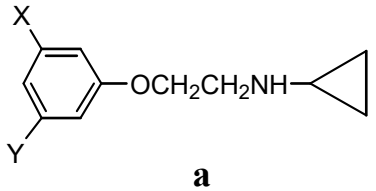
# CM5239 Tutorial – QSAR and 3D-QSAR

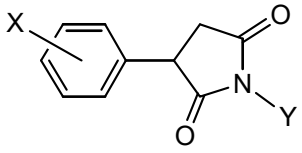
**(1)** The following QSAR equation was derived for the activity of monoamine oxidase inhibitor (**a**). Explain what the various terms mean and whether the equation is a valid one. Identify what kind of substituents would be best for activity.

$$\log (1/C) = 0.40\ \pi_x + 1.09\ \sigma + 1.03\ E_s + 4.54$$

$$(n = 12,\ r^2 = 0.96)$$




**a**



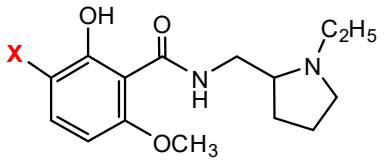

**b**

**(2)** The isonarcotic activities of two classes of compounds, (1) ketone, ester and ether and (2) alcohol in tadpoles are given below:

| **(1) ketones, esters, and ethers** | | | **(2) alcohols** | | |
|---|---|---|---|---|---|
| compound | log P | log(1/C) | compound | log P | log(1/C) |
| $CH_3COCH_3$ | –0.73 | 0.65 | $CH_3OH$ | –1.27 | 0.30 |
| $CH_3CO_2CH_3$ | –0.38 | 1.10 | $C_2H_5OH$ | –0.75 | 0.50 |
| $C_2H_5COCH_3$ | –0.27 | 1.10 | $(CH_3)_2CHOH$ | –0.36 | 0.90 |
| $HCO_2C_2H_5$ | –0.38 | 1.20 | $(CH_3)_3COH$ | 0.07 | 0.90 |
| $C_2H_5COC_2H_5$ | 0.59 | 1.20 | $C_3H_7OH$ | –0.23 | 1.00 |
| $CH_3CO_2C_2H_5$ | 0.14 | 1.50 | $(CH_3)_2C(C_2H_5)OH$ | 0.59 | 1.20 |
| $C_2H_5COC_3H_7$ | 0.31 | 1.50 | $C_4H_9OH$ | 0.29 | 1.40 |
| $C_3H_7COCH3$ | 0.31 | 1.70 | $(CH_3)_2CHCH_2OH$ | 0.16 | 1.40 |
| $CH_3CO_2C_3H_7$ | 0.66 | 2.00 | $C_5H_{11}OH$ | 0.81 | 1.60 |
| $C_2H_5CO_2C_2H_5$ | 0.66 | 2.00 | | | |
| $(CH_3)_2CHCO_2C_2H_5$ | 1.05 | 2.00 | | | |

Derive the QSAR equations, $\log(1/C) = k_1\log P + k_2$, from the two sets of data using Excel. Determine the value of statistical term $r^2$ of both equations. Which class have a stronger correlation? [If you are not familiar how to perform a linear regression using Excel, please view the following YouTube video: https://www.youtube.com/watch?v=L_a8Z0BVjyM]

**(3)** The influence of the aromatics substituent X in a series of monosubstituted salicylamides on their ability to displace [$^3$H]-spiperone (measures by $IC_{50}$ values ion nM) leads to the equation:




salicylamide derivative

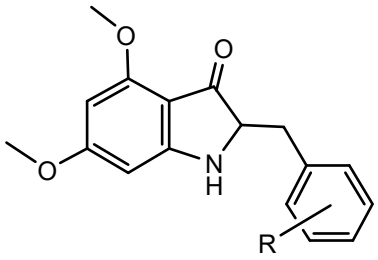
$$\log (1/IC_{50}) = 1.28\ \pi - 0.52\ \pi^2 + 1.54$$

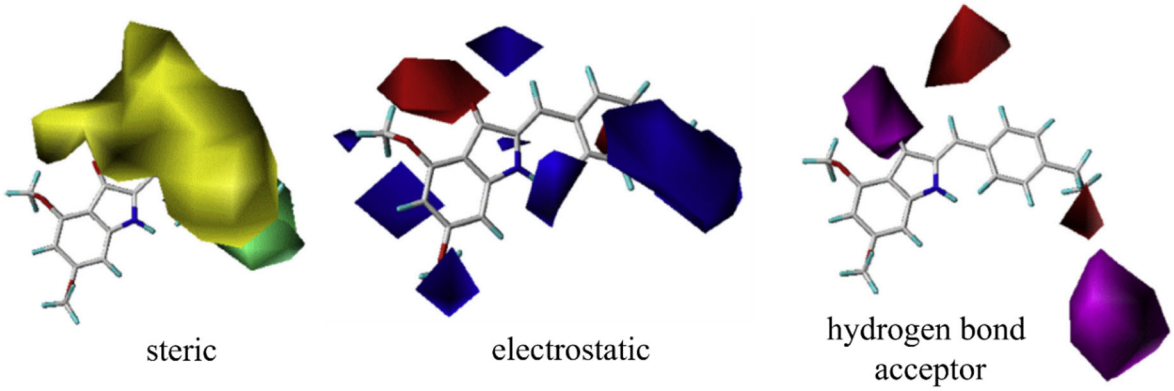
$$(n = 15,\ r^2 = 0.94,\ s = 0.40)$$

From the equation and table below, what is the best value for π, the corresponding $IC_{50}$ values, and the best substituent X.

| **substituent** | **H** | **F** | **Cl** | **Br** | **I** | **$NO_2$** | **$CH_3$** | **$C_2H_5$** | **$C_3H_5$** | **$C_3H_7$** |
|---|---|---|---|---|---|---|---|---|---|---|
| π | 0.00 | 0.14 | 0.71 | 0.86 | 1.12 | –0.28 | 0.56 | 1.02 | 1.14 | 1.55 |

**(4)** A 3D-QSAR study based on CoMSIA model was carried out for a test set of 35 aurone analogues (antimalarial drug candidate) acting as Qo site inhibitors at the cytochrome b. Excellent predictive power was observed for the test set. The CoMSIA contour maps (steric, electrostatic and hydrogen bond acceptor fields) with the most active compound (R = para-$CH_2CH_3$) of the series as a reference are shown below.



steric electrostatic hydrogen bond acceptor

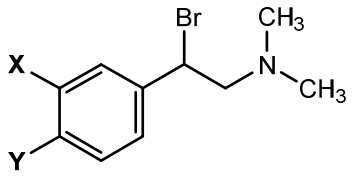
(a) Which activity information (field effects) can one derive from these colored contour maps?

(b) Predict the effect of bulky substitution at the *para* position on the biological activity.

(c) What statistical parameter can be used to quantify the significant statistical quality?

(d) What is a suitable method to confirm the predictive power by external validation?

**(5)** Based on the following enhancement table, construct a QSAR equation using Free-Wilson analysis.




| **position** | **F** | **Cl** | **Br** | **I** | **$CH_3$** |
|---|---|---|---|---|---|
| *meta* | -0.30 | 0.21 | 0.43 | 0.58 | 0.45 |
| *para* | 0.34 | 0.77 | 1.02 | 1.43 | 1.26 |

Which combination of *meta* and *para* substituents yields the highest predicted activity?

**(6)** The adrenergic blocking activity of β-halo-arylamines was related to two physiochemical parameters through the following QSAR equation.

$$\log (1/C) = 1.22\ \pi - 1.59\ \sigma + 7.89$$

From the following Craig plot of π verse σ, suggest 2 substituents which may yield high predicted biological activity. Which other physiochemical parameter would you consider to improve the QSAR model?