# HANSCH AND FREE WILSON ANALYSIS

## HANSCH ANALYSIS

**Introduction:-**

**Corwin Hansch said "Similar compounds behave similarly"** Corwin Hansch ( born October 6, 1918, Kenmare, North Dakota) is Professor of Chemistry at Pomona College in California. Hansch taught Organic Chemistry for many years at Pomona College . His course in Physical Bio-Organic Medicinal Chemistry was ground-breaking at an undergraduate level.

Hansch may be best known as the father of the concept of Quantitative Structure-Activity Relationship (Q.S.A.R.), the quantitative correlation of the physicochemical properties of molecules with their biological activities.

He is also noted for the Hansch equation, which is used in

(1) Multivariate Statistics - Multivariate statistics is a set of statistical tools to analyse data (e.g., chemical and biological) matrices using regression and/or pattern recognition techniques.

(2) Hansch Analysis - Hansch analysis is the investigation of the quantitative relationship between the biological activity of a series of compounds and their physicochemical substituent or global parameters representing hydrophobic, electronic, steric and other effects using multiple regression correlation methodology.

(3) Hansch-Fujita π constant - The Hansch-Fujita π constant describes the contribution of a substituent to the lipophilicity of a compound.

Hansch proposed the action of a drug as depending on two processes. Firstly the movement of drug from the point of entry in the body to the site of action and secondly the interaction with the receptor site. Hansch suggested the linear and non-linear dependence of biological activity on difference parameters.

**Importance of Lipophilicty:**- Hansch visualized that diffusion into cell is slow process so as it is also important one to determine .It is highly dependent on molecular structure of the drug.

Drug must pass out two barriers to put out their effect at site of action ,lipophilliic barrier(cell membrane) and aqueous barrier(cytoplasm) as we know that cytoplasm is made of fatty acids and membrane is made of glycolipids and phospholipids, they have two ends -

(i)Lipophilicity or Hydrophobic—steroids and hydrocarbons

(ii)Hydrophillic end— hydroxyl group in cholesterol, sugar in glycolipids ammonia moiety in phospholipids

The cell is in lipid bilayer structure it is generally found that increasing the lipophilicity of a lead compound results in an increase in biological activity. This means fact that drugs have to cross hydrophobic barriers such as cell membranes to reach their target. But in vitro studies there is no such barriers have to be crossed and interact with a target system such as an enzyme or receptor where the binding site is usually hydrophobic. Therefore, increasing hydrophobicity must give positive results but it is not like we hope because highly hydrophobic means poorly soluble in the aqueous phase,Alternatively, it may got 'trapped' in fat depots and unable to reach the target site.

Conclusion-The drug must have a balance between hydrophilic and lipophilic properties to cross these barriers, that cell structure or target site approach make lipophilicity a crucial factor.

**Assumptions in Hansch analysis:-**

i. Conformational changes takes place in target site can be ignored

ii. Metabolism doesn't interferes in it

iii. Linear free energy terms are relevant to receptor's affinity and additive in nature

iv. Relationship between potency and lipophilicity is linear or parabolic

v. Hansch proposed the action of a drug as depending on two processes. Firstly the movement of drug from the point of entry in the body to the site of action and secondly the interaction with the receptor site. Hansch suggested the linear and non-linear dependence of biological activity on difference parameters.

**Linear Hansch model:**

The correlation of biological activity with physicochemical properties is often termed an Extrathermodynamic 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.

Interestingly, the concurrent considerations of π and σ (Hammett's constant) has evolved gainful vital correlations existing between the biological activities of quite a few drug substances with their corresponding physical properties and chemical structures.

Therefore, Hansch's correlations piece together valuable informations of a newly designed 'drug molecule' in a more plausible, predictive and quantifiable manner than before and apply it to a biological system more logistically and judiciously. This particular concept and idea was further substantiated and expanded by assuming that all the three substituents viz., π, σ and Es, exert a significant effect on the efficacy and hence the potency of a 'drug substance'; and are found to be additive in nature independently. Therefore, it has given rise to the underlying linear Hansch equation also called extrathermodynamic approach.

$$\log 1/C = a \log P + b\,\sigma + c\,E_S + d \qquad \text{(Equation 2)}$$

Where, C = Concentrations of drug producing the biological response being measured,

Log P = Substituent constant for solubility (i.e., π),

$E_S$ = Taft constant (for steric effects),

σ = Hammett substitution constant

a, b, d = Constants of the system (which are determined by computer to obtain the 'best fitting line').

It is pertinent to state at this juncture that not all the parameters shall necessarily be significant.

**Nonlinear Hansch models**

Later on experiences results that

(i) Increase in log P value from log $P_o$ does not linearly cause increase in biological activity some time its decreases.

(ii) If P values spread over a large range.

Thus, Hansch et al suggested that the compounds could be involved in a random-walk process: low hydrophobic molecules had a tendency to remain in the first aqueous compartment, whereas highly hydrophobic analogs sequestered in the first lipoidal phase that they encountered. This led to the formulation of a parabolic equation, relating biological activity and hydrophobicity

$$\log 1/C = -a ( \log P)^2 + b. \log P + \text{constant (k)} \qquad \text{(equation 3)}$$

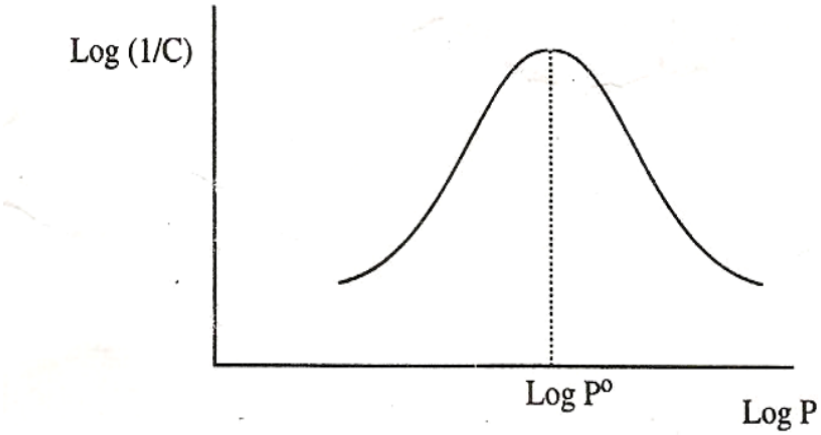
In the random-walk process, the compounds partition in and out of various compartments and interact with myriad biological components in the process. To deal with this conundrum, Hansch proposed a general, comprehensive equation for QSAR

$$\log 1/C = -a ( \log P)^2 + b. \log P + \rho \sigma + \delta E_S + k \quad \text{(equation 4)}$$

Where, P = n-octanol/ water partition coefficient, σ = Hammett electronic parameter, a,b,c = regression coefficients, $E_S$ = taft's steric factor, k = constant term, ρ (rho) and δ are Proportionality constant designating the sensitivity of the reaction to electron density.

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.

The coefficients (a, b, c, d, e) are determined by multi-regression analysis.




**Graph-** When P is small, the $(\log P)^2$ term is very small and the equation is dominated by the log P term. This represents the first part of the graph where activity increases with increasing P. When P is large, the $(\log P)^2$ term is more significant and eventually 'overwhelms' the log P term. This represents the last part of the graph where activity drops with increasing P

Because of importance of electronic factors, steric effect and shape of molecules for receptor interaction, electric factor (σ) steric factor (Es) and topography terms (S) added to equation:-

If Linear, $\log 1/C = a \log P + b\,\sigma + c\,E_S + d$

If Parabolic, $\log 1/C = -a(\log P)^2 + b.\log P + c\,\sigma + d\,E_S + k$

Equation 4, was developed from the concept that the transport of a drug from the site of application to its site of action depends in a nonlinear manner on the lipophilicity of the drug, and that the binding affinity to its biological counterpart, such as an enzyme or a receptor, depends on the lipophilicity, the electronic properties and other linear free-energy-related properties. Equation 4, combines the description of both processes in one mathematical model. In addition to the introduction of a parabolic term for the nonlinear lipophilicity dependence and the combination of different physicochemical properties in one equation, Hansch and Fujita defined lipophilicity parameters π of substituents 'X'(Eqn. 5), in the same manner as Hammett had defined the electronic parameter σ (Eqn. 6). The partition coefficient P in equation 5 is an equilibrium constant, similar to the dissociation or reaction constants K in Eqn. 6. The absence of a 'reaction term' π in Eqn. 5 is explained by the fact that all π values refer to the n-octanol/water system.

$\Pi_x = \log P_x - \log P_{yH}$ (Equation 5)

$\rho\,\sigma = \log K_{RX} - \log K_{RH}$ (Equation 6)

Where, log P = logarithm of 1-octanol-water partition coefficient.

ρ (rho) = Proportionality constant designating the sensitivity of the reaction to electron density.

y = A parent compound (i.e., an unsubstituted reference compound/drug).

The hydrophobic characterictic, designated by $\pi_x$, may be correlated to a drug's distribution pattern, within which a given substituent 'x' affects molecular behavior and conduct with regard to its:

1. Distribution and transport, and

2. Drug-receptor activities.

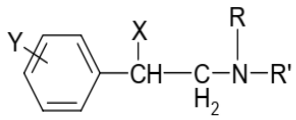
**Salient Features:** The various salient features of **Hansch equation** are as enumerated under:

(1)Value of π is indicative, to a certain extent, the behavioural pattern of a 'substituent' contributing to the solubility behavior of a molecule under investigation. It also reflects upon the manner it gets partitioned between lipoidal and aqueous interfaces in the reputed compartments it happens to cross as a 'drug' so as to reach the 'site of action' ultimately.

(2) It is, however, not very clear and definite whether the solid surface of a drug undergoes adsorption on colloidally suspended plasma proteins while establishing the hydrophobic characteristics π.

Example: β-Halo-arylamines: The adrenergic blocking profile of β-halo-arylamines was observed to be solely related to the constants, π and σ; and specifically excluded the steric factor altogether.

i.e., $\log 1/C = 1.22\,\pi - 1.59\,\sigma + 7.89$

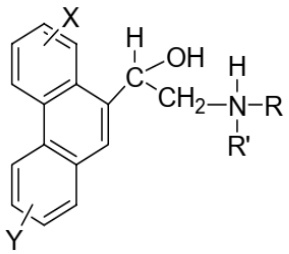



The aforesaid equation offers a dictum that the 'biological response'gets enhanced if the substituents possess a positive π value and a negative σ value; or more explicitely the substituents must preferentially be both hydrophobic in nature and electron donating in character.

It has been established beyond any reasonable doubt that there exists no correlation between the π factor and the P value; therefore, it is quite feasible to have Hansch equations essentially comprising of these two stated components:

Examples: Phenanthrene aminicarbinols: An analogues series of more than one hundred phenanthrene aminocarbinols were successfully synthesized and subsequently screened for their antimalarial profile. Interestingly, the analogues series fitted appropriately into the following version of Hansch equation:

$\text{Log } 1/C = -0.015\,(\log P)^2 + 0.14 \log P + 0.27\,\square\,\pi_x + 0.40\,\square\,\pi_y + 0.65\,\square\,\sigma_x + 0.88\,\square\,\sigma_y + 2.34$




PHENANTHRENE AMINOCARBINOL

**Salient features:** The various salient features that may be derived from the above equation are, namely:

(1) As the hydrophobicity of the molecule (P) enhances there exists a very nominal increase in the antimalarial activity.

(2) The corresponding constant is low (0.14) which reflects that the increase in antimalarial activity is also low.

(3) The value of $(\log P)^2$ evidently reveals that there prevails a maximum P value for activity.

(4) Further the above equation suggests that the antimalarial activity gets enhanced appreciably when the hydrophobic moieties are strategically located either on ring 'X' or more specifically on ring 'Y'. It further ascertains that the hydrophobic interactions(s) are virtually taking place at these sites.

(5) The electron- withdrawing substituents on rings 'X' and 'Y' contribute enormously to the antimalarial activity; however, the effect is more on ring 'Y' than in ring 'X'.

## Free-Wilson analysis (structure-property relationship)

In 1964, Free and Wilson derived a mathematical model that describes the presence and absence of certain structural features, i.e. those groups that are chemically modified, by values of 1 or 0 and correlate the resulting structural matrix with biological activity values, following Eqn. 7; the values $a_i$ in Eqn.7 are the biological activity group contributions of the substituents $X_1$, $X_2$,…$X_i$ in the different positions P of compound, the most often the unsubstituted parent structure of a series . The method of Free and Wilson is based upon an additive mathematical model in which a particular substituent in a specific position is assumed to make an additive and constant contribution to the biological activity of a molecule in a series of chemically related molecules. This method is based on the assumption that the introduction of a particular substituent at a particular molecular position always leads to a quantitatively similar effect on biological potency of the whole molecule, as expressed by the equation

$$\log 1/C = \Sigma a_i + \mu \qquad \text{(Equation7)}$$

$a_i$ = substituent group contributions, $\mu$ = activity contribution of reference compound

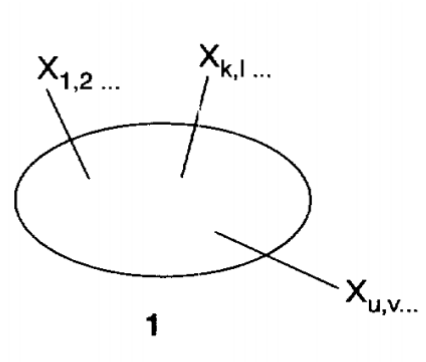



Figure 1: Schematic presentation of a molecule for Free Wilson analysis. A common skeleton bears subsituents Xi in different positions p; the presence or absence of these substituents is coded by the values 1 and 0, respectively.

## Mixed Hansch/Free-Wilson model

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:

Log $1/C = \square\ a_i + \square\ c_j \Phi_j$ + constant (Equation 8)

The term $a_i$ denotes the contribution for each ith substituent, whereas $\Phi_j$ is any physicochemical property of a substituent $X_j$.
Equation 8 combines the advantages of Hansch and FreesWilson analyses and widens the applicability of both methods. Physicochemical parameters describe parts of the molecules with broad structural variation, whereas indicator variables $a_i$ (Free Wilson type variables) encode the effects of structural variations that cannot be described otherwise . 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

**Conclusion:** It involves the mathematical and statistical analysis of SAR-data which helps to reduce the number of educated guesses in molecular modification. QSAR is thus a scientific achievement and an economic necessity to reduce an empiricism in drug design to ensure that every drug synthesized and pharmacologically tested should be as meaningful.

REFERENCES


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

2. A Crum-Brown, TR Fraser, Trans *R* Soc Edinburgh 1868–1869; *25:* 151.
3. Richardson BJ. Physiological research on alcohols. Med Times Gaz 1868; 2: 703-706.
4. Mills EJ. On melting point and boiling point as related to composition. *Philos Mag* 1884; *17:* 173-187.
5. Richet C. On the relationship between the toxicity and the physical properties of substances. Compt. Rendus Seances Soc Biol 1893; 9: 775-776.
6. Overton E. Osmotic properties of cells in the bearing on toxicology and pharmacology. *Z* Physik Chem 1897; *22:* 189-209.
7. Meyer H. On the theory of alcohol narcosis I. Which property of anesthetics gives them their narcotic activity? Arch Exp Pathol Pharmakol 1899; *42:* 109-118.
8. Hammett LP. Some relations between reaction rates and equilibrium constants. Chem Rev 1935; 17: 125-136.
9. Hammett LP. The effect of structure upon the reactions of organic compounds. benzene derivatives. J Am Chem Soc 1937; *59:* 96-103.
10. Ferguson J. The Use of Chemical Potentials as Indices of Toxicity. Proc R Soc Lond B 1939; 127: 387-404.
11. Bell PH, Roblin RO. Studies in chemotherapy. vii. a theory of the relation of structure to activity of sulfanilamide type compounds1. J Am Chem Soc 1942; 64: 2905-2917.
12. Albert A, Goldacre R, Phillips J. The strength of heterocyclic bases. J Chem Soc 1948: 2240-2249.
13. Taft RW. Polar and steric substituent constants for aliphatic and o- Benzoate groups from rates of esterification and hydrolysis of esters1. J Am Chem Soc 1952; *74:* 3120-3128.
14. Hansch C, Maloney PP, Fujita T, Muir RM. Correlation of biological activity of phenoxyacetic acids with hammett substituent constants and partition coefficients. Nature 1962; 194: 178-180.