Exercise in QSAR – determination of dependence of anti-fungal or antibacterial activity of parabens against *Aspergillus niger* or *Staphylococcus aureus* respectively; or toxicity of parabens on their lipophilicity expressed by various parameters and also on other structure parameters

•repeat the chapter about physico-chemical properties and mainly QSAR from MC I

We are looking for a dependence = a function

log (1/MIC) = f (parametr´) in the form log (1/MIC) = ax + b or log (1/MIC) = $ax^2 + bx + c$

a) log k´ from RP-HPLC: stationary phase: a column eg. C18; mobile phase e.g. methanol : water 6:4, stream 0.6 ml/min., injected 20 μl of mixture of parabens of concentration 0.1 mg/ml each, also containing a compound for dead retention time determination t_0 (eg. acetone)

$$\log k' = \log\left(\frac{(t_r - t_0)}{t_0}\right),$$

b) further hydrophobic, electronic, steric and other parameters either experimentally determined (from the literature) or calculated: log P for various systems, log D at given pH, molar volume, Taft steric constant E_s, parachor, molar refractivity, Dreiding energy...

Calculation & results

- log P_{exp}: ChemId; primary literature resources (=articles, patents); for a chance to acquire a good correlation, it´s desirable to have values for all the compounds from the same resource (a journal article) when it is assumed that they have been determined in the same way
- log P_{calc}: ChemSketch, Marvin Beans
- molar volume: ChemSketch
- other calculated parameters
- with values R_m , log P, and log k´ and others, perform a **linear regression** in MS Excel or an equivalent, or in a suitable statistics software; where
  - the activity represented by log (1/MIC) is a **dependent quantity;**
  - log P, R_m , log k´etc. - **independent quantity**
- compose the **equation** from resulted values, present also the linear correlation coefficient R, and the value of Fisher-Snedecor test F
- input data in the table and results into the report
- upload the report into the subject homework vault **QSAR of parabens** in IS **until May 15th**

Individual tasks:

1. log P from ChemId vs. activity against *A. niger*; calculate activity of hexylparaben (MIC) provided that the linear dependence also applies in the homologous series after pentyparaben https://chem.nlm.nih.gov/chemidplus/
2. log P from the article Kobayashi Y. et al.: Eur J Pharm Sci **21**, 471-477 (2004) vs. activity against *A. niger*; calculate activity (MIC) of hexylparaben asuming the linear dependence also applies in the homologous series after pentyparaben.
3. log k´ from RP-HPLC on C_{18} reverse phase vs. activity against *A. niger.* Peak assignation: 1 - t_0, 2 – methylparaben, 3 – ethylparaben, 4 – isopropylparaben, 5 – propylparaben, 6 – *iso*butylparaben, 7 – *sec*butylparaben, 9 – butylparaben, 10 – pentyparaben. Assume that the linear dependence also applies for branched esters, and calculate their MICs.
4. logP calculated in Marvin Beans by the "Consensus" algorithm vs. activity against *S. aureus*; calculate activity (MIC) of heptylparaben.
5. logP caculated in ChemSketch vs. activity against *S. aureus*; calculate activity (MIC) of undecylparaben.
6. Molar volume calculated in ChemSketch vs. activity against *S. aureus*; if a linear dependence is demonstrated ($R > 0.75$), calculate MIC of nonylparaben.

Procedure of calculation in Excel (no problem with use of other software)

- data of activity and a structure parameter along into two adjascent columns of the table
- select a rectangle of 2 x 4 cells at other place other than the input data are
- select the function LINEST (linear regression)
- select the column of the **structure parameter** as **data x**
- column of **activity** as **data y**
- type B = true, Stat = true
- “perform calculation”(OK)
- you´ll get the results in the table placed likewise in the table at the bottom, if not, the value **a** appears in the top left corner cell of the selected rectangle; press F2, and then CTRL+ SHIFT + ENTER
- now, the required results should appear; you can compose now the equation in the form y = ax + b (eg. log (1/MIC) = a.logP + b); r^2 … square of the linear correlation coefficient; s_y … standard deviation of the estimate, F … Fischer-Snedecor test value

a	**b**
s_a	s_b
$\mathbf{r^2}$	$\mathbf{s_y}$
F	d_f

Procedure of linear regression calculation in LibreOffice Calc

•similar as in Excel but 3 individual functions must be step-wisely used:

- SLOPE for calculation of **a**
- INTERCEPT for calculation of **b**
- CORREL for the correlation coefficient **r**

Marvin Sketch

- available from https://chemaxon.com/products/marvin/download#download
- the free license needs to be renewed every 2 months
- for Windows, Mac, Linux (Debian, RedHat)

Calculations in Marvin

- draw the structure
- 2D optization (Structure-Clean 2D)
- Calculations
- Partitioning-logP-OK-OK(-logD-OK-OK)
- Charge-Polarizability
- Geometry-Geometrical Descriptors-OK-OK
- Other-Refractivity-OK-OK

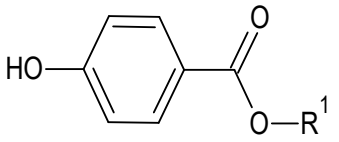
ACD ChemSketch

- available for free 30days trial from webstore.acdlabs.com
- registration required

Calculations in ChemSketch

- log P: draw the structure, select Log P right on the top bar
- other parameters: draw the structure, select Tools – Calculate – (appropriate parameter) on the top bar

Values of log(1/MIC) of parabens against *Aspergillus niger* and *Staphyloccoccus aureus*



R^1	log (1/MIC) *A.niger*	log (1/MIC) *S.aureus*	Structure parameter
CH_3	-0,11	-2.818	
C_2H_5	0,12	-	
C_3H_7	0,33	-2.745	
C_4H_9	0,56	-2.411	
C_5H_{11}	0,57	-	
C_8H_{15}	-	-2.301	
$C_{10}H_{21}$	-	-0.748	
$CH_3(CH_2)_5CH(OH)(CH_2)_2$	-	-2.230	
$(CH_3)_2CH(CH_2)_8$	-	-1.932	
$(CH_3)_2CH(CH_2)_5CH(OH)(CH_2)_2$	-	-1.909	
$C_{12}H_{25}$	-	-2.213	
$CH_3(CH_2)_5CH=CH(CH_2)_4$	-	-1.613	

Upload your report into the homework vault before May 15th.