Structure–Activity Correlations for Antibacterial Agents on Gram-Positive and Gram-Negative Cells[1]

ERIC J. LIEN, CORWIN HANSCH, AND SUSAN M. ANDERSON

Department of Chemistry, Pomona College, Claremont, California

Received December 19, 1967

The antibacterial activity of sets of alcohols, amines, phenols, alkyl-β-naphthols, aromatic and aliphatic isothiocyanates, diguanidines, diamidines, phenyl methacrylates, N[1]-alkylnikethamide chlorides, arylnitroalkenes, ureas, benzyl alcohols, alkyl sulfates, α-bromo and α-hydroxy soaps, and quinine derivatives has been correlated with their chemical structure. It is shown by means of substituent constants and regression analysis that the lipophilic character of the molecule or substituent as expressed by log P or π is the most important factor in determining the activities of the compounds examined. The ideal lipophilic character (log P_0) for gram-negative cells has been found to be about 4, but that for gram-positive cells is about 6. Where linear dependence on log P or π is observed (less than supraoptimal lipophilic character was studied), the slope relating log BR and log P or π is about 0.7. This is very close to that found for the equation correlating the binding of phenols by bovine serum albumin. This work clearly shows the great advantage in using the octanol-water reference system for comparing the dependence of biological activity on hydrophobic character of work of different investigators using different sets of drugs in different biological systems.

In extending our use of a mathematical model for the correlation of chemical structure with biological activity,[2-5] we turn our attention in this report to antibacterial agents. Since the classic work of Meyer and Overton, considerable effort has been made to find linear relations between the nonspecific toxicity of organic compounds and their lipophilic character. Often oil–water partition coefficients have been used to define lipophilic character and $1/C$ to define relative toxicity in a standard test. C is the molar concentration of the drug necessary to cause a standard biological response (BR). Equation 1 represents a way we have found convenient for formulating the relationship

$$\log \mathrm{BR} \equiv \log (1/C) = k \log P + k' \qquad (1)$$

In eq 1, k and k' are constants best evaluated by the method of least squares, and P represents the partition coefficient.

Although there have been many scattered attempts to correlate structure and activity using partition coefficients, there has been no serious reported attempt outside of our laboratory to study many different sets of drugs acting on different biological systems using a *single* reference system. The biggest deterrent to such studies has been of course the large effort necessary to measure the many hundreds of partition coefficients. Our discovery of the additive character of log P[6] makes it possible to calculate many partition coefficients from relatively few base values. This has greatly expedited our work. We have used the 1-octanol-water system as our standard reference. This then allows one to compare the lipophilic properties of different sets of congeners acting in different systems. For example, we have found[7] that a large variety of inhibitors of oxidative metabolism in a variety of different biological systems (bacteria, brain tissue, tadpoles, mitochondria, etc.) all show the same relative dependence on P for their toxic action. For 15 different examples conforming to eq 1, we found a range of slopes of only 0.80–1.3 with a mean of 1.04.

It has been our hypothesis that for the general case (where there is a very wide range of lipophilic character in a set of congeners), one should not expect a linear relationship between log $(1/C)$ and log P, but instead, one should look for a parabolic relationship. This has led to the development of eq 2. In eq 2 $\rho\sigma$ are the

$$\log (1/C) = -k(\log P)^2 + k' \log P + \rho\sigma + k'' \qquad (2)$$

Hammett constants.[8] In deriving eq 2 we assumed that, in general, in the testing of drugs one does not reach a true equilibrium between drug in the exobiophase and drug at the sites of action. In other words, a molecule of drug has a certain amount of time during the test interval to find the sites of action *via* a random-walk process. The course of the random walk will be highly dependent on the lipophilic character of the drug. Consider the extremes; if P is near zero, then the drug will be so water soluble it will not easily cross a lipophilic membrane and the drug will be localized in the first aqueous phase. As P approaches ∞, the drug becomes so tightly bound to lipophilic phases that it cannot cross aqueous barriers. Somewhere between $P = 0$ and $P = \infty$ there will be an ideal value such that the drug having this partition coefficient will have maximum freedom in the random-walk process. Its probability of reaching the reaction site in the standard test interval will be greatest. We have found, within the possible experimental range of P values, that organic compounds are bound by bovine serum albumin and bovine hemoglobin according to eq 1.[9] We have also found that various body tissues bind barbiturates in much the same fashion.[7] This means that the movement of very lipophilic compounds through biological tissue is severely restricted.

If the partial derivative of eq 2, $\partial \log (1/C)/\partial \log P$, is taken and set equal to zero, we can solve for the constant we call log P_0. This gives the apex of the parabola defined by the first two terms on the right

(1) This work was supported by Research Grant GM 07492 from the National Institutes of Health.

(2) C. Hansch and T. Fujita, *J. Am. Chem. Soc.*, **86**, 1616 (1964).

(3) C. Hansch, A. R. Steward, J. Iwasa, and E. W. Deutsch, *Mol. Pharmacol.*, **1**, 205 (1965).

(4) C. Hansch and E. W. Deutsch, *Biochim. Biophys. Acta*, **126**, 117 (1966).

(5) C. Hansch and E. J. Lien, *Biochem. Pharmacol.*, in press.

(6) (a) T. Fujita, J. Iwasa, and C. Hansch, *J. Am. Chem. Soc.*, **86**, 5175 (1964); (b) C. Hansch and S. M. Anderson, *J. Org. Chem.*, **32**, 2583 (1967); (c) J. Iwasa, T. Fujita, and C. Hansch, *J. Med. Chem.*, **8**, 150 (1964).

(7) C. Hansch and S. M. Anderson, *ibid.*, **10**, 745 (1967).

(8) H. H. Jaffé, *Chem. Rev.*, **53**, 191 (1953).

(9) K. Kiehs, C. Hansch, and L. Moore, *Biochemistry*, **5**, 2602 (1966).

side of eq 2. This log P_0 represents the ideal lipophilic character for a set of congeneric drugs. We have postulated[3,9,10] that, steric and electronic factors remaining constant, different sets of congeneric drugs acting in the same way on the same receptor sites should have the same log P_0 constants. Once log P_0 is found for a given set of compounds, this becomes a useful constant for the design of completely new sets of congeners to act on the same centers. The purpose of this work was to take data from the studies of antibacterial agents and fit them to eq 2, and its simpler forms, in order to explore our thesis concerning log P_0. From some preliminary results[2] it was felt that log P_0 would depend on the type of organism used in the test. Since considerable quantitative work has been carried out in the field of antibacterial agents using a variety of microorganisms, this seemed to be a good field in which to make a comparative study. We are of course quite interested in the differential susceptibility of gram-positive and gram-negative microorganisms to various agents.[11]

Methods

The biological data[12–33] and physicochemical parameters are assembled in Table I. We have used two methods of expressing relative biological activity. One, using log $(1/C)$, is defined above. The other, using PC', refers to the phenol coefficient converted to a molar basis.

To derive the equations in the section on results, we have used the method of least squares and an IBM 360/40 computer. The values of σ were taken from the compilation of Jaffé.[8]

The log P values refer to the neutral molecules.[2] Some of the values in Table I were obtained experimentally and others were calculated, taking advantage of the additive–constitutive nature of log P.[6] For compounds used for eq 3 and 30, 0.5 was added for each CH_2 unit to the experimental value of 1.34 found for 4-methoxyphenol. For a branch in a chain, 0.2 unit was subtracted. For example, 1.50 was used for *n*-propyl and 1.30 for isopropyl. The same procedure was followed for the molecules used in eq 4, 5, 31, and 30. For 4-phenoxyphenol we subtracted 0.5 from 4-methoxyphenol and added 2.13 for the phenyl moiety. For eq 6 and 29, the value of 0.62 for the CH_3S group was taken from the phenoxyacetic acid system[6a] and added to 1.46 for phenol to obtain 4-methylthiophenol. Hydrogen is defined as zero. The higher members of the series were then calculated by adding CH_2 unit values. For eq 12, the experimental value for cyclohexanol is 1.23. The methyl derivative values were obtained by adding 0.5 to 1.23. The log P of diethylcarbinol was obtained by adding 0.5 to 0.61, the experimental value for 2-butanol. The figure for triethylcarbinol was obtained by adding 1.50 for three CH_2 units to log P for *t*-butyl alcohol (0.37). The values used for benzyl alcohol and phenethyl alcohol are the experimentally found ones. For eq 13, 14, 37, 38, 47, 49, 52, 54, and 55 where the charged functional group makes it almost impossible to obtain accurate log P values in the octanol–water system, we have taken the functional group as an unknown constant and simply used 0.5 for each CH_2 unit. This allows us to determine the dependence of biological activity on lipophilic character in terms of the slope but not the intercept. The log P values for the molecules used to find eq 9, 10, 19-21, and 40 were based on the experimentally found values for *n*-butylamine (0.81), di-*n*-propylamine (1.73), triethylamine (1.44), aniline (0.90), N-methylaniline (1.66), N,N-dimethylaniline (2.31), and quinoline (2.03). Tetrahydroquinoline was calculated by adding 4 × 0.41 for the four cyclic CH_2 units[6b] to 0.65 for pyridine. Log P for naphthylamine was found by adding 1.35 for the $(CH)_4$ moiety to 0.90 for aniline. The log P values for alcohols of eq 11, 35, and 36 were based on the value of −0.66 for methanol, 0.37 for *t*-butyl, and 0.89 for *t*-amyl alcohol. For the thiocyanates of eq 16, log P for the phenyl derivative was measured. For the congeners in this set π values from the benzene system[6a] were used except for 4-I which was taken from the phenoxyacetic acid system. The phenoxy group was calculated by subtracting 0.5 from 2.11 for anisole. For the 2-naphthyl derivative, 1.35 was added for $(CH)_4$. For eq 17 and 46 we elected to hold the common functional group (NCS) constant and use π values for the rest of the molecules. Where a functional group is attached to an alkyl moiety, aliphatic values are used.[6b] For example, to estimate π for the group $CH_3CH(CN)CH_2$, we add −0.84 for aliphatic CN to 1.30 for isopropyl to obtain 0.46. For those mixed aliphatic aromatic compounds, $4\text{-}NCC_6H_5CH_2$ serves as an example. To the value of 2.69 for toluene we add −0.57 for an aromatic CN to obtain 2.12. The other members were calculated in the same way. For the 2-Cl function we used the value of 0.59 from the phenoxyacetic acid system.

For eq 18 and 45 we used π values from the phenoxyacetic acid system except for $4\text{-}(Et)_2N$ which was based on $(CH_3)_2N$ (0.18) from the benzene system. Log P values for eq 7, 15, 28, and 48 were based on the ex-

(10) C. Hansch, A. R. Steward, S. M. Anderson, and D. Bentley, *J. Med. Chem.*, **11**, 1 (1968).
(11) R. J. Dubos, *Ann. Rev. Biochem.*, **11**, 659 (1942).
(12) (a) E. Klarmann, L. W. Gates, and V. A. Shternov, *J. Am. Chem. Soc.*, **54**, 298 (1932); (b) *ibid.*, **54**, 1204 (1932).
(13) C. M. Sutter, *Chem. Rev.*, **28**, 269 (1941).
(14) E. Klarmann, V. A. Shternov, and L. W. Gates, *J. Am. Chem. Soc.*, **55**, 2576 (1933).
(15) E. Klarmann, L. W. Gates, V. A. Shternov, and P. H. Cox, Jr., *ibid.*, **55**, 4657 (1933).
(16) M. L. Khorana, S. Y. Pandit, and A. D. Pishawikar, *J. Pharm. Sci.*, **56**, 993 (1967).
(17) (a) F. W. Tilley and J. M. Schaffer, *J. Bacteriol.*, **16**, 279 (1928); (b) *ibid.*, **12**, 303 (1926).
(18) I. J. Kligler, *J. Exp. Med.*, **27**, 463 (1918).
(19) J. M. Schaffer and F. W. Tilley, *J. Bacteriol.*, **14**, 259 (1927).
(20) A. T. Fuller, *Biochem. J.*, **36**, 548 (1942).
(21) D. Vlachová and L. Drobnica, *Collect. Czech. Chem. Commun.*, **31**, 997 (1966).
(22) R. Woodside, M. Zief, and G. Sumrell, *Antibiot. Chemotherapy*, **9**, 470 (1959).
(23) A. F. McKay, L. D. Garmaise, R. Gaudry, H. A. Baker, G. Y. Paris, R. W. Kay, G. E. Just, and R. Schwartz, *J. Am. Chem. Soc.*, **81**, 4328 (1959).
(24) O. Schales and H. A. Graefe, *ibid.*, **74**, 4486 (1952).
(25) D. E. Burton, K. Clarke, and G. W. Gray, *J. Chem. Soc.*, 1314 (1964).
(26) D. V. Carter, P. T. Charlton, A. H. Fenton, J. R. Housley, and B. Lessel, *J. Pharm. Pharmacol.*, **10**, suppl. 149T (1958).
(27) D. J. Beaver, D. P. Roman, and P. J. Stoffel, *J. Am. Chem. Soc.*, **79**, 1236 (1957).
(28) C. O. Wilson and O. Gisvold, "Textbook of Organic Medicinal and Pharmaceutical Chemistry," 4th ed, J. B. Lippincott Co., Philadelphia, Pa., 1962: (a) p 15; (b) p 221.
(29) G. Wirgin, *Z. Hyg. Infektionskrankh.*, **46**, 149 (1904).
(30) P. B. Cowles, *Yale J. Biol. Med.*, **11**, 33 (1938).
(31) H. Braun and H. Schaeffer, *Berlin. Klin. Wochschr.*, **54**, 885 (1917).
(32) (a) A. H. Eggerth, *J. Exp. Med.*, **49**, 53 (1929); (b) *ibid.*, **50**, 229 (1929).
(33) W. Ciusa and A. Buccell, *Gazz. Chim. Ital.*, **95**, 1455 (1965).

TABLE I

DATA USED IN DERIVATION OF EQUATIONS IN TABLES II AND III

Compd	Log P	σ	Log PC' obsd[12b] Eq 3	Eq 30
4-HOC_6H_5OH	0.59	−0.36	...	−0.30
4-MeO	1.34	−0.27	0.12	0.02
4-EtO	1.84	−0.25	0.34	0.34
4-n-PrO	2.34	−0.27	0.94	0.82
4-i-PrO	2.14	−0.29	0.76	0.57
4-n-BuO	2.84	−0.32	1.39	1.21
4-n-AmO	3.34	−0.34	1.74	1.76
4-s-AmO	3.14	−0.36	1.56	1.70
4-n-HexO	3.84	−0.34	1.57	2.31
4-n-HepO	4.34	−0.34	1.58	2.65
4-n-OctO	4.84	−0.34	...	2.93
4-PhO	2.95	−0.03	1.91	1.74
4-$PhCH_2O$	3.47	−0.42	1.65	1.47

Compd	Log P	σ	Log PC' obsd[12b] Eq 4	Eq 30
C_6H_5OH	1.46	0.00	0.00	0.00
3-HO	0.80	0.00	−0.33	−0.33
3-MeO	1.58	0.12	0.23	0.20
3-EtO	2.08	0.15	0.72	0.64
3-n-PrO	2.58	0.15	1.05	0.94
3-n-BuO	3.08	0.15	1.55	1.50
3-n-AmO	3.58	0.15	1.86	1.84
3-s-AmO	3.38	0.10	1.70	1.77
3-n-HexO	4.08	0.15	1.98	2.41
3-n-HepO	4.58	0.15	1.67	2.86
3-PhO	3.21	0.25	1.90	1.86

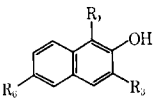
Compd	Log P	σ	Log PC' obsd[13] Eq 5	Eq 33
C_6H_5OH	1.46	0.00	0.00	0.00
4-HO	0.80	0.00	−0.40	...
3-HO-4-n-Pr	2.30	−0.13	0.95	0.82
3-HO-4-n-Bu	2.80	−0.16	1.63	1.29
3-HO-4-i-Bu	2.60	−0.15	1.46	...
3-HO-4-n-Am	3.30	−0.16	1.84	2.80
3-HO-4-i-Am	3.10	−0.15	1.70	...
3-HO-4-n-Hex	3.80	−0.16	2.06	2.34
3-HO-4-i-Hex	3.60	−0.15	1.78	...
3-HO-4-n-Hep	4.30	−0.16	1.85	2.82
3-HO-4-n-Oct	4.80	−0.16	...	3.24
3-HO-4-n-Non	5.30	−0.16	...	3.42

Compd	Log P	σ	Log PC' obsd[3] Eq 6	Eq 29
C_6H_5OH	1.46	0.00	0.00	0.00
4-MeS	2.08	−0.05	0.87	0.77
4-EtS	2.58	−0.05	1.29	1.29
4-n-PrS	3.08	−0.05	1.65	1.65
4-n-BuS	3.58	−0.05	2.16	2.06
4-n-AmS	4.08	−0.05	2.16	2.49
4-n-HexS	4.58	−0.05	1.95	2.65
2-Me-4-MeS	2.64	−0.22	1.33	1.29
2-Me-4-EtS	3.14	−0.22	1.55	1.95
2-Me-4-n-PrS	3.64	−0.22	1.65	2.19
2-Me-4-n-BuS	4.14	−0.22	1.43	2.28
2-Me-4-n-AmS	4.64	−0.22	1.22	2.71

Compd	Log P	Log PC' obsd[3] Eq 7	Eq 15	Eq 28	Eq 48
4-ClC_6H_4OH	2.39	0.77	0.81	0.77	0.78
4-Cl-2-Me	2.89	1.28	1.34	1.28	1.23
4-Cl-2-Et	3.39	1.68	1.73	1.76	1.72
4-Cl-2-n-Pr	3.89	2.23	2.26	2.23	2.15
4-Cl-2-n-Bu	4.39	2.44	2.52	2.70	2.69
4-Cl-2-n-Am	4.89	2.52	2.63	3.03	3.07
4-Cl-2-s-Am	4.69	2.00	2.23	2.82	2.82
4-Cl-2-n-Hex	5.39	1.72	2.88	3.45	3.48
4-Cl-2-c-Hex	4.90	...	2.25	2.99	2.91
4-Cl-2-Hep	5.89	...	2.51	3.56	3.73
4-Cl-2-n-Oct	6.39	...	1.83	3.65	...
4-Cl-2-s-Oct	6.19	...	...	3.41	...
2-Cl	2.15	0.53	0.50	0.60	0.43
2-Cl-4-Me	2.65	0.98	0.91	1.06	0.93
2-Cl-4-Et	3.15	1.46	1.35	1.42	1.40
2-Cl-4-n-Pr	3.65	1.83	1.86	1.77	1.80
2-Cl-4-n-Bu	4.15	2.23	2.20	2.27	2.24
2-Cl-4-n-Am	4.65	2.22	2.23	2.78	2.67
2-Cl-4-t-Am	4.33	1.83	2.00	2.42	2.47
2-Cl-4-n-Hex	5.15	...	1.91	3.21	3.15
2-Cl-4-n-Hep	5.65	...	2.96	2.93	2.68
4-Cl-3-Me	2.95	1.21	...	1.24	1.24
4-Cl-3,5-Me_2	3.51	1.70	...	1.63	1.66
4-Cl-6-Et-3-Me	3.95	2.07	...	1.96	2.00
4-Cl-6-n-Pr-3-Me	4.45	2.42	...	2.60	2.54
4-Cl-6-i-Pr-3-Me	4.25	2.32	...	2.47	2.43
4-Cl-2-Et-3,5-Me_2	4.51	1.96	...	2.32	2.27
4-Cl-6-s-Bu-3-Me	4.77	1.96	...	2.86	2.85
4-Cl-2-i-Pr-3,5-Me_2	4.81	2.24	...	2.82	2.82
4-Cl-6-Et_2Me-3-Me	5.25	1.78	...	3.19	3.10
4-Cl-6-i-Pr-2-Et-3-Me	5.25	2.11	...	2.66	2.60
4-Cl-2-s-Bu-3,5-Me_2	5.31	1.81	...	3.11	3.10
4-Cl-2-s-Am-3,5-Me_2	5.81	...	...	3.26	3.43
4-Cl-2-Et_2Me-3,5-Me_2	5.81	...	...	3.44	3.38
4-Cl-2-s-Oct-3-Me	6.75	...	...	...	2.52
4-Cl-2-s-Oct-3,5-Me_2	7.31	...	...	2.46	...

TABLE I (*Continued*)

Compd	Log P	Log PC^r obsd[15] Eq 8	Eq 32
4-Br-C_6H_4OH	2.59	1.04	0.96
4-Br-2-Me	3.09	1.39	1.35
4-Br-2-Et	3.59	1.82	1.73
4-Br-2-*n*-Pr	4.09	2.15	2.15
4-Br-2-*n*-Bu	4.59	2.58	2.88
4-Br-2-*n*-Am	5.09	3.21	3.17
2-Br	2.35	0.78	0.75
2-Br-4-*t*-Am	4.53	1.93	2.59
4-Br-2-*s*-Am	4.89	...	2.59
4-Br-2-*n*-Hex	5.59	...	3.53
4-Br-2-*c*-Hex	5.10	...	3.06
2-Br-4-*n*-Hex	5.35	...	3.23
2-Br-4-*n*-Pr-3,5-Me_2	4.97	...	2.96



R_1	R_3	R_6	Log P	Log $(1/C)$ obsd[16] Eq 26	Eq 42
H	H	H	2.84	3.16	2.86
H	H	*i*-Am	5.14	3.33	4.33
H	H	*i*-Hex	5.64	3.36	4.66
H	H	*i*-Hep	6.14	3.38	4.68
H	Me	H	3.34	3.20	3.50
H	Et	H	3.84	3.94	3.94
H	Pr	H	4.34	4.27	4.27
H	Bu	H	4.84	4.00	5.00
Me	H	H	3.34	3.20	3.20
Et	H	H	3.84	3.24	3.24
Pr	H	H	4.34	3.27	3.27
Me	H	Me	3.84	3.24	3.94
Me	H	Et	4.34	3.97	4.27
Et	H	Et	4.84	4.30	4.30
Et	H	Pr	5.34	4.33	5.03
Et	H	Bu	5.84	...	5.06
Et	H	Am	6.34	...	5.38
Et	H	Hex	6.84	...	5.19
Me	H	Pr	4.84	3.00	4.60
Me	H	Bu	5.34	3.03	5.03
Me	H	Am	5.84	3.06	5.06
Me	H	Hex	6.34	3.08	5.38

Compd	Log P	pK_a	Log PC^r obsd[17a] Eq 9	Eq 40
n-$PrNH_2$	0.31	10.53[b]	0.00	−0.85
n-Bu	0.81	10.60[b]	0.20	−0.40
n-Am	1.31	10.60[a]	0.45	0.08
n-Hex	1.81	10.60[a]	0.68	0.52
n-Hep	2.31	10.60[a]	0.90	...
Et_2NH	0.73	10.93[b]	0.23	−0.80
$(n\text{-Pr})_2$	1.73	10.93[a]	0.34	...
$(n\text{-Bu})_2$	2.73	10.93[a]	0.81	...
Et_3N	1.44	10.87[b]	0.20	−0.57
$PhCH_2NH_2$	1.09	9.34	0.04	...

Compd	Log P	Log PC^r obsd[17b] Eq 11	Eq 36
MeOH	−0.66	−2.05	−2.00
EtOH	−0.16	−1.70	−1.72
n-PrOH	0.34	−1.19	−1.28
n-BuOH	0.84	**−0.67**	−0.76
n-AmOH	1.34	−0.13	−0.23
n-HexOH	1.84	0.40	...
n-HepOH	2.34	0.92	...
n-OctOH	2.84	1.46	...
i-PrOH	0.14	−1.39	−1.47
s-BuOH	0.61	−0.92	−0.99
s-AmOH	1.11	−0.44	−0.52
s-HexOH	1.51	0.04	...
t-BuOH	0.37	−1.19	−1.30
t-AmOH	0.89	−0.77	−0.88
t-HexOH	1.39	−0.31	...

Compd	Log P	pK_a	Log $(1/C)$ obsd[18] Eq 10	Eq 19	Eq 20	Eq 21
$C_6H_5NH_2$	0.90	4.58[b]	4.51	4.37	4.45	4.62
2-$MeC_6H_4NH_2$	1.40	4.39[b]	4.71	4.71	4.71	4.81
4-$MeC_6H_4NH_2$	1.39	5.12[b]	4.77	4.71	4.71	4.81
C_6H_5NHMe	1.66	4.85[b]	4.84	4.81	4.84	5.01
C_6H_5NHEt	2.16	5.11[b]	5.12	5.08	5.08	5.21
2-MeC_6H_4NHMe	2.16	5.13[a]	5.12	5.06	5.12	5.23
4-MeC_6H_4NHMe	2.15	5.03[a]	5.12	5.06	5.12	5.23
$C_6H_5N(Me)_2$	2.31	5.06[b]	5.26	5.21	5.21	5.33
2-Me-$C_6H_4NMe_2$	2.81	5.86[b]	5.37	5.32	5.32	5.45
4-Me-$C_6H_4NMe_2$	2.80	4.94[c]	5.59	5.55	5.53	5.67
$C_6H_5NEt_2$	3.31	6.56[b]	5.95	5.83	5.83	6.02
Quinoline	2.03	4.94[b]	5.32	5.15	5.26	5.32
Tetrahydroquinoline	2.29	5.13[a]	5.33	5.27	5.33	5.33
2-Me-Quinoline	2.53	5.87[d]	5.57	5.48	5.57	5.57
1-Naphthylamine	2.25	3.92[b]	5.57	5.57	5.57	5.65

Compd	Log P	Log PC^r obsd[19] Eq 12
Cyclohexanol	1.23	−0.26
o-Me	1.73	0.10
m-Me	1.73	0.18
p-Me	1.73	0.20
Diethylcarbinol	1.11	−0.47

$(CH_2)_n\begin{pmatrix} CNH_2 \\ \| \\ NH \end{pmatrix}_2$

n	π	Log $(1/C)$ obsd[20] Eq 13	Eq 37
8	4.00	2.20	2.77
9	4.50	2.41	3.17
10	5.00	2.97	3.66
11	5.50	3.23	4.08
12	6.00	3.88	4.40

TABLE I (*Continued*)

Compd	Log P	Log PC' obsd[19] Eq 12
Triethylcarbinol	1.87	0.07
Benzyl alcohol	1.10	−0.06
Phenethyl alcohol	1.60	0.07

$(CH_2)_n$(C(=NH)NH$_2$)$_2$

n	s	Log (1/C) obsd[20] Eq 13	Eq 37
13	6.50	4.13	4.73
14	7.00	4.33	4.75
16	8.00	4.47	4.77

$(CH_2)_n$(NHC(=NH)NH$_2$)$_2$

Compd	Log P	σ	Log PC' obsd[21] Eq 16	n	π	Log (1/C) obsd[20] Eq 14	Eq 38	Eq 49
C_6H_5NCS	3.28	0.00	4.20	5	2.50	1.75	2.45	2.27
4-Cl	3.99	0.23	4.95	6	3.00	2.00	2.85	2.78
3-Br	4.14	0.39	5.05	8	4.00	3.03	3.51	3.39
4-Br	4.14	0.23	4.95	10	5.00	3.96	4.26	4.36
4-I	4.54	0.28	5.05	12	6.00	4.24	5.23	4.78
4-EtOOC	3.73	0.45	5.25	14	7.00	4.90	5.71	6.11
4-PhO	4.89	−0.03	4.20	16	8.00	5.28	5.13	5.73
4-NO_2	3.00	0.78	4.15	18	9.00	5.15	5.15	6.06
2-Naphthyl	4.63	0.17	4.70					

CH_2=C(CH_3)COO–C$_6$H$_4$–X

X	Log P	$\Sigma\sigma$	Log (1/C) obsd[22] Eq 24, 25	Eq 41	Eq 51	Eq 56	Eq 57	Eq 60
H	1.99	0.00	...	2.89	2.89	2.89	2.89	2.89
o-Cl	2.58	0.21	3.08	3.08	3.08	3.08	3.08	3.08
p-Cl	2.69	0.23	...	3.12	3.25	3.25	3.25	3.25
m-Cl	2.75	0.37	3.12	2.91	2.77	2.91	2.91	3.12
o,*p*-Cl_2	3.28	0.44	3.28	3.49	3.49	3.80	3.49	3.49
2,4,6-Cl_3	3.87	0.65	2.96	3.76	3.76	3.76	3.76	3.24
2,4,5-Cl_3	4.04	0.81	3.09	3.84	3.84	3.84	3.84	3.54
2,4,5,6-Cl_4	4.63	1.02	3.21	4.97	3.97	4.67	4.45	3.32
Cl_5	5.54	1.39	...	5.19	4.71	5.00	5.00	3.71
Br_5	6.29	1.41	...	5.32	3.64	5.32	3.85	3.64

RNCS

R	$\Sigma\pi$	Log (1/C) obsd[23] Eq 17	Eq 46
$NCCH_2CH_2$	0.16	3.65	3.65
$MeCHCNCH_2$	0.46	4.00	3.40
$NCCH_2MeCH$	0.46	3.70	3.70
$MeOOCCH_2CH_2$	0.73	3.46	3.46
$NC(CH_2)_4$	1.16	3.75	3.75
$PhCH_2$	2.69	4.68	4.68
4-$NCPhCH_2$	2.12	4.57	4.57
4-$ClPhCH_2$	3.40	4.77	4.77
4-NO_2PhCH_2	2.41	4.79	4.79
2,4-Cl_2PhCH_2	3.99	4.54	5.15
3,4-Cl_2PhCH_2	4.11	5.15	5.75
3-NO_2-4-$ClPhCH_2$	3.12	5.04	5.34
2-$ClPhCH_2$	3.28	...	4.77

XPhCH=C(Me)NO_2

X	$\Sigma\pi$	$\Sigma\sigma$	Log (1/C) obsd[23] Eq 18	Eq 45
H	0.00	0.00	...	1.30
2-MeO	−0.33	−0.27	−0.67	1.15
2-EtO	0.17	−0.25	...	1.40
3-MeO	0.12	0.12	−0.25	1.52
4-MeO	−0.04	−0.27	−0.28	1.22
2,3-$(MeO)_2$	−0.21	−0.15	...	1.40
2-Cl	0.59	0.23	−0.20	1.70
2,4-Cl_2	1.29	0.46	...	2.40
3,4-Cl_2	1.46	0.60	0.46	2.52
2-NO_2	−0.23	0.78	0.46	1.10
4-NO_2	0.24	0.78	0.19	1.70
4-MeCONH	−0.79	−0.02	−0.40	0.89
4-Et_2N	1.18	−0.60	0.07	...

Compd	Log P	$\Sigma\sigma$	Log (1/C) obsd[25] Eq 22
C_6H_5OH	1.46	0.00	1.08
2-F	1.71	0.06	1.11
3-F	1.93	0.34	1.26
2-Cl	2.15	0.23	1.28
3-MeO	1.58	0.12	1.34
2-MeO	1.58	−0.27	1.48
2-I	2.65	0.28	1.51
4-F	1.77	0.06	1.53
3-Me	2.02	−0.07	1.60
4-Me	1.94	−0.17	1.60
2-Me	1.96	−0.14	1.70
3-Cl	2.50	0.37	1.70
4-Cl	2.39	0.23	1.75
4-MeO	1.34	−0.27	1.79
2,4-Br_2	3.48	0.46	1.92
3-Br	2.63	0.39	1.96
4-Br	2.59	0.23	1.98
3-I	2.93	0.35	2.23
4-I	2.91	0.28	2.31
4-Cl-3-Me	2.95	0.16	2.34
2,4-I_2	4.10	0.56	2.42

Compd	Log P	$\Sigma\sigma$	Log (1/C) obsd[25] Eq 23	Eq 61
$C_6H_5CH_2OH$	1.10	0.00	4.33	4.51
4-Cl	1.96	0.23	4.76	4.76
2,4-Cl_2	2.55	0.46	5.85	5.55
3,4-Cl_2	2.80	0.60	5.85	5.85
2,4,5-Cl_3	3.39	0.83	...	6.32
3,4,5-Cl_3	3.64	0.97	6.32	6.63
2-Br	1.85	0.23	5.15	5.15
4-Br	2.12	0.23	5.27	5.57
4-I	2.36	0.28	...	5.75
4-Me	1.58	−0.17	4.79	4.79
2,4-Me_2	2.26	−0.34	5.13	5.13
4-Cl-3,5-Me_2	2.96	0.09	5.83	6.05
4-I-3,5-Me_2	3.36	0.14	...	6.72
4-MeO	1.10	−0.27	...	4.84
4-NO_2	1.26	0.78	5.00	5.00
4-CN	0.78	0.63	4.67	4.67
2-NO_2	0.87	0.78	5.18	5.49
4-COOH	0.82	0.27	4.73	4.88

TABLE I (*Continued*)

$XC_6H_4NHCON-HC_6H_3$-3,4-Cl_2 X	Log P	$\Sigma\sigma$	Log $(1/C)$ obsd[27] Eq 27
H	4.71	0.00	3.39
3-Cl	5.47	0.37	6.92
3,4-Cl_2	6.17	0.60	6.50
3,4,5-Cl_3	6.93	0.97	6.54
4-MeO	4.67	−0.27	4.44
4-Me	5.23	−0.17	3.41
4-Ph	6.84	0.01	3.51
4-NMe_2	4.87	−0.27	3.46
4-NH_2	3.48	−0.66	2.42
4-NHPh	5.62	−0.40	3.53
4-SO_2NH_2	2.89	0.62	4.55
4-OH	4.10	−0.36	3.42

p-HOC_6H_4COOR R	Log P	Log PC' obsd[28] Eq 34
Me	1.88	0.62
Et	2.38	1.10
n-Pr	2.88	1.46
i-Pr	2.68	1.40
n-Bu	3.38	1.88
Allyl	2.58	1.36
Benzyl	4.01	2.30
Phenol	1.46	0.00

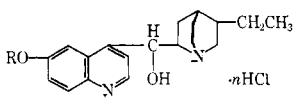
Compd	Log P	Log $(1/C)$ obsd[29] Eq 35
MeOH	−0.66	−0.35
EtOH	−0.16	−0.08
n-PrOH	0.34	0.28
i-PrOH	0.14	0.29
n-BuOH	0.84	0.49
i-BuOH	0.64	0.49
t-BuOH	0.37	0.28
n-AmOH	1.14	1.03
t-AmOH	0.89	0.57

RNH_2 R	Log P	Log $(1/C)$ obsd[20] Eq 39	Eq 50	Eq 58	Eq 59
n-C_9H_{19}	3.31	2.85	4.33	3.55	3.16
n-$C_{12}H_{25}$	4.81	4.44	5.44	4.81	4.44
n-$C_{14}H_{29}$	5.81	4.51	5.81	5.23	4.63
n-$C_{16}H_{33}$	6.81	4.26	5.38	4.98	4.93
n-$C_{18}H_{37}$	7.81	3.61	4.61	5.03	4.31

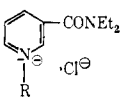
$RCHOHCOO^-K^+$ R	π	Log $(1/C)$ obsd[32b] Eq 52
n-C_8H_{17}	4.00	1.60
n-$C_{10}H_{21}$	5.00	2.51
n-$C_{12}H_{25}$	6.00	3.41
n-$C_{14}H_{29}$	7.00	4.01
n-$C_{16}H_{33}$	8.00	3.71

$ROSO_3^-Na^+$ R	π	Log $(1/C)$ obsd[30] Eq 47
n-Bu	2.00	0.08
n-Am	2.50	0.36
n-Hex	3.00	0.51
n-Hep	3.50	0.94
n-Oct	4.00	1.27
n-Non	4.50	1.79
n-Dec	5.00	2.25
Lauryl	6.00	3.27
Myristyl	7.00	3.88
Cetyl	8.00	3.58

$XPhCH{=}CHNO_2$ X	$\Sigma\pi$	$\Sigma\sigma$	Log $(1/C)$ obsd[24] Eq 44
H	0.00	0.00	0.42
3-MeO	0.12	0.12	0.80
4-MeO	−0.04	−0.27	0.35
3,4-Cl_2	1.46	0.60	1.22
3-NO_2	0.11	0.71	0.62
4-NO_2	0.24	0.78	0.43
2-MeO	−0.33	−0.27	0.54
2-Cl	0.59	0.23	0.85
2,4-Cl_2	1.29	0.46	1.30
2-NO_2	−0.23	0.78	0.68
2-EtO	0.17	−0.25	0.82
4-MeCOO	−0.64	0.31	0.15



R	Log P^f	Log $(1/C)$ obsd[31] Eq 54
Me·HCl	2.03	3.60
Et·HCl	2.53	3.62
i-Pr·HCl	2.83	3.73
i-Bu·HCl	3.33	4.34
H·2HCl	1.53	3.62
Et·2HCl	2.53	3.65
i-Pr·2HCl	2.83	4.06
i-Bu·2HCl	3.33	4.47
i-Am·2HCl	3.83	4.99
n-Hex·2HCl	4.53	5.18
n-Hep·2HCl	5.03	5.41
n-Oct·2HCl	5.53	5.43
s-Oct·2HCl	5.33	5.61
n-Dec·2HCl	6.53	5.46
n-Dodec·2HCl	7.53	4.78
n-Cet·2HCl	9.53	3.52
Quinine·HCl	1.73	3.60



R	π	Log $(1/C)$ obsd[33] Eq 43
CH_3	0.50	3.82
C_2H_5	1.00	3.70
n-C_3H_7	1.50	3.82
n-C_4H_9	2.00	4.00
n-C_5H_{11}	2.50	5.12
n-C_6H_{13}	3.00	5.40
n-C_7H_{15}	3.50	5.12
n-C_8H_{15}	4.00	5.12
n-C_9H_{19}	4.50	6.00
n-$C_{10}H_{21}$	5.00	6.00
n-$C_{11}H_{23}$	5.50	6.30
n-$C_{12}H_{25}$	6.00	6.30
n-$C_{13}H_{27}$	6.50	6.60
n-$C_{14}H_{29}$	7.00	6.60
n-$C_{15}H_{31}$	7.50	6.30
n-$C_{16}H_{33}$	8.00	6.30
n-$C_{17}H_{35}$	8.50	6.30
n-$C_{18}H_{37}$	9.00	6.00
n-$C_{19}H_{39}$	9.50	6.00
n-$C_{20}H_{41}$	10.00	6.30

$RCHBrCOO^-K^+$ R	π	Log $(1/C)$ obsd[32a] Eq 53	Eq 55
n-C_6H_{13}	3.00	2.20	1.90
n-C_8H_{17}	4.00	3.11	2.51
n-$C_{10}H_{21}$	5.00	3.71	2.81
n-$C_{12}H_{25}$	6.00	4.61	3.71
n-$C_{14}H_{29}$	7.00	4.91	4.61
n-$C_{16}H_{33}$	8.00	5.21	4.31
n-$C_{18}H_{37}$	9.00	2.81	...
n-$C_{20}H_{41}$	10.00	2.81	...

[a] Estimated from values of other closely related congeners. [b] From A. Albert and E. P. Serjeant, "Ionization Constants of Acids and Bases," John Wiley and Sons, Inc., New York, N. Y., 1962, p 144. [c] From E. A. Braude and F. C. Nachod, "Determination of Organic Structures by Physical Methods," Vol. 1, Academic Press Inc., New York, N. Y., 1955, p 596. [d] From G. Harris, "Dictionary of Organic Compounds," 4th ed, Oxford University Press, New York, N. Y., 1965. [e] From M. Yoshioka, K. Hamamoto, and T. Kubota, *Bull. Chem. Soc. Japan*, **35**, 1725 (1962). [f] The log P value of the un-ionized base was used.

perimentally determined values for 4- and 2-chlorophenol. For the 4-chloro-3-methyl congeners the value of 0.56 for the 3-methyl group in phenol was employed.[6a] Most of the values for eq 22 are experimental values.[6a] Log P values for eq 23 and 61 are based on experimental values for benzyl alcohols where possible.[6a] Where not possible, the π value for the function in the phenoxyacetic acid system was added to log P for benzyl alcohol. For eq 27, log P for the basic structure N-phenyl-N′-3,4-dichlorophenylurea (4.71) was measured experimentally and log P values for its derivatives were obtained by adding π values from the phenoxyacetic acid system except for phenyl (2.13) and $(CH_3)_2N$, NH_2, and SO_2NH_2 obtained from the benzene system. For eq 34 we have added π_{COOCH_3} (0.42) to log P of phenol to get log P of 1.88 for methyl 4-hydroxybenzoate. The value of 0.42 was calculated from eq 17 of Fujita, Iwasa, and Hansch.[6a] Log P values for the quinine derivatives of eq 47 were found by adding CH_2 units to the value of 2.03 for dihydroquinine. Log P for dihydroquinine was found by adding 0.3 (the difference between ethyl and vinyl) to 1.73 for quinine.

In estimating log P for the phenyl methacrylate of eq 24, 25, 41, 51, 56, 57, and 60 we have used π of -0.64 for the CH_3COO moiety.[6a] Subtracting 0.5 for the methyl group yields π of -1.14 for $-COO-$. To this we added 2.13 for benzene and 1.00 for $CH_2{=}C(CH_3)$ to get 1.99. For naphthol derivatives of eq 26 and 42 we started with the value of 2.84 for β-naphthol.

In some of the equations we have used π instead of log P. This constant is defined as: $\pi = \log P_X - \log P_H$. P_X is the partition coefficient of a derivative and P_H that of a parent molecule. For example, π_{CH_3} = log P toluene − log P benzene. Thus π is the logarithm of the partition coefficient of a molecular part, whereas log P refers to the whole molecule. Since π is known to be additive in nature, we can explore the lipophilic role of substituents in a series of drugs without actually measuring any partition coefficients. Equations of the form of 1 will have the same slope whether we use π or log P. They will differ only in intercept. When strong electron-withdrawing groups are placed on molecules with acidic protons, significant changes in ionization may result. In one way or another, these changes in ionization may affect the biological activity of the compound. Each set of data has been fit to eq 2 and then, by stepwise regression analysis, we have omitted first the $\rho\sigma$ term and then the $(\log P)^2$ term. Only those terms are included in the equations in Table II which are justified at >0.90 level of significance by an F test.

In several of the equations we have used pK_a instead of σ to account for electronic effects on activity. Since σ is defined as $\sigma = \log (K_X/K_H)$ where K_H is the ionization constant of benzoic acid and K_X that of a derivative, either pK_a or σ may be used as a measure of relative acidities of members of a set of congeners.

For substituents in *ortho* positions we have used σ_p, assuming this to give a rough approximation of the electronic effect. Fortunately, electronic effects are small for most of the compounds under consideration so that this is not a serious problem. If electronic effects are large enough to make big changes in ionization constants, then special corrections must be made.[34,35]

The σ values of 4-*n*-Hex-O- and higher homologs are assumed to be the same as the σ of 4-*n*-Am-O-. The σ value of 4-*sec*-Am-O- is obtained by adding -0.02 to the σ value of 4-*n*-Am-O- ($-0.02 - 0.34 = -0.36$) since the difference between the σ values of 4-*i*-Pr-O- and 4-*n*-Pr-O- is -0.02. The same approximation is used for the σ value of 4-*n*-Am- and higher members. The σ values of the substituents of alkylchlorophenols and alkyl sulfates are practically constant and can be neglected in the analysis.

Results

In Table II we have summarized the statistically most significant equations correlating the structure-activity relationship in gram-negative bacteria. In these equations, n is the number of data points used in the regression analysis, r is the correlation coefficient, and s is the standard deviation. The figures under log P_0 define the 90% confidence interval on this constant.

The correlations with the 24 different systems in Table II are, on the whole, quite satisfying. Of the 24, eleven have correlation coefficients above 0.95, nine have r between 0.95 and 0.83, and three have very poor correlations. Twelve of the equations are linear in nature. We assume this is because in these investigations an insufficient number of molecules with log P greater than 4 was studied and hence the apex of the parabola relating log PC' or log $(1/C)$ and log P could not be defined with any degree of statistical assurance.

The most interesting result from the equations of Table II are the eight cases where log P_0 could be established. These structures are summarized in Table III. The range for these eight values is 3.8-5.1 with a mean of 4.4. Omitting the highest value (5.1), we find a range of only 3.8-4.6 with a mean of 4.3. Unfortunately, all of the examples but one where log P_0 could be calculated were studies employing phenols. The one exception is that embodied in eq 16 for phenyl isothiocyanates. It is of special interest that for this set we find log P_0 = 4.10, very close to the mean value found for the phenols. This means that the ideal lipohydrophilic character required for maximum toxicity is the same for phenols and phenyl isothiocyanates and that the sites of action must be the same.

For the equations in Table II showing a linear dependence on log P (8-12, 17-23), we find a rather limited range of slopes.

Omitting equations 9, 11, 12, 17, and 18 we find a range of slopes of 0.54-0.77 for seven sets of alcohols, amines, and phenols. The mean value is 0.65. Considering the wide variety of compounds included and the fact that the investigations were carried out in several different laboratories using different gram-negative bacteria as test organisms, the similarity in slopes of the seven equations is remarkable. Equation 11 for alcohols has a slope of 1.02 which is considerably higher than the others. It was observed by the investigators[19] that the bacteria used in testing these alcohols was unusually sensitive and, in fact, there was considerable doubt as to its identity. The same problem of uncertainty applies to eq 9 and 12.

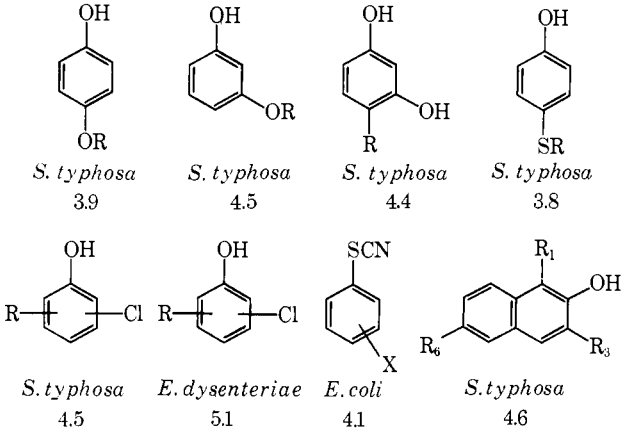
(34) T. Fujita, *J. Med. Chem.*, **9**, 797 (1966).
(35) T. Fujita and C. Hansch, *ibid.*, **10**, 991 (1967).

TABLE II
EQUATIONS DESCRIBING THE STRUCTURE-ACTIVITY RELATIONSHIP IN GRAM-NEGATIVE BACTERIA

Drug *vs.* *S. typhosa*	Equation	n	r	s	Log P_0 or π_0	Eq no.
Hydroquinone monoethers	Log $PC' = -0.280(\log P)^2 + 2.199 \log P + 1.219\sigma - 2.215$	11	0.972	0.169	3.93 (3.54-4.87)[a]	3
Resorcinol monoethers	Log $PC' = -0.180(\log P)^2 + 1.628 \log P - 1.777$	11	0.975	0.208	4.52 (3.84-6.63)	4
4-Alkylresorcinols	Log $PC' = -0.204(\log P)^2 + 1.771 \log P - 1.871$	10	0.982	0.180	4.35 (3.72-6.06)	5
p-Hydroxyphenyl alkyl sulfides	Log $PC' = -0.407(\log P)^2 + 3.082 \log P + 2.460\sigma - 3.649$	12	0.971	0.168	3.79 (3.62-4.03)	6
Alkylchlorophenols	Log $PC' = -0.334(\log P)^2 + 2.991 \log P - 4.540$	26	0.936	0.190	4.48 (4.33-4.70)	7
Alkylbromophenols	Log $PC' = 0.765 \log P - 0.998$	8	0.954	0.259	...	8
Aliphatic amines[b]	Log $PC' = 0.375 \log P - 0.151$	10	0.880	0.159	...	9
Arylamines	Log $(1/C) = 0.589 \log P + 3.949$	15	0.940	0.137	...	10
Alcohols[b] (prim-tert)	Log $PC' = 1.024 \log P - 1.536$	15	0.996	0.090	...	11
Alcohols (cyclohexanols, etc.)	Log $PC' = 0.614 \log P - 0.949$	8	0.826	0.142	...	12
Diamidines	Log $(1/C) = -0.115\pi^2 + 2.001\pi - 4.127$	8	0.989	0.152	8.73 (7.63-13.23)	13
Diguanidines	Log $(1/C) = -0.081\pi^2 + 1.483\pi - 1.578$	8	0.996	0.156	9.20 (8.26-11.12)	14
Drug *vs.* ***E. dysenteriae***						
Alkylchlorophenols	Log $PC' = -0.219(\log P)^2 + 2.251 \log P - 3.396$	19	0.937	0.241	5.14 (4.86-5.64)	15
Drug *vs.* *E. coli*						
Phenyl isothiocyanates	Log $(1/C) = -1.040(\log P)^2 + 8.531 \log P + 0.774\sigma - 12.629$	9	0.967	0.138	4.10 (4.01-4.22)	16
RNCS	Log $(1/C) = 0.367\pi + 3.582$	12	0.890	0.284	...	17
1-Aryl-2-nitropropenes	Log $(1/C) = 0.401\pi - 0.269$	9	0.825	0.212	...	18
Arylamines	Log $(1/C) = 0.694 \log P - 0.158\text{p}K_a + 4.462$	15	0.962	0.114	...	19
Drug *vs.* ***B. aerogenes***						
Arylamines	Log $(1/C) = 0.662 \log P - 0.136\text{p}K_a + 4.452$	15	0.948	0.130	...	20
Drug *vs.* ***B. dysenteriae F.***						
Arylamines	Log $(1/C) = 0.648 \log P - 0.119\text{p}K_a + 4.504$	15	0.961	0.110	...	21
Drug *vs.* *Ps. aeruginosa*						
Substituted phenols	Log $(1/C) = 0.684 \log P - 0.921\sigma + 0.268$	21	0.847	0.222	...	22
Drug *vs.* *P. vulgaris*,[c] *E. coli*, and *Ps. pyocyanea*						
Benzyl alcohols	Log $(1/C) = 0.539 \log P + 0.531\sigma + 4.001$	14	0.939	0.212	...	23
Drug *vs.* *K. pneumoniae*						
Phenyl methacrylates	Log $(1/C) = 0.009 \log P + 3.093$	6	0.068	0.124	...	24
Phenyl methacrylates	Log $(1/C) = 0.034(\log P)^2 - 0.286 \log P + 0.113\sigma + 3.606$	6	0.190	0.173	...	25
Drug *vs.* *S. typhosa*						
Alkyl-β-naphthols	Log $(1/C) = -0.226(\log P)^2 + 2.088 \log P - 1.126$	19	0.479	0.438	4.62 (3.93-5.28)	26

[a] 90% confidence interval. [b] Identity of the organism was doubtful; see text. [c] In this example three different microorganisms were used simultaneously.

TABLE III
SUMMARY OF PARENT STRUCTURES AND LOG P_0 VALUES FOR GRAM-NEGATIVE BACTERIA



Equations 17 and 18 have slopes of 0.37 and 0.40, respectively. The different slope for the RNCS derivatives of eq 17 points to a different mode of action for these compounds. The low slope with the arylnitropropenes (eq 18) may reflect the fact that this set of congeners has log P values rather near log P_0. We did not place this set on a log P basis since we did not have log P for the parent compound. Log P for the parent compound would be near 3. Different susceptibilities among different species to a group of congeners are clearly illustrated by the study of phenyl methacrylates. In this case no correlation with log P and σ could be made for the gram-negative bacteria *Klebsiella pneumoniae* (eq 24, 25). Inspection of the data in Table I shows almost no difference in susceptibility of this organism to compounds having *greatly* different log P and σ values. The unusual resistance of this organism to structural variations to which other bacteria respond

TABLE IV

EQUATIONS DESCRIBING THE STRUCTURE-ACTIVITY RELATIONSHIP IN GRAM-POSITIVE BACTERIA

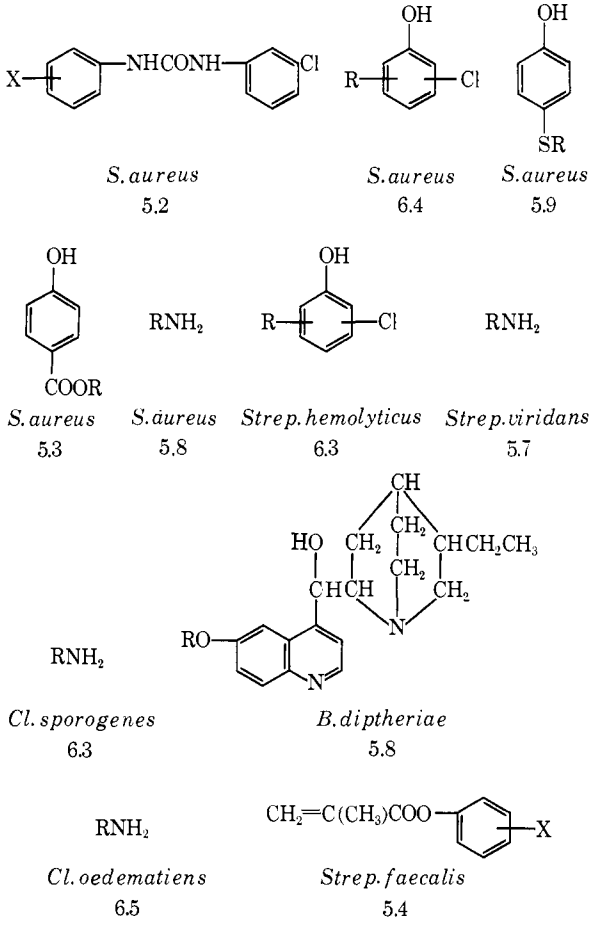
Drug *vs. S. aureus*	Equation	n	r	s	Log P_0 or π_0	Eq no.
Substituted ureas	$\log (1/C) = -0.335(\log P)^2 + 3.453 \log P + 2.995\sigma - 4.200$	12	0.899	0.770	5.15 (4.49-11.95)[a]	27
Alkylchlorophenols	$\log PC' = -0.167(\log P)^2 + 2.121 \log P - 3.498$	35	0.961	0.236	6.36 (5.98-6.94)	28
p-Hydroxyphenyl alkyl sulfides	$\log PC' = -0.147(\log P)^2 + 1.733 \log P - 2.211$	12	0.995	0.093	5.90 (5.15-7.46)	29
Hydroquinone monoethers	$\log PC' = 0.823 \log P - 1.020$	13	0.982	0.196	...	30
Resorcinol monoethers	$\log PC' = 0.871 \log P - 1.164$	11	0.994	0.115	...	31
Alkylbromophenols	$\log PC' = 0.847 \log P - 1.258$	13	0.991	0.126	...	32
4-Alkylresorcinols	$\log PC' = 0.912 \log P - 1.108$	8	0.952	0.409	...	33
Esters of *p*-hydroxybenzoic acid	$\log PC' = -0.167(\log P)^2 + 1.784 \log P - 2.201$	8	0.996	0.066	5.34 (4.54-7.37)	34
Alcohols (prim-tert)	$\log (1/C) = 0.671 \log P + 0.069$	9	0.964	0.112	...	35
Alcohols (prim-tert)	$\log PC' = 0.888 \log P - 1.543$	10	0.988	0.089	...	36
Diamidines	$\log (1/C) = -0.165\pi^2 + 2.500\pi - 4.680$	8	0.997	0.073	7.60 (7.30-8.05)	37
Diguanidines	$\log (1/C) = -0.112\pi^2 + 1.736\pi - 1.363$	8	0.979	0.296	7.75 (7.00-9.71)	38
Aliphatic amines	$\log (1/C) = -0.264(\log P)^2 + 3.081 \log P - 4.416$	5	0.991	0.131	5.84 (5.63-6.10)	39
Aliphatic amines	$\log PC' = 0.834 \log P - 1.574pK_a + 15.590$	6	0.944	0.229	...	40
Phenyl methacrylates	$\log (1/C) = 0.668 \log P + 1.342$	10	0.966	0.262	...	41
Alkyl-β-naphthols	$\log (1/C) = 0.626 \log P + 1.316$	22	0.898	0.347	...	42
N[1]-Alkylnikethamide chlorides	$\log (1/C) = -0.060\pi^2 + 0.909\pi + 2.920$	20	0.961	0.291	7.63 (7.06-8.53)	43
β-Nitrostyrenes	$\log (1/C) = 0.489\pi + 0.570$	12	0.885	0.167	...	44
1-Aryl-2-nitropropenes	$\log (1/C) = 0.746\pi + 1.384$	12	0.976	0.114	...	45
RNCS	$\log (1/C) = 0.516\pi + 3.330$	13	0.947	0.258	...	46
$ROSO_3^-Na^+$	$\log (1/C) = 0.694\pi - 1.365$	10	0.976	0.325	...	47
Drug *vs. Strep. hemolyticus*						
Alkylchlorophenols	$\log PC' = -0.171(\log P)^2 + 2.146 \log P - 3.576$	33	0.956	0.251	6.29 (5.78-7.22)	48
Drug *vs. Strep. viridans*						
Diguanidines	$\log (1/C) = -0.068\pi^2 + 1.387\pi - 0.848$	8	0.984	0.314	10.21 (8.08-31.76)	49
Aliphatic amines	$\log (1/C) = -0.247(\log P)^2 + 2.815 \log P - 2.301$	5	0.994	0.094	5.69 (5.54-5.87)	50
Drug *vs. Strep. faecalis*						
Phenyl methacrylates	$\log (1/C) = -0.125(\log P)^2 + 1.359 \log P + 0.415$	10	0.861	0.334	5.42 (4.64-109.90)	51
Drug *vs. D. pneumoniae*						
$RCHOHCOO^-K^+$	$\log (1/C) = -0.194\pi^2 + 2.903\pi - 6.990$	5	0.990	0.201	7.47 (6.72-13.68)	52
$RCHBrCOO^-K^+$	$\log (1/C) = -0.199\pi^2 + 2.672\pi - 4.264$	8	0.893	0.596	6.73 (6.25-7.34)	53
Drug *vs. B. diphtheriae*						
Hydrocupreines[b]	$\log (1/C) = -0.123(\log P)^2 + 1.431 \log P + 1.161$	17	0.936	0.300	5.81 (5.55-6.12)	54
$RCHBrCOO^-K^+$	$\log (1/C) = 0.550\pi + 0.283$	6	0.961	0.330	...	55
Drug *vs. B. subtilis*						
Phenyl methacrylates	$\log (1/C) = 0.617 \log P + 1.530$	10	0.976	0.204	...	56
Drug *vs. B. cereus*						
Phenyl methacrylates	$\log (1/C) = 0.400 \log P + 2.144$	10	0.815	0.420	...	57
Drug *vs. Cl. oedematiens*						
Aliphatic amines	$\log (1/C) = -0.159(\log P)^2 + 2.072 \log P - 1.529$	5	0.981	0.185	6.50 (5.94-8.87)	58
Drug *vs. Cl. sporogenes*						
Aliphatic amines	$\log (1/C) = -0.189(\log P)^2 + 2.373 \log P - 2.631$	5	0.985	0.164	6.27 (5.87-7.22)	59
Drug *vs. Sarcina lutea*						
Phenyl methacrylates	$\log (1/C) = 0.161 \log P + 2.721$	10	0.849	0.148	...	60
Drug *vs. S. aureus*[c] *S. albus* and *Strep. faecalis*						
Benzyl alcohols	$\log (1/C) = 0.599 \log P + 0.421\sigma + 4.069$	18	0.906	0.307	...	61

[a] 90% confidence interval. [b] The log P values of the free alkaloids were used. [c] In this case three different microorganisms were used simultaneously.

in a predictable manner is most interesting and its protecting structural features merit careful study. The alkyl-β-naphthols also gave a very poor correlation with the gram-negative bacteria *Salmonella typhosa* (eq 26), although a typically good correlation for this set of drugs in gram-positive bacteria (*Staphylococcus aureus*) was found (eq 42). The poor correlation with gram-negative cells reflects their more complex structure.[36]

In Table IV we find that where it can be defined, log P_0 for gram-positive bacteria is much higher than for gram-negative organisms. For eleven examples, summarized in Table V, we find a range of 5.2–6.5 with a mean of 5.9 for a heterogeneous group of ureas, phenols, esters, amines, and quinine derivatives. Thus the ideal partition coefficient for antibacterial agents for gram-positive organisms is much higher than for gram-negative organisms.

TABLE V
SUMMARY OF PARENT STRUCTURES AND LOG P_0 VALUES FOR GRAM-POSITIVE BACTERIA



The difference in log P_0 for gram-positive and gram-negative bacteria (6 *vs.* 4) indicates that micelle formation[37] cannot account for the loss of biological activity in the upper part of a homologous series since, if it were the reason, the log P_0 would depend upon the type of compound regardless of the organism and this is not the case.

(36) J. Bracket and A. E. Mirsky, "The Cell," Vol. II, Academic Press Inc., New York, N. Y., 1960, p 121.

(37) A. Albert, "Selective Toxicity," 3rd ed, John Wiley and Sons, Inc., New York, N. Y., 1965, p 170.

For the 17 equations with linear dependence on log P or π, with $r > 0.90$, we find slopes of 0.52–0.91 with a mean of 0.73. This is not far from the mean of 0.65 found for the gram-negative bacteria. In fact, for comparative purposes, both values should probably be rounded off to 0.7.

These findings indicate that the toxic action, when electronic effects can be separated or held constant, is due to the relative lipophilic character of the drugs. Since data are limited and since part of the work was reported in terms of PC' and part in terms of $1/C$, we cannot make any useful comparisons of intrinsic activity of the different sets of congeners by comparing intercepts. We are only able to compare Δ log BR with Δ log P or $\Delta\pi$.

The fact that most of the equations in Table IV are linear with respect to log P is explained by the high log P_0 found for gram-positive bacteria. In none of the systems described by eq 30–33, 35, 36, 40, and 61 where log P was used were data points for log P as high as 6 available. We have not attempted to estimate log P for the ions used in eq 13, 14, 37, 38, 43, 47, 49, 52, 53, and 55; therefore we used π values.

For eq 54, the log P values are for the free base rather than the salt. The base strength of all of the amines in this series will be constant and so the percentage of free base present in each case will be the same. It seems most likely that it is the free base that is the active species in this example; however, insufficient data are at hand to be certain of this point.[34]

Discussion

Considering first the linear equations in Tables II and IV, it is instructive to compare the mean slope of 0.65 for gram-negative bacteria and the slope of 0.73 for gram-positive bacteria with the slope in eq 62. Equation 62 correlates the binding of phenols to bovine serum albumin.[9] In eq 62, C stands for the molar

$$\log (1/C) = 0.681 \log P + 2.489 \qquad (62)$$
$$n = 19 \quad r = 0.962 \quad s = 0.133$$

concentration of phenol producing a 1:1 phenol-protein complex *via* equilibrium dialysis. The dependence of antibacterial action on lipophilic character very closely parallels the dependence of protein binding on lipophilic character. This of course explains why phenols and long-chain amines are inactive or much less active in the presence of serum.[20,28]

The relatively nonspecific nature of the toxic action indicated by the equations in Tables II and IV is apparent from the fact that a variety of different sets of phenols, alkyl-β-naphthols, phenyl methacrylates, amines, alkyl sulfates, and alcohols give good linear correlations between log BR and log P with slopes near 0.7. It is interesting to compare this type of toxic action with that for a variety of compounds inhibiting oxidative metabolic processes. As mentioned above, for 15 different sets of drugs acting in different biochemical systems (whole animals, isolated tissue, bacteria, etc.) we found a linear relationship between log BR and log P; however, in these examples where inhibition of oxidative metabolism appeared to be the critical reaction, the mean slope was found to be 1. Thus the

slope associated with log P can be used to characterize the biochemical process.

The lower log P_0 of about 4 for gram-negative bacteria may be attributed to the higher lipid content of the cell wall (up to 25% dry weight) compared to that of the gram-positive species (0–2.6%).[38,39] There is some evidence in *Escherichia coli* of a lipoprotein membrane on both sides of the cell wall.[40] Recently it has been shown that when three species of gram-positive microorganisms were grown under conditions in which their cellular lipid content was increased, a corresponding increase in their resistance to penicillins was produced. Cell-wall lipid depletion increased their sensitivity.[41]

Before a molecule can reach the cytoplasmic membrane or the interior of the cell, it must cross the cell wall. Here it will be more or less tightly bound depending on the nature of the wall and its own chemical constitution. If the cell wall is rich in lipid, as in the case of gram-negative cells, the adsorption of highly lipophilic molecules would be very strong. As one increases the lipophilic character of a given function in the kind of activity considered above, biological response tends to follow in a linear fashion up to a point. This is the point where binding of the drugs by the first lipophilic material with which they come into contact is so strong that the random walk by which these drugs reach their sites of action becomes quite strongly time dependent. This departure from linearity is probably exaggerated by the popular method of characterizing biological activity in terms of $1/C$ or PC'. As one goes to lower and lower concentrations to obtain the equivalent biological response with the more active, more lipophilic members, one reaches very low concentrations of the highly lipophilic drugs. Loss of a small amount of material to very lipophilic binding sites results in an increasingly large percentage loss of drug.

The gram-negative organisms have a protective layer of lipid which protects them from lipophilic compounds as well as very hydrophilic compounds. The evidence seems strong that the difference between the susceptibility of gram-negative compared to gram-positive bacteria to the more hydrophobic anionic and cationic detergents, higher alkyl sulfates, amines, phenols, chloroforms, ethers, esters, penicillins, etc.,[42] is due to the lipid content of the cell wall.

The appearance of a σ or pK_a term in 10 out of the 58 equations indicates that the electronic effect of the substituent does play a significant role. The positive coefficient with σ (except eq 22 where the correlation is not as good as others) indicates that electron withdrawal promotes activity. Part of this effect may simply be to make the molecules more lipophilic.[6a] Electron withdrawal also increases the hydrogen bonding power of acidic hydrogens as well as their degree of ionization. Not enough information is present to enable us to sort out the primary role of the electronic effect of substituents.

It is noteworthy that 3,5,3′,4′-tetrachlorosalicylanilide (TCS), a substituted phenol with a calculated log P of greater than 6, localizes on the cytoplasmic membrane of bacteria and causes leakage of cell contents inhibiting the accumulation of nutrients from the medium.[43] It has also been reported that bacteriostasis results from the action of TCS either on the energy-producing systems of the cell or on a mechanism coupling this energy to energy-requiring processes[44] which is to be expected from our previous results.[5] Unfortunately, we do not have a value of PC' for TCS so that we can make direct comparison of it with the other phenols. It seems likely that the mechanism of action would be the same.

There seems to be a basic difference in the mechanism of action of aliphatic and aromatic isothiocyanates. The aromatic ones of eq 16 yield results comparable to the phenols, but the aliphatic compounds of eq 17 and 46 show a low dependence on lipophilic character, especially in eq 17. Inspection of the $1/C$ values for those derivatives not having a benzene ring shows a small degree of variance in relative activity. Two possible reasons for the much lower dependence on log P are apparent. It might be that the sites of action are located so that movement through lipophilic material to reach them is not necessary (*e.g.*, in the cell wall), or it might be that the mechanism of action at the site is not influenced by the lipophilic portion of the drug. At present it is not possible to decide between the two.

Although the above equations with their attendant log P_0 values enable us to pull together a massive amount of miscellaneous antibacterial structure-activity study so that a relatively coherent view is possible, more uniform work should permit more detailed analysis. Since we have no idea what level of precision the various research groups were striving for in collecting the data, we are not sure just how precisely the slopes associated with log P and the log P_0 values can be defined. For example, it is tempting to think that very careful testing under uniform conditions might indicate a single log P_0 (or very narrow range) for Table III and another for Table V. On the other hand, the difference in log P_0 for each of the sets in Tables III and V may be quite real and characteristic of certain cellular structural features. The results so far obtained indicate that time spent in very careful testing could pay off by revealing through regression analysis small but significant differences in the mechanism of action which, when fully appreciated, could be more consciously exploited in drug design.

One must not conclude that all antibacterial agents will have a log P_0 near 4 or 6. The value of log P_0 is quite dependent on the total test system *as well as* the molecular mechanism of action. The mechanism of action of the sulfonamide drugs is quite different from the molecules considered in this report and their log P_0 values are also quite different.[35]

For very long chain aliphatic molecules there is some doubt about the strict additivity of 0.50 for each CH_2 unit. Intramolecular hydrophobic bonding[6b] could lower the value of 0.50. How serious this problem is

(38) I. G. Gunsalus and R. Y. Stanier, "The Bacteria," Vol. I, Academic Press Inc., New York, N. Y., 1960, p 121.
(39) M. R. J. Salton, "The Bacterial Cell Wall," Elsevier Publishing Co., Amsterdam, 1964.
(40) P. H. Clarke and M. D. Lilly, *Nature*, **195**, 516 (1962).
(41) W. B. Hugo and R. J. Stretton, *J. Gen. Microbiol.*, **42**, 133 (1966).
(42) J. W. Bartholomew and T. Mittwer, *Bacteriol. Rev.*, **16**, 1 (1952).
(43) (a) R. C. S. Woodroffe and B. E. Wilkinson, *J. Gen. Microbiol.*, **44**, 343 (1966); (b) R. C. S. Woodroffe and B. E. Wilkinson, *ibid.*, **44**, 353 (1966).
(44) W. A. Hamilton, *Biochem. J.*, **103**, 73P (1967).

is difficult to estimate since it would only be apparent in the correlations or in the determination of log P values. The rather good agreement between log P_0 and π_0 for both aliphatic and the more inflexible aromatic compounds does not reveal any discontinuity. From some preliminary work measuring partition coefficients, it would appear that at least for some systems π for each CH_2 unit is constant up to at least 10 carbon atoms. Of course this holds only as long as no electronic or dipolar interactions promote intramolecular hydrophobic bonding.[6b] The extreme difficulty in measuring partition coefficients of apolar groups larger than this leaves some uncertainty about the very large aliphatic compounds in Table I. This presents no problem for the results with gram-negative bacteria shown in Table III. While there are several instances where folding could occur with long chains of the molecules on which the data of Table V are based, comparison of the log P_0 for the rigid phenols with the flexible aliphatic amines does not reveal a significant difference in log P_0. For the six more rigid structures we find a mean log P_0 of 5.8 and, for the five flexible examples (including the quinine derivatives), we find a mean log P_0 of 6.0.

As mentioned above, it must be strictly borne in mind that the log P values we have used are for the neutral un-ionized form of the molecules. This poses no problem for the compounds of Table III; however, for the molecules of Table V we are comparing quite basic amines, of which only a very small fraction would be in the neutral form under test conditions, with relatively un-ionized phenols. The fact that we find the same log P_0 for these amines as we do for the phenols and ureas would indicate that the un-ionized form is more suitable to consider in correlation studies. The partition coefficient of the ionized molecule would be greatly different from that of the un-ionized form. Exactly why one finds very similar log P_0 values for highly ionized and un-ionized molecules as well as rather rigid aromatic and flexible aliphatic compounds is not apparent and suggests an important area for further study.

In summary, one can say that octanol-water partition coefficients constitute a very useful reference system for comparative biochemical and pharmacological studies where hydrophobic bonding is involved. Log P_0 also appears to be a useful constant for the study of the movement of organic compounds through biophases.

Acknowledgment.—C. Hansch wishes to express his gratitude to the Guggenheim Foundation for a supporting fellowship, and to Smith Kline and French for a research associateship for Susan Anderson.

Molecular Orbital Calculations of the Preferred Conformations of Histamine and a Theory on Its Dual Activity

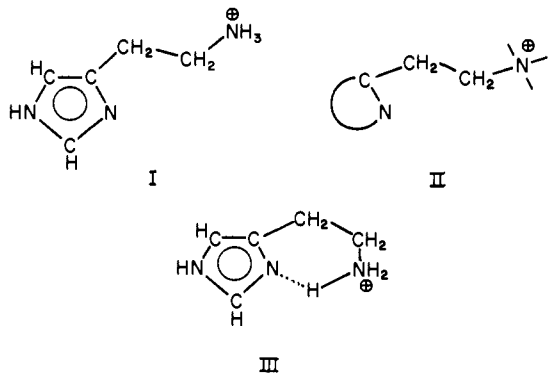
LEMONT B. KIER

Battelle Memorial Institute, Columbus Laboratories, Biochemistry Division, Columbus, Ohio

Received January 12, 1968

Extended Hückel molecular orbital calculations on the histamine molecule reveal two conformations of nearly equal preference, on the basis on calculated minimum energy. Neither conformation involves intramolecular hydrogen bonding. Population analyses reveal the charge-density pattern of the imidazole ring. The dual activity of histamine is proposed to be a consequence of the existence of two preferred conformations in equilibrium. One of these conformations places the quaternary nitrogen and the ($tr^2trtr\pi$) nitrogen of histamine 4.55 Å apart, which is quite comparable to the 4.8 Å estimated for the internitrogen distance in the antihistaminic triprolidine. An assignment of each histamine conformation to one of two histamine effects is provisionally made on this basis. This explanation of dual activity is comparable with that offered for a similar situation found in previous calculations on acetylcholine, muscarine, and nicotine.

Histamine (I) is known to produce a series of well-characterized biological responses when it is released from storage cells by the influence of trauma or chemical agents. A number of other molecules are known to produce these responses, but histamine is the most active compound known and remains the prototype of histaminic activity. It is evident that the histamine molecule must present near-optimal electronic features to its receptor. To date, several studies have been directed toward elucidating the features of the molecule that are necessary to elicit biological activity. Lee and Jones[1] have suggested that an important structural feature is the fragment II, in which the ring is a small aromatic nucleus. Neimann and Hays[2] have suggested that the univalent cation (the predominant form at body pH) will exist in a hydrogen-bonded form, III. These authors felt that the ability to form this hydrogen bond is a necessary condition for histaminic activity. Lee and Jones,[1] however, observed that, although all of the active compounds they studied were



I II III

(1) H. M. Lee and R. G. Jones, *J. Pharmacol. Exptl. Therap.*, **95**, 71 (1949).

(2) C. C. Nieimann and J. T. Hays, *J. Am. Chem. Soc.*, **64**, 2288 (1942).