Metabolic modelling, Spring 07, Exercise 2, Friday, 30.3.2007

1. Consider the metabolic network consisting of the following enzymatic reactions: $A \rightarrow B, B \rightleftharpoons C, B \rightarrow D, D \rightarrow C$. Draw the graph representation of the network.

   What is/are the

   (a) Stoichiometric matrix

   (b) Enzyme subsets (solve either manually or by using MATLAB)

   of the network?

2. Systems biology research group at University of California, San Diego, keeps the repository of genome scale metabolic network models of different organisms. Go to their www site *http://gcrg.ucsd.edu/organisms/hpylori.html* and download a stoichiometric matrix corresponding the metabolic network of *Helicobacter pylori* *i*IT341 GSM/GPR. Write a MATLAB function that

   (a) loads the stoichiometric matrix to MATLAB (hint: function *load*),

   (b) computes the degree distribution of metabolites in the network, that is, the distribution of the number of reactions connected to each metabolite in the network (hint: functions *find, sum, histc*),

   (c) draws a histogram corresponding the distribution (*plot, hist*). Use both normal and log-log scale (*loglog*).

   Prepare to demo the function in front of the class using MATLAB.

3. Fit the scale-free model $P(k) \approx k^{-\gamma}$ to the degree distribution of metabolites of the previous assignment by selecting γ that minimizes the sum of squared errors between the model and the observed distribution. Use for example MATLAB function *fminsearch*. Write a MATLAB function, that performs the fitting. Is the fit good? Try to improve the fit by excluding "outliers" from the observed distribution, that is, select some subinterval for k such that most values k_i where an empirical probability $P(k_i) \approx 0$ in the data are excluded. Did the fit improve? Was it a good idea to exclude the outliers? Prepare to demo the processing steps you took in front of the class using MATLAB.

4. Read the article **M. Arita: The metabolic world of Escherichia coli is not small. PNAS 101 (2004), 1543 - 1547** so that you can explain the main points of the article.