

Structure and Reactivity

Yasuhide Yukawa

Institute of Industrial Science Research, Osaka University, Osaka, Japan

(Received January 9, 1976)

One of the most important aims of organic chemistry is to establish a relationship between the structures of organic compounds and their reactivities. Such a quantitative type of a relationship was first proposed by L.P. Hammett who discovered linear relationships involving logarithms of rate or equilibrium constants for a number of systems. As a relationship between $\log K$ or $\log k$ at a constant temperature is essentially a relationship between Gibbs' free energies, such relationships are usually called as linear free energy relationships.

The Hammett equation (1) describes the influence of meta- or para-substituents on the side-chain reactions of benzene derivatives.

$$\log (k/k_0) = \rho\sigma \dots\dots\dots(1)$$

The symbol k is the rate or equilibrium constant and k_0 denotes k for the 'unsubstituted' or 'parent' compound.

The substituent constant, σ , measures the polar effect relative to hydrogen of the substituent in a given position, meta or para and is, in principle, independent of the nature of the reaction. The reaction constant, ρ , depends on the nature of the reaction including conditions such as solvent and temperature and measures the susceptibility of the reaction to polar effects.

The ionization of benzoic acids in water at 25° is chosen as a standard process, for which ρ is defined as unity. The value of σ for a given substituent is the logarithm of the ratio of the ionization constant of the substituted benzoic acid to that of benzoic acid itself. Figure 1 is an example of Hammett relationship which shows the correlation between the Hammett σ and the basic hydrolysis rates of ethyl benzoates bearing meta- or para-substituents in aqueous ethanol.

Just like as this example, the Hammett equation can be applied successfully to a considerably large number of reactions, but you can find also the failure of the application for almost equal or may be larger numbers of reactions. However, we have to note that the deviations from the correlation appears to be rather systematic and to follow some general rules.

the same scale of the resonance contribution. In this point of view, we attempted to correlate various sets of the apparent substituent constants in terms of a simple but theoretically reasonable relationship.

Following the changes of para substituent constants independently derived from various electrophilic reactions, we found that the increments of respective sets of σ^+ from the Hammett σ could be correlated linearly one another. They are empirically represented by the equation,

$$(\sigma_A^+ - \sigma) = r(\sigma_B^+ - \sigma)$$

where σ_A^+ represents one of various sets of σ^+ and σ_B^+ another set σ^+ . The increments $\sigma^+ - \sigma$ can be replaceable with any other set of the increments using an appropriate r value.

The increment can be regarded as a measure of the capacity of respective substituents for the resonance interaction with the reaction site. r-Parameter may be referred to the varied degree of their resonance interaction of reactions.

Combining with the Hammett equation, we can derive the equation,

$$\log k/k_0 = \rho(\sigma + r\Delta\sigma_R^+) \dots \dots \dots (2)$$

where γ is a reaction constant describing the degree of the resonance interaction between substituents and the reaction center at the transition state and $\Delta\sigma_R^+$ corresponds to a proper set of the difference $\sigma^+ - \sigma$.

Here, we defined the resonance substituent constant, $\Delta\sigma_R^+$, conveniently using the set of $\sigma^+ - \sigma$ in Brown and Okamoto's scale.

This equation was employed successfully for more than 30 representative electrophilic reactions and the generality was demonstrated in the original paper. Thereafter, this relationship has been utilized satisfactorily for the elucidation of various reactivities by a number of authors. That is the history of the development of our empirical relationship.

Now, it is well known that the Hammett's σ -values themselves involve the σ^+ -type resonance contribution. The substituent constant free from the mutual conjugation between substituent and the reaction site was proposed by Taft as σ^0 and independently by Wepster as σ^n on the different basis.

The resonance substituent constant, $\Delta\sigma_R^+$, in our equation might be defined by the difference between any two sets of substituent constants of electrophilic reactions. It should be no matter whether both standard reactions involve the exalted resonance contribution or not. This is evidently supported by the general applicability of the above equation. However, in view of further development and analysis of our treatment and parameters, we have changed the standard from the Hammett σ to σ^0 constants.

The increment of Brown's σ^+ from the σ^0 runs parallel to $\Delta\sigma_R^+$ for respective para electron-releasing substituents. The correlation can be represented by

$$\Delta\sigma_R^+ = \sigma^+ - \sigma = 0.74 (\sigma^+ - \sigma^0)$$

Furthermore, the difference $\sigma - \sigma^0$, the resonance exaltation in the Hammett σ , can be cor-

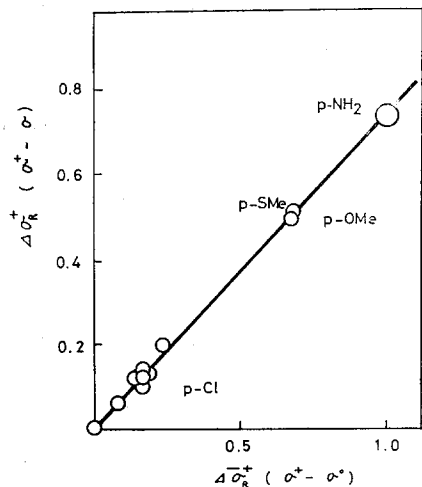


Fig. 2—Correlation of $\Delta\sigma_R^+$ vs $\Delta\sigma_R^+$.

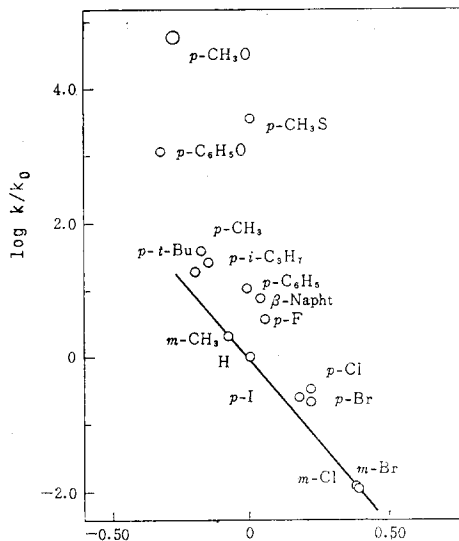


Fig. 3—Hammett plot of the hydrolysis of α -phenylethyl chlorides.

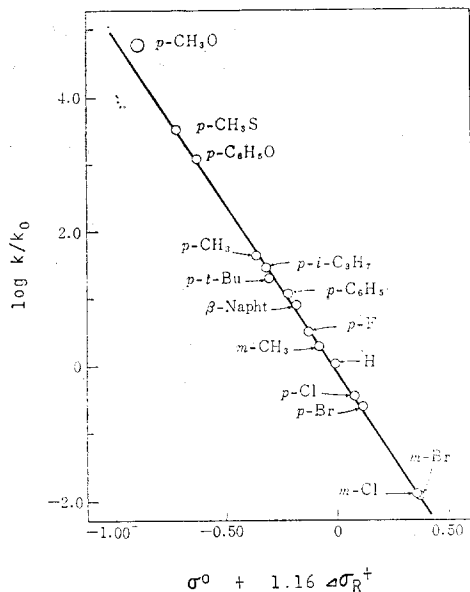


Fig. 4—LArSR treatment of the hydrolysis of α -phenylethyl chlorides.

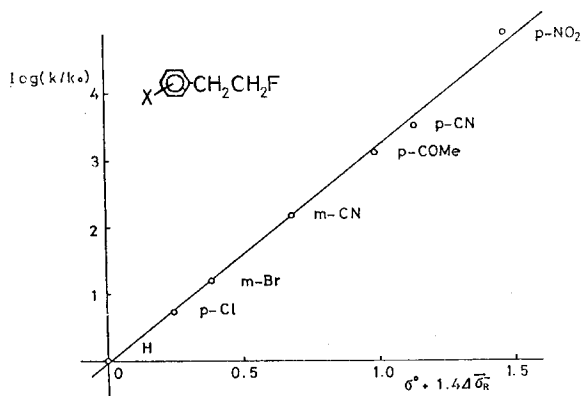


Fig. 5—LArSR treatment of the basic elimination of 2-phenylethyl fluorides.

related linearly with $\sigma^+ - \sigma^0$ as shown in Figure 2.

Now we can say that the variation of the resonance interaction can be described by a new set of resonance parameters defined by $\sigma^+ - \sigma^0$ in Brown and Okamoto's scale instead of $\sigma^+ - \sigma$. Thus, we derive the equation

$$\log k/k_0 = \rho(\sigma^0 + \gamma\Delta\sigma_R^+) \quad \dots\dots\dots (3)$$

where $\Delta\sigma_R^+$ corresponds to $\sigma^+ - \sigma^0$ in Brown's scale.

Similarly, nucleophilic exalted resonance reactivities were described by the same type of an equation, using $\Delta\bar{\sigma}_R^-$ instead of $\Delta\bar{\sigma}_R^+$. $\Delta\bar{\sigma}_R^-$ corresponds to the difference between σ^- and σ° . σ^- was defined as the relative pKas of substituted phenols.

$$\log k/k_0 = \rho(\sigma^\circ + r\Delta\bar{\sigma}_R^-) \dots \dots \dots (4)$$







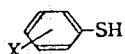
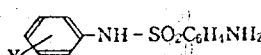
Recently, these treatments have been applied quite successfully to more than hundreds of reactivities, solvolytic reactions, aromatic substitution, and physical properties.

Figure 3 shows the Hammett plot of our results of the hydrolysis of meta- and para-substituted α -phenylethyl chlorides. This scattered pattern can be improved as in Figure 4 to an excellent linear plot, using the above equation. This correlation requires an r -value of 1.16. Similarly, the acetolysis of the same chlorides gave an excellent correlation with an r -value of 1.075.

Table I—The correlation of electrophilic reactions

Reaction	ρ	r	s	R	n
Molecular bromination	-9.49	1.42	0.192	0.999	8
Molecular chlorination	-7.95	1.22	0.35	0.993	14
Mercuration	-3.97	1.050	0.120	0.997	13
Acetylation	-9.52	0.813	0.336	0.992	14
Ethylation	-2.65	0.774	0.090	0.991	10
Protodesilylation, H ₂ SO ₄	-4.942	0.761	0.021	1.000	9
HClO ₄	-5.318	0.734	0.014	0.999	11
Protodegermylation, H ₂ SO ₄	-4.41	0.74	0.100	0.984	15
HClO ₄	-4.629	0.727	0.040	1.000	12
Protodestanylation	-3.723	0.561	0.015	0.999	12
Protodesilylation, 4'-substd. diphenyl	-1.325	0.424	0.24	0.999	7
Brown and Okamoto's σ^+	1.00	1.000			
pK _R ⁺ Benzhydrols	-9.14	1.22	0.43	0.999	7
Solv. Aryl-anisylchloro-methane, 85% aq. acetone,	-2.64	1.18	0.105	0.998	7
Solv. Arylethyl chloride 80% aq. acetone	-4.95	1.161	0.044	1.000	13
Solv. Benzhydryl chloride 70% aq. acetone	-4.58	1.14	0.102	0.993	10
Dissoc. Tritylchloride in liq. SO ₂	-4.042	0.885	0.048	0.998	12
pK _R ⁺ Trityl alcohol	-11.39	0.787	0.49	0.999	8
pK _{BH} ⁺ Acetophenone	-2.200	0.760	0.07	0.998	12
Decomp. Diazoacetophenone	-0.818	0.667	0.011	1.000	11
Acetolysis neophyl-OBs	-3.745	0.612	0.025	1.000	7
pK _{BH} ⁺ Benzoic acid	-1.167	0.558	0.057	0.991	13
pKa Benzoic acid	1.000	0.269	0.009	1.000	17
Benzoic acid with DDM	0.913	0.280	0.023	0.998	12

Table II—The correlation of nucleophilic reactions

Reaction	ρ	r	R	S	
 + picryl chloride	-2.19	1.14	0.9964	0.072	
 + picryl chloride	-3.89	0.561	0.9962	0.119	
 + imidazole-H ₂ O	1.736	0.610	0.9973	0.071	
 + MeO ⁻ -MeOH	4.824	0.621	0.9988	0.090	
	+ C ₆ H ₅ S ⁻	5.190	0.736	0.9941	0.227
	+ piperidine	4.287	0.738	0.9948	0.181
 + C ₂ H ₅ O ⁻	2.930	1.113	0.9984	0.081	
 + C ₂ H ₅ O ⁻	2.58	0.98	0.999	0.08	
pK _a 	2.558	0.649	0.9974	0.072	
pK _a 	1.843	0.627	0.9994	0.026	

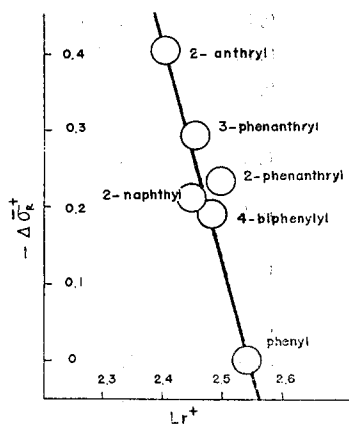
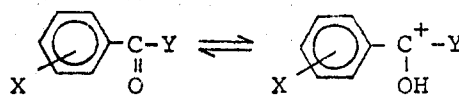
Fig. 6—Correlation of $\Delta\sigma_R^+$ vs. the localization energies, L_r^+ .

Table III—LArSR parameters for protonation of benzoyl derivatives



Y	ρ	r	log K _o
NH ₂	1.24	0.36	-2.16
OH	1.17	0.56	-7.26
CH ₃	2.20	0.76	-6.15
H	1.76	1.16	-7.10

Figure 5 shows the correlation of our data of the basic elimination of meta- and para-substituted 2-phenylethyl fluoride. This requires the nucleophilic parameter $\Delta\sigma_R^-$ and an r value as high as 1.40.

Some examples of the application of our equation (3) to electrophilic reactions are summarized in Table I, and those of the equation (4) to nucleophilic reactions are in Table II.

The polycyclic aromatic compounds can be regarded as a kind of substituted phenyl derivatives. Then, we may expect that the present treatment can be applicable to these derivatives just in the same way. Therefore, we derived the $\Delta\sigma_R^+$ values from the solvolysis rates of arylmethylcarbinyl chlorides, using ρ and r values for meta- and para-substituted α -phenylethyl chlorides, and σ° values obtained from hydrolysis of arylcarbinyl benzoates.

These parameters can be applied successfully to correlate the reactivities of polycyclic aromatic derivatives, for example, the protonation equilibria, pK_{BH^+} of arylmethyl ketones. All the points fall on a straight line, giving an r -value of 0.76. However, these treatment can be applied only to the unhindered derivatives do. Peri-derivatives not satisfy the correlation at all, because of their steric effects.

Another fact of interest is shown by the comparison of these resonance parameters with the reactivity indices of MO theory. Figure 6 is the linear correlation between $\Delta\sigma_R^+$ and the localization energies of aromatic compounds. This may be a theoretical support of our empirical equation.

Now, the properties of r -parameter should be a subject of interest. In our concept, r must be a measure of the degree of resonance interaction between conjugative substituents and the reaction center in the transition state.

Theoretically, both r and ρ should be responsible to the variation of the electron density at the reaction site, therefore, both parameters must be correlated each other. However, the variation of ρ is not always parallel to the variation of r . ρ appears too much sensitive to the conditions of reactions.

Table III shows the parameters for the correlations of the protonation of benzoyl derivatives. The conditions of each reaction are nearly the same for respective series. As Y changes from H to amino group, that is, from benzaldehyde to benzamide, r changes gradually from 1.16 to 0.36. This can be reasonably explained by the cross conjugation of aryl and Y groups. ρ value appears rather less appropriate measure than r -parameter for describing the nature of the reactions.

Meanwhile, the occurrence of steric as well as polar effects in aliphatic systems and ortho-substituted aromatic systems complicates the devising of correlation equations. Little progress was made until the early 1950s, when Taft made an excellent start in satisfactory correlation analysis in this area. Taft proposed E_s for steric parameter and σ^* for aliphatic polar parameter. These E_s and σ^* are very useful for aliphatic series. Moreover, he proposed σ_I for inductive-effect parameter and several σ_R , which are not unique, for resonance-effect parameter.

Taft assumed σ_I -meta equals to σ_I -para and σ_R is proportional to $\sigma_p \cdot \sigma_m$. σ_R has different values in respective reactions and they have no correlations with each other.

$$\begin{array}{ll} \sigma_p^\circ = \sigma_I + \sigma_R^\circ & \sigma_m^\circ = \sigma_I + 0.5 \sigma_R^\circ \\ \sigma_p = \sigma_I + \sigma_R & \sigma_m = \sigma_I + 0.3 \sigma_R \end{array}$$

$$\sigma_p^+ = \sigma_I + \sigma_R^+ \quad \sigma_m^+ = \sigma_I + 0.1 \sigma_R^+$$

We assumed, on the other hand, that the inductive parameter σ_i should be proportional to Roberts' σ' , which is then parallels to Taft's σ^* , and that the π -electronic effect parameter σ_π should be proportional to our $\Delta\bar{\sigma}_R^+$ or $\Delta\bar{\sigma}_R^-$. Following results were derived statistically from many reactions.

σ_i -meta is 1.17 times σ_i -para in all cases, whereas the proportional coefficient of σ_π -para varies markedly depending on the susceptibility of π -electronic effect in reactions and the one of σ_π -meta is constant. We defined σ° -para $\equiv \sigma_i + \sigma_\pi$ and σ° -meta corresponds to $0.87\sigma_i + 0.5\sigma_\pi$. Related equations are summarized below.

$$\log(k/k_0)_m = (0.87\sigma_i + 0.5\sigma_\pi)\rho$$

$$\log(k/k_0)_p = (\sigma_i + \alpha\sigma_\pi)\rho$$

$$\sigma_p^\circ = \sigma_i + \sigma_\pi \quad \sigma_m^\circ = \sigma_m = \sigma_m^+ = 0.87\sigma_i + 0.5\sigma_\pi$$

$$\sigma_p = \sigma_i + 1.65\sigma_\pi$$

$$\sigma_p^+ = \sigma_i + 3.41\sigma_\pi$$

Table IV shows the proposed values of σ_i and σ_π .

Here, I have no space to discuss about these topics in detail and just want to refer to the generalized form of our equation.

$$\delta\Delta F \text{ (or } \log k/k_0) = \rho(C_i\sigma_i + C_R^+\sigma_R^+ + C_R^-\sigma_R^-) \dots \dots \dots (4)$$

An example of a correlation only with σ_π is shown in Fig. 7, in which para-C-13 NMR chemical shifts of substituted benzenes are plotted to σ_π .

These substituent constants have found application in fields very different from those of

Table IV—Selected values of σ_i and σ_π

Subst.	σ_i	σ_π	Subst.	σ_i	σ_π^-
NMe ₂	0.06	-0.54	NO ₂	0.478	0.332
NH ₂	0.06	-0.42	CN	0.420	0.226
OH	0.19	-0.34	COMe	0.22	0.266
MeO	0.185	-0.281	COOMe	0.233	0.219
Me	-0.045	-0.078	CF ₃	0.341	0.143
Et	-0.045	-0.069	SO ₂ Me	0.458	0.220
t-Bu	-0.060	-0.054			
F	0.37	-0.16			
Cl	0.348	-0.070			
Br	0.337	-0.054			
Ph	0.11	-0.090			

$$\log(k/k_0)_p = \rho(\sigma_i + C_R^+\sigma_R^+ + C_R^-\sigma_R^-)$$

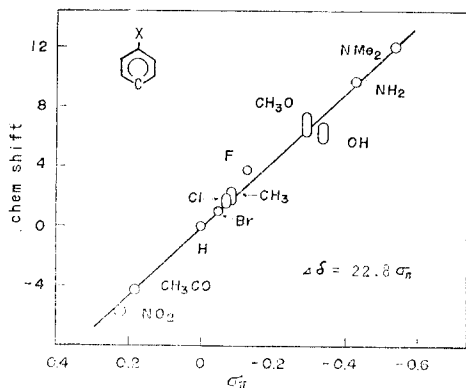


Fig. 7—Correlation of C-13 NMR chemical shifts at para positions vs. σ_p .

Table V—Selected values of π^\dagger

Substituent	π_o	π_m	π_p
F	0.01	0.13	0.15
Cl	0.59	0.76	0.70
Br	0.75	0.94	1.02
I	0.92	1.15	1.26
Me	0.68	0.51	0.52
Et	1.22	0.97	—
Pr ⁱ	—	1.30	1.40
Bu ^t	—	1.68	—
COMe	0.01	-0.28	-0.37
OH	—	-0.49	-0.61
OMe	-0.33	0.12	-0.04
NO ₂	-0.23	0.11	0.24

[†] Based on substituted phenoxyacetic acids. From Fujita, Iwasa, and Hanch(1964).

rates and equilibria of organic reactions. They have been applied extensively in optical spectroscopy, such as infrared, visible, and ultraviolet spectroscopies, nuclear magnetic resonance spectroscopy, such as ^1H , ^{19}F , and other nuclei, and mass spectrometry of organic compounds. Further, with appropriate additional considerations, the fields of enzymology and of structure-activity relationships for drugs have been considerably illuminated by the application of parameters derived from organic chemistry.

As pharmacologists must have interests in drug actions or biological activities of organic compounds, some comments about them shall be added in this lecture. This part was not my original works but is quoted from Hansch and Fujita's works. Two excellent reviews on this topic were recently published by Fujita and Shorter. I will quote many expressions from them and I would like to express grateful acknowledgement to these authors.

A biological response is the outcome of a complex series of events. The drug is administered in an aqueous phase outside the biophase and is then transported through various fluids and membranes in the biological material to the site of action. It becomes bound to the appropriate receptor and perturbs it. The perturbation issues in a biological response. In some studies (*in vitro*) the response is in an organ excised from its proper environment, for example, muscular or nerve tissue. Alternatively a living organism may be used (*in vivo* study). In all cases, particularly *in vivo*, the biological response is separated from the administration of the drug by many and varied processes. When we apply correlation analysis, and particularly if we use LFER parameters from organic chemistry, we assume hopefully that somewhere along the line there will be a process (or processes) whose free-energy change governs the biological response and which may depend on electronic, steric, and other factors similar to those operating in inanimate matter. Remarkably this hope often appears to be justified.

Pharmacologists, however, have long realised that processes of membrane penetration and

Table VI—Correlation of biological activities with P or π^+

Biological activity	Equation	n [†]	r [§]
Inhibition of α -chymotrypsin by substituted phenols	$\log(1/K_i)=0.95 \log P-1.88$	10	0.99
Inhibition of <i>Avena</i> cell elongation by substituted phenylacetic acids	$\log(1/C)=0.73\pi+3.01$	18	0.97
Inhibition of guinea pig ileum contractility by n-alkanols	$\log(1/C)=1.06\log P+0.62$	8	0.99
Narcotic action of miscellaneous aliphatic compounds on tadpoles	$\log(1/C)=1.17 \log P+0.68$	17	0.98

† From Cammaria and Rogers(1972).

‡ Number of compounds studied.

§ Correlation coefficient.

transport are often of dominant importance, that is, the passage of the drug from the aqueous phase into the macromolecular material of the biophase. The structural factors influencing this are not those usual to organic chemistry but involve the hydrophobic-lipophilic character of molecules.

Recently, Hansch approach has been widely accepted and recognized as a versatile way to understand drug action by analyzing the structure-activity relationship in various biological systems.

The partition of an organic solute between an organic solvent and water provides a possible measure of hydrophobic lipophilic character. The partition coefficient for distribution between n-octanol and water is most commonly used. This is denoted by P, and as an equilibrium constant, is used in correlation analysis as $\log P$. For aromatic systems, a substituent index π is defined by this equation,

$$\pi = \log P_X - \log P_H$$

where P_X is the partition coefficient for a compound with X as substituent, and P_H is that for the parent compound. For various reasons the most commonly used π -scale is based on substituted phenoxyacetic acids. Selected values are in Table V. P and π characterize the whole molecule, and for any extended range of compounds or substituents there are no clear relationships to ordinary polar or steric parameters.

For many biological activities, hydrophobic-lipophilic character by itself accounts for most of the observed structure-activity dependence. In such cases either the transport of the drug through membranes and biological fluids or the binding of the drug to the biological site is dominant. The binding of small molecules to biomacromolecules can sometimes be studied in isolation and is then found to depend on hydrophobic-lipophilic character.

In Table VI are summarized correlations with P or π for various biological activities, with examples from enzymes, cells, tissues, and simple intact organisms.

In certain other systems the plot of $\log(1/C)$ against $\log P$ or π is parabolic, that is, there is an optimum hydrophobic-lipophilic character. If the lipophilic character is too

marked the molecules tend to be taken up by the first biomacromolecule they encounter and get no further. Such systems can often be correlated if terms in $(\log P)^2$ or π^2 are included. The procedure was originally highly empirical but Hansch has justified it in terms of a kinetic model.

There are, however, many systems for which consideration of electronic and/or steric effects seems essential.

The LD₅₀ values of fourteen meta- or para-substituted phenyl diethyl phosphates towards houseflies are given by the first equation with a correlation coefficient of 0.97.

$$\log(1/C) = 2.08\sigma - 0.35$$

A parallel study of fly-head cholinesterase inhibition by six of the para-substituted compounds yielded I₅₀ values which were correlated by the second one with a correlation coefficient of 0.98.

$$\log(1/C) = 2.37\sigma + 4.38$$

The similar ρ -values suggest that the toxicity of the compound is a reflection of cholinesterase inhibition.

A related type of analysis uses pK_a values of drugs as polar parameters. Successful use of pK_a values can arise because at the pH value prevailing in the biological study, there are appreciable amounts of both ionized and unionized forms of drug present, but only the unionized form can enter the appropriate part of the biophase. There can then be a multiple correlation of biological activity with pK_a and the partition coefficient of the unionized form. For instance, the antibacterial activities (MIC) of fifteen arylamines towards *Escherichia coli* are expressed by this third equation with a correlation coefficient of 0.962.

$$\log(1/C) = -0.158 \text{ pK}_a + 0.694 \log P + 4.462$$

More generally, the free-energy related parameters are used such as π or $\log P$, σ or σ^* , and E_s . The most widely used equation in this approach is formulated as in equation (6), where $a (>0)$, b , ρ , δ , and c are constants which are determined by the regression analysis using the least-squares method.

$$\log 1/C = -a(\log P)^2 + b \log P + \rho\sigma + \delta E_s + c \dots \dots \dots (6)$$

In this way it is possible to analyze how each of the physicochemical properties of the molecule is concerned with the drug action.

The examples I showed you are only a few of numerous instances which have been found so far. In many cases it is possible to find clues to develop the studies on the mechanism of drug action by examining differences or common features among structure-activity and structure reactivity relationships. Even if detailed microscopic mechanisms on the overall processes of drug action are not identified explicitly, explanations of substituent effects and sometimes information about critical process or processes determining the activity can be obtained by comparison with those of simpler and similar model reactions and/or drug actions with the use of free-energy related parameters.

Even more important, the equations are a convenient way to store retrieve information on structure-activity relationships when combined with the use of a large computer. If the enormous amount of information accumulated so far is stored in the computer, one can sort out a set of equations according to their characteristic features such as the magnitude and sign of the coefficient of each terms and intercept. In this way the structure-activity correlations obtained for a certain series of congeneric drugs on various biological systems and for various series of drugs on a certain biological system or for various series of drugs on various biological systems can be classified and compared in terms of their physicochemical significance. The computerized approach should be a serious step toward quantitative comparative pharmacodynamic and may be useful for drug design.