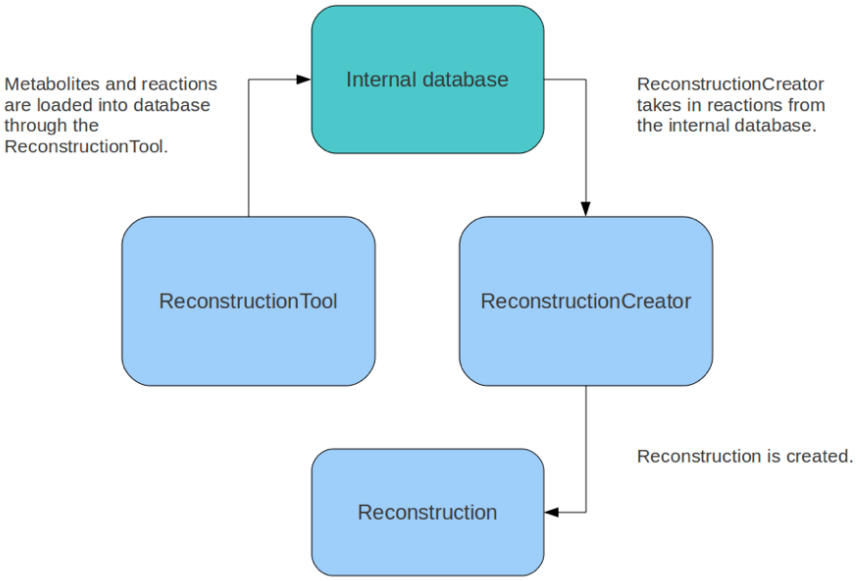
Generation and manipulation of reconstructions with rBioNet

Author(s): Ines Thiele, Ronan M. T. Fleming, Systems Biochemistry Group, LCSB, University of Luxembourg.

Reviewer(s): Catherine Clancy, LCSB, University of Luxembourg.

rBioNet is a reconstruction tool that lets you assemble reconstruction in a user friendly environment. In this tutorial you shall learn how you can use this tool to either start a new reconstruction or load in an existing one, followed by, its analysis. The tool consists of 3 main parts, i.e., metabolite creator, reaction creator and reconstruction creator. The metabolite creator is used to add in metabolites and its associated information, i.e., its elemental formula, charge, identifiers (for e.g., KEGG ID, PubChem ID etc.) and other associated attributes. Alternatively, a text file containing all the necessary information in the same order as in the metabolite database can be loaded directly. The reaction creator is used to formulate reactions and as stated before a text file containing all the necessary information about the reaction abbreviation, description, formula, reversibility, confidence score, notes, references. Alternatively, a text file containing all the necessary information in the same order as in the reaction database can be loaded on to the reaction creator directly. The reconstruction creator is used to load in reactions from the reactions database and then assign GPRs (gene-protein-reaction association), subsystem, add in more information in the notes and reference section. Once you have completed your reconstruction you can look at the S-matrix, identify dead ends, look for neighboring reaction to a particular reaction and plot metabolite connectivity in the reconstruction creator with its statistics function. Below we give you step by step approach on how to do this.



Features of rBioNet:

Environment to assemble reconstruction that consists of 3 parts

1. Metabolite creator
2. Reaction creator
3. Reconstruction creator

Metabolite creator

- Associated with a metabolite database.
- Used to create a new metabolite in one of three possible ways:

1. Uploading from a text file that contains all the information in the same order as in the database.
2. Manually filling in all the information.
3. Loading from other COBRA reconstructions.

- Checks for duplicate entries.
- Checks the metabolite abbreviation and charged formula.
- Metabolites are organism and compartment independent.

Reaction creator

- Associated with a reaction database.
- Methods to create a reaction is same as for metabolites.
- Reactions contain metabolites pre-existing in the metabolite database.
- Checks for duplicate entries, mass and charge balance.
- Reactions are organism independent but compartment specific. The same reaction can occur in different compartments.
- Either start from scratch or load pre-existing reconstruction.
- Primarily used to assign GPRs.
- Also to add notes, subsystem etc.

Add-ons

- Reconstruction analyzer.
- Checks for dead-end metabolites.
- Provides suggestions for exchange reactions.
- S-matrix visualization.
- Neighbor Reactions & Metabolite connectivity.

EQUIPMENT SETUP

Initialize the COBRA Toolbox.

Initialize The Cobra Toolbox using the `initCobraToolbox` function.

```
%initCobraToolbox
```

Setting the optimization solver.

This tutorial will be run with a `'glpk'` package, which is a linear programming (`'LP'`) solver. The `'glpk'` package does not require additional instalation and configuration.

```
solverName='glpk';
solverType='LP';
changeCobraSolver(solverName,solverType,1);
```

However, for the analysis of large models, such as Recon 3, it is not recommended to use the `'glpk'` package but rather an industrial strength solver, such as the `'gurobi'` package. For detailed information, refer to The Cobra Toolbox solver instalation guide.

A solver package may offer different types of optimization programmes to solve a problem. The above example used a LP optimization, other types of optimization programmes include; mixed-integer linear programming ('MILP'), quadratic programming ('QP'), and mixed-integer quadratic programming ('MIQP').

```
warning off MATLAB:subscripting:noSubscriptsSpecified
```

```
if usejava('desktop') % This line of code is to avoid execution of this turorial in non gui-en
```

Steps to load and initiate the Reconstruction Tool

Start up: rBioNet needs a pre-existing database to start up. Stored in the rBioNet is a database. The database consists of the ten reactions of the glycolysis pathway. Hence, you will see the glycolysis reactions in the reaction creator window, the metabolites participating in these reactions in the metabolite creator window.

0. Open MATLAB

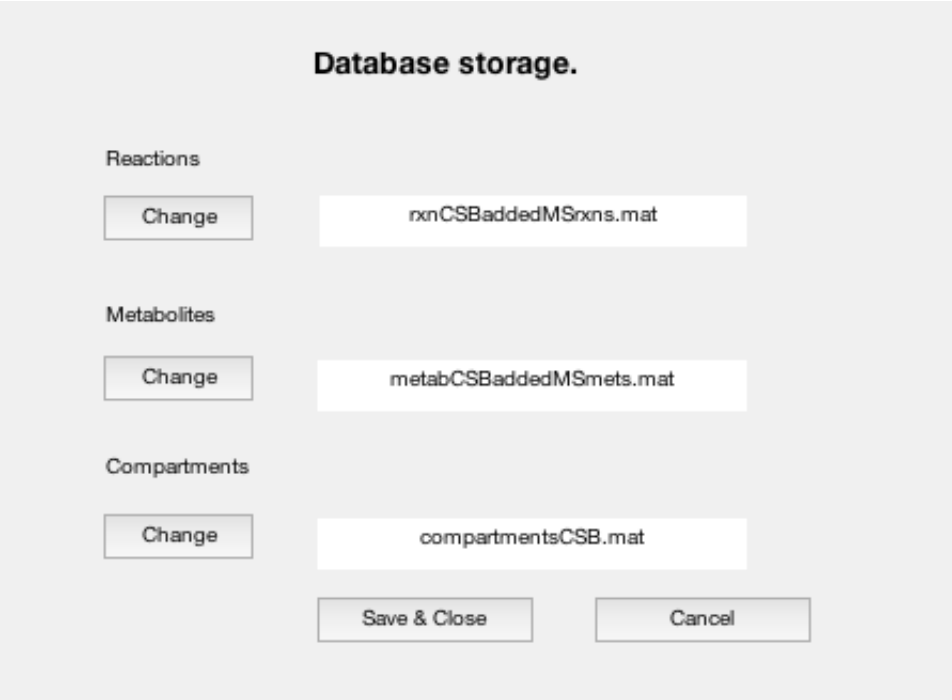
1. Add The CobraToolbox provided to you in your path:

In MATLAB window go to *file-> setpath-> add with subfolders-> cobra toolbox folder -> save & close*

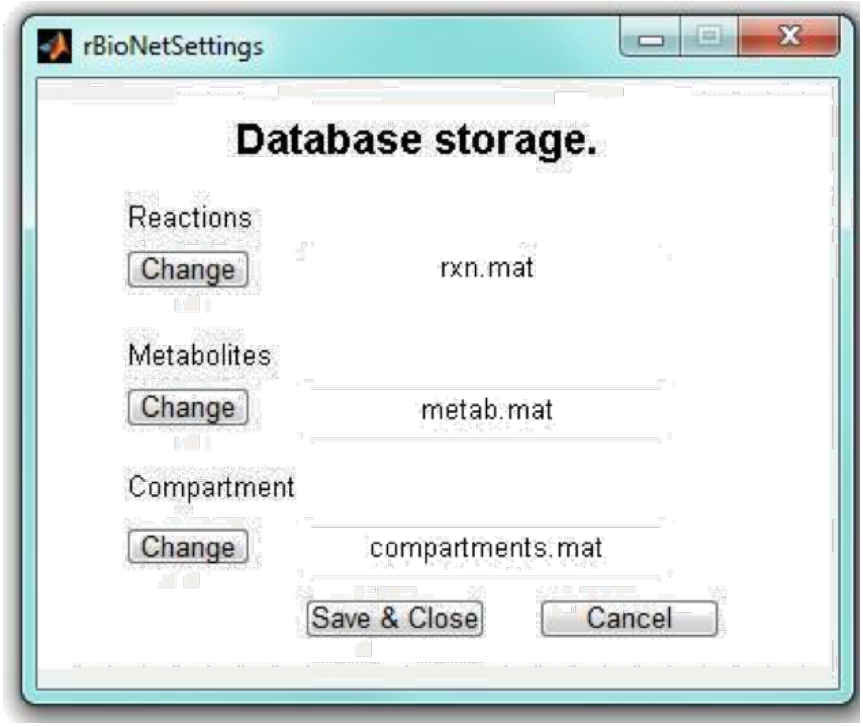
2. Initiate rBioNet by linking the database files.

- Note there are three files, corresponding to the reaction database (rxn.mat), metabolite database (metab.mat), and the compartment database (compartments.mat).

```
rBioNetSettings
```



- This window below appears:



Click on the change tab for reactions and locate to the rxn.mat file, which is provided to you in the rBioNet folder of The Cobra Toolbox.

Reactions-> Change-> rxn.mat->save

Click on the change tab under metabolites and locate to the metab.mat file, which is provided to you in the rBioNet folder of The Cobra Toolbox.

Metabolites->Change-> metab.mat->save

Click on the change tab under compartment and locate to the compartments.mat file, which is provided in the rBioNet folder of The Cobra Toolbox.

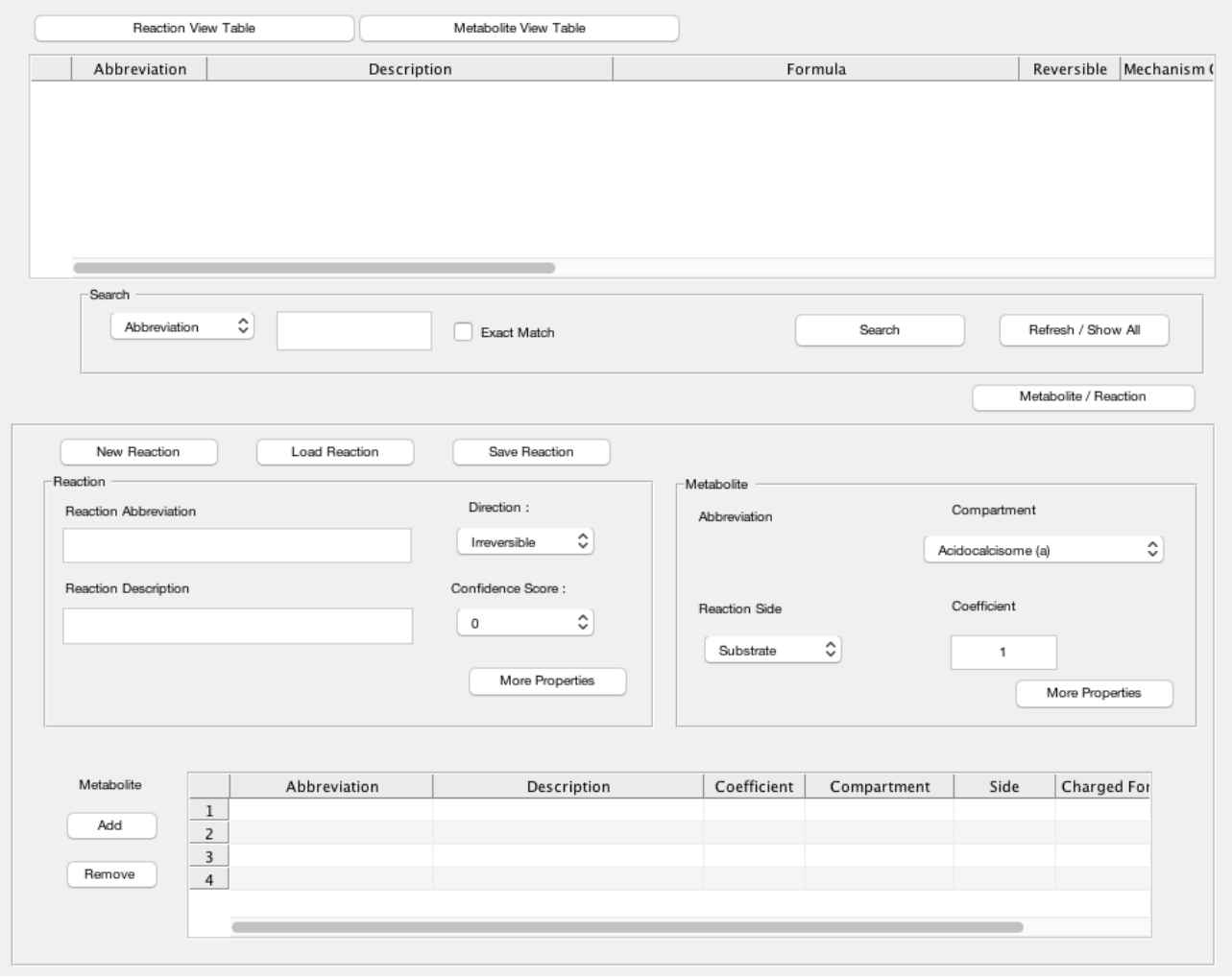
Compartment->Change-> compartments.mat -> save

This is the pre-existing database with glycolysis reactions and you saved it as your current database. You can modify it or remove the reactions as per your requirements.

Open and nagivating in rBioNet

Let's open the rBioNet tool:

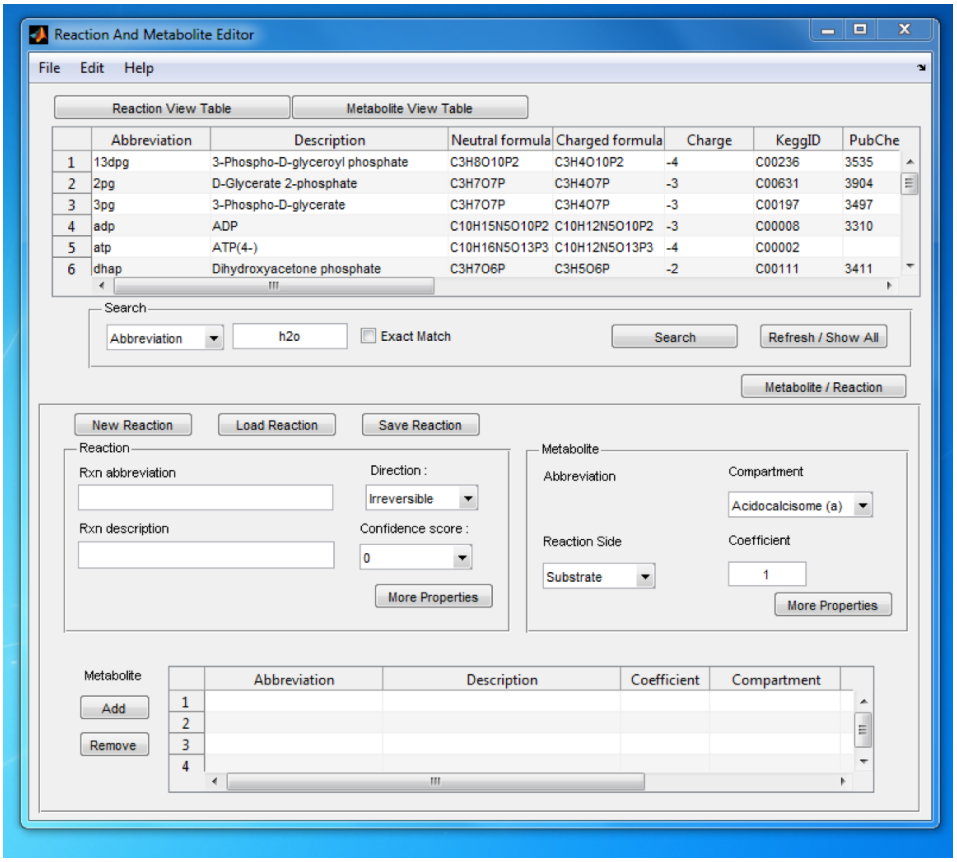
```
ReconstructionTool
```



A window appears called the 'Reaction and Metabolite Editor'.

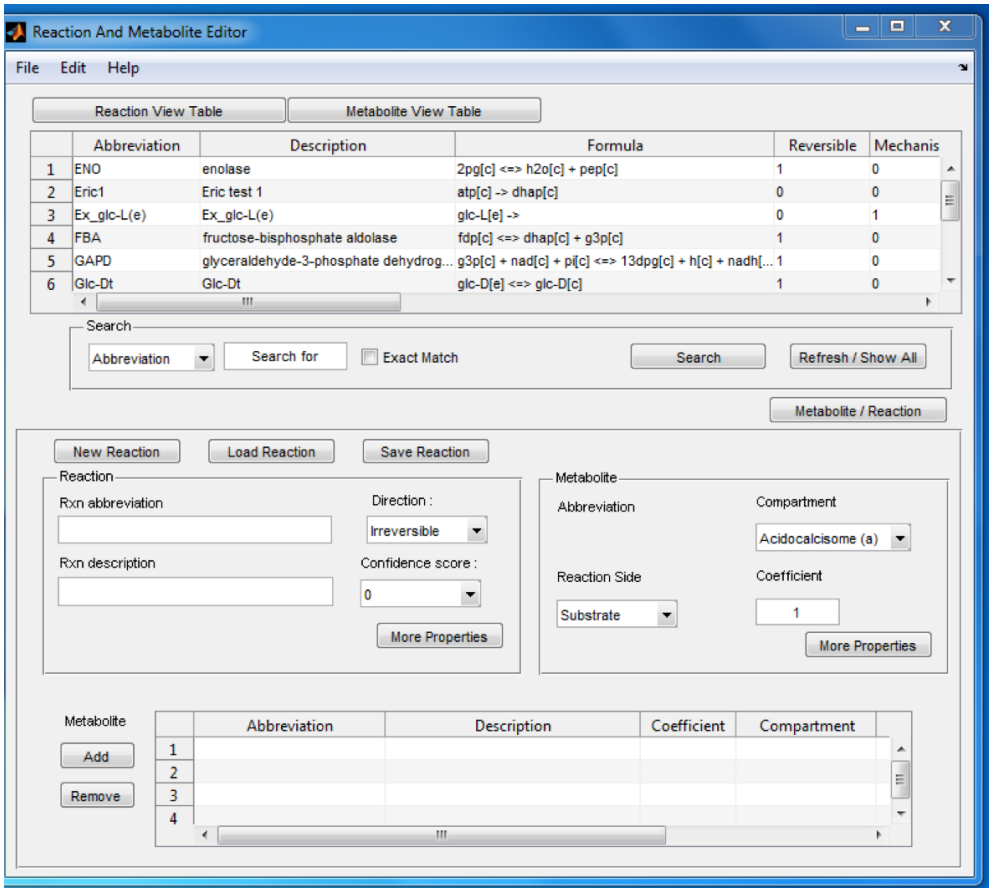
Click on the *Reaction View Table* and then S*how All*, which shall show all the glycolysis reactions.

Reaction view table: *Reaction creator -> Refresh/Show All*



Click on the *Metabolite View Table* and then S*how All*, which shall let you see all the glycolysis metabolites

Metabolite view table: *Metabolite creator -> Refresh/Show All*



To visualize the 'Reconstruction Creator' window go to File and Open the model creator.

File -> Open Model Creator -> Reconstruction Creator

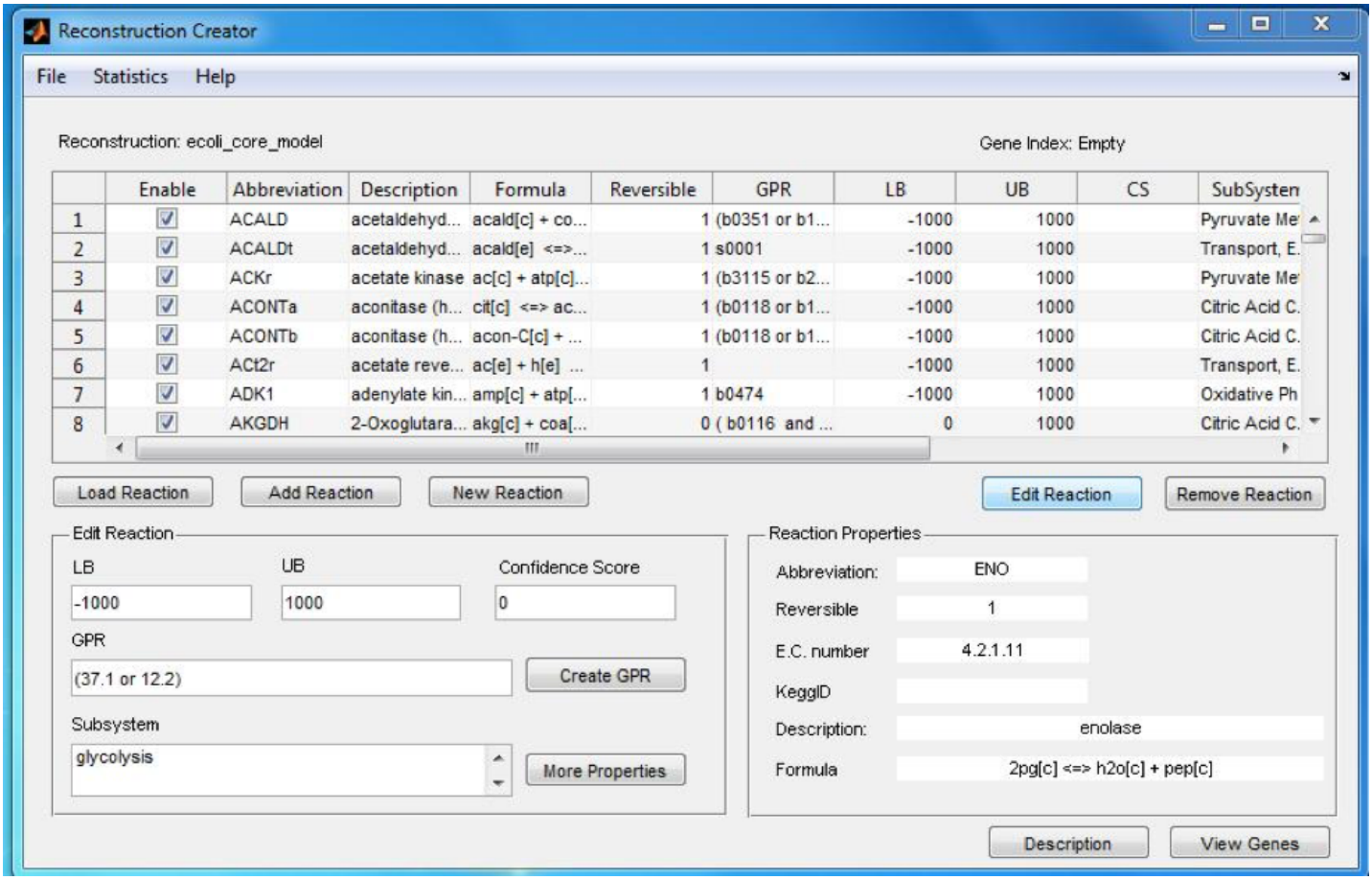


Figure: Reconstruction Creator with *E. coli_* core model

Load in the *E.coli* core model into the reconstruction creator

In the Reconstruction Creator do as follows:

File -> open model -> complete reconstruction -> select the E. coli_core_model (stored in the Examples folder of the rBioNet) -> click yes on the reconstruction description bar –> click no on the load gene index bar.

Now, we see the content of the *E. coli*_core model in the reconstruction creator (as shown in the Figure above).

On the reconstruction creator, click on S*tatistics -> Reconstruction analyze.* A window called the 'Reconstruction Analyzer′ appears and is used to visualize the S-matrix and identifies dead end metabolites.

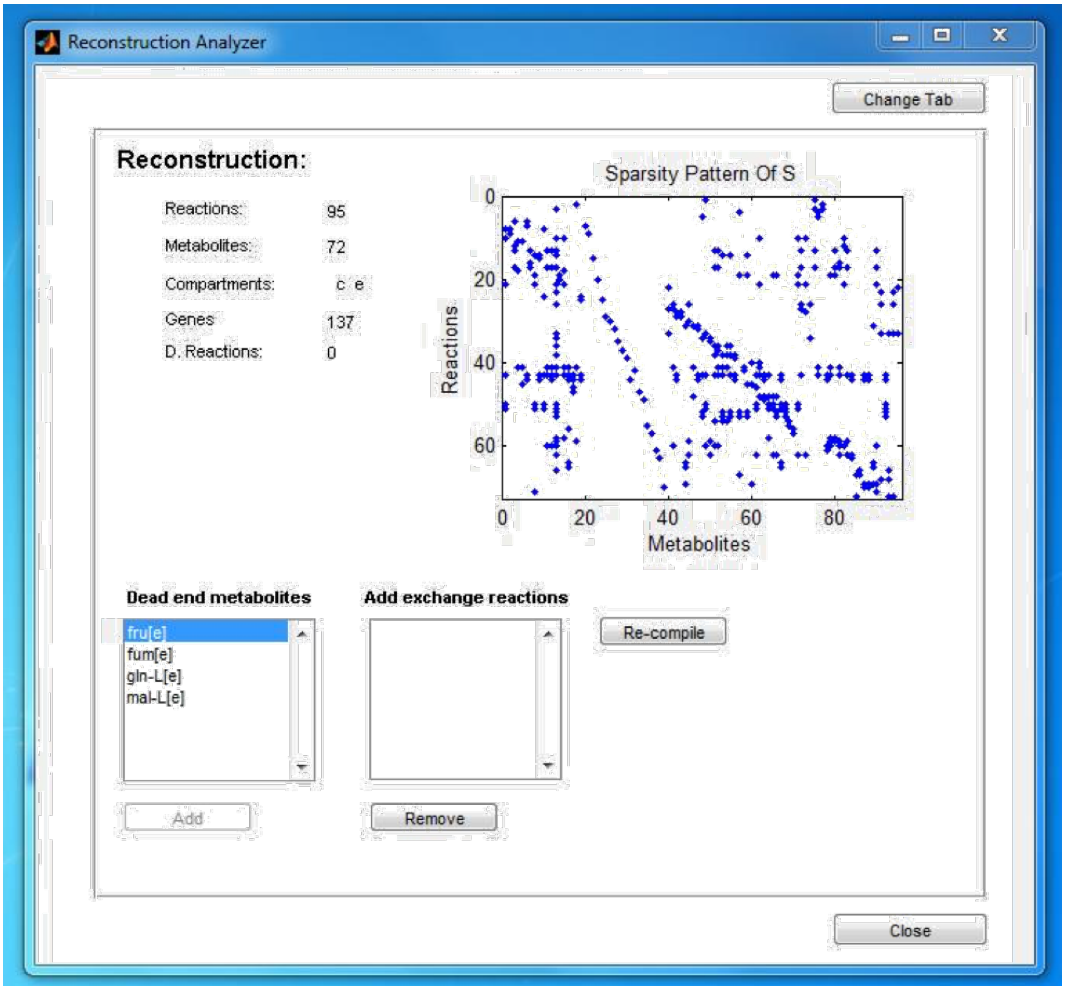
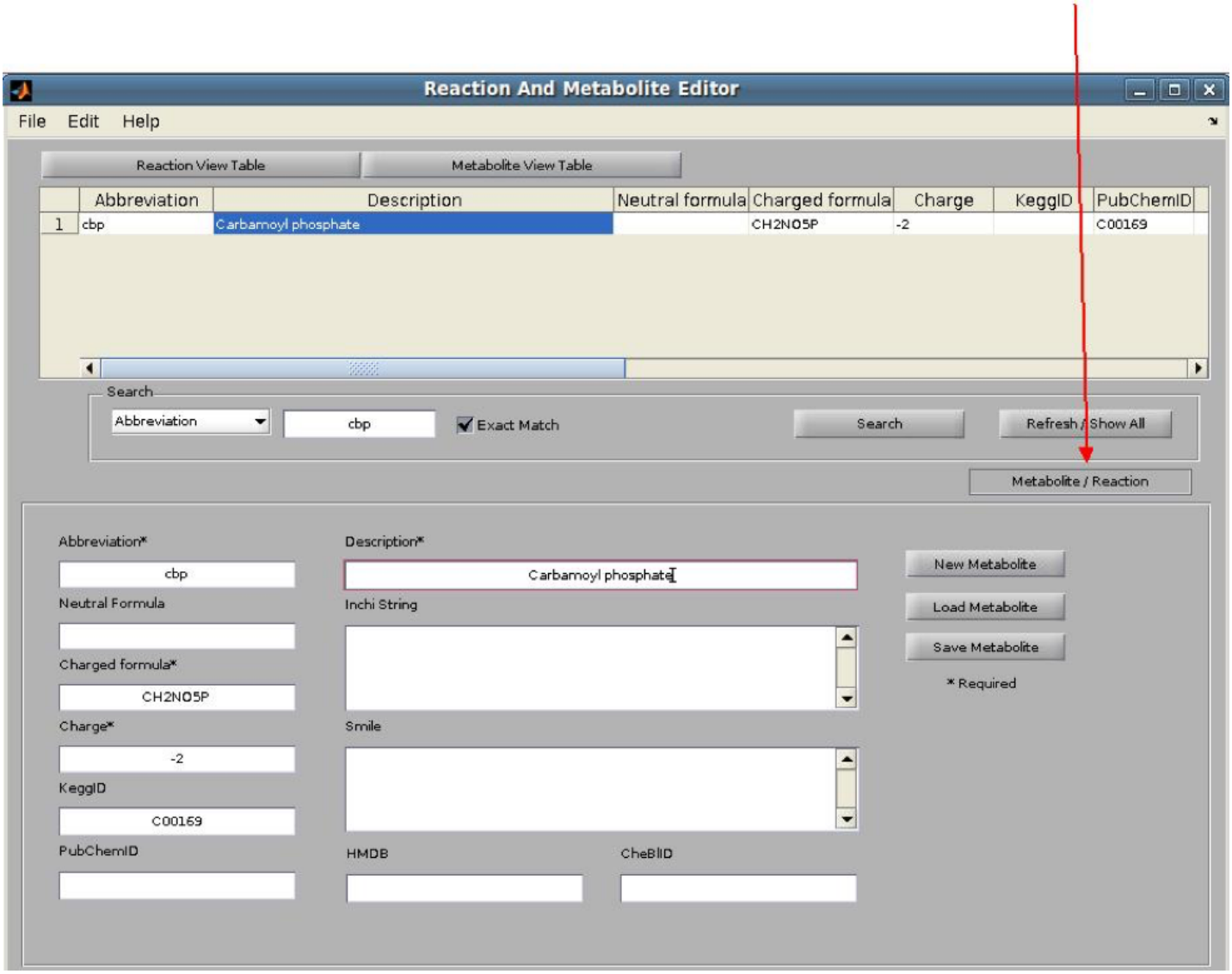


Figure: S-matrix and dead end metabolites for the *E. coli* core model

Adding a new metabolite

Go to the Reaction and Metabolite Editor window and click on Metabolite/Reaction.



- Enter the required information for the new metabolite, including: abbreviation, description, formula, charge and a metabolite ID.
- Then click Save metabolite.
- Repeat for all of the metabolites provided in the 'ureacycle_mets.txt' *(stored in the* TODO:name *folder)*. %TODO: add the ureacycle_mets.txt file to the tutorial (or preferably rBioNet) folder.

ALTERNATIVE

Alternatively, load a text file directly into the Reaction and Metabolite Editor.

Go to file -> add text file -> with metabolite -> select the file ureacycle_mets.txt *(stored in the TODO folder)*

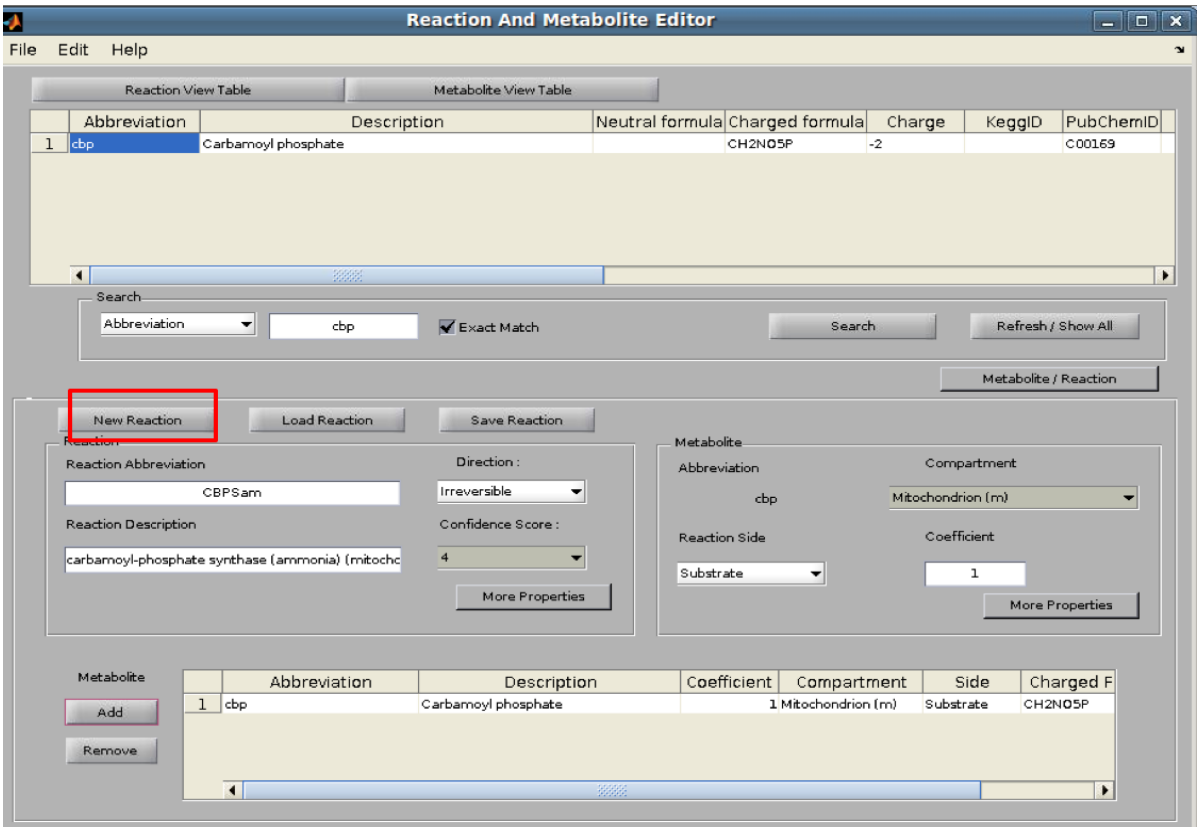
Click yes on each window that appears.

- When you are using this approach to create your reconstruction, make sure that all the information is **absolutely correct** and thoroughly checked before you make the addition.

Adding new reaction

Go to the Reaction and Metabolite Editor window, click on Metabolite/Reaction and yes for the window that appears.

Then, click 'New Reaction' and enter the reaction information including: the reaction abbreviation, description, direction and confidence score.

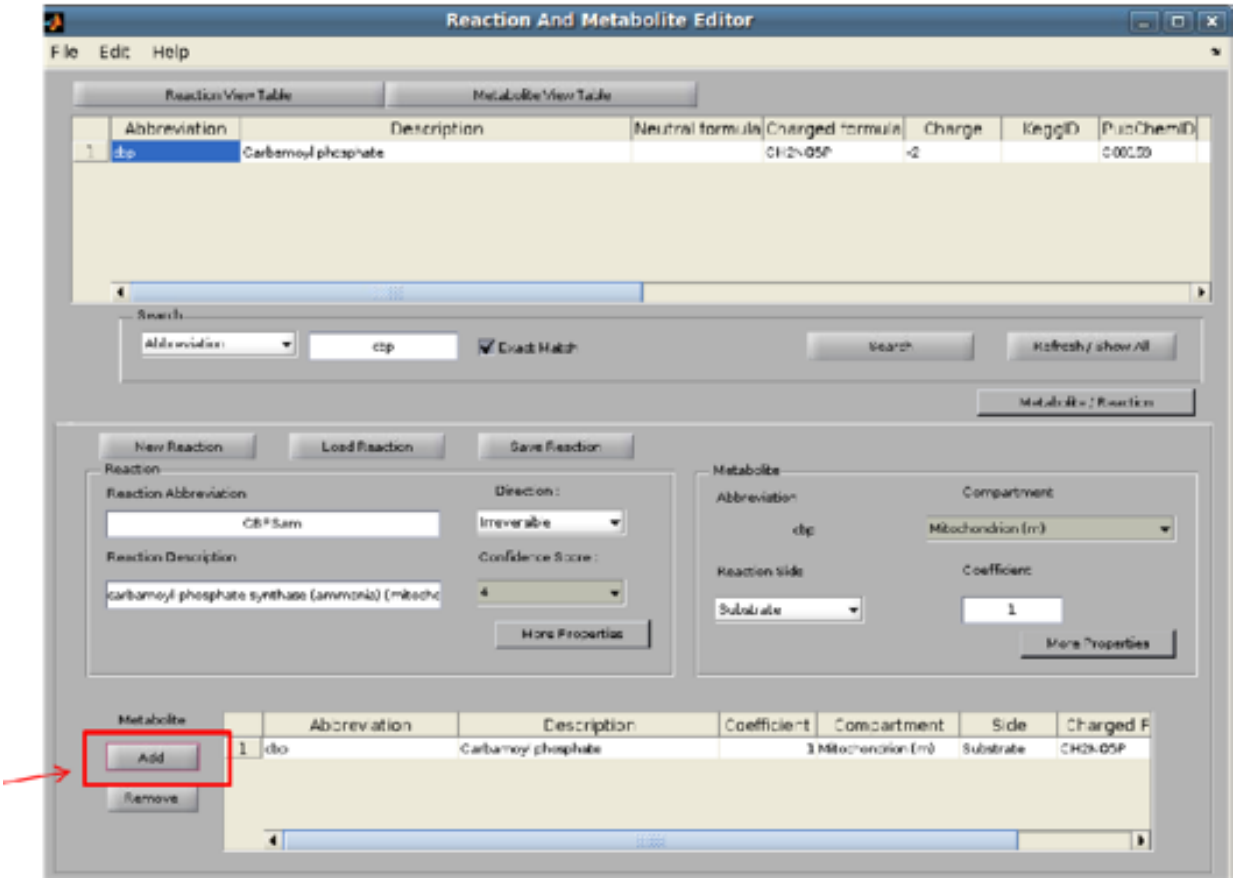


Next go the 'Metabolite View Table' and select a metabolite belonging to the reaction.

Enter the metabolite's coefficient, compartment and reaction side (substrate or product).

Then click Add to add the metabolite to the reaction.

Repeat for all metabolites in the reaction, and when finished, click save.



- For the assignment of GPRs to reaction ('Create GPR') please refer to the next section.

Now, the tool checks for elemental and charge balancing, and provides a warning if there is an error.

If everything is correct in the follow up window, click yes to save your reaction.

Repeat this for the five reactions given in the hand out. %TODO What reactions? There is not a hand out with this tutorial.

ALTERNATIVE

Alternatively, load a text file directly into the Reaction and Metabolite Editor.

Go to file -> add text file -> with reactions -> select the file *ureacycle_mets.txt* (stored in the TODO:name folder)

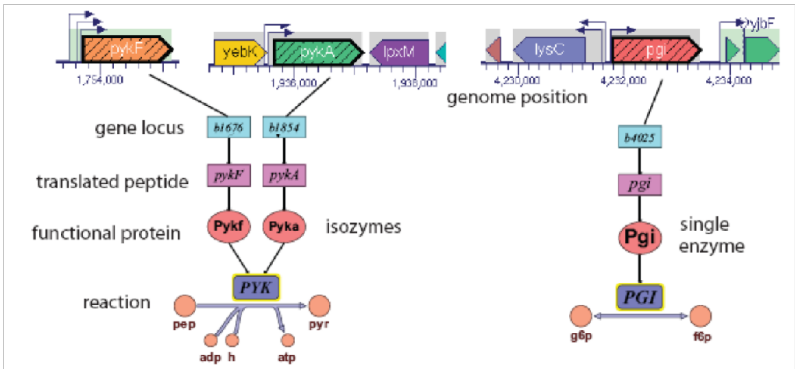
Click yes on each window that appears.

- When you are using this approach to create your reconstruction, make sure that all the information is **absolutely correct** and thoroughly checked before you make the addition.

Assigning GPRs to the reactions

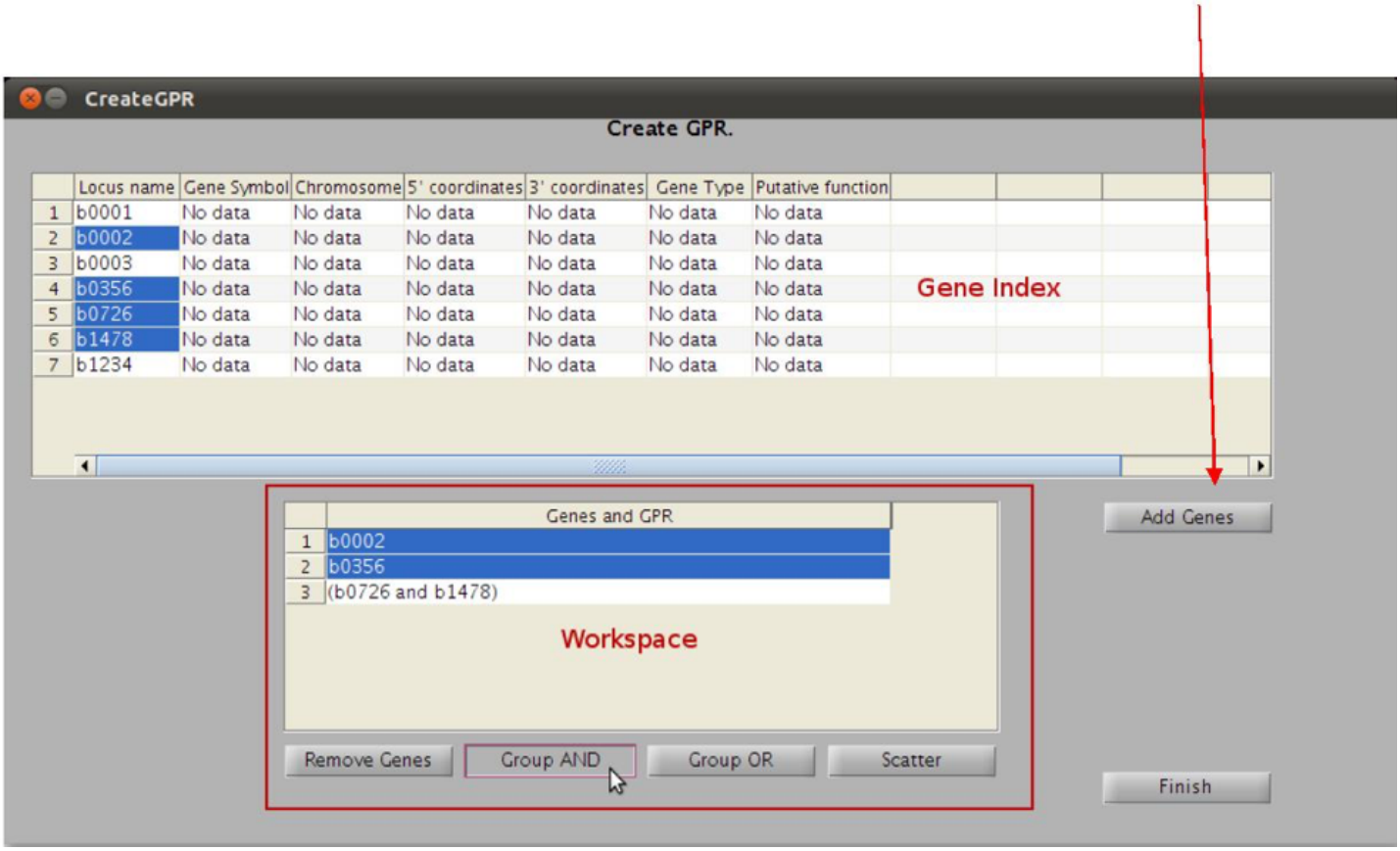
Gene protein reaction associations

Some reactions are catalyzed by more than one enzyme



- On the 'Reaction and Metabolite Editor' go to File -> Open 'Model creator'
- On the 'Reconstruction creator' window go to File -> Click on 'Load gene index' -> select the file *ureacycle_genes.txt* (stored in the folder) %TODO add the gene index file to the tutorial folder.
- Select 'Load reaction' - > select a reaction that you want to assign a GPR, and then click 'Load reaction'.
- Click on create GPR
- Click on genes you want to assign -> click add genes.
- Add a AND/OR rule to a group of genes: click 'Group AND' for groups selected genes with and, or click 'Group OR' for groups selected genes with or.

Add Genes tab



- Once the GPR has been assigned and other edits done (for e.g., addition of subsystem, further notes and references), click the add reaction button in the reconstruction creator.
- A window appears asking if you want to replace the current reaction -> click yes.
- Now you should be able to see the new edited (i.e., GPR associated) reaction in the list of reactions on the reconstruction creator.
- Once you have finished with all the reactions, go to file and save your reconstruction as a model.
- If you made an errors while adding metabolites or reactions manually, you can load in the reaction/ metabolite database (i.e., rxn.mat/ metab.mat files) directly into the MATLAB workspace and delete/ edit it in the variable editor in the MATLAB and save it. This shall be your new database from now on.
- Make sure that you never make errors while filling in the information in the database.

Explore reconstruction analyzer

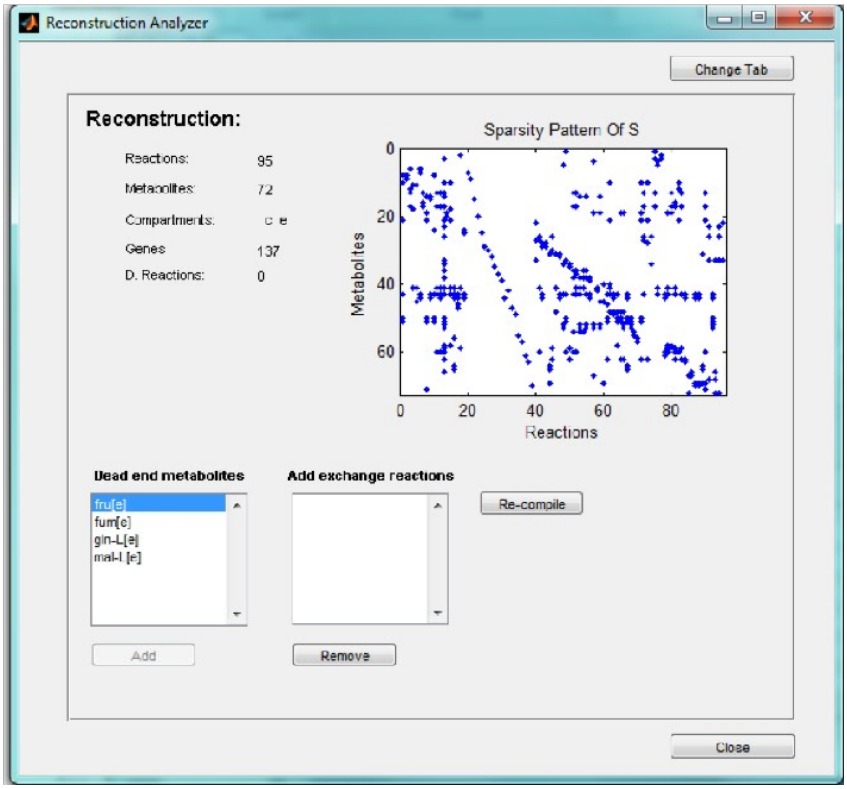
Load in *E. coli* core model

Open reconstruction analyzer

-> statistics
-> reconstruction analyzer

Visualize
the S-matrix
dead end-metabolites

Change tab
-> neighbor reactions
-> search Metabolite connectivity
-> plot




```
end
```

Reference

[1] Thorleifsson SG, Thiele I. Bioinformatics. 2011 Jul 15;27(14):2009-10.