J. Pharm. Pharmacol. 1996, 48: 883–886
Received December 12, 1995
Accepted March, 22 1996

A QSAR Study of Anti-inflammatory *N*-Arylanthranilic Acids

J. C. DUFFY, J. C. DEARDEN AND C. ROSTRON

School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK

Abstract

A detailed quantitative structure-activity relationship (QSAR) analysis of a series of 112 anti-inflammatory *N*-arylanthranilic acids has been performed to determine which physicochemical properties of these compounds are responsible for their anti-inflammatory activity. The results indicate that activity is modelled best by molecular shape parameters. The angle between the planes of the two benzene rings, dictated by the substitution pattern of the compounds, also appears relevant to activity. Dipole moments show some significance, but log P and other physicochemical parameters correlate poorly with activity. The best QSAR obtained was:

$$\log(1/\text{MED}) = 1{\cdot}601B_{1(2)} + 0{\cdot}576B_{1(3)} - 1{\cdot}187B_{3(4)} + 0{\cdot}522B_{1(6)} - 1{\cdot}681\mu(\text{bond})_{(2)} - 0{\cdot}208\mu(\text{bond})_{(5)} - 0{\cdot}265\mu(\text{bond})_{(6)} - 0{\cdot}226$$

$$n = 112, r = 0{\cdot}855, r^2(\text{adj}) = 0{\cdot}716, F = 40{\cdot}33, s = 0{\cdot}511, r(CV)^2 = 0{\cdot}665$$

where B_1 and B_3 are Verloop substituent width parameters and μ(bond) is bond dipole (position in parentheses).

Anti-inflammatory drugs represent a structurally diverse range of compounds. Although many mechanisms have been proposed to account for their activity, as yet little is known about their specific drug-receptor interactions. Analysis of existing data sets of anti-inflammatory drugs may provide information as to which specific molecular features are responsible for their biological activity.

Within the literature there are many examples of QSAR analyses for anti-inflammatory drugs, obtained using both in-vivo and in-vitro data. In these reports various physicochemical parameters are shown to correlate with biological activity. Hydrophobic parameters are expressed most frequently in the regression equations, and electronic parameters are also common. For example, Van den Berg et al (1975) reported that the ability of 2-aryl-1,3-indandiones to inhibit prostaglandin synthesis in-vitro was increased in compounds with more hydrophobic, electron-withdrawing substituents:

$$\log(1/\text{IC50}) = 3{\cdot}474 + 0{\cdot}379\pi + 1{\cdot}635\sigma \quad (1)$$
$$n = 24, r = 0{\cdot}912, s = 0{\cdot}239, F = 52{\cdot}023$$

where π is the hydrophobic substituent constant and σ is the Hammett constant, representing electron-directing ability.

Kuchař et al (1990) studied the activity of arylalkanoic acids in the in-vivo kaolin-induced rat paw oedema assay. They too showed that the potency of the compounds correlated with their hydrophobic and electronic properties:

$$\log I^k = 1{\cdot}167\sum\pi - 0{\cdot}278\left(\sum\pi\right)^2 + 0{\cdot}289\sum\sigma - 1{\cdot}237 \quad (2)$$
$$n = 39, r = 0{\cdot}940, s = 0{\cdot}071, F = 89{\cdot}3$$

Correspondence: J. C. Dearden, School of Pharmacy and Chemistry, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, UK.

On the other hand, Dearden & George (1979) found that in-vivo anti-inflammatory activity of a series of aspirin derivatives in the rat could be correlated with hydrophobicity and the shape of the substituent in the 4-position:

$$\log(1/\text{ED50}) = 2{\cdot}285 + 1{\cdot}031\log P - 0{\cdot}195(\log P)^2 - 0{\cdot}045L_{(4)} - 0{\cdot}244B_{2(4)} \quad (3)$$
$$n = 28, r = 0{\cdot}966, s = 0{\cdot}113$$

where L and B_2 are Verloop substituent length and width parameters respectively.

In the present study a detailed QSAR analysis was performed on a large data set, containing a wide variety of substituents in several different positions. The data set was that published by Kaltenbronn et al (1983), comprising 112 anti-inflammatory *N*-arylanthranilic acids. In their experiments, compounds were administered by gavage, and their ability to suppress erythema developing in the skin of depilated albino guinea-pigs, 2 h after exposure to UV radiation, was measured on an all-or-nothing basis. Compounds which were significantly more active than vehicle were administered at half of the previous dose, until a dose was reached for which the response was significantly less than that of a reference level of phenylbutazone. This dose was recorded as the minimal effective dose (MED). The structure of the parent compound is shown in Fig. 1.

The position and nature of the substituents, along with the MED of the respective derivatives, are given in Table 1. Previously, for this data set, we reported (Dearden & Duffy 1993) that of 135 physicochemical parameters generated, only indicator variables for the position and nature of the substituents correlated with activity. This was represented by the equation:

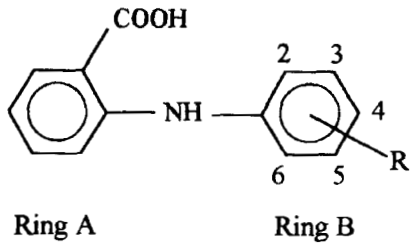

FIG. 1. Structure of *N*-arylanthranilic acids. Parent; R = H

$$\log(1/\text{MED}) = 1{\cdot}052I_2 + 0{\cdot}679I_3 - 0{\cdot}503I_4 - 0{\cdot}459I_5 - 0{\cdot}881I_{N_A} + 0{\cdot}342N_{N_A} - 0{\cdot}161$$
$$n = 112, r = 0{\cdot}84, r^2(\text{adj}) = 0{\cdot}71, s = 0{\cdot}529, F = 45{\cdot}1 \quad (4)$$

MED = minimal effective dose (mmol kg^{-1}). I_2, I_3, I_4, I_5 = indicator variables for the presence of substituents in the 2, 3, 4 and 5 positions, respectively, I_{NA} = indicator variable for the presence of a non-alkyl substituent in the 4 position, N_{NA} = total number of non-alkyl substituents.

In QSAR analysis one does not know a priori which physicochemical parameters are going to be significant. This leads to a dilemma as, if many parameters are investigated, there is an increased risk of chance correlations occurring. Conversely, if this is avoided by generating only a few parameters, an important parameter may be overlooked. The use of a large number of parameters in this study is justified because, as the data set is large, (consisting of 112 compounds), the risk of chance correlations is low.

Most published QSARs show the importance of particular physicochemical parameters in describing activity. However, this was not the case for this data set. Although the use solely of indicator variables in a regression equation is open to criticism on statistical grounds, the equation does provide meaningful information. The main factor governing activity is the specific position of the substituents which will determine the fit at the receptor site. This supports the conclusion drawn by Kaltenbronn et al (1983), that it is the overall shape of the compounds which is important. Hence a more detailed analysis of molecular shape was carried out.

Method

The analysis was performed using a series of programs supplied by Oxford Molecular Limited, running on a Silicon Graphics Personal IRIS 4D/25. Each of the compounds was subjected to conformational analysis using the COBRA package. The output files were analysed using PIMMS, an interactive molecular modelling system. Scherrer (1985) reported that the carboxylic acid function must be present and ortho to the diarylamine for this type of compound to show anti-inflammatory activity. This group was therefore treated as the fixed group. The carboxylic acid function of each molecule in turn was fitted onto the carboxylic acid function of the parent molecule and the structure optimized using the COSMIC energy minimization routine. Visual analysis of the compounds showed superimposition of the carboxylic acid functions and their respective A-rings, with the B-rings containing the substituents lying at various angles to the A-rings. The angles between the planes of the two benzene rings were obtained using the MAD program. The compounds, in the conformations obtained from PIMMS, were then read into the TSAR program for further analysis. TSAR was used to generate many physicochemical parameters which were analysed using the stepwise regression routine within the same program. The Verloop shape parameters L and B_{1-4} were generated for each of the substituents. B_5 values were also obtained, representing the maximum width of the substituent in any direction perpendicular to the axis of L. Other parameters generated included log P, bond lipoles, bond dipoles, shape similarity indices, molecular connectivities and other topological indices, and dipole moment.

MED values were converted to mmol kg^{-1} before being subjected to QSAR analysis.

Results

Stepwise regression analysis of the Verloop parameters alone resulted in the following equation:

$$\log(1/\text{MED}) = 1{\cdot}615B_{1(2)} - 0{\cdot}508B_{5(2)} + 0{\cdot}631B_{1(3)} - 1{\cdot}225B_{3(4)} + 1{\cdot}048B_{1(6)} - 1{\cdot}177$$
$$n = 112, r = 0{\cdot}818, r^2(\text{adj}) = 0{\cdot}657, F = 42{\cdot}88, s = 0{\cdot}561, r(\text{CV})^2 = 0{\cdot}613 \quad (5)$$

where CV = cross validated. The subscripts in parentheses relate to the position of the substituent considered.

Although the correlation coefficient r = 0·818 is not as high as desired it is still significant in this case as only 5 parameters account for two thirds of the variation in the data. The cross validated r^2 of 0·613 confirms that equation 5 is of some value in predicting activity. As Lozano et al (1993) had reported that dipole moments may be important in aligning the drug with the receptor, the values for total dipole moment were included in the regression equation with the above parameters. The correlation was slightly improved:

$$\log(1/\text{MED}) = 1{\cdot}718B_{1(2)} - 0{\cdot}55B_{5(2)} + 0{\cdot}545B_{1(3)} - 1{\cdot}211B_{3(4)} + 1{\cdot}045B_{1(6)} + 0{\cdot}147\mu - 1{\cdot}519$$
$$n = 112, r = 0{\cdot}835, r^2(\text{adj}) = 0{\cdot}683, F = 40{\cdot}24, s = 0{\cdot}540, r(\text{CV})^2 = 0{\cdot}666 \quad (6)$$

where μ = total dipole moment.

Finally all of the parameters generated were included in the stepwise regression routine. The parameters which appeared in the resulting regression equation were analysed for intercorrelation and highly correlated variables were removed, giving the equation:

$$\log(1/\text{MED}) = 1{\cdot}601B_{1(2)} + 0{\cdot}576B_{1(3)} - 1{\cdot}187B_{3(4)} + 0{\cdot}522B_{1(6)} - 1{\cdot}681\mu(\text{bond})_{(2)} - 0{\cdot}208\mu(\text{bond})_{(5)} - 0{\cdot}265\mu(\text{bond})_{(6)} - 0{\cdot}226$$
$$n = 112, r = 0{\cdot}855, r^2(\text{adj}) = 0{\cdot}716, F = 40{\cdot}33, s = 0{\cdot}511, r(\text{CV})^2 = 0{\cdot}665 \quad (7)$$

Table 1. Anti-inflammatory activities of *N*-arylanthranilic acids.

Substituent (s)	MED (mg kg^{-1})	Substiuent (s)	MED (mg kg^{-1})
H	200	2-CH_3,3-CN	3·1
2-Cl	50	2,3-Br_2	3·1
3-Cl	25	2-Br,3-CF_3	1·6
4-Cl	200	2-CH_3,3-OCH_3	6·2
2-CH_3	200	2-Br,3-CN	1·5
3-CH_3	100	2-CH_3,3-C_2H_5	3·1
4-CH_3	400	2-CH_3,3-CF_3	1·0
3-NH_2	400	2-CH_3,3-$SO_2N(CH_3)_2$	6·2
3-n-C_4H_9	200	2-CH_3,3-$N(CH_3)_2$	6·2
3-SCH_3	100	2,4-Cl_2	100
3-OC_2H_5	100	2,5-Cl_2	12·5
3-OCH_3	50	2,6-Cl_2	3·1
3-n-C_3H_7	50	3,4-Cl_2	100
3-Br	50	3,5-Cl_2	50
3-$SO_2N(CH_3)_2$	50	2,4-$(CH_3)_2$	400
3-C_2H_5	25	2,5-$(CH_3)_2$	200
3-CN	25	2,6-$(CH_3)_2$	50
3-NO_2	100	3,4-$(CH_3)_2$	200
3-C(=O)CH_3	200	3,5-$(CH_3)_2$	100
3-CF_3	3·3	3,5-$(CF_3)_2$	100
3-$N(CH_3)_2$	100	2-Cl,6-CH_3	12·5
2,3-Cl_2	2·1	2-CH_3,6-Cl,3-$N(CH_3)_2$	1·3
2-F,3-Cl	3·1	2,3,6-Cl_3	0·3
2-CH_3,3-Cl	5·3	2,6-Cl_2,3CH_3	0·4
2-Cl,3-CH_3	6·2	2,6-(CH_3)2,3-Cl	0·4
2-CH_3,3-NO_2	3·1	2-CH_3,3,6-Cl_2	0·8
2-CH_3,3-NH_2	50	2,3-$(CH_3)_2$,6-Cl	12·5
2,3-$(CH_3)_2$	10·4	2,6-$(CH_3)_2$,3-$SO_2N(CH_3)_2$	0·5
2,6-Cl_2,3-C_2H_5	0·8	2,6-$(C_2H_5)_2$,3-C(=O)CH_3	25
6-$(CH_3)_2$,3-NO_2	1·6	2-$(CH_3)_2$,3CF_3	0·8
2-NH_2,3-Cl,6-CH_3	25	2-Cl,3-$N(CH_3)_2$,6-CH_3	1·6
2-NH_3,3,6-$(CH_3)_2$	25	2,6-Cl,3-CN,6-CH_3	0·5
2-CH_3,3-Cl,6-NH_2	6·2	2,6-Cl_2,3-$SO_2N(CH_3)_2$	1·3
2-CH_3,3-NO_2,6-Cl	1·6	2,6-$(CH_3)_2$,3-$SOCH_3$	0·5
2-CH_3,3-NH_2,6-Cl	6·2	2,6-$(CH_3)_2$,3-SO_2CH_3	0·6
2,6-$(CH_3)_2$,3-NH_2	50	2,3,4-Cl_3	200
2-Cl,3,6-$(CH_3)_2$	3·1	2,3,5-Cl_3	3·1
2,6-$(CH_3)_2$,3-$N(CH_3)_2$	1·6	2,4,6-Cl_3	100
2,3-Cl_2,6-CH_3	0·8	2,4,5-Cl_3	>400
2,6-Cl_2,3-OCH_3	0·3	3,4,5-Cl_3	200
2,3,6-$(CH_3)_3$	6·2	2,3,5-$(CH_3)_3$	25
2,6-$(CH_3)_2$,3-Br	1·6	2,4,6-$(CH_3)_3$	400
2,6-$(C_2H_5)_2$,3-NO_2	12·5	2,4,5-$(CH_3)_3$	100
2,6-Cl_2,3-NH_2	3·1	2-CH_3,3,5-Cl_2	1·6
2,6-Cl_2,3-$N(CH_3)_2$	0·6	2,3-Cl_2,5-CH_3	6·2
2,6-$(CH_3)_2$,3-C_2H_5	1·6	3,5-Cl_2,4-CH_3	100
2,6-$(CH_3)_2$,3-C(=O)CH_3	0·9	2,5-$(CH_3)_2$,3-Cl	1·6
2,6-$(CH_3)_2$,3-n-C_3H_7	6·2	2,3-$(CH_3)_2$,5-Cl	25
2,6-$(CH_3)_2$,3-CN	0·4	2,3,4,5-Cl_4	100
2-CH_3,3-OCH_3,6-Cl	0·6	2,3,4,6-Cl_4	12·5
2,6-Cl_2,3-OC_2H_5	0·8	2,3,5,6-Cl_4	1·6
2,6-$(CH_3)_2$,3-SCH_3	0·4	2,3,5,6-$(CH_3)_4$	100
2,6-$(C_2H_5)_2$,3-$SO_2N(CH_3)_2$	12·5	2,3,4-Cl_3,6-CH_3	12·5
2,6-Cl_2,3-CF_3	0·8	2,4,6-Cl_3,3-CH_3	12·5
2,6-Cl_2,3-CN	1·6	2,3,4,5,6-Cl_5	25
2-CH_3,3-$SO_2N(CH_3)_2$,6-Cl	0·7	2,3,5-$(CH_3)_3$,4,6-Cl_2	100

Figures in subscript parentheses after the Verloop parameters relate to substituent positions; μ(bond) represents the bond dipole for the individual substituent positions given in subscript parentheses.

Although no correlation was obtained which related the angle between the planes of the two benzene rings to the activity of the molecules, Fig. 2 shows the relationship.

The molecules which are most active are those substituted in the 2, 3 and 6 positions. The angles between the planes of the two benzene rings for this subgroup all lie within the region 94·97–112·08°. Only 11 of the remaining 72 compounds have angles within this range. For the majority of the compounds not 2, 3, 6-substituted, the angles range from 112·11 to 124·12°; that is, the molecules are more planar.

Discussion

These results indicate that for this series of *N*-arylanthranilic acids the shape of a particular molecule is most significant in determining activity. Clearly any substituent in position 4 greatly reduces activity, the effect being independent of the nature of the substituent. It is, therefore, most probably due to direct steric hindrance interfering with drug receptor binding (Dearden & George 1979). The importance of the substituent

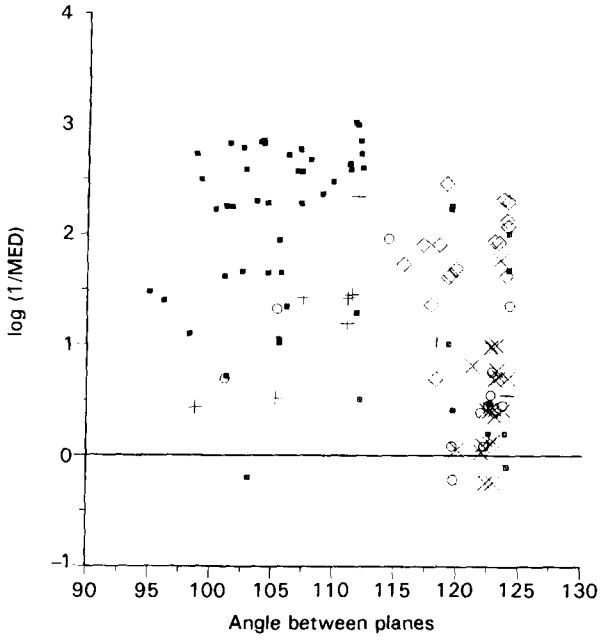

FIG. 2. Plot of activity versus angle between planes. ■ Tri-substituted (positions 2,3,6), □ tri-substituted (not 2,3,6), ◇ di-substituted (2 and 3), ○ di-substituted (not 2 and 3), + tetra- and penta-substituted, × mono-substituted.

shape and position appears to be their ability to maintain the molecule in an orientation conducive to receptor binding. This is inferred from the fact that shape parameters are consistently expressed in stepwise regression analysis. The equation highlights the importance of substituents in positions 2, 3 and 6 for anti-inflammatory activity. It is noteworthy also that substituent B_1 is the key shape parameter at each position, suggesting that a certain minimum bulk is required for good activity. The selection of bond dipoles suggests that dipole interaction controls the receptor binding. It is interesting that μ(bond) for position 3 is not selected, suggesting that substituents in the 3 position simply exert a buttressing effect on the 2-substituent, and do not participate in binding.

For more effective binding it is better for the ring B to be twisted with respect to ring A. As reported by Kaltenbronn et al (1983), 2-substitution results in a twisting effect, which is favourable to activity. 2, 3-Di-substituted molecules are highly active, due probably to the buttressing effect of the 3 substituent increasing the twisting effect of the 2 substituent. 2, 3, 6-Tri-substituted compounds of this series are the most effective anti-inflammatories in the assay used. From Fig. 2, it can clearly be seen that the presence of the additional substituent in the 6 position adds to the twisting effect, holding ring B in a more perpendicular orientation with respect to ring A. This also supports the theory of Kaltenbronn et al (1983) that these derivatives will, accordingly, show a better fit to the receptor.

Although not as significant as the steric effects, electronic effects are also involved (eqns 3 and 4), possibly, as Lozano et al (1993) suggested, in aligning the drug with the receptor.

Hydrophobic parameters are very often important in correlations with biological activity, as log P plays a significant role in the ability of molecules to cross biological membranes and may be important in hydrophobic binding at the receptor site. Despite *N*-arylanthranilic acids being predominantly ionized at physiological pH, their high log P values enable them to partition readily into lipid membranes. As log P is not expressed in any of the regression equations it is likely that transport is not a limiting factor to drug activity in this case.

Although it is disappointing that the correlation coefficient is quite low in equation 4, it is not entirely surprising, as in-vivo data are always subject to considerable biological variation. In addition, as can be seen from Table 1, the method employed by Kaltenbronn et al (1983) to determine MED led to inaccurate values almost certainly being recorded for some compounds (i.e. many values of 400, 200 and 100 mg kg^{-1}). It is expected that a congeneric series of compounds will show similar pharmacokinetic properties and act via the same mechanism of action. The biological systems involved are very complex and it is quite possible that some of the molecules are metabolized more rapidly than others, thus reducing the concentration at the site of action. Metabolites of some of the compounds also show anti-inflammatory activity, which will affect the measurement of biological activity. This makes quantitative analysis of receptor interactions more difficult.

Being the correct shape to fit into the receptor appears to be the most important factor in determining potency in this case, as the receptor interactions appear to be quite tolerant of very diverse groups. This would suggest that new compounds of this type may show little improvement in activity. However, some improvements in activity may be brought about by incorporating larger substituents in positions 2, 3 and 6 and no substituents in positions 4 and 5. Activity may be further improved if substituents with greater electron-attracting ability (greater negative bond dipoles) are considered. New drugs of a very different basic structure which possess a similar overall shape may also be more effective.

References

Dearden, J. C., Duffy, J. C. (1993) QSAR study of anti-inflammatory activity of a series of *N*-arylanthranilic acids. J. Pharm. Pharmacol. 45 (Suppl. 2): 1142

Dearden, J. C., George E. (1979) Anti-inflammatory potencies of some aspirin derivatives: a quantitative structure-activity study. J. Pharm. Pharmacol. 31(Suppl.): 45P

Kaltenbronn, J. S., Scherrer, R. A., Short, F. W., Jones, E. M., Beatty, H. R., Winder, C. V., Wax, J., Williamson, W. R. N. (1983) Structure-activity relationships in a series of anti-inflammatory *N*-arylanthranilic acids. Arzneim. Forsch. 33: 621–626

Kuchař, M., Grimova, J. Rejholec, V., Tomkova, H., Jelinkova, M., Holubek, J. (1990) Use of QSAR in design of anti-inflammatory fluorinated arylalkanoic acids. Collect. Czech. Chem. Commun. 55: 296–306

Lozano, J. J., Lopez, M., Ruiz, J., Vasquez, I. J., Pouplana, R. **(1993)** QSAR in the non-steroidal anti-inflammatory agents: the fenamic acids. In Wermuth, C. G. (ed.) Trends in QSAR and Molecular Modelling 92, ESCOM, Leiden, pp 560–561

Scherrer, R. A. (1985) Fenamic acids. In: Rainsford, K. D. (ed.) Anti-Inflammatory and Anti-Rheumatic Drugs, Volume II. CRC Press Inc., Boca Raton FL, pp 65–85

Van den Berg, G., Bultsma, T., Nauta, W. T. (1975) Inhibition of prostaglandin biosynthesis by 2-aryl-1,3-indandiones. Biochem. Pharmacol. 24: 1115–1119