Estimation of the Quantitative Structure–Activity Relationship Descriptor σ_{S° for Di- and Tri-Substituted Benzene Derivatives

Yoshio SASAKI,[a] Hideko KAWAKI,[b] and Tatsuya TAKAGI,*[,a]

Faculty of Pharmaceutical Sciences, Osaka University,[a] 1–6 Yamada-Oka, Suita, Osaka 565, Japan and Faculty of Pharmacy, Kinki University,[b] 3–4–1 Kowakae, Higashi-Osaka, Osaka 577, Japan. *Received January 22, 1990*

The quantitative structure–activity relationship descriptor σ_{S°, representing the contributions from dispersion and repulsion interactions can be expressed as follows: (1) for disubstituted benzene derivatives, $\sigma_{S^\circ}(1,2)=0.831\sum\sigma_{S^\circ}(\text{mono})-0.004$, $\sigma_{S^\circ}(1,3)=0.855\sum\sigma_{S^\circ}(\text{mono})-0.007$, and $\sigma_{S^\circ}(1,4)=0.874\sum\sigma_{S^\circ}(\text{mono})-0.018$; (2) for trisubstituted benzene derivatives, $\sigma_{S^\circ}(1,3,5)=0.752\ \sum\sigma_{S^\circ}(\text{mono})-0.017$, $\sigma_{S^\circ}(1,2,4)=0.715\sum\sigma_{S^\circ}(\text{mono})+0.006$, and $\sigma_{S^\circ}(1,2,3)=0.723\times\sum\sigma_{S^\circ}(\text{mono})-0.006$.

These are revealed to be successful for the estimation of the descriptors σ_{S° for optional di- and tri-substitued benzene derivatives, and meet the needs of the practical quatitative structure–activity relationships.

Keywords Quantitative structure–activity relationship; entropy constant σ_{S°; disubstituted benzene derivative; trisubstituted benzene derivative; regression analysis

Introduction

In a previous report[1a] on the weak intermolecular forces from the modified Lennard–Jones 12,6 potential, the novel quantitative structure–activity relationship (QSAR) descriptors σ_{S° and μ^2/α were presented, and the former were revealed to be effective in representing both dispersion and repulsion interactions. They are derived from the absolute entropy $S^\circ_{298}(g)$, but are inadequate for poly-substituted benzene derivatives. In another earlier work[1b] in this series, in order to overcome this weak point, we especially determined the descriptors for disubstituted benzene derivatives, together with trisubstituted benzene series having the same kind of substituent group.

Here, from the linear relation between σ_{S°(mono) and the increment $\Delta\sigma_{S^\circ}$, given by subtracting σ_{S° of monosubstituted benzenes from those of *o*-, *m*-, and *p*-disubstituted series, the descriptors σ_{S°(AB) were revealed to be linear against $\sum\sigma_{S^\circ}$ (mono). Furthermore, the authors extended the above method for trisubstituted benzene derivatives having optional substituent groups, and acknowledge successful results. Namely, the descriptors σ_{S°(ABC) and $\sigma_{S^\circ}(A_2B)$ are all linear against $\sum\sigma_{S^\circ}$(mono).

This simple procedure affords an available means for the estimation of the descriptor σ_{S° for optional di- and trisubstituted benzene derivatives, and can be expected to meet the needs of QSAR, gas-liquid chromatography (GLC), ecological toxicology, *etc.*, assuming a weak intermolecular interaction.

Experimental

Increment $\Delta\sigma_{S^\circ}$ from Mono- and Di-Substituted Benzene Derivatives Data of descriptor σ_{S° used in this work were all cited from our previous reports,[1a,b] and the increments $\Delta\sigma_{S^\circ}$ on *o*-, *m*-, and *p*-positions of disubstituted benzene series were calculated by subtracting σ_{S° of monosubstituted benzenes from those of their corresponding disubstituted series.

Regression Analysis The details are the same as given in our preceding report.[1c]

Results and Discussion

Presentation of Increment $\Delta\sigma_{S^\circ}$, and Descriptors σ_{S°(AB or A_2) for Disubstituted Benzene Derivatives The descriptors σ_{S°(AB or A_2) for *o*-, *m*-, and *p*-disubstituted benzene derivatives summarized in our previous report[1b] can be given by Eq. 1:

$$\sigma_{S^\circ}(AB)=\sigma_{S^\circ}(A)+\Delta\sigma_{S^\circ}(B) \quad (A\neq B \text{ or } A=B) \tag{1}$$

The term $\Delta\sigma_{S^\circ}$(B or A) means an increment given by the second substituent. Namely, the increment $\Delta\sigma_{S^\circ}$ on *o*-, *m*-, and *p*-positions of disubstituted benzene series could be given by subtracting σ_{S°(A) of monosubstituted benzene series from their *o*-, *m*-, and *p*-disubstituted ones. Furthermore, as written in Table I, the descriptors σ_{S°(AB) are also linear against the sum of $\sigma_{S^\circ}(A)+\sigma_{S^\circ}(B)$ (*cf.* Table I), because the relation $\Delta\sigma_{S^\circ}(B)\propto\sigma_{S^\circ}(B)$ is established. This evidence is supported by the examples summarized in Table

TABLE I. Values σ_{S° and $\Sigma\sigma_{S^\circ}$ (mono) for Disubstituted Benzene Derivatives $C_6H_4R_1R_2$

R_1	*o*-NO_2–		*m*-F–		*p*-COOEt–	
R_2	σ_{S°	$\Sigma\sigma_{S^\circ}$ (mono)	σ_{S°	$\Sigma\sigma_{S^\circ}$ (mono)	σ_{S°	$\Sigma\sigma_{S^\circ}$ (mono)
NO_2	0.183	0.230	0.125[a]	0.166	0.283	0.325
CN	0.152	0.191	0.102	0.127	0.251	0.286
COMe	0.208	0.256	0.158	0.192	0.306	0.351
COEt	0.248	0.295	0.198	0.231	0.346	0.399
CF_3	0.208	0.256	0.158	0.192	0.301	0.351
COOMe	0.251	0.300	0.201	0.236	0.348	0.404
COOEt	0.283	0.334	0.233	0.270	0.380	0.438
F	0.122[a]	0.166	0.076[a]	0.102	0.229	0.261
Cl	0.139[a]	0.181	0.100[a]	0.117	0.242	0.276
Br	0.153[a]	0.196	0.106	0.132	0.254	0.291
I	0.166[a]	0.209	0.117	0.145	0.265	0.304
Me	0.152	0.191	0.102	0.127	0.250	0.286
Et	0.196	0.242	0.146	0.178	0.293	0.337
OMe	0.195	0.241	0.146	0.178	0.294	0.337
OEt	0.235	0.285	0.185	0.221	0.333	0.389
OH	0.146	0.184	0.096	0.120	0.244	0.279
NH_2	0.150	0.189	0.098	0.125	0.248	0.284
NMe_2	0.201	0.249	0.151	0.185	0.299	0.344

a) Observed values. Regression analyses afford the following results:

$\sigma_{S^\circ}(o\text{-}NO_2\text{–}R_2\text{–})=0.927(\pm0.023)\Sigma\sigma_{S^\circ}(\text{mono})-0.028(\pm0.005)$
$n=18$, $r=0.999$, $F=7267$, S.D. $=0.002$

$\sigma_{S^\circ}(m\text{-F–}R_2\text{–})=0.905(\pm0.041)\Sigma\sigma_{S^\circ}(\text{mono})-0.014(\pm0.007)$
$n=18$, $r=0.996$, $F=2215$, S.D. $=0.004$

$\sigma_{S^\circ}(p\text{-COOMe–}R_2\text{–})=0.839(\pm0.015)\Sigma\sigma_{S^\circ}(\text{mono})+0.009(\pm0.005)$
$n=18$, $r=0.999$, $F=13403$, S.D. $=0.002$

II, where $\Delta\sigma_{S^\circ}$(R) of *o*-NO_2–, *m*-Cl– and *p*-NMe_2–series are expressed as coordinate variables of σ_{S°(mono).

Consequently, for the 171 congeners referred to in our previous report,[1b] the three regression equations were confirmed as given below:

$$\sigma_{S^\circ}(12)=0.831(\pm0.005)\sum\sigma_{S^\circ}(\text{mono})-0.004(\pm0.283) \quad (2)$$
$$n=171,\ r=0.999,\ F=1.3\times10^5,\ \text{S.D.}=0.002$$

$$\sigma_{S^\circ}(13)=0.835(\pm0.004)\sum\sigma_{S^\circ}(\text{mono})-0.007(\pm0.156) \quad (3)$$
$$n=171,\ r=1.000,\ F=1.3\times10^5,\ \text{S.D.}=0.002$$

$$\sigma_{S^\circ}(14)=0.874(\pm0.005)\ \sum\sigma_{S^\circ}(\text{mono})-0.018(\pm0.065) \quad (4)$$
$$n=171,\ r=0.999,\ F=1.4\times10^5,\ \text{S.D.}=0.002$$

Equations 2, 3, and 4 enable us to estimate the descriptors σ_{S° for optional disubstituted benzene derivatives.

Descriptors σ_{S°(ABC or AB_2) for Trisubstituted Benzene Derivatives The observed data[2] of S°_{298}(g) for only 13 congeners of the same kind are known among trisubstituted benzene derivatives, and are transformed to σ_{S° and summarized in Table III, together with the estimated values given in our previous report.[1b]

When we continue and extend our procedures adopted in the previous section for the trisubstituted benzene series, the following Eq. 5 should be established:

$$\sigma_{S^\circ}(\text{ABC})=\sigma_{S^\circ}(\text{A})+\Delta\sigma_{S^\circ}(\text{B})+\Delta\sigma_{S^\circ}(\text{C}) \quad (5)$$
$$\sigma_{S^\circ}(\text{AB}_2)=\sigma_{S^\circ}(\text{A})+\Delta\sigma_{S^\circ}(\text{1-B})+\Delta\sigma_{S^\circ}(\text{2-B})$$

where (1-B) and (2-B) indicate the first and the second substituent B.

For example, as written in Chart 1, the descriptor σ_{S° for 1-NO_2–2-OMe–4-Me–C_6H_3 could be estimated from 1,2,4-Me_3–C_6H_3 *via* 1-NO_2–2,4-Me_2–C_6H_3 by the successive displacement of Me by NO_2 and OMe. As shown in Chart 1, the estimations are carried out for the three cases, where one of the three kinds of substituents is regarded as a

Table II. Increment Values $\Delta\sigma_{S^\circ}$ for Disubstituted Benzene Derivatives $C_6H_4R_1R_2$ from σ_{S°(mono) Reference

R_1 / R_2	σ_{S° (mono)	*o*-NO_2– σ_{S°	*o*-NO_2– $\Delta\sigma_{S^\circ}$	*m*-Cl– σ_{S°	*m*-Cl– $\Delta\sigma_{S^\circ}$	*p*-NMe_2– σ_{S°	*p*-NMe_2– $\Delta\sigma_{S^\circ}$
NO_2	0.115	0.183	0.068	0.142[a]	0.076	0.199	0.068
CN	0.076[a]	0.152	0.037	0.116	0.050	0.166	0.032
CF_3	0.141[a]	0.208	0.093	0.171	0.105	0.223	0.089
COMe	0.141[a]	0.208	0.093	0.171	0.105	0.223	0.089
COOMe	0.185	0.251	0.136	0.214	0.148	0.267	0.133
F	0.051[a]	0.122[a]	0.007	0.100[a]	0.034	0.144	0.010
Cl	0.066[a]	0.139[a]	0.024	0.106[a]	0.040	0.157	0.023
Br	0.081[a]	0.153[a]	0.038	0.120	0.054	0.170	0.036
I	0.094[a]	0.166[a]	0.051	0.130	0.064	0.181	0.047
Me	0.076[a]	0.152	0.037	0.116	0.050	0.167	0.033
Et	0.127[a]	0.196	0.081	0.159	0.093	0.211	0.077
OMe	0.127[a]	0.195	0.080	0.159	0.093	0.211	0.077
OH	0.069[a]	0.146	0.031	0.110	0.044	0.160	0.026
NH_2	0.074[a]	0.150	0.035	0.113	0.047	0.164	0.030
NMe_2	0.134[a]	0.201	0.086	0.164	0.098	0.216	0.082

a) Observed values. Regression analyses afford the following results:

$\Delta\sigma_{S^\circ}$(*o*-NO_2–R_2–) = 0.914(±0.034)σ_{S°(mono) − 0.035(±0.004)
$n=15,\ r=0.998,\ F=3381,\ \text{S.D.}=0.002$

$\Delta\sigma_{S^\circ}$(*m*-Cl–R_2–) = 0.856(±0.045)σ_{S°(mono) − 0.015(±0.005)
$n=15,\ r=0.996,\ F=1673,\ \text{S.D.}=0.003$

$\Delta\sigma_{S^\circ}$(*p*-NMe_2–R_2–) = 0.900(±0.020)σ_{S°(mono) − 0.037(±0.002)
$n=15,\ r=0.999,\ F=9320,\ \text{S.D.}=0.001$

1-NO_2–2-OMe–4-Me– basic data	σ_{S°	1-NO_2–2,4-Me_2–	σ_{S°	1,2,4-Me_3–	σ_{S°
OMe	0.127[a]	NO_2	0.115	1,2,4-Me_3	0.167[a]
o-NO_2, OMe	0.195	*o*-NO_2, Me	0.152	Me	0.076[a]
p-NO_2, Me	0.148	*p*-NO_2, Me	0.148	*o*-Me_2	0.117[a]
m-Me, OMe	0.168			*m*-Me_2	0.123[a]
				p-Me_2	0.117[a]
method of calculation					
(1)		(1)		(1)	
σ_{s° 1-NO_2	0.115	1-NO_2	0.115	1-Me	0.076[a]
$\Delta\sigma_{s^\circ}$ 2-OMe	0.080	2-Me	0.037	2-Me	0.041
$\Delta\sigma_{s^\circ}$ 4-Me	0.033	4-Me	0.033	4-Me	0.041
Σ	0.228		0.185		0.158
(2)		(2)		(2)	
σ_{s° 2-OMe	0.127[a]	2-Me	0.076[a]	2-Me	0.076[a]
$\Delta\sigma_{s^\circ}$ 1-NO_2	0.068	1-NO_2	0.076	1-Me	0.041
$\Delta\sigma_{s^\circ}$ 4-Me	0.041	4-Me	0.047	4-Me	0.047
Σ	0.236		0.199		0.164
(3)		(3)		(3)	
σ_{s° 4-Me	0.076[a]	4-Me	0.076[a]	4-Me	0.076[a]
$\Delta\sigma_{s^\circ}$ 1-NO_2	0.072	1-NO_2	0.072	1-Me	0.041
$\Delta\sigma_{s^\circ}$ 2-OMe	0.092	2-Me	0.047	2-Me	0.047
Σ	0.240		0.195		0.164
average	0.235		0.193		0.162
σ_{s°(cal)	0.242		0.199		

Chart 1. Estimation of σ_{s° for 1-NO_2–2-OMe–4-Me–C_6H_3 from 1,2,4-Me_3–C_6H_3 *via* 1-NO_2–2,4-Me_2–C_6H_3

a) Observed values. (1), (2), and (3) denote the case number.

Table III. Observed and Calculated Descriptors σ_{S° for Trisubstituted Benzene Derivatives

	$\Sigma\sigma_{S^\circ}$	σ_{S° (123)	σ_{S° (124)	σ_{S° (135)
F_3	0.153	0.095[a]	0.109[a]	0.088[a]
Cl_3	0.198	0.138[a]	0.145[a]	0.129[a]
Br_3	0.243	0.172	0.182	0.171[a]
Me_3	0.228	0.155[a]	0.167[a]	0.156[a]
Et_3	0.381	0.275[a]	0.284[a]	0.276[a]
Me_2, NO_2	0.267	0.185/2,3-Me_2	0.199/2,4-Me_2	0.183/3,5-Me_2
		0.187/2,6-Me_2	0.200/2,5-Me_2	
			0.197/3,4-Me_2	
Me, OMe, NO_2	0.318	0.266/2-OMe–3-Me	0.242/2-OMe–4-Me	0.222/3-Me–5-OMe
		0.227/2-Me–3-OMe	0.243/2-Me–4-OMe	
		0.227/2-Me–6-OMe	0.243/3-Me–4-OMe	
			0.242/3-OMe–4-Me	
			0.242/2-OMe–5-Me	
			0.243/2-Me–5-OMe	

a) Observed values.

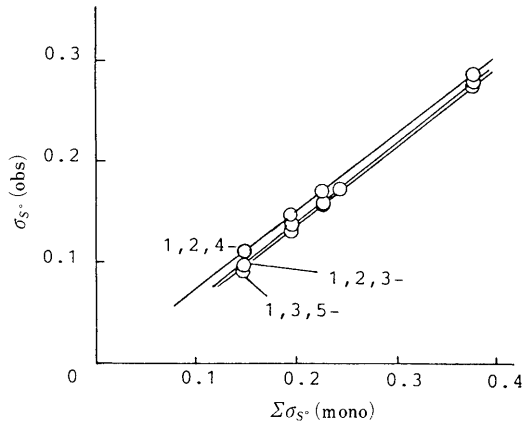

Fig. 1. Linear Relations between σ_{S°(obs) and $\Sigma\sigma_{S^\circ}$(mono) for Trisubstituted Benzene Series

TABLE IV. Values σ_{S° for Trisubstituted Benzene Derivatives $C_6H_3R_3$ and Their Descriptors $\Sigma\sigma_{S^\circ}$(mono)

R	σ_{S° 1,2,3-	1,2,4-	1,3,5-	$\Sigma\sigma_{S^\circ}$(mono)
NO_2	0.250	0.259	0.250	0.345
CN	0.161	0.171	0.157	0.228
COMe	0.309	0.318	0.312	0.423
COEt	0.399	0.407	0.405	0.540
CF_3	0.309	0.317	0.311	0.423
COOMe	0.405	0.413	0.410	0.555
COOEt	0.470	0.477	0.477	0.640
F	0.095[a]	0.109[a]	0.088[a]	0.153
Cl	0.138[a]	0.145[a]	0.129[a]	0.198
Br	0.172	0.182	0.171[a]	0.243
I	0.203	0.213	0.202	0.282
Me	0.155[a]	0.167[a]	0.156[a]	0.228
Et	0.275[a]	0.284[a]	0.276[a]	0.381
OMe	0.279	0.288	0.280	0.381
OEt	0.372	0.380	0.376	0.510
OH	0.142	0.153	0.137	0.207
NH_2	0.154	0.165	0.151	0.222
NMe_2	0.293	0.302	0.295	0.402

a) Observed values.

reference, respectively, and the mathematical means are presented. As a result, the reference 1,2,4-Me_3–C_6H_3 having σ_{S°(obs)=0.167 afforded the calculated σ_{S°(cal)=0.162, consequently, the calculated averages of 0.193 and 0.235 for 1-NO_2–2,4-Me_2–and 1-NO_2–2-OMe–4-Me–C_6H_3 become 0.199 and 0.242.

And, as the relations $\Delta\sigma_{S^\circ}$(1-B or 2-B) $\propto\sigma_{S^\circ}$(B) and $\Delta\sigma_{S^\circ}$(C) $\propto\sigma_{S^\circ}$(C) are already established, we are able to realize Eq. 6 as below:

$$\sigma_{S^\circ}(\mathrm{ABC}) \quad \text{or} \quad \sigma_{S^\circ}(\mathrm{AB_2}) \propto \Sigma\sigma_{S^\circ}(\mathrm{mono}) \tag{6}$$

The validity of Eq. 6 is supported by the data summarized in Table III, where the observed data are linear against $\Sigma\sigma_{S^\circ}$(mono) as depicted in Fig. 1, and the calculated data summarized in Tables III and IV are also plotted without exception on the reference lines.

In the next stage, we tried to express the values of σ_{S° for trisubstituted benzene derivatives $C_6H_3R_3$,[1b] having the same kind of substituent, by regression analyses using $\Sigma\sigma_{S^\circ}$(mono) as independent variable (*cf.* Table IV).

The results are as follows:

$$\sigma_{S^\circ}(123)=0.762(\pm0.0010)\Sigma\sigma_{S^\circ}(\mathrm{mono})-0.015(\pm0.004) \tag{7}$$
$$n=18,\ r=1.000,\ F=24701,\ \text{S.D.}=0.003$$

$$\sigma_{S^\circ}(124)=0.753(\pm0.008)\Sigma\sigma_{S^\circ}(\mathrm{mono})-0.002(\pm0.003) \tag{8}$$
$$n=18,\ r=1.000,\ F=38549,\ \text{S.D.}=0.002$$

$$\sigma_{S^\circ}(135)=0.790(\pm0.013)\Sigma\sigma_{S^\circ}(\mathrm{mono})-0.025(\pm0.005) \tag{9}$$
$$n=18,\ r=1.000,\ F=18802,\ \text{S.D.}=0.003$$

When we proceed on a line given by Eq. 6, the regression analyses for 54 congeners having substituent groups ABC and AB_2 at random afforded Eqs. 10, 11, and 12.

$$\sigma_{S^\circ}(123)=0.723(\pm0.004)\Sigma\sigma_{S^\circ}(\mathrm{mono})-0.006(\pm0.266) \tag{10}$$
$$n=54,\ r=1.000,\ F=121117,\ \text{S.D.}=0.002$$

$$\sigma_{S^\circ}(124)=0.715(\pm0.003)\Sigma\sigma_{S^\circ}(\mathrm{mono})-0.006(\pm0.224) \tag{11}$$
$$n=54,\ r=1.000,\ F=171840,\ \text{S.D.}=0.001$$

$$\sigma_{S^\circ}(135)=0.752(\pm0.005)\Sigma\sigma_{S^\circ}(\mathrm{mono})-0.017(\pm0.120) \tag{12}$$
$$n=54,\ r=1.000,\ F=80638,\ \text{S.D.}=0.002$$

In conclusion, the descriptor σ_{S° for di- and tri-substituted benzene derivatives are now easily prepared from Eqs. 10, 11, and 12, and the merit of this successful result can be reflected on the progress of QSAR, GLC, partition, solubility, *etc.*, assuming a weak intermolecular interaction. As written in a previous report,[1b] the substituent located on the *ortho*-position does not exert the so-called "*ortho* effect" on the descriptor σ_{S°. This evidence is approved for σ_{S°(123), as well as σ_{S°(124) and σ_{S°(12). In other words, the *ortho* effect observed in a weak intermolecular interaction should be reconciled with the charge-transfer, electrostatic or polarization effect, assuming the perturbations on both π-electron delocalization and dipole moment.

Acknowledgement Financial assistance from Hoan-sha is gratefully acknowledged.

References

1) *a*) Y. Sasaki, T. Takagi, and H. Kawaki, *Chem. Pharm. Bull.*, **36**, 3743 (1988); *b*) H. Kawaki, Y. Sasaki, T. Takagi, S. Fujii, and F. Masuda, *ibid.*, **37**, 3268 (1989); *c*) H. Kawaki, F. Masuda, and Y. Sasaki, *ibid.*, **36**, 4814 (1988).
2) *a*) D. R. Stull, E. F. Westrum, Jr., and G. C. Sinke, "The Chemical Thermodynamics of Organic Compounds," Wiley, New York, 1969; *b*) J. B. Butler and J. Lielmezs, *J. Chem. Eng. Data*, **14**, 335 (1969); *c*) H. Aleman and J. Lielmezs, *Thermochimica Acta*, **3**, 391 (1972).