16.9.2024

Exercises for Module BIO03

Solutions in red

In some of the exercises, you will need to try and find some of the required information online. The exercise will require you to make some assumptions and estimates. This approach is a preparation for the quantitative analysis you need to do in the project report.

Question 1: Energetics of enzyme catalysis
(Question courtesy of Prof. Peter Westh)
We consider a simple reaction where a substrate S is converted to a product P. In other words, the net reaction is

$$S \xrightarrow{k_{uncat}} P$$

The kinetics of this reaction is governed by the first order rate constant k_{uncat}. Experiments have suggested that the following mechanism applies for the enzyme catalyzed conversion of S to P

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} EP \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} E + P$$

The experiments also revealed an energy diagram for this enzyme reaction (Fig. 1).

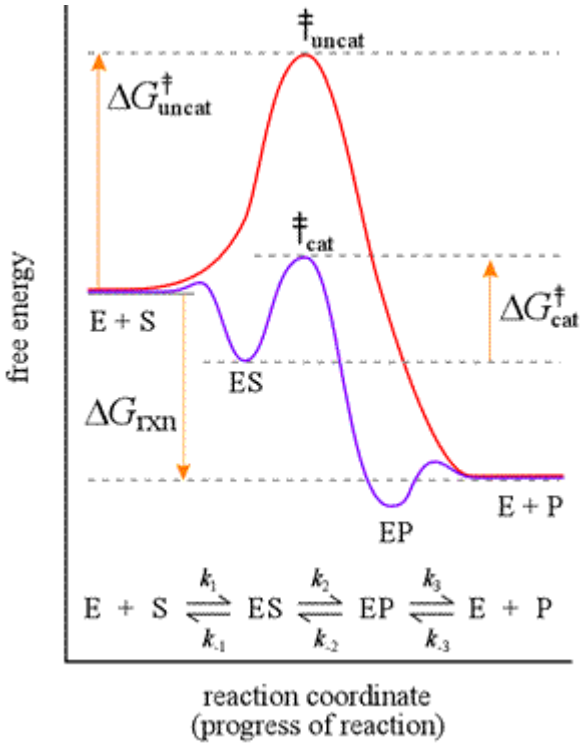

Fig. 1. Energy diagram for enzyme reaction.

1. Identify the rate limiting (i.e. slowest) step for the enzymatic conversion of S to P (explain your answer).

Enzymes always catalyze the reaction in both directions, and this is indicated by the double-arrows in the above mechanism. In this problem, we consider both directions of the second step.

2. Which rate constant is larger, k_2 or k_{-2}? (again, you need to explain your answer).
3. Approximately how much larger is the larger of the two rate constants, if $\Delta G^{\ddagger}_{cat}$ in the scheme above is 10 kJ/mol?

Solution:

1. In the forward direction (i.e. S->P) the activation barrier for the step ES->EP is far larger than the other activation barriers. Hence this is the rate limiting step (and $k_{cat} \sim k_2$)
2. $k_2 > k_{-2}$. To rationalize this, we note that the step in free energy from EP to the transition state is far larger than the step from ES to the transition state. It also appears as DG for the ES->EP step is negative. It follows that the equilibrium constant – and hence the ratio of rate constants k_2/k_{-2} is larger than 1.
3. The energy barrier for k_{-2} is approximately 2.5 times higher than that for k_2. This is purely by estimate, not measured. The rate constant is exponentially modulated by the free energy barrier, according to:

   $k \propto e^{-\frac{\Delta G^{\ddagger}}{RT}}$. $\Delta G^{\ddagger}_{cat}$ corresponds to $\Delta G^{\ddagger}_{2}$. Therefore we have

$$\frac{k_2}{k_{-2}} = e^{\frac{\Delta G^{\ddagger}_{-2} - \Delta G^{\ddagger}_{2}}{RT}} = e^{\frac{25\frac{kJ}{mol} - 10\frac{kJ}{mol}}{RT}} \approx e^{\frac{15\frac{kJ}{mol}}{2.5\frac{kJ}{mol}}} = e^{6} \approx 400$$

   Therefore k_2 is ca. 400 times faster than k_{-2}.

Question 2:

Entropic elasticity of biomolecules. In the lecture (slide 6) an expression is given for the force that is needed to generate a certain end-to-end distance of a (bio-)polymer. Use this formula to answer the following question.

a) A piece of dsDNA (persistence length 50 nm) of 10 µm total length ('contour length') is brought from an average end-to-end distance of 0 (at equilibrium) to an end-to-end-distance of 5 µm. This requires a force of f = -6e-14 N = -60 fN. What force is needed to stretch a piece of ssDNA (persistence length of 2.2 nm) of the same contour length by the same amount?
b) What happens when the temperature is increased in such a stretching experiment?

Solution:

a) The Kuhn length is twice the persistence length and the full chain can be thought of as being a sequence of independent segments of a length equal to the Kuhn length. Therefore the 10 µm long piece of dsDNA can be though of as a polymer of 100 independent Kuhn segments of 100 nm length each. WE are told that the force needed to stretch this DNA molecule such that its two ends are on average 5 µm apart is 60 pN. The equation to be used in this case is the following:

$\vec{f}$=-$k_B T \frac{3\vec{R}}{Nl^2}$

The denominator Nl^2 changes from 1e-12 m^2 for dsDNA to ca. 4.4e-14 m^2 for ssDNA. This means that the force required for pulling the ssDNA by the same distance is ca. 23 times larger than for dsDNA. This result may be counter-intuitive, as one might think that one should need more force to stretch the stiffer polymer (dsDNA). However, the less stiff polymer (ssDNA) has a much higher conformational entropy at equilibrium, and stretching it is therefore associated with a much higher entropy loss than in the case of the stiffer polymer.

b) The force constant increases with temperature and therefore stretching by the same amount at a higher temperature requires higher force.

Question 3:
Enzymes are astonishing molecular machines and in some cases are able to catalyze reactions at ambient conditions that we are only able to implement as complex and energy-intensive industrial processes. Engineering enzymes and incorporating them into our industrial processes is therefore a major aim of biotechnology and bioengineering. In this exercise, we will discuss two enzymes that are able to catalyze reactions that we are extremely interested in: capture of carbon dioxide (CO_2) and capture of nitrogen (N_2) from air. CO_2 capture is relevant in the context of mitigating climate change and capturing N_2 is a requirement for fertilizer production.

Let us start with the enzyme *rubisco*, the most abundant enzyme in the world, which is the main enzyme responsible for converting carbon dioxide, CO_2, into energy-rich molecules in plants.

a) What kind of reaction does carbon dioxide conversion into energy-rich carbon compounds, such as sugars correspond to, and why is it considered a difficult reaction? (Hint: consider the oxidation state of the carbon)
Carbon dioxide is the most strongly oxidized from of the element carbon, with all 4 valance electrons formally belonging to the oxygens. Because oxygen attracts the electrons strongly, CO_2 is a very stable molecule. Its conversion to other carbon compounds, such as sugars corresponds formally to a reduction. Oxidation state of carbon in CO_2 is +4 and in a typical sugar, it is 0. Therefore carbon takes up electrons and is reduced.

b) Rubisco is a very slow enzyme, reflecting the energetic difficulties associated with carbon dioxide fixation. Each molecule can only convert ca. 3 molecules of CO_2 per second. How long (order of magnitude) does it take for a rubisco protein molecule to fix enough carbon for the production of another rubisco molecule? How does this time compare to a typical half life of rubisco in a leaf (https://bionumbers.hms.harvard.edu/bionumber.aspx?id=107781)?

Typical rubiscos contain more than 1000 carbon atoms per subunit, which translates into several minutes during which the rubisco simply fixes the carbon needed for its successor. The half life is ca. 7 days, so this calculation shows that the rubisco most of the time (99.9%) fixes carbon for other uses, which makes it a very resource-efficient machine.

c) Rubisco has a low catalytic rate, k_{cat}, a low affinity for its substrate CO_2 and a low specificity for its substrate compared to another possible substrate, O_2. Why did Rubisco not evolve to be highly selective for CO_2 compared to O_2?

When photosynthesis and rubisco first evolved, there was no oxygen in the atmosphere, and therefore no need to discriminate between oxygen and carbon dioxide as substrates.

d) Much efforts are being undertaken to engineer improved rubisco enzymes, because rubisco's low turnover can be a serious bottleneck in carbon fixation and therefore biomass generation of plants. Many different approaches are being tried. In a recent study (Lin *et al*., Science Advances 2022), the authors re-constructed ancient rubiscos based on phylogenetic trees and among those rubiscos that occurred during periods of higher CO_2 concentration on earth, several were found with improved kinetic characteristics. What do you think are the hopes associated with better performing rubiscos? What consequences could they have for world food supply and climate change? Why are natural rubiscos not evolving to exploit anthropogenic increased CO_2 levels?

The main hopes are to create faster growing crops, which will contribute to increased food supply. It is not clear if climate change would be affected, even if plants with improved rubiscos were to be cultivated/released at scale. Faster photosynthesis does not necessarily lead to higher overall biomass at steady state, Perhaps such plants could be useful to re-plant zones of forest fires etc., to quickly repair the damage.
Anthropogenic CO_2 concentration change happens too rapidly on an evolutionary time scale for many higher organisms. Adaptation takes thousands to millions of years, whereas human-made climate change happens on a time scale of tens to hundreds of years.

Next, we will study the enzyme (or rather enzyme complex) nitrogenase, which is the only enzyme that is able to convert atmospheric nitrogen, N_2, a very inert gas, into the reduced form ammonia, NH_3. Nitrogen is only soluble as reduced or oxidized (e.g. nitrate NO_3^-) compound, and plants need soluble nitrogen to be able to take it up. Nitrogenase is used by microorganisms that live in close symbiosis with plants and provide them with soluble nitrogen in exchange for nutrients.

e) The Haber-Bosch process is, still today, the major process by which humanity fixes nitrogen and makes it available for agriculture, according to:

$N_2 + 3H_2 \rightarrow 2NH_3$

The process uses an iron-based catalyst (Mittasch catalyst) and has an energy consumption of more than 30 GJ/ton of NH_3 and 1% of total annual world energy production is used for this process. The hydrogen for the process comes in large parts from natural gas (methane), leading to large CO_2 emissions during the process. Given that the formation of ammonia from nitrogen and hydrogen is energetically favorable (the reaction being both exothermic and exergonic), why are such large energies needed to produce ammonia? Why is the reaction carried out at high temperature and what is problematic about that, from a thermodynamic point of view?
The energy is needed to help the reagents overcome the very high activation energy, which is mostly needed to break the stable nitrogen-nitrogen triple bond. Carrying out the reaction at high temperature therefore helps the system to overcome the activation barrier. However, an exothermic reaction (one that releases heat) becomes less favorable if the temperature is increased, according to the principle of Le Chatelier and Braun. The equilibrium position therefore becomes less and less favorable as temperature is increased. Without a catalyst, the reaction would have to be carried out at such high temperatures that basically no ammonia would be formed because the equilibrium would be shifted very much towards the reagents. The actual temperature of 400-500ºC is a compromise between the contradictory requirements of reaction rate and yield.

f) Biological Ammonia production by microorganisms occurs at ambient conditions of pressure and temperature (albeit in the absence of oxygen), catalyzed by the enzyme nitrogenase, according to the following equation:

$N_2 + 8H^+ + 16\ MgATP + 8e^- \rightarrow 2\ NH_3 + H_2 + 16\ MgADP + 16\ P_i$

Where MgATP/MgADP signify the magnesium salts of adenosine triphosphate/diphosphate and P_i is a phosphate ion. The necessary 8 electrons (e^-) are shuttled by several other proteins towards the catalytically active center of the main nitrogenase component in which the reduction of nitrogen happens.
The formation of 2 molecules of ammonia therefore requires 16 molecules of ATP. Assuming an energy release per ATP molecule of ~12 k_BT, compare the energy efficiencies of the Haber-Bosch process with that of biological ammonia production.
30 GJ/ton of ammonia corresponds to 510 kJ/mol of ammonia. 8 molecules of ATP per molecule of ammonia corresponds to 240 kJ/mol of ammonia ($k_BT \approx$ 2.5 kJ/mol at ambient conditions). Therefore biological ammonia fixation is about twice as energy efficient as industrial fixation by the Haber-Bosch process, and obviously does not obtain its energy requirements from fossil fuels.

g) It is interesting to consider that industrial and enzymatic nitrogen fixation have one thing in common: they both involve an iron catalyst. In the case of the Haber Bosch process, it is the Mittasch catalyst and in the case of nitrogenase it is the MoFe (molybdenum-iron) cluster in the main enzyme. Let us compare the efficiencies in terms of catalytic activity per unit mass of the two types of catalysts. It turns out that both are actually numerically quite similar, at ca. 30 000 µmol of ammonia per g of catalyst per hour (see Rapson *et al.,* Catalysts 2022). How much catalyst/nitrogenase do you need if you want to produce 1 ton of ammonia per hour? Consider that the industrial catalyst is only active at the surface of catalyst particles and the enzyme only at the active center. How would you increase the efficiency of the catalyst/nitrogenase per unit mass?
1 ton of ammonia is ca. 59 000 moles. At a rate of 30 000 µmol = 0.03 mol per hour per gram of catalyst, we would need ca. 2 tons of catalyst to produce 1 ton of ammonia per hour. Making the catalyst particles smaller would be a way to increase efficiency of the Mittasch catalyst, but there are limits to that, because at the high temperatures of the process, sintering could be a problem. In some sense, the nitrogenase enzyme has pushed this strategy to the absolute limit by using iron/molybdenum nanoparticles. However, the rest of the enzyme around is necessary to stabilize the catalytic centres and create a conducive environment for the reaction, and decreases the nominal efficiency per mass unit. It will be hard to change the mass significnatly, so the only way to improve the efficiency will be to increase the catalytic rate constant, e.g. by directed evolution of the enzyme.

h) Now we are trying to estimate what fraction of the ammonia produced for fertilizers ends up actually being used for human food. Assume for this calculation that protein is the exclusive source of nitrogen in our food. Approach this question systematically and try and find the relevant information online to perform this estimate.
Yearly ammonia production is of the order of 200 million tons, 80% of which is used for fertilizer. This corresponds to roughly 130 million tons of pure nitrogen. The average daily consumption of protein per capita is 83 g (http://www.anthroponumbers.org/catalog/search/?q=protein ; this website is curated by Rob Phillips from Caltech and is an excellent source of numbers related to human impact on the planet), corresponding to ca. 250 million tons of protein per year. Average nitrogen content of protein by mass is 16%, corresponding to 40 million tons per year. Therefore a quarter of the total nitrogen used as fertiliser ends up in human food. Much potential to improve efficiency!

Question 4: Neutralisation of venom toxins with the help of antivenom
(Question courtesy of Prof. Andreas Laustsen)
Snake venoms are highly complex biochemical cocktails that contain a variety of toxic proteins (toxins). To neutralize them, one needs to understand many different variables, such as how toxic the overall venom is, how much venom is injected, and how many injected toxin molecules does this translate to.

You are on safari in Africa and, whilst trying to get the perfect photograph of a majestic lion in the distance, you accidentally step on a Puff Adder (*Bitis arietans*) and are bitten. The Puff Adder can inject up to 290 mg of toxin per bite and has an LD_{50} (median lethal dose, i.e. at this dose 50% of test animals/subjects die) of 0.389 mg/kg, which means that without antivenom, your chance of survival is low.

a) How much weight would you have to gain, if you weighed 80 kg, in order to increase your chance of survival to above 50%?
$\frac{290\ \text{mg}}{0.389\ \text{mg/kg}} = 745.5\ \text{kg}$; 745.5 kg – 80 kg = **665 kg**

b) How many people weighing 80 kg could this bite theoretically kill with a probability of 50%, and how many mice (25 g)?
$\frac{745.5\ \text{kg}}{80\ \text{kg}} = 9.32$; $\frac{745.5\ \text{kg}}{0.025\ \text{kg}} = 29{,}820$; This bite could kill 9 people and 29,820 mice

c) Name at least one animal that you can see on your safari large enough to survive this bite.
Many animals work, but e.g. a Hippo (3750kg), white Rhino (2175kg), or a giraffe (800kg).

Fortunately, you make it to the hospital, but since antivenom is so expensive, the doctors want to know how much antivenom to administer. Assuming that antivenom doses for treatment are based on the collective average masses of toxins, and that from this, the mass of antivenom required for neutralizing is calculated. The doctors know the following:

- The venom composition of the Puff Adder:

Molecular weight	SVMPs	SVSPs	PLA_2s	Disintegrins	Lectins	Non-relevant
kDa	50	30	14	9.5	28	NA
Composition (%)	38.46	19.48	4.3	17.8	13.2	6.76

- Normally the Puff Adder only injects 50% of its maximum venom yield
- Number of antibodies needed per toxin to be neutralized = 4
- MW of the IgG antibodies 150 kDa or 150,000 g/mol
- A single vial of antivenom contains 1 gram of antibodies

d) How many vials of antivenom should the doctors give you?

Calculate M_{venom}:

$$\frac{50 \times 38.46 + 30 \times 19.48 + 14 \times 4.3 + 9.5 \times 17.8 + 28 \times 13.2}{93.24} = 33.31510082\ \text{kDa}$$

Convert M_{venom} from kDa to g/mol:

$$33.31510082 \times 1000 = \mathbf{33315.10082}\ \frac{\mathbf{g}}{\mathbf{mol}}$$

Calculating the amount of antibodies required for neutralization:

$$\frac{0.29\ \text{g} \times 0.9324 \times 0.5}{33315.10082\ \text{g/mol}} \times \frac{150000\ \text{g}}{\text{mol}} \times 4 = 2.43\ \text{g}$$

Since there is 1g of antibody in a single vial the doctors need to administer at least three vials.

Question 5: Creation of improved proteins

a) You are joining Novozymes as a new enzyme engineer, and you are tasked with improving a lipase (i.e. fat-degrading) enzyme that is added to washing powder. The aim is to make this enzyme work faster at room temperature, so that the washing can be done in a more energy efficient manner. Describe briefly the most suitable approach to solve this problem.

You would likely be using **directed evolution**. You start with the enzyme that is active at a higher temperature and then you prepare a large library of sequence variants. Then you need an assay to evaluate the efficiency of the enzymes at high throughput and you simply check whether any members of your library are active at room temperature. Then you take those promising candidates and repeat the process until you have found an enzyme that fulfils the requirements.

b) You are joining Novo Nordisk as a new antibody engineer, and you are tasked with creating an antibody that binds to human insulin at pH 4 with high affinity. You know that there is an antibody that binds to insulin at pH 7 with high affinity, but only weakly at pH 4. What do you do?

Here you would likely use **phage display** to improve the binding affinity at pH 4. You start with the antibody that binds to insulin at pH 7 and you prepare a large library of sequence variants of that antibody. Then you incorporate this library into the phages and then you do your selection experiment. In this case, you coat your plate with insulin and you wash it at pH 4, so that only those phages remain bound that have an antibody on their surface that is able to bind at pH 4. You repeat the variation and amplification steps and at the end you sequence the phages that bind more and more strongly to the target under the desired conditions.

c) Assume that the binding region on the antibody contains 25 amino acids. How big would your phage library have to be if you want it to contain all possible single point mutations in this region. What about if you want all possible double mutants?

The number of all possible single point mutations is 25x19=475, because at every one of the 25 positions, 19 other amino acids can be placed.
The number of all possible double mutations is 19x19x24x25/2=108300
This is because there are 25x24/2 different pairs (the factor of 2 is to avoid double counting), and every pair has 19x19 double mutant combinations.