"Aromatic" Substituent Constants for Structure-Activity Correlations†

Corwin Hansch,* Albert Leo, Stefan H. Unger, Ki Hwan Kim, Donald Nikaitani, and Eric J. Lien

Department of Chemistry, Pomona College, Claremont, California 91711. *Received June 18, 1973*

Aromatic substituent constants (lipophilic π, electronic σ_m and σ_p, and steric MR, molar refractivity) have been collected for 236 substituents including 128 π values and 191 values for which both σ_m and σ_p were found. Swain and Lupton's $\mathcal{F}$ and $\mathcal{R}$ values could then be calculated for these 191 substituents by a corrected procedure. The mutual correlation of σ_m and σ_p is high, $r = 0.903$, while $\mathcal{F}$ and $\mathcal{R}$ are essentially orthogonal.

Interest in the use of substituent constants for correlating structure with reactivity continues to grow rapidly in both simple organic[1] as well as complex biochemical[2] and biomedicinal[3a-d] systems. The use of σ constants (Hammett and Taft) for the electronic and E_s constants (Taft[3e]) for the steric effects of substituents has greatly facilitated understanding of organic reaction mechanisms.[1,3e,4] The hydrophobic parameters[5] π and log P have bridged the gap between simple organic and biochemical-biomedicinal systems. To further advance the extrathermodynamic approach[4b] to quantitative structure-activity relationships (QSAR),[3a-d] we have made a search of the literature to assemble as many substituents as possible for which both σ_m and σ_p have been determined. These appear in Table I together with measured π values and other parameters calculated as described below.

There are two reasons for focusing on these two particular electronic parameters[6] instead of any of the others. First, far more of these constants are available than others such as σ_I or σ_R.[7] Second, it is possible, using only σ_m and σ_p, to factor the electronic effect into resonance and nonresonance components.[6] Since Taft and Lewis[7] first explored factoring of this kind, it has become increasingly evident that it can lead to greater insight into substituent effects. Although there are a variety of ways in which such factoring can be accomplished,[6-8,‡] the general approach of Swain and Lupton[6] has the advantage that $\mathcal{F}$ and $\mathcal{R}$ can be calculated directly from σ_m and σ_p, avoiding nongeneral procedures.[8,‡] A further improvement by way of an optimization-orthonormalization procedure will be published shortly.‡ This modification has the advantage of depending upon a large amount of data instead of an "arbitrarily" selected model reaction(s); it contains only a single general chemical constraint and it guarantees completely independent (orthogonal) and equiscalar (normal) substituent vectors. Actually, $\mathcal{F}$ and $\mathcal{R}$ are remarkably orthogonal, as will be discussed below, and therefore they largely avoid the common pitfall of multicollinearity (even though they are neither "optimum" nor precisely orthogonal).

To provide a measure of hydrophobic character, Table I lists the known[5b] values of π from the benzene *solute* system partitioning in the octanol-water *solvent* system and also includes many new π values not previously reported. π constants have previously been measured using a variety of *solute* systems.§ The most common "parent" solutes (besides benzene) on which the substituents are placed are phenol, phenoxyacetic acid, benzoic acid, aniline, and nitrobenzene. The variation from system to system can often be predicted with acceptable precision,[5a] and it is generally not so great that it precludes using the π constants from benzene to serve in the design of new drugs and enzyme substrates where the variable substituent is attached to *any* aromatic ring. The π-/benzene values are much more independent of electronic contributions than, say, π-/nitrobenzene and are thus more likely to represent a true "lipophilicity" for aromatic substituents. Where the attachment is at an aliphatic site, the substituent π constant from benzene requires a correction for electronic effects and possibly for "folding" over the ring.[5b] Additivity of both π and σ on benzene rings is considered below.

The most widely used parameter for steric substituent effects in organic reaction mechanism studies is E_s.[3e] This parameter is useful for studying intramolecular steric effects, particularly in reactions where the substituent is near the reaction center. However, since E_s constants have not been determined for the majority of substituents listed in Table I and since biochemical-biomedical "steric" requirements are often (but not always) of the "bulk" type, we have sought another measure, albeit approximate, of the general "steric" bulk. Fortunately, there are two parameters which are readily available for each substituent: molar refraction (MR) and molecular weight (MW). Van der Waals molar volumes as calculated by Bondi[9] might also have been used but, for the full range of substituents listed in Table I, a great number of approximations and assumptions would have been necessary.

MR values have been used previously in some biological QSAR.[10] For liquids, the MR can be calculated (in units of volume) from the Lorentz-Lorentz relation=

$$MR = MW(n^2 - 1)/d(n^2 + 2) \quad (cm^3/mol) \quad (1)$$

where MW = molecular weight, n = index of refraction (normally at 20°, Na D line), and d = density (normally at 20°).

The tolerance by enzymes and receptors for the "bulkiness" of the substrates and drugs to which they are exposed is a problem of great concern[11] in biomedicinal-biochemical studies. We believe that MR and/or MW may be crude but useful measures of "bulk." *We wish to emphasize, however, that we consider these parameters as only a possible interim solution. MR contains an electronic contribution (it is directly proportional to the polarizability); therefore, its use in multiple regression analysis must be viewed with caution.* The MR values in Table I were included partly for the purpose of the cluster analysis which appears in the following paper.[12] In this they are somewhat less critical than in a multiple regression.

Methods

1. **Hydrophobic Parameters.** Log P values have been determined as previously described.[5a] In general, at least

†This work was supported by Grant CA 11110 from the National Institutes of Health and in part by Contract DRR-70-4115 from the National Institutes of Health.

‡C. G. Swain, S. H. Unger, P. Strong, and N. R. Rosenquist, unpublished results.

§To avoid confusion in use of the term "system" in partitioning studies, the following symbolism will be followed in this and subsequent papers from this laboratory. (1) π or log P is followed by a dash and the formula for the substituent or the solute, respectively. For π values, the substituent formula can be followed by a slash (/) and then the solute name or formula. (2) The *solvent* system is given in parentheses following the solute, but only the organic phase need be specified. If *no* solvent is specified, it is assumed to be octanol-water. For example, π-4-Cl/phenol (oleyl alcohol) = 0.82 refers to the π value for the chloro substituent in the para position from the phenol solute system measured in oleyl alcohol-water.

=References to MR are found in Table I, footnote *d*.

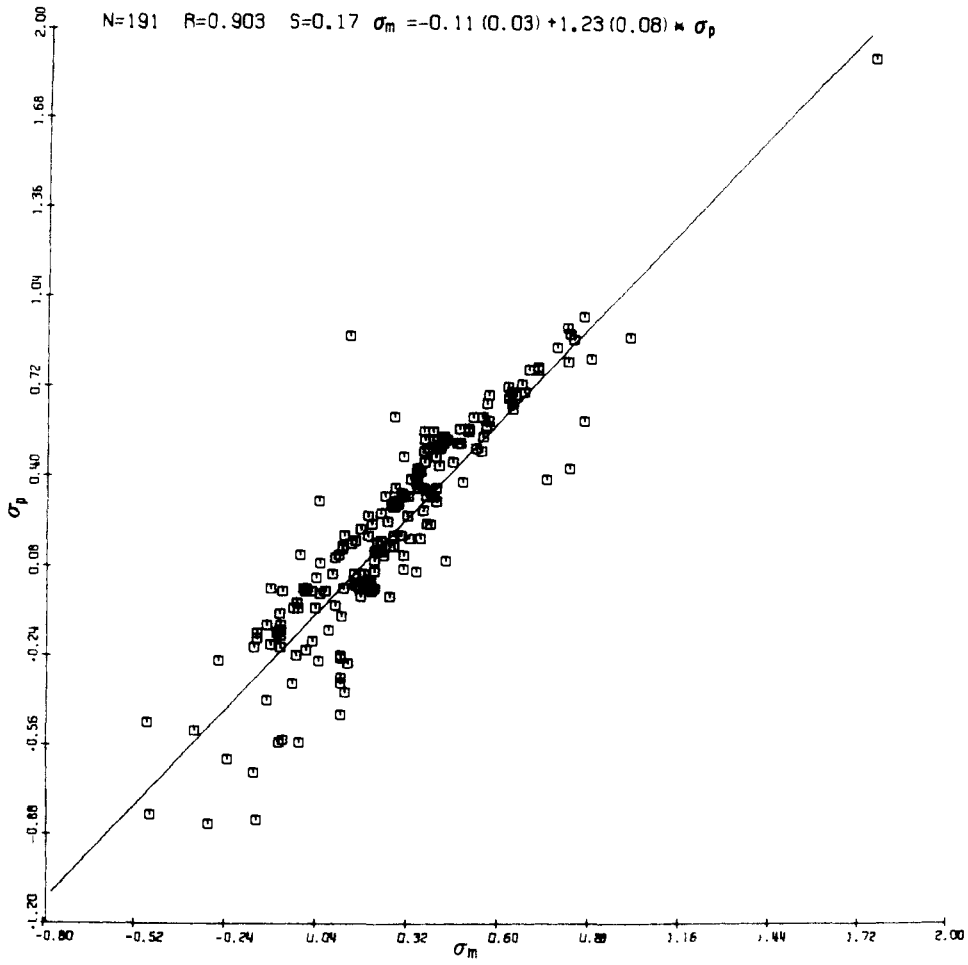


Figure 1. Plot of σ_m *vs.* σ_p for 191 substituents.

four determinations of P have been made at varying concentrations. Where P is a function of concentration, we have extrapolated to infinite dilution. We have used the value of 2.13 for log P-benzene. Boček and Tichý** have obtained a value of 2.15. In a few instances, as noted, π-X/benzene[8] was not available, and a constant from another solute system was given. Ionic substituents, such as $-CO_2^-$ and $-NMe_3^+$, are so hydrophilic that, when attached to benzene, the resulting log P is so low that it is quite difficult to measure. In such cases, the more lipophilic biphenyl solute system was used.

2. Electronic Parameters. σ_m and σ_p values were taken from our larger (unpublished) collection of substituent constants†† and represent both primary and secondary values of varying quality. We have attempted to select the "best" values available for each substituent, using updated values and discarding inconsistent values. However, we urge that the original sources be consulted because of the variety of methods represented.

We have completely repeated Swain and Lupton's procedure for obtaining $\mathcal{F}$ and $\mathcal{R}$ since the factor $\rho = 1.65$ was omitted from the calculation of σ' (ionization of 4-X-bicyclooctanecarboxylic acids in 50 wt % EtOH, Set 5)[6] with the effect that their $\mathcal{F}$ values are out of scale with $\mathcal{R}$. Furthermore, our selected σ_m and σ_p do not always agree with those of Swain and Lupton. We have consistently rounded to two decimal places, and our $\mathcal{F}$ and $\mathcal{R}$ are self-consistent with the σ_m and σ_p in Table I. Thus

$$\mathcal{F} = 1.369\ (\pm 0.186)\ \sigma_m - 0.373\ (\pm 0.142)\ \sigma_p - 0.009\ (\pm 0.038)$$

$$n \quad r \quad s$$
$$14 \quad 0.9915 \quad 0.0417 \quad (2)$$

The figures in parentheses are 95% confidence limits; the overall F statistic if $F_{2,11} = 318.77$ ($F_{2,11;\alpha=0.005} = 8.91$). All the coefficients are evaluated from 14 data points of Baker, *et al.*[13] (corrected for $\rho = 1.65$), but all $\mathcal{F}$ are calculated from eq 2. Then

$$\mathcal{R} = \sigma_p - \alpha\mathcal{F} \quad (3)$$

which gives $\alpha = 0.921$ under the assumption[6] that $R(NMe_3^+) = 0.0$ ($\sigma_p = 0.82$, $\mathcal{F} = 0.89$). Independent evaluation of new substituent constants has confirmed the general validity of this assumption.[8,‡] Results given below further show that $\mathcal{F}$ and $\mathcal{R}$ are remarkably orthogonal, considering the relative simplicity of the assumptions.

3. Steric Parameters. There are various systems for calculating molar refractivity, but the atom-group-structure constants of Vogel and the bond values of Vogel and others[5] are the ones most commonly used. The atom-group-structure system of Vogel could be applied to the greatest number of substituent structures of Table I, and so it was chosen for the sake of consistency. However, exaltations between aliphatic and aromatic values can be rather large (as much as 10%), and for substituents containing unsaturation or a lone electron pair which could interact with the benzene ring, and for which Vogel did not list separate aromatic values, we have used Ingold's special values. We have ignored the slight variation

**M. Tichý and K. Boček, private communication.

††Pomona College Medicinal Chemistry Project, Claremont, Calif. 91711.

Table I. "Aromatic" Substituent Constants

No.	Function[a]	π[b]	σ_m	σ_p	$\mathcal{F}$[c]	$\mathcal{R}$[c]	MR[d]	MW[e]	Wiswesser line notation[f]	Ref σ_m	Ref σ_p[g]
1	$B(OH)_2$	−0.55	−0.01	0.12	−0.07	0.18	11.04A[h]	44.8	*BQQ	1	1
2	Br	0.86	0.39	0.23	0.44	−0.17	8.88B	79.9	*E	2	2
3	CBr_3		0.28	0.29	0.27	0.04	28.81	251.7	*XEEE	3	3
4	CCl_3		0.32	0.33	0.31	0.05	20.12	118.4	*XGGG	3	3
5	CF_3	0.88	0.43	0.54	0.38	0.19	5.02	69.0	*XFFF	2	2
6	CN	−0.57	0.56	0.66	0.51	0.19	6.33B	26.0	*CN	2	2
7	COO^-	−4.36W	−0.10	0.00	−0.15	0.13	6.05BC	44.0	*VO	2	2
8	CHO	−0.65	0.35	0.42	0.31	0.13	6.88B	29.0	*VH	4	5
9	COOH	−0.32	0.37	0.45	0.33	0.15	6.93B	45.0	*VQ	2	2
10	CH_2Br	0.79	0.12	0.14	0.10	0.05	13.39	93.9	*1E	1	1
11	CH_2Cl	0.17	0.11	0.12	0.10	0.03	10.49	49.5	*1G	1	1
12	CH_2I		0.10	0.11	0.09	0.03	18.60	140.9	*1I	1	1
13	$CONH_2$	−1.49	0.28	0.36	0.24	0.14	9.81B	44.0	*VZ	4	6
14	CH=NOH	−0.38	0.22	0.10	0.25	−0.13	10.28B	44.0	*1UNQ	7	7
15	C=O(NHOH)	−1.87					11.22B	60.0	*VMQ		
16	CH_3	0.56	−0.07	−0.17	−0.04	−0.13	5.65	15.0	*1	2	2
17	CH_2OH	−1.03	0.00	0.00	0.00	0.00	7.19	31.0	*1Q	8	8
18	CH_2NH_2	−1.04					9.09	30.1	*1Z		
19	$C{=}O(CF_3)$	0.02					11.17B	97.0	*VXFFF		
20	3,4-(CF_2OCF_2)		0.81	0.81	0.80	0.08	10.19	116.0	*XFFOX*FF(C,D)	9	9
21	C≡CH	0.40	0.21	0.23	0.19	0.05	9.55B	25.0	*1UU1	10	10
22	CH_2SCF_3		0.12	0.15	0.10	0.06	17.59	115.1	*1SXFFF	11	11
23	$CH_2SO_2CF_3$		0.29	0.31	0.27	0.06	17.51[i]	147.1	*1SWXFFF	11	11
24	CH_2CN	−0.57	0.16	0.01	0.21	−0.18	10.11	40.0	*1CN	1	4
25	$CH{=}CHNO_2$ (trans)	0.11	0.32	0.26	0.33	−0.05	16.42B	72.0	*1U1NW -T	12	12
26	$CH{=}CH_2$	0.82	0.05	−0.02	0.07	−0.08	10.99B	27.1	*1U1	5	5
27	$COCH_3$	−0.55	0.38	0.50	0.32	0.20	11.18B	43.1	*V1	2	2
28	CO_2CH_3	−0.01	0.37	0.45	0.33	0.15	12.87B	59.0	*VO1	2	13
29	CH_2COOH	−0.72		−0.07			11.88	59.0	*1VQ		8
30	$C{=}O(NHCH_3)$	−1.27	0.35	0.36	0.34	0.05	14.57B	58.1	*VM1	14	14
31	CH_2CONH_2	−1.68		0.07			14.41	58.1	*1VZ		15
32	$C{=}S(NHCH_3)$		0.30	0.34	0.27	0.09	22.33	74.1	*YUS&M1	14	14
33	C_2H_5	1.02	−0.07	−0.15	−0.05	−0.10	10.30	29.1	*2	2	2
34	1-(1,2-$B_{10}H_{10}C_2H$) α-carboranyl		0.48	0.52	0.45	0.10		143.2	*?	16	16
35	3-Barenyl		0.20	0.19	0.19	0.01		143.2	*?	17	17
36	1-Neobarenyl		0.25	0.33	0.21	0.14		143.2	*?	17	17
37	$C{\equiv}CCF_3$		0.41	0.51	0.36	0.18	14.13B	93.0	*1UU1XFFF	18	18
38	$CF(CF_3)_2$		0.37	0.53	0.30	0.25	13.44	169.0	*XFXFFFXFFF	19	19
39	$C(OH)(CF_3)_2$		0.29	0.30	0.28	0.05	15.18	167.0	*XQXFFFXFFF	19	19
40	$CH{=}CHCF_3$ (trans)		0.24	0.27	0.22	0.07	15.57B	95.0	*1U1XFFF -T	3	3
41	$CH{=}CHCF_3$ (cis)		0.16	0.17	0.15	0.03	15.57B	95.0	*1U1XFFF -C	3	3
42	CH=CHCN	−0.17[j]	0.24	0.17	0.26	−0.07	16.23B	52.1	*1U1CN	20	20
43	$C{\equiv}CCH_3$			0.09			14.14B	39.1	*1UU2	21	21
44	CH=CHCHO		0.24	0.13	0.27	−0.12	16.88B	55.1	*1U1VH	20	20
45	CH=CHCOOH	0.00	0.14	0.90	−0.15	1.04	17.91B[k]	71.1	*1U1VQ	22	8
46	$CH_2CH{=}CH_2$	1.10					14.49	41.1	*2U1		
47	Cyclopropyl		−0.07	−0.21	−0.03	−0.19	13.53	41.1	* AL3TJ	23	23
48	CH_2COCH_3	−0.69					15.06	57.1	*1V1		
49	$CO_2C_2H_5$	0.51	0.37	0.45	0.33	0.15	17.47B	73.1	*VO2	2	2
50	$CH_2OC{=}O(CH_3)$	−0.17		0.05			16.48	73.1	*1OV1		15
51	$CH_2CH_2CO_2H$	−0.29	−0.03	−0.07	−0.02	−0.05	16.52	73.1	*2VQ	22	22
52	3,4-$(CH_2CH_2CH_2)$	1.20	−0.26	−0.26	−0.27	−0.01	13.94	42.1	*3*(C,D)	4	4
53	$CH_2CH(NH_3^+)$-COO^-	−3.56[l]						88.1	*1YZVQ		
54	C_3H_7	1.55	−0.07	−0.13	−0.06	−0.08	14.96	43.1	*3	5	4
55	$CH(CH_3)_2$	1.53	−0.07	−0.15	−0.05	−0.10	14.98	43.1	*Y	24	2
56	$CH_2N(CH_3)_2$	−0.15[m]		0.01			18.74	58.1	*1N1&1		25
57	$CF_2CF_2CF_2CF_3$		0.47	0.52	0.44	0.11	17.65	219.0	*/XFF/ 4F	19	19
58	2-Thienyl	1.61	0.09	0.05	0.10	0.04	24.04A[n]	83.1	* BT5SJ	26	26
59	3,4-(CH=CHCH=CH)	1.32	0.04	0.04	0.03	0.01	17.47A[o]	52.1	R A* B*(C,D)	2	2
60	$CH{=}CHCOCH_3$	−0.06[j]	0.21	−0.01	0.28	−0.27	21.10B	69.1	*1U1V1	20	20
61	Cyclobutyl			−0.15			17.88	55.1	* AL4TJ		27
62	3,4-$(CH_2)_4$	1.39X	−0.48	−0.48	−0.49	−0.03	18.59	56.1	*4*(C,D)	4	4
63	C_4H_9		−0.08	−0.16	−0.06	−0.11	19.59	57.1	*4	5	4
64	$C(CH_3)_3$	1.98	−0.10	−0.20	−0.07	−0.13	19.62	57.1	*X	2	2
65	$CH_2Si(CH_3)_3$		−0.16	−0.21	−0.15	−0.07	29.61D	87.2	*1-SI-1&1&1	2	2
66	4-Pyridyl	0.32					23.03A[e]	78.1	* DT6NJ		
67	$CH{=}CHCO_2C_2H_5$	0.86[j]	0.19	0.03	0.24	−0.19	27.21B	99.1	*1U1VO2	20	20
68	Cyclopentyl	2.14X		−0.02			22.02	69.1	* AL5TJ		27

Table I *(Continued)*

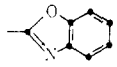
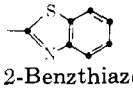
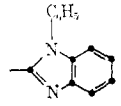
No.	Function[a]	π[b]	σ_m	σ_p	$\mathcal{F}$[c]	$\mathcal{R}$[c]	MR[d]	MW[e]	Wiswesser line notation[f]	Ref σ_m	Ref σ_p[g]
69	C_5H_{11}		−0.08	−0.15	−0.06	−0.09	24.25	71.2	*5	5	5
70	$(CH_2)_3N(CH_3)_2$	0.60		−0.13			28.04	86.2	*3N1&1		25
71	C_6Cl_5		0.25	0.24	0.24	0.02	49.53B	249.3	*R-/G 5	29	29
72	C_6F_5		0.34	0.41	0.30	0.13	23.98B	167.1	*R-/F 5	29	29
73	C_6H_2[2,4,6-$(NO_2)_3$]		0.26	0.30	0.24	0.08	42.21B	212.1	*R BNW DNW FNW	1	1
74	C_6H_5	1.96	0.06	−0.01	0.08	−0.08	25.36[r]	77.1	*R	2	2
75				0.15			24.80	81.2	* AL35TJ		27
76	Cyclohexyl	2.51X		−0.22			26.69	83.2	* AL6TJ		27
77	$(CH_2)_3N(CH_3)_3^+$	−4.15		0.02				101.2	*3K		25
78	 2-Benzoxazolyl		0.30	0.33	0.28	0.07	32.74	118.1	* CT56 BN DOJ	30	30
79	 2-Benzthiazolyl	2.13	0.27	0.29	0.25	0.06	38.88D[n]	134.2	* CT56 BN DSJ	30	30
80	$C{=}O(C_6H_5)$	1.05	0.34	0.43	0.30	0.16	30.33B	105.1	*VR	31	31
81	$CH{=}NC_6H_5$	−0.29	0.35	0.42	0.31	0.13	33.01D[n]	104.1	*1UNR	32	33
82	$CH_2C_6H_5$	2.01	−0.08	−0.09	−0.08	−0.01	30.01	91.1	*1R	1	1
83	$CH(OH)C_6H_5$	0.54		−0.03			31.52	107.1	*YQR		34
84				0.01			29.44	95.2	* AL36TJ		27
85	$C{\equiv}CC_6H_5$		0.14	0.16	0.12	0.05	33.21B	101.1	*1UU1R	1	1
86	$CH{=}CHC_6H_5$		0.03	−0.07	0.06	−0.12	34.17B	103.1	*1U1R	35	35
87	$CH_2CH_2C_6H_5$	2.66		−0.12			34.65	105.2	*2R		36
88	$CH{=}CHCOC_6H_4$-(4-NO_2)		0.15	0.05	0.18	−0.11	45.68B	176.2	*1U1VR DNW	20	20
89	$CH{=}CHCOC_6H_5$	0.95[j]	0.18	0.05	0.22	−0.15	40.25B	131.2	*1U1VR	20	20
90	Ferrocenyl	2.46	−0.15	−0.18	−0.15	−0.04	48.24A[q]	185.0	* AL50J 0-FE--0L50J	37	37
91	Adamantyl	3.30Y	−0.12	−0.13	−0.12	−0.02	40.63A[r]	135.3	* BL66 B6 A B-C 1B ITJ	38	39
92	 1-Phenyl-2-benzimidazolyl		0.17	0.21	0.15	0.08	59.08D[n]	193.2	* CT56 BN DNJ BR	30	30
93	$CO_2CH(C_6H_5)_2$		0.36	0.56	0.27	0.31	60.37B	211.3	*VOYR&R	13	34
94	Cl	0.71	0.37	0.23	0.41	−0.15	6.03B	35.4	*G	2	2
95	F	0.14	0.34	0.06	0.43	−0.34	0.92B	19.0	*F	2	2
96	$GeBr_3$		0.66	0.73	0.62	0.16	36.35D	312.3	*-GE-EEE	40	40
97	$GeCl_3$		0.71	0.79	0.67	0.17	25.85D	178.9	*-GE-GGG	40	40
98	GeF_3		0.85	0.97	0.79	0.24	6.95D	129.6	*-GE-FFF	40	40
99	H	0.00	0.00	0.00	0.00	0.00	1.03	1.0	*H	22	22
100	$HgCH_3$		0.43	0.10	0.54	−0.40	19.43D	215.6	*-HG-1	41	41
101	I	1.12	0.35	0.18	0.40	−0.19	13.94B	126.9	*I	2	2
102	IO	−3.74					39.06C[s]	142.9	*IO		
103	IO_2	−3.46	0.68	0.78	0.63	0.20	63.51CD[s]	158.9	*IW	1	1
104	NO	−0.12		0.12			5.2	30.0	*NO		4
105	NO_2	−0.28	0.71	0.78	0.67	0.16	7.36[t]	46.0	*NW	2	2
106	$N{\equiv}N^+$		1.76	1.91	1.69	0.36		28.0	*NN &J	42	42
107	NNN	0.46	0.27	0.15	0.30	−0.13	10.2B[u]	42.0	*NNN	43	43
108	NH_2	−1.23	−0.16	−0.66	0.02	−0.68	5.42B	16.0	*Z	2	2
109	NHOH	−1.34	−0.04	−0.34	0.06	−0.40	7.22	32.0	*MQ	4	4
110	NH_3^-		0.86	0.60	0.94	−0.27		17.0	*Z &H	44	44
111	$NHNH_2$	−0.88	−0.02	−0.55	0.17	−0.71	8.44	31.0	*MZ	4	4
112	$NHSO_2NHSO_2$-NH_2	−2.11[v]					28.40[t]	174.2	*MSWMSZW		
113	5-Cl-1-tetrazolyl	−0.65	0.60	0.61	0.58	0.07	23.16D[w]	103.5	* AT5NNNNJ EG	3	3
114	$N{=}CCl_2$	0.41	0.21	0.13	0.23	−0.08	18.35D	96.9	*NUYGG	3	3
115	$N{=}C{=}O$		0.27	0.19	0.29	−0.08	8.82D	42.0	*NCO	3	3
116	$N{=}C{=}S$	1.15	0.48	0.38	0.51	−0.09	17.24D	58.1	*NCS	43	43
117	5-Azido-1-tetrazolyl		0.54	0.54	0.53	0.05	26.85CD[x]	110.1	* AT5NNNNJ ENNN	3	3
118	NHCN		0.21	0.06	0.26	−0.18	10.14	41.0	*MCN	28	28
119	1-Tetrazolyl	−1.04	0.52	0.50	0.52	0.02	18.33D[w]	69.1	* AT5NNNNJ	3	3
120	5-OH-1-tetrazolyl		0.39	0.33	0.40	−0.04	19.77D[w]	85.1	* AT5NNNNJ EQ	3	3
121	5-SH-1-tetrazolyl		0.45	0.45	0.44	0.05	26.06D[w]	101.1	* AT5NNNNJ ESH	3	3

Table I *(Continued)*

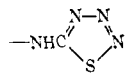
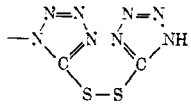
No.	Function[a]	π[b]	σ_m	σ_p	$\mathcal{F}$[c]	$\mathcal{R}$[c]	MR[d]	MW[e]	Wiswesser line notation[f]	Ref σ_m	Ref σ_p[g]
122			0.30	0.19	0.33	−0.11		101.1	*M- ET5NNNSJ	28	28
123	NHCHO	−0.98	0.19	0.00	0.25	−0.23	10.31	44.0	*MVH	43	43
124	$NHCONH_2$	−1.30	−0.03	−0.24	0.04	−0.28	13.72	59.1	*MVZ	43	43
125	$NHCSNH_2$	−1.40	0.22	0.16	0.23	−0.05	22.19	75.1	*MYZUS	28	28
126	$NHCH_3$	−0.47	−0.30	−0.84	−0.11	−0.74	10.33	30.1	*M1	4	2
127	$NHSO_2CH_3$	−1.18	0.20	0.03	0.25	−0.20	18.17[i]	94.1	*MSW1	43	43
128	$N(CF_3)_2$		0.40	0.53	0.34	0.22	14.28	152.0	*NXFFFXFFF	45	45
129	$NHCOCF_3$	0.08	0.30	0.12	0.36	−0.21	14.30	112.0	*MVXFFF	43	43
130			0.63	0.64	0.61	0.07	49.17D[w]	201.2	* AT5NNNNJ ESS- ET5MNNNJ	3	3
131	$NHCOCH_2Cl$		0.17	−0.03	0.23	−0.25	19.77	92.5	*MV1G	43	43
132	$NHCOCH_3$	−0.97	0.21	0.00	0.28	−0.26	14.93	58.1	*MV1	2	2
133	$NHCSCH_3$		0.24	0.12	0.27	−0.13	23.40	74.1	*MYUS	14	14
134	NHC_2H_5	0.08[y,z]	−0.24	−0.61	−0.11	−0.51	14.98	44.1	*M2	4	6
135	$N(CH_3)_2$	0.18	−0.15	−0.83	0.10	−0.92	15.55	44.1	*N1&1	46	2
136	$N(SO_2CH_3)_2$	−1.51					31.22[i]	172.2	*NSW1&&SW1		
137	$N{=}NN(CH_3)_2$	0.46					20.88D	72.1	*NUNN1&1		
138	$NHCOC_2H_5$	−0.47					19.58	72.1	*MV2		
139	$NHCO_2C_2H_5$	0.17	0.07	−0.15	0.14	−0.28	21.18	88.1	*MVO2	43	43
140	$NHCONHC_2H_5$		0.04	−0.26	0.14	−0.39	23.19	87.1	*MVM2	14	14
141	$NHCSNHC_2H_5$		0.30	0.07	0.38	−0.28	31.66	103.2	*MYUS&M2	14	14
142	$N(CH_3)_3^+$	−5.96W	0.88	0.82	0.89	0.00		59.1	*K	2	2
143	$NHCOCH(CH_3)_2$		0.11	−0.10	0.18	−0.26	24.25	86.1	*MVY	43	43
144	$NHCH_2CO_2C_2H_5$		−0.10				25.82	102.1	*M1VO2	47	
145	NHC_4H_9	1.45[y]	−0.34	−0.51	−0.28	−0.25	24.26	72.15	*M4	4	6
146	$N{=}NC_6H_5$	1.69	0.32	0.39	0.28	0.13	31.31	105.1	*NUNR	13	13
147	NHC_6H_5	1.37	−0.12	−0.40	−0.02	−0.38	30.04	92.1	*MR	48	48
148	$NHSO_2C_6H_5$	0.45	0.16	0.01	0.21	−0.18	37.88[i]	156.2	*MSWR	43	43
149	$N{=}CHC_6H_5$	−0.29	−0.08	−0.55	0.09	−0.63	33.01D	104.1	*NU1R	43	43
150	$NHCOC_6H_5$	0.49	0.02	−0.19	0.09	−0.27	34.64	120.1	*MVR	43	43
151	$N{=}NC_6H_3$-(2-OH)(5-CH_3)		0.27	0.31	0.24	0.08	37.45	135.2	*NUNR BQ E1	43	43
152	$N{=}CHC_6H_4$-(4-OCH_3)		−0.07	−0.54	0.10	−0.63	39.29D	134.2	*NU1R DO1	43	43
153	$NHCOC_6H_4$-(4-OCH_3)		0.09	−0.06	0.14	−0.19	41.03	150.2	*MVR DO1	43	43
154	$N(C_6H_5)_2$		0.00	−0.22	0.07	−0.29	54.96	168.2	*NR&R	49	50
155	O^-	−3.87W	−0.47	−0.81	−0.35	−0.49		16.0	*O	44	44
156	OH	−0.67	0.12	−0.37	0.29	−0.64	2.85B	17.0	*Q	2	2
157	3,4-(OCF_2O)		0.36	0.36	0.35	0.04	8.95B	82.0	*OXFFO*(C,D)	9	9
158	OCF_3	1.04	0.38	0.35	0.38	0.00	7.86B	85.0	*OXFFF	45	45
159	$OCHF_2$		0.31	0.18	0.35	−0.14	7.86B	57.0	*OYFF	51	51
160	$OCONH_2$	−1.05					11.28B	60.0	*OVZ		
161	3,4-(OCH_2O)	−0.05	−0.16	−0.16	−0.17	0.00	8.96B	46.0	*O1O*(C,D)	22	22
162	OCH_3	−0.02	0.12	−0.27	0.26	−0.51	7.87B	31.0	*O1	2	2
163	OSO_2CH_3	−0.88	0.39	0.36	0.39	0.00	16.99	95.1	*OSW1	43	43
164	OCF_2CHFCl		0.35	0.28	0.37	−0.06	17.30B	133.5	*OXFFYGF	51	51
165	$OCOCH_3$	−0.64	0.39	0.31	0.41	−0.07	12.47B	59.0	*OV1	2	2
166	OCH_2COOH	−0.87		−0.33			13.99B	75.0	*O1VQ		52
167	OEt	0.38	0.10	−0.24	0.22	−0.44	12.47B	45.1	*O2	2	2
168	$OPO(OCH_3)_2$			0.04			22.02B	125.0	*OPO&O1&O1		53
169	$OCH(CH_3)_2$		0.10	−0.45	0.30	−0.72	17.06B	59.1	*OY	2	2
170	OC_3H_7	1.05	0.10	−0.25	0.22	−0.45	17.06B	59.1	*O3	2	2
171	OC_4H_9		0.10	−0.32	0.25	−0.55	21.66B	73.1	*O4	2	2
172	OC_5H_{11}		0.10	−0.34	0.25	−0.57	26.26B	87.2	*O5	2	2
173	OC_6H_5	2.08	0.25	−0.03	0.34	−0.35	27.68B	93.1	*OR	2	4
174	$OSO_2C_6H_5$	0.93	0.36	0.33	0.36	0.00	36.70[i]	157.2	*OSWR	43	43
175	O-β-glucose	−2.84[aa]					36.53D	179.0	*O- BT6OTJ CQ DQ EQ F1Q		
176	$OCOC_6H_5$	1.46	0.21	0.13	0.23	−0.08	32.33B	121.1	*OVR	43	43
177	$POCl_2$		0.80	0.43	0.93	−0.42	20.16D	117.9	*PO&GG	54	53
178	PCl_2		0.53	0.61	0.49	0.16	21.42D	101.9	*PGG	54	41
179	POF_2		0.81	0.89	0.77	0.18	9.58D	85.0	*PO&FF	41	41
180	PF_2		0.26	0.61	0.12	0.50	11.02D	69.0	*PFF	41	41
181	$PSCl_2$		0.73	0.39	0.84	−0.39	28.29D	133.9	*PS&GG	54	53
182	PO_3H^-		0.20	0.26	0.17	0.11		80.0	*PWQ	2	2
183	PH_2		0.05				12.19D	33.0	*PHH	54	
184	$P(Cl)N(CH_3)_2$		0.38	0.56	0.30	0.28	27.01D	110.5	*PGN1&1	41	41
185	$PO(CH_3)_2$		0.42				19.93D	77.0	*PO&1&1	54	

Table I (Continued)

No.	Function[a]	π[b]	σ_m	σ_p	$\mathcal{F}$[c]	$\mathcal{R}$[c]	MR[d]	MW[e]	Wiswesser line notation[f]	Ref σ_m	Ref σ_p[g]
186	$PO(OCH_3)_2$		0.42	0.53	0.37	0.19	21.87D	109.0	*PO&O1&O1	55	55
187	$P(CH_3)_2$	0.44	0.03	0.31	−0.08	0.39	21.19D	61.1	*P1&1	54	41
188	$P(OC_2H_5)_2$			0.33			32.42D	121.1	*PO2&O2		53
189	$PO(OC_2H_5)_2$		0.55	0.60	0.52	0.12	31.16D	137.1	*PO&O2&O2	56	56
190	$PO(Cl)C_6H_4$-3-F		0.65				39.49D[h,i]	177.5	*PO&GR CF	54	
191	$P(Cl)C_6H_4$-3-F		0.42				40.75D[h,i]	161.5	*PGR CF	54	
192	$PS(Cl)C_6H_4$-3-F		0.56				47.62D[h,i]	193.6	*PS&GR CF	54	
193	$P(Cl)C_6H_5$			0.44			40.99D	143.5	*PGR		53
194	$P(H)C_6H_4$-3-F		0.09				36.14D[h,i]	127.1	*PHR CF	54	
195	$PO(OC_6H_5)_2$		0.38	0.50	0.32	0.20	40.46D	165.2	*PO&O3&O3	55	55
196	$P(OCH_3)C_6H_4$-3-F		0.33				41.68D[h,i]	157.1	*PO1&R CF	54	
197	$PO(CH_3)C_6H_4$-3-F		0.40				39.37D[h,i]	157.1	*PO&1&R CF	54	
198	$P(CH_3)C_6H_4$-3-F		0.20				40.63D[h,i]	141.1	*P1&R CF	54	
199	$PO(C_6H_5)_3$		0.35	0.49	0.29	0.23	47.81D	161.2	*PO&4&4	55	55
200	$PO(C_6H_5)_2$	0.70	0.38	0.53	0.31	0.24	59.29D	201.2	*PO&R&R	57	57
201	$P(C_6H_5)_2$		0.11	0.19	0.07	0.12	60.55D	185.2	*PR&R	57	57
202	$PS(C_6H_5)_2$		0.29	0.47	0.21	0.27	67.42D	217.2	*PS&R&R	57	57
203	$SO_2(F)$	0.05Z	0.80	0.91	0.75	0.22	8.65	83.1	*SWF	58	58
204	SF_5	1.23	0.61	0.68	0.57	0.15	9.89A[i,j]	127.1	*SFFFFF	59	59
205	SO_2		−0.02	−0.05	−0.02	−0.03		64.1	*SW	60	60
206	SO_3	−4.76W	0.05	0.09	0.03	0.07		80.1	*SWO	2	2
207	SH	0.39	0.25	0.15	0.28	−0.11	9.22B	33.1	*SH	2	2
208	$SO_2(NH_2)$	−1.82	0.46	0.57	0.41	0.19	12.28[i]	80.1	*SZW	2	2
209	$SCCl_3$	1.65					28.34B	150.4	*SXGGG		
210	$S(=O)(CF_3)$		0.63	0.69	0.60	0.14	13.07[i]	117.1	*SO&XFFF	18	18
211	$SO_2(CF_3)$	0.55	0.79	0.93	0.73	0.26	12.86[i]	133.1	*SWXFFF	61	61
212	SCF_3	1.44	0.40	0.50	0.35	0.18	13.81B	101.1	*SXFFF	61	61
213	SCN	0.41	0.41	0.52	0.36	0.19	13.40	58.1	*SCN	6	2
214	$SCHF_2$		0.33	0.37	0.30	0.09	13.81B	83.1	*SYFF	1	1
215	$SOCHF_2$		0.54	0.58	0.51	0.11	13.28[i]	99.1	*SO&YFF	62	62
216	SO_2CHF_2		0.75	0.86	0.70	0.22	13.08[i]	115.1	*SWYFF	1	1
217	$SOCH_3$	−1.58	0.52	0.49	0.52	0.01	13.70[i]	63.1	*SO&1	2	2
218	SO_2CH_3	−1.63	0.60	0.72	0.54	0.22	13.49[i]	79.1	*SW1	2	2
219	SCH_3	0.61	0.15	0.00	0.20	−0.18	13.82B	47.1	*S1	2	2
220	SCF_2CHF_2		0.38	0.47	0.34	0.16	18.40B	133.1	*SXFFYFF	61	61
221	$SCOCH_3$	0.10	0.39	0.44	0.36	0.11	18.42B	75.1	*SV1	2	2
222	SC_2H_5	1.07	0.18	0.03	0.23	−0.18	18.42B	61.1	*S2	5	2
223	$S(CH_3)_2$[k]		1.00	0.90	1.02	−0.04		62.1	*S1&1	2	2
224	$SO_2(C_6H_5)$	0.27	0.61	0.70	0.56	0.18	33.20[i]	141.2	*SWR	5	5
225	SC_6H_5	2.32		0.18			34.29B	109.2	*SR		36
226	$S(CH_3)(=NSO_2(C_6H_4$-4-$CH_3))$		0.65	0.70	0.62	0.13		216.3	*S1&UNSWR D1	63	63
227	$SeCF_3$		0.32	0.38	0.29	0.12	16.32D	148.0	*-SE-XFFF	64	64
228	SeCN		0.61	0.66	0.58	0.13	16.82D	105.0	*-SE-CN	48	4
229	$SeCH_3$	0.74	0.10	0.00	0.13	−0.12	17.03D	94.0	*-SE-1	2	2
230	$SiBr_3$		0.48	0.57	0.44	0.17	32.76D	267.8	*-SI-EEE	40	40
231	$SiCl_3$		0.48	0.56	0.44	0.16	23.85D	134.4	*-SI-GGG	40	40
232	SiF_3		0.54	0.69	0.47	0.25	7.62D	85.1	*-SI-FFF	41	41
233	$Si(CH_3)_3$	2.59[l,m]	−0.04	−0.07	−0.04	−0.04	24.96D	73.2	*-SI-1&1&1	2	2
234	$Si(CH_3)_2[OSi(CH_3)_3]$		0.00	−0.01	−0.01	−0.01	43.64D	147.4	*-SI-1&1&O-SI-1&1&1	65	65
235	$Si(CH_3)[OSi(CH_3)_3]_2$		−0.02	−0.01	−0.03	0.02	62.32D	221.6	*-SI-1&/O-SI-1&1&1 2	65	65
236	$Si[OSi(CH_3)_3]_3$		−0.09	−0.01	−0.13	0.11	80.99D	295.7	*-SI-/O-SI-1&1&1 3	65	65

[a] Function begins with attachment atom, sorted alphabetically on attachment atom and within each such grouping: first, if no C or H, then alphabetically on remainder; second, if no C, then on H and alphabetically on remainder; third, C then H then alphabetically on remainder. [b] All π values from partition coefficients measured in this laboratory using octanol–water solvent system and substituted benzene solutes unless footnoted to give other sources or suffixed to give other solute systems: W = from substituted biphenyl solutes; X = from substituted phenoxyacetic acid solutes; Y = calculated from OH derivative; Z = from substituted toluene solutes. [c] Calculated from σ_m and σ_p given in this table according to the procedure outlined in the text. [d] Molar refraction using A. I. Vogel's [*J. Chem. Soc.*, 1833 (1948)] atom, group, or structural R_D (yellow line) values unless suffixed: A = calculated [usually from index of refraction, density, and molecular weight from Lorentz Lorentz formula (eq 1)] using Vogel's (1948) values for corrections; B = atom, group, or structural H_α (= R_C red line) values from Ingold ("Structure and Mechanism in Organic Chemistry," 2nd ed, Cornell University Press, Ithaca, N. Y., 1969, pp 142–152). Note: Table 10.1 "alcohol" and "ether" values inverted; C = approximate; D = bond values (including bond to C of substrate) from A. I. Vogel, W. T. Cresswell, G. H. Jeffery, and J. Leicester, *J. Chem. Soc.*, 514 (1952), and earlier references cited therein (general); A. I. Vogel, W. T. Cresswell, and J. Leicester, *J. Phys. Chem.*, **58**, 174 (1954) (Sn, Si, Ge, and Hg bonds); A. A. Foxton, G. H. Jeffery, and A. I. Vogel, *J. Chem. Soc. A*, 249 (1966) (P bonds); R. G. Gillis, *Rev. Pure Appl. Chem.*, **10**, 21 (1960) (bonds to C, H, O, and self), updated by P. M. Christopher and T. L. Patterson, *Aust. J. Chem.*, **21**, 2373 (1968), and earlier references cited therein; C. Stölzer and A. Simon, *Chem. Ber.*, **96**, 1335 (1963) (P bonds to F, Cl, N); R. Sayre, *J. Amer. Chem. Soc.*, **80**, 5438 (1958) (P bonds to S). [e] From "Handbook of Chemistry and Physics," 53rd ed, Chemical Rubber Publishing Co., Cleveland, Ohio, 1972. [f] The WLN follow as closely as possible the rules in "The

Wiswesser Line-Formula Chemical Notation," E. C. Smith, Ed., McGraw-Hill, New York, N. Y., 1968, with these additions. (1) The WLN begins at the point of attachment: (a) if the substituent group becomes part of an aromatic fused ring system, the substituent is cited as a closed ring and the attachment locants (for the substituent ring) are marked with asterisks. The notation is followed by a parentheses showing attachment locants on the parent ring; (b) if the substituent completes a single saturated ring on an aromatic ring it is treated as a linear chain with a two-point attachment; (c) the WLN for a carbocyclic or heterocyclic ring as a substituent begins with a space and then a locant showing the attachment point on the substituent ring. (2) Methyl contractions are made on "X," "Y," and "K" symbols but not on rings. (3) Multipliers are used according to normal rules. (4) The # symbol denotes a saturated alkyl chain of undetermined length. (5) If a "?" begins the notation, the structure is not definable by WLN. [g] The following refrences refer to σ_m and σ_p, respectively: (1) O. Exner, *Collect. Czech. Chem. Commun.*, **31,** 65 (1966); (2) D. H. McDaniel and H. C. Brown, *J. Org. Chem.*, **23,** 420 (1958); (3) W. A. Sheppard. *Trans. N. Y. Acad. Sci.*, [II] **29,** 700 (1967); (4) H. H. Jaffé, *Chem. Rev.*, **53,** 191 (1953); (5) M. Charton, *J. Org. Chem.*, **30,** 552 (1965); (6) M. Charton, *ibid.*, **28,** 3121 (1963); (7) P. Cecchi, *Ric. Sci.*, **28,** 2526 (1958); (8) P. Zuman, "Substituent Effects in Organic Polarography," Plenum Press, New York, N. Y., 1967, p 76; (9) L. M. Yagupol'skii and L. N. Yagupol'skaya, *Dokl. Chem.*, **134,** 1207 (1960); (10) J. A. Landgrebe and R. H. Rynbrandt, *J. Org. Chem.*, **31,** 2585 (1966); (11) V. V. Orda, L. M. Yagupol'skii, V. F. Bystrov, and A. U. Stepanyants, *J. Gen. Chem. USSR*, **35,** 1631 (1965); (12) R. Stewart and L. G. Walker, *Can. J. Chem.*, **35,** 1561 (1957); (13) W. F. Little, C. N. Reilley, J. D. Johnson, K. N. Lynn, and A. P. Sanders, *J. Amer. Chem. Soc.*, **86,** 1376 (1964); (14) T. Nishiguchi and Y. Iwakura, *J. Org. Chem.*, **35,** 1591 (1970); (15) O. Exner and J. Jonáš, *Collect. Czech. Chem. Commun.*, **27,** 2296 (1962); (16) M. F. Hawthorne, T. E. Berry, and P. A. Wegner, *J. Amer. Chem. Soc.*, **87,** 4746 (1965); (17) L. I. Zakharkin, V. N. Kalinin, and I. P. Shepilov, *Dokl. Chem.*, **174,** 484 (1967); (18) V. F. Bystrov, L. M. Yagupol'skii, A. U. Stepanyants, and Yu. A. Fialkov, *ibid.*, **153,** 1019 (1963); (19) W. A. Sheppard, *J. Amer. Chem. Soc.*, **87,** 2410 (1965); (20) G. B. Ellam and C. D. Johnson, *J. Org. Chem.*, **36,** 2284 (1971); (21) M. Charton and H. Meislich, *J. Amer. Chem. Soc.*, **80,** 5940 (1958); (22) L. P. Hammett, "Physical Organic Chemistry," McGraw-Hill, New York, N. Y., 1940, p 188; (23) J. Šmejkal, J. Jonáš, and J. Farkaš, *Collect. Czech. Chem. Commun.*, **29,** 2950 (1964); (24) M. Charton, *J. Chem. Soc.*, 1205 (1964); (25) J. H. Smith and F. M. Menger, *J. Org. Chem.*, **34,** 77 (1969); (26) F. Fringuelli, G. Marino, and A. Taticchi, *J. Chem. Soc. B*, 1595 (1970); (27) Yu. S. Shabarov, T. P. Surikova, and R. Ya. Levina, *J. Org. Chem. USSR*, **4,** 1131 (1968); (28) J. C. Kauer and W. A. Sheppard, *J. Org. Chem.*, **32,** 3580 (1967); (29) W. A. Sheppard, *J. Amer. Chem. Soc.*, **92,** 5419 (1970); (30) V. F. Bystrov, Zh. N. Belaya, B. E. Gruz, G. P. Syrova, A. I. Tolmachev, L. M. Shulezhko, and L. M. Yagupol'skii, *J. Gen. Chem. USSR*, **38,** 963 (1968); (31) W. N. White, R. Schlitt, and D. Gwynn, *J. Org. Chem.*, **26,** 3613 (1961); (32) C. Weiss, W. Engewald, and H. Müller, *Tetrahedron*, **22,** 825 (1966); (33) J. J. Ryan and A. A. Humffray, *J. Chem. Soc. B*, 842 (1966); (34) W. F. Little, C. N. Reilley, J. D. Johnson, and A. P. Sanders, *J. Amer. Chem. Soc.*, **86,** 1382 (1964); (35) J. K. Kochi and G. S. Hammond, *ibid.*, **75,** 3452 (1953); (36) L. M. Litvinenko, *Bull. Acad. Sci. USSR, Div. Chem. Sci.*, **10,** 1653 (1962); (37) A. N. Nesmeyanov, E. G. Perevalova, S. P. Gubin, K. I. Grandberg, and A. G. Kozlovsky, *Tetrahedron Lett.*, **22,** 2381 (1966); (38) H. Alper, E. C. H. Keung, and R. A. Partis, *J. Org. Chem.*, **36,** 1352 (1971); (39) P. Zuman, "Substituent Effects in Organic Polarography," Plenum Press, New York, N. Y., 1967, p 325; (40) N. V. Kondratenko, G. P. Syrova, V. I. Popov, Yu. N. Sheinker, and L. M. Yagupol'skii, *J. Gen. Chem. USSR*, **41,** 2075 (1971); (41) M. G. Hogben and W. A. G. Graham, *J. Amer. Chem. Soc.*, **91,** 283 (1969); (42) E. S. Lewis and M. D. Johnson, *ibid.*, **81,** 2070 (1959); (43) O. Exner and J. Lakomý, *Collect. Czech. Chem. Commun.*, **35,** 1371 (1970); (44) J. Hine, *J. Amer. Chem. Soc.*, **82,** 4877 (1960); (45) F. S. Fawcett and W. A. Sheppard, *ibid.*, **87,** 4341 (1965); (46) J. C. Howard and J. P. Lewis, *J. Org. Chem.*, **31,** 2005 (1966); (47) E. P. Serjeant, *Aust. J. Chem.*, **22,** 1189 (1969); (48) M. Charton, *J. Org. Chem.*, **30,** 557 (1965); (49) E. N. Tsvetkov, D. I. Lobanov, and M. I. Kabachnik, *Chem. Abstr.*, **66,** 28299r (1967); (50) E. N. Tsvetkov, D. I. Lobanov, and M. I. Kabachnik, *ibid.*, **64,** 12523f (1966); (51) R. Van Poucke, R. Pollet, and A. De Cat, *Bull. Soc. Chim. Belg.*, **75,** 573 (1966); (52) P. J. Bray and R. G. Barnes, *J. Chem. Phys.*, **27,** 551 (1957); (53) H. L. Retcofsky and C. E. Griffin, *Tetrahedron Lett.*, **18,** 1975 (1966); (54) H. Schindlbauer and W. Prikoszovich, *Chem. Ber.*, **102,** 2914 (1969); (55) E. N. Tsvetkov, D. I. Lobanov, L. A. Izosenkova, and M. I. Kabachnik, *J. Gen. Chem. USSR*, **39,** 2126 (1969); (56) L. D. Freedman and H. H. Jaffé, *J. Amer. Chem. Soc.*, **77,** 920 (1955); (57) E. N. Tsvetkov, D. I. Lobanov, M. M. Makhamatkhanov, and M. I. Kabachnik, *Tetrahedron*, **25,** 5623 (1969); (58) K. Kalfus, M. Večeřa, and O. Exner, *Collect. Czech. Chem. Commun.*, **35,** 1195 (1970); (59) W. A. Sheppard, *J. Amer. Chem. Soc.*, **84,** 3072 (1962); (60) B. J. Lindberg, *Acta Chem. Scand.*, **24,** 2852 (1970); (61) W. A. Sheppard, *J. Amer. Chem. Soc.*, **85,** 1314 (1963); (62) L. N. Sedova, L. Z. Gandel'sman, L. A. Alekseeva, and L. M. Yagupol'skii, *J. Gen. Chem. USSR*, **39,** 2011 (1969); (63) Á. Kucsman, F. Ruff, S. Sólyom, and T. Szirtes, *Acta Chim. Acad. Sci. Hung.*, **57,** 205 (1968); (64) V. G. Boloshckuk, L. M. Yagupol'skii, G. P. Syrova, and V. F. Bystrov, *J. Gen. Chem. USSR*, **3,** 105 (1967); (65) Z. Plzák, F. Mareš, J. Hetflejš, J. Schraml, Z. Papoušková, V. Bažant, E. G. Rochow, and V. Chvalovský, *Collect. Czech. Chem. Commun.*, **36,** 3115 (1971). [h] From M. M. Otto, *J. Amer. Chem. Soc.*, **57,** 1476 (1935). [i] SO_2 from $SO_2(OR)_2$ = 7.838; SO from $SO(OR)_2$ = 8.046; Vogel's values (1948). [d] [j] From C. E. Lough, Suffield, Memorandum No. 17/71. [k] Aromatic oxygen from –OH = 1.76. [l] pH = 6.0; phosphate buffer. [m] T. Fujita, private communication. [n] From G. H. Jeffery, R. Parker, and A. I. Vogel, *J. Chem. Soc.*, 570 (1961), average for thenoates and thienyl ketones. [o] From H. H. Huang and S. C. Ng, *J. Chem. Soc. B*, 582 (1968), report 41.8 for naphthalene. [p] Vogel's group value (1948). [d] Bond value is 26.387 – 1.676 = 24.711 where 26.387 is from Vogel's (1948) 25.359 phenyl value plus his 1.028 hydrogen value. Ingold's value 25.93 – 1.092 = 24.838. [q] From T. P. Vishnyakova, Y. M. Paushkin, and T. A. Sokolinskata, *J. Gen. Chem. USSR*, **33,** 3615 (1963), from monosubstituted isooctyl and cyclohexyl derivatives. Disubstituted showed –0.90 "exaltation." [r] From S. Landa and V. Machacek, *Collect. Czech. Chem. Commun.*, **5,** 1 (1933). [s] From Vogel, *et al.* (1952), [d] which gives a factor of 2.3 for conversion of single to double bonds. Christopher and Patterson (1968) [d] estimate (I–O) bond as 10.63, thus we obtain 24.449 for (I=O). Gillis (1960) [d] gives (C–I) as 14.61. [t] Vogel, *et al.* (1952), [d] give $PhNO_2$ = 32.72 in Table 53. Thus 32.72 – 25.36 = 7.36. This procedure was not used for other values because of better agreement with standard method. [u] From E. Lieber, C. N. R. Rao, T. S. Chao, and W. H. Wohl, *J. Sci. Ind. Res., Sect. B*, **16,** 95 (1957), for Ph derivative. [v] $PhNHSO_2NHSO_2NH_2$ was the gift of Dr. R. Appel, University of Bonn, Germany. [w] Using C_{ar}–NR_2 = 3.22 or C_{ar}–NHR = 2.96 bond values from Vogel, *et al.* (1952). [d] Other values of C_{ar}–X were not used since these agreed closely with normal values. [z] Approximate: group value for N_3 plus bond value for tetrazolyl. [y] From Y. Ichikawa, T. Yamano, and H. Fujishima, *Biochim. Biophys. Acta*, **171,** 32 (1969). [z] From O. E. Schultz, C. Jung, and K. E. Moller, *Z. Naturforsch., B*, **25,** 1024 (1970). [aa] From R. Poretz and I. Goldstein, *Arch. Biochem. Biophys.*, **125,** 1034 (1968). [bb] Bond: 24.711 – H + F = 24.475 for C_6H_4-3-F. [cc] From W. A. Sheppard, *J. Amer. Chem. Soc.*, **84,** 3064, 3072 (1962). Calculated from data on $PhSF_5$. [dd] From V. Lee, M.S. Thesis, San Jose State College, Aug 1967.

(<1%) between the R_C (red line) values reported by Ingold and the R_D (yellow line) values of Vogel.

Very few atom or group refractivities have been reported for organometallic substituents, and so the more common *bond* values were used. Our MR values always include the bond to the carbon atom of the parent benzene ring. We have not calculated the refractivity for charged substituents as no suitable values are available (except the octet values given by Ingold, which have not been widely used).

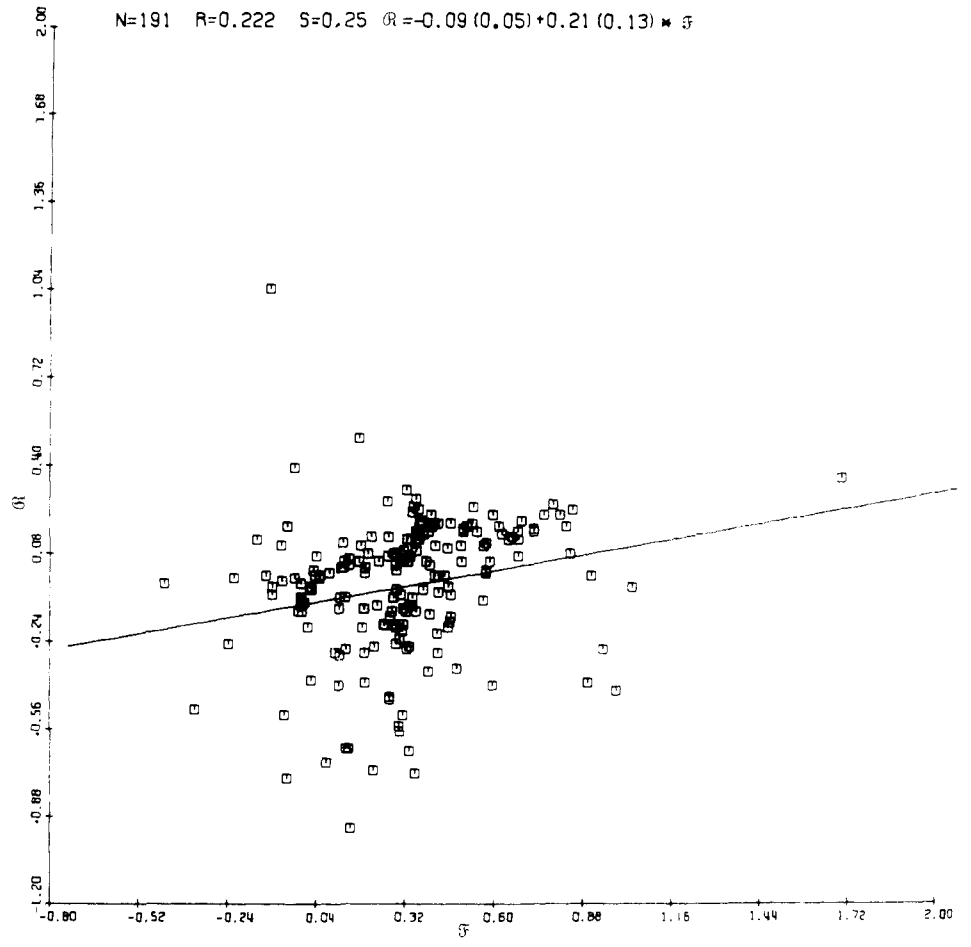


Figure 2. Plot of ℑ *vs.* ℜ for 191 substituents.

The following examples illustrate these methods.

(1) $-NHCSNH_2$ (by Vogel's group values)

$$\underset{4.678}{\underline{\text{NH}\\ \text{secondary aromatic amine}}} + \underset{13.07}{\underline{\text{CS}\\ \text{xanthates}}} + \underset{4.438}{\underline{\text{NH}_2\\ \text{primary aliphatic amine}}} = 22.19$$

(2) $-OCOC_6H_5$ (by Ingold's group values with exaltation)

$$\underset{2.18}{\underline{-\text{O}-\\ \text{aromatic PhOR}}} + \underset{2.413}{\underline{-\text{C}-}} + \underset{2.90}{\underline{=\text{O}\\ \text{PhC(OR)}=\text{O}}} + \underset{24.84}{\underline{\text{C}_6\text{H}_5\\ \text{footnote } p\text{, Table I}}} = 32.33$$

We have used the symbol MR for molar refractivity instead of R_M or P_E in order to avoid confusion with the chromatographic $R_M = \log [(1/R_f) - 1]$ and to avoid confusion between electronic polarization (P_E) and polarizability (α) since $P_E = MR = \frac{4}{3}\pi N\alpha$.[10a,c]

4. Structural Notation. In planning for computerized multiple parameter regression analysis with potentially thousands of sets of data and hundreds of different substituents and parameters, it is highly expedient (and much more accurate) to machine load various parameters. This further facilitates large-scale comparisons of parameter types. Machine loading requires the storage and retrieval of values classified both as to substituent and parameter type. Substituents require a suitable system for the handling of chemical structures and substructures. For these reasons Wiswesser line notation has been given for the functions in Table I. A number of such systems have been thoroughly investigated,[14] and it is evident that the Wiswesser line notation (WLN)[14d] combines three important attributes: low-cost file construction and searching, readability, by the chemist (*cf.* connectivity tables), and the widest acceptance thus far by both industry and the literature searching services.

We have modified the canonical WLN rules[14d] to a very slight extent to better specify the structure of a molecular fragment (*i.e.*, substituent) rather than the whole molecule. A simple subroutine will suffice to make them compatible with a file encoded in canonical WLN. These modifications are explained in Table I. This system will, furthermore, interface with our extensive structure-activity and physical parameter data bases.††

Discussion

There are two general aspects of the lipophilic constant which must be kept in mind when comparing π from different *solute* systems. The value of π for a given function is, for the most part, determined by the intrinsic lipophilic or hydrophilic character of the substituent. However, this intrinsic value can be influenced by the environment which includes both substrate and solvent; that is, π is defined operationally. The electronic effect on π can be easily appreciated by noting that log *P*-OH/benzene[8] is −0.67, while the corresponding value from aliphatic alcohols is −1.16. The strong influence of electron withdrawal on the π values of alkyl groups has been shown to be quite general.[15] This effect can be seen when a methylene group is placed between two electron-withdrawing functions: π-CN = −0.57 and π-CH_2CN = −0.57; π-$COCH_3$ =

Table II. Octanol–Water π Constants from the Benzene System and Hammett σ Constants for Multiple Substituents

No.	Substituents[a]	$\Sigma\pi$ Obsd	$\Sigma\pi$ Calcd[b]	$\Sigma\sigma$ Obsd[c]	$\Sigma\sigma$ Calcd[b]	No.	Substituents[a]	$\Sigma\pi$ Obsd	$\Sigma\pi$ Calcd[b]	$\Sigma\sigma$ Obsd[c]	$\Sigma\sigma$ Calcd[b]
1	3-Cl, 4-Cl	1.25[d]	1.42	0.52	0.60	35	3-NO_2, 5-NO_2	−0.64	−0.56	1.39	1.42
2	3-Cl, 4-OH	0.02	0.04	−0.05	0.00	36	3-NO_2, 5-Cl	0.33	0.43	1.07	1.08
3	3-Cl, 4-CH_3	1.29[d]	1.27	0.23	0.20	37	3-NO_2, 5-OH	−0.13	−0.95		0.83
4	3-Cl, 4-OCH_3		0.69	0.27	0.10	38	3-NO_2, 5-NH_2	−0.76	−1.51		0.55
5	3-Cl, 4-NH_2	−0.23	−0.52		−0.29	39	3-NO_2, 5-OCH_3	0.03	−0.30		0.83
6	3-Br, 4-Br	1.51	1.72		0.62	40	3-F, 5-OH	−0.20	−0.53		0.46
7	3-Br, 4-CH_3		1.42	0.15	0.22	41	3-F, 5-NH_2	−0.83	−1.09		0.18
8	3-Br, 4-OCH_3		0.84	0.09	0.12	42	3-Cl, 5-Cl	1.25[d]	1.42	0.75	0.75
9	3-Br, 4-OH	0.22	0.19		0.02	43	3-Cl, 5-OH	0.37	0.04		0.49
10	3-Br, 4-NH_2	0.16[e]	−0.37		−0.27	44	3-Cl, 5-NH_2	−0.25	−0.52		0.21
11	3-I, 4-OH	0.52	0.45		−0.02	45	3-Br, 5-Br	1.62	1.72	0.72	0.78
12	3-F, 4-OH	−0.42	−0.53		−0.03	46	3-Br, 5-OH	0.50	0.19		0.51
13	3-CH_3, 4-CH_3	0.99	1.12	−0.30	−0.24	47	3-Br, 5-NO_2	0.51	0.58		1.10
14	3-CH_3, 4-OCH_3		0.54	−0.26	−0.34	48	3-I, 5-OH	0.80	0.45		0.47
15	3-CH_3, 4-NO_2	0.17[d]	0.28	0.69	0.71	49	3-CH_3, 5-CH_3	1.07[d]	1.12	−0.17	−0.14
16	3-CH_3, 4-$N(CH_3)_2$	0.68[f]	0.74	−0.30	−0.90	50	3-CH_3, 5-Cl	1.15[d]	1.27	0.35	0.30
						51	3-CH_3, 5-NO_2	0.32	0.28		0.64
17	3-CH_3, 4-Cl	1.29	1.27	0.17	0.16	52	3-CH_3, 5-OH	−0.17	−0.11		0.05
18	3-CH_3, 4-NH_2	−0.81[d]	−0.67	−0.72	−0.73	53	3-CH_3, 5-NH_2	−0.73	−0.67		−0.23
19	3-CH_3, 4-OH	−0.18	−0.11		−0.44	54	3-OCH_3, 5-OCH_3	0.08	−0.04	0.05	0.24
20	3-OCH_3, 4-OCH_3	0.08	−0.04	−0.12	−0.15						
						55	3-OCH_3, 5-Cl		0.69	0.44	0.49
21	3-OCH_3, 4-Cl		0.69	0.34	0.35	56	3-OCH_3, 5-OH	−0.55	−0.69		0.24
22	3-OCH_3, 4-OH		−0.69	−0.33	−0.25	57	3-OH, 5-OH	−1.33	−1.34	0.16	0.24
23	3-NO_2, 4-NO_2	−0.55[d]	−0.56	1.38	1.49	58	3-OH, 5-NH_2	−1.96	−1.90		−0.04
24	3-NO_2, 4-Cl	0.11[d]	0.43	0.90	0.94	59	3,4,5-$(OCH_3)_3$	−0.60	−0.06	0.07	−0.03
25	3-NO_2, 4-Br		0.58	0.83	0.94	60	1,3,5-$(CH_3)_3$	1.29	1.68		
26	3-NO_2, 4-OCH_3		−0.30	0.41	0.44	61	1,3,5-$(OH)_3$	−1.97	−2.01		
27	3-NO_2, 4-CH_3	0.17[d]	0.28	0.50	0.54	62	1,3,5-$(NO_2)_3$	−0.95	−0.84		
28	3-NO_2, 4-NH_2	−0.30	−1.51		0.04	63	1,3-$(OH)_2$, 2-NO_2	−0.57	−1.62		
29	3-NO_2, 4-OH	−0.34	−0.95		0.34						
30	3-OH, 4-OH	−1.25	−1.34	−0.28	−0.25	64	2,4-$(NO_2)_2$, 1-CH_3	−0.15	0.00		
31	3-NH_2, 4-NH_2	−1.98[d]	−2.46		−0.82						
32	3-NH_2, 4-OH	−1.51	−1.90		−0.53	65	3-OCH_3, 4-OH, 5-NO_2		−0.97	0.43	0.46
33	3-NH_2, 4-CH_3	−0.73[f]	−0.67	−0.21	−0.33						
34	3-$N(CH_3)_2$, 4-CH_3	0.68[f]	0.74	−0.18	−0.32	66	3-OH, 4-OCH_3, 5-NO_2		−0.97	0.63	0.56

[a] For π constants 3,4 represents the ortho derivative and 3,5 the meta derivative. [b] The calculated value is simply the sum of the values from Table I for the monosubstituted benzenes (π) or benzene derivatives (σ). [c] From H. H. Jaffé, Table I, footnote *g*, ref 4. [d] See footnote **. [e] From Y. Ichikawa, T. Yamano, and H. Fujishima, *Biochim. Biophys. Acta*, **171**, 32 (1969). [f] From O. E. Schultz, C. Jung, and K. E. Moller, *Z. Naturforsch., B*, **25**, 1024 (1970).

−0.55 and π-CH_2COCH_3 = −0.69. A methylene unit normally increases π by 0.50 units.[5b] In the above two examples and many others of this type,[15] no increase in π is observed. Generally, however, there is no correlation of π-/benzene with electronic effects; for 93 substituents, π is completely uncorrelated with $\mathcal{F}$ and/or $\mathcal{R}$ ($r \sim 0.21$, $s \sim 1.5$). Furthermore, there is no correlation of π with MR ($n = 116$, $r = 0.159$, $s = 1.30$).

The phenyl group is not quite additive (differing by about −0.2) when attached to another aromatic ring. In Table I, π-C_6H_5 is 1.96, while π-C_6H_5 from 2-phenylquinoline is 1.87, from 4-phenylpyridine it is 1.80, and from 2-phenylthiophene it is 1.93. The mean deviation from 2.13 is ±0.24. Whether this is due to a steric or electronic effect (or both) is not yet clear.

There is a range in π in Table I of −5.96–3.30 (over nine powers of ten). Many other functions, more lipophilic than those listed, can be constructed with confidence simply by adding CH_2, halogen, or phenyl groups to a known function and following additivity rules in calculating π.[5b,15] It is relatively easy to obtain variation in π of 10–12 log units; that is, one can design molecules having only one functional group but varying as much as 100 billion in their relative hydrophobicity.

Exner[16] has commented upon the dominance of MW in giving apparent additivity in MR. This will undoubtedly be true of individual substituents; however, we find that for 220 substituents the correlation of MR with MW is not high ($r = 0.759$, $s = 8.97$). Thus, less than 60% of the overall variation in MR is accounted for by MW, and "volume" (approximated by MR) is only roughly correlated by "mass." In other words, the density of the substituents is not very constant. This is interesting because $(MW)^{1/3}$ is sometimes used as an approximation to molecular radius in diffusion studies.[17]

Figure 1 shows a plot of σ_p *vs.* σ_m for 191 substituents. Exner[18] and Taft, *et al.*,[19] have studied subsets of this group, but Figure 1 is an impressive demonstration of the general interdependence of these two substituent parameters, especially considering the variety of sources from which the data were obtained. (The single outlier corresponds to CH=CHCOOH. Since this substituent contains no special structural features compared to others, it is most likely that either or both of the parameter values are in error; they come from two different sources. No other values were found in the literature for confirmation.) Note that the full range of charged, uncharged, and even organometallic substituents is accommodated with surprising precision ($r = 0.903$, $s = 0.17$). Considered as vectors, σ_m and σ_p are separated by an angle of only 25° (arcosine 0.903).[8b] On the other hand, the simple factorization of Swain and Lupton achieves a remarkably high separation of effects as shown in Figure 2. The correlation of $r = 0.222$ corresponds to an angle of 77° (arcosine 0.222). Similarly, the angle between σ_p and $\mathcal{F}$ is 39° ($r = 0.780$), and the angle between σ_m and $\mathcal{F}$ is 13° ($r = 0.973$).

Since $\mathfrak{F}$ and $\mathfrak{R}$ were calculated directly from σ_m and σ_p, they are coplanar with them; that is, the angle between any other vector **Z** and the plane of $\mathfrak{F}$ and $\mathfrak{R}$ or σ_m and σ_p (or any two of the four) is the same; that is, the correlation coefficients will be the same. Thus it is clear that an increasing angle (decreasing correlation coefficients) between the vectors will be reflected in plots such as Figures 1 and 2 by increasing scatter. Plotting a parameter against a simple multiple of itself will yield a perfect straight line. As the second parameter becomes more and more unlike the first, the points on the original line will disperse, eventually forming a perfect scatter plot. Hence, limited relationships found by Exner[18] and Taft, *et al.*,[19] are, in this sense, artifactitious. This is confirmed by employing truly orthogonal vectors,[8a] in which many of the special linear relations are eliminated. Subgroups on the basis of uniting atoms generally tend to have constant resonance values and differ only in field-inductive effects. On this scale,[8a] σ_m has about 30% resonance and σ_p 50% resonance, not a very large difference. These results further confirm the dual nature of electronic substituent effects.[6,8,‡] Further discussion of the interrelations of parameters will be given in the accompanying paper on clustering.[12]

Table II provides some insight into the additivity of π and σ. For disubstitution and three examples of trisubstitution (consult Jaffé's survey, Table I, footnote *g*, ref 4), σ shows consistently good additivity. There are only four examples (**16, 34, 54,** and **65**) where the difference between observed and calculated values is much greater than experimental error. In the case of π, additivity is not nearly so good. The most pronounced effects occur when strong electron-withdrawing groups are placed on the ring with groups having lone pair electrons ($-OH$, $-NH_2$). The $-NO_2$ function or even halogen functions, when combined with $-OH$ and $-NH_2$, give higher than expected π values.

Acknowledgment. We wish to thank Dr. Rolf Appel of the University of Bonn for a generous sample of $PhNHSO_2NHSO_2NH_2$ and Dr. W. A. Sheppard of the Du Pont Company for a sample of $3\text{-}NH_2PhSF_5$ from which we made $PhSF_5$.

References

(1) N. B. Chapman and J. Shorter, Ed., "Advances in Linear Free Energy Relationships," Plenum Press, New York, N. Y., 1972.
(2) (a) C. Hansch, J. Schaeffer, and R. Kerley, *J. Biol. Chem.*, **247**, 4703 (1972); (b) J. Fastrez and A. R. Fersht, *Biochemistry*, **12**, 1067 (1973); (c) M. Stockdale and M. J. Selwyn, *Eur. J. Biochem.*, **21**, 565 (1971).
(3) (a) C. Hansch in "Drug Design," Vol. 1, E. J. Ariëns, Ed., Academic Press, New York, N. Y., 1971, p 271; (b) W. P. Purcell, J. A. Singer, K. Sundaram, and G. L. Parks in "Medicinal Chemistry," A. Burger, Ed., Wiley-Interscience, New York, N. Y., 1970, p 164; (c) B. Duperray, *Chim. Ther.*, **6**, 305 (1971); (d) *Advan. Chem. Ser.*, **No. 114** (1972); (e) R. W. Taft in "Steric Effects in Organic Chemistry," M. S. Newman, Ed., Wiley, New York, N. Y., 1956, p 556.
(4) (a) L. P. Hammett, "Physical Organic Chemistry," 2nd ed, McGraw-Hill, New York, N. Y., 1970; (b) J. E. Leffler and E. Grunwald, "Rates and Equilibria of Organic Reactions," Wiley, New York, N. Y., 1963.
(5) (a) T. Fujita, J. Iwasa, and C. Hansch, *J. Amer. Chem. Soc.*, **86**, 5175 (1964); (b) A. Leo, C. Hansch, and D. Elkins, *Chem. Rev.*, 71, 525 (1971).
(6) C. G. Swain and E. C. Lupton, *J. Amer. Chem. Soc.*, **90**, 4328 (1968).
(7) R. W. Taft and I. C. Lewis, *J. Amer. Chem. Soc.*, **80**, 2436 (1958).
(8) (a) S. H. Unger, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, Mass., Sept 15, 1970; (b) S. H. Unger and C. Hansch, *J. Med. Chem.*, **16**, 745 (1973).
(9) A. Bondi, *J. Phys. Chem.*, **68**, 441 (1964).
(10) (a) D. Agin, L. Hersh, and D. Holtzman, *Proc. Nat. Acad. Sci. U. S.*, **53**, 952 (1965); (b) C. Hansch and E. Coats, *J. Pharm. Sci.*, **59**, 731 (1970).
(11) B. R. Baker, "Design of Active-Site-Directed Irreversible Enzyme Inhibitors," Wiley, New York, N. Y., 1967.
(12) C. Hansch, S. Unger, and A. B. Forsythe, *J. Med. Chem.*, **16**, 1217 (1973).
(13) F. W. Baker, R. C. Parish, and L. M. Stock, *J. Amer. Chem. Soc.*, **89**, 5678 (1967).
(14) (a) M. F. Lynch, J. M. Harrison, W. G. Towne, and J. E. Ash, "Computer Handling of Chemical Structure Information," Macdonald/American Elsevier, New York, N. Y., 1971; (b) L. H. Campey, E. Hyde, and A. R. H. Jackson, *Chem. Brit.*, **6**, 427 (1970); (c) G. W. Gibson and C. E. Granito, *Amer. Laboratory*, **4**, 27 (1972); (d) E. G. Smith, "The Wiswesser Line-Formula Chemical Notation," McGraw-Hill, New York, N. Y., 1968.
(15) C. Hansch, A. Leo, and D. Nikaitani, *J. Org. Chem.*, **37**, 3090 (1972).
(16) O. Exner, *Collect. Czech. Chem. Commun.*, **31**, 3222 (1966).
(17) S. Kitahara, E. Heinz, and C. Stahlmann, *Nature (London)*, **208**, 187 (1965).
(18) O. Exner and J. Lakomý, *Collect. Czech. Chem. Commun.*, **35**, 1371 (1970), and earlier references cited therein.
(19) R. W. Taft, E. Price, I. R. Fox, I. C. Lewis, K. K. Anderson, and G. T. Davis, *J. Amer. Chem. Soc.*, **85**, 3146 (1963).