REI601M **Assignment 7**

Given: 21.2.2018 **Due: Wed 28.2.2018**

In this assignment you will use Matlab COBRA or COBRApy. Hand in by sending an email to steinng@hi.is. Your report should be **at most 4 A4** pages so you need to aim for quality instead of quantity. Provide all Matlab/Python code that you write in an appendix.

1. (Growth coupling via knockouts, 50%) Find *E. coli* mutants which couple the secretion of a target compound to growth by knocking out appropriate genes or reactions (note 1). Test the following compounds, one by one: i) acetate, ii) lactate, iii) succinate and iv) ethanol using the *E. coli* core model, aerobic conditions and glucose as the carbon source. Your results should be in the form of graphs illustrating the maximum and minimum amount of product as a function of growth for selected mutants, together with the names of the reactions or genes identified. Include the wild-type (non-modified model) as well (note 2).

2. (Production of a non-native compound, 30%) Create an *E. coli* mutant which produces 3-hydroxy propionic acid, 3-HPA), used in the manufacturing of biodegradable plastics, from glycerol or some other inexpensive carbon source. You need to select a suitable starting compound in the model together with enzymes of the new pathway leading to 3-HPA. Use the iJO1366 model and select a pathway from one of the following resources

http://www.sciencedirect.com/science/article/pii/S1096717614001256
http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0156286

Add reactions of the new pathway to the model, taking care to balance the chemical equations, as well as an exchange reaction for 3-HPA. List the reactions involved in the pathway. How much can be produced under i) aerobic conditions and ii) anaerobic conditions? Carry out sensitivity analysis of the production w.r.t. oxygen uptake (see the paper by Borodina cited above) using "phenotypic phase plances" in Cobra. What can be said of co-fermentation products (products secreted alongside 3-HPA)? How does the pathway you selected compare to the pathway below which was discussed in class, in terms of theoretical yields (i.e. maximum amount of 3-HPA for a fixed glucose uptake)?

```
DHAB : glyc[c] -> 3hpald[c] + h2o[c]
HPALDD :  3hpald[c] + nad[c] + h2o[c] -> 3hpa[c] nadh[c] + 2 h[c]
DM_3hpa : 3hpa[c] ->
```

3. (Flux modulation, 20%) Identify bottlenecks in succinate production in the *E. coli* core model using flux variability analysis. Consider the following conditions: i) 90% of maximum biomass is produced and ii) 90% of maximum succinate is produced. Perform flux variability analysis for both conditions and plot the min/max intervals for each reaction for both conditions in a single figure (you may need to experiment a bit until you get a nice figure), e.g. biomass in red and succinate in blue. Identify non-overlapping flux intervals. These intervals indicate fluxes which must decrease/increase or be knocked out for succinate production to be increased. Exclude exchange and transport reactions from the list.

Notes

1. There exist several algorithms which search systematically for reaction (or gene) knockouts that couple chemical production to growth (c.f. OptKnock, RobustKnock, GDSL, OptGene, OptFlux). Most of these algorithms formulate the search as a Mixed Integer Linear Program which in turn requires an industrial-strength optimizer such as Gurobi (Academic licenses can be obtained for free via Gurobi.com).

You should write a Cobra/CobraPy function which tests all single reaction knockouts, by fixing flux through the corresponding reactions to zero. For each knockout

```
Determine the maximum growth rate, v_biomax
If v_biomax == 0 then proceed to next knockout (lethal knockout)
Select exchange reaction for the target compound as an objective
v_bio= linspace(0, v_biomax, 10)
for v in v_bio
   Fix the growth rate to v
   Maximize to determine upper bounds for the target
   Minimize to determine lower bounds for the target
Plot both the minimum and maximum as a function of v.
```

It is likely that more than one knockout is needed for growth-coupling so you need to modify your code to include a double for-loop over all pairs of reactions.

Prior to performing the iterations, it helps to reduce the model as much as possible by removing blocked reactions under the given conditions (e.g. aerobic growth on glucose), identified by flux variability analysis (look for minFlux ≈ maxFlux ≈ 0). Exclude exchange and transport reactions as well using findExcRxns and findTransRxns to reduce the set of candidate reactions even further.

2. Such figures are conveniently obtained with the productionEnvelope function in Matlab. The following example calculates the product-growth relationship for succinate when the lactate dehydrogenase and glucose-6-phosphate-isomerize enzymes knocked out (LDH_D and PGL).

```
>> productionEnvelope(model,{'LDH_D', 'PGL'},'k','EX_succ(e)', ...
                  model.rxns(find(model.c)),false);
```