



Project no. : **200767**  
Project acronym: **APO-SYS**  
Project title: **Apoptosis systems biology applied to cancer and AIDS.**

Instrument: Collaborative project

Thematic Priority: Large-scale integrating project

### **Deliverable reference number and title**

**DL12: BIOPAX converters for GINsim and ConsensusPathDB**

Due date of deliverable: **31/01/2009**  
Actual submission date: **29/01/2009**

Start date of project:  
**1 of February 2008**

Duration:  
**4 years**

Organization name of lead contractor for this deliverable:  
**Tel Aviv University (Beneficiary no 14)**

<b>Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)</b>		
<b>Dissemination Level</b>		
<b>PU</b>	Public	PU
<b>PP</b>	Restricted to other programme participants (including the Commission Services)	
<b>RE</b>	Restricted to a group specified by the consortium (including the Commission Services)	
<b>CO</b>	Confidential, only for members of the consortium (including the Commission Services)	

## Description of DL12 with completion date “Month 12”

**Editor:** Denis Thieffy (Université de la Méditerranée-Partner 20).  
Emmanuel Barillot (Institut Curie- Partner 3)  
Ralf Herwig (Max-Planck Gesellschaft zur Förderung der Wissenschaften - Partner 15A)

## BIOPAX converters for *GINsim* and ConsensusPathDB

### Introduction

*GINsim* (Gene Interaction Network simulation) is a computer software for the qualitative (multilevel logical) modeling and simulation of genetic regulatory networks, developed at UnivMed [Gonzalez *et al*, 2006 ; Naldi *et al*, 2007].

*BiNoM* is a *Cytoscape* plugin, developed at IC to facilitate the manipulation of biological networks represented in standard systems biology formats (*SBML*, *SBGN*, *BioPAX*) and to carry out studies on the network structure [Zinovyev *et al*, 2008].

Developed and maintained at the MPG, *ConsensusPathDB* is a database system for the integration of human functional interactions. *ConsensusPathDB* currently integrates the content of 12 different interaction databases with heterogeneous foci comprising a total of 26,133 distinct physical entities and 74,289 distinct functional interactions (protein–protein interactions, biochemical reactions, gene regulatory interactions), and covering 1738 pathways. The *ConsensusPathDB* web interface enables users to search and visualize interaction networks, upload, modify and expand networks, or carry out over-representation analyses with uploaded identifier lists with respect to substructures derived from the integrated interaction network [Kamburov *et al*, 2009]

The Deliverable 12 of *APO-SYS* project aims at easing the exchanges of networks between these different tools and take advantage of their complementary functionalities. Originally, *BioPAX* was selected as a target exchanged format in this respect. Since the beginning of *APO-SYS* project, interactions between the partners involved have led to a broader and more generic approach to the problem of network and model exchanges, yet keeping *BioPAX* as one exchange format between our tools. In brief, *BioPAX* support has been achieved by computational developments enabling:

- 1) the import, export and handling of interaction networks encoded in *BioPAX*, as well as *SBML* (only import) and *PSI-MI* (only import) formats by *ConsensusPathDB* (MPG);
- 2) the compatibility of *ConsensusPathDB* and *Reactome* using *BioPAX* format;
- 3) the export and import of *BioPAX* files (among other formats) from *Cytoscape*, through *BiNoM*, conceived as a plugin of *Cytoscape* (IC);
- 4) the conversion of regulatory graphs encoded in *GINsim* into *Cytoscape* format, which can thus be further converted into *BioPax* using *Cytoscape/BiNoM* (UnivMed);
- 5) the conversion of *Cytoscape* files (and thus also of *BioPax* files via *BiNoM*) into *GINsim* format (*GINML*), through the use of a novel *Cytoscape* plugin (*Cytoscape2GINML*) (UnivMed).

This solution currently enables various types of exchanges between *ConsensusPathDB*, *Reactome*, *Cytoscape/BiNoM*, and *GINsim*, as these tools each already support different network/model formats (see Table below). As supports for other network/model formats are

currently being implemented in these different tools, exchanges with other modeling software will soon be made available.

**Table 1.** Different formats supported by *ConsensusPathDB*, *Cytoscape/BiNoM* and *GINsim*.

<b>Software</b>	<b>Input format</b>	<b>Ouput format</b>
<b><i>ConsensusPathDB</i></b> http://cpdb.molgen.mpg.de	<i>BioPAX</i> <i>SBML</i> <i>PSI-MI</i>	<i>BioPAX</i> <i>ConsensusPathDB</i> model dump Graphical formats
<b><i>Cytoscape/BiNoM</i></b> http://bioinfo.curie.fr/projects/binom	<i>Cytoscape</i> <i>BioPAX</i> <i>SBML</i> <i>CellDesigner</i>	<i>Cytoscape</i> <i>BioPAX</i> <i>SBML</i>
<b><i>GINsim (version 2.3)</i></b> http://gin.univ-mrs.fr	<i>GINML</i>	<i>GINML</i> <i>PNML (Petri nets)</i> <i>INA (Petri nets)</i> <i>NuSMV (Model Checking)</i> Graphical formats
<b><i>Cytoscape2GINML</i></b> http://gin.univ-mrs.fr	<i>Cytoscape</i>	<i>GINML</i>

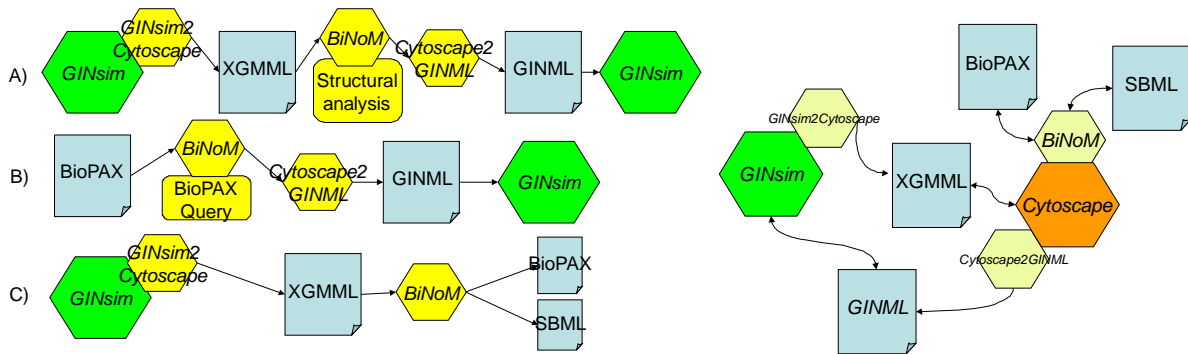
## 1. Case study with *ConsensusPathDB* and *Reactome*

Interaction data in *BioPAX* (level 2), *PSI-MI* (level 2.5) or *SBML* (level 2) format can be uploaded through the web-interface of *ConsensusPathDB*. If the entities (proteins, metabolites, etc.) from the uploaded dataset are sufficiently annotated with accession numbers (*UniProt*, *Ensembl*, *KEGG*, etc.), these entities and their interactions are mapped to the integrated interaction data in *ConsensusPathDB*. An interactive image of the uploaded dataset is displayed as an interaction network. For each interaction in this network, it is shown whether it is also present in *ConsensusPathDB* and if so, which source database it originates from.

Moreover, with the web interface the user can interactively construct interaction networks based on the interaction data in the database. Such networks can be downloaded in *BioPAX* level 2 format (a *BioPAX* level 3 export functionality will be available soon). The combination of import and export functionality of the web interface enables the user to upload an interaction dataset, validate it in means of interaction evidence, expand it with further interactions from *ConsensusPathDB*, and download the expanded dataset in *BioPAX* format. For example, Annex 1 graphically depicts three reactions of a glycolysis model in the *ConsensusPathDB* web server view. This model has been imported in *BioPAX* format. This format was exported from the *Reactome* database (see supplementary file *Reactome70171.owl*). Thus, the case study shows the compatibility of *ConsensusPathDB* and *Reactome*.

## 2. Case studies for the assessment of file conversions between *GINsim* and *Cytoscape/BiNoM*

Here, we document the ability of *Cytoscape/BiNoM* and *GINsim* to exchange structural network information through dedicated software extensions. We considered three different scenarios, presented in the Figure 1.



**Figure 1.** Scheme of the case-study to test the integration of *GINsim* and *Cytoscape/BiNoM* softwares. **Left panel:** Three scenarios combining *GINsim* and *Cytoscape* functionalities. **(A)** the use of *BiNoM* algorithms to analyse the structure of a regulatory graph defined with *GINsim*; **(B)** the delineation of a logical model using a *BioPAX* file obtained from a pathway database (e.g., *ConsensusPathDB*) ; **(C)** the export of a regulatory graph from *GINsim* into *BioPAX* and *SBML* formats through *BiNoM*. **Right Panel:** Flow-chart describing the different possible exchanges between *Cytoscape/BiNoM* and *GINsim*.

### Material for the case studies

#### A) Models

As a test case, we have first used the high-level cell fate decision model developed by the Systems Biology team at IC in collaboration with the group of Denis Thieffry at *UnivMed* (cf. APO-SYS deliverable 24). Designed with *GINsim*, this model is stored in the form of *GINML* file (master\_model200109.zginml).

We have further performed a conversion of the semantic layer of the *BioBase* pathway database ([www.biobase-international.com/files/BiobaseSept2007.xgmml](http://www.biobase-international.com/files/BiobaseSept2007.xgmml)) and *BiobaseSept2007.owl*, prepared at the IC).

#### B) Software

We have used the Windows version of the *Cytoscape* version 2.6.0, *GINsim* version 2.3\_2009-01-28, *BiNoM* version dated 23/01/2009, and *Cytoscape2GINML* plugin version 0.1. All the software is provided as a Supplementary zip-package for this report.

All tests have been performed by A Zinovyev at IC, in collaboration with D Thieffry at *UnivMed*.

## Results of the case studies

### A) Use of BiNoM algorithms to analyze GinSim network structure

**Step1.** The file `master_model200109.zginml` was opened by *GINsim* and exported to `test.xgmml` file.

**Step2.** The `test.xgmml` file was imported to *Cytoscape*. The result is presented on Figure 2.

Note 1: Some edges do not appear in the *Cytoscape* version of the network (namely, 'CASP3 (activate) CASP8' and 'ROS (activate) MPT'). This is probably a bug in *Cytoscape*, caused by identifying an edge with the same nodes and the same 'interaction' attribute value. One solution is to change the value of the 'interaction' attribute for these edges for 'activate1'.

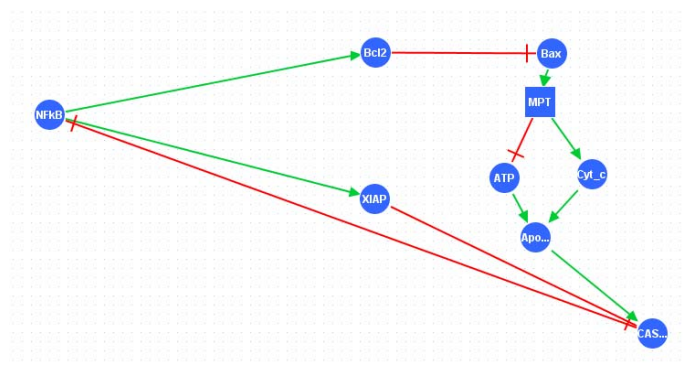
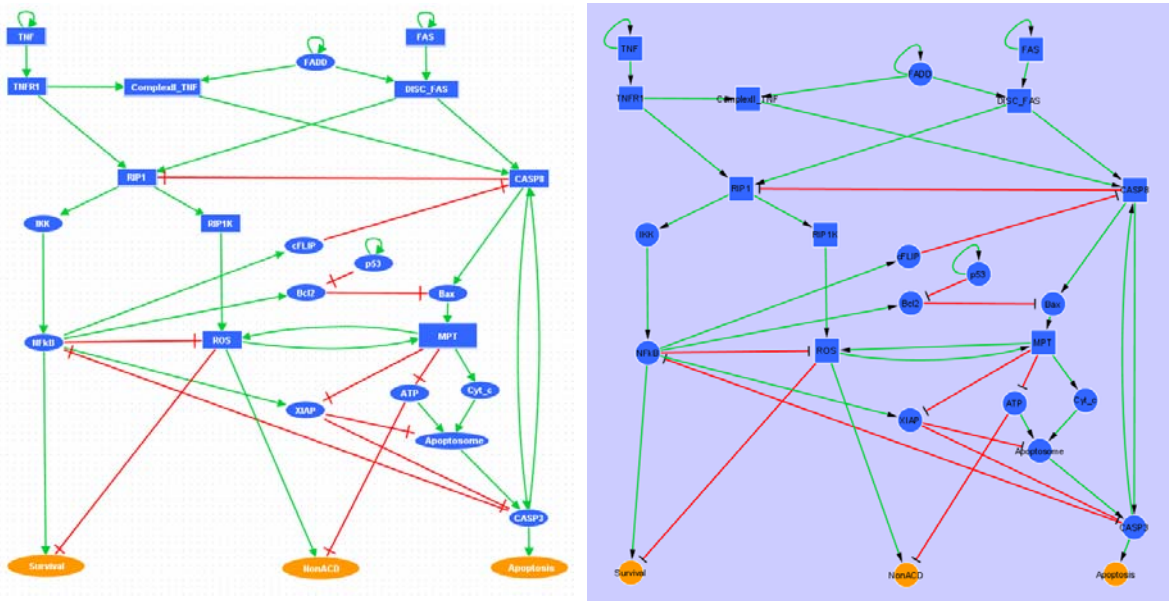
Note 2: The target arrows of the edges do not appear and should be defined manually through *VizMapper*.

**Step 3.** *BiNoM* plugin is used to extract the strongly connected components (SCC) of the network, along with all cycles from each SCC. The cycles are then clustered using *BiNoM* 'Cluster Networks' algorithm with 65% for the 'Intersection threshold' parameter. The subnetwork corresponding to the biggest cluster consists of 9 nodes and 11 edges (see Fig.2).

**Step 4.** The subnetwork is exported to *GINML* through the *Cytoscape2GINML* plugin.

**Step 5.** The *GINML* file is opened in *GINSim*.

**Note:** Depending on the original file, *GINsim* may reports some problems during the import such as 'InvalidGraphName' (due to the use of non standard characters in the graph name).



**Figure 2. Upper-left panel:** cell-fate decision network defined into *GINsim*. **Upper-right panel:** Result of the import of this network into *Cytoscape* using *GINsim2Cytoscape* plugin. **Lower panel:** subnetwork extracted from the cell-fate decision network using *BiNoM* and exported back into *GINsim* (only the graphs are shown).

### B) Delineation of a logical model draft from a pathway database dump in BioPAX format

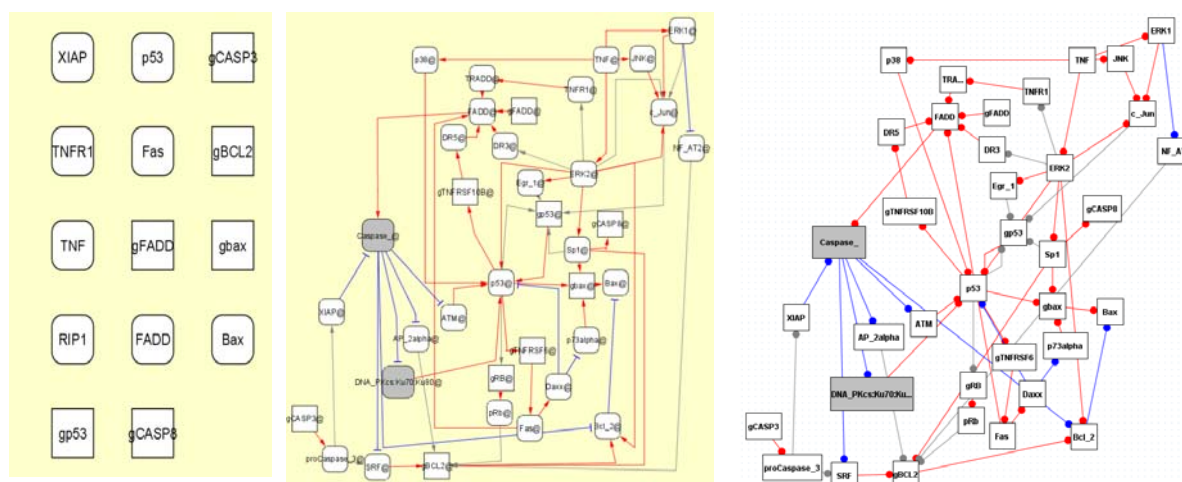
**Step 1.** The test.xgmml file generated in (A) was imported to *Cytoscape*.

**Step 2.** The *BioBase BioPAX* index together with the list of synonyms was loaded via *BiNoM*.

**Step 3.** The names of the nodes in the network were used to make a query from the *BioBase*. 14 genes and proteins were identified. *BiNoM* was asked to find all possible shortest paths between the pairs of genes/proteins in the semantic reaction network of *BioBase* with maximum of 2 intermediates. The result was transformed into the influence diagram.

**Step 4.** Using *Cytoscape2GINML* plugin, the diagram was exported in *GINML* format and the resulting file was opened with *GINsim*. The result is shown on Figure 3.

**Note:** Activation and inhibition arrows changed their appearance in *GINsim* after the conversion.



**Figure 2. Left panel:** Identified genes (squares) and proteins (rounded) from the *BioPAX* file; 'g' before the protein name denotes the corresponding gene. **Middle panel:** Result of *BioPAX* query in *Cytoscape*. **Right panel:** Result of *BioPAX* query imported into *GINsim*.

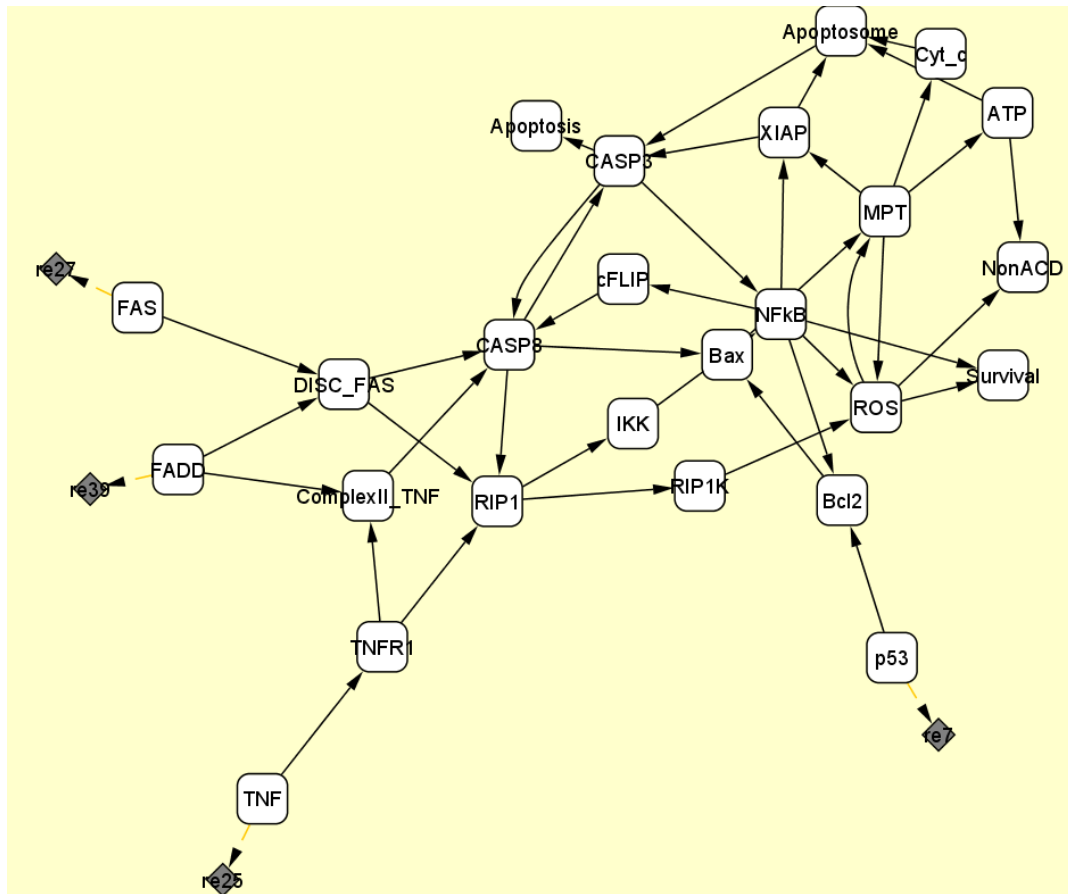
### C) Exporting the network structure from GINsim to BioPAX and SBML through BiNoM

**Step 1.** The test.xgmml file generated in (A) was imported to *Cytoscape*.

**Step 2.** The 'Export current network to SBML...' operation was performed. The corresponding generated *SBML* file was re-opened using *BiNoM* and we checked that all network connections were preserved in the transformation (after applying 'Mono-molecular reaction to edges' operation).

**Note:** The auto-regulations cannot be treated in *SBML* (since the product and the reactant are the same) and are thus translated as catalyses of some 'unknown' reactions. 'Inhibition' and 'Activation' annotations are also lost (see Figure 4).

**Step 3.** The imported *SBML* file was exported to *BioPAX*. The resulting *BioPAX* file was opened using *BiNoM* to check that all connections are conserved in the transformation.



**Figure 3.** View of the network after exporting to *SBML* and re-importing back to *BiNoM*.

### Software availability

All the tools used for these case studies are freely available at the addresses listed in Table 1 and the APO-SYS web-page. They are also provided along with all the files generated in a zip archive provided as supplementary material.

### Conclusions and prospects

The main objective of **deliverable 12**, namely facilitating the exchanges of network and model files between different relevant software (in particular *ConsensusPathDB*, *Cytoscape/BiNoM*, and *GINsim*) has been largely attained, as summarized in Figure 1.

However, at the time of this report, exchanges between *GINsim* and *BiNoM* remain mainly limited to structural (topological) information. Annotations (such as standard protein/gene identifiers, or publication references) are lost during the transfer. This limitation is presently

being addressed by the *UnivMed* team (updates of the plugins *GINsim2Cytoscape* and *Cytoscape2GINML* implementing solutions to these problems will be released in the course of 2009).

Our tests showed several places where the conversion requires manual intervention to deal with varying ways of representing regulatory networks. These aspects will also be automatized in forthcoming software releases.

We also consider the possibility to convert the information encoded in terms of logical rules with *GINsim* into *SBML*. This point is currently being addressed by the *UnivMed* team in collaboration with Nicolas Le Novère (*SBML* editor, EBI, UK) in the form of an extension of *SBML* Level 3, currently under development.

### Action points

- Development of an interface for *GINsim* to enable the capture of regulatory signs encoded using different attribute/value associations in *Cytoscape* files (in particular to cover different ways to define signs or effects of regulatory interactions).
- Extension of *Cytoscape2GINML* plugin to ease the correction of graph and component names defined in *Cytoscape* format for compliance with *GINML* format.
- Modification of *GINML* to integrate the annotation format used by *Cytoscape*.
- Integration of *Cytoscape2GINML* into *BiNoM*.

### References

González AG, Naldi A, Sánchez L, Thieffry D, Chaouiya C (2006). *GINsim*: a software suite for the qualitative modelling, simulation and analysis of regulatory networks. *Biosystems* **84**: 91-100.

Kamburov A, Wierling C, Lehrach H, Herwig R (2009). ConsensusPathDB-a database for integrating human functional interaction networks. *Nucleic Acids Res* **37**: D623-8.

Naldi A, Thieffry D, Chaouiya C (2007). Decision Diagrams for the Representation and Analysis of Logical Models of Genetic Networks. *Lect Notes Comput Sci* **4695**: 233-47.

Naldi A, Lopez F, Thieffry D, Chaouiya C (in preparation). Qualitative modelling and analysis of biological regulatory networks with *GINsim* 2.3.

Zinovyev A, Viara E, Calzone L, Barillot E (2008). *BiNoM*: a *Cytoscape* plugin for manipulating and analyzing biological networks. *Bioinformatics* **24**: 876-7.