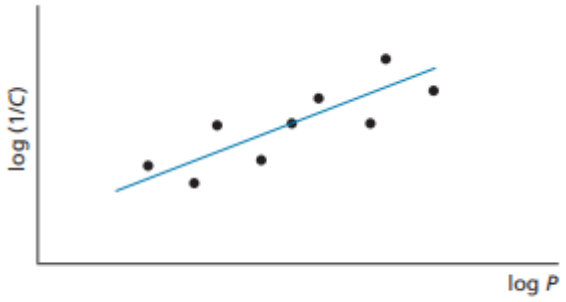
QSAR

The QSAR approach attempts to identify and quantify the physicochemical properties of a drug and to see whether any of these properties has an effect on the drug's biological activity by using a mathematical equation.

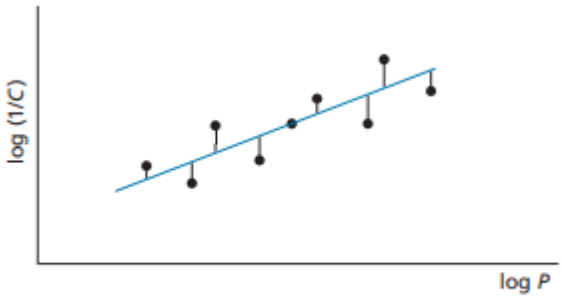
GRAPH AND EQUATIONS

A range of compounds is synthesized in order to vary one physicochemical property (e.g. log P) and to test how this affects the biological activity .A graph is then drawn to plot the biological activity on the y-axis versus the physicochemical feature on the x-axis

It is then necessary to draw the best possible line through the data points on the graph. Th is done by a procedure known as linear regression analysis by the least squares method. If we draw a line through a set of data points, most of the points will be scattered on either side of the line. Th e best line will be the one closest to the data points. To measure how close the data points are, vertical lines are drawn from each point



Biological activity versus log *P*.



Proximity of data points to line of best fit.

These verticals are measured and then squared in order to eliminate the negative values. Th e squares are then added up to give a total (the sum of the squares). The best line through the points will be the line where this total is a minimum.

PHYSICOCHEMICAL PROPERTIES

Many physical, structural, and chemical properties have been studied by the QSAR approach, but the most common are hydrophobic, electronic, and steric properties

HYDROPHOBICITY

Th e hydrophobic character of a drug is crucial to how easily it crosses cell membranes and may also be important in receptor interactions

THE PARTITION COEFFICIENT (P)

The hydrophobic character of a drug can be measured experimentally by testing the drug's relative distribution in an n -octanol/water mixture. Hydrophobic molecules will prefer to dissolve in the n -octanol layer of this two phase system, whereas hydrophilic molecules will prefer the aqueous layer. The relative distribution is known as the partition coefficient (P) and is obtained from the following equation:

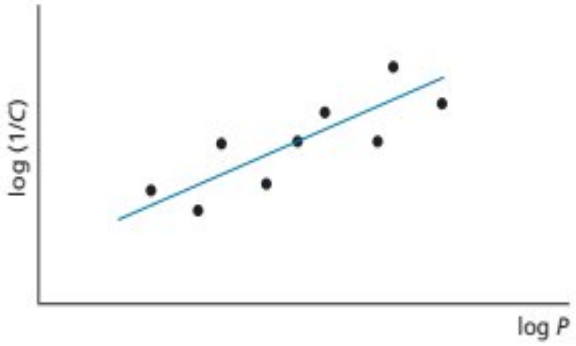
$$P = \frac{\text{Concentration of drug in octanol}}{\text{Concentration of drug in aqueous solution}}$$

Hydrophobic compounds have a high P value, whereas hydrophilic compounds have a low P value. Varying substituents on the lead compound will produce a series of analogues having different hydrophobicities and, therefore, different P values. By plotting these P values against the biological activity of these drugs, it is possible to see if there is any relationship between the two properties. Th e biological activity is normally expressed as 1/ C , where C is the concentration of drug required to achieve a defined level of biological activity. Th e reciprocal of the concentration (1/ C) is used, as

more active drugs will achieve a defined biological activity at lower concentration. Th e graph is drawn by plotting log (1/ C) versus log P . In studies where the range of the log P values is restricted to a small range (e.g. log P = 1–4), a straight-line graph is obtained showing that there is a relationship between hydrophobicity and biological activity. Such a line would have the

$$\log\left(\frac{1}{C}\right) = -k_1 \log P + k_2$$

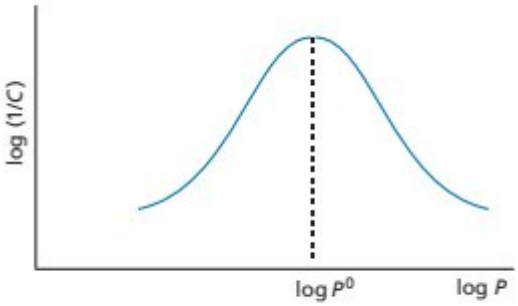
following equation:



Biological activity versus log *P*.

If these studies were to be extended to include compounds with very high log P values, then we would see a different picture. The graph would be parabolic. Here, the biological activity increases as log P increases until a maximum value is obtained. Th e 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. If the partition coefficient is the only factor influencing biological activity, the parabolic curve can be expressed by the equation:

$$\log\left(\frac{1}{C}\right) = -k_1(\log P)^2 + k_2 \log P + k_3$$



Parabolic curve of log (1/*C*) vs. log *P*.

Note that the $(\log P)^2$ term has a minus sign in front of it. When P is small, the $(\log P)^2$ term is very small and the equation is dominated by the log P term. Th is represents the first part of the graph where activity increases with increasing P . When P is large, the $(\log P)^2$ term is more.

The substituent hydrophobicity constant (π)

Partition coefficients can be calculated by knowing the contribution that various substituents make to hydrophobicity. Th is contribution is known as the substituent hydrophobicity constant (π) and is a measure of how hydrophobic a substituent is relative to hydrogen. Th e hydrophobicity constant (πX) for the substituent (X) is then obtained using the following equation:

$\pi_X = \log P_X - \log P_H$ where P_H is the partition coefficient for the standard compound and P_x is the partition coefficient for the standard compound with the substituent. A positive value of π indicates that the substituent is more hydrophobic than hydrogen; a negative value indicates that the substituent is less hydrophobic. These π values are characteristic for the substituent and can be used to calculate how the partition coefficient of a drug would be affected if these substituents were present.

As an example, consider the log P values for benzene (log P = 2.13), chlorobenzene (log P = 2.84), and benzamide (log P = 0.64). Benzene is the parent compound, and the substituent constants for Cl and CONH 2 are 0.71 and −1.49 respectively. Having obtained these values, it is now possible to calculate the theoretical log P value for meta -chlorobenzamide:

$$\begin{aligned}\log P_{(\text{chlorobenzamide})} &= \log P_{(\text{benzene})} + \pi_{Cl} + \pi_{CONH_2} \\ &= 2.13 + 0.17 + (-1.49) \\ &= 1.35\end{aligned}$$

Th e observed log P value for this compound is 1.51.

ELECTRONIC EFFECTS

Th e electronic effects of various substituents will clearly have an effect on a drug's ionization or polarity. This, in turn, may have an effect on how easily a drug can pass through cell membranes or how strongly it can interact with a binding site. It is, therefore, useful to measure the electronic effect of a substituent.

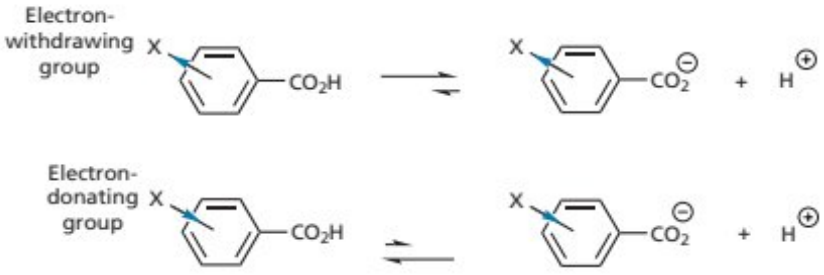
Hammett substituent constant (σ)

This is a measure of the electronwithdrawing or electron-donating ability of a substituent, and has been determined by measuring the dissociation of a series of substituted benzoic acids compared with the dissociation of benzoic acid itself.

Benzoic acid is a weak acid and only partially ionizes in water .An equilibrium is set up between the ionized and non-ionized forms, where the relative proportion of these species is known as the equilibrium or dissociation constant KH (the subscript H signifies that there are no substituents on the aromatic ring).

$$K_H = \frac{[PhCO_2^-]}{[PhCO_2H]}$$

When a substituent is present on the aromatic ring, this equilibrium is affected. Electron-withdrawing groups, such as a nitro group, result in the aromatic ring having a stronger electron-withdrawing and stabilizing influence on the carboxylate anion, and so the equilibrium will shift more to the ionized form. Therefore, the substituted benzoic acid is a stronger acid and has a larger KX value (X represents the substituent on the aromatic ring)



Position of equilibrium dependent on substituent group X.

If the substituent X is an electron-donating group such as an alkyl group, then the aromatic ring is less able to stabilize the carboxylate ion. Th e equilibrium shift s to the left indicating a weaker acid with a smaller KX value

The Hammett substituent constant (σ_x) for a particular substituent (X) is defined by the following equation:

$$\sigma_X = \log \frac{K_X}{K_H} = \log K_X - \log K_H$$

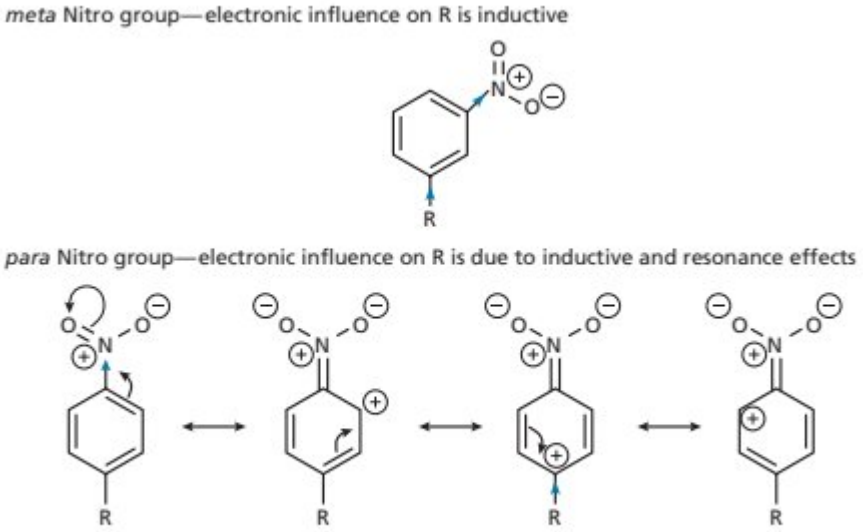
Benzoic acids containing electron-withdrawing substituents will have larger K X values than benzoic acid itself (K_H) and, therefore, the value of σ_x for an electronwithdrawing substituent will be positive. Substituents such as Cl, CN, or CF 3 have positive σ values.

Benzoic acids containing electron-donating substituents will have smaller K_x values than benzoic acid itself and, hence, the value of σ_x for an electron-donating substituent will be negative.

Substituents such as Me, Et, and t -Bu have negative values of σ. Th e Hammett substituent constant for H is zero.

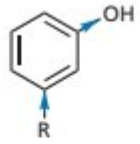
The Hammett substituent constant takes into account both resonance and inductive effects. Therefore, the value of σ for a particular substituent will depend on whether the substituent is meta or para . This is indicated by the subscript m or p after the σ symbol. For example, the nitro substituent has σ_P = 0.78 and σ_M = 0.71. In the meta position, the electron-withdrawing power is due to the inductive influence of the substituent, whereas at the para position inductive and resonance both play a part and so the σ_P value is greater.

For the hydroxyl group σ_M= 0.12 and σ_P = −0.37. At the meta position, the influence is inductive and electron-withdrawing. At the para position, the electron donating influence due to resonance is more significant than the electron-withdrawing influence due to induction. Hammett substituent constants cannot be measured for ortho substituents as such substituents have an important steric, as well as electronic, effect.

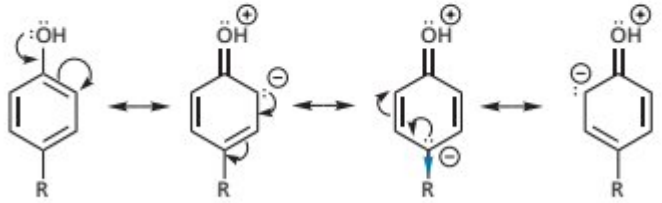


Substituent effects of a nitro group at the *meta* and *para* positions.

meta Hydroxyl group—electronic influence on R is inductive



para Hydroxyl group—electronic influence on R dominated by resonance effects



Substituent effects of a phenol at the *meta* and *para* positions.

STERIC FACTORS

The bulk, size, and shape of a drug will infl uence how easily it can approach and interact with a binding site. A bulky substituent may act like a shield and hinder the ideal interaction between a drug and its binding site. Alternatively, a bulky substituent may help to orientate a drug properly for maximum binding and increase activity. Steric properties are more diffi cult to quantify than hydrophobic or electronic properties.Different steric factors include

1. Molar refractivity- This is a measure of the volume occupied by an atom or a group of atoms. Th e MR is obtained from the following equation:

$$MR = \frac{(n^2-1)}{(n^2+2)} \times \frac{MW}{d}$$

2. Verloop steric parameter: Another approach to measuring the steric factor involves a computer program called Sterimol , which calculates steric substituent values (Verloop steric parameters) from standard bond angles, van der Waals radii, bond lengths, and possible conformations for the substituent. Unlike Es , the Verloop steric parameters can be measured for any substituent.

3 Taft's steric factor (Es)

The value for Es can be obtained by comparing the rates of hydrolysis of substituted aliphatic esters against a standard ester under acidic conditions. Thus

$$E_s = \log k_X - \log k_O$$

where kx represents the rate of hydrolysis of an aliphatic ester bearing the substituent X and ko represents the rate of hydrolysis of the reference ester.

The substituents that can be studied by this method are restricted to those which interact sterically with the tetrahedral transition state of the reaction and not by resonance or internal hydrogen bonding. For example, unsaturated substituents which are conjugated to the ester cannot be measured by this procedure. Substituents such as H and F, which are smaller than a methyl group, result in a faster rate of hydrolysis (kx > ko), making Es positive. Substituents which are larger than methyl reduce the rate of hydrolysis (kx < ko), making Es negative.

A disadvantage of Es values is that they are a measure of an intramolecular steric effect, whereas drugs interact with target binding sites in an intermolecular manner. For example, consider the Es values for i -Pr, n -Pr, and n -Bu. Th e Es value for the branched isopropyl group is signifi cantly greater than that for the linear n -propyl group since the bulk of the substituent is closer to the reaction centre. Extending the alkyl chain from n -propyl to n -butyl has little effect on Es . The larger n -butyl group is extended away from the reaction centre and so has little additional steric eff ect on the rate of hydrolysis. As a result, the Es value for the n -butyl group undervalues the steric eff ect which this group might have if it was present on a drug approaching a binding site.

Values of E_s for various substituents

Substituent	H	F	Me	Et	n-Pr	n-Bu	i-Pr	i-Bu	Cyclopentyl
E_s	1.24	0.78	0	–0.07	–0.36	–0.39	–0.47	–0.93	–0.51

HANCH ANALYSIS

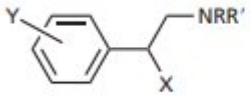
The biological activity of most drugs, however, is related to a combination of physicochemical properties. In such cases, simple equations involving only one parameter are relevant only if the other parameters are kept constant. These equations are known as Hansch equations and they usually relate biological activity to the most commonly used physicochemical properties (log P , π, σ, and a steric factor). If the range of hydrophobicity values is limited to a small range then the equation will be linear, as follows:

$$\log\left(\frac{1}{C}\right) = k_1 \log P + k_2\sigma + k_3 E_s + k_4$$

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

$$\log\left(\frac{1}{C}\right) = -k_1(\log P)^2 + k_2 \log P + k_3\sigma + k_4 E_s + k_5$$

Th e constants k 1 – k 5 are determined by computer soft - ware in order to get the best-fi tting equation. Not all the parameters will necessarily be significant.



β-Halo-arylamines

$$\log\left(\frac{1}{C}\right) = 1.22\pi - 1.59\sigma + 7.89$$

$(n = 22, r^2 = 0.841, s = 0.238)$

For example, the adrenergic blocking activity of β-halo-arylamines was related to π and σ and did not include a steric factor. Th is equation tells us that biological activity increases if the substituents have a positive π value and a negative σ value. In other words, the substituents should be hydrophobic and electron donating.

APPLICATIONS

Used to predict the activity of an as yet unsynthesized compound

To give an indication of the importance of influence of parameters on mechanism by which drug acts

Applied to various problems in order to co relate biological activity with chemical structure

It serves as a guide in future testing and synthesis of new compound and to play role of hydrophobic,electronic and steric factors in drug receptor interaction