A REVIEW ON- QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIP

Dr. Shashikant R. Pattan[1], Gaurav B. Ahirrao[1], Shashikant S. Pawar[1], Sonali.B.Pawar [1]

R.K.Godge[1] G.K.Kapse[2], H.S.Bhawar[1], Utkarsha D.Gharate[1], Sunita Kavade[3]

1) Department of Medicinal Chemistry, Pravara Rural College of Pharmacy, Pravaranagar, 413736, (MS) India.

2) Department of Medicinal Chemistry, Luqman College of Pharmacy, Gulbarga, (KA)

3) BPHE's, Ahemadnagar College, Ahemadnagar, (MS) India.

Summary

QSAR has done much to enhance our understanding of fundamental processes and phenomena in medicinal chemistry and drug design (251). The concept of hydrophobicity and its calculation has generated much knowledge and discussion as well as spawned a mini-industry. QSAR has refined our thinking on selectivity at the molecular and cellular level. Hydrophobic requirements vary considerably between tumor-sensitive cells and resistant ones. It has allowed us to design more selectivity into antibacterial agents that bind to dihydrofolate reducates. QSAR studies in the pharmacokinetic arena have established different hydrophobic requirements for renal no renal clearance, whereas the optimum hydrophobicity for CNS penetration has been determined by Hansch et al. (252). QSAR has helped delineate allosteric effects in enzymes such as cyclooxygenase, trypsin, and in the well-defined and complex hemoglobin system (253, 254). QSAR has matured over the last few decades in terms of the descriptors, models, methods of analysis, and choice of substituent and compounds. Embarking on a QSAR project may be a daunting and confusing task to a novice. However, there are many excellent reviews and tomes (1, 4, 19, 58Ð60) on this subject that can aid in the elucidation of the paradigm. Dealing with biological systems is not a simple problem and in attempting to develop a QSAR, one must always be cognizant of the biochemistry of the system analyzed and the limitations of the approach used.

Address for Correspondence:
Dr. Shashikant R Pattan
Principal, Pravara Rural College of Pharmacy, Pravaranagar
Department of Medicinals Chemistry, Pravara Rural College of Pharmacy, Pravaranagar.
A/P- Loni Bk. Taluka -Rahata, Dist-Ahmednagar 413736, India (MS).

E-mail- shashipattan@yahoo.co.

Introduction

It has been nearly 40 years since the quantitative structure-activity relationship (QSAR) found into the practice of agro chemistry, pharmaceutical chemistry, toxicology, and eventually most facets of chemistry (1). Its staying power may be attributed to the strength of its initial postulate that activity was a function of structure as described by electronic attributes, hydrophobicity, and steric properties as well as the rapid and extensive development in methodologies and computational techniques that have ensued to delineate and retain the many variables approaches in this. The overall goals of QSAR retain their original essence and remain focused on the predictive ability of the approach and its receptiveness to mechanistic interpretation. Rigorous analysis and of independent variables has led to an expansion in development of molecular and atom based descriptors, as well as descriptors derived from quantum chemical calculations and spectroscopy (2). The improvement in high-throughput screening procedures allows for rapid screening of large numbers of compounds under similar test conditions and thus minimizes the risk of combining variable test data from many sources. The formulation of thousands of equations using QSAR methodology attests to a validation of its concepts and its utility in the elucidation of the mechanism of action of drugs at the molecular level and a more complete understanding of physicochemical phenomena such as hydrophobicity. It is now possible not only to develop a model for a system but also to compare models from a biological database and to draw analogies with models from a physical organic database (3). This process is dubbed *model mining* and it provides a sophisticated approach to the study of chemical-biological interactions. QSAR has clearly matured, although it still has a way to go. The previous review by Kubinyi has relevant sections covering portions of this chapter as well as an extensive bibliography recommended for a more complete overview (4).

Historical Development of QSAR

More than a century ago, Crum-Brown and Fraser expressed the idea that the physiological action of a substance was a function of its chemical composition and constitution (5). A few decades later, in 1893, Richet showed that the cytotoxicities of a diverse set of simple organic molecules were inversely related to their corresponding water solubility (6). At the turn of the 20th century, Meyer and Overton independently suggested that the narcotic (depressant) action of a group of organic compounds paralleled their olive oil/water partition coefficients (7, 8). In 1939 Ferguson introduced a thermodynamic generalization to the correlation of depressant action with the relative work of Albert, and Bell and Robin established the importance of ionization of bases and weak acids in bacteriostatic activity (10Ð12). Meanwhile on the physical organic front, great strides were being made in the delineation of substituent effects on organic reactions, led by the seminal work of Hammett, which gave rise to the specific culture (13, 14). Taft devised a way for separating polar, steric, and resonance effects and introducing the first steric parameter, *ES* (15). The contributions of Hammett and Taft together laid the mechanistic basis for the development of the QSAR paradigm by Hansch and Fujita. In 1962 Hansch and Muir published their brilliant study on the structure-activity relationships of plant growth regulators and their dependency on Hammett constants and hydrophobicity (16). Using the octanol/water system, a whole series of partition coefficients were measured, and thus a new hydrophobic scale was introduced (17). The parameter p, which is the relative hydrophobicity of a substituent, was defined in a manner analogous to the definition of sigma (18). pX 5 log *P*X 2 log *P*H (1.1) *P*X and *P*H represent the partition coefficients of a derivative and the parent molecule, respectively. Fujita and Hansch then combined these hydrophobic constants with Hammetts electronic constants to yield the linear Hansch equation and its many extended forms (19). Log 1/*C* 5 *a*s 1 *b*p 1 *ck* (1.2) Hundreds of equations later, the failure of linear equations in cases with extended hydrophobicity ranges led to the development of the Hansch parabolic equation (20): The delineation of these models led to explosive development in QSAR analysis and related approaches. The bilinear model is a parabolic model and, in many cases, it has proved to be superior (21). Besides the Hansch approach, other methodologies were also developed to tackle structure- activity questions. The Free-Wilson approach addresses structure-activity studies in a congener series as described in Equation 1.5 (22).

$$BA = \sum a_i x_i + u$$

BA is the biological activity, *u* is the average contribution of the parent molecule, and *also* the contribution of each structural feature; *xi* denotes the presence *Xi* 5 1 or absence *Xi* 5 0 of a particular structural fragment. Limitations

in this approach led to the more sophisticated Fujita-Ban equation that used the logarithm of activity, which brought the activity parameter in line with other free energy-related terms (23).

$$\text{Log } BA = \sum G_i X_i + u$$

In Equation 1.6, *u* is defined as the calculated biological activity value of the un-substituted parent compound of a particular series. *Gi* represents the biological activity contribution of the substituent, whereas *Xi* is ascribed with a

value of one when the substituent is present or zero when it is absent. Variations on this activity-based approach have been extended by Klopman et al. (24) and Enslein et al. (25). Topological methods have also been used to address the relationships between molecular structure and physical/biological activity. The minimum topological difference (MTD) method of Simon and the extensive studies on molecular connectivity by Kier and Hall have contributed to the development of quantitative structure property/activity relationships (26, 27). Connectivity indices based on hydrogen- suppressed molecular structures are rich in information on branching, 3-atom fragments, and the degree of substitution, proximity of substituent and length, and heteroatom of substituted rings. A method in its embryonic stage of development uses both graph bond distances and Euclidean distances among atoms to calculate E-state values for each atom in a molecule that is sensitive to conformational structure. Recently, these electro topological indices that encode significant structured information on the topological state of atoms and fragments as well as their valence electron content have been applied to biological and toxicity data (28). Other recent developments in QSAR include approaches such as HQSAR, Inverse QSAR, and Binary QSAR (29-32). Improved statistical tools such as partial least square (PLS) can handle situations where the number of variables overwhelms the number of molecules in a data set, which may have collinear *X* variables (33).

Tools and Techniques of QSAR

In QSAR analysis, it is imperative that the biological data be both accurate and precise to develop a meaningful model. It must be realized that any resulting QSAR model that is developed is only as valid statistically as the data that led to its development. The equilibrium constants and rate constants that are used extensively in physical organic chemistry and medicinal chemistry are related to free energy values D*G*. Thus for use in QSAR, standard biological equilibrium constants such as *K*i or *K*m should be used in QSAR studies. Likewise only standard rate constants should be deemed appropriate for a QSAR analysis. 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 on QSAR studies. Biological data are usually expressed on a logarithmic scale because of the linear relationship between response and log dose in the mid region 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. A few common endpoints are outlined in Table 1.2. Biological data should pertain to an aspect of biological/biochemical function that can be measured. The events could be occurring in enzymes, isolated or bound receptors, in cellular systems, or whole animals. Because there is considerable variation in biological responses, test samples should be run in duplicate or preferably triplicate, except in whole animal studies where assay conditions (e.g. plasma concentrations of a drug) preclude such measurements.

It is also important to design a set of molecules that will yield a range of values in terms of biological activities. It is understandable that most medicinal chemists are reluctant to synthesize molecules with poor activity, even though these data points are important in developing a meaningful QSAR. Generally, the larger the range (.2 log units) in activity, the easier it is to generate a predictive QSAR. This kind of equation is more forgiving in terms of errors of measurement. A narrow range in biological activity is less forgiving in terms of accuracy of data. Another factor that merits consideration is the time structure. Should a particular reading be taken after 48 or 72 h? Knowledge of cell cycles in cellular systems or biorhythms in animals would be advantageous. Each single step of drug transport, binding, and metabolism involves

some form of partitioning between an aqueous compartment and a non-aqueous phase, which could be a membrane, serum protein, receptor, or enzyme. In the case of isolated receptors, the endpoint is clear-cut and the critical step is evident. But in more complex systems, such as cellular systems or whole animals, many localized steps could be involved in the random-walk process and the eventual interaction with a target. Usually the observed biological activity is the slow step or the rate-determining step. To determine a biological response (e.g., IC50), a dose-response curve is established.

Usually six to eight concentrations are tested to yield percentages of activity or inhibition between 20 and 80%, the linear portion of the curve. Using the curves, the dose responsible for an established effect can easily be determined. This procedure is meaningful if, at the time the response is measured, the system is at equilibrium, or at least under steady-state conditions. Other approaches have been used to apply additively the concept and ascertain the binding energy contributions of various substituent (R) groups. Fersht et al. have measured the binding energies of various alkyl groups to aminoacyl-tRNA synthetases (54). Thus the D*G* values for methyl, ethyl, isopropyl, and substituent were determined to be 3.2, 6.5, 9.6, and 5.4 kcal/mol, respectively. An alternative, generalized approach to determining the energies of various drug-receptor interactions was developed by Andrews et al. (55), who statistically examined the interactions of a diverse set of molecules in aqueous solution. Using Equation 1.9, a relationship was established between D*G* and *E*X (intrinsic binding energy), *E*DOF (energy of average entropy loss), and the D*Sr*,*t* (energy of rotational and translational entropy loss).

$$\Delta G = T\,\Delta S_{r,t} + n_{\mathrm{DOF}} E_{\mathrm{DOF}} + n_X E_{\mathrm{X}}$$

*E*X denotes the sum of the intrinsic binding energy of each functional group of which *N*x are present in each drug in the set. Using Equation 1.9, the average binding energies for various functional groups were calculated. These energies followed a particular trend with charged groups showing stronger interactions and non-polar entities, such as sp2, sp3 carbons, contributing very little. The applicability of this approach to specific drug-receptor interactions remains to be seen.

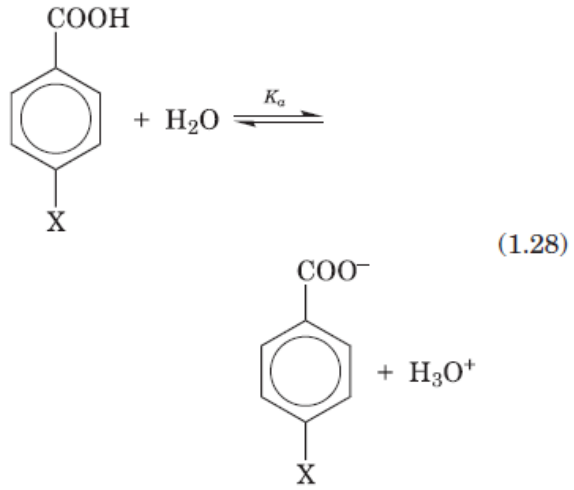
Parameters Used In QSAR

Electronic Parameters

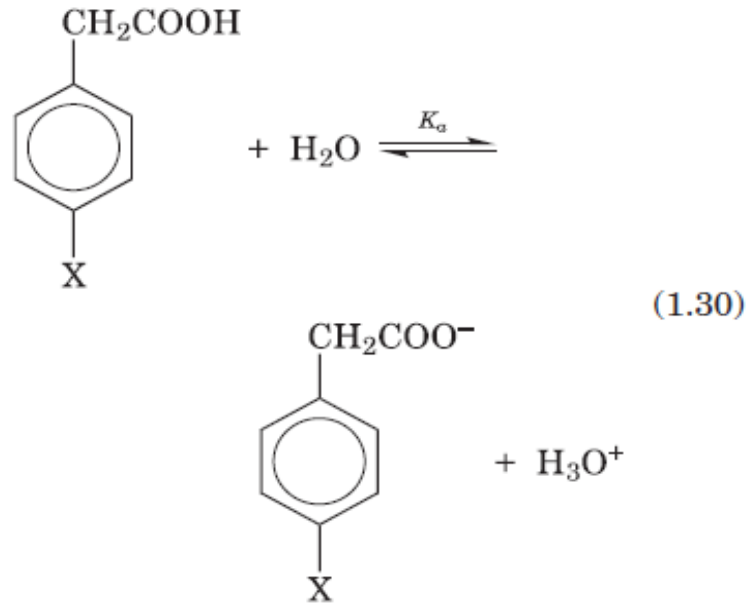
Parameters are of critical importance in determining the types of intermolecular forces that drug-receptor interaction. The three major types of parameters that were initially suggested and still hold sway are electronic, hydrophobic, and steric in nature (20, 75). Extensive studies using electronic parameters reveal that electronic attributes of molecules are intimately related to their chemical relativities and biological activities. A search of a computerized QSAR database reveals the following: the common Hammett constants (s, s1, s2) account for 7000/8500 equations in the Physical organic chemistry (PHYS) database and nearly 1600/8000 in the Biology (BIO) database, whereas quantum chemical indices such as HOMO, LUMO, BDE, and polarizability appear in 100 equations in the BIO database (76). The extent to which a given reaction responds to electronic perturbation constitutes a measure of the electronic demands of that reaction, which is determined by its mechanism. The introduction of substituent groups into the framework and the subsequent alteration of reaction rates helps delineate the overall mechanism of reaction. Early work examining the electronic role of substituent's on rate constants was first tackled by Burckhardt and established by Hammett (13, 14, 77, 78). Hammett employed, as a model reaction, the ionization in water of substituted benzoic acids and determined their equilibrium constants *Ka*. See Equation 1.28. This led to an operational definition of s, the substituent constant. It is a measure of the size of the electronic effect for a given substituent and represents a measure of electronic charge distribution in the benzene nucleus.

$$\sigma_X = \log K_X - \log K_H \quad or$$

$$\log(K_X/K_H) = -pK_X + pK_H$$



Electron-withdrawing substituents are thus characterized by positive values, whereas electron-donating ones have negative values. In an extension of this approach, the ionization of substituted phenyl acetic acids was measured.



The effect of the 4-Cl substituent on the ionization of 4-Cl phenyl acetic acid (PA) was found to be proportional to its effect on the ionization of 4-Cl benzoic acid (BA).

ϱ (rho) is defined as a proportionality or reaction constant, which is a measure of the susceptibility of a reaction to substituent effects.

$$\log \frac{k_X}{k_H} \propto \log \frac{K_X}{K_H} = \rho \cdot \sigma$$

Hydrophobicity Parameters

More than a hundred years ago, Meyer and Overton made their seminal discovery on the correlation between oil/water partition coefficients and the narcotic potencies of small organic molecules (7, 8). Ferguson extended this analysis by placing the relationship between depressant action and hydrophobicity in a thermodynamic context; the relative saturation of the depressant in the biophase was a critical determinant of its narcotic potency (9). At this time, the success of the Hammett equation began to permeate structure-activity studies and hydrophobicity as a determinant was relegated to the background. In a landmark study, Hansch and his colleagues devised and used a multiparameter approach that included both electronic and hydrophobic terms, to establish a QSAR for a series of plant growth regulators (16). This study laid the basis for the development of the QSAR paradigm and also established the importance of lipophilicity in biosystems. Over the last 40 years, no other parameter used in QSAR has generated more interest, excitement, and controversy than hydrophobicity (96). Hydrophobic interactions are of critical importance in many areas of chemistry. These include enzyme- ligand interactions, the assembly of lipids in biomembranes, aggregation of surfactants, coagulation, and detergency (97Ð100). The integrity of biomembranes and the tertiary structure of proteins in solution are determined by polar-type interactions. Molecular recognition depends strongly on hydrophobic interactions between ligands and receptors. Excellent treatises on this subject have been written by Taylor (101) and Blokzijl and Engerts (51). Despite extensive usage of the term *hydrophobic bond*, it is well known that there is no strong attractive force between polar molecules (102). Frank and Evans were the first to apply a thermodynamic treatment to the salvation of polar molecules in water at room temperature (103). Their iceberg model suggested that a large entropic loss ensued after the dissolution of a polar compounds and the increased structure of water molecules in the surrounding a polar solute. The quantization of this model led to the development of the ßickering cluster model of N. Scheraga, which emphasized the formation of hydrogen bonds in liquid water Hydrophobicity of solutes can readily be determined by measuring partition coefficients designated as *P*. Partition coefficients deal with neutral species, whereas distribution ratios incorporate concentrations of charged and/or polymeric species as

well. By convention, *P* is defined as the ratio of concentration of the solute in octanol to its concentration in water.

$$P = [\text{conc}]_{\text{octanol}}/[\text{conc}]_{\text{aqueous}}$$

Steric Parameters

The quantization of steric effects is complex at best and challenging in all other situations, particularly at the molecular level. An added level of confusion comes into play when attempts are made to delineate size and shape. Nevertheless, steric are of overwhelming importance in ligand-receptor interactions as well as in transport phenomena in cellular systems. The steric parameter to be used in QSAR studies was TaftÕs *E*S constant (157). *E*S is defined as

$$E_S = \log(k_X/k_H)_A$$

Where *k*X and *k*H represent the rates of acid hydrolysis of esters, XCH2COOR and CH3COOR, respectively. To correct for hyper conjugation in the hydrogen of the acetate moiety, Hancock devised a correction on *E*S such That the failure of the MR descriptor to adequately address three-dimensional shape issues led to Overlook development of STERIMOL parameters (162), which the steric constraints of a given substituent along several fixed axes. Five parameters were deemed necessary to definite shape: *L*, B1, B2, B3, and B4. *L* represents the length of a substituent along the axis of a bond between the parent molecule and the substituent; B1 to B4 represent four different width parameters. However, the high degree of co linearity between B1, B2, and B3 and the large number of training set members needed to establish the statistical validity of this group of parameters led to their demise in QSAR studies. Overlook subsequently established the adequacy of just three parameters for QSAR analysis: a slightly modified length *L*, a minimum width B1, and a maximum width B5 that is orthogonal to *L* (163). The use of these insightful parameters has done much to enhance correlations with biological activities. Recent analysis in our laboratory has established that in many cases, B1 alone is superior to Tafton ES and a combination of B1 and B5 can adequately replace *E*S (164). Molecular weight (MW) terms have also been used as descriptors, particularly in cellular systems, or in distribution/transport studies where diffusion is the mode of operation. According to the Einstein-Sutherland equation, molecular weight affects the diffusion rate. The Log MW term has been used extensively in

some studies (159-161) and an example of such usage is given below. In correlating permeability (Perm) of non electrolytes through cells, Lien et al. obtained the following QSAR (168):

PARAMETERS	SYMBOL
Hydrophobic parameters	
a) Partition coefficient	Log P.(log P)2
b) PI substituent constant	$\pi 2, \pi$
c) Rm chromatographic parameter	log Rm
d) Solubility	δ
e) elution time in HPTLC	Log K
Electronic parameter	
experimental parameters	
b) ionization constant	pKa
c) sigma substituent constant	σ2,σ
d) spectroscopic chemical shift	Δ Fr, ppm
e) resonance effect	F
f) field effect	I
g) ionization potential	E
Theoretical quantum mechanical indices	
a) Automic charge densities	
b) Striper delocalizabitily	QT
c) Energy of molecular orbit	pKa
	ELEMO
Steric parameters	
a) Taft's steric substituent constant	Es
b) Vander walls radii	γ
c) Interatomic distances	B,L

Quantitative Models

Linear Models

The correlation of biological activity with physicochemical properties is often termed an *extra thermodynamic relationship*. Because it follows in the line of Hammett and Taft equations that correlate thermodynamic and related parameters, it is appropriately labeled. The Hammett equation represents relationships between the logarithms of rate or equilibrium constants and substituent constants. The linearity of many of these relationships led to their designation as linear free energy relationships. The Hansch approach represents an extension of the Hammett equation from physical organic systems to a biological milieu. It should be noted that the simplicity of the approach belies the tremendous complexity of the intermolecular interactions at play in the overall biological response. Biological systems are a complex mix of heterogeneous phases. Drug molecules usually traverse many of these phases to get from the site of administration to the eventual site of action. Along this random-walk process, they perturb many other cellular components such as organelles, lipids, proteins, and so forth. These interactions are complex and vastly different from organic reactions in test tubes, even though the eventual interaction with a receptor may be chemical or physicochemical in nature. Thus, depending on the biological system involved isolated receptor, cell, or whole animal one expects the response to be multifactorial and complex. The overall process, particularly *in vitro* or *in vivo*, studies a mix of equilibrium and rate processes a situation that gives easy separation and delineation.

$$C = kA^m$$

$$\text{Log } 1/C = m \text{ Log}(1/A) + \text{constant}$$

C represents the equipotent concentration, k and m are constants for a particular system, and A is a physicochemical constant representative of phase distribution equilibria such as aqueous solubility, oil/water partition coefficient, and vapor pressure. In examining a large and diverse number of biological systems, Hansch and coworkers depend a relationship (Equation 1.62) that expressed biological activity as a function of physicochemical parameters (e.g., partition coefficients of organic molecules) (19).

Nonlinear Models

Extensive studies on development of linear models led Hansch and coworkers to note that a breakdown in the linear relationship occurred when a greater range in hydrophobicity was assessed with particular emphasis placed on test molecules at extreme ends of the hydrophobicity range. Thus, Hansch et al. suggested that the compounds could be involved in a Òrandom-walkÓ process: low hydrophobic molecules had a tendency to remain in the aqueous compartment, whereas highly hydrophobic analogs sequestered in the lipoid phase that they encountered. This led to the formulation of a parabolic equation, relating biological activity and hydrophobicity (187).

$$\text{Log } 1/C = -a(\log P)^2 + b \cdot \log P + \text{constant} \quad (1.70)$$

In the random-walk process, the compounds partition in and out of various compartments and interact with myriad biological components in the process. To deal with this conundrum, Hansch proposed a general, comprehensive equation for QSAR 1.71 (188).

$$\text{Log } 1/C = -a(\log P)^2 + b \cdot \log P + \rho\sigma + \delta E_S + \text{constant} \quad (1.71)$$

The optimum value of log *P* for a given system is log *P*O and it is highly influenced by the number of hydrophobic barriers a drug encounters in its walk to its site of action. Hansch and Clayton formulated the following parabolic model to elucidate the narcotic action of alcohols on tadpoles (189).

Other QSAR Approaches

The similarity in approaches of Hansch analysis and Free-Wilson analysis allows them to be used within the same framework. This is based on their theoretical consistency and the numerical equivalencies of activity contributions. This development has been called the mixed approach and can be represented by the following equation:

$$\text{Log } 1/C = \sum a_i + \sum c_j \varnothing_j + \text{constant} \quad (1.80)$$

The term *ai* denotes the contribution for each substituent, whereas any physicochemical property of a substituent X*j*. For a thorough review of the relationship between Hansch and Free-Wilson

analysis. A recent study of the P-glycoprotein inhibitory activity of 48 propafenone-type modulators of multidrug resistance, using a combined Hansch/Free-Wilson approach was deemed to have higher predictive ability than that of a stand-alone Free-Wilson analysis (201). Molar refractivity, which has high co linearity with molecular weight, was a significant determinant of modulating ability. It is of interest to note that molecular weight has been shown to be an omnipresent parameter in cross-resistance probes in multidrug-resistance phenomena (167).

A) Hansch's Approach (Extra thermodynamic or linear free energy relationship)

Hansch analysis is the investigation of the quantitative relationship between the biological activity of a series of compounds and their physicochemical substituent's or global parameters representing hydrophobic, electronic, steric & other effects using multiple regression correlation methodology. According to Hansch analysis,

Biological activity = a(Hydrophobic parameter) + b(Electronic parameter
+c(Steric descriptor) + d(Other descriptor) + e

Where,

a,b,c,d & e are constants determined by least square regression analysis (i.e. Hansch analysis). If the hydrophobicity values are limited to a small range then the equation will be linear as follows,

$$\text{Log } 1/c = K1\log P + K2\sigma + K3Es + K4$$

If the P values are spread over a large, range then the equation will be parabolic i.e.

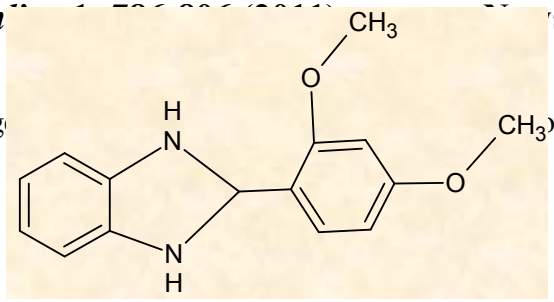
$$\text{Log}1/c = K1(\text{LogP})^2\ K2\text{LogP} + K3\sigma + K4Es + K5\ E$$

Applications of Hansch analysis:

1. in Pharmacokinetics:

Pharmacokinetics describes the time dependence of 'transport and distribution of drugs in different compartment of biological system'. eg. rate constant at blood & tissue level absorption, metabolism & elimination rate constants.

According to Hansch rule 'to get compound into the CNS, design it in such a way that LogP is near to 2' and 'to keep the compound out of the CNS & to avoid possible unwanted CNS side-effects such as drowsiness, design it in such a way that LogP is not near to 2'

eg. Cardio-tonic ag  f Sulmazole.

Stru.: 2,4-dimethoxy analogue of Sulmazole(LogP = 2.59)

A CNS side effect described as *bizarre* (seeing bright vision) was reported in the analogue of Sulmazole having Log P=2.59

2. Activity of drugs:

Bell & Roblin observe QSAR of sulfa drugs that a logarithmic plot of the bacteriostatic activities of some 40 sulfanilamide against parabolic equation was formulated.

$$\text{Log } 1/c = 2.103(\pm 29)\ PKa - 0.155(\pm 0.02)\ PKa^2 - 1.351(\pm 0.96)$$

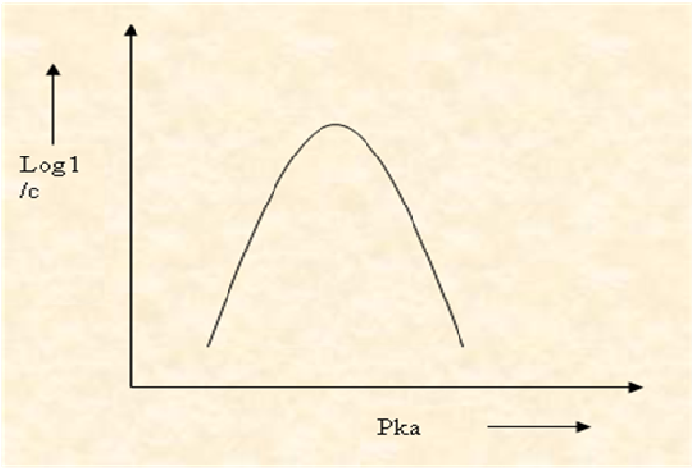
$$N=39;\quad r=0.939\quad ;s=0.321$$

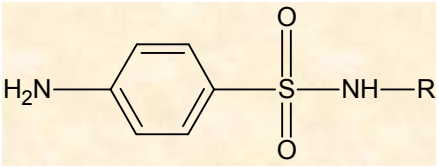
Where= Minimum inhibitory concentration.

N= Number of compounds utilized.

r= Correlation coefficient.

They considered that, the more negative the sulfomyl (SO2NH)group of sulfanilamide derivative, the more closely they will resemble the P.aminobenzoate anion with which sulfanilamide compete for the dihydrofolate synthetase enzyme. Being affected by negative charge on the adjacent amide nitrogen the sulfomyl group of sulfanilamide in the ionized form is much negative than in non-ionized form.





SULFANILAMIDE

Fig: Shows relationship between bacteriostatic activity & acidity of N1 substituted sulfanilamide.

3. In vitro determination of activity:

Hansch analysis can be applicable:

A. In determination of potency of new analogue for muscarinic receptor antagonist.

B. For studying protein binding affinity of thyroid hormone.

C. in QSAR studies of receptor agonist & antagonist of benzodiazepine receptor &estrogen receptor.

D. For in-vitro study of nonspecific, hemolytic, antibacterial & antifungal activities.

4. Enzyme inhibition:

Hansch analysis can be applicable in studying enzyme inhibition by analysis of enzyme inhibitors especially in combination with protein 3D structures & molecular graphics.

B) Free Wilson Analysis:

This method is preferred when nothing is known about the mode of action or when the physicochemical properties of the substituents used are unknown.
This method is based on the assumption that the introduction of a particular substituent at a particular position always leads to quantitatively similar effect on biological potency of the molecules expressed by equation:

Log BA = Contribution of un-substituted parent comd. + Contribution of corresponding substituent.

$= \mu + \varepsilon aij$

Where,

μ = the overall biological activity.

i = the no. of the position at which substitution occurs.

j = the no. of the substituent at that position.

The approach can be applied to a congeneric series having a common skeleton. Various derivatives must have been prepared by using different substituent at the same distinct positions of the parent skeleton. The substituents have to contribute to the biological activity additively at the same position. When choosing derivatives for the synthesis care has to be taken that every substituent appears at least twice at the same position. It is stated that the no. of derivatives for the solution of regression analysis must be at least ten, equal to the no. of increments.

Applications of QSAR

Over the last 40 years, the glut in scientific information has resulted in the development of thousands of equations pertaining to structure-activity relationships in biological systems. In its original definition, the Hansch equation was depend to model drug-receptor interactions involving electronic, steric, and hydrophobic contributions. Nonlinear relationships helped this approach in cellular systems and organisms where pharmacokinetic constraints had to be considered and tackled. They have also found increased utility in addressing the complex QSAR of some receptor-ligand interactions. In many cases the Kubinyi bilinear model has provided a sophisticated approach to delineation of steric effects in such interactions. Examples of ligand-receptor interactions will be drawn from receptors such as the much-studied dihydrofolate reeducates (DHFR), a-chymotrypsin and 5areductase (202-204).

Isolated Receptor Interactions

The critical role of DHFR in protein, purine, and pyrimidine synthesis; the availability of crystal structures of binary and ternary complexes of the enzyme; and the advent of molecular graphics combined to make attractive target for well-designed heterocyclic ligands generally incorporating a 2,4-diamino-1,-3-diazapharmacophore (205). The earliest study focused on the inhibition of DHFR by 4, 6-diamino-1, 2-dihydro- 2, 2-dimethyl-1*R*-*s*-triazines.

Interactions at the Cellular Level

QSAR analysis of studies at the cellular level allows us to get a handle on the physicochemical parameters critical to pharmacokinetics processes, mostly transport. Cell culture systems offer an ideal way to determine the optimum hydrophobicity of a system that is more complex than an isolated receptor. Extensive QSAR have been developed on the toxicity of 3-X-triazines to many mammalian and bacterial cell lines (202, 209). A comparison of the cytotoxicities of these analogs vs. sensitive murine leukemia cells (L1210/S) and methotrexate- resistant murine leukemia cells (L1210/R) reveals some startling differences.

Comparative QSAR

There are literally dozens of databases containing information about chemical structures, Synthetic methods and reaction mechanisms. The C-QSAR database is a database for QSAR models (164, 234). It was designed to organize QSAR data on physical (PHYS) organic reactions as well as chemical-biological (BIO) interactions, in numerical terms, to bring cohesion and understanding to mechanisms of chemical-biodynamic. The two databases are organized on a similar format, with the emphasis on reaction types in the PHYS database. The entries in the BIO database are sequestered into six main groups: macromolecules enzymes, organelles, single-cell organisms, organs/tissues, and multicellular organisms (e.g., insects). The combined databases or the separate PHYS or BIO databases can be searched independently by a string search or searching using the SMILES notation. A SMILES search can be approached in three ways: one can identify every QSAR that contains a specific molecule, one can use a MERLIN search that locates all derivatives of a given structure, or one can search on single or multiple parameters.

Progress in QSAR

The last four decades have seen major changes in the QSAR paradigm. In tandem with developments in molecular modeling and X-ray crystallography, it has impacted drug design and development in many ways. It has also spawned 3D QSAR approaches that are routinely used in computer-assisted molecular design. In terms of ligand design, it shares center stage with other approaches such as structure- based ligand design and other rational drug design approaches including docking methods and genetic algorithms (243). Success stories in QSAR have been recently reviewed (244, 245). Bioactive compounds have emerged in agro chemistry, pesticide chemistry, and medicinal chemistry. Bifenthrin, a pesticide, was the product of a design strategy that used cluster analysis (244) (Fig. 1.7). Guided by QSAR analysis, the chemists at Kyorin Pharmaceutical Company designed and developed Norflooxacin, a 6-ßuoro quinolone, which heralded the arrival of a new class of antibacterial agents (246) (Fig. 1.7). Two azoles-containing fungicides, metconazole (Fig. 1.8) and ipconazole were launched in 1994 in France and Japan, respectively (247). Lomerizine, a 4-*F*-benzhydryl-4- (2, 3, 4-trimethoxy benzyl) piperazine, was introduced into the market in 1999 after extensive design strategies using QSAR (248) (Fig. 1.8).

Flobufen, an anti-inflammatory agent was designed by Kuchar et al. as a long acting agent without the usual gastric toxicity

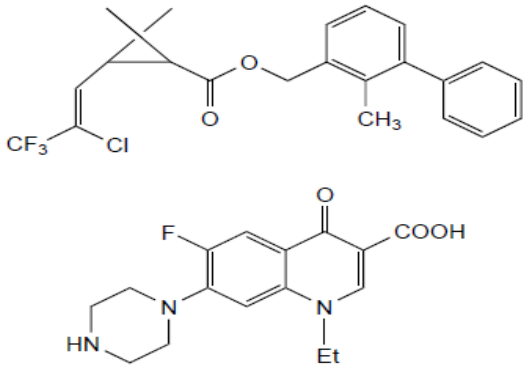


Figure 1.7. Bifenthrin and Norfloxacin.

(249) (Fig. 1.8). It is currently in clinical trial. Other examples of the commercial utility of QSAR include the development of metamitron and bromobutide (250). In most of these examples, QSAR was used in combination with other rational drug-design strategies, which is a useful and generally fruitful approach. In addition to these commercial successes, the QSAR paradigm has steadily evolved into a science. It is empirical in nature and it seeks to bring coherence and rigor to the QSAR models that are developed. By comparing models one is able to more fully comprehend scientific phenomena with a global perspective; trends in patterns of reactivity or biological activity become self-evident.

Structure-Activity Relationship for Anti-inflammatory Effect of Luteolin and its Derived Glycosides

Flavonoids are common secondary metabolites in the plant kingdom. Over 4000 flavonoids have been isolated from plants and structurally identified (Harborne, 1994). They play a significant role in plant metabolism and are considered relatively non-toxic bio-active substances. It is known that flavonoids differing by the type and numbers of substitution patterns show anti-inflammatory and free radical scavenging activities (Lewis, 1989; Moroney *et al.*, 1988; Arora *et al.*, 1998; Bors and Saran, 1987). Flavonoids have been shown to inhibit cyclooxygenase, lipoxygenase, microsomal monooxygenase, glutathione S-transferase, mitochondrial

succinoxidase and NADH oxidase, all involved in reactive oxygen species generation (Pieta, 2000; Robak and Gryglewski, 1996; Chang and Hsu, 1992; Hirano *et al.*, 2004). Thromboxane B2 (TXB2), a product of arachidonic acid metabolism, has attracted a great deal of attention as a potential pro-inflammatory mediator in diseases such as liver cirrhosis, systemic lupus erythematosis and thrombosis, while leukotriene B4 (LTB4) in the case of asthma, psoriasis, gout and inflammatory bowel diseases. The release of abundant amounts of free radicals in the human body also because the initiation of various diseases associated with inflammation. Thus, the determination of substances with active anti-inflammatory activity from natural sources might be a basis for new effective drugs. The present study assessed the hydrogen peroxide scavenging activity and the inhibitory effect on the enzymes for TXB2 and LTB4 of the arachidonate pathway synthesis of structurally related flavones such as luteolin, and its derived glycosides luteolin-7-O-β –D-Glycoside, luteolin-7-O-β -D-primeveroside , luteolin-6-C-β-D-glucoside and luteolin-6-O-β -D-glycosides which previously were isolated and identified from the different kinds of medicinal plant sources. The structure identification of these five compounds has been described previously.

Conclusion

QSAR - It is a mathematical calculation for predicting the new drug. Drug design a continues process of discovering a new regimen. Various parameters, models have been employed to search for new drug having least toxicity and high test efficacy QSAR is guidance for knowing its three dimension role in the biological system. Hansch analysis, amend constant, free Wilson model are the best parameters for QSAR both electronic, steric factor have been considered for QSAR research need to know guidelines for QSAR in the regard we attempted to review the QSAR and its application.

References

1. C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, John Wiley & Sons, New York, 1979.

2. C. D. Selassie, Z. X. Fang, R. Li, C. Hansch, T. Klein, R. Langrage, and B. T. Kaufman, *J. Med. Chem.*, 29, 621 (1986).

3. J. M. Blaney and C. Hansch in C. A. Ramsden, Ed., *Comprehensive Medicinal Chemistry. The*

Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds, Vol. 4, Quantitative Drug Design, Pergamon, Elmsford, NY, 1990, p. 459.

4. G. C. K. Roberts, *Pharmacochem. Libr.*, 6, 91 (1983).

5. A. A. Kumar, J. H. Mangum, D. T. Blankenship, and J. H. Freisheim, *J. Biol. Chem.*, 256, 8970 (1981).

6. G. D. Rose and R. Wolfenden, *Annu. Rev. Biophys. Biomol. Struct.*, 22, 381 (1993).

7. A. T. Hagler, P. Dauber, and S. Lifson, *J. Am. Chem. Soc.*, 101, 5131 (1979).

8. W. Kauzmann, *Adv. Protein Chem.*, 14, 1 (1959).

9. A. Ben-Naim, *Pure Appl. Chem.*, 69, 2239 (1997).

10. W. Blokzijl and J. B. F. N. Engberts, *Angew. Chem. Int. Ed. Engl.*, 32, 1545 (1993).

11. G. Odontuya,1* J. R. S. Hoult2† and P. J. Houghton3 1Natural Product Chemistry Laboratory, Research Institute of Chemistry and Chemical Technology, Mongolian Academy of Sciences, Zhukov Street-51, Ulaanbaatar 51, Mongolia 2Messengers and Signalling Research Group, Department of Pharmacy, King's College of London, 150 Stamford Street, London SE1 8WA, UK 3Pharmacognosy Research Laboratories, Department of Pharmacy, King's College of London, 150 Stamford Street, London SE1 8WA, UK

12. A. R. Fersht, J. S. Shindler, and W. C. Tsui, *1Biochemistry*, 19, 5520 (1980).

13. P. R. Andrews, D. J. Craik, and J. L. Matin, *J. Med. Chem.*, 27, 1648 (1984).

14. N. R. Draper and H. Smith, *Applied Regression Analysis*, 2nd ed., John Wiley & Sons, New York, 1981.

15. Y. Martin in G. Grunewald, Ed., *Quantitative Drug Design*, Marcel Dekker, New York, 1978, p. 167.

16. H. Kubinyi in R. Mannhold, P. Krogsgaard- Larsen, and H. Timmerman, Eds., *QSAR: Hansch Analysis and Related Approaches*, VCH, New York, 1993, p. 91.

17. R. Franke in W. Th. Nauta and R. F. Rekker, Eds., *Theoretical Drug Design Methods*, Elsevier Science, Amsterdam/New York, 1983, p. 395.

18. C. Hansch in C. J. Cavallito, Ed., *Structure Activity Relationships*, Vol. 1, Pergamon, Oxford.