High-Performance Liquid Chromatographic Determination of the π Values of Azol-*N*-yl Substituents

JULIO ALVAREZ-BUILLA*, TERESA CRESPO*, RAQUEL LOPEZ*, JOSE ELGUERO‡x, CATHERINE TOIRON‡, AND GLORIA I. YRANZO‡

Received January 8, 1991, from the *Departamento de Química Orgánica, Universidad de Alcalá de Henares, 28871 Alcalá de Henares, Madrid, Spain, and the ‡Instituto de Química Médica (CSIC), Juan de la Cierva 3, 28006 Madrid, Spain. Accepted for publication August 6, 1991.

Abstract □ The π (hydrophobic constant) values for 16 parent azoles (pyrrole, imidazole, pyrazole, four triazoles, two tetrazoles, indole, benzimidazole, 1*H*- and 2*H*-indazoles, 1*H*- and 2*H*-benzotriazoles, and carbazole) were calculated from the logarithms of the capacity factors (log k') determined by HPLC. The values thus obtained are discussed according to an additive model in which the number and position of pyridinelike nitrogen atoms and the annelation effect are considered.

The logarithm of the octanol–water partition coefficient of a compound, log $P_{o/w}$, is a parameter frequently used as a measure of hydrophobicity in the study of quantitative structure–activity relationships (QSARs) in drug design and pharmacology.[1] Traditionally, log $P_{o/w}$ has been determined by the shake-flask method,[2] but lately, many attempts have been made to simplify this determination with different techniques.[3] Techniques based on HPLC[4] with hydroorganic mobile phases continue to be the most popular and the best studied alternatives so far.

The discovery of the additive and constitutive character of partition coefficients has led to the development of several systems of substituent or fragmental constants, which have been used as hydrophobicity descriptors in parametric QSAR models. By analogy with Hammett's equation, Hansch and co-workers[5] used a linear free-energy relationship to define a hydrophobic constant (π) according to the following equation: $\pi_X = \log P_{RX} - \log P_{RH}$. In this equation, π_X is the hydrophobic constant of substituent X, log P refers to the partition coefficient in an octanol–water system, and RH and RX are the unsubstituted compound and its X-substituted derivative, respectively. With few exceptions, the method has been used to calculate π values of substituents on the benzene ring, and the Pomona College group has developed a large data base[6,7] that includes most of the simple benzene substituents. However, few examples of heterocycles as substituents are included in the data base. Paucity of information about the electronic or hydrophobic properties of heterocycles limits the usefulness of the Pomona College data base in QSAR design.

To date, few papers have dealt with π values for compounds with heterocyclic substituents (heterocyclic π values), and only in the azine field is there a major revision of partition coefficients,[8] a study on lipophilicity of nicotinate esters,[9] and a determination of π values in substituted pyridines.[10] A recent publication[11] about the use of propylene glycol dipelargonate (PGDP) as a solvent in the determination of log P values lists log P values of simple heterocyclic derivatives taken from reference 7.

In the field of nitrogen heterocycles, the importance of azoles in medicinal chemistry is clear. Some azoles (pyrrole, indole, imidazole, and benzimidazole) are present in endogenous substances, whereas others (pyrazole, 1,2,4-triazole, tetrazole, and carbazole) belong to important xenobiotic compounds. As nitrogen-linked moieties, azol-*N*-yl residues can be found in drugs as important as antifungal and anti-inflammatory agents, modulators of the central nervous system, and so on. This paper provides a complete set of π values for azol-*N*-yl substituents, except for the rather unusual and unstable isoindolyl group.

Because of the satisfactory results obtained for other series,[10,12–15] HPLC was used again to determine the logarithms of the capacity factor (log k') of all the *N*-methyl- and *N*-benzylazoles studied.

Experimental Section

Materials—Analytical-grade chemicals were either commercially available or prepared according to standard procedures. The identities of all compounds were verified by IR and NMR spectroscopy, and the purities were determined by HPLC.

Apparatus—All experiments were performed on a series 10 Perkin-Elmer liquid chromatograph equipped with a fixed-wavelength detector that was set at 254 nm. A 10-μm Spherisorb ODS2 column (250 × 4.6-mm i.d.; Phase Separations Ltd.) was used throughout. Retention times were computed by a Spectra-Physics SP 4290 integrator.

Methods—*Chromatographic Determination of log k'*—Isocratic elution was performed with HPLC-grade methanol:0.010 M triethylamine (TEA; 65:35 or 55:45, v/v; pH 12.0).[10] All determinations were made at a constant flow rate of 1.0 mL/min at room temperature. Samples were dissolved in methanol at concentrations yielding similar peak areas with injections of up to 10 μL (typically 5 μL). The column dead time (t_o) was determined by injection of a methanol solution of $NaNO_2$. From the solute retention time (t_R), log k' was calculated as follows: $\log k' = \log [(t_R - t_o)/t_o]$. All log k' values are the means of at least three measurements.

Results and Discussion

All the available information on azoles, classed according to IUPAC seniority rules, has been included in Table I. (In Table I and henceforth in the text, *N*-methylazoles are designated **a**, *N*-benzylazoles are designated **b**, and *N*-adamantylazoles are designated **c**.) In a recent paper, Taylor et al.[11] collected log P values for azoles and *N*-methylazoles (shake-flask method). However, to avoid tautomerism problems, only the log P values of the *N*-methylazole derivatives of **1**, **2**, **5**, **7**, **10**, and **15** ($\log P_{octanol}$; the derivatives are designated **1a**, **2a**, **5a**, **7a**, **10a**, and **15a**, respectively) as well as the corresponding value for **14a** (taken from reference 6) have been included in Table I. The values for **3a** and **16a** can be estimated from the linear equations relating $\log P_{octanol}$ to $\log P_{PGDP}$ and $\log P_{octanol}$(*N*-methylazoles) to $\log P_{octanol}$(NH-azoles) (Table I, footnotes). These values (shown in parentheses in Table I) proved to be reasonable estimates of the missing values.

The log P values of a few *N*-adamantylazoles[15] are also shown in Table I. Some of these compounds (the *N*-adamantylazole derivatives of **4**, **5**, and **9**, or **4c**, **5c**, and **9c**) do not

0022-3549/92/0600-0577$02.50/0

Table I—Experimental log *P* and log *k'* of Azol-*N*-yl Substituents

Compound	Substituent	log *P*		log *k'* in TEA:Methanol (35:65)		log *k'* in TEA:Methanol (55:45)	
		N-CH_3 (**a**)	*N*-Adamantyl (**c**)	*N*-CH_3 (**a**)	*N*-Benzyl (**b**)	*N*-CH_3 (**a**)	*N*-Benzyl (**b**)
1	Pyrrol-1-yl	1.21[a]	—[b]	−0.069 ± 0.008	0.607 ± 0.01	—	—
2	Imidazol-1-yl	−0.03[a]	2.90	(−0.423)[c]	0.044 ± 0.008	−0.278 ± 0.03	0.545 ± 0.007
3	Pyrazol-1-yl	(0.28)[d]	3.05	−0.257 ± 0.001	0.135 ± 0.02	—	—
4	1,2,4-Triazol-4-yl	—	—	−0.534 ± 0.004	−0.386 ± 0.01	—	—
5	1,2,4-Triazol-1-yl	−1.08[a]	—	(−0.394)[c]	−0.194 ± 0.006	−0.220 ± 0.003	0.231 ± 0.01
6	1,2,3-Triazol-2-yl	—	—	−0.003 ± 0.004	0.024 ± 0.02	0.363 ± 0.008	—
7	1,2,3-Triazol-1-yl	−1.20[a]	—	−0.183 ± 0.003	−0.122 ± 0.01	0.125 ± 0.03	0.245 ± 0.02
8	Tetrazol-1-yl	—	2.95	(−0.538)[c]	−0.075 ± 0.005	−0.512 ± 0.007	0.485 ± 0.01
9	Tetrazol-2-yl	—	—	(−0.688)[c]	−0.240 ± 0.02	−0.814 ± 0.01	0.207 ± 0.007
10	Indol-1-yl	2.72[a]	—	0.542 ± 0.007	0.991 ± 0.03	—	—
11	Benzimidazol-1-yl	—	3.64	0.097 ± 0.001	0.443 ± 0.01	—	—
12	Indazol-2-yl	—	3.68	0.130 ± 0.03	0.512 ± 0.04	—	—
13	Indazol-1-yl	—	4.33	0.232 ± 0.05	0.644 ± 0.02	—	—
14	Benzotriazol-2-yl	1.64[e]	4.30	0.097 ± 0.01	0.635 ± 0.008	—	—
15	Benzotriazol-1-yl	1.13[a]	3.52	−0.133 ± 0.01	0.352 ± 0.01	—	—
16	Carbazol-9-yl	(4.56)[f]	—	1.115 ± 0.006	1.536 ± 0.001	—	—

[a] From ref 11. [b] —, Not determined. [c] Calculated from data obtained with TEA:methanol (55:45; see text). [d] From the equation log $P_{octanol}$ = 0.402 + 0.746 log P_{PGDP} (n = 6, r^2 = 0.945) and the value for log P_{PGDP} (**3a**) (from ref 11) of −0.16 (the six compounds used in the regression were **1a**, **2a**, **5a**, **7a**, **10a**, and **15a**). [e] From appendix II (compound 4707) of ref 6. [f] From the equation log $P_{octanol}$ (*N*-methylazoles) = −0.054 + 1.234 log $P_{octanol}$ (NH-azoles) (n = 6, r^2 = 0.954) and the value log $P_{octanol}$ (carbazole) = 3.74 (compounds 9588 and 9589 of ref 7); the six compounds used in the regression were **1**, **2**, **3**, **5**, **10**, and **15**.

absorb in the UV range and were not detected. The log k' values of *N*-methylazoles (**a**) and *N*-benzylazoles (**b**), determined under two different conditions (Table I), indicate that, with an eluant richer in aqueous TEA, the more hydrophilic compounds gave peaks closer to the sodium nitrite standard that were better resolved, and the log k' values were determined with smaller errors. For **9a** however, the log k' values were so imprecise that the value shown (−0.688) was calculated from the value (−0.814) determined in 55:45 TEA:methanol. If **9a** is excluded, there exists a linear relationship between the log k' values for *N*-methylazoles (**a**) and *N*-benzylazoles (**b**) in both 35:65 and 55:45 mixtures of TEA:methanol (eq 1):

$$\log k'_{(35:65)} = -0.264 + 0.501\ (\pm 0.051) \log k'_{(55:45)} \qquad (1)$$

$$(r^2 = 0.924, F = 96.932, s = 0.053, n = 10)$$

In eq 1, r is the correlation coefficient, F is the variance ratio, s is the standard deviation, and n is the number of observations.

The log *P* values of *N*-adamantylazoles (log $P_{c(N\text{-}Ad)}$) are consistent with the log $k'_{(35:65)}$ values determined here. The correlation is better with the benzyl derivatives (log $k'_{(35:65)b}$ in eq 3) than with the methyl derivatives (log $k'_{(35:65)a}$ in eq 2), a fact indicating that, for the latter, the values were less accurate:

$$\log P_{c(N\text{-}Ad)} = 3.721 + 1.802\ (\pm 0.364) \log k'_{(35:65)a} \qquad (2)$$

$$(r^2 = 0.803, F = 24.530, s = 0.270, n = 8)$$

$$\log P_{c(N\text{-}Ad)} = 2.880 + 1.981\ (\pm 0.241) \log k'_{(35:65)b} \qquad (3)$$

$$(r^2 = 0.919, F = 67.722, s = 0.174, n = 8)$$

The observed linearity gives confidence both to the values of log k' and to the additivity of hydrophobic properties, at least for the eight compounds (**2**, **3**, **8**, and **11–15**) included in the set.

Because eqs 2 and 3 have similar slopes, there should be a proportionality between log k'_b and log k'_a. When these values are represented in a two-dimensional plot, it appears that **4–7** (all triazole isomers) are outsiders. Why triazoles behave differently is difficult to ascertain, but it is almost sure that the anomaly lies in the hydrophilic *N*-methylazole (**a**) series and not in the *N*-benzylazole (**b**) derivatives. For the remaining compounds, eq 4 was used:

$$\log k'_b = 0.454 + 0.980\ (\pm 0.055) \log k'_a \qquad (4)$$

$$(r^2 = 0.969, F = 312.111, s = 0.089, n = 12)$$

Finally, the relationship between log k' values and the literature log *P* values was calculated. Figure 1 represents the points whose log *P* values are known (including those of **3a** and **17a**). Fortunately, relationships are linear if **7b** is excluded from the *N*-benzylazoles and if **5a** and **7a** are excluded from the *N*-methylazoles. The curvature in the left part of Figure 1 is produced because errors increase as the peak of the substance becomes closer to the nitrite standard in the chromatogram. Because the two lines in Figure 1 are almost parallel, it was assumed that both slopes are identical. With this hypothesis, all the points of Figure 1 (except for those of **5a**, **7a**, and **7b**) led to eq 5:

$$\log P_a = 1.257 + 3.030\ (\pm 0.108) \log k' - 1.508\ (\pm 0.118)\mathrm{Ph} \qquad (5)$$

$$(r^2 = 0.985, F = 397.45, s = 0.213, n = 15)$$

In eq 5, log k' represents log k' for both *N*-methylazoles (log k'_a) and *N*-benzylazoles (log k'_b), and Ph is a dummy variable (Ph = 0 for *N*-methylazoles and Ph = 1 for *N*-benzylazoles). The value of 1.51 for Ph corresponds to the separation of the parallel lines in the log *P* axis. The physical meaning of this value is clear: it is the averaged increase in hydrophobicity when a hydrogen of the *N*-methyl group is changed by a phenyl (i.e., it is the π value of the phenyl group as a

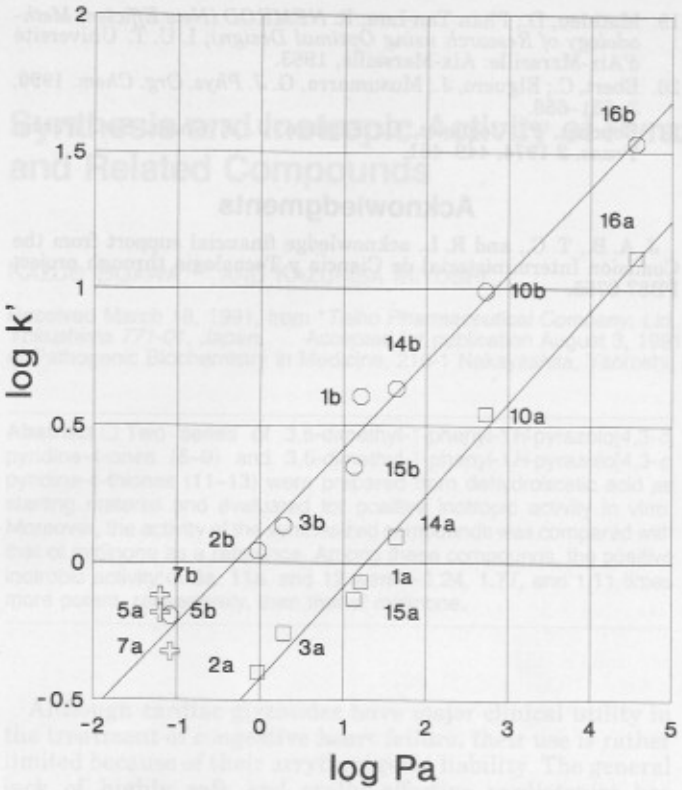

Figure 1—Graphical representation of the regression model relating log P_a with log k' of compounds **a** (□) and **b** (○) (eq 5); **5a**, **7a**, and **7b** (⊕) appear as calculated from experimental values.

substituent on an sp^3 carbon). This value depends on the molecules used for its determination;[16] for instance, between methane (1.09) and toluene (2.69), π_{phenyl} is 1.60, and between toluene and diphenylmethane (4.14),[6] π_{phenyl} is 1.45. The determination of the π_{phenyl} value of 1.51 in eq 5 is consistent with these values.

Equation 5 also yields an estimation of the log P of N-methylazoles. By subtracting the values for methane (log P = 1.09),[6] these log P values led to π values for azol-N-yl residues. The π values for azol-N-yl substituents, obtained by the shake-flask method [log P_a (from Table I) − 1.09], HPLC, and the model [log P_a (from eq 5) − 1.09], are compared in Table II. As noted previously, the HPLC method cannot be used for **7** and **8**. If it is assumed that the hydrophobic nature of an azolyl substituent is linearly related to the number of α-nitrogen (Nα) atoms (0, 1, 2), to the number of β-nitrogen atoms (Nβ) (0, 1, 2), to the number of adjacent lone pairs (LPs; 0, 1, 2)[17] (the so-called α-effect[11]), to the number of benzene rings (benzo; 0, 1, 2; the so-called annelation effect[18]), and to the number of isoaromatic rings (isobenzo; 0, 1; **12** and **14**), then a multiple regression,[19] with π values determined by both the shake-flask and HPLC methods, leads to eq 6:

$$\pi = 0.18 - 0.87\,N\alpha - 1.39\,N\beta + 0.61\,LP + 1.52\,benzo + 2.09\,isobenzo \quad (6)$$

$$(r^2 = 0.99,\ F = 302.97,\ s = 0.19,\ n = 21)$$

In eq 6, π values for **7** (shake-flask) and **12** (HPLC) are excluded. The π values for **7** and **8**, which were not attainable from HPLC determinations, can be estimated with eq. 6. For 1,2,3-triazol-1-yl (**7**), the estimated π value of −1.46 is quite different from that determined by the shake-flask method (−2.29[11]). For indazol-2-yl (**12**), the model value of 1.40 is very different from the HPLC value of 0.21. Given the empirical nature of the model, the experimental values should be preferred.

Table II—Hydrophobicity Values of Azol-*N*-yl Substituents

Compound	π			σp^a
	Shake-flask	HPLC	Model	
1	0.12[b]	0.23	0.18	0.20
2	−1.12[b]	−1.17	−1.20	0.30
3	−0.81[c]	−0.77	−0.68	0.18
4	—[d]	−2.51	−2.59	0.33[e]
5	−2.17[b]	−1.93	−2.07	0.29
6	—	−1.27	−1.55	0.42
7	−2.29[b]	—	(−1.46)	0.32
8	—	—	(−2.33)	0.59[e]
9	—	−2.07	−2.24	0.52
10	1.63[b]	1.73	1.70	0.23
11	—	0.00	0.32	0.31
12	—	0.21	(1.40)	0.18
13	—	0.61	0.84	0.16
14	0.55[f]	0.52	0.53	—
15	0.04[b]	−0.25	0.06	0.27
16	3.47[g]	3.43	3.22	0.40

[a] Hammett substituent constant, from ref 20. [b] From ref 11. [c] See footnote "d" in Table I. [d] —, Not determined. [e] From ref 21. [f] See footnote "e" in Table I. [g] See footnote "f" in Table I.

The more "available" Nβ produces more hydrophilic effects than the more "shielded" Nα. The presence of two adjacent LPs, (LP/LP) diminishes these effects. The annelation effect is strongly hydrophobic, and for the isobenzo effects, the loss of aromaticity contributes to the value of the coefficient.

Conclusions

Retention (log k') and hydrophobicity (log $P_{o/w}$) parameters have been determined for N-methyl- and N-benzylazoles by HPLC. The linear correlation observed between these parameters demonstrates that both are intrinsic (free-energy) parameters of the azole set. The π values of the azoles as substituents were calculated from the log k' and log $P_{o/w}$ values and, through an empirical model, were explained in terms of the structural features of every ring. The Hammett (σp) values of azolyl substituents, which were recently determined,[20] together with the π values obtained here, provide a suitable ground for new QSAR studies.

References and Notes

1. Van de Waterbeemd, H.; Testa, B. *Adv. Drug Res.* **1987**, *16*, 85–225.
2. Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525–616.
3. Hersey, A.; Hill, A. P.; Hyde, R. M.; Livingstone, D. J. *Quant. Struct.–Act. Relat.* **1989**, *8*, 288–296.
4. Kaliszan, R. *Quant. Struct.–Act. Relat.* **1990**, *9*, 83–87, and references cited therein.
5. Fujita, T.; Iwasa, J.; Hansch, C. *J. Am. Chem. Soc.* **1964**, *86*, 5175–5180.
6. Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; J. Wiley: New York, 1979.
7. Hansch, C.; Leo, A. *Pomona College Medicinal Chemistry Project log P Database*; Comtex Scientific Corporation: New York, 1983.
8. Lewis, S. J.; Mirrlees, M. S.; Taylor, P. J. *Quant. Struct.–Act. Relat.* **1983**, *2*, 100–111.
9. Reymond, D.; Chung, G. N.; Mayer, J. M.; Testa, B. *J. Chromatogr.* **1987**, *391*, 97–109.
10. Gago, F.; Alvarez-Builla, J.; Elguero, J. *J. Chromatogr.* **1988**, *449*, 95–101.
11. Leahy, D. E.; Taylor, P. J.; Wait, A. R. *Quant. Struct.–Act. Relat.* **1989**, *8*, 17–31.
12. Gago, F.; Alvarez-Builla, J.; Elguero, J. *J. Chromatogr.* **1986**, *360*, 247–251.

13. Gago, F.; Alvarez-Builla, J.; Elguero, J.; Diez-Masa, J. C. *Anal. Chem.* **1987**, *59*, 921–923.
14. Gago, F.; Alvarez-Builla, J.; Elguero, J. *J. Liq. Chromatogr.* **1987**, *10*, 1031–1047.
15. Claramunt, R. M.; Sanz, D.; Elguero, J.; Alvarez-Builla, J.; Gago, F. *Farmaco Ed. Sci.* **1987**, *42*, 915–919.
16. Lewis, S. J.; Mirrlees, M. S.; Taylor, P. J. *Quant. Struct.–Act. Relat.* **1983**, *2*, 1–6.
17. Taft, R. W.; Anvia, F.; Taagepera, M.; Catalán, J.; Elguero, J. *J. Am. Chem. Soc.* **1986**, *108*, 3237–3239.
18. Catalán, J.; Claramunt, R. M.; Elguero, J.; Laynez, J.; Menendez, M.; Anvia, F.; Quian, J. H.; Taagepera, M.; Taft, R. W. *J. Am. Chem. Soc.* **1988**, *110*, 4105–4111.
19. Mathieu, D.; Phan Tan Luu, R. *NEMROD (New Efficient Methodology of Research using Optimal Design)*; I. U. T. Université d'Aix-Marseille: Aix-Marseille, 1983.
20. Ebert, C.; Elguero, J.; Musumarra, G. *J. Phys. Org. Chem.* **1990**, *3*, 651–658.
21. Bouchet, P.; Coquelet, C.; Elguero, J. *J. Chem. Soc., Perkin Trans. 2* **1974**, 449–451.

Acknowledgments

J. A. B., T. C., and R. L. acknowledge financial support from the Comisión Interministerial de Ciencia y Tecnologia through project PB87-0755.