BRIEFINGS IN BIOINFORMATICS. VOL 7. NO 4. 331–338 doi:10.1093/bib/bbl039

Representing, storing and accessing molecular interaction data: a review of models and tools

Lena Strömbäck, Vaida Jakoniene, He Tan and Patrick Lambrix

Submitted: 30th May 2006; Received (in revised form): 18th September 2006

Abstract

One important aim within systems biology is to integrate disparate pieces of information, leading to discovery of higher-level knowledge about important functionality within living organisms. This makes standards for representation of data and technology for exchange and integration of data important key points for development within the area. In this article, we focus on the recent developments within the field. We compare the recent updates to the three standard representations for exchange of data SBML, PSI MI and BioPAX. In addition, we give an overview of available tools for these three standards and a discussion on how these developments support possibilities for data exchange and integration.

Keywords: *molecular interactions; cellular pathways; standardization; SBML; PSI MI; BioPAX*

INTRODUCTION

Standards and standardization of information within systems biology are currently active research fields and new standards are rapidly being developed. The main reasons for this are a dramatic increase of the amount of experimental results and the desire of researchers to exchange and integrate results to gain a better understanding of interactions between different substances in living organisms. A better understanding of these complex interaction networks is one of the main goals for genomics and proteomics [1, 2]. As a means to reaching these goals the articles mention the development of reusable software modules, new ontologies and improved technologies for database and knowledge management, which today are active research fields within the area. The focus of this article is to review these developments.

In a previous paper [3], we evaluated the three standards SBML [4], PSI MI [2] and BioPAX [5] regarding their underlying models, content and support for creation of tools. In this article, we update our earlier evaluation to reflect recent changes and follow the further developments of the standards, and discuss how the above goals have been met. Furthermore, we extend our previous work by putting more emphasis on tools for creation, analysis and management of data represented in these standards. There are many other related standards in the area. These will not be further discussed here but are instead described in a companion paper [6].

The article consists of three parts. In the first part, we describe the new features of SBML [4], PSI MI [2] and BioPAX [5], and how these can be used to

Corresponding author. Lena Strömbäck, Department of Computer and Information Science, Linköpings Universitet, Sweden. E-mail: lestr@ida.liu.se

Lena Strömbäck is an Assistant Professor at Linköpings Universitet, Sweden. She holds a PhD degree in computer science from Linköpings Universitet. Her background includes working with standards and standardization at Nokia. Her current research interests are in XML standards, especially within the area of life sciences.

Vaida Jakoniene is a PhD student at Linköpings Universitet, Sweden. She holds an MSc degree in informatics from Kaunas University of Technology, Lithuania. Her research interests cover issues related to the integration of biological data sources.

He Tan is a PhD student at Linköpings Universitet, Sweden. She holds an MSc degree in computer science from Linköpings universitet. Her current interests include ontologies and Semantic Web for the life sciences.

Patrick Lambrix is an Associate Professor at Linköpings Universitet, Sweden. He holds a PhD degree in computer science from Linköpings Universitet and MSc degrees in mathematics and computer science from KU Leuven, Belgium. His current main research interests are in Semantic Web, ontologies and knowledge representation for the life sciences.

meet future needs. The second part of the article gives an overview of tools enabling analysis and manipulation of data represented in the standards. In the last part, we focus on general tools for knowledge management and integration and discuss how developments in the field have enabled greater support for integration. We conclude the article with a short summary and some pointers to future directions.

SBML, PSI MI AND BIOPAX

This section presents the recent developments in each of the three different standards, SBML, PSI MI and BioPAX. Table 1 updates the table from [3] with recent additions in the standards and gives a summary of the main features for each of the formats.

For *Systems Biology Markup Language* (SBML) (www.sbml.org) [4], the main aim is to represent several kinds of pathways, biochemical reactions and

Table I: Main features of SBML, PSI MI and BioPAX, updated from [3]

	SBML version 2.2	**PSI MI version 2.5**	**BioPAX version 2**
Environment for the specification: Developers Existing tools Used by	Systems Biology Workbench Development group. Tools for validation, visualization and conversion. Used by around 90 systems, including simulation environments and databases.	Proteomics Standards Initiative. Tools for viewing and analysis. Data sets available from IntAct, DIP and MINT. More databases, for instance BIND and HPRD, accept data in PSI MI.	BioPAX working group. The implementation in OWL can use tools supporting OWL, such as, Protégé. Collaboration with BioCYC, BIND and WIT. Data sets available from Reactome, BioCyc and PathCase.
Representation of interactors: Used notation Description of parts of interactor	Species, can be grouped into types by the user. No current representation of parts of molecules but a proposal exists for next release of SBML.	Interactor, can be further specified by external ontology. Protein, DNA and RNA sequences can be described as strings.	PhysEnity, with subclasses: complex, small molecule, DNA, RNA and protein. Description of parts dependent on type. Sequence description adopted from PSI MI.
Representation of interaction: Used notation Role of interactor Number of interactors	Reaction, can be further specified by external ontology. Each reaction allows interactors of three predefined roles (reactants products or modifiers). Unbounded number of interactors for each role.	Interaction, can be further specified by external ontology. Each interactor can be given both an experimental and biological role, specified by external ontology. Unbounded number of interactors.	Interaction, with subclasses control and conversion, each of these subclasses have several other subclasses. Roles and number of interactors dependent on subtype.
Representation of pathway	There is no specific construct to describe pathways, but each SBML model is supposed to represent a pathway of selected interactions.	PSI MI is not intended to describe pathways.	Pathways can be described by putting together interactions Complex and flexible structure of pathways allowed.
Other predefined entities: Environment for interaction Experimental data Mathematical relations	Compartment defined as the environment for interactions. No data about experiments. Mathematical relations for reactions.	It is possible to refer to compartments for interactions. Data about experiments verifying the interaction. No mathematical relations.	No environment for interactions. Experimental evidence can be described. No mathematical relations, but chemical details around interactions.
Expressivness: Main structure Inheritance Definition of new attributes and entities	All entities are defined separately. References between them indicate the structure of interactions. A hierarchy between the predefined entities but no possibility for the user to define types. The note and annotation fields can be used for extra information.	Entities can be defined separately, but it is also possible to structure information around interactions. No inheritance. A specific list of attributes can be used to add information that does not fit into the format.	Entities are defined in an inheritance hierarchy. Predefined inheritance hierarchy of subclasses for all entities. The user is expected to use the concepts defined by the ontology.
Referencing to publications and databases	An addition from December 2005 defines a format for external references.	Links to publications, vocabularies and other databases.	Links to publications and other sources.

gene regulation. The main concepts in SBML are the interacting substances (*Species*), how these substances interact (*Reaction*) and where the reaction takes place (*Compartment*). In addition, the user can specify mathematical properties describing the reaction's behavior, sizes of compartments, concentrations of substances and similar information. The new proposal for version 2 of level 2 [7] contains several interesting new features. One is a framework for linking SBML descriptions to complementary information about the objects in available databanks. Another interesting addition is the ability to place restrictions on the type of objects. For some of the main concepts, such as *Reactions* and participants in the reactions, the user can refer to controlled vocabularies, thereby providing a more detailed specification of the concept. These vocabularies are provided as a controlled vocabulary, included in Open Biomedical Ontologies (OBO) [8]. For *Species* and *Compartments*, there is another solution. Here the user can group concepts in a way that is useful for him by a type specification that is specific to each model. In the proposal for level 3 [9], future versions of the standard will enable the encoding of protein states from protein structure.

For *Proteomics Standards Initiative Molecular Interaction* format (PSI MI) (http://psidev.sourceforge.net/) [2] the main aim is to provide a mean for representation and exchange of experimental data. The main objects in PSI MI are *Interactors* (SBML substances), *Interactions* (SBML reactions) and *Experiments*. In addition, information about the type of experiment, methods for detecting a substance, statistical evidence for an interaction and the participating *Interactors* can be stored. The new version 2.5 of PSI MI was released in December 2005 and provides a means for a more fine-grained representation of *Interactor* and *Interactions*. It also allows representation of both the biological and the experimental role of a participant in a detected interaction. There are also additions to *Interaction* allowing the user to represent deduced interactions and experiments made on species other than the one the interaction is reported from. The new version also contains a generalization of types and naming allowing representation of interactions between substances other than proteins. Another major addition is the use of controlled vocabularies, providing means of referring to *Interactortypes*, *Interactiontypes*, *Experimenttypes* and different kinds of experimental methods in a consistent way. As with SBML, these vocabularies are part of OBO [8].

The aim of the *BioPAX Data Exchange* (www.biopax.org) standard is to define a unified framework for sharing pathway information. It uses Web Ontology Language (OWL) as the representation format. In BioPAX, information is centered on substances, called *Physical Entities* and *Interactions*. For each of these main concepts, a number of subclasses are defined specifying many types of substances, such as proteins and DNA, together with different kinds of interactions. BioPAX also includes an interesting means for the user to combine single *Interactions* into *Pathways* in various ways. Version 2 of BioPAX was finalized in December 2005. The focus has been to extend the standard to represent metabolic pathways and molecular interactions. Among the major extensions is the ability to represent molecular binding interactions. Another interesting addition is an import of the PSI MI features for representation of sequences for *proteins*, *DNA* and *RNA*. BioPAX is adapted toward experimental data by the ability to represent information about experimental evidence of an interaction. Finally, the concept of pathways has been extended to allow hierarchical pathways.

To conclude this section, we can see that many of the features that were listed in [3] as important for a standard in this field are accounted for in the new versions of the standards. *Identification of entities between data sources* is well supported in PSI MI and BioPAX, and with version 2.2, SBML has also included this feature. Furthermore, PSI MI and BioPAX have further extended the *representation of protein structure* by an enhanced possibility to refer to subsequences. SBML on the other hand, has a proposal on how to include this in future versions. Also, the *granularity of reactions and roles of the reactants* have been further increased by the addition of links to external ontologies by PSI MI and SBML. In PSI MI's case both experimental and biological roles can be defined for each participant. For *representation of pathways*, BioPAX has added new mechanisms for structuring pathways in their latest release. The remaining features listed in [3], *pathway presentation, information for reasoning* and *user-defined entities and attributes* have remained unchanged, with SBML being the format most suitable for simulations while SBML and PSI MI allow the user to add information not specified by the standard.

TOOLS FOR CREATION AND ANALYSIS OF STANDARDIZED DATA

The importance of the area and worldwide interest in the field have resulted in the development of a large number of tools for working with molecular interactions and interaction networks. The main tasks supported by the tools are computational model building and model analysis. With respect to the purpose of modeling, the tools can be divided into two categories: structural modeling and dynamic modeling. The former group focuses on building network diagrams and analyzing their structure. The latter group builds kinetic models of the networks and enables analysis of their behavior. In addition, there is a large group of tools that provide utilities enabling faster development of new tools and support for work with the existing tools. The functions provided by the tools vary a lot. Some specialize in solving a single task while others provide general-purpose environments supporting users in solving multiple tasks. It is a great advantage that several of these tools are open-source applications or provide Application Programming Interfaces (APIs) that allow extension and adaptations of the tools to specific needs. Some of the tools have been developed as web applications enabling task performance online.

Already, a large number of tools recognize data represented in one or several of the discussed standards, i.e. the tools support import and export of data formatted according to these standards. Some tools use one of the standards as a local model, while other tools use a different model as their main data structure. When some information is not relevant to a task, some tools may import only parts of a standard. The tools may also extend the standards using specialized data types, e.g. diagram layout information. Currently, SBML is much better supported by the tools than PSI MI and BioPAX. The SBML web page enumerates over 90 software systems compatible with the standard. At the same time, only a few modeling tools recognize data in the PSI MI and BioPAX formats. For all the standards, there are some tools for structural modeling, while only SBML is supported by tools for modeling of network behavior. References to most of the relevant tools can be found at the web pages of each standard (Tools for BioPAX are listed in the BioPAX Wiki web pages.). As the research community is very active, we expect that the number of tools compatible with the standards will grow.

Below, we provide an overview of some of the characteristic features provided by the existing tools. The discussion is organized according to the types of tasks supported and references are provided to some of the relevant tools. The list of tools is not meant to be exhaustive.

Model building

This task covers the creation and editing of computational models. Based on the type of data stored, there are two types of models. Qualitative models describe the structure of reaction networks that can be specified as network diagrams via Graphical User Interface (GUI)-based tools, or can be automatically generated, for instance, from a set of given interactions, e.g. Cytoscape [10] (BioPAX, PSI MI, SBML). Quantitative models typically support SBML and describe the behavior of a model that is expressed by reactions, together with state variables and reaction rate laws. For some tools, quantitative models can be specified as text-based script files, e.g. Jarnac [11], while other tools provide GUI-based interfaces that can either support the specification of data describing model behavior, e.g. Gepasi [12], or support modeling of network structure and behavior, e.g. CellDesigner [13] and JDesigner [14]. The adapted graphical notation differs among the tools, and the type of data expressible in the diagrams may vary. For instance, NetBuilder [15] bases the notation on electronic circuit design principles to represent genetic regulatory networks, while CellDesigner uses process diagrams to model gene-regulatory and biochemical networks. Some tools allow the representation of additional information, e.g. descriptions of experiments or references to other sources.

Model analysis

To gain insights about biological systems, tools have been developed to support the analysis of the structure and the analysis of the behavior of models. *Analysis of structure* is enabled by the tools supporting visual exploration of network structure, search capabilities in the networks and different statistical and analytical algorithms to extract topological properties of the network. Tools for visualizing interaction networks range from those that represent models as still pictures to highly flexible interactive environments, e.g. Cytoscape, PATIKA [16] (BioPAX, SBML), PIMWalker [17] (PSI-MI), ProViz [18] (PSI MI) and VisANT [19] (BioPAX, PSI MI). The tools may provide the ability to

represent models at different levels of detail or to hide or expand certain parts of the network. Also, different representation layouts may be available. Some of the tools extend models by integrating information from other sources, e.g. molecular descriptions, expression profiles and ontological annotations (Cytoscape, VisANT). Visualization tools also differ in the size of the supported network and how efficiently it can be processed. Types of searches supported by the tools may range from simple searches to more complex searches, for instance, to find interactions in which a molecule is involved, to find paths among given molecules or to detect common targets or regulators for a given group of proteins. Broader types of searches may be supported when the models are integrated with annotations. While performing statistical and analytical analysis, tools may explore the degree of the nodes in the network or may compare networks of different organisms (Cytoscape, VizANT). To perform *analysis of behavior*, a large variety of simulation tools have been developed. Examples of the supported analyses are: time course simulation, steady-state analysis and metabolic control analysis. Gepasi, Jarnac and PySCeS [20] are examples of simulators solving ordinary differential equations. Virtual Cell [21] is an example of a system enabling exploration of system dynamics in space and time by solving partial differential equations. JWS Online [22] and WebCell [23] are examples of tools supporting simulation online.

Utilities

Various tools in the form of independent applications, libraries and plug-ins, are available to facilitate faster development of new tools and to support work with the existing tools. For instance, to help developers to read, write and manipulate data, generic libraries like Jena [24], and specialized libraries like libSBML [25] (SBML) can be used. To guarantee correctness of the specified models with respect to their schema, validation tools are available for BioPAX and SBML. To enable the use of models from different data sources, tools for converting data between different formats are of great importance. For instance, tools exist for converting data from CellML [26] representations to SBML [27]. In the presence of the variety of tools developed in different programming languages and running on different platforms, Systems Biology Workbench [11] and BIO-SPICE [28] are important initiatives that develop software frameworks enabling integration of heterogeneous applications into a single environment where the applications can reuse each others' capabilities.

DATA MANAGEMENT AND INTEGRATION

The discussion in the previous sections shows that the recent developments of the three standards as well as the increase in available tools for creating and analyzing models makes these standards very powerful tools for addressing the goals of gaining knowledge within systems biology. However, the tools addressed in the previous section are often specific for one standard and in this section we put focus on reusable software for data management and integration.

For a user who needs to work with larger quantities of data, management and storage capabilities are important. Some of the tools mentioned above include facilities for storage of larger quantities of data. However, in many cases, traditional database technology is a better option. For the XML implemented standards, SBML and PSI MI, there are in principle two options, either translation of data to a traditional relational database, or use of the newer XML-database approach. The latter has the benefit of allowing direct access to the data represented in XML via the query language XQuery. The two storage approaches are further evaluated and compared for SBML and PSI MI in [29, 30], and XML databases are found interesting to use for both standards. The more complex relationships expressed by BioPAX inheritance hierarchy in OWL are harder to capture in the above database approaches and special purpose tools developed for OWL are needed.

The use of XML technology requires knowing the representation of standards in detail. This is a drawback, especially for a user who is dependent on working on data represented in several of the above standards. An illustrative example is given in Table 2. In this table, we give examples of how *Succinate dehydrogenase* and *Succinate dehydrogenase catalysis* can be represented in each of the three formats. To make the examples easier to read and understand, we have simplified the examples focusing on the parts of the descriptions that illustrate our point. For the sake of readability, for BioPAX we have chosen to show the

Table 2: Examples of representations within the three formalisms

BioPAX	PSI MI	SBML
Complex Name: Succinate dehydrogenase Components: Protein . . . Complex . . . Organism:	‹Interactor id = "Succdeh"› ‹names› ‹shortLabel›Succinate dehydrogenase‹/shortLabel› ‹/names› ‹Interactortype› Protein complex ‹/Interactortype› ‹Organism›. . . . .‹/Organism› ‹/Interactor›	‹species name = "Succinate dehydrogenase" compartment = "MM" id = "Succdeh" speciesType = "complex"/›
Catalysis Name: Succinate dehydrogenase catalysis Controller: Succinate dehydrogenase . . . Controlled: Conversion: Name: . . . Left: Succinate Right: Fumarate	‹interaction id="RI "› ‹names› ‹shortLabel›Succinate dehydrogenase catalysis ‹/shortLabel› ‹/names› ‹interactiontype›Enzymatic reaction ‹/interactiontype› ‹participantList› ‹Participant› ‹proteinInteractorRef ref="Succinate"/› ‹/proteinParticipant› ‹Participant› ‹proteinInteractorRef ref="Fumarate"/› ‹Participant› ‹proteinParticipant› ‹proteinInteractorRef ref="Succdeh"/› ‹biologicalRole›enzyme ‹/biologicalRole› ‹/proteinParticipant› ‹/participantList› ‹/interaction›.	‹reaction name ="Succinate dehydrogenase catalysis" id = "RI "› ‹listOfReactants› ‹speciesReference species = "Succinate"/› ‹/listOfReactants› ‹listOfProducts› ‹speciesReference species = "Fumarate"/› ‹/listOfProducts› ‹listOfModifiers› ‹modifierSpeciesReference species = "Succdeh"/ sboTerm = "enzyme"› ‹/listOfModifiers› ‹/reaction›

structure in an indented format, where the header represents a class and indented attributes represent properties. The examples clearly show the differences between the choices of representation in the three formats. In PSI MI, we use the general *Interactor* and *Interaction* and the substructures they provide; the more specific type is only indicated by the attribute *Interactortype*. For BioPAX, we can use the more specific classes, *Complex* and *Catalysis*, for the representation. This gives access to more specific properties, such as *Components, Controller* and *Controlled*, to describe the information we need. For SBML, we note that *SpeciesType* is a reference to an internal type in the current model while *sboTerm* in the modifier refers to an external vocabulary. The user who wants to use XML tools directly for accessing data represented in the standards must know all these differences.

The examples also illustrate the difficulty with and possibilities for data integration. Data integration is an important step toward future research within the area. For integration between different standards, a first step is to identify the identical concepts in each of the standards. This is straightforward for the main concepts, but is not obvious regarding more detailed information such as types of interactions and roles of participants, as illustrated by the examples. This problem is analogous to schema matching for databases, where several approaches exist, e.g. [31].

To enable integration of datasets at the level of data, data entries representing the same real world objects must be found in the datasets. In the context of heterogeneous data sets, this is a difficult task. Recent results [32] show that additional information received through links to other databases or from relevant vocabularies can be used to support the task and to improve the integration results. This shows the importance of the recent improvements presented in the standards where the amount of this kind of information is increased.

However, as links to external ontologies become very important for all standards discussed, we also have the problem of management and integration of multiple ontologies. Many tools that support ontological integration exist today, e.g. [33], most of which propose a semi-automatic approach to the problem. As our previous studies show [32], this task is helpful for integration of data sets, but it can also provide important information for the process of integration of concepts in different standards.

CONCLUSION

The recent developments within the standards SBML, PSI MI and BioPAX add many interesting features to the standards. In particular, we would like to mention the development allowing addition of external links and ontologies, which is very important to support integration of data sets within the area. In addition, the development of tools capable of importing and exporting data in either of the standards makes the use of the standards very promising for the future, and it is an important step in the direction of supplying software modules for understanding the complex interaction of substances within living organisms. However, efficient data integration is still an open issue where further improvements are needed.

All these developments within systems biology are very important for reaching the goal of a complete understanding of the interactions of genes, proteins and other substances. However, to reach this final goal, there is a need for transparent flow of data from experiments to larger models. For the future, that means further developments for ontologies and the ability to link data sets between various sources. In terms of tools, there is a further need for development of format-independent tools as well as tools for integration of data represented in the same or different formats.

Key Points

- Recent developments of the standards SBML, PSI MI and BioPAX extend their capabilities for fine-grained representation of information and emphasize the use of external links and ontologies.
- The developments also include a number of tools allowing modeling and analysis in the standards.
- The ability to use external references and links is important for the integration of data sets. However, within this area there is still a large need for new technology and tools.

Acknowledgments

We acknowledge the financial support of the Center for Industrial Information Technology, the Swedish Research Council, the Swedish national graduate school in computer science and the EU Network of Excellence REWERSE (Sixth Framework Programme project 506779).

References

1. Collins FS, Green ED, Guttmacher AE, *et al*. A vision for the future of genomics research: A blueprint for the genomic era. *Nature* 2003;**422**:835–47.
2. Hermjakob H, Montecchi-Palazzi L, Bader G, *et al*. The HUPO PSI's Molecular Interaction format – a community standard for the representation of protein interaction data. *Nat Biotechnol* 2004;**22**:177–83.
3. Strömbäck L, Lambrix P. Representations of molecular pathways: an evaluation of SBML, PSI MI and BioPAX. *Bioinformatics* 2005;**21**:4401–7.
4. Hucka M, Finney A, Sauro HM, *et al*. The Systems Biology Markup Language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 2003;**19**:524–31.
5. BioPAX working Group, *BioPAX Biological Pathways Exchange Language. Level 2, Version 1.0 Documentation*. 2006. http://www.biopax.org (May 2006, date last accessed).
6. Strömbäck L, Hall D, Lambrix, P. A review of standards for data exchange within systems biology. *Proteomics*, to appear.
7. Finney A, Hucka M. *Systems Biology Markup Language (SBML) Level 2: Structures and facilities for model definitions. Final draft for public review*, 2 April 2006. http://sbml.org/specifications/sbml-level-2/version-2/revision-1/(May 2006, date last accessed).
8. *Open Biomedical Ontologies*. http://obo.sourceforge.net/ (May 2006, date last accessed).
9. Finney A. *Systems Biology Markup Language (SBML) Level 3: Proposal: multi-component species features. Proposal manuscript*. March 2004. http://www.cds.caltech.edu/~afinney/multi-component-species.pdf (May 2006, date last accessed).
10. Shannon P, Markiel A, Ozier O, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;**13**:2498–504.
11. Sauro HM. Jarnac: a system for interactive metabolic analysis. In: Hofmeyr J-HS, Rohwer JM, Snoep JL (eds). *Animating the Cellular Map 9th International BioThermoKinetics Meeting*, Chapter 33, 2000, 221–8.
12. Mendes P.. GEPASI: a software package for modelling the dynamics, steady states and control of biochemical and other systems. *Comput Appl Biosci* 1993;**9**:563–71.
13. Funahashi A, Morohashi M, Kitano H. CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *BIOSILICO* 2003;**1**:159–62.
14. Sauro HM, Hucka M, Finney A, *et al*. Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration. *OMICS A Journal of Integrative Biology* 2003;**7**:355–72.
15. Schilstra MJ, Bolouri H. *Modelling the regulation of gene expression in genetic regulatory networks*. 2002. http://strc.herts.ac.uk/bio/maria/NetBuilder/Theory/NetBuilderModelling.htm (May 2006, date last accessed).

16. Dogrusoz U, Erson EZ, Giral E, *et al*. PATIKAweb: a Web interface for analyzing biological pathways through advanced querying and visualization. *Bioinformatics* 2006; **22**:374–5.
17. Meil A, Durand P, Wojcik J. PIMWalker™: Visualising protein interaction networks using the HUPO PSI molecular interaction format. *Appl Bioinfo* 2005;**4**:137–9.
18. Iragne F, Nikolski M, Mathieu B, *et al*. ProViz: Protein interaction visualization and exploration. *Bioinformatics* 2005;**21**:272–4.
19. Hu Z, Mellor J, Wu J, *et al*. VisANT: data-integrating visual framework for biological networks and modules. *Nucleic Acids Res* 2005;**33**:W352–7.
20. Olivier BG, Rohwer JM, Hofmeyr J-HS. Modelling cellular systems with PySCeS. *Bioinformatics* 2005;**21**:560–1.
21. Slepchenko BM, Schaff J, Macara IG, *et al*. Quantitative cell biology with the virtual cell. *Trends Cell Biol* 2003;**13**:570–6.
22. Olivier BG, Snoep JL.. Web-based kinetic modelling using JWS Online. *Bioinformatics* 2004;**20**:2143–4.
23. Lee D-Y, Yun C, Cho A, *et al*. WebCell: a web-based environment for kinetic modeling and dynamic simulation of cellular networks. *Bioinformatics* 2006;**22**:1150–1.
24. Jena: a semantic web framework for Java. http://jena.sourceforge.net/ (May 2006, date last accessed).
25. *LibSBML: A library for SBML*. http://www.sbml.org/software/libsbml/ (May 2006, date last accessed).
26. Lloyd CM, Halstead MDB, Nielsen PF. CellML: its future, present and past. *Prog Biophys Mol Bio* 2004;**85**:433–50.
27. Schilstra MJ, Li L, Matthews J, *et al*. CellML2SBML: conversion of CellML into SBML. *Bioinformatics* 2006;**22**: 1018–20.
28. Bio-SPICE: an open source framework and software toolset for Systems Biology. https://biospice.org/ (May 2006, date last accessed).
29. Strömbäck L. Possibilities and challenges using XML technology for storage and integration of molecular interactions. In: *Proceedings of the Sixteenth International Workshop on Database and Expert Systems Applications*, pp. 575–579. *22–26 August 2005*, Copenhagen, Denmark.
30. Strömbäck L, Hall D. An evaluation of the use of XML for representation, querying, and analysis of molecular interactions. In: *EDBT 2006 Workshops*, 2006 *LNCS* 4254, Springer Verlag, Munich, Germany.
31. Rahm E, Bernstein PA. A survey of approaches to semantic schema matching. *VLDB J* 2001;**10**:334–50.
32. Jakoniene V, Rundqvist D, Lambrix P. A method for similarity-based grouping of biological data. In: *Proceedings of the 3rd International Workshop on Data Integration in the Life Sciences, 2006*. pp. 136–51. LNBI 4075.
33. Lambrix P, Tan H. SAMBO – A system for aligning and merging biomedical ontologies. *Journal of Web Semantics, special issue on Semantic Web for the Life Sciences* 2006;**4**: 196–206.

WEB RESOURCES FOR THE TOOLS

PATIKA. http://web.patika.org
CellDesigner. http://www.celldesigner.org/features.html
VisANT. http://visant.bu.edu/
ProViz. http://cbi.labri.fr/eng/proviz.htm
WebCell. http://webcell.org
PIMWalker. http://pim.hybrigenics.com/pimriderext/pimwalker/
GEPASI. http://www.gepasi.org/
PySCeS. http://pysces.sourceforge.net/
JWS Online. http://jjj.biochem.sun.ac.za/
Jarnac. http://sbw.kgi.edu/software/jarnac.htm
SBW. http://sbw.kgi.edu/
JDesigner. http://sbw.kgi.edu/software/jdesigner.htm
CellML2SBML. http://sbml.org/software/cellml2sbml/
Cytoscape. http://www.cytoscape.org/
Virtual Cell. http://www.nrcam.uchc.edu/