecular species. Also, it is sometimes difficult to discern the particular modification (e.g., phosphorylation) introduced by an enzyme (such as a protein kinase) because all of the potential modification sites of reactant and product must be compared; in large diagrams the modification symbols tend to be small and difficult to read. Another difficulty, 'combinatorial explosion', occurs when multiple paths lead from one point to another in a network. Such a situation is difficult to depict in process diagrams, because all of the paths must be explicitly shown. Thus, in addition to the full graphics, the authors define a reduced notation, in which the graphics are simplified by inserting text into the molecular symbols to indicate modification states.

The authors discuss the complementarity between process diagrams and 'entity-relationship diagrams', such as the molecular interaction maps that they refer to as 'Kohn diagrams' (Kohn, 1999, 2001, 2005). The latter type of diagrams differ from Kitano's process diagrams mainly in two ways: (1) they show each named molecular species only once on a map; (2) they do not specify particular event sequences, but instead show all of the interactions that can occur if potentially interacting species are in the same place at the same time (Figure 1B). The authors argue that the different types of diagrams can serve different purposes.

Using epidermal growth factor receptor (EGFR) as an example, the authors show how their process diagram notation can describe a complex signal transduction pathway. In several on-line supplements, they also show simple examples of how the notation can depict a variety of different circumstances.

To illustrate the scalability of the notation, Kitano and co-workers used the CellDesigner software to construct (manually) a comprehensive pathway map of signaling from EGFR-family receptors, comprising 211 reactions and 322 molecular species, all stored in SBML (Oda *et al*, 2005). The map is copiously referenced, but it is not easy to find the source of the evidence supporting particular reactions. Although the network is not ready for computer simulation, the authors sought insights into the network's architecture. They noted a global feature that Kitano has described as a 'bow-tie' or 'hourglass' architecture that may contribute to network robustness

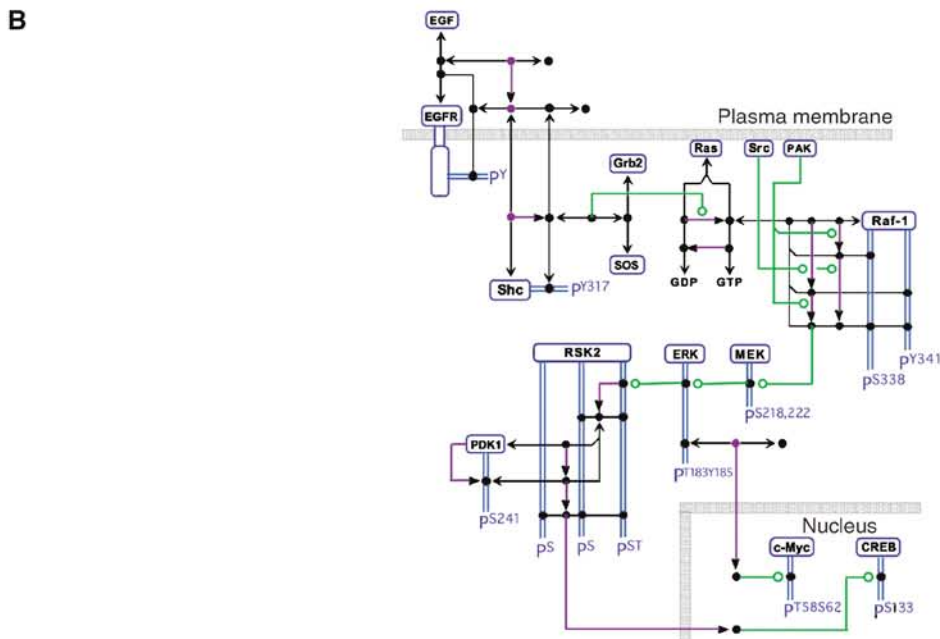
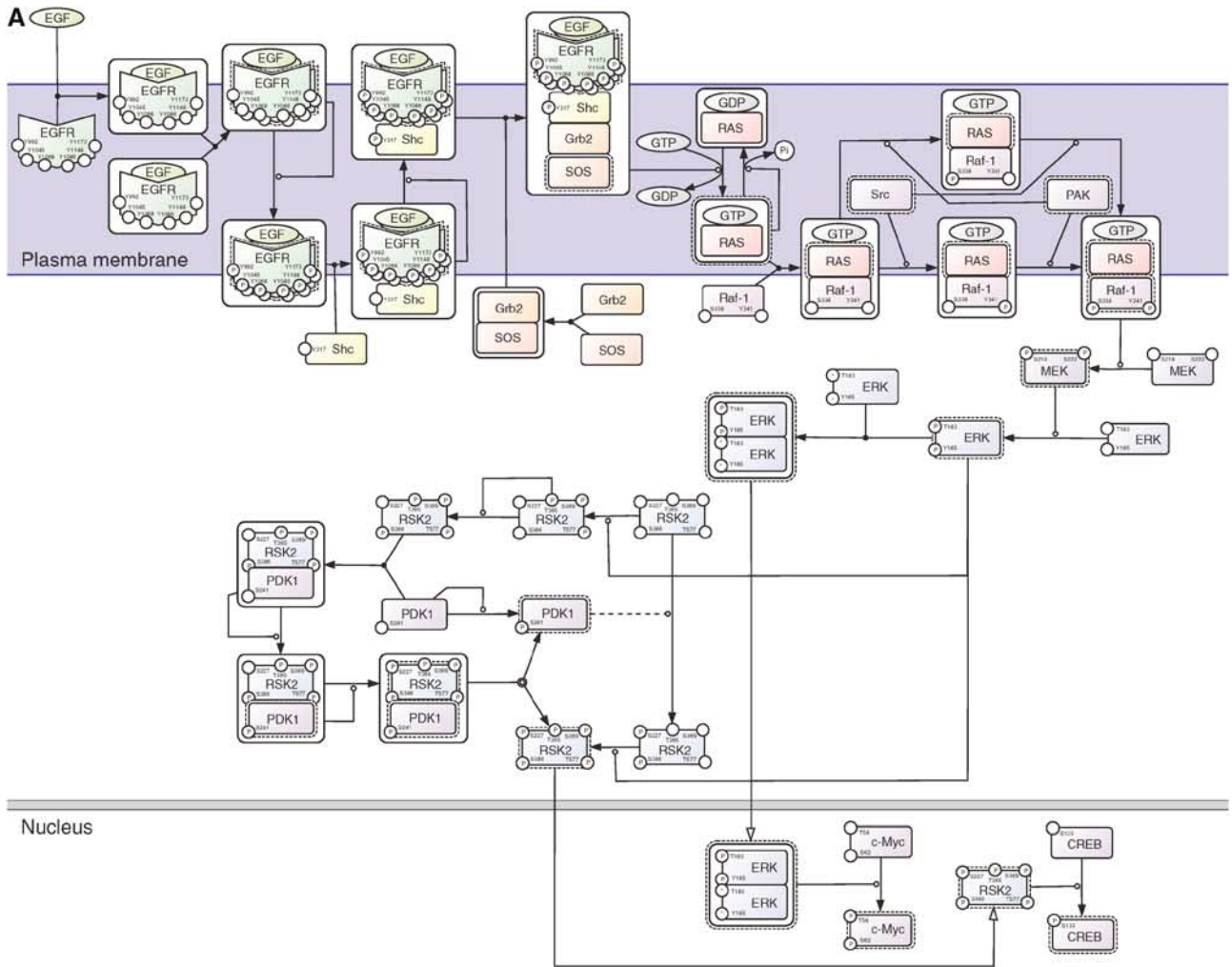


Figure 1 An SBGN process diagram (**A**) and a molecular interaction map (**B**) of a signaling pathway triggered by ligand binding to EGFR molecules. Both diagrams depict ligand-activated phosphorylation and dimerization of EGFR, followed by a phosphorylation cascade involving membrane-bound and cytoplasmic kinases. Phosphorylated kinases (ERK, Rsk2) can transport to the nucleus, where they may activate transcription via the phosphorylation of nuclear transcription factors (myc and CREB). Panel A depicts this pathway as a process diagram (for symbols, see Kitano *et al*, 2005; Oda *et al*, 2005); panel B depicts the same pathway as a molecular interaction map (for symbols, see Kohn, 1999, 2001; Kohn *et al*, 2005).

(Kitano, 2004). They emphasized this feature in an abbreviated diagram that shows many inputs from EGFRs to a relatively few 'core' molecules (such as Ras, phosphatidylinositol phosphates, and the MAP kinases), which then transmit the signal to many outputs (such as transcription factors controlling many genes). This architecture essentially comprises signals from many sources, passed via a limited number of transmission lines to many outputs. Functional versatility may arise by linkage of inputs to outputs in different ways for different cell states.

Although the SBML format is already accepted by many simulation software, it remains to be seen how SBGN process diagrams will be used. The process diagram notation—like all of the notation schemes that have been proposed—has advantages and disadvantages. Different notations may be best suited to different purposes. It may be hoped, however, that a common graphical language will develop that will be as useful in systems biology as circuit diagrams are in electronics.

Acknowledgements

Research in the authors' laboratories was supported by the Intramural Research Program of the National Cancer Institute, CCR, NIH.

References

- Funahashi A, Morohashi M, Kitano H (2003) CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *Biosilico* **1**: 159–162
- Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novère N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J, SBML Forum (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**: 524–531
- Kitano H (2004) Biological robustness. *Nat Rev Genet* **5**: 826–837
- Kitano H, Funahashi A, Matsuoka Y, Oda K (2005) Using process diagrams for the graphical representation of biological networks. *Nat Biotechnol* **23**: 961–966
- Kohn KW (1999) Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Mol Biol Cell* **10**: 2703–2734
- Kohn KW (2001) Molecular interaction maps as information organizers and simulation guides. *Chaos* **11**: 84–97
- Kohn KW, Aladjem MI, Weinstein JN, Pommier Y (2005) Molecular interaction maps of bioregulatory networks: a general rubric for systems biology. *Mol Biol Cell* (E-pub ahead of print)
- Oda K, Matsuoka Y, Funahashi A, Kitano H (2005) A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol Syst Biol* 25 May 2005; doi: 10.1038/msb4100014