
**ACIDITY IN WATER (pK_a VALUES) OF CARBOXYLIC ACIDS
DERIVED FROM SIMPLE HETEROCYCLES (AZOLES AND AZINES)**

Carlos CATIVIELA^a, Jean-Louis DEJARDIN^b, José ELGUERO^c, José I. GARCIA^a,
Emmanuel GONZALEZ^b and José A. MAYORAL^a

^a *Instituto de Ciencia de los Materiales de Aragón,
Universidad de Zaragoza, CSIC, 50009 Zaragoza, Spain*

^b *Laboratoires de Physique Appliquée et de Chimie Organique,
Université de Perpignan, 66025 Perpignan, France and*

^c *Instituto de Química Médica, CSIC, 28006 Madrid, Spain*

Received February 13, 1989

Accepted March 9, 1989

Dedicated to Professor Otto Exner on the occasion of his 65th birthday.

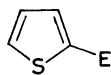
The pK_a of eighteen heterocyclic carboxylic acids derived from thiophene, furan, pyrrole, benzothiophene, benzofuran, indole, pyridine and quinoline have been determined by spectrophotometry and potentiometry. The values thus obtained are discussed using additive models of the Free-Wilson type. Some theoretical calculations within the AM1 approximation have been carried out in order to understand why annelation effects are dependent on the position, α or β , of the carboxylic group. In the case of pyridine and quinoline carboxylic acids, the problem of tautomerism between neutral and zwitterionic forms has been considered.

The use of different parameters to describe the effects of substituents on a molecule has arisen increasing interest in recent years, especially in QSAR methodology¹. Nevertheless, there are no systematic studies concerning these parameters when the substituent is an heterocyclic group. We have previously reported^{2,3} the hypothesis that structural characteristics in a Free-Wilson type model could be used to describe heterocyclic substituents and now we want to report the use of this approach to study the pK_a of a set of eighteen simple heterocyclic acids containing only one heteroatom.

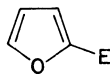
In order to avoid the dispersion of values of the literature⁴ we decided to determine the pK_a 's of all the above reported acids.

EXPERIMENTAL

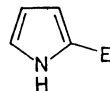
Heterocyclic compounds I–XVIII. These compounds were either commercially available or prepared according to standard procedures. Only 3-pyrrolicarboxylic acid *VI* was obtained using a new synthetic method⁵.



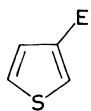
I



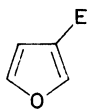
II



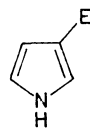
III



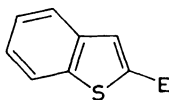
IV



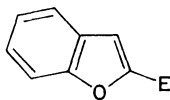
V



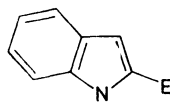
VI



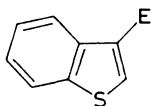
VII



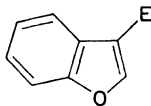
VIII



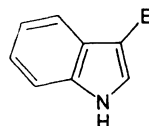
IX



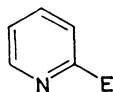
X



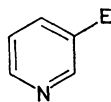
XI



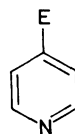
XII



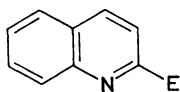
XIII



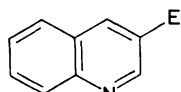
XIV



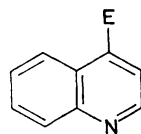
XV



XVI



XVII



XVIII

E = COOH

*pK*_a determinations. The *pK*_a of the heterocyclic acids at 20°C in water have been calculated by means of the relationship

$$pK_a = pH - \log \frac{D - D_{A^-}}{D_{AH} - D}$$

using the spectrometric technique previously described^{6,7}. D_{AH} , D_{A^-} and D are the optical densities of the acid, its conjugated base and the equilibrium mixture for a given pH, all of them measured at a chosen wavelength. Measurements have been carried out with a Pye Unicam spectrophotometer equipped with thermostated cells and a Tacussel pH-meter (1/100) coupled with a 12/HS electrode standardized with Titrisol buffers.

Acid concentration was comprised between 10^{-4} and 10^{-5} mol l⁻¹. The ionic strength was maintained constant by means of potassium chloride (0.1 mol l⁻¹) and the titrations were carried out under nitrogen atmosphere with decarbonated sodium hydroxide.

The curve $D = f(\text{pH})$ present an inflexion point I (see Fig. 1a) in which $D_I = 1/2(D_{AH} + D_{A^-})$ (Eq. (1)) allowing the calculation of the $\text{p}K_a$. However, the precision in the determination of this point is considerably increased using the intersection point of the $D - D_{A^-} = f(\text{pH})$ (Eq. (2)) and $D_{AH} - D = f(\text{pH})$ (Eq. (3)) curves (Fig. 1b). The solution of the system of equations (2) and (3) led to a $\text{p}K_a$ value identical, but more precise, than that obtained from Eq. (1).

When the solubility in water is sufficient (about 10^{-2} mol . l⁻¹), $\text{p}K_a$'s have been determined by potentiometry using decinormal sodium hydroxide with the classical procedure described by Albert and Serjeant⁸. The values thus obtained differ from the spectrophotometrically ones by less than 0.05 $\text{p}K_a$ units. In both methods the temperature was keep constant at $20.0 \pm 0.1^\circ\text{C}$. The experimental results are gathered in Table I.

AM1 calculations. AM1 calculations⁹ were performed on a Vax 11/780 computer (Zaragoza). All molecules were fully optimized unless the planarity which was imposed.

RESULTS AND DISCUSSION

Five-Membered Heterocyclic Acids

These acids, I–XII, can be described in a Free–Wilson way (0, 1 matrix) using three descriptors:

- 1) number of cycles: one (X_1^1) or two (benzoderivatives) (X_2^1)
- 2) position of the COOH group: α (X_3^1) or β (X_4^1)

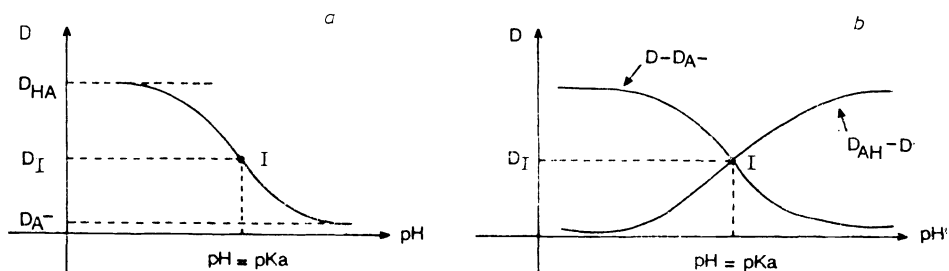


FIG. 1

Determination of $\text{p}K_a$ by spectrophotometric titration: a inflexion point method, b intersection method

3) nature of the heteroatom: S (X_5^1), O (X_6^1) or NH (X_7^1)
 leading to the following model:

$$pK_a = a_0^1 + a_1^1(X_1^1) + a_2^1(X_2^1) + a_3^1(X_3^1) + \dots + a_7^1(X_7^1). \quad (\text{Model 1})$$

The corresponding Free-Wilson matrix is then constructed in the usual way¹⁰, each heterocyclic substituent being unambiguously codified by a row vector. Using the averaged pK_a values of Table I and following the usual procedure^{2,3} the coefficients of the model can be calculated:

$$pK_a = 3.96 + 0.055(X_1^1) - 0.055(X_2^1) - 0.505(X_3^1) + 0.505(X_4^1) - 0.23(X_5^1) - 0.43(X_6^1) + 0.66(X_7^1); \quad r^2 = 0.94, r_A^2 = 0.91. \quad (4)$$

Although this empirical model seems quite satisfactory, the very weak effect of annelation ($a_2^1 - a_1^1 = 0.11$) results from averaging two opposite effects: (i) when the carboxylic group is in α -position, the annelation effects amounts to -0.43 pK_a

TABLE I
 Experimental pK_a values (water, 20°C)

Carboxylic acid	Spectrophotometry	Potentiometry
2-Thiophene <i>I</i>	3.48	3.50
2-Furan <i>II</i>	3.17	3.15
2-Pyrrole <i>III</i>	4.38	4.34
3-Thiophene <i>IV</i>	—	4.08
3-Furan <i>V</i>	4.03	3.97
3-Pyrrole <i>VI</i>	—	5.01
2-Benzothiophene <i>VII</i>	3.10	insoluble
2-Benzofuran <i>VIII</i>	2.79	~2.9
2-Indole <i>IX</i> ^a	3.84	3.79
3-Benzothiophene <i>X</i>	4.24	insoluble
3-Benzofuran <i>XI</i>	4.18	insoluble
3-Indole <i>XII</i> ^b	5.29	insoluble
2-Pyridine <i>XIII</i>	—	5.42
3-Pyridine <i>XIV</i>	—	4.80
4-Pyridine <i>XV</i>	—	4.87
2-Quinoline <i>XVI</i>	—	4.96
3-Quinoline <i>XVII</i> ^c	—	4.68
4-Quinoline <i>XVIII</i> ^d	4.78	4.74

In alcohol-water 50 : 50 mixtures, the following pK_a's were measured: ^a 4.97; ^b 7.00; ^c 4.57; ^d 4.52.

units (pairs *I* and *VII*, *II* and *VIII*, *III* and *IX*) and (ii) when it is in β -position, the effect amounts to +0.21 p*K*_a units (pairs *IV* and *X*, *V* and *XI*, *VI* and *XII*).

Therefore, we tried another Free-Wilson model which uses two descriptors:

1) number of cycles/COOH position: one cycle, $\alpha(X_1^2)$, one cycle, $\beta(X_2^2)$, two cycles, $\alpha(X_3^2)$ or two cycles, $\beta(X_4^2)$

2) nature of the heteroatom: S(X_5^2), O(X_6^2) or NH(X_7^2)
the model being:

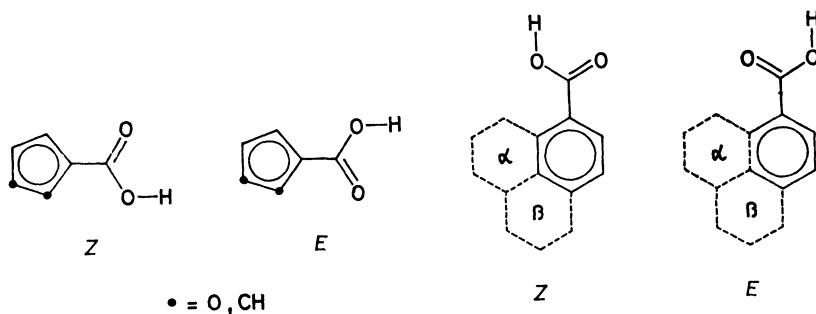
$$\text{p}K_a = a_0^2 + a_1^2(X_1^2) + a_2^2(X_2^2) + a_3^2(X_3^2) + \dots + a_7^2(X_7^2). \quad (\text{Model 2})$$

The same procedure leads to Eq. (5):

$$\begin{aligned} \text{p}K_a = & 3.96 - 0.29(X_1^2) + 0.40(X_2^2) - 0.72(X_3^2) + 0.61(X_4^2) - \\ & - 0.23(X_5^2) - 0.43(X_6^2) + 0.66(X_7^2); \quad r^2 = 0.992, r_A^2 = 0.986. \end{aligned} \quad (5)$$

This model, not only is better in a statistical sense, but it describes correctly the dependence of the annelation on the position of the carboxylic group: $a_3^2 - a_1^2 = -0.43$, $a_4^2 - a_2^2 = +0.21$. The heteroatom effect is identical in both models and independent of annelation and on the COOH position (Scheme 1). The acidity decreases from furan (-0.43) to thiophene (-0.23), and pyrrole (+0.66), following the electron release ability of the heteroatom which stabilizes the undissociated acid.

In an attempt to understand why the annelation effect increases the acidity of the carboxylic group when located in the α -position (-0.43) but decreases it when in the β -position (+0.21) we have carried out some calculations (AM1 method⁹). As we have shown previously¹¹, there is a linear relationship between p*K*_a values and the charge of the acidic hydrogen. Table II shows the heats of formation and the charges of the COOH hydrogen for the two most stable conformations of furan, benzofuran, benzene and naphthalene carboxylic acids.



SCHEME 1

Considering the averaged values of q_H for both conformations (since they are quite similar in energy), we have found a correlation of the form $\Delta pK_a = -0.17 - 234.4 \Delta q_H$; $r^2 = 0.94$, i.e. when annelation increases the charge on the hydrogen (2-benzofuran, +0.00105; α -naphthalene, +0.0016; β -naphthalene, +0.00015) the pK_a decreases (-0.37, -0.50 and -0.04, respectively), whereas when annelation decreases the charge on the hydrogen (3-benzofuran, -0.0016), the pK_a increases (+0.18).

The annelation effects we have been discussing (-0.43 and +0.21) are averaged values for thiophene, furan and pyrrole. A more detailed examination shows that these effects are larger for pyrrole-indole pairs (-0.54 and +0.28) as if the pyrrole ring could transmit better the annelation effect.

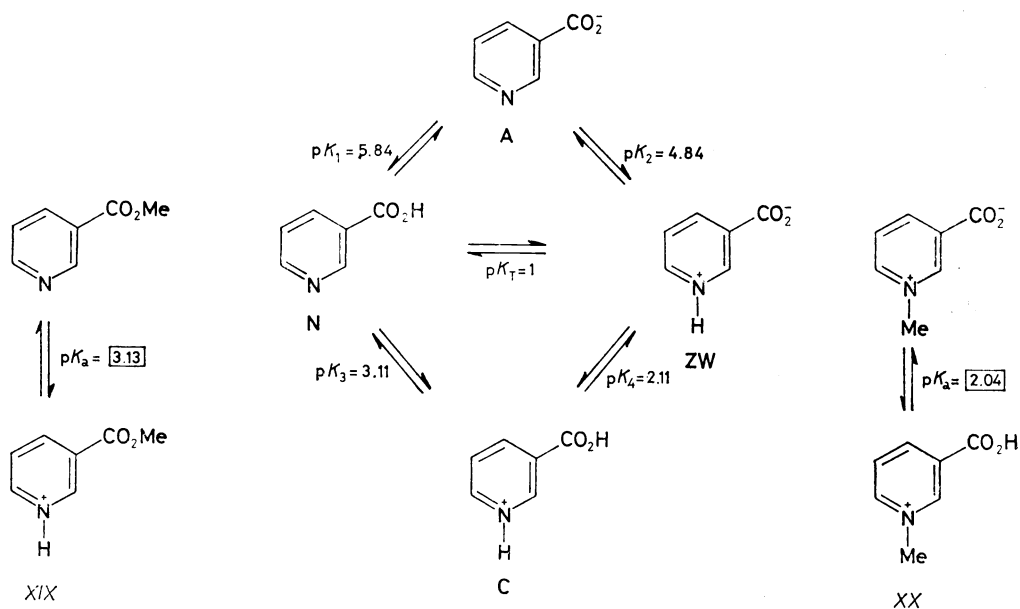
TABLE II
Calculated (AM1) heats of formation (kcal mol⁻¹) and total charges of acid hydrogens

Carboxylic acid		ΔH_f	q_H	pK_a	ΔpK_a	
2-Furan <i>II</i>	<i>Z</i>	-82.3	0.2468	3.16	-0.37	
	<i>E</i>	-81.8	0.2499			
2-Benzofuran <i>VIII</i>	<i>Z</i>	-63.5	0.2478	2.79		
	<i>E</i>	-62.9	0.2510			
3-Furan <i>V</i>	<i>Z</i>	-82.3	0.2467	4.00		+0.18
	<i>E</i>	-81.7	0.2499			
3-Benzofuran <i>XI</i>	<i>Z</i>	-69.2	0.2475	4.18		
	<i>E</i>	-68.9	0.2459			
Benzene		-67.9	0.2455	4.20		
α -Naphthalene	<i>Z</i>	-45.5	0.2471	3.70	-0.50	
	<i>E</i>	-47.1	0.2471			
β -Naphthalene	<i>Z</i>	-49.3	0.2458	4.16	-0.04 ^a	
	<i>E</i>	-49.3	0.2455			

^a $\Delta pK_a = pK_a(\beta\text{-naphthalene}) - pK_a(\text{benzene})$.

Six-Membered Heterocyclic Acids

To discuss the pK_a of compounds XIII–XVIII it is necessary to take into account the existence of a tautomer equilibrium between neutral and zwitterionic forms^{12,13}. Consider, for instance, the case of nicotinic acid XIV which is the best known¹² (Scheme 2).



$$K_{\text{exp}}^a = K_3 + K_4 ; pK_a^a = 2.07$$

$$K_{\text{exp}}^b = K_1 + K_2 ; pK_a^b = 4.80$$

SCHEME 2

The microscopic constants we have calculated (different from those reported elsewhere¹²) are in agreement with the experimental values for model compounds XIX and XX (in our hands, the pK_a of compound XX is 1.87), with the apparent values of nicotinic acid (proton gained, 2.07 (ref.¹²) and proton lost, 4.80 – Table I) and with the tautomeric constant K_T determined by UV, which corresponds to 91% of zwitterion. Since the other acids, XIII and XV, also exist at >90% in this form^{12,13} and presumably the same occurs in quinoline derivatives, XVI–XVIII, it is possible to study an additive Free-Wilson model (Model 3) for the proton loss of zwitterions (microscopic equilibrium constant K_2) using Table I values:

$$pK_a = a_0^3 + a_1^3(X_1^3) + a_2^3(X_2^3) + \dots + a_5^3(X_5^3). \quad (\text{Model 3})$$

Model 3, in which the number of cycles [one (X_1^3) or two (X_2^3)] and the position of the COOH group [$\alpha(X_3^3)$, $\beta(X_4^3)$ or $\gamma(X_5^3)$] are taken into account was treated in the usual way (Free–Wilson matrix, multiregression), leading to Eq. (6):

$$\begin{aligned} \text{p}K_a &= 4.715 + 0.115(X_1^3) - 0.115(X_2^3) + 0.27(X_3^3) - 0.17(X_4^3) - 0.10(X_5^3); \\ r^2 &= 0.887, r_A^2 = 0.718. \end{aligned} \quad (6)$$

The quality of fitness is less good in part due to the fact that we are dealing with apparent pK_a's, but also, probably, because the annelation effect is not independent of the position of the COOH group.

REFERENCES

1. Hansch C., Leo A.: *Substituent Constants for Correlation Analysis in Chemistry and Biology*. Wiley, New York 1979.
2. Cativiela C., Elguero J., Mathieu D., Meléndez E., Phan Tan Luu R.: *Eur. J. Med. Chem.* **18**, 359 (1983).
3. Cativiela C., García J. I., Elguero J., Mathieu D., Phan Tan Luu R.: *Quant. Struct. Act. Relat.* **6**, 173 (1987).
4. Palm V. A. in: *Tables of Rate and Equilibrium Constants of Heterolytic Organic Reactions*, Vol. 2, Part 1. VINITI, Moscow 1976; Serjeant E. P., Dempsey B.: *Ionization Constants of Organic Acids in Aqueous Solution*. Pergamon Press, Oxford 1979; Perrin D. D., Dempsey B., Serjeant, E. P.: *pK_a Prediction for Organic Acids and Bases*. Chapman and Hall, London 1981; Catalán J., Abboud J. L. M., Elguero J.: *Adv. Heterocycl. Chem.* **41**, 187 (1987).
5. Cativiela C., García J. I.: *Org. Prep. Proced. Int.* **18**, 283 (1986).
6. Gonzalez E., Jacquier R.: *Bull. Soc. Chim. Fr.* **1968**, 5006.
7. Gonzalez E., Alcalde E., Dinares I., Elguero J.: *Bull. Soc. Chim. Fr.* **1987**, 604.
8. Albert A., Serjeant E. P.: *The Determination of Ionization Constants*. Chapman and Hall, London 1971.
9. Dewar M. J. S., Zoebisch E. G., Healy E. F., Stewart J. J. P.: *J. Am. Chem. Soc.* **108**, 8075 (1985).
10. Free S. M., Wilson J. W.: *J. Med. Chem.* **7**, 395 (1964).
11. Catalán J., Elguero J.: *J. Chem. Soc., Perkin Trans. 2*, **1983**, 1869; Catalán J., de Paz J. L. G., Yáñez M., Elguero J.: *Chem. Scr.* **24**, 84 (1984).
12. Albert A.: *Phys. Meth. Heterocycl. Chem.* **1**, 28 (1963).
13. Katritzky A. R., Lagowski J. M.: *Adv. Heterocycl. Chem.* **1**, 435 (1963).