3D PHAMARCOPHOMORE MODELLING PART II

OBJECTIVES OF QSAR

Mostly all the QSAR methods focus on the following goals:

1. Quantitative relationship between the structure and physiochemical properties of substances and their biological activity are being used as the foundation stone in search of new medicines. The mathematical and statistical analysis helps us to predict the drug activity.
2. QSAR makes it easy now to reach the conclusion for any of the congener that still not in process, in way that whether it will optimal and profitable or not.
3. To quantitatively correlate and recapitulate the relationships between trends in chemical structure alterations and respective changes in biological endpoint for comprehending which chemical properties are most likely determinants for their biological activities.
4. To optimize the existing leads so as to improve their biological activities.
5. To predict the biological activities of untested and sometimes yet unavailable compounds.

TECHNIQUES AND TOOLS OF QSAR

1. Compound Selection: In setting up to run a QSAR analysis, compound selection is an important angle that needs to be addressed. One of the earliest manual methods was an approach which involves two-dimensional plots of important physicochemical properties. Care is taken to select substituents from all four quadrants of the plot. The Topliss operational scheme allows one to start with two compounds and construct a potency tree that grows branches as the substituent set is expanded in a stepwise fashion. Topliss later proposed a batchwise scheme including certain substituents such as the 3,4-Cl2, 4-Cl, 4-CH3, 4-OCH3, and 4-H analogs.

2. Biological Parameters: In QSAR analysis, it is vital important that the

biological data be both accurate and precise to develop a meaningful model. The equilibrium constants and rate constants that are used extensively in physical organic chemistry and medicinal chemistry are related to free energy values ΔG. Thus, for use in QSAR, standard biological equilibrium constants such as Ki or Km should be used in QSAR studies. Percentage activities (e.g., % inhibition of growth at certain concentrations) are not appropriate biological endpoints because of the nonlinear characteristic of dose-response relationships. These types of endpoints may be transformed to equieffective molar doses. Only equilibrium and rate constants pass muster in terms of the free-energy relationships or influence on QSAR studies. Biological data are usually expressed on a logarithmic scale because of the linear relationship between response and log dose in the midregion of the log dose-response curve. Inverse logarithms for activity (log $1/C$) are used so that higher values are obtained for more effective analogs. Various types of biological data have been used in QSAR analysis.

PARAMETERS USED IN QSAR

1. Hydrophobicity Parameters: More than a hundred years ago, Meyer and Overton made their discovery on the correlation between oil/water partition coefficients and the narcotic potencies of small organic molecules.

1.1 Estimation of hydrophobicity: -

Hansch established a model to measure the lipophilicity in term of partition coefficient. Drug travels to the site of action that means solubility in 1- octanol that simulate the lipid membrane then it goes to via cytoplasm that is simulated by Aqueous buffer “water”. Hansch proposed the lipophilicity measurement in term of partition coefficient “P”

$P = [C]_{octanol} / [C]_{water}$

It is called “Distribution coefficient**”.**

With help of the partition coefficient, we can determine the hydrophobic (π) like the difference caused in the partition coefficient of substituted and unsubstantiated compounds is relevant to the attached new substituent in it Formula is

π = log Px - logPH

Px denotes for substituted compound by "x"

PH denotes unsubstituted "x = H"

E.g.- Consider the log P values for benzene (logP = 2.13), chlorobenzene logP = 2.84), and benzamide (logP = 0.64), Since benzene is the parent compound, the substituents constants for Cl and CONH2 are 0.71 and -1.49 respectively. Having obtained these values, it is now possible to calculate the theoretical log P value for media **Chlorobenzamide.**

1.2 Partition coefficient: - It is the ratio of concentrations of a compound in the two phases of a mixture of two immiscible solvents at equilibrium.

$$P = \frac{\text{Concentration of drug in organic phase}}{\text{Concentration of drug in aqueous phase}}$$

Hydrophobic compounds will have a high P value, whereas hydrophilic compounds will have a low P value. The hydrophobic character of a drug can be measured experimentally by testing the drug's relative distribution coefficient. Octanol is a suitable solvent for the measurement of partition coefficients for many reasons. It is cheap, relatively nontoxic, and chemically unreactive. The hydroxyl group has both hydrogen bond acceptor and hydrogen bond donor features capable of interacting with a large variety of polar groups. Despite its hydrophobic attributes, it is able to dissolve many more organic compounds than can alkanes, cycloalkanes, or aromatic hydrocarbons. It is UV transparent over a large range and has a vapour pressure low enough to allow for reproducible measurements. Varying substituents on the lead compound will

produce a series of analogues having different hydrophobicities and therefore different P values. Various substituents make to hydrophobicity. This contribution is known as the substituent hydrophobicity constant. These π value are characteristic for the substituents and can be used to calculate how the partition coefficient of a drug would be affected by adding these substituents. QSAR would allow us to predict the most promising and satisfying structures (closest to the optimum value log P_0). The substituent hydrophobicity constant is a measure of how hydrophobic a substituent is, relative to hydrogen. A positive value of π indicates that the substituent is more hydrophobic than hydrogen. A negative value indicates that the substituent is less hydrophobic. By plotting these P value against the biological activity of these drugs, it is possible to see if there is any relationship between the two properties. The biological activity is normally expressed as 1/C so a graph is drawn by plotting log 1/C versus log P values to correlate the activity and partition coefficient or hydrophobicity. In studies where the range of the log P values are ranges between 1 to 4 and a straight-line graph is obtained i.e. there is an existence of relation between hydrophobicity and biological activity.

As per the equation is

$\log (1/C) = K1 \log P + K2$

E.g., Binding of drug to serum albumin. It can be determined by their hydrophobicity. In a study of 40 compounds, they resulted in the following equation:

(i) Serum albumin binding increases as log P increases that mean hydrophobic drugs bind more strongly to serum albumin than hydrophilic drug.

(ii) It helps us to know how strongly a drug binds to serum albumin that can be important in estimating effective dose levels for that drug and drugs of similar structure and predict Some time log P Values increases over the given ranges results in decreased activity

what about drugs which are independent of cell target action like the General Anaesthetics.

These are related to the log P factor alone to operate in cell membrane only no receptor interaction. These function by entering the central nervous system (CNS) and 'dissolving' into cell membranes where they affect membrane structure and nerve function.

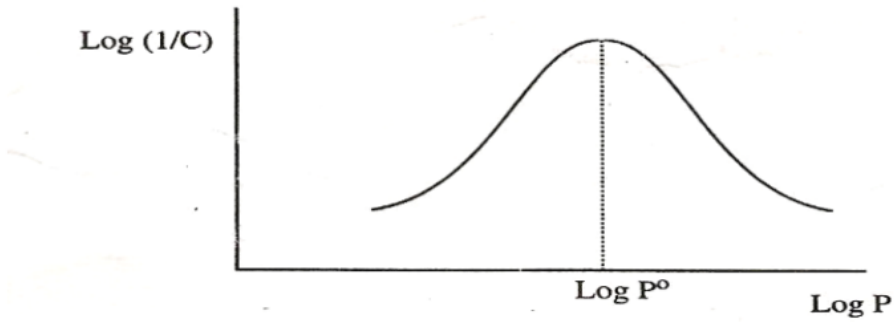
Conclusion:-

(i) Anaesthetic activity increases with increasing hydrophobicity

(ii) they depend upon lipophilicity only

(iii) There is an optimum value for log P (log P_0), beyond which increasing hydrophobicity causes a decrease in anaesthetic activity.

Finally, hydrophobic drugs are often more susceptible to metabolism and subsequent elimination. Lipophilicity has a relationship with concentration and indirectly with biological activity. It can be concluded by this graph



[Transition of linear equational effect of "log P" into parabolic equational effect "log P2"] The value of log P at the maximum (log P_0) represents the optimum partition coefficient for biological activity. Beyond that point, an increase in log P results in a decrease in biological activity.

The Shake flask method.

The shake-flask method is most commonly used to measure partition

coefficients with great accuracy and precision and with a log *P* range that extends from -3 to +6. The procedure calls for the use of pure, distilled, deionized water, high-purity octanol, and pure solutes. At least three concentration levels of solute should be analysed and the volumes of octanol and water should be varied according to a rough estimate of the log *P* value. The classical and most reliable method of log *P* determination is the, which consists of dissolving some of the solute in a volume of octanol and water, then measuring the concentration of the solute in each solvent.

Advantages:

1. Most accurate method
2. Accurate for broadest range of solutes (neutral and charged compounds applicable)
3. Chemical structure does not have to be known beforehand

Disadvantages:

1. Time consuming (>30 minutes per sample)
2. Octanol and water must be premixed and equilibrated (takes at least 24 hours to equilibrate)
3. Complete solubility must be attained, and it can be difficult to detect small amounts of undissolved material.
4. The concentration vs. UV-VIS response must be linear over the solute's concentration range.
5. If the compound is extremely lipophilic or hydrophilic, the concentration in one of the phases will be exceedingly small, and thus difficult to quantify.
6. Relative to chromatographic methods, large amounts of material are required.

Qsar method: The QSAR method involves recognition that a molecule (organic, peptide, protein, etc.) is really a three-dimensional distribution of

properties. The most important of these properties are steric (eg shape and volume), electronic (e.g. electric charge and electrostatic potential) and lipophilic properties (how polar or non-polar the sections of molecular are, usually exemplified by the log of the octanol-water partition coefficient, log P). Scientists are used to visualizing mainly steric properties of molecules. However, molecules look different when viewed in electrostatic or lipophilic space.

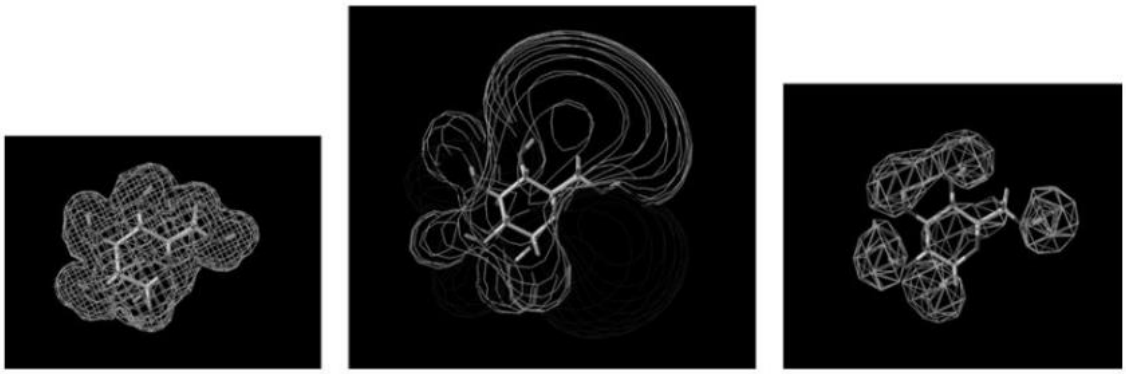


Figure: A small organic molecule (glucopyranose) viewed in steric (left), electrostatic (centre) and lipophilic (right) space

The QSAR method (and analogously QSTR And QSPR) involves a number of key steps:

1. Converting molecular structures into mathematical descriptors that encapsulate the key properties of the molecules relevant to the activity or property being modelled.
2. Selecting the best descriptors from a larger set of accessible, relevant descriptors.
3. Mapping the molecular descriptors into the properties, preferably using a model-free mapping system in which no assumptions are needed as to the functional form of the structure–activity relationship. These relationships are often complex, unknown and non-linear.
4. Validating the model to determine how predictive it is, and how well it will

generalise to new molecules not in the date set used to generate the model (the training set).

References

1. Kim, K.H., Greco, G. and Novellino, E., *A critical review of recent CoMFA applications*, In Kubinyi, H., Folkers, G., and Martin, Y.C., (Eds.) 3D QSAR in drug design: Vol. 3, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 257–316.
2. Dunn III, W.J. and Hopfinger, A.J., *3D QSAR of flexible molecules using tensor representation*, In Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 3, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 167–182.

3. Hahn, M. and Rogers, D., *Receptor surface models*, in Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 3, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 117–134.
4. Heritage, T.W., Ferguson, A.M., Turner, D.B. and Willett, P., *EVA — a novel theoretical descriptor for QSAR studies*, In Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 2, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 381–398.
5. Klebe, G., *Comparative molecular similarity indices analysis — CoMSIA*, In Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 3, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 87–104.
6. Walters, D.E., *Genetically evolved receptor models (GERM) as a 3D QSAR tool*, In Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 3, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 159–166.
7. Wade, R.C., Ortiz, A.R. and Gago, F., *Comparative binding energy analysis*, In Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 2, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 19–34.
8. Holloway, M.K., *A priori predicition of ligand affinity by energy minimization*, In Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 2, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 63–84.
9. Todeschini, R. and Gramatica, P., *New 3D molecular descriptors: The WHIM theory and QSAR applications*, In Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 2, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 355–380.
10. Silverman, B.D., Platt, D.E., Pitman, M. and Rigoutsos, I., *Comparative molecular moment analysis (COMMA)*, in Kubinyi, H., Folkers, G. and Martin, Y.C. (Eds.) 3D QSAR in drug design: Vol. 3, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 183–196.
11. Jain, A.N., Koile, K. and Chapman, D., *Compass: Predicting biological activities from molecular surface properties — performance comparisons on a steroid benchmark*, J. Med. Chem., 37 (1994) 2315–2327.
12. Martin, Y.C., Kim, K.-H. and Lin, C.T., *Comparative molecular field analysis: CoMFA*, In Charton, M. (Ed.) Advances in quantitative structure property relationships, JAI Press, Greenwich, CT, 1996, pp. 1–52.
13. Greco, G., Novellino, E. and Martin, Y.C., *Approaches to 3D-QSAR*, In Martin, Y.C. and Willett, P. (Eds.) Designing bioactive molecules: Three-dimensional techniques and applications, America Chemical Society, Washington, DC, 1997 (in press).
14. Ajay and Murcko, M.A., *Computational methods to predict binding free-energy in ligand–receptor complexes*, J. Med. Chem., 38 (1995) 4953–4967.
15. Kollman, P.A., *Advances and continuing challenges in achieving realistic and predictive*

simulations of the properties of organic and biological molecules, Acc. Chem. Res., 29 (1996) 461–469.
16. Bush, B.L. and Nachbar Jr., R.B., *Sample-distance partial least-squares — PLS optimized for many variables, with application to CoMFA*, J. Comput.-Aided Mol. Design, 7 (1993) 587–619.
17. Burger, A., *Medical chemistry — the first century*, Med. Chem. Res., 4 (1994) 3–15.
18. Willett, P., *Similarity and clustering techniques in chemical information systems*, Research Studies Press, Letchworth, 1987.
19. Hodgkin, E.E. and Richards, W.G., *Molecular similarity based on electrostatic potential and electric field*, Int. J. Quantum Chem., 14 (1987) 105–110.
20. Kier, L.B., *Molecular orbital theory in drug research*, Academic Press, New York, 1971, p. 258.