

# Chapter 12

## Visualising Metabolic Pathways and Networks: Past, Present, Future



Falk Schreiber, Eva Grafahrend-Belau, Oliver Kohlbacher, and Huaiyu Mi

**Abstract** Visualisations of metabolites and metabolic pathways have been used since the early years of research in biology, and pathway maps have become very popular in biochemistry textbooks, on posters, as well as in electronic resources and web pages about metabolism. Visualisations help to present knowledge and support browsing through chemical structures, enzymes, reactions and pathways. In addition, visual and immersive analytics of metabolism connects network analysis algorithms and interactive visualisation methods to investigate structures in the network such as centralities, motifs and paths, or to compare pathways for finding differences between species or conditions. The graphical depiction of networks supports the mapping and investigation of additional data such as metabolomics, enzyme activity, flux and transcriptomics data, and the exploration of the data in the network context. It builds a foundation for investigating the dynamics of metabolic processes obtained either experimentally or via modelling and simulation. Here we discuss past, present and future of the visualisation of metabolic networks and pathways and provide links to several resources.

**Keywords** Visualisation · Metabolism · Metabolic networks · Visual analytics · Immersive analytics · Bioivis

---

F. Schreiber (✉)

University of Konstanz, Konstanz, Germany  
Monash University, Clayton, VIC, Australia  
e-mail: [falk.schreiber@uni-konstanz.de](mailto:falk.schreiber@uni-konstanz.de)

E. Grafahrend-Belau

Martin Luther University Halle-Wittenberg, Halle (Saale), Germany  
e-mail: [eva.grafahrend-belau@pharmazie.uni-halle.de](mailto:eva.grafahrend-belau@pharmazie.uni-halle.de)

O. Kohlbacher

University of Tübingen, Tübingen, Germany  
e-mail: [oliver.kohlbacher@uni-tuebingen.de](mailto:oliver.kohlbacher@uni-tuebingen.de)

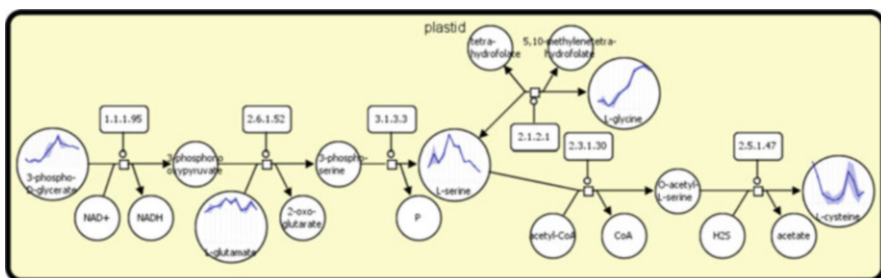
H. Mi

University of Southern California, Los Angeles, CA, USA  
e-mail: [huaiyumi@usc.edu](mailto:huaiyumi@usc.edu)

## 12.1 Introduction

Visualisations of metabolic pathways and pathway components such as enzymes and compounds have been used since the early years of research in biology, and metabolic pathway maps have become very popular in biochemistry textbooks, on posters, as well as in electronic resources and web pages. One example is Gerhard Michal's famous poster *Biochemical Pathways* (Michal, 1968, 1998) which has been printed over a million times. The first example of computational representation of pathways was the EcoCyc database (Karp and Mavrouniotis, 1994; Keseler et al., 2016). Another well-known example is the KEGG pathway database (Kanehisa and Goto, 2000; Kanehisa et al., 2012), the largest collection of manually curated pathway maps and related metabolic information.

Visualisations are commonly used in biology (Gehlenborg et al., 2010; Kerren et al., 2017). Metabolic pathway visualisations help to present knowledge and to support browsing through chemical structures, enzymes, reactions and pathways. Visual and immersive analytics of metabolism connects network analysis algorithms and (interactive and/or immersive) visualisation methods to investigate hubs, motifs, paths and so on in the network, or to compare pathways for finding differences between species or conditions. In addition, network visualisation also supports the mapping and investigation of further data such as metabolomics, proteomics, transcriptomics, enzyme activity and flux data, and the exploration of the data in the network context. It builds a foundation for exploring and navigating the dynamics of metabolic processes obtained either experimentally or via modelling and simulation. In conclusion, visualising and visually exploring metabolic pathways and networks helps in understanding them, is important in making sense of the complex biological data and knowledge that is being produced these days, and is an important research area. A simple visualisation example is shown in Fig. 12.1.



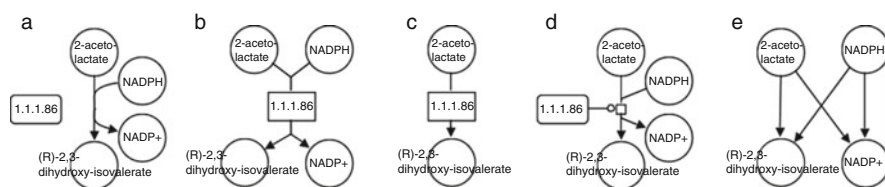
**Fig. 12.1** A metabolic pathway example with some time series data of metabolite concentrations shown within the vertices representing metabolites (excerpt from a MetaCrop pathway (Schreiber et al., 2012) rendered by the Vanted tool (Junker et al., 2006))

### 12.1.1 Network Representation

A network representing metabolic processes consists of a set of elements (called vertices or nodes) and their connections (called edges or arcs) which have a defined appearance (e.g. size of vertices) and are placed in a specific layout (e.g. coordinates of vertices). Typical representations of metabolic reactions as graphs with different interpretations of vertices and edges are shown in Fig. 12.2. Although initiatives for a uniform representation of metabolic pathways have been presented earlier (Kitano, 2003; Kitano et al., 2005; Michal, 1998), no graphical representation became a standard to represent metabolic processes. In 2006 an international consortium started developing a standard for the graphical representation of cellular processes and biological networks including metabolism called the Systems Biology Graphical Notation (SBGN) (Le Novère et al., 2009). SBGN allows the visualisation of complex biological knowledge, including metabolic networks (see also the information in Fig. 12.3). Within this chapter, we will use SBGN for representing metabolic pathways and networks where possible.

### 12.1.2 Network Layout

Metabolic pathway maps have been produced manually for a long time. These drawings are manually created (usually with help of computer programs) long before their actual use and provide a static view of the data defined by the creator. They show the knowledge at the time of the map's generation and an end-user cannot change the visualisation. Some navigation may be supported in electronic systems using such pre-drawn pictures, but the result of an action (the new picture) either



**Fig. 12.2** Different representations of biochemical reactions: **(a)** hypergraph (vertices denote metabolites and enzymes, edges denote reactions); **(b)** bipartite graph with enzymes represented within reactions (vertices denote metabolites and reactions including enzymes, edges connect metabolites with reactions); **(c)** simplified representation of **(b)** without co-substances (as used in KEGG), **(d)** bipartite graph with enzymes represented as separate entities (SBGN notation, vertices denote metabolites, enzymes and reactions, edges connect metabolites with reactions (consumption, production) and enzymes with reactions (catalysis)); **(e)** simplified metabolite network (vertices denote metabolites, edges connect metabolites transformed by reactions). Note that the classical representations such as the ones in Michal's poster (Michal, 1998) and in Stryer's biochemistry textbook (Stryer, 1988) are similar to **(a)**

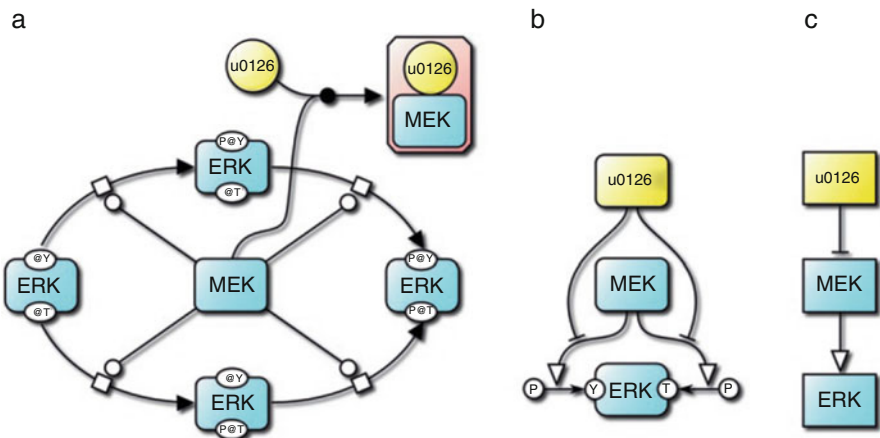
The **Systems Biology Graphical Notation** aims at an unambiguous graphical representation of biological networks and cellular processes. To encode levels of detail and different aspects of the interactions, SBGN provides three orthogonal and complementary languages:

**Process description (PD)** focuses on temporal dependencies of transformations and interactions in a network and describes them in a mechanistic way. It is used to represent networks of events which convert biological entities into each other, change their state or transport them to another location. Entity pool nodes represent pools of simple chemicals, macromolecules, etc.; process nodes encode their transformation or transport. For details see Rougny et al. (2019). SBGN PD is commonly used to graphically represent metabolic pathways and networks.

**Entity relationship (ER)** focuses on the relationship and influences in which the entities are involved or which they have onto each other. It does not consider temporal aspects but describes relationships in a mechanistic manner and shows all possible relationships at once. Entity nodes in ER represent entities that exist; relationships are rules that decide whether an entity node exists. Details can be found in the technical specification (Sorokin et al., 2015).

**Activity flow (AF)** focuses on the biological activity. It shows the sequential influence of activities and can be ambiguous when it comes to the underlying mechanism. Activity nodes represent the biological activities; modulation arcs represent the influence of activities onto other activities. For details see Mi et al. (2015).

Detailed information on SBGN is available at [www.sbgng.org](http://www.sbgng.org), as part of the yearly special issue on standards in systems and synthetic biology (Schreiber et al., 2020) and here (Junker et al., 2012). Software support for SBGN maps is described by van Iersel et al. (2012) and on [www.sbgng.org](http://www.sbgng.org).



**Fig. 12.3** Box SBGN: Explanation of SBGN and SBGN languages, the image shows as example protein phosphorylation catalysed by an enzyme and modulated by an inhibitor in all three SBGN languages: (a) PD, (b) ER, (c) AF (image from Le Novère et al. (2009))

replaces the current image or it is visualised in a new view, and the visualisations are not interactive.

However, the automatic computation of visualisations and interactive exploration methods are highly desirable, due to size and complexity of biological networks, a

steady growth of knowledge and the derivation of user-specific parts of networks from databases. The computer-based generation (layout) of a network map on demand at the time it is needed is called dynamic visualisation. These visualisations are created by the end-user from up-to-date data with help of a layout algorithm. They can be modified to provide specific views of the data, and several navigation methods such as the extension of an existing drawing or map with new additional parts are supported.

The automatic layout of networks, that is the computation of maps from a given network, is called graph drawing. Graph drawing methods take a network (or graph) and compute a layout consisting of coordinates for the vertices and routings of the edges. See the books by Di Battista et al. (1999) and Kaufmann and Wagner (2001) for general graph drawing algorithms. Although standard graph drawing algorithms can be used for laying out metabolic networks, domain specific network visualisations that conform to biological representational conventions are advantageous (Bourqui et al., 2011; Schreiber, 2002).

In the following sections we will discuss major resources for metabolites, reactions and pathways (Sect. 12.2); the visualisation of metabolites and enzymes, which are the building blocks of metabolic pathways (Sect. 12.3); and the visualisation of the pathways themselves (Sect. 12.4). We will focus on key questions which can be addressed using visualisation, present important graph drawing algorithms in brief (both standard and domain specific algorithms) and discuss a selection of useful tools. Next we will discuss the exploration and analytics of pathways and data, in particular visual analytics and immersive analytics of metabolic pathways and related data (Sect. 12.5). We conclude with perspectives and research questions in this field. Boxes contain additional background information regarding, for example, standards for metabolic network representation and layout algorithms.

## 12.2 Resources for Metabolites, Reactions and Pathways

Large amounts of knowledge about metabolites, enzymes, metabolic reactions, pathways and networks have been accumulated and are derived at increasing speed. Several databases and information systems have been developed to provide a comprehensive way to manage, explore and export this knowledge in meaningful ways. We will concentrate on the most important primary databases and briefly discuss their typical content and important features.

Databases in this area can be divided into *metabolite/compound databases* providing information about the chemical compounds used or produced in biochemical reactions; *reaction/enzyme databases* containing information about enzymes and the reactions catalysed by them; and *pathway databases* providing information about metabolic pathways. See also Table 12.1 for more information and a comparison of relevant databases.

**Table 12.1** Databases. Abbreviations: W - Web, S - Web services, F - FTP

Database	Ease of use	Level of detail	Data access	Support of exchange formats	URL
<i>Metabolite/compound databases</i>					
ChEBI	++	+	W, S, F	TSV, XML, OBO, OWL, MOL, SDF	bit.ly/zWCpY1
KEGG COM- POUND	+	+	W, S, F	MDL/MOL, KCF, Jmol, KegDraw, RDF	bit.ly/wxHkVi
PubChem	++	++	W, S, F	ASNT, XML, CSV, SDF, JSON, PNG	bit.ly/2JsSyOm
<i>Reaction/enzyme databases</i>					
BRENDA	++	++	W, S	SBML, Fasta, CSV	bit.ly/xGVzgz3
ExPASy-ENZYME	+		W, F		bit.ly/33yii31
KEGG ENZYME			W, S, F	RDF	bit.ly/3fSXiZE
Rhea			W	TSV, BioPAX, RDF, RXN	bit.ly/AfG1d4
Sabio-RK	+	++	W, S	SBML, BioPAX, SBPAX, XLS, MatLab, TSV	bit.ly/zZ7Ax0
<i>Metabolic pathway databases</i>					
MetaCyc	+	++	W, S	SBML, BioPAX, DB, TSV, ...	bit.ly/2JfJigV
KEGG PATHWAY	++	+	W, S, F	KGML	bit.ly/w2urRG
PANTHER Pathway	+	++	W, S, F	SBML, SBGN, BioPAX	bit.ly/x6aQ9n
Reactome	++	++	W, S	SBML, BioPAX, DB, SBGN, PSI, PPTX, PNG, ...	bit.ly/9RtvaZ

### 12.2.1 Metabolite/Compound Databases (Chemical Databases)

The typical content is information about metabolites (compounds) and their properties such as name, synonyms, molecular weight, molecular formula and structure. Often associated information such as chemical reactions, metabolic pathways, publications and various links to other databases can also be found.

Important resources are PubChem (Kim et al., 2015, 2020; Wang et al., 2009) is a comprehensive source of compound and substance information (consisting of the three primary databases: Compounds, Substances and BioAssay). KEGG COMPOUND (Goto et al., 2002) is a database of small molecules, biopolymers and other chemical substances of biological interest. ChEBI (Hastings et al., 2015) is a database of small molecules with detailed information about nomenclature, molecular structure, formula and mass.

The visualisation and visual analysis of data from these databases is discussed in Sect. 12.3.

### ***12.2.2 Reaction/Enzyme Databases***

The typical content is information about enzymes and their properties such as nomenclature, enzyme structure, functional parameters and specificity. Often additional information about the reactions catalysed by the given enzyme, metabolic pathways, references and links to other databases can also be found.

Important resources are: BRENDA (Chang et al., 2020; Scheer et al., 2011) is a comprehensive enzyme information system providing detailed molecular and biochemical information on enzymes based on primary literature. ExPASy-ENZYME (Gasteiger et al., 2003) is an enzyme database which covers information related to the nomenclature of enzymes. Rhea (Lombardot et al., 2018) is an expert-curated reaction database with information about biochemical reactions and reaction participants. The KEGG databases ENZYME and REACTION (Kanehisa and Goto, 2000; Kanehisa et al., 2004) provide enzyme- and reaction-specific information about chemical reactions in the KEGG metabolic pathway database. Sabio-RK (Wittig et al., 2012, 2017) is an expert-curated biochemical reaction kinetics database with detailed kinetic information.

The visualisation and visual analysis of data from these databases is discussed in Sect. 12.3.

### ***12.2.3 Metabolic Pathway Databases***

The typical content is information about metabolic pathways and their single reactions, involved enzymes and reactants and associated information such as organism-specific information about genes, their related gene products, protein functions and expression data. Often several types of information are provided in the context of the graphical representation of pathways.

Major databases are: KEGG PATHWAY (Kanehisa et al., 2002, 2020), a multi-organism pathway database which contains metabolic pathways, represented as curated, manually drawn pathway maps consisting of links to information about compounds, enzymes, reactions and genes. BioCyc/MetaCyc (Caspi et al., 2012, 2019; Krieger et al., 2004) is a collection of organism-specific pathway databases including MetaCyc, a curated multiorganism pathway database, which contains metabolic pathways curated from the literature, lists of compounds, enzymes, reactions, genes and proteins associated with the pathways. Reactome (Croft et al., 2014; Matthews et al., 2009) is a curated multi-organism pathway database initially focussing on human biology. PANTHER pathway (Mi and Thomas, 2009; Mi et al., 2020) is an expert-curated multi-organism pathway database.

In addition to the mentioned primary databases, there are secondary pathway databases and collaborative databases. Secondary metabolic pathway database systems are collecting and presenting information from various sources. Examples are NCBI BioSystems (Geer et al., 2010) and Pathway Commons (Cerami et al., 2011; Rodchenkov et al., 2019). The former is a centralised repository for metabolic pathway information containing biological pathways from multiple databases (e.g. KEGG, Human Reactome, BioCyc and the National Cancer Institute's Pathway Interaction database). The latter provides access to various public metabolic pathway databases, such as Reactome, HumanCyc and IMID. A well-known community-driven collaborative platform dedicated to the curation and representation of biological pathways is WikiPathways (Martens et al., 2021).

There is also BioModelsDB (Chelliah et al., 2013; Malik-Sheriff et al., 2019), a database of mathematical models representing biological processes including metabolism, and the BioModelsDB part Path2Models (Büchel et al., 2013), an automatic translation of metabolism from databases such as KEGG into biological models using the SBML and SBGN standards. In addition, there are also special metabolic pathway databases covering specific species or groups of species, for example, for plants PlantCyc (Schlöpfer et al., 2017; Zhang et al., 2010), MetaCrop (Grafahrend-Belau et al., 2008; Weise et al., 2006) and Plant Reactome (Naithani et al., 2016, 2019).

The visualisation of data from these databases is discussed in Sect. 12.4. The above mentioned databases also provide static (e.g. KEGG) or dynamic (e.g. BioCyc) visualisations of pathways and networks. Furthermore, they often come with integrated analysis tools to support high-throughput experimental data analysis. For example, Reactome visualises pathways and maps expression data using colour-coding onto pathway maps. A Cytoscape plugin enables to generate new pathways based on database queries and to perform some graph analysis on these networks. KegArray is a light-weight data mapping utility, good for easily mapping expression data (csv) onto KEGG pathways to colour-code vertices and provides also some scatter plots of the data. And PANTHER Pathways allow users to view results in both SBGN process view and an automatically converted activity flow view. However, these tools are specifically developed for specific databases and often provide less functionality than the best general purpose tools presented in Sect. 12.4.3; therefore we will not present them in detail here.

### 12.2.4 *Exchange Formats*

To represent metabolic pathway information in a unified way and to support the exchange of pathway models between software tools, exchange formats have been proposed. Two exchange formats which focus on the exchange of information between software tools and databases are SBML (Hucka et al., 2003) and BioPAX (Demir et al., 2010), see the box in Fig. 12.4. Although they also partly support the exchange of graphical information, they are mainly relevant for software



**SBML (Systems Biology Markup Language)** is a machine-readable format for representing pathway models [Keating et al., 2020]. SBML has been developed by an international community of software developers and systems biologists to provide a common format for data sharing between various computer-modelling software applications. It is neutral with respect to software encoding and programming languages, and oriented towards enabling XML-encoded models. Software tools which use SBML as a format for writing and reading models can exchange the same computable representation of those models. Today, around 230 software packages support SBML. Detailed documentation on the SBML format is available online at [www.sbml.org](http://www.sbml.org).

**BioPAX** is a collaborative effort to create a data-exchange format for biological pathway data [Demir et al., 2010]. The aim is to support accessing, sharing and integrating data from multiple pathway databases. BioPAX supports the representation of metabolic and signalling pathways, molecular and genetic interactions as well as gene regulation. It describes relationships between genes, small molecules, complexes and their states, including the results of events. Detailed documentation on the BioPAX format is available online at [www.biopax.org](http://www.biopax.org). BioPAX is complementary to other standard pathway information-exchange languages, such as SBML, focusing on qualitative, large networks and their integration rather than on mathematical modelling of quantitative, small models.

SBML, BioPAX as well as other standards such as SBGN are part of the COMBINE initiative [Waltemath et al., 2020] and are used from single reaction to small pathway to larger and whole cell models [Waltemath et al., 2016].

**Fig. 12.4** Box BioPAX and SBML

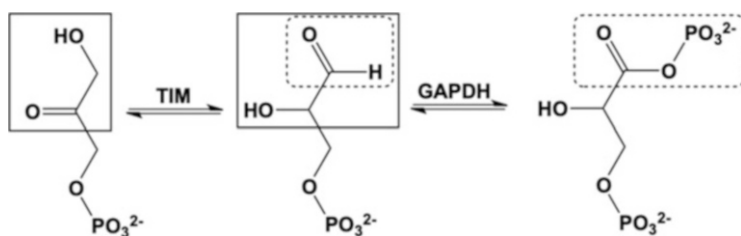
developers and modellers. The exchange format relevant for transferring graphical information and most relevant for users only interested in the visual representation of pathways is SBGN, the Systems Biology Graphical Notation (Le Novère et al., 2009), see the box in Fig. 12.3. Exchange of metabolite structures relies on several formats popularised in cheminformatics and drug design. MDL Molfiles are suitable for storing individual structures; collections of these files can then be assembled into SD Files (Dalby et al., 1992). While some pathway databases provide small molecule structures as MOL or SD files (e.g. KEGG, ChEBI), others provide the structures as SMILES (Weininger, 1988). SMILES, a so-called line notation, encodes the structure as a string and thus does not provide atom coordinates, but provides a more compact representation. A multitude of other file formats exists; most of these formats can be easily accessed and interconverted by cheminformatics toolkits and libraries (e.g. CDK (Steinbeck et al., 2003)) or conversion utilities (e.g. OpenBabel (O'Boyle et al., 2011)).

## 12.3 Visualising Metabolites and Enzymes

Textbook views of metabolic pathways often illustrate the underlying biochemical mechanisms. To this end it is essential not just to provide the name of the metabolites. Structural drawings are much better suited to illustrate the molecular details of an enzymatic reaction. The example in Fig. 12.5 shows the reactions catalysed by triosephosphate isomerase and glyceraldehyde 3-phosphate dehydrogenase, the isomerisation of dihydroxyacetone phosphate to D-glyceraldehyde-3-phosphate to D-glycerate 1,3-bisphosphate. Visualisation of the metabolites by their names, IDs or abbreviation makes it hard to understand the mechanism, while the structural drawings immediately reveal the conversion of the hydroxyl group to an aldehyde and of the keto group to a hydroxyl group and subsequent introduction of a phosphate group. The layout of the structural formulas has been designed to highlight the fact that the larger part of the structure remains unchanged during the two reactions. Only parts of the structure (highlighted by the boxes) are modified in the reaction. Manual layouts of metabolic pathways typically found in biochemistry textbooks are thus careful with the layout of both the structures and the pathway to maximise the mental map preservation between adjacent structures.

While drawing structural formulas comes natural to chemists and biochemists, the automated generation of structural formulas is a difficult task. The drawings have to adhere to numerous conventions developed since their initial conception by Kekulé towards the end of the 1800s. While many of these conventions have been standardised by the International Union of Pure and Applied Chemistry (IUPAC), there is no unique way for drawing a chemical structure; it can be adapted depending on the context, the level of detail required, and the information that needs to be conveyed.

Most small molecule chemical structures can be represented by planar graphs and thus can be laid out in 2D without major issues (Rücker and Meringer, 2002). Specific conventions have to be followed with respect to angles, representation of stereochemistry or bond orders, to name just a few. Ring systems pose particular



**Fig. 12.5** Triosephosphate isomerase (TIM) catalyses the conversion of dihydroxyacetone phosphate to D-glyceraldehyde-3-phosphate, which in turn can be converted to D-glycerate 1,3-bisphosphate by glyceraldehyde 3-phosphate dehydrogenase (GAPDH). A consistent layout of the three metabolites involved makes it easier to grasp the structural changes entailed by each metabolic reaction (highlighted by the boxes)

challenges, since they are typically drawn in very specific ways and more often than not projection of the three-dimensional shape is preferred over a non-crossing planar embedding of the final structure. It could be shown that already the drawing of planar graphs with fixed edge lengths (as is the case for structural diagrams) is NP-hard (Eades and Wormald, 1990), most of these algorithms have to resort to heuristics to generate good layouts.

Several algorithms have been proposed over the years to layout structures in an aesthetic manner (Clark et al., 2006; Helson, 1990). In addition, a number of algorithms have been implemented in commercial tools for structure editing and cheminformatics, for example, in the ChemDraw suite,<sup>1</sup> in Accelrys Draw,<sup>2</sup> or in MOE.<sup>3</sup> Also academic cheminformatics projects such as CACTVS (Ihlenfeld et al., 1994) or the more recent Chemistry Development Kit, CDK, (Steinbeck et al., 2003) permit the layout of molecular structures. Based on the structure stored in pathway databases (see Sect. 12.2) these tools permit the rendering of the structure into a 2D image. Another option for the retrieval of structure drawings is PubChem,<sup>4</sup> which contains pre-computed structural formulas. These can be downloaded in PNG format.

A challenge in the visualisation of metabolic networks is currently the joint layout of the metabolic network and its constituent metabolites. While it is in principle possible to layout metabolic networks and simultaneously display the structural formulas of its metabolites, current pathway visualisation tools do not consider this problem (see, for example, Fig. 12.6). Not all tools are able to display structural formulas at all. Those that do, resort to pre-rendered images of the structures. If the structures are drawn individually, their orientation depends mostly on the algorithm used—there is no canonical orientation of a molecular structure. The orientation, size and general layout of any two structures adjacent in a metabolic network are thus mostly random, and it becomes very difficult to match the conserved common substructure between the two structures and thus to comprehend the underlying mechanism. Mental map preservation between any two adjacent structures would of course be preferable and clearly enhance readability of the pathway. Hand-made pathway diagrams found in textbooks are thus so far vastly superior to automatically drawn pathways with structural formulas. The simultaneous constrained drawing of metabolite structures and metabolic pathways is one of the more difficult problems in this area. Some algorithms for the constrained drawing of structures that should be suitable to solve this problem have been suggested in the literature in different contexts (Boissonnat et al., 2000; Fricker et al., 2004).

For small molecules (metabolites) 2D visualisation is the method of choice, because the structures are easier to comprehend and—to the schooled eye—the three-dimensional aspects of the structures are typically obvious. The same does

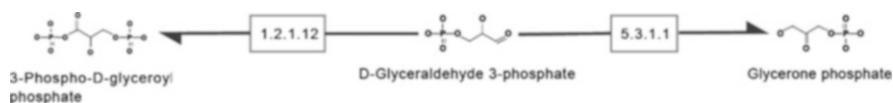
---

<sup>1</sup> CambridgeSoft Corp., Cambridge, MA, USA.

<sup>2</sup> Accelrys, Inc., San Diego, CA, USA.

<sup>3</sup> Chemical Computing Group, Montreal, Canada.

<sup>4</sup> [pubchem.ncbi.nlm.nih.gov](http://pubchem.ncbi.nlm.nih.gov).



**Fig. 12.6** Visualisation of metabolic pathways usually relies on pre-computed metabolite structures. As a consequence, the resulting pathway layout does not match the (usually random) orientation of the structural formulas embedded in the pathway and hampers understanding of the pathway mechanisms; excerpt from a KEGG (Kanehisa et al., 2014) pathway rendered by BiNA (Gerasch et al., 2014)

not apply to proteins, however. Representing proteins as structural formulas is not only impractical, but the function of proteins can only be understood from their three-dimensional structure.

## 12.4 Visualising Reactions and Pathways

### 12.4.1 Visualising the Structure of Metabolic Reactions and Pathways

Visual representations of metabolic pathways are widely used in the life sciences. They help in understanding the interconnections between metabolites, analysing the flow of substances through the network, and identifying main and alternative paths. Important visualisation requirements are (Schreiber, 2002):

- For *parts of reactions*: The level of detail shown concerning specific substances and enzymes is very much dependent on the goal of the visualisation, see also Sect. 12.3. Often for main substances their name and/or structural formula should be shown, for co-substances the name or abbreviation, and for enzymes the name or EC-number.
- For *reactions*: The reaction arrows should be shown from the reactants to the products with enzymes placed on one side of the arrow and co-substances on the opposite side. Both sides of a reaction as well as their reversibility should be visible.
- For *pathways*: The main direction of reactions should be visible to show their temporal order. Few exceptions to the main direction are used to visualise specific pathways such as the fatty acid biosynthesis and the citric acid cycle. The arrangement of these cyclic reaction chains should be emphasised: a repetition of a reaction sequence in which the product of the sequence re-enters as reactant in the next loop, either as cycle (the reactant and the product of the reaction sequence are identical from loop to loop, e. g. citric acid cycle) or as spiral (the reactant of the reaction sequence varies slightly from the product, e. g. fatty acid biosynthesis).

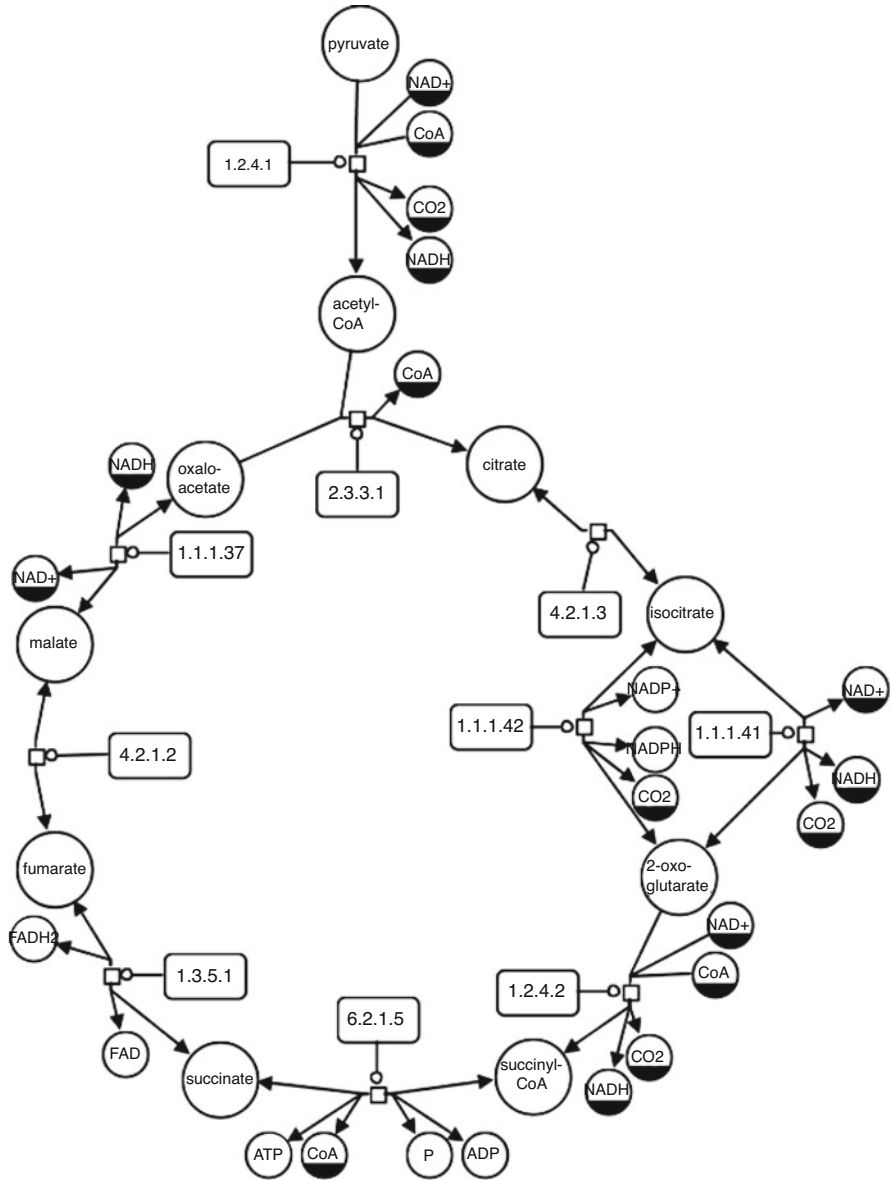
Besides specific visualisation requirements, reaction and pathway visualisations should meet the usual quality criteria of network layouts such as low number of edge crossings and a good usage of the overall area. See Fig. 12.7 for an example which meets these requirements.

### ***12.4.2 Layout Algorithms for Visualising Metabolic Pathways and Networks***

Metabolic networks are usually represented as directed graphs. Common approaches to automatically layout these networks are force-directed and hierarchical (or layered) layout methods. Although quite common as visualisation principal, for example, in the manual KEGG maps layout, automatic orthogonal (or grid) methods are less often used. See the box in Fig. 12.8 as well as the images in Fig. 12.9 for these layout methods. Force-directed methods are widely used, and several network analysis tools support such layouts. However, these approaches do not meet common visualisation requirements. Different vertex sizes, the special placement of co-substances and enzymes, the partitioning of substances into reactants and products and the general direction of pathways are not considered. A few approaches extend the force-directed layout method to deal with application specific requirements. An example is implemented in the PATIKA system (Demir et al., 2002; Dogrusöz et al., 2006) where the layout algorithm considers directional and rectangular regional constraints which can be used to enforce layout directions and sub-cellular locations.

Layered layout methods are often used as they emphasis the main direction within a network. Tools which support such layered layout methods are often based on existing layout libraries. These approaches show the main direction of reactions and are sometimes able to deal with different vertex sizes. However, there is no special placement of co-substances or specific pathways (e. g. cycles). Some improved approaches consider cyclic structures or depict pathways of different topology with different layouts, e. g. the algorithm by Becker and Rojas (2001) which emphasises cyclic structures, and PathDB (Mendes, 2000; Mendes et al., 2000) which visualises metabolic networks based on hierarchical layout allowing co-substances to be represented in a smaller font on the side of the reaction arrow.

There are some advanced methods for the automatic layout of metabolic pathways and networks such as the mixed, the extended layered and the constraint layout. The mixed layout approach (Karp and Mavrouniotis, 1994) depicts (sub-)pathways of different topology with suitable layout algorithms such as linear, circular, tree and hierarchical layout, and places co-substances and enzymes beside reaction arrows. It is used in the MetaCyc/BioCyc database system. The extended layered approach (Schreiber, 2002) extends the hierarchical layout for different vertex sizes, consideration of co-substances and enzymes, and special layout of open and closed cycles; it is implemented in BioPath system (Brandenburg et al.,



**Fig. 12.7** Example of metabolic pathway visualisation which meets the requirements outlined in Section 4.1 (citric acid cycle, including reversible and irreversible reactions and circular shape of the pathway; excerpt from a MetaCrop (Hippe et al., 2010) pathway rendered by Vanted (Colmsee et al., 2013))

**Force-directed methods**

The idea of force-directed layout methods is to simulate a system of physical forces (e.g. spring or magnetic forces) on the graph and compute a distribution of the vertices with low overall force or energy. Standard force-directed graph drawing algorithms look like:

1. Compute the initial layout  $d_0$  of graph  $G$   
 $d_0(G) = \text{init}(G)$
2. While not finished
  - a) for all vertices  $v$  from vertex set  $V$   
 Compute the current force to this vertex, e.g. energy, magnetic, distribution, penalties  
 $F(v) = \text{force}(v, d_i(G))$
  - b) for vertex / vertices from vertex set  $V$  (chosen one vertex, vertex set depending of strategy  $s$ , or all vertices)  
 Compute a new layout  $d_{i+1}(G)$   
 $d_{i+1}(G) = \text{move } v \text{ according to } F(v)$

For specific algorithms the parameters of this general method have to be specified; several methods could be used to compute the initial layout (e.g. random placement or pre-defined from an application), the current force, the strategy for choosing a vertex and how the vertex should be moved, the criteria to finish the computation and so on. Force-directed methods can be applied to undirected and directed graphs.

**Layered methods**

Layered graph drawing algorithms (also called hierarchical layouts) consist of the following four phases:

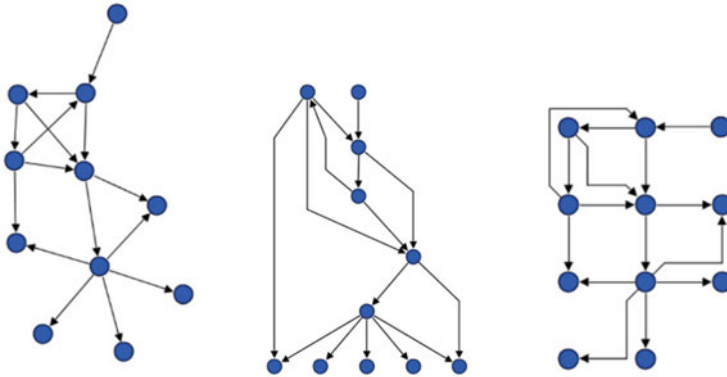
1. The removal of cycles in the graph as a result of changing temporarily the direction of some edges (decycling)
2. The assignment of vertices to layers in a way that all edges obtain the same direction, e.g. from top to bottom (layering)
3. The permutation of vertices within each layer with the aim to reduce edge crossings (crossing reduction)
4. The assignment of coordinates to vertices and of bend points (coordinates) to edges (coordinate assignment)

Layered methods work on directed graphs. To apply them to undirected graphs these graphs have to be changed into directed graphs, e.g. by choosing a start vertex  $v$ , applying breath first search (BFS) starting with  $v$  and directing each edge depending on the order given by the BFS algorithm.

**Fig. 12.8** Box layout algorithms

2004). Finally, the constraint layout approach (Schreiber et al., 2009) allows the expression of visualisation requirements including positions of co-substances and specific pathways as constraints and produces a layout by solving these constraints. This approach is particularly well suited in cases when parts of the layout are predefined as shown in Czauderna et al. (2013).

Figure 12.10 shows examples of visualisations computed by these layout algorithms.



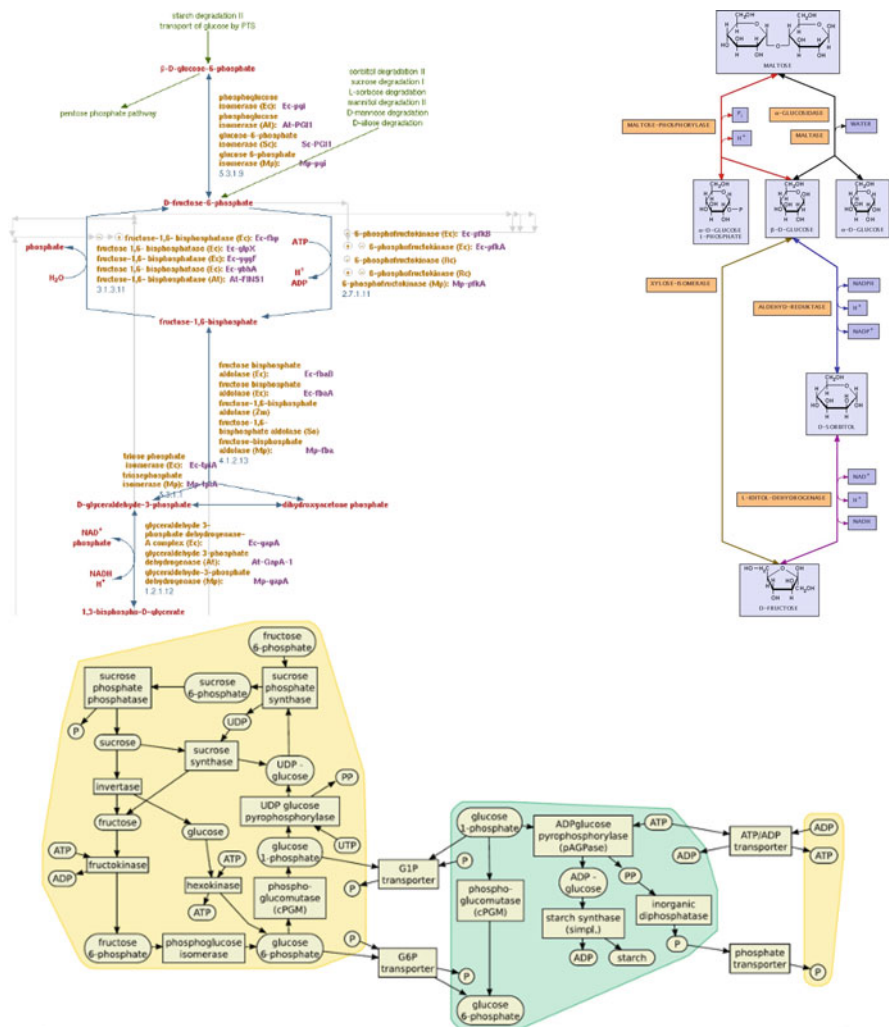
**Fig. 12.9** The same network with three different layouts: (from left to right) force-directed, layered (top to bottom) and orthogonal layout

### 12.4.3 Tools

There are more than 170 tools available, and previous reviews have already compared a number of them. Kono et al. focus in their comparison on pathway representation, data access, data export and exchange, mapping, editing and availability (Kono et al., 2009). Suderman and Hallett compare more than 35 tools relevant in 2007 regarding several aspects of network and data visualisation (Suderman and Hallett, 2007). Rohn et al. present a comparison of 11 non-commercial tools for the network-centred visualisation and analysis of biological data (Rohn et al., 2012). And Gehlenborg et al. present visualisation tools for interaction networks and biological pathways including tools for multivariant omics data visualisation (Gehlenborg et al., 2010). It should be noted that progress in this field is fast, many new tools appeared and old tools obtained new features since then. Well-known tools supporting network visualisation and analysis are:

- *BiNa* (Gerasch et al., 2014; Küntzer et al., 2007) (<http://bit.ly/y6ix9i>)
- *BioUML* (Kolpakov, 2002; Kolpakov et al., 2006) (<http://bit.ly/yIETIt>)
- *CellDesigner* (Funahashi et al., 2003, 2006) (<http://bit.ly/A0FQiF>)
- *CellMicrocosmos* (Sommer and Schreiber, 2017a; Sommer et al., 2010) (<http://bit.ly/WJ8cnE>)
- *Cytoscape* (Shannon et al., 2003; Smoot et al., 2011) (<http://bit.ly/wY2sbG>)
- *MapMan* (Thimm et al., 2004; Usadel et al., 2005) (<http://bit.ly/3yaa6UE>)
- *OMIX* (Droste et al., 2011) (<http://bit.ly/wY2sbG>)
- *Ondex* (Köhler et al., 2006) (<http://bit.ly/AetZjz>)
- *Pathway Projector* (Kono et al., 2009) (<http://bit.ly/zo5x2M>)
- *PathVisio* (van Iersel et al., 2008) (<http://bit.ly/zunwxW>)
- *SBGN-ED* (Czauderna et al., 2010) (<http://bit.ly/17m7KfW>)
- *Vanted* (Junker et al., 2006; Rohn et al., 2012) (<http://bit.ly/Aigr0T>)
- *VisAnt* (Hu et al., 2004, 2009) (<http://bit.ly/agZBni>)





**Fig. 12.10** Example visualisations computed by layout algorithms specifically tailored to metabolic networks: (top left) mixed layout (from the MetaCyc webpage), (top right) extended layered layout (from BioPath) and (bottom) constraint layout (from a prototype implementing of the constraint layout approach; note that these networks are not in SBGN notation)

These tools often provide a selection of standard and partly specific layout algorithms for metabolic pathways, the possibility to map additional data onto pathways as well as analysis algorithms.

Note that for a specific metabolic database or pathway collection often several different visualisation methods exists. For example, the visualisation of KEGG pathways can be done with tools and layout methods such as implemented in Pathway projector (Kono et al., 2009), KEGGgraph (Zhang and Wiemann, 2009)

and Vanted (Rohn et al., 2012), can be rebuilt and visualised as in Gerasch et al. (2014), or can be even translated into SBML or SBGN and then layouted and visualised (Czuderna et al., 2013; Wrzodek et al., 2011).

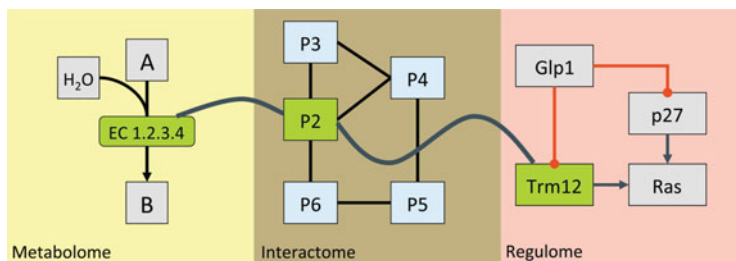
## 12.5 Visual and Immersive Analytics of Metabolic Pathways and Related Data

For a fast and automatic production of pictures or maps of metabolic networks layout algorithms are very useful. However, a layout is just the first step, and in interactive systems many additional requirements exist, for example, for interactive exploration, structural analysis of the networks, visualisation of experimental data (transcriptomics, metabolomics, fluxes, etc.) in the network context, studying networks in their spatial (3D) context and so on. Here we discuss some typical examples.

### 12.5.1 *Multiscale Representation of Metabolism and Navigation Through Metabolic Networks*

Metabolic networks can be huge, and a visualisation may become unreadable due to the large number of objects and connections. Several abstraction and exploration techniques have been transferred from the field of information visualisation to navigate in metabolic networks. As metabolic pathways are hierarchically structured (e. g. carbohydrate metabolism includes a number of sub-pathways such as TCA cycle and glycolysis) this information can be used to help navigating through the network. Often used navigation techniques include clickable overview maps (in many databases and tools, e. g. KEGG (Kanehisa et al., 2002) and iPath (Letunic et al., 2008)), maps showing increasing levels of detail (e. g. the MetaCyc website (Caspi et al., 2012)), interconnected maps (e. g. in GLIEP (Jusufi et al., 2012)), overview and detail diagrams (e. g. method by Garkov et al. (2019)) and interactive extension of pathways within a map (e. g. the method in KGML-ED (Klukas and Schreiber, 2007)).

It should be noted that there is a major obstacle for simple interactive visualisation methods including automatic layout: the mental map of the user (Misue et al., 1995). When browsing through pathways the user builds a mental representation of the objects, their relative position and connections. Basically the user's mental map is its understanding of the network based on the current view. However, sudden or large changes between the current and the next view destroy the user's mental map and therefore hinder interactive understanding of networks. So far there are only few approaches which address this problem.



**Fig. 12.11** Multivariate networks: Different networks are connected through shared entities (from Kohlbacher et al. (2014))

Metabolic networks can be part of multivariate networks (Kohlbacher et al., 2014) (see Fig. 12.11) and heterogeneous networks (Schreiber et al., 2014), both increase the complexity for representation and navigation. The development of interactive layout algorithms for these structures is still an open research problem, and so far only some initial approaches exist such as the previously mentioned constraint layout approach (Schreiber et al., 2009).

### 12.5.2 Visual Analytics of the Structure of Metabolic Networks

Analysing structural properties in biological networks can help in gaining new insights, and there are several structural properties of interest in metabolic networks: shortest paths between metabolites which may represent preferred routes, network motifs within the network which may indicate functional properties, different centralities of metabolites and reactions which may correspond to their importance, and clusters or communities within the metabolic network which may structure the network into functional modules. Many network analysis algorithms which can be used for the investigation of structural properties in networks have been developed; overviews can be found in the book of Brandes and Erlebach (2005) and Junker and Schreiber (2008).

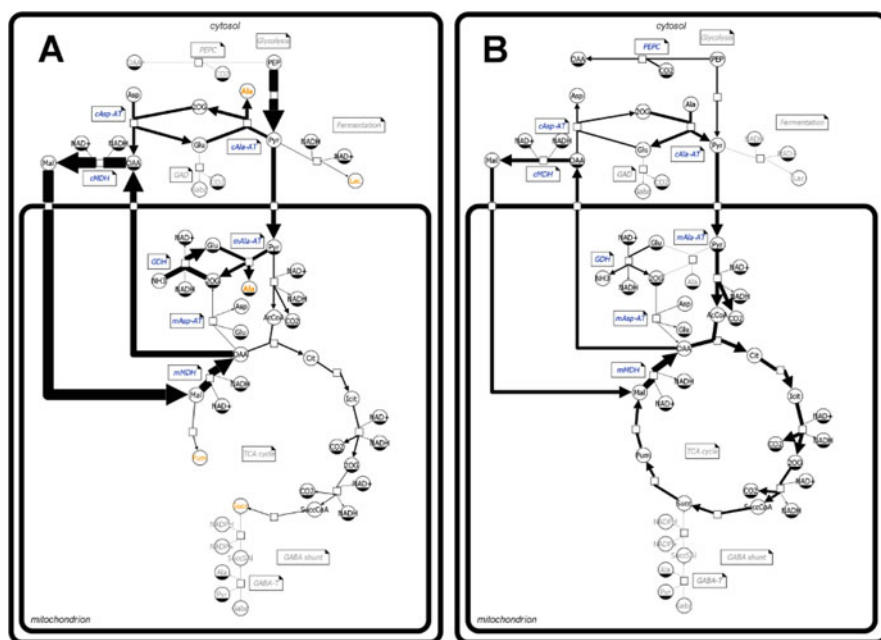
“Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces” (Thomas and Cook, 2006). An important aspect of this field is that data analysis is combined with interactive visualisation methods. Here, analytics includes structural analysis of networks as well as investigating additional data as discussed in the following Sect. 12.5.3. Also for metabolic networks interaction plays an important role in visual analytics (Kerren and Schreiber, 2012).

Several tools implement visual analytics methods, for example, Cytoscape, Ondex, Vanted, and VisAnt, often provided via additional Plugins/Add-ons (see also Sect. 12.4.3 for details and references). Some tools also allow the integration of a wide range of other data into the analysis (Rohn et al., 2012). To better understand the analysis results, visualisation algorithms can help by highlighting the relevant

structures such as straightening the shortest path in the map, putting central elements in the centre of the image or laying out the same motifs in the same way. A few specialised layout algorithms have been developed for a better visualisation and graphical investigation of structures and connections in networks such as coordinated perspectives for the analysis of network motifs (Klukas et al., 2006) or visually comparing pathways, for example, to understand metabolic pathways across organisms using two and a half dimensional layout (Brandes et al., 2004).

### 12.5.3 Integration and Visualisation of Omics Data in Metabolic Networks

Data mapping deals with the integration of additional data into metabolic networks. Examples are metabolomics, transcriptomics and fluxomics measurements, which can be mapped on different network elements (such as metabolites, enzymes and reaction edges), see also Figs. 12.1 and 12.12. A common problem for data mapping and subsequent analysis such as correlation analysis and clustering is the usage of correct identifiers, that is having the correct name in both the data and the network. To help the user several tools support mapping tables which translate



**Fig. 12.12** An example of flux visualisation showing the flux distribution in a metabolic network under two scenarios (rendered by Vanted, data from Rolletschek et al. (2011))

identifiers in the data into identifiers in the network, and translation services such as BridgeDB (van Iersel et al., 2010) exist. Depending on the data different diagram types are desired in the vertices or at the edges of the metabolic pathway. Examples are bar charts, pie charts, line charts, box plots and heat maps.

Whereas most tools support the visualisation of data connected to vertices of the network, only few tools provide mapping of data onto edges. Metabolomics data, in particular the results of stable isotope tracer experiments, yield important details on the dynamics of networks, and flux visualisation is important as it provides insights on the integrated response of a biochemical reaction network to environmental changes or genetic modifications. Thus, such representations are also important tools in metabolic engineering (Wiechert, 2001). To support the analysis and understanding of simulated or experimentally measured flux distributions, the visualisation of flux information in the network context is essential and is mainly performed by scaling the width of the reaction edges according to the flux data or by displaying the flux values in the corresponding reaction vertices, see Fig. 12.12. Tools such as FBASimViz (Grafahrend-Belau et al., 2009), MetaFluxNet (Lee et al., 2003), Omix (Droste et al., 2011) and OptFlux (Rocha et al., 2010) support such visualisations.

### ***12.5.4 Immersive Analytics of Metabolic Networks***

The visualisation of structures and pathways in 3D has advantages and disadvantages. A good 2D visualisation may be easier to understand and is directly printable on paper. For small molecules 2D visualisation is the method of choice, because the structures are easier to comprehend and the three-dimensional aspects of the structures are typically obvious to an expert. However, the same does not apply to proteins. Visualising proteins as structural formulas is not only impractical, but the function of proteins can only be understood from their three-dimensional structure. This provides arguments for an integration of 2D (mainly Information Visualisation) and 3D (mainly Scientific Visualisation) techniques (Kerren and Schreiber, 2014).

Early work of representing metabolic pathways in 3D by Qeli et al. (2004) and Rojdestvenski et al. (2003; 2002) goes back to the early 2000. In the last years the novel research field of immersive analytics (Chandler et al., 2015) is developed which is concerned with “the use of engaging, embodied analysis tools to support data understanding and decision making” with a focus on immersive (3D) environments (Dwyer et al., 2018). It builds on and combines ideas from the fields of data visualisation, visual analytics, virtual reality, computer graphics and human–computer interaction. The key idea is to get immersed into the data and employ all senses, not only vision. This area has many potential applications in the life and health sciences (Czauderna et al., 2018). Some initial applications for the visualisation and exploration of metabolism in immersive environments include MinOmics, an immersive tool for multi-omics analysis (Maes et al., 2018) and



**Fig. 12.13** Exploration of metabolic pathways within the spatial context using an immersive environment based on CAVE2 (stereoscopic 3D) and zSpace (stereoscopic fishtank 3D) (from Sommer and Schreiber (2017b))

the integration and exploration of pathways in a cell environment (Sommer and Schreiber, 2017b) as shown in Fig. 12.13.

## 12.6 Perspectives

The visual exploration and analytics of metabolic networks is a fast developing field. Although there are already several methods and tools that help in understanding metabolic networks, continuous development is imminent. Here we outline some current directions of research and tool developments in this area:

- Connection to other networks: Metabolism is strongly linked to other biological processes represented, for example, by protein interaction or gene regulatory networks (see also Sect. 12.5). The combined visualisation and easy visual travelling from one network to the next may help in better understanding effects such as regulation of metabolism.
- Context for combined omics data: Although several tools support integration and visualisation of omics data within metabolic networks, the visualisation of complex data sets covering several domains (networks, images, sequences, omics data, etc.) is not yet sufficiently solved. Initial solutions have been presented (e.g. Rohn et al. (2011)), but as more and more such data sets are produced in experiments, there is an increasing need for better analysis and visualisation approaches.

- Mental map preserving layouts: A mental map of a layout is a mental picture of the structure of the layout and helps understanding changing maps (Misue et al., 1995), see also Sect. 12.5.1. It is often used to measure the quality of a dynamic network layout (Archambault et al., 2011), and has been shown to be important in dynamic layouts (Purchase and Samra, 2008; Purchase et al., 2007). Most existing layout algorithms are not mental map preserving and often the same algorithm would produce different visualisations when applied to the same network. Also, there are only a few studies regarding mental map preserving network layouts in visual and immersive analytics (e. g. Kotlarek et al. 2020). However, acceptance of visualisation and exploration methods also depends on better support of the user’s mental map and this is an important area for future research.

Biological network visualisation and the layout of metabolic networks is an interesting area in graph drawing (Binucci et al., 2019). More open questions and major problems arising in biological network visualisation are also discussed in Albrecht et al. (2009). Metabolic network and pathway visualisation is only a small aspect of biological data visualisation. As biology aims to provide insights into the overall system, that is into processes on cellular, tissue, organ and even organism levels, visualisation of metabolism has to be embedded into broader visualisation frameworks. Beside networks and related data, other data modalities are also important, for example, imaging data and phenotypical data.

This chapter presented history and state-of-the-art of visualisation and visual analysis of metabolic pathways and networks, provided descriptions of important metabolic network databases and exchange formats, gave a brief overview of often used tools and discussed future research directions including immersive analytics. Methods and tools presented here are a building block of such a broader visualisation framework for biological data.

## References

- Albrecht M, Kerren A, Klein K, Kohlbacher O, Mutzel P, Paul W, Schreiber F, Wybrow M (2010) On open problems in biological network visualization. In: Eppstein D, Gansner ER (eds) Graph Drawing 17th International Symposium GD 2009. LNCS, vol 5849. Springer, Berlin, pp 256–267
- Archambault D, Purchase H, Pinaud B (2011) Animation, small multiples, and the effect of mental map preservation in dynamic graphs. *IEEE Trans Visual Comput Graphics* 17(4):539–552
- Becker MY, Rojas I (2001) A graph layout algorithm for drawing metabolic pathways. *Bioinformatics* 17(5):461–467
- Binucci C, Brandes U, Dwyer T, Gronemann M, von Hanxleden R, van Kreveld M, Mutzel P, Schaefer M, Schreiber F, Speckmann B (2019) 10 reasons to get interested in Graph Drawing. In: Computing and software science—state of the art and perspectives. LNCS, vol 10000, pp 85–104
- Boissonnat JD, Cazals F, Flötotto J (2000) 2D-structure drawings of similar molecules. In: Proceedings of the Graph Drawing. Lecture Notes in Computer Sciences, vol 1984. Springer, Berlin, pp 115–126

- Bourqui R, Purchase HC, Jourdan F (2011) Domain specific vs generic network visualization: an evaluation with metabolic networks. In: Proceedings of the Australasian conference on user interface (AUIC '11). Australian Computer Society Inc, New York, pp 9–18
- Brandenburg FJ, Forster M, Pick A, Raitner M, Schreiber F (2004) BioPath—exploration and visualization of biochemical pathways. In: Jünger M, Mutzel P (eds) Graph Drawing Software. Springer, Berlin, pp 215–236
- Brandes U, Erlebach T (2005) Network analysis: methodological foundations. In: Springer book series lecture notes in computer science tutorial
- Brandes U, Dwyer T, Schreiber F (2004) Visual understanding of metabolic pathways across organisms using layout in two and a half dimensions. *J Integr Bioinf* 1:e2
- Büchel F, Rodriguez N, Swainston N, Wrzodek C, Czauderna T, Keller R, Mittag F, Schubert M, Glont M, Golebiewski M, van Iersel MP, Keating, SM, Rall, M, Wybrow M, Hermjakob H, Hucka M, Kell DB, Müller W, Mendes P, Zell A, Chaouiya C, Saez-Rodriguez J, Schreiber F, Laibe C, Dräger A, Le Novère N (2013) Path2models: large-scale generation of computational models from biochemical pathway maps. *BMC Syst Biol* 7:116
- Caspi R, Altman T, Dreher K, Fulcher C, Subhraveti P, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Pujar A, Shearer AG, Travers M, Weerasinghe D, Zhang P, Karp PD (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 40(1):D742–753
- Caspi R, Billington R, Keseler IM, Kothari A, Krummenacker M, Midford PE, Ong WK, Paley S, Subhraveti P, Karp PD (2019) The MetaCyc database of metabolic pathways and enzymes—a 2019 update. *Nucleic Acids Res* 48(1):D445–D453
- Cerami EG, Gross BE, Demir E, Rodchenkov I, Babur O, Anwar N, Schultz N, Bader GD, Sander C (2011) Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res* 39(1):D685–690
- Chandler T, Cordeil M, Czauderna T, Dwyer T, Glowacki J, Goncu C, Klapperstück M, Klein K, Marriott K, Schreiber F, et al (2015) Immersive analytics. In: Big data visual analytics (BDVA), vol 2015. IEEE, New York, pp 1–8
- Chang A, Jeske L, Ulbrich S, Hofmann J, Koblitz J, Schomburg I, Neumann-Schaal M, Jahn D, Schomburg D (2020) BRENDA, the ELIXIR core data resource in 2021: new developments and updates. *Nucleic Acids Res* 49(1):D498–D508
- Chelliah V, Laibe C, Le Novère, N (2013) BioModels Database: A repository of mathematical models of biological processes. *Methods Mol Biol* 1021:189–199
- Clark AM, Labute P, Santavy M (2006) 2D structure depiction. *J Chem Inf Model* 46(3):1107–1123
- Colmsee C, Czauderna T, Grafarend-Belau E, Hartmann A, Lange M, Mascher M, Weise S, Scholz U, Schreiber F (2013) Optimas-DW, MetaCrop and Vanted: a case study for data integration, curation and visualisation in life sciences. In: Proceedings of ontologies and data in life sciences, vol LNI P-220, pp 1834–1840
- Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, Jassal B, Jupe S, Matthews L, May B, Palatnik S, Rothfels K, Shamovsky V, Song H, Williams M, Birney E, Hermjakob H, Stein L, D'Eustachio, P (2014) The Reactome pathway knowledgebase. *Nucleic Acids Res* 42(1):D472–D477
- Czauderna T, Klukas C, Schreiber F (2010) Editing, validating, and translating of SBGN maps. *Bioinformatics* 26(18):2340–2341
- Czauderna T, Wybrow M, Marriott K, Schreiber F (2013) Conversion of KEGG metabolic pathways to SBGN maps including automatic layout. *BMC Bioinf* 14:250
- Czauderna T, Haga J, Kim J, Klapperstück M, Klein K, Kuhlen T, Oeltze-Jafra S, Sommer B, Schreiber F (2018) Immersive analytics applications in life and health sciences. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T, Stuerzlinger W, Thomas BH (eds.) Immersive Analytics. Springer, Berlin, pp 289–330
- Dalby A, Nourse JG, Hounshell WD, Gushurst AKI, Grier DL, Leland BA, Laufer J (1992) Description of several chemical structure file formats used by computer programs developed at Molecular Design Limited. *J Chem Inf Comput Sci* 32(3):244–255



- Demir E, Babur O, Dogrusöz U, Gürsoy A, Nisanci G, Çetin Atalay R, Ozturk M (2002) PATIKA: an integrated visual environment for collaborative construction and analysis of cellular pathways. *Bioinformatics* 18(7):996–1003
- Demir E, Cary MP, Paley S, Fukuda K, Lemer C, Vastrik I, Wu G, D'Eustachio P, Schaefer C, Luciano J, Schacherer F, Martinez-Flores I, Hu Z, Jimenez-Jacinto V, Joshi-Tope G, Kandasamy K, Lopez-Fuentes AC, Mi H, Pichler E, Rodchenkov I, Splendiani A, Tkachev S, Zucker J, Gopinath G, Rajasimha H, Ramakrishnan R, Shah I, Syed M, Anwar N, Babur O, Blinov M, Brauner E, Corwin D, Donaldson S, Gibbons F, Goldberg R, Hornbeck P, Luna A, Murray-Rust P, Neumann E, Reubenacker O, Samwald M, van Iersel M, Wimalaratne S, Allen K, Braun B, Whirl-Carrillo M, Cheung KH, Dahlquist K, Finney A, Gillespie M, Glass E, Gong L, Haw R, Honig M, Hubaut O, Kane D, Krupa S, Kutmon M, Leonard J, Marks D, Merberg D, Petri V, Pico A, Ravenscroft D, Ren L, Shah N, Sunshine M, Tang R, Whaley R, Letovksy S, Buetow KH, Rzhetsky A, Schachter V, Sobral BS, Dogrusoz U, McWeeney S, Aladjem M, Birney E, Collado-Vides J, Goto S, Hucka M, Le Novère NL, Maltsev N, Pandey A, Thomas P, Wingender E, Karp PD, Sander C, Bader GD (2010) The BioPAX community standard for pathway data sharing. *Nat Biotechnol* 28(9):935–942
- Di Battista G, Eades P, Tamassia R, Tollis IG (1999) Graph drawing: algorithms for the visualization of graphs. Prentice Hall, New Jersey
- Dogrusöz U, Erson EZ, Giral E, Demir E, Babur O, Cetintas A, Colak R (2006) Patikaweb: a web interface for analyzing biological pathways through advanced querying and visualization. *Bioinformatics* 22(3):374–375
- Droste P, Miebach S, Niedenführ S, Wiechert W, Nöh K (2011) Visualizing multi-omics data in metabolic networks with the software Omix: a case study. *Biosystems* 105(2):154–161
- Dwyer T, Marriott K, Isenberg T, Klein K, Riche N, Schreiber F, Stuerzlinger W, Thomas BH (2018) Immersive analytics: An introduction. In: Marriott K, Schreiber F, Dwyer T, Klein K, Riche NH, Itoh T, Stuerzlinger W, Thomas BH (eds) Immersive analytics. Springer, Berlin, pp 1–23
- Eades P, Wormald N (1990) Fixed edge length graph drawing is NP-hard. *Discrete Appl Math* 28:111–134
- Fricker PC, Gastreich M, Rarey M (2004) Automated drawing of structural molecular formulas under constraints. *J Chem Inf Comput Sci* 44(3):1065–1078
- Funahashi A, Morohashi M, Kitano H (2003) CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *Biosilico* 1(5):159–162
- Funahashi A, Matsuoka Y, Jouraku A, Kitano H, Kikuchi N (2006) CellDesigner: a modeling tool for biochemical networks. In: Proceedings of the 38th conference on Winter simulation. Winter Simulation Conference, pp 1707–1712
- Garkov D, Klein K, Klukas C, Schreiber F (2019) Mental-map preserving visualisation of partitioned networks in Vanted. *J Integr Bioinf* 16(3):e0026
- Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExpASY: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31(13):3784–3788
- Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH (2010) The NCBI BioSystems database. *Nucleic Acids Res* 38(1):D492–D496
- Gehlenborg N, O'Donoghue SI, Baliga NS, Goemann A, Hibbs MA, Kitano H, Kohlbacher O, Neuweger H, Schneider R, Tenenbaum D, Gavin AC (2010) Visualization of omics data for systems biology. *Nat Methods* 7:S56–S68
- Gerasch A, Faber D, Küntzer J, Niermann P, Kohlbacher O, Lenhof HP, Kaufmann M (2014) BiNA: a visual analytics tool for biological network data. *PLoS One* 9(2):e87397
- Gerasch A, Kaufmann M, Kohlbacher O (2014) Rebuilding KEGG maps: algorithms and benefits. In: 2014 IEEE Pacific Visualization Symposium (PacificVis), pp 97–104
- Goto S, Okuno Y, Hattori M, Nishioka T, Kanehisa M (2002) LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Res* 30(1):402–404

- Grafahrend-Belau E, Weise S, Koschützki D, Scholz U, Junker BH, Schreiber F (2008) MetaCrop—a detailed database of crop plant metabolism. *Nucleic Acids Res* 36(1):D954–D958
- Grafahrend-Belau E, Klukas C, Junker BH, Schreiber F (2009) FBASimViz: interactive visualization of constraint-based metabolic models. *Bioinformatics* 25(20):2755–2757
- Hastings J, Owen G, Dekker A, Ennis M, Kale N, Muthukrishnan V, Turner S, Swainston N, Mendes P, Steinbeck C (2015) ChEBI in 2016: improved services and an expanding collection of metabolites. *Nucleic Acids Res* 44(1):D1214–D1219
- Helson HE (1990) Structure diagram generation. In: Lipkowitz, B, Boyd DB (eds) *Reviews in computational chemistry*. Wiley-VCH, New York, pp 313–398
- Hippe K, Colmsee C, Czauderna T, Grafahrend-Belau E, Junker BH, Klukas C, Scholz U, Schreiber F, Weise S (2010) Novel developments of the metacrop information system for facilitating systems biological approaches. *J Integr Bioinf* 7(3):125
- Hu Z, Mellor J, Wu J, DeLisi C (2004) VisANT: an online visualization and analysis tool for biological interaction data. *BMC Bioinf* 5(1):e17
- Hu Z, Hung JH, Wang Y, Chang YC, Huang CL, Huyck M, DeLisi C (2009) VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology. *Nucleic Acids Res* 37:W115–W121
- 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 I, 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 (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19:524–531
- Ihlenfeld WD, Takahashi Y, Abe H, Sasaki S (1994) Computation and management of chemical properties in CACTVS: an extensible networked approach toward modularity and flexibility. *J Chem Inf Comput Sci* 34:109–116
- Junker BH, Schreiber F (2008) *Analysis of Biological Networks*. In: *Wiley Series on Bioinformatics, Computational Techniques and Engineering*. Wiley, New York
- Junker BH, Klukas C, Schreiber F (2006) Vanted: A system for advanced data analysis and visualization in the context of biological networks. *BMC Bioinf* 7(109):1–13
- Junker A, Rohn H, Czauderna T, Klukas C, Hartmann A, Schreiber F (2012) Creating interactive, web-based and data-enriched maps using the systems biology graphical notation. *Nat Protoc* 7:579–593
- Jusufi I, Klukas C, Kerren A, Schreiber F (2012) Guiding the interactive exploration of metabolic pathway interconnections. *Inf. Visualization* 11(2):136–150
- Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28(1):27–30
- Kanehisa M, Goto S, Kawashima S, Nakaya A (2002) The KEGG databases at GenomeNet. *Nucleic Acids Res* 30(1):42–46
- Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res* 32(1):D277–280
- Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* 40(1):D109–D114
- Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 42(1):D199–D205
- Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M (2020) KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res* 49(D1):D545–D551
- Karp PD, Mavrovouniotis ML (1994) Representing, analyzing, and synthesizing biochemical pathways. *IEEE Expert* 9(2):11–21
- Kaufmann M, Wagner D (2001) *Drawing graphs: methods and models*. In *Lecture Notes in Computer Science Tutorial*, vol 2025. Springer, Berlin

- Keating SM, Waltemath D, König M, Zhang F, Dräger A, Chaouiya C, Bergmann FT, Finney A, Gillespie CS, Helikar T, Hoops S, Malik-Sheriff RS, Moodie SL, Moraru II, Myers CJ, Naldi A, Olivier BG, Sahle S, Schaff JC, Smith LP, Swat MJ, Thieffry D, Watanabe L, Wilkinson DJ, Blinov ML, Begley K, Faeder JR, Gomez HF, Hamm TM, Inagaki Y, Liebermeister W, Lister AL, Lucio D, Mjolsness E, Proctor CJ, et al (2020) SBML Level 3: an extensible format for the exchange and reuse of biological models. *Mol Syst Biol* 8(16):e9110
- Kerren, A, Schreiber F (2012) Toward the role of interaction in visual analytics. In: Rose O, Uhrmacher AM (eds) *Proceedings winter simulation conference*, p 420
- Kerren, A, Schreiber F (2014) Why integrate infovis and scivis? an example from systems biology. *IEEE Comput Graphics Appl* 34(6):69–73
- Kerren A, Kucher K, Li YF, Schreiber F (2017) BioVis Explorer: a visual guide for biological data visualization techniques. *PLoS One* 12(11):1–14
- Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, Fulcher C, Gama-Castro S, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Peralta-Gil M, Subhraveti P, Velázquez-Ramírez DA, Weaver D, Collado-Vides J, Paulsen I, Karp PD (2016) The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res* 45(1):D543–D550
- Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2015) PubChem Substance and Compound databases. *Nucleic Acids Res* 44(1):D1202–D1213
- Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B, Zaslavsky L, Zhang J, Bolton EE (2020) PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res* 49(1):D1388–D1395
- Kitano H (2003) A graphical notation for biochemical networks. *Biosilico* 1(5):169–176
- Kitano H, Funahashi A, Matsuoka Y, Oda K (2005) Using process diagrams for the graphical representation of biological networks. *Nat Biotechnol* 23:961–966
- Klukas C, Schreiber F (2007) Dynamic exploration and editing of KEGG pathway diagrams. *Bioinformatics* 23(3):344–350
- Klukas C, Schreiber F, Schwöbbermeyer H (2006) Coordinated perspectives and enhanced force-directed layout for the analysis of network motifs. In: Misue K, Sugiyama K, Tanaka J (eds) *Proceedings Asia-Pacific symposium on information visualization (APVis'06)*. CRPIT, vol. 60. ACS, New York, pp. 39–48
- Köhler J, Baumbach J, Taubert J, Specht M, Skusa A, Rüegg A, Rawlings C, Verrier P, Philipp S (2006) Graph-based analysis and visualization of experimental results with ONDEX. *Bioinformatics* 22(11):1383–1390
- Kohlbacher O, Schreiber F, Ward MO (2014) Multivariate networks in the life sciences. In: *Multivariate network visualization*, Springer, Berlin, pp 61–73
- Kolpakov FA (2002) BioUML—framework for visual modeling and simulation of biological systems. In: *Proceedings of the international conference on bioinformatics of genome regulation and structure*, pp 130–133
- Kolpakov F, Puzanov M, Koshukov A (2006) BioUML: Visual modeling, automated code generation and simulation of biological systems. In: *Proceedings of the fifth international conference on bioinformatics of genome regulation and structure (BGRS)*, pp 281–284
- Kono N, Arakawa K, Ogawa R, Kido N, Oshita K, Ikegami K, Tamaki S, Tomita M (2009) Pathway projector: web-based zoomable pathway browser using KEGG atlas and Google Maps API. *PLoS One* 4(11):e7710
- Kono N, Arakawa K, Ogawa R, Kido N, Oshita K, Ikegami K, Tamaki S, Tomita M (2009) Pathway projector: web-based zoomable pathway browser using KEGG atlas and Google maps API. *PLoS One* 4(11):e7710
- Kotlarek J, Kwon OH, Ma KL, Eades P, Kerren A, Klein K, Schreiber F (2020) A study of mental maps in immersive network visualization. In: *Proceedings of the IEEE Pacific visualization*. IEEE, New York, pp 1–10

- Krieger CJ, Zhang P, Müller LA, Wang A, Paley S, Arnaud M, Pick J, Rhee SY, Karp PD (2004) MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res* 32:438–442
- Küntzer J, Backes C, Blum T, Gerasch A, Kaufmann M, Kohlbacher O, Lenhof HP (2007) BNDB—the biochemical network database. *BMC Bioinf* 8:367.1–9
- Lee SY, Lee DY, Hong SH, Kim TY, Yun H, Oh YO, Park S (2003) MetaFluxNet, a program package for metabolic pathway construction and analysis, and its use in large-scale metabolic flux analysis of *escherichia coli*. *Genome Inform* 14:23–33
- Le Novère N, Hucka M, Mi H, Moodie S, Schreiber F, Sorokin A, Demir E, Wegner K, Aladjem M, Wimalaratne SM, Bergman FT, Gauges R, Ghazal P, Kawaji H, Li L, Matsuoka Y, Villéger A, Boyd SE, Calzone L, Courtot M, Dogrusoz U, Freeman T, Funahashi A, Ghosh S, Jouraku A, Kim S, Kolpakov F, Luna A, Sahle S, Schmidt E, Watterson S, Wu G, Goryanin I, Kell DB, Sander C, Sauro H, Snoep JL, Kohn K, Kitano H (2009) The systems biology graphical notation. *Nat Biotechnol* 27:735–741
- Leticic I, Yamada T, Kanehisa M, Bork P (2008) iPath: interactive exploration of biochemical pathways and networks. *Trends Biochem Sci* 33:101–103
- Lombardot T, Morgat A, Axelsen KB, Aimo L, Hyka-Nouspikel N, Niknejad A, Ignatchenko A, Xenarios I, Coudert E, Redaschi N, Bridge A (2018) Updates in Rhea: SPARQLing biochemical reaction data. *Nucleic Acids Res* 47(1):D596–D600
- Maes A, Martinez X, Druart K, Laurent B, Guégan S, Marchand CH, Lemaire SD, Baaden M (2018) MinOmics, an integrative and immersive tool for multi-omics analysis. *J Integr Bioinf* 15(2):20180006
- Malik-Sheriff RS, Glont M, Nguyen TVN, Tiwari K, Roberts MG, Xavier A, Vu MT, Men J, Maire M, Kananathan S, Fairbanks EL, Meyer JP, Arankalle C, Varusai TM, Knight-Schrijver V, Li L, Dueñas-Roca C, Dass G, Keating SM, Park YM, Buso N, Rodríguez N, Hucka M, Hermjakob H (2019) BioModels—15 years of sharing computational models in life science. *Nucleic Acids Res* 48(1):D407–D415
- Martens M, Ammar A, Riutta A, Waagmeester A, Slenter DN, Hanspers K, Millerand RA, Digles D, Lopes EN, Ehrhart F, Dupuis LJ, Winckers LA, Coort SL, Willighagen EL, Evelo CT, Pico AR, Kutmon M (2021) WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res* 49(1):D613–D621
- Matthews L, Gopinath G, Gillespie M, Caudy M, Croft D, de Bono B, Garapati P, Hemish J, Hermjakob H, Jassal B, Kanapin A, Lewis S, Mahajan S, May B, Schmidt E, Vastrik I, Wu G, Birney E, Stein L, D’Eustachio P (2009) Reactome knowledge-base of human biological pathways and processes. *Nucleic Acids Res* 37(1):D619–622
- Mendes P (2000) Advanced visualization of metabolic pathways in PathDB. In: Proceedings of the conference on plant and animal genome
- Mendes, P, Bulmore DL, Farmer AD, Steadman PA, Waugh ME, Wlodek ST (2000) PathDB: a second generation metabolic database. In: Hofmeyr JHS, Rohwer JM, Snoep JL (eds) Proceedings of the International BioThermoKinetics Meeting. Stellenbosch University Press, Stellenbosch, pp 207–212
- Mi H, Thomas P (2009) PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol Biol* 563:123–140
- Mi H, Schreiber F, Moodie SL, Czauderna T, Demir E, Haw R, Luna A, Le Novère N, Sorokin AA, Villéger A (2015) Systems biology graphical notation: activity flow language level 1 version 1.2. *J Integr Bioinf* 12(2):265
- Mi H, Ebert D, Muruganujan A, Mills C, Albu LP, Mushayamaha T, Thomas PD (2020) PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res* 49(1):D394–D403
- Michal G (1968) Biochemical pathways (Poster). Boehringer Mannheim (latest version 2005)
- Michal, G (1998) On representation of metabolic pathways. *BioSystems* 47:1–7
- Misue K, Eades P, Lai W, Sugiyama K (1995) Layout adjustment and the mental map. *J Vis Lang Comput* 6:183–210

- Naithani S, Preece J, D'Eustachio P, Gupta P, Amarasinghe V, Dharmawardhana PD, Wu G, Fabregat A, Elser JL, Weiser J, Keays M, Fuentes AMP, Petryszak R, Stein LD, Ware D, Jaiswal P (2016) Plant Reactome: a resource for plant pathways and comparative analysis. *Nucleic Acids Res* 45(1):D1029–D1039
- Naithani S, Gupta P, Preece J, D'Eustachio P, Elser JL, Garg P, Dikeman DA, Kiff J, Cook J, Olson A, Wei S, Tello-Ruiz MK, Mundo AF, Munoz-Pomer A, Mohammed S, Cheng T, Bolton E, Papatheodorou I, Stein L, Ware D, Jaiswal P (2019) Plant Reactome: a knowledgebase and resource for comparative pathway analysis. *Nucleic Acids Res* 48(1):D1093–D1103
- O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open Babel: an open chemical toolbox. *J Cheminf* 3:33
- Purchase HC, Samra A (2008) Extremes are better: Investigating mental map preservation in dynamic graphs. In: *Diagrammatic representation and inference*. LNCS, vol 5223. Springer, Heidelberg, pp. 60–73
- Purchase HC, Hoggan E, Görg C (2007) How important is the “Mental Map”?—an empirical investigation of a dynamic graph layout algorithm. In: *Graph Drawing*, LNCS, vol 4372. Springer, Heidelberg, pp. 184–195
- Qeli E, Wiechert W, Freisleben B (2004) 3D visualization and animation of metabolic networks. In: *Proceedings of the 18th European simulation multicongress*, pp 1–4
- Rocha I, Maia P, Evangelista P, Vilaca P, Soares S, Pinto JP, Nielsen J, Patil KR, Ferreira EC, Rocha M (2010) OptFlux: an open-source software platform for in silico metabolic engineering. *BMC Syst Biol* 4:45
- Rodchenkov I, Babur O, Luna A, Aksoy BA, Wong JV, Fong D, Franz M, Siper MC, Cheung M, Wrana M, Mistry H, Mosier L, Dlin J, Wen Q, O'Callaghan C, Li W, Elder G, Smith PT, Dallago C, Cerami E, Gross B, Dogrusoz U, Demir E, Bader GD, Sander C (2019) Pathway commons 2019 update: integration, analysis and exploration of pathway data. *Nucleic Acids Res* 48(1):D489–D497
- Rohn H, Klukas C, Schreiber F (2011) Creating views on integrated multidomain data. *Bioinformatics* 27(13):1839–1845
- Rohn H, Hartmann A, Junker A, Junker BH, Schreiber F (2012) Fluxmap: a Vanted add-on for the visual exploration of flux distributions in biological networks. *BMC Syst Biol* 6:33.1–9
- Rohn H, Junker A, Hartmann A, Grafahrend-Belau E, Treutler H, Klapperstück M, Czauderna T, Klukas C, Schreiber F (2012) Vanted v2: a framework for systems biology applications. *BMC Syst Biol* 6(139):1–13
- Rojdestvenski I (2003) Metabolic pathways in three dimensions. *Bioinformatics* 19(18):2436–2441
- Rojdestvenski I, Cottam M (2002) Visualizing metabolic networks in VRML. In: *Proceedings of the international conference on information visualisation (IV'02)*, pp 175–180
- Rolletschek H, Melkus G, Grafahrend-Belau E, Fuchs J, Heinzel N, Schreiber F, Jakob PM, Borisjuk L (2011) Combined non-invasive imaging and modeling approaches reveal metabolic compartmentation in the barley endosperm. *Plant Cell* 23(8):3041–3054
- Rougnay A, Toure V, Moodie S, Balaur I, Czauderna T, Borlinghaus H, Dogrusoz U, Mazein A, Dräger A, Blinov ML, Villeger A, Haw R, Demir E, Mi H, Sorokin A, Schreiber F, Luna A (2019) Systems biology graphical notation: Process description language level 1 version 2. *J Integr Bioinf* 16(3):22
- Rücker C, Meringer M (2002) How many organic compounds are graph-theoretically nonplanar? *MATCH Commun Math Comput Chem* 45:153–172
- Scheer M, Grote A, Chang A, Schomburg I, Munaretto C, Rother M, Söhngen C, Stelzer M, Thiele J, Schomburg D (2011) BRENDA, the enzyme information system in 2011. *Nucleic Acids Res* 39(1):D670–D676
- Schläpfer P, Zhang P, Wang C, Kim T, Banf M, Chae L, Dreher K, Chavali AK, Nilo-Poyanco R, Bernard T, Kahn D, Rhee SY (2017) Genome-wide prediction of metabolic enzymes, pathways, and gene clusters in plants. *Plant Physiol* 173(4):2041–2059
- Schreiber F (2002) High quality visualization of biochemical pathways in BioPath. In *Silico Biol* 2(2):59–73

- Schreiber F, Dwyer T, Marriott K, Wybrow M (2009) A generic algorithm for layout of biological networks. *BMC Bioinf* 10:375
- Schreiber F, Colmsee C, Czauderna T, Grafahrend-Belau E, Hartmann A, Junker A, Junker BH, Klapperstück M, Scholz U, Weise S (2012) MetaCrop 2.0: managing and exploring information about crop plant metabolism. *Nucleic Acids Res* 40(1):D1173–D1177
- Schreiber F, Kerren A, Börner K, Hagen H, Zeckzer D (2014) Heterogeneous networks on multiple levels. In: *Multivariate network visualization*. Springer, Berlin, pp 175–206
- Schreiber F, Sommer B, Czauderna T, Golebiewski M, Gorochowski TE, Hucka M, Keating SM, König M, Myers C, Nickerson D, Waltemath D (2020) Specifications of standards in systems and synthetic biology: status and developments in 2020. *J Integr Bioinf* 17(2–3):22
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–2504
- Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27(3):431–432
- Sommer B, Schreiber F (2017a) Integration and virtual reality exploration of biomedical data with CmPI and VANTED. *Inf Technol* 59(4):181–190
- Sommer B, Schreiber F (2017b) Integration and virtual reality exploration of biomedical data with CmPI and Vanted. *Inf Technol* 59(4):181–190
- Sommer B, Künsemöller J, Sand N, Husemann A, Rummig M, Kormeier B (2010) CELLmi-crocosmos 4.1—an interactive approach to integrating spatially localized metabolic networks into a virtual 3D cell environment. In: Fred ALN, Filipe J, Gamboa H (eds) *Bioinformatics 2010—Proceedings of the International Conference on Bioinformatics*, pp 90–95
- Sorokin AA, Le Novère N, Luna A, Czauderna T, Demir E, Haw R, Mi H, Moodie SL, Schreiber F, Villéger A (2015) Systems biology graphical notation: Entity relationship language level 1 version 2. *J Integr Bioinf* 12(2):264
- Steinbeck C, Han Y, Kuhn S, Horlacher O, Luttmann E, Willighagen EL (2003) The Chemistry Development Kit (CDK): an Open-Source Java Library for Chemo- and Bioinformatics. *J Chem Inf Comput Sci* 43(2):493–500
- Stryer L (1988) *Biochemistry*. W H Freeman, New York
- Suderman, M, Hallett M (2007) Tools for visually exploring biological networks. *Bioinformatics* 23(20):2651–2659
- Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J* 37:914–939
- Thomas JJ, Cook KA (2006) A visual analytics agenda. *IEEE Comput Graph Appl* 26(1):10–13
- Usadel B, Nagel A, Thimm O, Redestig H, Blaesing OE, Palacios-Rojas N, Selbig J, Hannemann J, Piques MC, Steinhauser D, Scheible WR, Gibon Y, Morcuende R, Weicht D, Meyer S, Stitt M (2005) Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiol* 138:1195–1204
- van Iersel MP, Kelder T, Pico AR, Hanspers K, Coort S, Conklin BR, Evelo C (2008) Presenting and exploring biological pathways with PathVisio. *BMC Bioinf* 9:399.1–399.9
- van Iersel M, Pico A, Kelder T, Gao J, Ho I, Hanspers K, Conklin B, Evelo C (2010) The BridgeDb framework: standardized access to gene, protein and metabolite identifier mapping services. *BMC Bioinf* 11:5
- Van Iersel MP, Villéger AC, Czauderna T, Boyd SE, Bergmann FT, Luna A, Demir E, Sorokin A, Dogrusoz U, Matsuoka Y, Funahashi A, Aladjem M, Mi H, Moodie S, Kitano H, Le Novère N, Schreiber F (2012) Software support for SBGN maps: SBGN-ML and LibSBGN. *Bioinformatics* 28(15):2016–2021
- Waltemath D, Karr J, Bergmann F, Chelliah V, Hucka M, Krantz M, Liebermeister W, Mendes P, Myers C, Pir P, Alaybeyoglu B, Aranganathan N, Baghalian K, Bittig A, Burke P, Cantarelli M, Chew YH, Costa R, Cursons J, Schreiber F (2016) Toward community standards and software for whole-cell modeling. *IEEE Trans Biomed Eng* 63(10):2007–2014

- Waltemath D, Golebiewski M, Blinov ML, Gleeson P, Hermjakob H, Hucka M, Inau ET, Keating SM, König M, Krebs O, Malik-Sheriff RS, Nickerson D, Oberortner E, Sauro HM, Schreiber F, Smith L, Stefan MI, Wittig U, Myers CJ (2020) The first 10 years of the international coordination network for standards in systems and synthetic biology (COMBINE). *J Integr Bioinf* 17(2–3):20200005
- Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Bryant SH (2009) PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res* 37(7):W623–W633
- Weininger D (1988) SMILES, a chemical language and information system 1. Introduction to methodology and encoding rules. *J Chem Inf Model* 28:31
- Weise S, Grosse I, Klukas C, Koschützki D, Scholz U, Schreiber F, Junker BH (2006) Meta-All: a system for managing metabolic pathway information. *BMC Bioinf* 7:e465
- Wiechert W (2001) <sup>13</sup>C metabolic flux analysis. *Metab Eng* 3(3):195–206
- Wittig U, Kania R, Golebiewski M, Rey M, Shi L, Jong L, Algae E, Weidemann A, Sauer-Danzwith H, Mir S, Krebs O, Bittkowski M, Wetsch E, Rojas I, Müller W (2012) SABIO-RK—database for biochemical reaction kinetics. *Nucleic Acids Res* 40(1):D790–D796
- Wittig U, Rey M, Weidemann A, Kania R, Müller W (2017) SABIO-RK: an updated resource for manually curated biochemical reaction kinetics. *Nucleic Acids Res* 46(1):D656–D660
- Wrzodek C, Dräger A, Zell A (2011) KEGGtranslator: visualizing and converting the KEGG PATHWAY database to various formats. *Bioinformatics* 27(16):2314–2315
- Zhang JD, Wiemann S (2009) KEGGgraph: a graph approach to KEGG PATHWAY in R and bioconductor. *Bioinformatics* 25:1470–1471
- Zhang P, Dreher K, Karthikeyan A, Chi A, Pujar A, Caspi R, Karp P, Kirkup V, Latendresse M, Lee C, Mueller LA, Muller R, Rhee SY (2010) Creation of a genome-wide metabolic pathway database for *populus trichocarpa* using a new approach for reconstruction and curation of metabolic pathways for plants. *Plant Physiol* 153(4):1479–1491