Communications to the Editor

Chem. Pharm. Bull. 30(3)1091—1093(1982)

CHARACTERIZATION OF NOVEL SUBSTITUENT ENTROPY CONSTANT BY CLUSTER ANALYSIS

Tatsuya Takagi,[a] Akihiro Iwata,[a] Yoshio Sasaki,*,[a] and Hideko Kawaki[b]

Faculty of Pharmaceutical Sciences, Osaka University,[a] Yamadaoka 1-6, Suita 565, Japan and Faculty of Pharmacy, Kinki University,[b] Kowakae 3-4-1, Higashi-Osaka 577, Japan

Eighteen kinds of current QSAR parameters were analysed by the cluster analysis, and divided into 4 groups, namely, cluster 1 = σ_i, $|\sigma_i|$, σ_m, F; cluster 2 = σ_π, σ_p, R; cluster 3 = MR, Pr, Vw, σ_{s°, ΔS°, π, Es, Mw; cluster 4 = $|\sigma_\pi|$, ι, μ. Of the eight kinds of parameters belonging to cluster 3, MR, Pr, and Vw can be substituted by σ_{s°, whereas π (aromatic group) and Es are expressed by the linear combination of σ_i, σ_π and σ_{s°.

KEYWORDS ——— substituent entropy constant σ_{s°; QSAR analysis; cluster analysis; substituent constant; regression analysis; nonlinear mapping

In the previous works[1] on the studies of QSAR analysis, we have proposed a substituent entropy constant σ_{s° and approved the utility of this parameter as a quadratic form, where the biological responce BR of the drug was expressed as Eq.1.

$$BR = a(\sigma_{s^\circ})^2 + b\sigma_{s^\circ} + c|\sigma_i| + d|\sigma_\pi| + e \qquad (1)$$

This communication dealt with the characterization of σ_{s° using the cluster analysis and nonlinear mapping (NLM) of the eighteen kinds of current QSAR parameters including σ_{s°. The results show that they are classifiable into four clusters (Table 1). Eighteen-space coordinates of QSAR parameters are displayed in two-dimensional space using the NLM method[2] (Fig.1).

Table I. Cluster Analysis of the Eighteen Kinds of QSAR Parameters

cluster 1	σ_i[3] $\|\sigma_i\|$ σ_m F[4]	σ-inductive or field effect
cluster 2	σ_π[3] σ_p R[4]	π-resonance effect
cluster 3	MR[5] Pr[5] Vw[5] Es Mw[5] ΔS° σ_{s° π[5]	steric effect (entropy term)
cluster 4	$\|\sigma_\pi\|$ ι[6] μ[6]	others

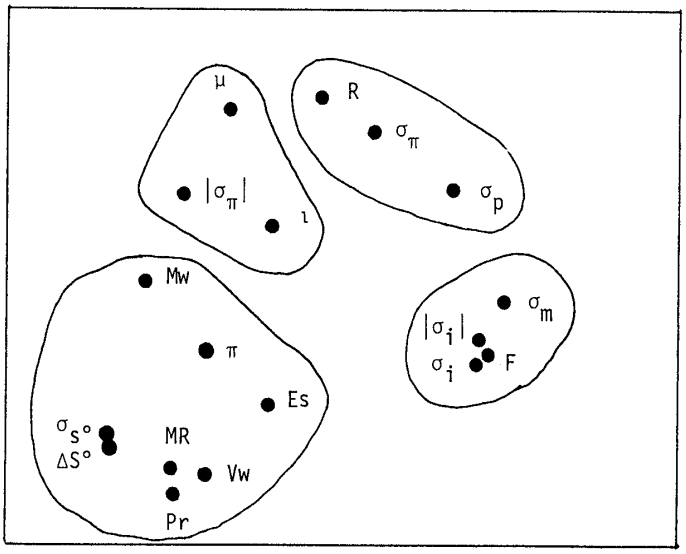

Fig.1. Nonlinear Map of Eighteen Kinds of Current QSAR Parameters

Of the eight kinds of empirical parameters belonging to cluster 3, π and Es are taken as the parameters related to the energy term, but others are not. Consequently, the two parameters should be analysed by the linear combination of the enthalpy and entropy parameters, e.g. σ_i, σ_π[3] and $\sigma_{s\circ}$. π of aromatic compounds was analysed in the previous works.[1] In this communication, we report the analysis of the aliphatic π.

The results are as follows[7];

$$\pi = 11.42\,\sigma_{s\circ} - 0.90 \qquad r = 0.937^{**} \quad SE = 0.098 \qquad (2)$$

n = 22 MeR R = H, F, OH, OMe, Cl, Br, Me, Et, n-Pr, n-Bu, n-Am, n-Oct, i-Bu, t-Bu, i-Pr, CF_3, Ph, $PhCH_2$, $PhCH_2CH_2$, Cyclopentyl, Cyclohexyl, $Ph(CH_2)_3$

$$Es = -10.56\,\sigma_{s\circ} - 2.35\,\sigma_\pi - 0.08 \qquad (3)$$

$r = 0.925$ $F = 20.99^{**}$ SE = 0.14

n = 10 PhR R = H, Me, Et, t-Bu, OH, OMe, NH_2, NMe_2, F, NO_2

$|\sigma_\pi|$ denotes the contribution of a weak interaction, σ_π indicates a strong one.

On the other hand, the entropy parameter $\sigma_{s\circ}$ or ΔS° is thermodynamically most applicable to cluster 3, so we analysed the non-energy parameters with respect to $\sigma_{s\circ}$ and obtained these results:

$$Vw = 3.855\,\sigma_{s\circ} - 0.020 \qquad r = 0.941^{**} \quad SE = 0.025 \qquad (4)$$

n = 12 PhR R = H, Me, Et, n-Pr, i-Pr, t-Bu, Ph, $CH{=}CH_2$, F, Cl, Br, I

$$MR = 87.82\,\sigma_{s\circ} + 0.26 \qquad r = 0.853^{**} \quad SE = 0.95 \qquad (5)$$

n = 13 PhR R = H, Me, Et, n-Pr, NMe_2, t-Bu, F, Cl, Br, I, C≡CH, OMe, NO_2

$$Pr = 888.8\,\sigma_{s\circ} - 3.8 \qquad r = 0.876^{**} \quad SE = 6.8 \qquad (6)$$

n = 12 PhR R = H, Me, Et, Ph, C≡CH, $CH{=}CH_2$, F, Cl, Br, I, CN, NO_2

The details of this work will be published in a subsequent paper.

EXPERIMENTAL

1. Cluster Analysis

Cluster Analysis was carried out by the farthest neighbor method using Euclidean distances calculated from the contribution ratio matrix of eighteen kinds of QSAR parameters.

2. Nonlinear Mapping (NLM)

NLM was carried out using the same data as for the cluster analysis. Two-dimensional space coordinates were obtained by our own programs using the SIMPLEX method[8] and the Davidson-Fletcher-Powell method[9] for minimization of the stress and by the NEC GDSP-6 system for plotting two-dimensional space coordinates.

3. Regression Analysis

The regression analysis were carried out using the program NEC TSS LIBRARY TSS/LIB-6 and our own programs coded in BASIC language. The values of SE are obtained from the equation $SE = [S_{se}/(n-k)(n-k-1)]^{1/2}$, where n = number of observations, k = number of independent variables and S_{se} = sum of squres of residuals.

4. Calculation

All numerical treatments were carried out with a NEAC S-900 computer at the Computation Center, Osaka University; a FACOM M-200 computer at the Computation Center, Kyoto University; a HITAC M-200H computer at the Computation Center, Tokyo University and a personal computer NEC PC8001.

REFERENCES AND NOTES

1) a) Y.Sasaki, T.Takagi, Y.Yamazato, A.Iwata and H.Kawaki, Chem.Pharm.Bull., 29, 3073 (1981); b) *Idem*, Abstract Papers of the 8th Symposium on Structure-Activity Relationships, p.65, Tokyo (1981); c) *Idem*, Abstract Papers of 3rd. Symposium on Medicinal Chemistry, p.1, Osaka (1981).

2) a) J.M.Samon,Jr., IEEE Trans Comput., C-18, 401 (1969); b) B.R.Kowalski and C.F.Bender, J.Am.Chem.Soc., 94, 5632 (1974); c) B.R.Kowalski and C.F.Bender, J.Am.Chem.Soc., 95, 686 (1975).

3) a) Y.Yukawa and Y.Tsuno, Nippon Kagaku Zasshi, 86, 873 (1965); b) M.Sawada, M.Ichihara, Y.Yukawa, T.Nakachi and Y.Tsuno, B.C.S.Jpn., 53, 2055 (1980).

4) C.G.Swain and E.C.Lupton,Jr., J.Am.Chem.Soc., 90, 4328 (1968).

5) MR:molar refructivity, Pr:parachor, Vw:van der Waals volume, Mw:molecular weight, π:hydrophobic substituent constant. All the values of these parameters, Es, σ_p and σ_m are cited from references.[10]

6) a) N.Inamoto and S.Masuda, Tetrahedron Lett., 1977, 3287; b) N.Inamoto and S.Masuda, Kagaku no Ryoiki, 33, 105 (1979).

7) All regression coefficients in this communication passed the t-test (p=0.01). **: The null hypothesis, H_0, is abandoned (p=0.01). H_0: $r = 0$ or $a_1 = a_2 = \cdots = 0$.

8) J.A.Neider and R.Nead, Computer J., 6, 163 (1963).

9) R.Fletcher and M.J.D.Powell, Computer J., 7, 308 (1965).

10) C.Hansch and A.Leo, "Substituent Constants for Correlation Analysis in Chemistry and Biology", John Wiley and Sons, Inc., New York, 1979 and references therein.

(Received January 5, 1982)