R task: Analyzing enzyme kinetic data

The enzyme 5-α reductase type 2 is present in small amounts in the prostate and converts testosterone into dihydrotestosterone (DHT) following Michaelis-Menten kinetics. DHT is an androgen that plays a role in maintaining the proper functioning of the prostate. The drug finasteride inhibits the activity of this enzym, and thus inhibits the conversion of testosterone into DHT. Such a drug one only uses when the androgenic or masculinizing effects are too much like prostate problems or suffering from male baldness

With regard to reaction kinetics is 5-α reductase type 2 the enzyme E, testosterone the substrate S, and DHT the product P in the reaction mechanism

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} SE \overset{k_2}{\rightarrow} E + P$$
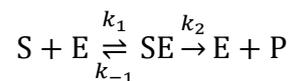
Hereto belongs the following Michaelis-Menten formula for the overall reaction rate r of the conversion $S \rightarrow P$:

$$r = \frac{v_{max} \cdot s}{K_m + s}$$

where s of the conversion

You are going to determine from experimental data the best fit of kinetic parameters v_{max} and K_m You are going to determine from experimental data the best fit of kinetic parameters.

Assigment 1: The meaning of kinetic parameters.

What is the meaning of v_{max} in the Michaelis-Menten formula $r = \frac{v_{max} \cdot s}{K_m + s}$?

What is the meaning of K_m in the Michaelis-Menten formula?

Is it possible to determine the values of these kinetic parameters from experimental data? Why or why not?

Hint at v_{max}. Can you approximate the Michaelis-Menten formula with a simpler formula for very high substrate concentration and, if so, what is this approximation? What does that say about v_{max}?

Hint at K_m. Can you approximate the Michaelis-Menten formula with a simpler formula for very low substrate concentration and, if so, what is this approximation?

What can you conclude when the substrate concentration s equals the Michaelis-Menten constant K_m?

Assignment 2: Differential equations in chemical kinetics.

The Michaelis-Menten formula leads to a differential equation for the substrate concentration s. What is this differential equation in the language of mathematical fomulas?

Suppose that the substrate concentration in an experiment at time $t = 0$ equals s_0. What can you write down about the concentration p of the product in case you may assume that at time $t = 0$ no product has been formed yet (no DHT, testosterone present only)..

What is the differential equation for the product concentration p?

Measured data

In the following article we found data about the concentration of the product in the course of the enzymatic conversion of testosterone into DHT.

Moss, M.L, Kuzmic, P., Stuart, J.D., et al (1996). Inhibition of human steroid 5-alpha reductases type I and II by 6-aza-steroids: Structural determinants of one-step vs. two-step mechanism, *Biochemistry* **35** (11), 3457-3464.

t (min)	p (nM)	t (min)	p (nM)	t (min)	p (nM)
2	1.99	18	16.87	38	26.67
4	3.86	20	17.80	42	27.66
6	6.05	22	19.65	46	28.63
8	7.90	24	20.35	50	29.12
10	9.53	26	21.88	54	29.42
12	11.77	28	22.84	60	29.91
14	13.22	30	23.53		
16	14.70	32	25.09		

Assignment 3: Estimation of the initial substrate concentration.

Plot the product concentration p (in nM) via an R script against time t (in minutes) for the given data set.

How high do you estimate the initial concentration s_0 of the substrate on the basis of the graph?

Assignment 4: From nonlinear to linear regression: the Lineweaver-Burk plot.

The Michaelis-Menten formula for the reaction rate r as a function of substrate concentration s is the nonlinear relationship

$$r = \frac{v_{\max} \cdot s}{K_m + s}$$

When you know the reaction rate at given substrate concentrations s (and we still have to realise this), then you estimate the kinetic parameters $v_{\max}$ and K_m via a nonlinear regression method. But in order to achieve this, you must have adequate initial values of these parameters for the nonlinear regression method. One way to find these values is a transformation of the data to a linear relationship.

One way that you encounter in many textbooks is the creation of the **Lineweaver-Burk plot** (original source: Lineweaver, H. & Burk, D. (1934). The determination of enzyme dissociation constants, *Journal of the American Chemical Society* **56** (3), 658-666). In this graph you plot $\frac{1}{r}$ against $\frac{1}{s}$ for the given data set and you expect to see data points on a straight line.

What is the relationship between $\frac{1}{r}$ and $\frac{1}{s}$ in the language of mathematical formulas?

How do you see from the formula that its graphical representation correpsonds with a straight line?

What is the slope of this straight line?

What is the intercept of this straight line?

How can the kinetic parameters $v_{\max}$ and K_m be calculated from the slope and intercept?

Assignment 5: Numerical reaction rate and the Lineweaver-Burk plot.

The reaction rate r is equal to the rate of formation of the product. What do you think is an appropriate way to approximate the reaction rate numerically at several times based on the given data set. In other words, how are you going to calculate the numerical derivative $\frac{dp}{dt}$? Implement your method of R..

Note that here we do not have equidistant data, because at the beginning the time interval between two measurements is 2 minutes, later on it is equal to 4 minutes, and finally equals 6 minutes. Does your method for computing a numerical derivative take this into account?

To determine the quality of your numerical derivative and to estimate the kinetic parameter values you create the so-called **Lineweaver-Burk plot** in R. So you calculate $\frac{1}{r}$ and plot it against $\frac{1}{s}$ for the given data set. You expect points on a straight line.

Is the expectation of a straight line met for your numerical derivative r? If not, try to improve your method for numerical differentiation.

Hint Numerical derivatives can be calculated via forward finite differences, backward finite differences, or a combination of both.

In this task with non-equidistant data you are advised to compute a numerical derivative point-by-point in a repetition loop.

Assignment 6: Parameter estimation via the Lineweaver-Burk plot.

In the Lineweaver-Burk plot that you constructed in the previous task, only calculated points are plotted roughly on a a straight line. You do not have a formula for a linear

function as the best fit of these calculated points to go on with. Carry out a linear fit in R using the lm command (check the manual *Using R in mathematics* in the section on regression analysis).

Esitmate the values of the kinetic parameters $v_{\max}$ and K_m with your linear regression curve.

Do these estimates agree with the values estimated earlier from the graph of r versus s and the meaning of $v_{\max}$ and K_m herein (assignment 1)?

Assignment 7: Sensitivity of the estimation of kinetic parameters.

The calculated Lineweaver-Burk plot depends on the estimation of the initial concentration s_0 of the substrate, because these values have you indeed used to compute the substrate concentration s at other times.

Find out how the values of the kinetic parameters change when you vary s_0, say with ± 0.1 nM.

Assignment 8: Eadie-Hofstee plot and Hanes-Woolf plot.

Alternative nonlinear transformations of the Michaelis-Menten formula $r = \dfrac{V_{\max} \cdot s}{K_m + s}$ toward a linear model are

- the *Eadie-Hofstee plot*, in which the reaction rate r is plotted against the quotient $\dfrac{r}{s}$ of the reaction rate r and the substrate concentration s
  Original sources:
  Eadie, G.S. (1942). The Inhibition of Cholinesterase by Physostigmine and Prostigmine. *Journal of Biological Chemistry* **146** (1) 85–93.
  Hofstee, BHJ (1959). Non-Inverted Versus Inverted Plots in Enzyme Kinetics. *Nature* **184**, 1296–1298.
- the *Hanes-Woolf plot*, in which the quotient $\dfrac{s}{r}$ of the substrate concentration s and the reaction rate r is plotted against the substrate concentration s.
  Original sources:
  Hanes, C.S. (1932). Studies on plant amylases: The effect of starch concentration upon the velocity of hydrolysis by the amylase of germinated barley. *Biochemical Journal.* **26** (5) 1406–1421
  Haldane, J.B.S. (1957). Graphical methods in enzyme chemistry. *Nature* **179**, 832-835.

Determine the linear formulas for the Eadie-Hofstee plot, $r = a + b \cdot \dfrac{r}{s}$, and the Hanes-Woolf plot, $\dfrac{s}{r} = c + d \cdot s$. In other words, find out what the mathematical expressions for a, b, c, d are in terms of the kinetic parameters $V_{\max}$ and K_m.

Create these two plots in R and estimate the values of the kinetic parameters $V_{\max}$ and K_m using a linear regression curve.

Do the values found by you now correspond with the values found via the Lineweaver-Burk plot? Are these parameter values just as sensitive to the chosen initial concentration s_0 substrate?

Assignment 9: Nonlinear regression through the Michaelis-Menten formula.

Through the Lineweaver-Burk plot you have estimated the kinetic parameter values $v_{\max}$ and K_m on the basis of an estimate of s_0. Choose a value for s_0 and use the other estimated parameter values to find even better parameter values by a nonlinear regression method for the Michaelis-Menten formula

$$r = \frac{v_{\max} \cdot s}{K_m + s}$$

Assignment 10: Nonlinear regression using the differential equation.

Lineweaver-Burk, Eadie-Hofstee and Hanes-Woolf plots are in today's comptuer era only used to find appropriate initial values for kinetic parameters in a nonlinear regression method. The FME packet with its dependent R packages enables the user to fit parameters in nonlinear differential equations to measured data. Study the example in the section *Parameter estimation by differential equations* of the chapter *Using R in mathematics.* Then implement similarly a parameter estimation for the Michaelis-Menten model

$$r = \frac{V_{\max} \cdot s}{K_m + s}$$

wherein you introduce $V_{\max}$, K_m and the start concentration s_0 of the substrate as parameters to be optimized. You can also use the maximum concentration $p_{\max}$ of the product use as an alternative for s_0.

Draw in a single diagram, the solution of the differential equation for the concentration p of the product using the chosen initial values of the parameters according to the Lineweaver-Burk plot, and the solution at the optimized parameter values, with the measured data in the background.

Compare your calculated parameter values with those of the authors of the article where the data come from (Moss et al, 1996). These authors report the following results::

$$s_0 = p_{\max} = 31 \text{ nM}, \quad V_{\max} = 1.86 \pm 0.12 \text{ nM/min}, \quad K_m = 20.5 \pm 3.0 \text{ nM}$$