PERSPECTIVE

The Systems Biology Graphical Notation

Nicolas Le Novère[1], Michael Hucka[2], Huaiyu Mi[3], Stuart Moodie[4], Falk Schreiber[5,6], Anatoly Sorokin[7], Emek Demir[8], Katja Wegner[9], Mirit I Aladjem[10], Sarala M Wimalaratne[11], Frank T Bergman[12], Ralph Gauges[13], Peter Ghazal[4,14], Hideya Kawaji[15], Lu Li[1], Yukiko Matsuoka[16], Alice Villéger[17,18], Sarah E Boyd[19], Laurence Calzone[20], Melanie Courtot[21], Ugur Dogrusoz[22], Tom C Freeman[14,23], Akira Funahashi[24], Samik Ghosh[16], Akiya Jouraku[24], Sohyoung Kim[10], Fedor Kolpakov[25,26], Augustin Luna[10], Sven Sahle[13], Esther Schmidt[1], Steven Watterson[4,22], Guanming Wu[27], Igor Goryanin[4], Douglas B Kell[18,28], Chris Sander[8], Herbert Sauro[12], Jacky L Snoep[29], Kurt Kohn[10] & Hiroaki Kitano[16,30,31]

Circuit diagrams and Unified Modeling Language diagrams are just two examples of standard visual languages that help accelerate work by promoting regularity, removing ambiguity and enabling software tool support for communication of complex information. Ironically, despite having one of the highest ratios of graphical to textual information, biology still lacks standard graphical notations. The recent deluge of biological knowledge makes addressing this deficit a pressing concern. Toward this goal, we present the Systems Biology Graphical Notation (SBGN), a visual language developed by a community of biochemists, modelers and computer scientists. SBGN consists of three complementary languages: process diagram, entity relationship diagram and activity flow diagram. Together they enable scientists to represent networks of biochemical interactions in a standard, unambiguous way. We believe that SBGN will foster efficient and accurate representation, visualization, storage, exchange and reuse of information on all kinds of biological knowledge, from gene regulation, to metabolism, to cellular signaling.

"Un bon croquis vaut mieux qu'un long discours" ("A good sketch is better than a long speech"), said Napoleon Bonaparte. This claim is nowhere as true as for technical illustrations. Diagrams naturally engage innate cognitive faculties[1] that humans have possessed since before the time of our cave-drawing ancestors. Little wonder that we find ourselves turning to them in every field of endeavor. Just as with written human languages, communication involving diagrams requires that authors and readers agree on symbols, the rules for arranging them and the interpretation of the results. The establishment and widespread use of standard notations have permitted many fields to thrive. One can hardly imagine today's electronics industry, with its powerful, visually oriented design and automation tools, without having first established standard notations for circuit diagrams. Such was not the case in biology[2]. Despite the visual nature of much of the information exchange, the field was permeated with *ad hoc* graphical notations having little in common between different researchers, publications, textbooks and software tools. No standard visual language existed for describing biochemical interaction networks, inter- and intracellular signaling gene regulation—concepts at the core of much of today's research in molecular, systems and synthetic biology. The closest to a standard is the notation long used in many metabolic and signaling pathway maps, but in reality, even that lacks uniformity between sources and suffers from undesirable ambiguities (**Fig. 1**). Moreover, the existing tentative representations, however well crafted, were ambiguous, and only suitable for specific needs, such as representing metabolic networks or signaling pathways or gene regulation.

The molecular biology era, and more recently the rise of genomics and other high-throughput technologies, have brought a staggering increase in data to be interpreted. It also favored the routine use of software to help formulate hypotheses, design experiments and interpret results. As a group of biochemists, modelers and computer scientists working in systems biology, we believe establishing standard graphical notations is an important step toward more efficient and accurate transmission of biological knowledge among our different communities. Toward this goal, we initiated the SBGN project in 2005, with the aim of developing and standardizing a systematic and unambiguous graphical notation for applications in molecular and systems biology.

Historical antecedents

Graphical representation of biochemical and cellular processes has been used in biochemical textbooks as far back as sixty years ago[3], reaching an apex in the wall charts hand drawn by Nicholson[4] and Michal[5]. Those graphs describe the processes that transform a set of inputs into a set of outputs, in effect being process, or state transition, diagrams. This style was emulated in the first database systems that depicted metabolic networks, including EMP[6], EcoCyc[7] and KEGG[8]. More notations have been 'defined' by virtue of their implementation in specialized software tools such as pathway and network designers (e.g., NetBuilder[9], Patika[10], JDesigner[11], CellDesigner[12]). Those

---
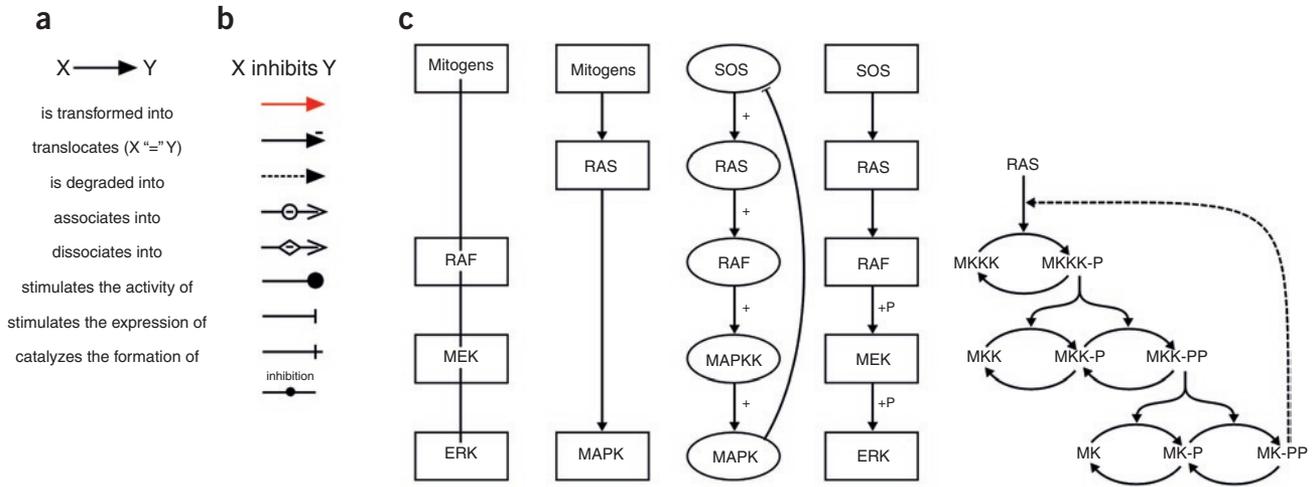
Figure 1 Inconsistency and ambiguity of current nonstandardized notations. (**a**) Eight different meanings associated with the same symbol in a chart describing the role of cyclin in cell regulations (http://www.abcam.com/ps/pdf/nuclearsignal/cell_cycle.pdf). (**b**) Nine different symbols found in the literature to represent the same meaning. (**c**) Five different representations of the MAP kinase cascade found in the scientific literature, depicting progressive levels of biological and biochemical knowledge. From left to right: relations[30], directionality of influence[31], directionality of effect[32], biochemical effect[33], chemical reactions[34]. In the last diagram, different instances of an identical arrowhead style represent catalysis, production and inhibition.

graphical notations were not standardized, and their understanding relied mainly on relating examples with one's preexisting knowledge of biochemical processes. Although the classical graphs adequately conveyed information about biochemistry, other types of diagrams were needed to represent signaling pathways, and incomplete or indirect information, as coming from molecular biology or genomics. Those conventions effectively mimicked the empirical notations used by biologists, describing either the relationships between elements[13,14] or the flow of activity or influence[15–17]. Lists of standard glyphs (**Box 1**) to represent identified concepts were then provided. The efforts to create rigidly defined schema were pioneered by Kurt Kohn with his Molecular Interaction Maps (MIM), which defined not only a set of symbols but also a syntax to describe interactions and relationships of molecules[18,19]. The MIM notation influenced other proposals[14]. Several proposals followed to describe process diagrams, not only with standard symbols but also defined grammars[20–23].

The SBGN project
Despite the popularity of some of the efforts mentioned above, none of the notations has acquired the status of a community standard. This can be attributed partly to the fact that the efforts only went as far as to propose notations, or implement them in software. Several of us have been involved in the development of the Systems Biology Markup Language (SBML)[24], from which we learned that establishing a standard is extremely difficult without an explicit, concerted, effort to engage a community and build a consensus among participants. We organized the SBGN project with this lesson in mind.

For SBGN to be successful, it must satisfy a majority of technical and practical needs and be embraced by a diverse community of biologists, biochemists, bioinformaticians, geneticists, theoreticians and software engineers. Early in the project's history, we established the following overarching principles to help steer SBGN toward those aims, ranked by rough hierarchical order of precedence.

The notation should
• be free of intellectual property restrictions to allow free use by the community;
• be syntactically and semantically consistent and unambiguous;

• support representation of diverse common biological objects, their properties and their interactions;
• keep the number of symbols and syntax to a minimum to help comprehension and learning by humans;
• be visually consistent and concise, using discriminable symbols;
• support modularity to help cope with diagram size and complexity;
• support the automated generation of diagrams by software starting from mathematical models.

Many of the design principles above resonate with research on visual languages[25,26] and studies aimed at understanding end-user needs in pathway visualization[27], although we derived them from our collective hands-on experiences with developing notations and software. In addition to these principles, we also sought to avoid many problems (**Table 1**) that affect some existing notations.

SBGN aims to specify the connectivity of the graphs and the types of the nodes and edges, but not the precise layout of the graphs. The semantics of an SBGN diagram does not depend on the relative position of the symbols. Furthermore, it does not depend on colors, patterns, shades, shapes and thickness of edges (**Fig. 2**). Similarly, the labels of symbols are not regulated and are only required to be unique within a map.

Finally, it was clear at the outset that it would be impossible to design a perfect and complete language from the beginning. Apart from the prescience this would require, it also would likely require a vast language that most newcomers would shun as being too complex. Thus, the SBGN community decided to stratify language development into levels. A level in SBGN represents a usable set of functionalities that the user community agrees is sufficient for a reasonable set of tasks and goals. Capabilities and features that cannot be agreed upon and are judged insufficiently critical to require inclusion in a given level are postponed to a higher level. In this way, SBGN development is envisioned to proceed in stages, with each higher SBGN level adding richness compared to the levels below it, while maintaining compatibility whenever possible. Furthermor only the actual usage of SBGN languages will tell us how well they work for the diverse communities involved, and this experience will certainly shape the evolution of the notation.

The three languages of SBGN

Molecular entities possess many properties that affect their interactions with other entities. Attempting to represent all the possible reactions and interactions in the same diagram is often futile, usually resulting in an incomprehensible jumble. The different styles of notations described above were attempts to control this complexity by presenting only what was needed in a specific context, or what was available through specific views of the system[14]. Each view focuses on only a portion of the semantics of the overall system, trading off diagram comprehensibility against completeness of biological knowledge.

SBGN follows this strategy and defines three orthogonal and complementary types of diagrams that can be seen as three alternative projections of the underlying more complex biological information. The process diagram draws its inspiration from process-style notations, borrowing ideas from the work of CellDesigner[28] and EPE[22]. By contrast, the entity relationship diagram is based to a large extent on Kohn's MIM notation[18,19]. The SBGN activity flow diagram depicts only the cascade of activity, thus making the notation similar to the reduced representations often used in the current literature to describe signaling pathways and gene regulatory networks. In **Figure 2**, we illustrate the three views applied to a very simple example. The characteristics of the SBGN languages are summarized in **Table 2**.

The idea of having three diagram types naturally begs the question of whether they could be merged into one, at least in paper form. The answer is no, for at least two reasons. First, a single diagram type would bring us back to the problem of dealing with unreasonable numbers of interactions as described above. Second, each SBGN language reflects fundamental differences in the underlying formal description of the phenomena. The meanings are so different that merging diagram types would compromise their representational robustness.

Having multiple visual languages is not uncommon in engineering (consider, for example, block diagrams and circuit diagrams in electronics, UML class, state sequence and deployment diagrams in software engineering), and this supports the idea that having three sublanguages in SBGN will be manageable in practice. In SBGN, the sharing of symbols representing identical concepts further reduces the differences between the three languages to differences in syntax and semantics. We believe that this, combined with careful design, will mitigate some of the difficulties of learning SBGN. However, it is to be noted that the clean orthogonality of the languages makes their overlap very limited, mostly to modulatory arcs, and node decorations.

Box 1 Glossary

SBGN diagrams are a specific set of graphs and thus make use of concepts from graph theory. The following list defines the terms used most often. We are aware of the unavoidable circularity of such definitions.

- **Arc**. A directed edge, that is, an edge that is not symmetrical in shape.
- **Edge**. A line joining two nodes.
- **Glyph**. A symbol that conveys information nonverbally.
- **Graph**. A set of nodes connected with edges.
- **Node**. A point that terminates a line or curve or comprises the intersection of two or more lines or curves.

SBGN process diagram

A process diagram represents all the molecular processes and interactions taking place between biochemical entities, and their results. This type of diagram depicts how entities transition from one form to another as a result of different influences; thus, it portrays the temporal qualities of molecular events occurring in biochemical reactions. In this way, the approach underlying process diagrams is the same as in the familiar textbook drawings of metabolic pathways. The main drawback of process diagrams is that a given entity must appear multiple times in the same diagram if it exists under several states; therefore, the notation is sensitive to the combinatorial explosion of possible entities and reactions, as is often the case in signaling pathways.

The SBGN process diagram level 1 specification defines six major classes of glyphs: entity pool nodes, process nodes, container nodes, reference nodes, connecting arcs and logical operators (**Supplementary Note 1**). In **Figure 3a**, we show a complete example of an SBGN process diagram. The number of symbols in level 1 of the SBGN process diagram notation has been purposefully limited so that they could be easily memorized. The notation may be enriched (perhaps using subclasses of symbols) in higher levels of SBGN.

Table 3 lists software projects that are already developing support for SBGN process diagram level 1 (see also **Supplementary Note 2**). Some of these rely on manual design of the pathways, whereas others, such as Arcadia, automatically generate SBGN PD from SBML models that have been annotated with terms from the Systems Biology Ontology[29]. The encoding of SBGN diagrams using computer-readable formats, a crucial step toward exchange and reuse of SBGN

Table 1 Features of *ad hoc* graphical notations, and the problems they create

Feature	Problem(s)
Different line thicknesses distinguish different types of processes or elements Dotted or dashed line styles distinguish different types of processes or elements	1. Rescaling a diagram can make line thicknesses and styles impossible to discern 2. Photocopying or faxing a diagram can cause differences in line thicknesses and styles to disappear 3. Differences in line thickness and style are difficult to make consistent in diagrams drawn by hand
Different colors distinguish different types of processes or elements	1. Photocopying or faxing a diagram will cause color differences to be indistinguishable 2. Color characteristics are difficult to achieve and keep consistent when drawing diagrams by hand
Identical line terminators (e.g., a single arrow) indicate different effects or processes depending on context	1. Greater ambiguity is introduced into a diagram 2. Interpreting a diagram requires more thought on the part of the reader 3. Automated verification of diagrams is more difficult due to lack of distinction between different processes or elements
Ad hoc symbols introduced at will by author	Interpreting a diagram requires the reader to search for additional information explaining the meaning of the symbols

diagrams, is currently supported in different formats such as SBML, GML and GraphML by different tools, and a general XML-based exchange format for SBGN is currently under discussion.

SBGN entity relationship diagram

The SBGN notation for entity relationships puts the emphasis on the influences that entities have upon each other's transformations rather than the transformations themselves. One can imagine that each of the relationships represents a specific conclusion of a scientific experiment or article. Their addition on a map represents the knowledge we have of the effects the entities have upon each other. Contrary to the process diagrams, where the different processes affect each other, the relationships are independent, and this independence is the key to avoiding the combinatorial explosion inherent to process diagrams. Unlike in process diagrams, a given entity may appear only once. Readers can better grasp at first sight all the possible influences and interactions affecting an entity, without having to explore the whole diagram to discover the different states an entity may be in, or to trace all the edges to find the relevant process nodes.

The relationship symbols in entity relationship diagrams support the representation of interactions and state variable assignments, thus allowing the notation to describe certain processes that cannot be expressed in process diagrams, such as allosteric modulation. In process diagrams, one can represent the formation of a ligand-receptor complex, but it is not possible to state that the complex is more active than the receptor alone without additional markup; entity relationships

support this by allowing the interaction with the ligand to modulate the assignment of the variable representing the activity. The trade-off is that the temporal course is difficult to follow in entity relationships, because the sequence of events is not explicitly described (**Fig. 2a,b**).

The specification of SBGN entity relationship diagram level 1 defines three major classes of glyphs: entity nodes, statements and influences (logical operators are entity nodes). We summarize the symbols and the rules for their assembly (**Supplementary Note 3**). In **Fig. 3b**, we show a complete example of an SBGN entity relationship diagram.

SBGN activity flow diagram

A strategy often used for coping with biochemical network complexity or with incomplete or indirect knowledge is to selectively ignore the biochemical details of processes, instead representing the influences between entities directly. SBGN's activity flow diagrams permit modulatory arcs to directly link different activities, rather than entities and processes or relationships as described previously. Instead of displaying the details of biochemical reactions with process nodes and connecting arcs, the activity flow diagrams show only influences such as 'stimulation' and 'inhibition' between the activities displayed by the molecular entities (**Fig. 2c**). For example, a signal 'stimulates' the activity of a receptor, and this activity in turn 'stimulates' the activity of an intracellular transducing protein (note that activity flow retains the sequential chains of influences). Because most signaling pathway diagrams in the current literature are essentially activity flow diagrams, we expect many biologists will find this type of diagram familiar.
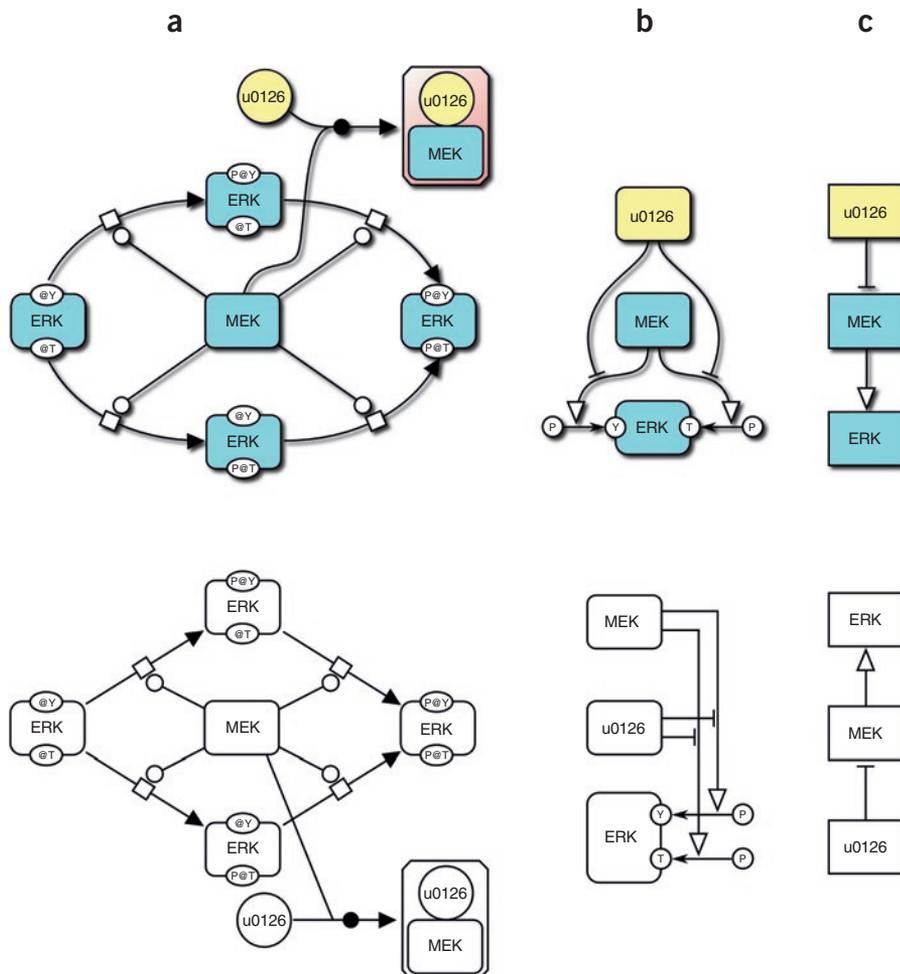
Figure 2 Simple example of protein phosphorylation catalyzed by an enzyme and modulated by an inhibitor. The semantics of an SBGN diagram does not depend on the relative position of the symbols, or on colors, patterns, shades, shapes and thickness of edges. Therefore, the upper and lower diagrams are identical as far as SBGN is concerned, and have to be interpreted exactly the same way. (**a**) Process diagrams, explicitly displaying the four forms of ERK, phosphorylated and nonphosphorylated on the tyrosine and the threonine, as well as the processes of phosphorylation by MEK and the inhibition of MEK by complexation with u0126. Note that the inhibition in this diagram emerges from the sequestration of MEK and is not explicitly represented. The phosphorylation sites are represented by variables, which in this example are labeled simply as 'Y' and 'T' (but in general could be anything desired by the diagram author), shown adorning the main symbols for ERK. (**b**) Entity relationship diagrams, showing ERK and the assignment of its phosphorylations (at the tyrosine and threonine residues), as well as the relationships between those and MEK and u0126. Note that ERK appears only once in this diagram; the different possible states are not explicitly depicted. (**c**) Activity flow diagrams depicting the activation of ERK by MEK and the inhibition of MEK by u0126. In this notation, only the relevant activities of u0126, MEK and ERK are represented, as well as abstract representations of the influences of activities upon each other, whereas the biochemical details are omitted.

Table 2 Comparison between the three languages of SBGN

	Process diagram	Entity relationship diagram	Activity flow diagram
Purpose	Represent processes that convert physical entities into other entities, change their states or change their location	Represent the interactions between entities and the rules that control them	Represent the influence of biological activities on each other
Building block	Different states of physical entities are represented separately	Physical entities are represented only once	Different activities of physical entities are represented separately
Ambiguity	Unambiguous transcription into biochemical events	Unambiguous transcription into biochemical events	Ambiguous interpretation in biochemical terms
Level of description	Mechanistic descriptions of processes	Mechanistic description of relationships	Conceptual description of influences
Temporality	Representation of sequential events	Absence of sequentiality between events	Representation of sequential influences
Pitfalls	Sensitive to combinatorial explosion of states and processes	Creation, destruction and translocation are not easily represented	Not suitable to represent association, dissociation, multistate entities
Advantages	The best for representing temporal/mechanistic aspects of processes such as metabolism	The best for representing signaling involving multistate entities	The best for functional genomics and signaling with simple activities

By ignoring processes and entity states, the number of nodes in an activity flow diagram is greatly reduced compared to an equivalent process diagram (**Fig. 2a,c**). Activity flow diagrams are also especially convenient for representing the effects of perturbations, whether genetic or environmental, because the complete mechanisms of the perturbations may not be known, or are irrelevant to the goals of a given study. The drawback is that activity flow diagrams may contain a high level of ambiguity. For instance, the biochemical basis of a positive or negative influence in a given system is left undefined. For this reason, this type of SBGN diagram should not exist alone; it should be associated, when possible, with detailed entity relationship and process diagrams, and used only for viewing purposes. We expect it will often be possible to generate activity flow diagrams mechanically from process diagrams and entity relationships, and have already performed preliminary work in that direction.

The SBGN activity flow diagram level 1 specification defines four major classes of glyphs: activity nodes, container nodes, modulating arcs and logical operators (**Supplementary Note 4**). **Figure 3c** shows a complete example of an SBGN activity flow diagram.

Participation and future prospects

The SBGN website (http://sbgn.org/) is a portal for all things related to SBGN. Interested persons can get involved in SBGN discussions by joining the SBGN discussion list (sbgn-discuss@sbgn.org). Face-to-face meetings of the SBGN community, generally held as satellite workshops of larger conferences, are announced on the website as well as the mailing list.

Standardizing a notation for depicting networks of biochemical interactions has so far remained an elusive goal, despite numerous but isolated efforts in that direction. Only with such a standardized notation will biologists, modelers and computer scientists be able to exchange accurate descriptions of complex systems—a task that continues to grow more demanding as our collective knowledge expands. SBGN blends many influences from past efforts, and also introduces many new ideas designed to overcome limitations of other notations.

Using a community-based approach involving many interested groups and individuals (including some who have been involved in previous efforts), we have developed and released the first version of the three languages of the SBGN, the process diagram, the entity relationship diagram and the activity flow diagram.

Future levels of the three languages should address major challenges currently faced by the systems biology community, as the field matures and diversifies. To cite but a few examples, the representation of spatial structures and spatial events, of composed and modular models, and of dynamic creation or destruction of compartments remains unchartered territory.

Table 3 List of software systems known to provide support, or to be in the process of developing support, for SBGN

Name	Organization	Link
Arcadia	Manchester Centre for Integrative Systems Biology, Manchester, UK	http://arcadiapathways.sourceforge.net/
Athena	University of Washington, Seattle, WA, USA	http://www.codeplex.com/athena/
BioModels Database	European Bioinformatics Institute, Cambridge, UK	http://www.ebi.ac.uk/biomodels/
BioUML	Institute of Systems Biology, Novosibirsk, Russia	http://www.biouml.org/
ByoDyn	Institut Municipal d'Investigació Mèdica, Barcelona, Spain	http://byodyn.imim.es/
CellDesigner	The Systems Biology Institute, Tokyo	http://www.celldesigner.org/
Dunnart	Monash University, Melbourne, Australia	http://www.csse.monash.edu.au/~mwybrow/dunnart/
Edinburgh Pathway Editor	Edinburgh Centre for Bioinformatics, Edinburgh, UK	http://www.pathwayeditor.org/
JWS Online	Stellenbosch University, Stellenbosch, South Africa	http://jjj.biochem.sun.ac.za/
NetBuilder	STRI, University of Hertfordshire, Hatfield, UK	http://strc.herts.ac.uk/bio/maria/Apostrophe/
PANTHER	Artificial Intelligence Center, SRI international, Menlo Park, CA, USA	http://www.pantherdb.org/pathway/
Reactome	European Bioinformatics Institute, Cambridge, UK	http://www.reactome.org/
Vanted	IPK Gatersleben, Gatersleben, Germany	http://vanted.ipk-gatersleben.de/
VISIOweb	Bilkent University, Ankara, Turkey	http://www.bilkent.edu.tr/~bcbi/pvs.html

Figure 3 Example of complete SBGN diagrams. (**a**) Process diagram representing the synthesis of the neurotransmitter acetylcholine in the synaptic button of a nerve terminal, its release in the synaptic cleft, degradation in the synaptic cleft, the post-synaptic stimulation of its receptors and the subsequent effect on muscle contraction. Colors are used to enhance the biological semantics, blue representing catalytic reactions, orange for transport between compartments (including unrepresented ions, through channels) and green for the function of contractile proteins. However, it is important to note that those colors are not part of SBGN process diagram notation, and must not change the interpretation of the graph. (**b**) SBGN entity relationship diagram representing the transduction, by calcium/calmodulin kinase II, of the effect of voltage-induced increase of intracellular calcium onto the long-term potentiation (LTP) of the neuronal synapses, triggered by a translocation of glutamate receptors. The diagram describes the various relationships between the phosphorylations of the kinase monomers and their conformation. Colors highlight the direction of the relationships relative to the phenotype; blue relationships enhance LTP whereas red ones preclude this enhancement. (**c**) SBGN activity flow diagram representing the cascade of signals triggered by the epidermal growth factor, and going from the plasma membrane to the nucleus. The diagram is derived from reference 30.

AUTHOR CONTRIBUTIONS
N.L.N., M.H., H.M., S.M., F.S. and A.S. contributed equally to the redaction of SBGN specifications.

Published online at http://www.nature.com/naturebiotechnology/.

Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Larkin, J.H. & Simon, H.A. Why a diagram is (sometimes) worth ten thousands words. *Cogn. Sci.* **11**, 65–100 (1987).
2. Lazebnik, Y. Can a biologist fix a radio?—Or,what I learned while studying apoptosis. *Cancer Cell* **2**, 179–182 (2002).
3. Gortner, R.A. *Outlines of Biochemistry*. (Wiley, New York, 1949).
4. Dagley, S. & Nicholson, D.E. *An Introduction to Metabolic Pathways* (Wiley, New York, 1970).
5. Michal, G. Biochemical Pathways (wall chart). (Boehringer Mannheim, Mannheim, Germany, 1984)
6. Sel'kov, E.E., Goryanin, I.I., Kaimatchnikov, N.P., Shevelev, E.L. & Yunus, I.A. Factographic data bank on enzymes and metabolic pathways. *Studia Biophysica* **129**, 155–164 (1989).
7. Karp, P.D. & Paley, S.M. Representations of metabolic knowledge: pathways. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 203–211 (1994).
8. Goto, S. *et al.* Organizing and computing metabolic pathway data in terms of binary relations. *Pac. Symp. Biocomput.* **PSB97**, 175–186 (1997).
9. Brown, C.T. *et al.* New computational approaches for analysis of cis-regulatory networks. *Dev. Biol.* **246**, 86–102 (2002).
10. Demir, E. *et al.* PATIKA: an integrated visual environment for collaborative construction and analysis of cellular pathways. *Bioinformatics* **18**, 996–1003 (2002).
11. Sauro, H.M. *et al.* Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration. *OMICS* **7**, 355–372 (2003).
12. Funahashi, A., Morohashi, M., Kitano, H. & Tanimura, N. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *BIOSILICO* **1**, 159–162 (2003).
13. Kohn, K.W. Functional capabilities of molecular network components controlling the mammalian G1/S cell cycle phase transition. *Oncogene* **16**, 1065–1075 (1998).
14. Kitano, H. A graphical notation for biochemical networks. *BIOSILICO* **1**, 169–176 (2003).
15. Pirson, I. *et al.* The visual display of regulatory information and networks. *Trends Cell Biol.* **10**, 404–408 (2000).
16. Cook, D.L., Farley, J.F. & Tapscott, S.J. A basis for a visual language for describing, archiving and analyzing functional models of complex biological systems. *Genome Biol.* **2**, research0012.1–0012.10 (2001).
17. Longabaugh, W.J.R., Davidson, E.H. & Bolouri, H. Visualization, documentation, analysis, and communication of large-scale gene regulatory networks. *Biochim. Biophys. Acta* **1789**, 363–374 (2009).
18. Kohn, K.W. Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Mol. Biol. Cell* **10**, 2703–2734 (1999).
19. Kohn, K.W., Aladjem, M.I., Weinstein, J.N. & Pommier, Y. Molecular interaction maps of bioregulatory networks: a general rubric for systems biology. *Mol. Biol. Cell* **17**, 1–13 (2006).
20. Demir, E. *et al.* An ontology for collaborative construction and analysis of cellular pathways. *Bioinformatics* **20**, 349–356 (2004).
21. Kitano, H., Funahashi, A., Matsuoka, Y. & Oda, K. Using process diagrams for the graphical representation of biological networks. *Nat. Biotechnol.* **23**, 961–966 (2005).
22. Moodie, S.L., Sorokin, A.A., Goryanin, I.I. & Ghazal, P. A graphical notation to describe the logical interactions of biological pathways. *J. Integr. Bioinform.* [Au: Correct publication is only one page?] **3**, 36–46 (2006).
23. Raza, S. A logic-based diagram of signalling pathways central to macrophage activation. *BMC Syst. Biol.* **2**, 36 (2008).
24. Hucka, M. *et al.* The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531 (2003).
25. Britton, C., Jones, S., Kutar, M., Loomes, M. & Robinson, B. Evaluating the intelligibility of diagrammatic languages used in the specification of software. in *Theory and Application of Diagrams: Diagrams 2000* (eds. Anderson, M., Cheng, P. & Haarslev, V.) 376–391 (Springer, New York, 2000).
26. Narayanan, N.H. & Hübscher, R. Visual language theory: towards a human-computer interaction perspective. in *Visual Language Theory* (eds. Marriott, K. & Meyer, B.) 85–127 (Springer, New York, 1998).
27. Saraiya, P., North, C. & Duca, K. Visualizing biological pathways: requirements analysis, systems evaluation and research agenda. *Inform. Visual* **4**, 191–205 (2005).
28. Funahashi, A. *et al.* CellDesigner 3.5: a versatile modeling tool for biochemical networks. *Proc. IEEE* **96**, 1254–1265 (2008).
29. Le Novère, N., Courtot, M. & Laibe, C. Adding semantics in kinetics models of biochemical pathways. in *Proc. 2nd Intl. Symp. Exp. Stand. Cond. Enzyme Characterizations*, Rüdesheim, Germany, March 19–23, 2006 (Beilstein Institute, Frankfurt, 2007). <http://www.beilstein-institut.de/index.php?id=196>
30. Anonymous. MAP kinases. Gene set bank. *Riken BioResource Center DNA Bank* <http://www.brc.riken.go.jp/lab/dna/en/GENESETBANK/index.html> (August 19, 2008).
31. Riechelmann, H. Cellular and molecular mechanisms in environmental and occupational inhalation toxicology. *GMS Cur. Topics. Otorhinolaryngol.—Head Neck Surg.* **3**, Doc02 (2004).
32. Schlessinger, J. Epidermal growth factor receptor pathway. *Sci. Signal.* (connections map in the database of cell signaling, as seen May 29, 2009). <http://stke.sciencemag.org/cgi/cm/stkecm;CMP_14987>
33. Anonymous. MAPK signaling pathway, *Homo sapiens. KEGG Pathway hsa04010* <http://www.genome.jp/kegg/pathway/hsa/hsa04010.html> (July 15, 2009).
34. Kholodenko, B. Negative feedback and ultrasensitivity can bring about oscillations in the mitogen-activated protein kinase cascades. *Eur. J. Biochem.* **267**, 1583–1588 (2000).

[1]*EMBL European Bioinformatics Institute, Hinxton, UK.* [2]*Engineering and Applied Science, California Institute of Technology, Pasadena, California, USA.* [3]*SRI International, Menlo Park, California, USA.* [4]*Centre for Systems Biology at Edinburgh, University of Edinburgh, Edinburgh, UK.* [5]*Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany.* [6]*Institute of Computer Science, University of Halle, Halle, Germany.* [7]*School of Informatics, University of Edinburgh, Edinburgh, UK.* [8]*Memorial Sloan Kettering Cancer Center - Computational Biology Center, New York, NY, USA.* [9]*Science and Technology Research Institute, University of Hertfordshire, Hatfield, UK.* [10]*National Cancer Institute, Bethesda, Maryland, USA.* [11]*Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand.* [12]*Department of Bioengineering, University of Washington, Seattle, Washington, USA.* [13]*BIOQUANT, University of Heidelberg, Heidelberg, Germany.* [14]*Division of Pathway Medicine, University of Edinburgh Medical School, Edinburgh, UK.* [15]*Riken OMICS Science Center, Yokohama City, Kanagawa, Japan.* [16]*The Systems Biology Institute, Tokyo, Japan.* [17]*School of Computer Science, University of Manchester, Manchester, UK.* [18]*Manchester Interdisciplinary Biocentre, Manchester, UK.* [19]*Clayton School of Information Technology, Faculty of Information Technology, Monash University, Melbourne, Victoria, Australia.* [20]*U900 INSERM, Paris Mines Tech, Institut Curie, Paris, France.* [21]*Terry Fox Laboratory, British Columbia Cancer Research Center, Vancouver, British Columbia, Canada.* [22]*Bilkent Center for Bioinformatics, Bilkent University, Ankara, Turkey.* [23]*The Roslin Institute, University of Edinburgh, Midlothian, UK.* [24]*Department of Biosciences and Informatics, Keio University, Hiyoshi, Kouhoku-ku, Yokohama, Japan.* [25]*Institute of Systems Biology, Novosibirsk, Russia.* [26]*Design Technological Institute of Digital Techniques SB RAS, Novosibirsk, Russia.* [27]*Ontario Institute for Cancer Research, Toronto, Ontario, Canada.* [28]*School of Chemistry, University of Manchester, Manchester, UK.* [29]*Department of Biochemistry, Stellenbosch University, Matieland, South Africa.* [30]*Sony Computer Science Laboratories, Tokyo, Japan.* [31]*Okinawa Institute of Science and Technology, Okinawa, Japan. Correspondence should be addressed to N.L.N. (lenov@ebi.ac.uk).*