Extension of the IsaViz Software for the Representation of Metabolic and Regulatory Networks

Diogo Fernando Veiga[1,2*]**, Pedro de Stege Cecconello**[2]**, José Eduardo de Lucca**[2] **and Luismar Marques Porto**[1]

[1]*Grupo de Engenharia Genômica; Departamento de Engenharia Química e Engenharia de Alimentos;* [2]*Departamento de Informática e Estatística; Universidade Federal de Santa Catarina; C. P. 476; 88040-900; Florianópolis - SC - Brasil*

ABSTRACT

In this work we developed an extension of IsaViz software, a RDF (Resource Description Framework) authoring tool, designed to be a graphical environment to build models of metabolic and regulatory networks. This environment, called Metabolic IsaViz, was linked to a genomic library of types and was modeled on the basis of ontologies. Biochemical pathways included data at sequence level (e.g., the amino acid sequence of enzymes), besides kinetic and thermodynamic parameters for the reactions. Models created with Metabolic IsaViz could be exported to pathways simulators through SBML (Systems Biology Markup Language), which allowed to analyze the pathway dynamics of target chemicals.

Key words: Metabolic and Regulatory Pathways, RDF/XML, Modeling

INTRODUCTION

Knowledge of the structure of metabolic pathways and their associated regulatory mechanisms may be used to design new process features, based on computational simulations that could improve the production of a commercial metabolite used in the food industry or as a drug precursor. A graphical software concerned with the modeling of dynamics and structure of metabolic pathways is a valuable tool in order to the analysis and comprehension of genomic data, that is being produced in large rates as the improvement of sequencing techniques.

In this work we combined a data representation technology known as Resource Description Framework, RDF (Klyne andCarrol, 2003; Manola andMiller, 2003), with the representation of metabolic pathways, both based on directed graphs. Logical and graphical extensions were implemented on a RDF authoring environment called IsaViz, developed by Pietriga (2003).

The Resource Description Framework (RDF)
RDF was created with the goal to become a standard language for the representation of data available on the Web. The RDF technology suggests a common representation for the web content, which would allow the integration of databases and the analysis of information by software agents. A web resource is conceived as anything that is either accessible by a browser, such as a document or a figure, or something that is not directly accessible but may be identified by a key called URI (Uniform Resource Identifier), for example an employee of a industry, a particular

* Author for correspondence

product in a storehouse, a chemical, or a metabolic pathway of interest.

A RDF model is a direct graph. The description of the resources is made by *Statements* or assertions, which graphically denote the link between two nodes of the graph. In RDF terminology, the node from which the edge leaves is called *Subject* while the target node is called *Object*. The edge itself is named *Predicate*. All the elements are labeled by their respective URIs. Figure 1 shows an example of a RDF graph with only one assertion. In this case, the resource gen:Glycolysis has the property rdf:type whose value denotes a metabolic pathway (gen:Metabolic_Pathway).

Figure 1 - A simple model in RDF.

Activators	ORI
CDS	PrimaryPolypeptide
Centromere	PrimaryTranscript
Chemicals	Products_List
Chemical_Alternate	Protein
Chromosome	Reactants_List
ChromosomeFrag	Reaction
Enzimatic_Data	Region
Gene	RegulatorySeq
Genome	rRNA
Inhibitors	Sequence
Intron	SequenceType
Kcat_List	snRNA
Kinetic_Data	SplicedTranscript
Km_List	SplicedTranscriptComponent
Metabolic_Pathway	Telomere
mRNA	Terminator
NTranscribedRegion	Thermodynamic_Data
ORF	TranscribedRegion

Box 1 - Classes of *RDFgenLib* used for model building with Metabolic IsaViz.

IsaViz: An authoring environment in RDF

IsaViz is a W3C (World Wide Web Consortium) recommended tool for RDF model authoring. The tool provides a visual environment for edition and navigation of RDF models, represented by directed graphs. IsaViz is RDF specifications compliant and offers a great number of resources such as graphics zooming, application of GSS style sheets (Graph Stylesheets) to the models, as well as RDF exporting to many formats, such as SVG (Scalable Vector Graphics), PNG (Portable Network

Graphics) and RDF/XML. The software is built in Java and its source code is freely available.

METHODOLOGY

We extended the logical and graphical functions of the original system in order to create the Metabolic IsaViz, a RDF tool that was able to model metabolic and regulatory pathways. In this section we describe how the available technologies were considered to modify the original IsaViz. The first modification made in IsaViz was its integration to a type library used to model metabolic pathways (Veiga and Porto, 2003). The genomic library *RDFgenLib* (Box 1), from which the user chooses the classes to build his own model, is written in RDFS (RDF Schema). The library classes were created to reflect sequencing terminology (intron, exon, UTR, etc) as well as the terminology involved in the modeling of metabolic routes (products, chemical, activators, and so on). The classes and properties from the library are used in Metabolic IsaViz when resources and properties

are inserted in the RDF model. When a resource is created, the class to which it belongs is provided. For example, *Glycolysis* belongs to gen:Metabolic_Pathway. On the other hand, when creating a property, the library helps modeling while showing the properties that may be inserted for a given resource. For a resource of the class gen:Activator (an activator in an enzymatic reaction), the library suggests the insertion of properties such as activation mechanism, description and structural formula.

Including a Resource

When inserting a resource, the user should provide, besides the URI, the class it belongs to, among those given by *RDFgenLib*. The software then associates it an icon, according to the class, and modifies the node layout of the graph (Fig. 2 and 3). The property rdf:type and the corresponding class resource still exist, but are then hidden from the user, and the icon therefore replaces the class of the resource.

Figure 2 - RDF graph in the original IsaViz. The property rdf:type refers to the class of the gen:Chorismate resource.
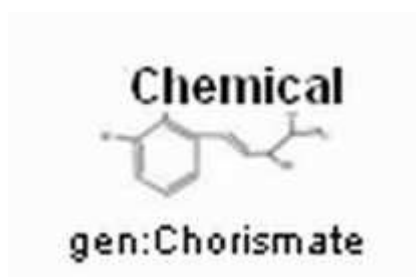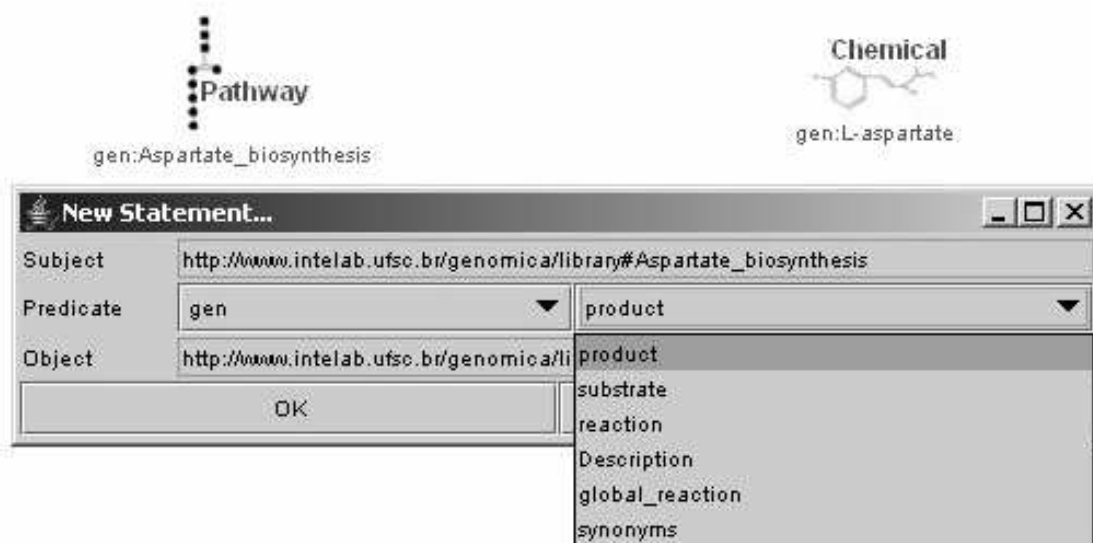

Figure 3 - RDF resource in Metabolic IsaViz. To improve visualization, the property rdf:type is hidden and an icon represents the class.

Including Properties

When a resource is selected, its properties are shown in the *Genomic Properties* tab of the *Definitions* window. By selecting an object of the class *Chemicals*, this tab shows its associated properties gen:inReaction, gen:Description, gen:MW gen:structuralFormula and the value of each one of them, *string* and *float*.

The properties that may be inserted for the selected resource, considering its class, are shown in the *New Statement* window. As can be seen from Figure 4, the resource gen:Aspartate_biosynthesis, which is a metabolic pathway, has the following suggested properties: *product*, *reaction*, *Description, global_reaction*, *synonymous*. To make up a *Statement* with the resource gen:L-aspartate the choice would be gen:product, which would produce the graph shown in Figure 5.

Figure 4 - Insertion of properties in Metabolic IsaViz. The properties of each resource are displayed to the user, according to what was defined in RDFgenLib.
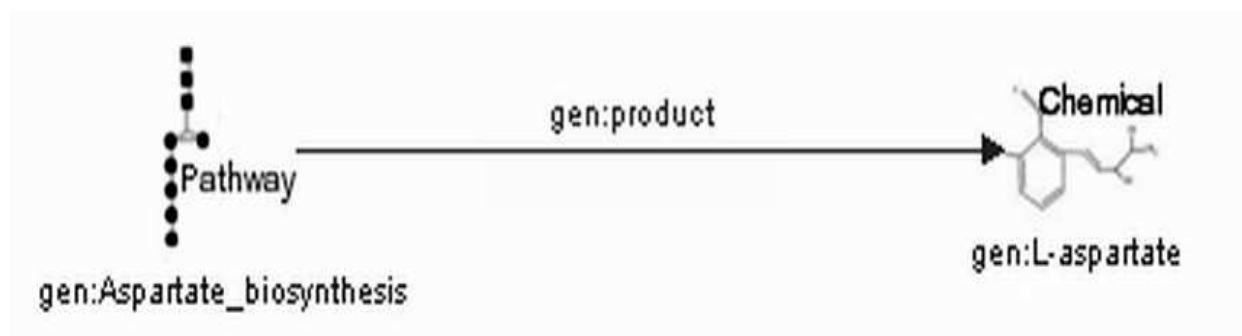
Figure 5 - Iconic results after insertion of the property gen:product linking the resources gen:Aspartate_biosynthesis and gen:L-aspartate.

RESULTS

With the modifications implemented in the IsaViz software, the extended suite Metabolic IsaViz is a new tool able to creating models of metabolic pathways, also including regulatory features and enzyme kinetics, as will be shown below. Modeling of metabolic pathways with Metabolic IsaViz is based on the creation of resources (nodes of a graph), providing the class they belong to, and assigning them properties, graphically represented as graph edges. When the resource class is provided, the system includes an icon in the visualization space that indicates the class just chosen. The modeling process assumes some knowledge of the *RDFgenLib* library and the metabolic routes of interest. Typical examples are distributed with the software to serve as starting point for the user.

Representing a reaction step in a pathway

Figure 6 shows the RDF model generated by Metabolic IsaViz to represent a metabolic pathway step, i.e., an enzymatic reaction. The major resource is gen:2.5.1.54, an object that represents the reaction. Two properties gen:left leave this resource to represent the substrates gen:D-Erythrose-4-phosphate (E4P) and gen:Phosphoenol-pyruvate (PEP). The gen:right property indicates the reaction product, the resource gen:7P-2-Dehydro-3-deoxy-D-arabino-heptonate. Those constitute the minimal information necessary to build a reaction step. The model depicted in Figure 6 also shows how to represent intersections among pathways. The resource PEP is the final product of the glycolysis pathway what is indicated by the property gen:product that comes from gen:glycolysis.

A reaction with enzymatic and kinetic parameters

Figures 7 and 8 present a reaction model that incorporates functional or kinetic data as well as data on the enzyme and its ligands, respectively. Figure 7 shows how ligands participating in a reaction are represented. The properties gen:inhibitor_elem are used to denote the inhibitors of the reaction; the properties gen:activators_elem, gen:prosthetic_group and gen:reaction_direction provide more detailed knowledge of the biochemical system under consideration.
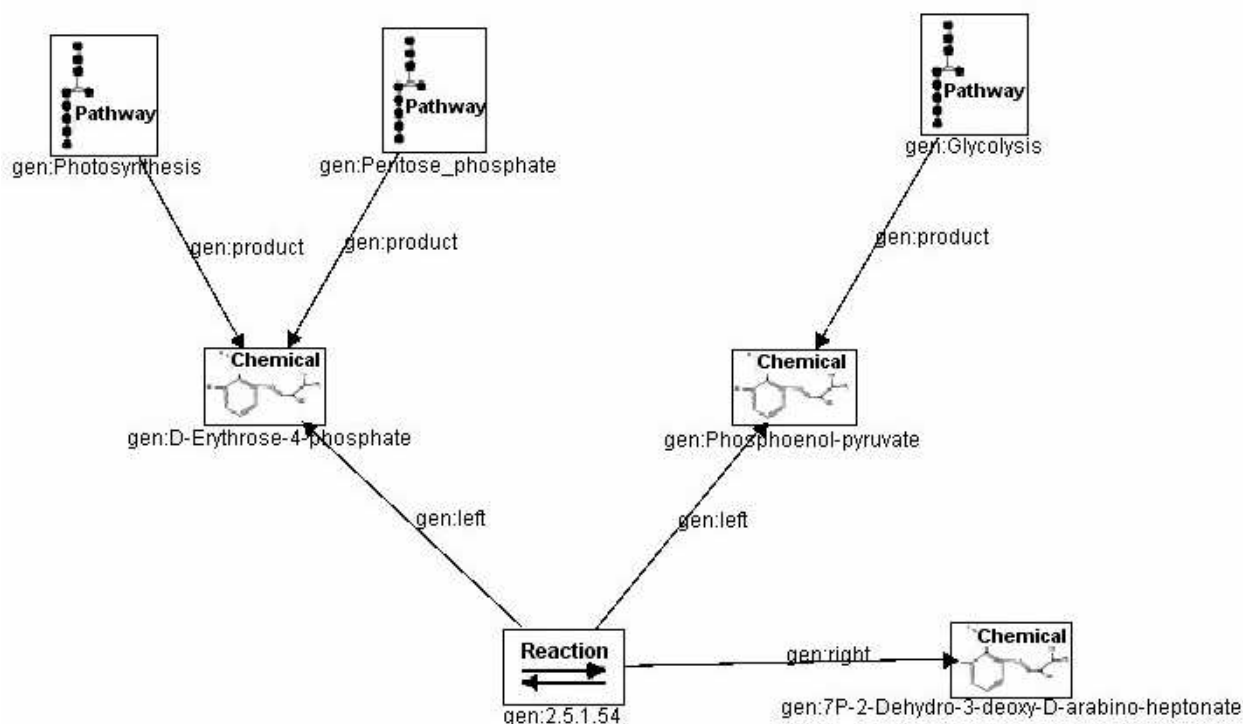
Figure 6 - Representation of an enzymatic reaction in Metabolic IsaViz.
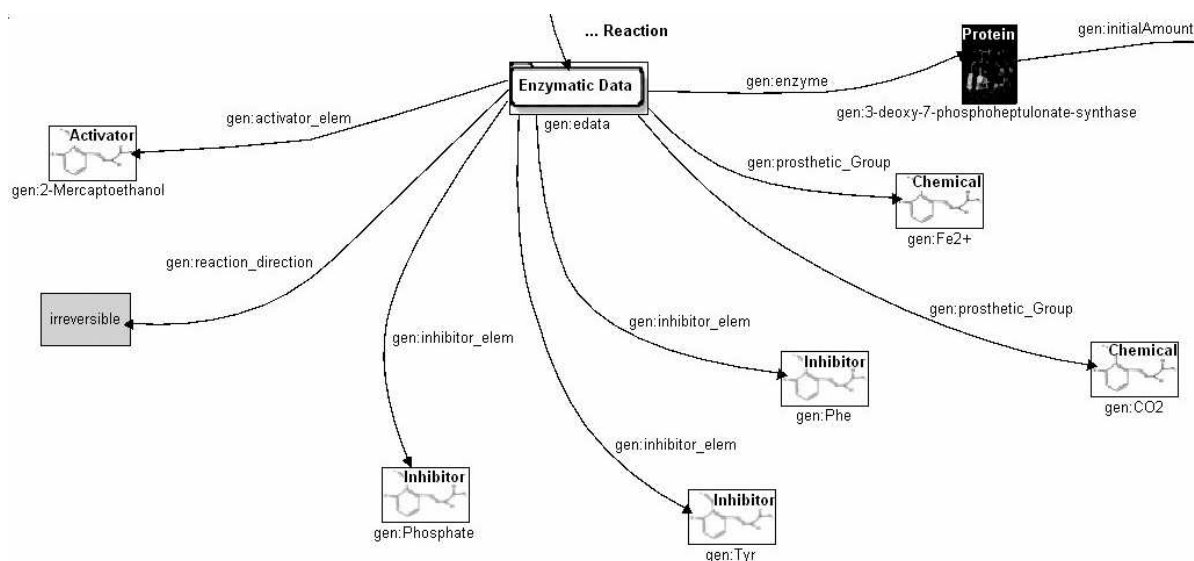
Figure 7 - Representation of activators, inhibitors, and prosthetic groups that are part of the reaction 2.5.1.54, and that are included using the resource gen:edata, of the class Enzymatic_Data.

To add parameters for the enzyme kinetics one may use a resource from the class *Kinetic_Data* (Figure 8). The gen:functional_param resource has the following properties: gen:pH_optimum, gen:Turnover_number, gen:Specific_activity gen:Ki and gen:Km. Those parameters may be used in the kinetic functional dependence (rate equation) such as a Michaelis-Menten kinetic model, and may be added to the model for dynamic simulation purposes.

Representing transcriptional regulation and inhibition by product
The *RDFgenLib* also provides classes to represent some regulatory levels. Figure 9 shows a model for reaction gen:2.3.1.46 involving some of these features. Regulation by product (feedback regulation), which occurs among intermediary products of the same pathway or even of different pathways, is represented in the graph by the resource gen:reg. The property gen:inhibited_By, that leaves that resource, specifies which compound regulates the production of gen:alpha-succinyl-L-homoserine and gen:HSCoA. The class *Regulation* has also the property gen:activated_By. The other regulatory mechanism shown refers to the transcription repression of the gene *metA*, done by a compound known as holorepressor. For that kind of regulation we may use the resource of the class gen:Expression combined with the property gen:repression_By, being also suitable the properties gen:activated_By and gen:inhibited_By.
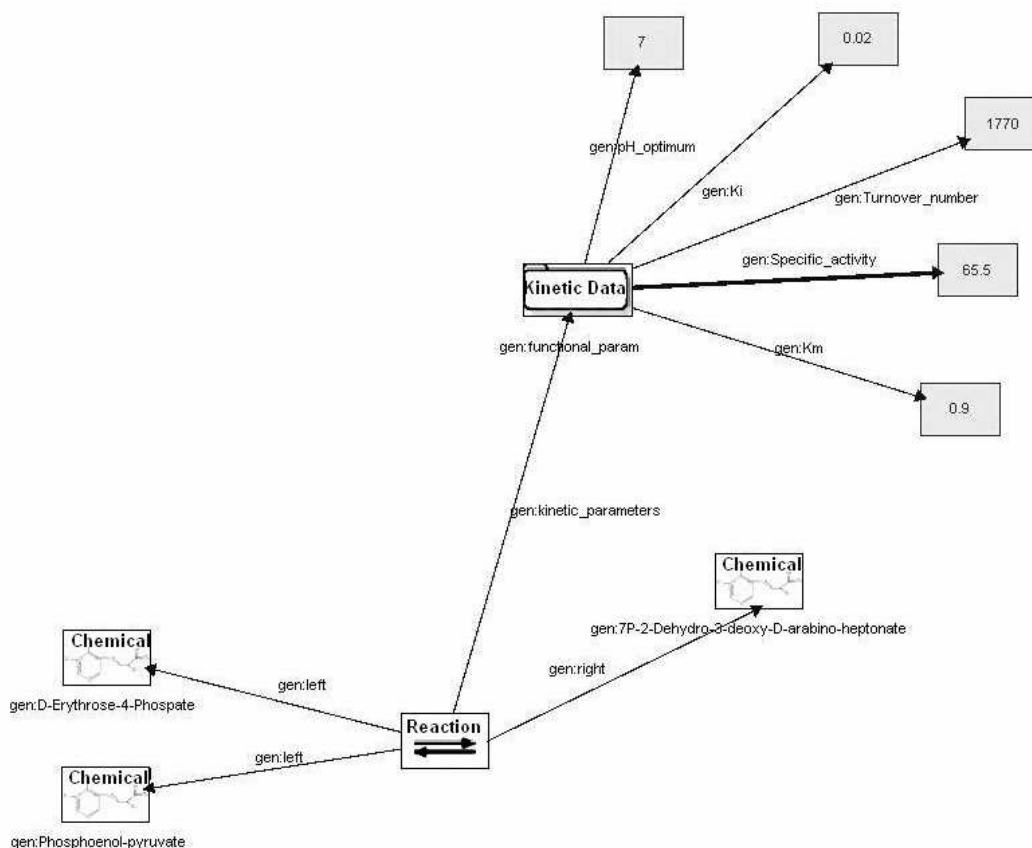
Figure 8 - Kinetic data may be modeled with a resource of the class Kinetic_Data. In the example shown, optimal pH, rate constant (Km), and inhibition (Ki) parameters are informed (See the properties that leave the resource gen:functional_param, class Kinetic_Data).

DISCUSSION

The models generated by Metabolic IsaViz were able to express enzymatic kinetic data, regulatory mechanisms, and interactions between enzymes and their ligands. The exportation of the models created by Metabolic IsaViz to SBML format allowed the integration with a variety of biochemical pathways simulators based on sets of ordinary differential.

Equations or based on Hybrid Petri Nets (HPN). Metabolic IsaViz is being used by our group as the basis of a genome information system called GEnIS (*Genome Engineering Information System*). GEnIS is a software suite that takes advantage of the Metabolic IsaViz representation to simulate biochemical networks, and is intended to rationalize the acquisition and organization of experimental data in order to better guide wet laboratory work.
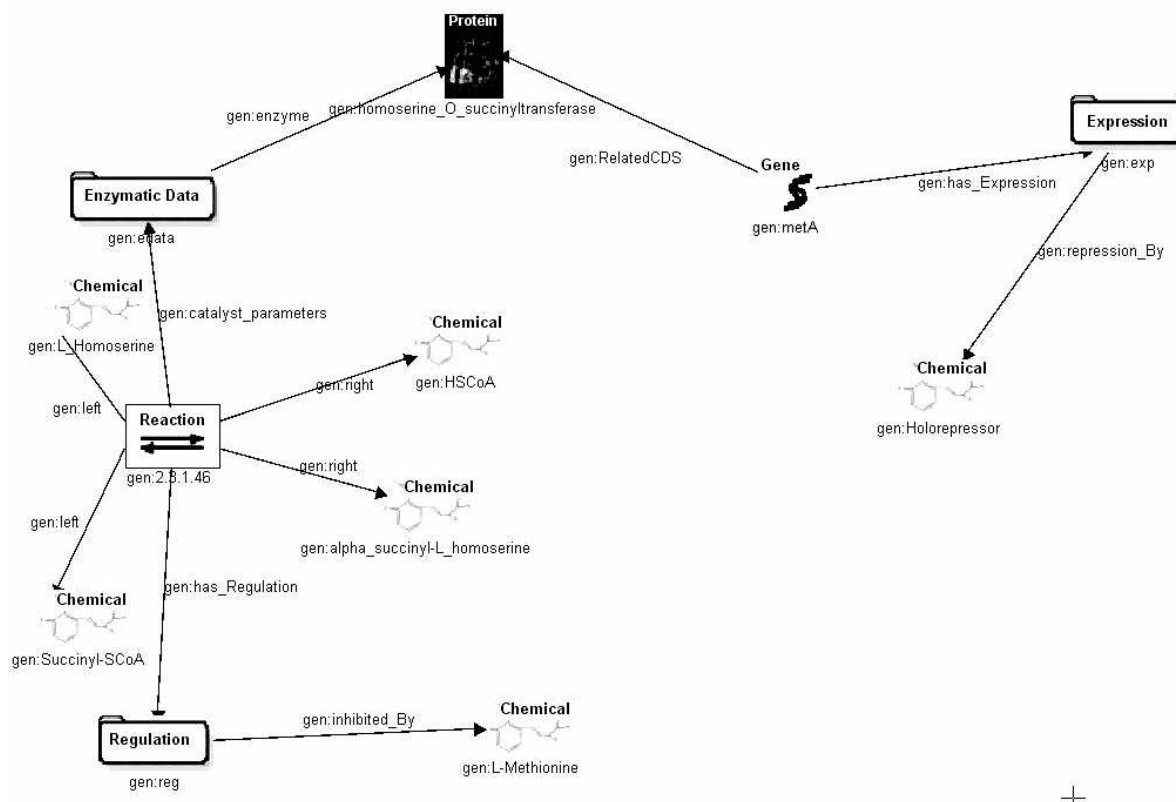
Figure 9 - Representation of regulatory action at the transcription level and among intermediary metabolites.

ACKNOWLEDGEMENTS

RESUMO

A determinação de vias metabólicas e regulatórias de microrganismos é essencial para estudos pós-sequenciamento de DNA, com aplicações diretas em várias áreas da biotecnologia, em especial em engenharia metabólica. Neste trabalho desenvolvemos uma extensão do software IsaViz, editor de grafos RDF (Resource Description Framework), com a finalidade de criar um ambiente gráfico para a construção de modelos de vias metabólicas e regulatórias. Este ambiente, o Metabolic IsaViz, foi integrado a uma biblioteca de tipos genômicos, modelada com base em ontologias, sendo que as vias bioquímicas podem incluir dados ao nível de seqüência (como a seqüência de aminoácidos das enzimas), além de parâmetros cinéticos e termodinâmicos. Os modelos criados com o Metabolic IsaViz podem ser exportados para simuladores de vias metabólicas através da linguagem SBML (Systems Biology Markup Language) para análise da formação de metabólitos de interesse.

REFERENCES

Klyne, G. and Carroll, J. J. (2003), Resource Description Framework (RDF): Concepts and Abstract Syntax. [S.l.]. *World Wide Web Consortium.* Availabe at: http://www.w3.org/TR/rdf-concepts/. Access on: 14 feb. 2003.

Manola, F. and Miller, E. (2003), RDF Primer. [S.l.]. *World Wide Web Consortium.* Available at: http://www.w3.org/TR/rdf-primer/. Access on: 14 feb. 2003.

Pietriga, E. (2003), Software. IsaViz: A Visual Authoring Tool for RDF. IsaViz 2.0 alpha - preview release (2003-05-12). Available at: http://www.w3c.org. Access on: 28 jun. 2003.

Veiga, D. F. and Porto, L. M. (2003), XML Schema Representation as a Way to Interchange and Customize Genomic and Metabolic Models. In: International Conference on Bioinformatics and Computational Biology, Ribeirão Preto, SP, Brazil.