

Mini-course 3: Systems Biology (Biochemistry 714)

Assessment 1: Tutorial test

Examiners: Dr DD van Niekerk and Prof JM Rohwer

Time: 2.5 hours (14h00–16h30)

Marks: 50

Instructions

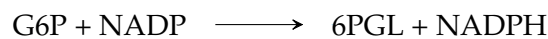
This is an **open-book** assessment. You may make use of your notes as well as online tutorial material including Jupyter notebook memos. However, **communication between candidates is not allowed**, you have to submit your own work.

Hand in a Jupyter notebook in which you answer the questions below. The notebook should be submitted via the dedicated link on SUNLearn, and the file should be named as follows: studentnumber_Lastname_Initials.ipynb (e.g. 12345678_Rohwer_JM.ipynb)

The assignment will be graded by self- and peer-assessment on Thursday 28 March. Further instructions will follow.

Background

Glucose 6-phosphate dehydrogenase (G6PDH) is an enzyme that catalyses the following reaction:



The enzyme was characterised kinetically by determining the initial rate of the formation of the product NADPH through spectrophotometric measurement at 340 nm. The experiment was repeated for various concentrations of both substrates.

The rate of the enzyme-catalysed reaction can be described by the following rate equation:

$$v = \frac{V_f ab}{(K_a + a)(K_b + b)} \quad (1)$$

where a and b refer to the concentrations of the substrates G6P and NADP, respectively, V_f is the maximal (limiting) rate, and K_a and K_b are the respective Michaelis constants for the substrates.

Tasks

All analyses should be performed in a Jupyter notebook, and at the end you should only submit this one file. You can insert headings as required and if you need to write text for explanations, you can use the “Markdown cell” functionality of the notebook.

In the Download section of the course website (folder Tutorial_assessment) you will find data files from the kinetic experiments. These are named AxxByy.csv, where xx refers to the concentration of A (in mM) for that particular experiment, and yy to the concentration of B.

The files are comma-delimited and contain two columns: time in minutes, and concentration of NADPH in mM.

1. Visualise the data by plotting the time courses as **scatter plots**, as follows: [15]
 - On the first graph, plot all the experiments where a was varied and b kept constant at 24 mM, i.e. $a = 0.5, 1, 2, 4$ and 8 mM. Plot these on the same set of axes, using different colours to distinguish the datasets. Label the axes and provide a legend.
 - On a second graph, do the same for the experiments where b was varied and a kept constant at 8 mM, i.e. $b = 1.5, 3, 6, 12$ and 24 mM. Format the graph in the same way as the first.
2. Perform linear regressions on each of the datasets to calculate the initial rate. [5]
3. Create a new pandas dataframe that combines the data. The columns of the dataframe should be $[a, b, \text{rate}]$. [10]

If you did not manage to perform step 2, you can use the following data to carry on with the rest of the assignment. Note that these are *not the correct values*, so only use these if you could not obtain the real values!

Backup initial rate data		
a (mM)	b (mM)	rate (mM min ⁻¹)
0	0	0
0.5	24	0.512
1	24	0.816
2	24	1.202
4	24	1.561
8	24	1.747
8	1.5	0.603
8	3	0.912
8	6	1.267
8	12	1.558

4. Fit the data to Equation (1) to obtain estimates for the parameters V_f , K_a and K_b . Perform a *global fit on all the data* simultaneously. Provide error estimates for each of the parameters and give their units! [10]
5. Plot the fit. The graphs should contain the fitted function (as a line) and the experimental data (as points). Provide two graphs: [10]
 - *rate vs. a* (with b fixed at its highest value)
 - *rate vs. b* (with a fixed at its highest value)

You will therefore use only a part of the data for each plot. Label the axes (including units) and provide a legend for the lines/points.