J. Pharm. Pharmacol. 1982, 34: 158–161
Received July 7, 1981

0022–3573/82/030158–04 $02.50/0

A comparison of log P and molecular connectivity in the structure-activity analysis of some antimicrobial agents

J. C. BOYD, J. S. MILLERSHIP AND A. D. WOOLFSON*

Department of Pharmacy, The Queen's University of Belfast, Medical Biology Centre, 97 Lisburn Road, Belfast BT9 7BL, U.K.

Several congeneric series of antimicrobial agents previously studied by Hansch analysis have been investigated using molecular connectivity as the descriptor of molecular structure. In all cases connectivity gave comparable or improved correlations compared with log P, particularly where the pattern of molecular substitution was more complex. It was concluded that the use of computer-generated connectivity terms had advantages over calculated log P values in its ease of application and should be considered at least for initial screening of structure-activity data.

The use of the linear free energy-related (LFER) model of Hansch & Fujita (1964) is well known in the development of quantitative structure-activity relationships (QSAR). In this model biological activity is related to one or more physicochemical parameters which generally include the logarithm of the octanol-water partition coefficient of the molecule (log P). A review of some applications of Hansch analysis to the study of congeneric series of biologically active molecules has been made by Tute (1971). Clearly, good quantitative biological data is of first importance. The more successful QSAR studies have therefore tended to be in such areas as antimicrobial activity, where reliable data can be readily obtained.

An alternative to the representation of molecular structure by physicochemical parameters is the use of a topologically derived index which directly represents chemical structure as a mathematical quantity. Such an index is molecular connectivity (Kier et al 1975), a development of the hydrocarbon branching index of Randic (1975). The various forms of the connectivity indices ($^{m}\chi_{t}$), where m and t refer respectively to order and type of molecular connectivity index, have been described by Kier & Hall (1976).

Molecular connectivity has been successfully correlated with a number of physicochemical properties including partition coefficient (Murray et al 1975), solubility and boiling point data (Hall et al 1975) and gas chromatographic retention data (Millership & Woolfson 1978, 1980; Kier & Hall 1976). Kier & Hall (1976) have described in their text on molecular connectivity some applications in the biological field.

* Correspondence.

In the present study connectivity has been applied to the correlation of biological data with chemical structure. The LFER equations obtained by Hansch analysis for various congeneric series of antimicrobial agents have been compared with those obtained using molecular connectivity as the sole descriptor of molecular structure.

METHOD

Calculations of molecular connectivity follow the method described by Kier & Hall (1976). The use of the computer program CFUNC* allowed calculation of all connectivity indices up to and including the sixth order. The set of indices thus generated was then used in a stepwise multiple regression procedure.

All biological data is expressed as log 1/c, where c is the concentration of compound necessary to elicit a fixed response in a standard test.

RESULTS AND DISCUSSION

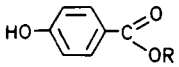
Hansch analysis of structure-activity data often involves the use of log P as a physicochemical parameter representative of molecular structure. In many cases log P values have been calculated by assuming additivity of substituent π values. Such calculated log P values often differ markedly from experimentally determined values due to steric factors and the effects of inter- and intramolecular hydrogen bonding. However, the experimental determination of log P can be a rigorous and time-consuming procedure (Purcell et al 1973). The use of a readily calculable alternative to log P may

* Supplied by Professor L. G. Hall, Eastern Nazarene College, Massachusetts.

therefore be advantageous. Molecular connectivity is one possible alternative which has been shown to correlate well with many physicochemical parameters such as log P (Murray et al 1975).

Various congeneric series of antimicrobial agents have been widely studied by Hansch analysis. In the simplest cases high quality equations have been obtained with log P. Hansch & Lien (1971) surveyed structure-activity relationships in a wide variety of antifungal agents. A one parameter equation in log P was obtained for the data (Table 1a) on seven *p*-hydroxy benzoates (eqn 1), a rather small group for QSAR work, whereas the sodium salts of fifteen

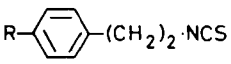
Table 1. Structures of compounds in the study and their Δ log 1/c values calculated from the appropriate connectivity equation.

1a 

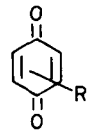
R	Log $1/c_{obs}$	Δ Log 1/c (from eqn 3)
CH_3	1·80	−0·049
C_2H_5	0·40	+0·126
C_3H_7	3·00	+0·080
C_4H_9	3·00	−0·048
C_5H_{11}	3·20	−0·191
C_6H_{13}	3·60	−0·110
C_7H_5	4·22	+0·190

1b R-COO⁻

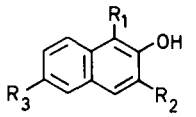
R	Log $1/c_{obs}$	Δ Log 1/c (from eqn 4)
$C_{13}H_{27}$	4·36	−0·062
$C_{13}H_{25}$	4·36	−0·062
$C_{12}H_{25}$	4·16	−0·052
$C_{11}H_{23}$	4·00	−0·001
$C_{11}H_{21}$	4·21	+0·119
$C_{10}H_{21}$	3·79	−0·001
$C_{10}H_{19}$	3·96	+0·170
C_9H_{19}	3·54	−0·040
C_8H_{17}	3·35	−0·019
C_8H_{15}	3·49	+0·121
C_7H_{15}	3·20	+0·041
C_6H_{13}	2·82	−0·179
C_5H_{11}	2·11	−0·072
C_5H_9	2·15	−0·032
C_2H_5	1·39	+0·070

1c 

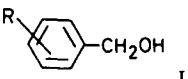
R	Log $1/c_{obs}$	Δ Log 1/c (from eqn 6)
H	4·32	−0·070
Cl	4·74	+0·094
I	4·85	−0·030
Me	4·60	−0·024
MeO	4·28	+0·008
NO_2	4·41	+0·002

1d 

R	Log $1/c_{obs}$	Δ Log 1/c (from eqn 8)
H	4·19	−0·14
2-Me	4·00	−0·40
2,5-Me_2	5·00	+0·55
3,6-$(OH)_2$-2,5-Cl_2	2·96	−0·64
2,5-Cl_2	5·52	+0·69
2,6-Cl_2	5·00	−0·24
2,3,5,6-Cl_4	5·00	−0·27
5,6-(C_4H_4)	5·10	+0·70
2-Me-5,6-(C_4H_4)	5·10	−0·21
2-3-Cl_2-5,6-(C_4H_4)	7·00	−0·04

1e 

R_1	R_2	R_3	Log $1/c_{obs}$	Δ Log 1/c (from eqn 10)
H	H	H	2·86	+0·045
H	H	i-Am	4·33	−0·285
H	H	i-Hex	4·66	−0·206
H	H	i-Hep	4·68	+0·081
H	Me	H	3·50	+0·185
H	Et	H	3·94	+0·128
H	Pr	H	4·27	+0·284
H	Bu	H	5·00	+0·415
Me	H	H	3·20	+0·138
Et	H	H	3·24	−0·121
Pr	H	H	3·27	−0·636
Me	H	Me	3·94	−0·104
Me	H	Et	4·27	−0·193
Et	H	Et	4·30	−0·283
Et	H	Pr	5·03	+0·266
Et	H	Bu	5·06	−0·025
Et	H	Am	5·38	−0·084
Et	H	Hex	5·19	−0·044
Me	H	Pr	4·60	−0·044
Me	H	Bu	5·03	+0·114
Me	H	Am	5·06	−0·116
Me	H	Hex	5·38	+0·266

1f 

R	Log $1/c_{obs}$	Δ Log 1/c (from eqn 12)
H	4·51	+0·17
4,Cl	4·76	−0·24
2,4-Cl_2	5·55	+0·04
3,4-Cl_2	5·85	+0·19
2,4,5-Cl_3	6·32	−0·20
3,4,5-Cl_3	6·63	+0·19
2-Br	5·15	−0·02
4-Br	5·57	+0·31
4-I	5·75	−0·01
4-Me	4·79	−0·14
2,4-Me_2	5·13	−0·09
4-Cl-3,5-Me_2	6·05	−0·06
4-I-3,5-Me_2	6·72	−0·05
4-MeO	4·84	−0·12
4-NO_2	5·00	+0·09
4-CN	4·67	−0·19
2-NO_2	5·49	+0·24
4-COOH	4·88	−0·12

aliphatic carboxylic acids (Table 1b) were considered to exhibit a parabolic dependency on log P (eqn 2).

$$\log 1/c = 0{\cdot}704 \log P + 0{\cdot}954 \quad (1)$$
$$r = 0{\cdot}971 \; s = 0{\cdot}205 \; n = 7$$

$$\log 1/c = 0{\cdot}059 (\log P)^2 + 0{\cdot}460 \log P + 3{\cdot}754 \quad (2)$$
$$R = 0{\cdot}985 \; s = 0{\cdot}170 \; n = 15$$

Both sets of data were analysed using molecular connectivity. Equations 3 and 4 respectively gave the best fit.

$$\log 1/c = 2{\cdot}553^5\chi_p - 2{\cdot}308 \quad (3)$$
$$r = 0{\cdot}985 \; s = 0{\cdot}150 \; n = 7$$

$$\log 1/c = 7{\cdot}062^5\chi_p - 7{\cdot}605^6\chi_p + 1{\cdot}320 \quad (4)$$
$$R = 0{\cdot}995 \; s = 0{\cdot}098 \; n = 15$$

In these simple cases connectivity gave a marginally improved fit to the data. The low coefficient of $(\log P)^2$ in equation 2 and the correspondingly good linear correlation provided by the connectivity equation 4 casts some doubt on the conclusion that the data in this set was parabolically dependent on log P over the range studied. Since in both data sets compounds only vary in the length of alkyl sidechain present the high quality correlations obtained are not surprising.

Two further sets of antifungal congeners studied by Hansch & Lien (1971) showed wider variation in both type and position of ring substituents. A set of six antifungal benzyl isothiocyanates with a single substituent para to the isothiocyanate group (Table 1c) gave almost identical good quality correlations with log P (eqn 5) and two higher order connectivity terms (eqn 6).

$$\log 1/c = 0{\cdot}462 \log P + 2{\cdot}789 \quad (5)$$
$$r = 0{\cdot}974 \; s = 0{\cdot}060 \; n = 6$$

$$\log 1/c = 2{\cdot}910^3\chi_c^v - 2{\cdot}280^4\chi_p^v + 6{\cdot}201 \quad (6)$$
$$R = 0{\cdot}970 \; s = 0{\cdot}073 \; n = 6$$

A set of ten antifungal quinones (Table 1d) showed much poorer correlation with log P (eqn 7).

$$\log 1/c = 0{\cdot}877 \log P + 3{\cdot}530 \quad (7)$$
$$r = 0{\cdot}859 \; s = 0{\cdot}579 \; n = 10$$

Here, the substituents varied in position on the ring, several being in the two position adjacent to the functional group. The relatively poor correlation probably reflects steric factors not accounted for by the calculated log P values. The correlation was enhanced using connectivity and the standard error of the regression(s) reduced by about 70% (eqn 8). The inclusion of $^4\chi_{pc}$ in equation 8 might be expected following the comments of Kier (1980) on the importance of this term in describing substitution in benzene-like rings.

$$\log 1/c = 16{\cdot}740^6\chi_p^v - 2{\cdot}342^4\chi_{pc} + 6{\cdot}910 \quad (8)$$
$$R = 0{\cdot}894 \; s = 0{\cdot}542 \; n = 10$$

The antibacterial activity against gram-negative organisms of two sets of compounds studied by Hansch et al (1967) were re-analysed using molecular connectivity. Twenty-two alkyl β-naphthols with three substituent positions (Table 1e) yielded equation 9 with log P.

$$\log 1/c = 0{\cdot}626 \log P - 1{\cdot}320 \quad (9)$$
$$r = 0{\cdot}898 \; s = 0{\cdot}347 \; n = 22$$

This association was improved considerably in the two parameter connectivity equation 10, with the standard error reduced by about 28%.

$$\text{Log } 1/c = 3{\cdot}217^6\chi_p - 1{\cdot}120 \, ^4\chi_p^v - 0{\cdot}490 \quad (10)$$
$$R = 0{\cdot}951 \; s = 0{\cdot}250 \; n = 22$$

Again, this improvement probably represents the influence on log P calculations of the steric effects of substituents R_1 and R_2 in proximity to the ring hydroxyl group.

A further series of eighteen benzyl alcohols (Table 1f) also active against gram negative bacteria were analysed by Hansch et al (1967). Five members of the series possessed *ortho* substituents. A reasonable correlation (r = 0·906) was obtained using a combination of log P and σ(eqn 11).

$$\log 1/c = 0{\cdot}599 \log P + 0{\cdot}421\sigma \quad (11)$$
$$R = 0{\cdot}906 \; s = 0{\cdot}307 \; n = 18$$

The best connectivity equation (12) involved three higher order terms with r rising to 0·971 and the standard error reduced by some 40%.

$$\log 1/c = 1{\cdot}733^5\chi_p + 2{\cdot}037^4\chi_p^v - 0{\cdot}706^4\chi_p + 2{\cdot}363 \quad (12)$$
$$R = 0{\cdot}971 \; s = 0{\cdot}181 \; n = 18$$

The fact that the σ term, highly significant in equation 11, is unnecessary in 12 suggests that connectivity may encode some measure of the electronic effects of substituents, probably via the valence connectivity term $^4\chi_p^v$ which takes into account the presence of hetero-atoms and the aromatic nature of the ring.

The comparison of connectivity and log P in the structure-activity analysis of antimicrobial agents indicates that the use of higher order connectivity terms can often improve on conventional correlations with physicochemical parameters. Since connectivity correlates well with many of these it may be a more fundamental representation of structure in structure-activity studies than log P (the use of which really denotes a property-activity relationship), although lacking the more easily interpreted physical

significance of the latter. Although improved methods for calculating log P are now available (Hansch & Leo 1979; Chou & Jurs 1980) the connectivity terms themselves are absolute and as such possess no inherent error. Perhaps the main advantage of connectivity lies in the fact that the various terms are computer-calculated and thus conveniently available for subsequent multiple regression analysis. Since no experimental determinations of χ are required, structure-activity analysis involving connectivity is relatively quick and therefore useful, at least for an initial screening of data.

Acknowledgements

One of us (J.C.B.) gratefully acknowledges receipt of a postgraduate studentship from the Department of Education, Northern Ireland.

REFERENCES

Chou, J. T., Jurs, P. C. (1980) Physical Chemical Properties of Drugs—Medicinal Research Series Vol. 10, 163–169. New York, Marcel Dekker, Inc.

Hall, L. H., Kier, L. B., Murray, W. J. (1975) J. Pharm. Sci. 64: 1974–1977

Hansch, C., Fujita, T. (1964) J. Am. Chem. Soc. 86: 1616–1626

Hansch, C., Leo, A. (1979) Substituent Constants for Correlation Analysis in Chemistry and Biology. New York, John Wiley and Sons

Hansch, C., Lien, E. J. (1971) J. Med. Chem. 14: 653–670

Hansch, C., Lien, E. J., Anderson, S. M. (1967) Ibid. 11: 430–441

Kier, L. B. (1980) J. Pharm. Sci. 64: 1034–1039

Kier, L. B., Hall, L. H. (1976) Molecular Connectivity in Chemistry and Drug Research, London, Academic Press Ltd

Kier, L. B., Hall, L. H., Murray, W. J., Randic, M. (1975) J. Pharm. Sci. 64: 1971–1974

Millership, J. S., Woolfson, A. D. (1978) J. Pharm. Pharmacol. 30: 483–485

Millership, J. S., Woolfson, A. D. (1980) Ibid., 32: 610–614

Murray, W. J., Hall, L. H., Kier, L. B. (1975) J. Pharm. Sci. 64: 1978–1981

Purcell, W. P., Bass, G. E., Clayton, J. M. (1973) Strategy of Drug Design, New York, Wiley-Interscience

Randic, M. (1975) J. Am. Chem. Soc., 97: 6609–6615

Tute, M. S. (1971) Advances in Drug Research 6: 1–77