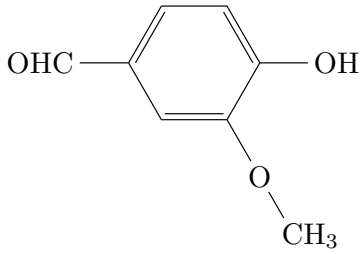
Exercise sheet 2: Naming molecules

Organic chemistry for bioinformatics, lecture 2. Exercises marked *(bonus)* are optional.

- **Use the handouts** *Functional groups* and *Nomenclature: practice and learn*: they contain all the rules you need, with worked examples.
- **Exercise 4 is self-study.** It introduces SMILES, which we did not cover in the lecture. Read the box at the start of exercise 4, then check every answer yourself in PubChem.

Exercise 1: Functional groups

This is **vanillin**, the main aroma of vanilla. Vanilla pods are the fruits of an orchid, and the flowers have to be pollinated by hand, which makes vanilla one of the most expensive spices. That is why most of the vanillin in food is not taken from vanilla at all: it is made from guaiacol or from lignin, a by-product of paper making.

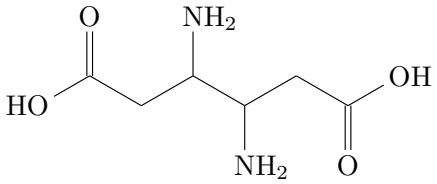


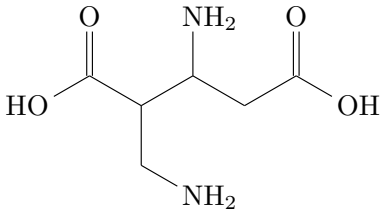
a) Circle all functional groups and name their classes.

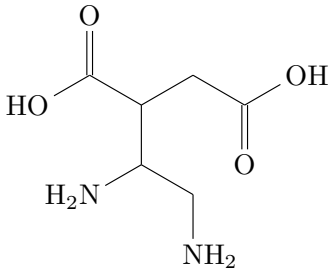
b) Which group has the highest priority? Which ending (suffix) would the name get?

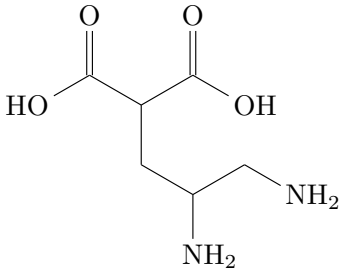
Exercise 2: One formula, many names

The "name" *dicarboxy-diaminobutane* seems to describe a butane with two COOH and two NH_2 groups. Here are four molecules A–D that fit it, and their four IUPAC names 1–4:

A 

B 

C 

D 

1 2-(2,3-diaminopropyl)propanedioic acid
2 3,4-diaminohexanedioic acid
3 2-(1,2-diaminoethyl)butanedioic acid
4 3-amino-2-(aminomethyl)pentanedioic acid

a) Pair each name 1–4 with its structure A–D.

b) Verify that all four molecules have the same molecular formula. Which one is it?

c) Mark the parent chain in each structure. Why does its length vary between the four molecules, although all four are "butanes with two COOH groups"?

d) Why is *dicarboxy-diaminobutane* not a correct name? Give two reasons.

Exercise 3: Find the mistake

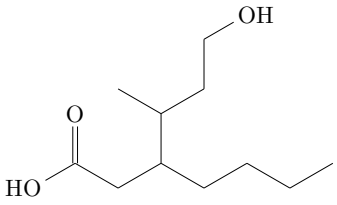
In a) and b), each name leads to the right structure, so it passes our test from the lecture (structure = structure′), but IUPAC writes it differently. Draw the structure, find the rule that is broken, and give the IUPAC name.

In c), a student gave the molecule below a name. Draw the structure that belongs to this name. Is it the same molecule? Find the mistake and give the correct name.

a) 3,3,2-trihydroxypropanal

b) 1-methylhexa-3,5-dien-1-amine

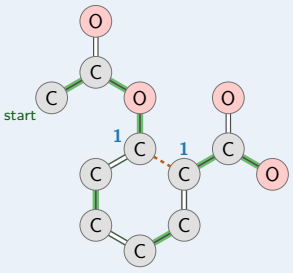
c) *3-(4-hydroxybutyl)heptanoic acid* for this molecule:



Exercise 4: SMILES and PubChem (self-study, check yourself)

SMILES in brief (Simplified Molecular Input Line Entry System)

1. Leave out the H atoms. Atoms are nodes, bonds are edges.
2. Break each ring once (one cut per ring) to get a spanning tree; mark both ends of the cut bond with the same digit.
3. Walk along the longest path and write the atoms (`C`, `N`, `O`, `S`) and bonds: single bond: nothing, double `=`, triple `#`. Side branches go in brackets `( )`.
4. cis/trans at a double bond: `/` and \ before the atoms on each side; `C/C=C/C` is trans, `C/C=C\C` is cis.



Example (aspirin): start at the CH_3, follow the green path, cut the ring at the dashed bond (digit 1):

`CC(=O)OC1=CC=CC=C1C(=O)O`

Many SMILES are correct for one molecule. To check yours, paste it into the search field of **pubchem.ncbi.nlm.nih.gov**: if PubChem shows the right molecule, your SMILES is right. The **PubChem Sketcher** (`pubchem.ncbi.nlm.nih.gov/edit3`) draws a SMILES for you.

a) Write a SMILES for vanillin (exercise 1) and check it in PubChem.

b) Draw the structure for `CC(=O)C(=O)O` and name it. Check in PubChem. (You will meet this molecule again in lecture 3.)

c) Do you recall exercise 6 on exercise sheet 1? Write the SMILES of the trans form of HO–CH_2–CH=CH–C≡CH. Find it in PubChem and open its 3D structure (section *3D Conformer*). Rotate it and check your answers from sheet 1: which atoms are tetrahedral, planar or linear? Which atoms lie in one plane?

d) *(bonus)* Write the SMILES of the cis form, too, and open its 3D structure in PubChem. What is the same in cis and trans, what is different?