JOURNAL OF PHYSICAL ORGANIC CHEMISTRY
J. Phys. Org. Chem. **12**, 447–454 (1999)

Quantum similarity approach to LFER: substituent and solvent effects on the acidities of carboxylic acids

Robert Ponec,[1] Lluís Amat[2] and Ramon Carbó-Dorca [2]*

[1]Institute of Chemical Process Fundamentals, Czech Academy of Sciences, Suchdol 2, 165 02, Prague 6, Czech Republic
[2]Institute of Computational Chemistry, University of Girona, 17071 Girona, Spain

Received 2 May 1998; revised 21 September 1998; accepted 20 October 1998

ABSTRACT: Quantum similarity measures were used to estimate dissociation constants for acid–base equilibria. It is proposed that the dissociation constant of a carboxylic acid may be described by the electronic density function of the COOH group and quantified by the self-similarity measure of this fragment. In addition, using a polarized continuum model, the inclusion of surrounding medium effects on the electronic density function of the solute is easily taken into account, and in this way the solvent effect on acid dissociation constants can be simply described. A successful correlation was obtained between dissociation constants of a variety of acids in different solvents and quantum self-similarity measures of the COOH fragment. Copyright © 1999 John Wiley & Sons, Ltd.

KEYWORDS: acid dissociation constant; LFER; molecular quantum similarity measures; atomic shell approximation; promolecular densities

INTRODUCTION

There is probably no other concept that has contributed to the development of chemistry so remarkably as the ill-defined, qualitative concept of similarity and, in recent years, the broad applicability of this concept has been repeatedly reviewed.[1–5] Not despite, but rather because of, certain fuzziness, the applicability of this concept is extremely broad and covers practically all areas of chemistry. An example is the Mendeleev periodic law, the disclosure of which was closely connected with efforts to classify and systematize the similarities in the properties of elements and their simple compounds. From the intuitively understood meaning of similarity arises also one of the most powerful chemical principles—the principle of analogy—which in the early days of chemistry served as the basis for the classification and systematization of molecules and reactions. Examples of the application of this principle are such fundamental chemical concepts as the homologous series, functional group, topicity, etc., which are the cornerstones of the contemporary organic chemistry system. The same principle underlies also the widely used idea that similar structures have similar properties which, in turn, is the basis for the existence of various empirical relationships between the structure and activity known as LFER and QSAR.

Because of the fundamental role which similarity plays in so many different situations, it is not surprising that its systematic investigation has become the focus of intense scientific interest. The main attention in this respect has been devoted to the design of new quantitative measures of molecular similarity. The philosophy underlying the development of quantitative similarity measures based on quantum theory arises from the idea that properties of molecules, whether chemical, physical or biological, are predetermined by the molecular structure. Thus, the quantities characterizing this electronic structure are often used in the design of new theoretical molecular descriptors.

The aim of this paper is to propose such a new theoretical molecular descriptor replacing empirical σ constants, which in empirical LFER equations characterize the effect of systematic substitution on the rate or equilibrium of chemical reactions. It will be shown that in addition to describing correctly the substituent effect on the acidities of a series of aliphatic carboxylic acids, the new descriptor also correctly describes the effect of solvent on the dissociation equilibrium.

THEORETICAL

As already said above, the rationalization of empirical structure–activity relationships is to a considerable extent

---

*Correspondence to: R. Carbó-Dorca, Institute of Computational Chemistry, University of Girona, 17071 Girona, Spain.
E-mail: director@iqc.udg.es

Contract/grant sponsor: EC; *Contract/grant number:* ENV4-CT97-0508.
Contract/grant sponsor: CICYT; *Contract/grant number:* SAF 96-0158.
Contract/grant sponsor: Fundació Maria Francisca de Roviralta.

CCC 0894–3230/99/060447–08 $17.50

connected with recent efforts in the design of new molecular descriptors based on quantum theory. The simplest of such quantities is the electron density function $\rho(\mathbf{r})$ and most of the theoretical molecular descriptors are derived just from this quantity. Among them, a privileged place belongs to the so-called molecular quantum similarity measures (MQSM) introduced by Carbó and co-workers.[6] These measures are generally based on the pairwise comparison of electron density functions $\{\rho_A(\mathbf{r}),\rho_B(\mathbf{r})\}$ of the corresponding molecules, and the simplest form of MQSM, the so-called overlap-like measure, is defined as

$$Z_{AB} = \int \rho_A(\mathbf{r})\rho_B(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{1}$$

When the similarity integral involves the same density function ($A = B$), the corresponding quantity, Z_{AA}, is referred to as the molecular quantum self-similarity measure (MQS-SM). As was shown in previous studies,[7,8] these self-similarity measures can be successfully used as the appropriate descriptors for a great number of molecular properties.

From the theoretical point of view, numerical MQSM may be obtained using *ab initio* molecular electronic density functions constructed with the usual LCAO approximation. However, such calculations are time consuming because of the great number of four-center integrals which have to be computed during the calculation of Z_{AB}. For that reason and in order to reduce the necessary computation time, a simplifying procedure, known as the atomic shell approximation (ASA), was recently proposed.[9,10] Within this technique, the electron density function is fitted as a linear combination of spherical Gaussian functions. Further reduction of computational costs can be obtained by using another approach known as the promolecular ASA,[10] where the molecular density function is represented by a simple sum of individual atomic densities:

$$\rho_A^{\mathrm{ASA}}(\mathbf{r}) = \sum_{a\in A} P_a \rho_a^{\mathrm{ASA}}(\mathbf{r}) \tag{2}$$

where the coefficients P_a represent the total electron densities on the atom a, and the atomic density functions $\rho_a^{\mathrm{ASA}}(\mathbf{r})$are expressed as a linear combination of square normalized 1S-type GTO:

$$\rho_a^{\mathrm{ASA}}(\mathbf{r}) = \sum_{i\in a} c_i |S_i(\mathbf{r} - \mathbf{R}_a; \zeta_i)|^2 \tag{3}$$

The coefficients c_i are restricted to be positive in order to assure the physical meaning of the density. In the present work, promolecular ASA density functions are characterized by $\{c_i\}$ and $\{\zeta_i\}$ parameters, which can be found in a WWW site.[11] The promolecular ASA density functions are constructed from the three-dimensional coordinates of a molecule, centering one function on H atoms, three functions on C, N and O atoms, four functions on Cl atoms and five functions on Br atoms. In this way, an accurate and reliable representation of molecular electron densities for the calculation of MQSM can be obtained.

THE MICROSCOPIC BASIS OF QSAR

The basis of all relationships between structure and activity is a simple model, according to which the molecules are represented by vectors, whose components are various molecular descriptors characterizing some particular molecular property relevant for a given process. These descriptors, such as the hydrophobicity parameter log$P$ or the substituent constants σ, are usually chosen empirically. One of the aims of the present study is to show that the above-introduced MQSM provide the appropriate theoretical background allowing us to formulate the general theory of QSAR. Next the basic ideas of this generalization will be briefly recalled.

Let $\mathbf{M} = \{m_I\}$ be a set of n structurally related molecules and let each of them be characterized by the density function ρ_I forming the set $\mathbf{P} = \{\rho_I\}$ in one-to-one correspondence with $\mathbf{M}$. Based on these densities one can construct the $(n \times n)$ similarity matrix $\mathbf{Z} = \{Z_{IJ}\}$, whose elements are the similarity measures Z_{IJ} for all possible pairs of molecules within a given set. If we adopt an alternative view of the similarity matrix, as the n-dimensional vector $\mathbf{Z} = \{\mathbf{z}_I\}$, whose elements are the columns of the matrix $\mathbf{Z}$, then the set of the elements $\mathbf{z}_I$ can be interpreted as the n-dimensional discrete representation of the elements of the set $\mathbf{P}$ and hence of $\mathbf{M}$. This result is very important since it demonstrates that a close parallel can be found between the original QSAR approach, in which the molecules are represented as vectors, whose components are empirical molecular descriptors, and rigorous theoretical model based on MQSM. In order to make this comparison even more explicit, one can use a well known result of quantum theory, consisting in that the expectation value π_I of any physico-chemical observable attached to molecule I, and represented by the operator $\Omega(\mathbf{r})$, is given by

$$\pi_I = \int \Omega(\mathbf{r})\rho_I(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{4}$$

This integral can be also interpreted as the scalar product:

$$\pi_I = \langle \Omega | \rho_I \rangle \tag{5}$$

which in the case of discrete, rather than continuous, representation of the set $\mathbf{P} = \{\rho_I\}$ can be rewritten in the form

$$\pi_I = \mathbf{w}^{\mathrm{T}}\mathbf{z}_I \tag{6}$$

where **w** is the *n*-dimensional vector representing the operator Ω in the *n*-dimensional discrete space of MQSM vectors. Generally, the form of the operator-vector **w** is unknown but if the pairs $\{\pi_I,\mathbf{z}_I\}$ are well defined, the elements of **w** can be obtained by least-squares fitting. This procedure is the theoretical basis of QSAR, and Eqn (6) represents the general theoretical form of QSAR.[12] In some cases, this general equation can be further simplified. For instance, it is often useful to separate the self-similarity term Z_{II} from the rest, obtaining the expression

$$\pi_I = w_I Z_{II} + \sum_{J\neq I} w_J Z_{IJ} = wZ_{II} + \nu \tag{7}$$

This separation is especially useful if the studied series forms a compact set of structurally related molecules. In this latter case it is plausible to accept that the ν term is more or less constant within a series, and a simple linear relationship between the molecular property and the self-similarity measure Z_{II} can be expected to hold. In a previous study,[7] it was shown that appropriately selected self-similarity measures satisfy the above-mentioned simplifying assumptions, so that they can be used as theoretical molecular descriptors. The property studied there[7] was the dissociation constant within a series of substituted benzoic and related heteroaromatic acids. The appropriate theoretical descriptor used to characterize the substituent effect in the acid dissociation constant was the self-similarity measure:

$$Z_{XX}^{\mathrm{COOH}} = \int \rho_X^{\mathrm{COOH}}(\mathbf{r})\rho_X^{\mathrm{COOH}}(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{8}$$

Integral (8) represents the resemblance of the reaction center, in this case a COOH group, for each particular substituent *X*. The function $\rho_X^{\mathrm{COOH}}(\mathbf{r})$denotes the fragment of the total density matrix associated with the COOH group.

The purpose of this study is to show that the same similarity measure is able to act as the descriptor of substituent effect also in a series of aliphatic carboxylic acids. Moreover, it will also be shown that in addition to describing the substituent effect, the same theoretical descriptor is also able to characterize correctly the effect of the solvent on the dissociation equilibrium of these acids. In the following section the results of the present calculations are reported.

RESULTS AND DISCUSSION

The previously proposed methodology was applied to the description of the dissociation of aliphatic carboxylic acids for which the p*K* values are available in water and in several dipolar aprotic solvents.[13–15] Using these data, the reliability of the similarity approach can be tested by correlating the experimental p*K* values for each particular solvent Σ with the corresponding self-similarity indices $Z_{XX,\Sigma}^{\mathrm{COOH}}$. Within this approach, the empirical Taft equation used for the statistical treatment of data by Bartnicka *et al.*,[14]

Table 1. MQS-SM Z_{XX}^{COOH} and gas-phase acidities of carboxylic acids XCOOH[a]

X	Z_{XX}^{COOH}	ΔH_0
1 Cl_2CH	210.496	12.0
2 $ClCH_2$	211.365	19.0
3 CH_3ClCH	211.182	21.2
4 CH_3CH_2ClCH	211.199	23.1
5 $ClCH_2CH_2$	211.890	25.7
6 $CH_3ClCHCH_2$	211.995	27.7
7 $CH_3CH_2CH_2$	212.228	29.7
8 CH_3CH_2	212.177	30.6
9 $ClCH_2CH_2CH_2$	211.984	31.0
10 CH_3	212.357	31.8

[a] Values taken from Refs 23–25.

$$\mathrm{p}K_{X,\Sigma} = \rho_\Sigma\sigma_X^{*} + \mathrm{p}K_{\mathrm{H},\Sigma} \tag{9}$$

is to be replaced by its theoretical counterpart,

$$\mathrm{p}K_{X,\Sigma} = w_\Sigma Z_{XX,\Sigma}^{\mathrm{COOH}} + \nu \tag{10}$$

The solvent effect enters into the formalism via the density functions $\rho_X^{\mathrm{COOH}}(\mathbf{r})$ used for the calculation of the descriptor Z_{XX}^{COOH}. Molecular density functions are affected by placing the molecule in the solvent so that the set of descriptors $Z_{XX,\Sigma}^{\mathrm{COOH}}$ can be calculated for each particular solvent Σ. The density functions in terms of which the self-similarity measures are defined were computed using the Gaussian 94 program[16] at the *ab initio* Hartree–Fock level of theory in the 3–21G* basis set. All molecules were considered in fully optimized molecular geometries. The environmental effects of a solvent were included in the molecular Hamiltonian using a polarized continuum model (PCM).[17,18] Within this model the solute molecule is put in a cavity surrounded by a homogeneous polarizable medium, characterized by the dielectric constant ϵ. This model can be expected to describe correctly the situations where the solvent effect is dominated by non-specific electrostatic solute–solvent interactions, but owing to ingenious parametrization even solvents as complex as water can surprisingly be described. The applicability of this model for the calculation of p*K* in water has been tested several times in recent years[19–22] and it was shown that even if the model is too crude to obtain absolute p*K* values, the relative acidities in a series of structurally related acids

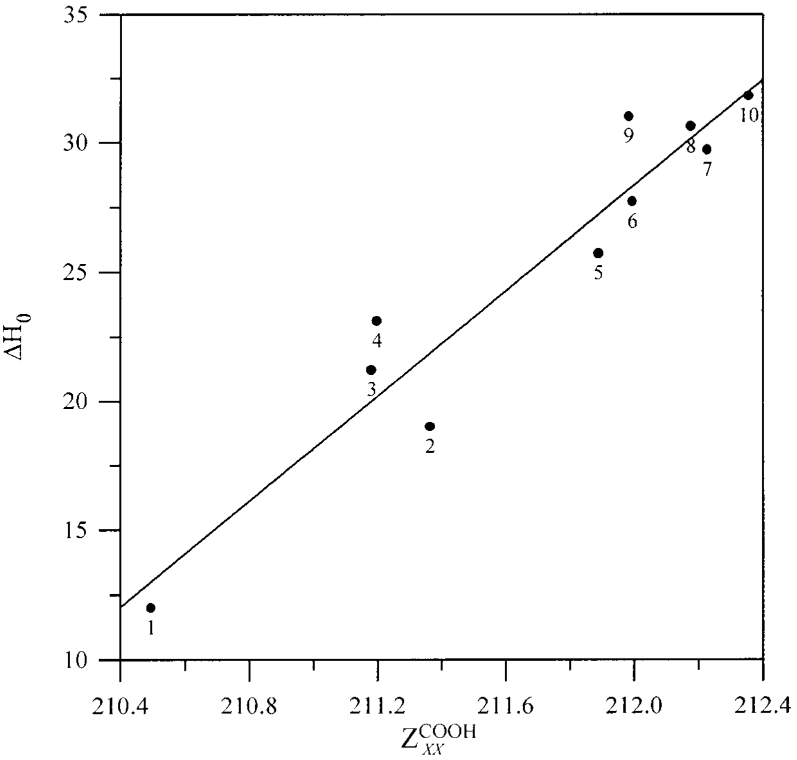

Figure 1. Dependence of calculated MQS-SM Z_{XX}^{COOH} for a series of carboxylic acids on gas-phase acidity. $r^2 = 0.915$; $n = 10$

are generally reproduced fairly well. However, this is just what is sufficient for the formulation of successful QSAR models and this is why the same approach based on the self-similarity indices, which within our similarity model are used as desriptors of substituent effect, was adopted.

Before starting the discussion of substituent and solvent effect on p*K*, it is first instructive to analyze the situation in the gas phase where the complicating solvent effects are absent. The set of acids for which the experimental gas phase acidities were available from the literature[23–25] is summarized in Table 1 and, in keeping with the philosophy of the similarity approach, these acidities were correlated with the corresponding gas-phase similarity measures. The resulting dependence is illustrated in Fig 1. The important feature of this correlation, which will be referred to in future discussions, is that a single correlation line involving both 'polar' and 'non-polar' alkyl-substituted acids is obtained. This result is especially interesting for non-polar alkyl-substituted acids where the existing correlation implies an increase in acidity with increasing length of the alkyl group. This is indeed observed in Table 1. However, this general trend is exactly the opposite of what is observed in water and the phenomenon was referred to as the anomalous gas-phase alkyl-substituent effect.[23–25] The existence of a single correlation line, which involves both 'polar' and 'non-polar' alkyl-substituted carboxylic acids, suggests that the self-similarity index Z_{XX}^{COOH} is inherently able to describe the substituent effect on the dissociation constant for all kinds of substituents. There is thus nothing unusual in the behavior of alkyl groups and their specific position, referred to as the anomalous alkyl-substituent effect, is simply the consequence of the fact that the order of gas-phase acidities is the opposite of the order of acidity in water which was known earlier.

Having presented the dependence of acidities on the self-similarity indices for the theoretically simplest case of the gas phase, it is now possible to estimate the effect of a transition from the gas phase to the solvent, and first we will discuss the substituent effect on p*K* values of carboxylic acids in water. The corresponding p*K* values and MQS-SM $Z_{XX,\text{water}}^{COOH}$ are summarized in Table 2 and the resulting dependence on the calculated self-similarity indices is given in Fig 2.

As can be seen, the situation here is more complex since instead of a single correlation line observed in the gas phase, there are in fact two separate relationships, one, with a positive slope, for 'polar' alkyl-substituted acids, and the other, with negative slope, for 'non-polar' alkyl-substituted acids. The explanation of this greater complexity is simple and there is in fact nothing unusual or surprising in it. The phenomenon which is responsible for the splitting of the correlation line into two is clearly the reversed order of acidity of alkyl-substituted acids in water. In this connection it is interesting to stress that the analogous Taft-like correlation of p*K* with empirical substituent constant σ^* is simpler and consists of only one regression line common to both 'polar' and 'non-polar' alkyl-substituted acids. This is due to the fact that

Table 2. MQS-SM $Z^{COOH}_{XX,\mathrm{water}}$ and p*K* values for carboxylic acids XCOOH in water[a]

X	$Z^{COOH}_{XX,\mathrm{water}}$	p*K*
1 Cl_3C	212.404	0.70
2 Cl_2CH	213.274	1.48
3 $CNCH_2$	213.712	2.45
4 $BrCH_2$	213.842	2.69
5 CH_3ClCH	214.343	2.83
6 $ClCH_2$	214.287	2.85
7 CH_3CH_2ClCH	214.401	2.86
8 $OHCH_2$	214.745	3.83
9 $ClCH_2CH_2$	215.143	3.98
10 $CH_3ClCHCH_2$	215.132	4.05
11 CH_2CH	215.613	4.25
12 $PhCH_2$	215.606	4.28
13 $ClCH_2CH_2CH_2$	215.543	4.52
14 CH_3	215.502	4.76
15 $CH_3CH_2CH_2$	215.546	4.83
16 CH_3CH_2	215.500	4.86
17 $(CH_3)_2CH$	215.304	4.88
18 $(CH_3)_3C$	214.758	5.03

[a] Values taken from Ref. 13.

the empirical scale of inductive constants was derived just from experimental data in water, which for alkyl-substituted acids reflect the $+I$ effect of alkyl substituents. In another words, the substituent constants σ^* are by definition chosen to be consistent with the concept of inductive effect. However, this implies that the same splitting into two separate regression lines would analogously be observed if the substituent constants were correlated with gas-phase acidity data for which the order of acidity of alkyl-substituted acids is the opposite. Consequently, it is of no importance what parameters, whether σ^* or $Z^{COOH}_{XX,\Sigma}$, are used as the descriptors of the substituent effect. The only difference is in what kind of effects are inherently described by these descriptors. Thus, whereas the classical substituent constants are chosen to reflect the order of the inductive effect, theoretically derived self-similarity indices apparently describe intrinsic 'structural' effects in which the inductive effect is only one of the components.

The present approach can be applied to other solvents in a similar manner to that reported above for water. This can easily be demonstrated by comparing the p*K* values of various alkyl-substituted carboxylic acids in several aprotic solvents, which again are consistent with the classical idea of the $+I$ effect. This implies that the same general pattern of correlation, with two separate regression lines for 'polar' and 'non-polar' alkyl-substituted acids can be expected. Representative series of such acids have been studied.[14,15] The experimental values of the corresponding dissociation constants and self-similarity indices are summarized in Tables 3 and 4, respectively. Additionally, a representative example of the correlation (10) corresponding to propylene carbonate is illustrated in Fig. 3. As can be seen, the form of this correlation is very reminiscent of what was found for water. There are again two separate correlation lines, one with a positive

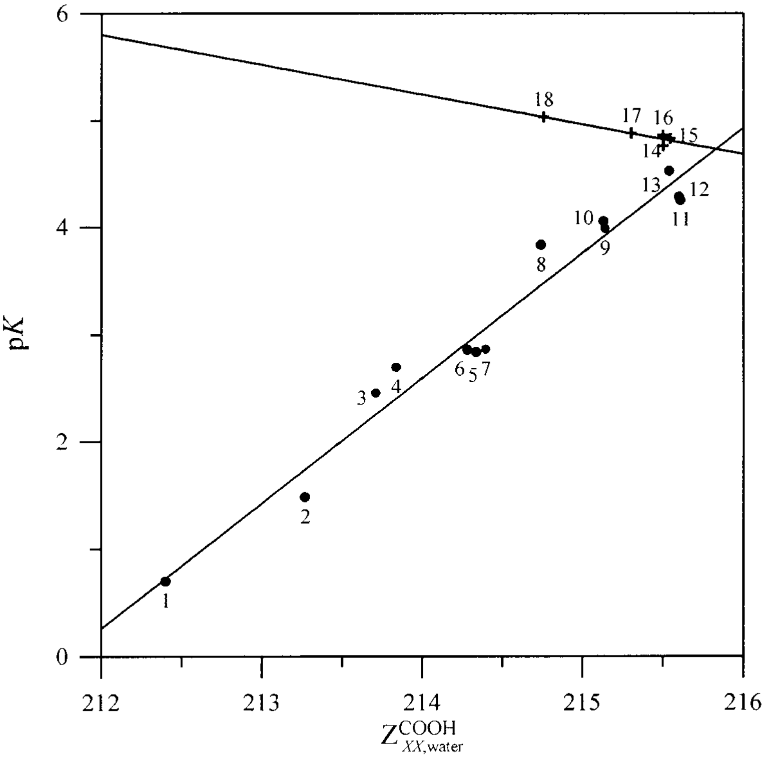

Figure 2. Dependence of calculated MQS-SM $Z^{COOH}_{XX,\mathrm{water}}$ for a series of carboxylic acids on p*K* values in water

Table 3. p*K* values for carboxylic acids XCOOH in different solvents

R	Acetonitrile[a]	Propylene carbonate[a]	Acetone[a]	*N*,*N*-Dimethylformamide[a]	Dimethyl sulfoxide[a]	Nitromethane[b]
1 Cl_3C						7.27
2 Cl_2CH	18.02	15.58	12.02	8.07	5.94	8.88
3 $CNCH_2$						10.68
4 $ClCH_2$	19.45	17.87	14.90	10.39	8.85	11.62
5 $PhCH_2$	20.73	20.03	17.17	12.93	11.69	13.80
6 CH_3	21.57	20.43	18.33	13.63	12.48	
7 CH_3CH_2	22.04	20.53	18.74	13.95	12.52	
8 $(CH_3)_2CH$	22.20	20.75	18.90	14.05	12.79	
9 $(CH_3)_3C$	22.28	20.88	19.00	14.27	12.92	

[a] From Ref. 14.
[b] From Ref. 15.

Table 4. MQS-SM $Z^{COOH}_{XX,\Sigma}$ for carboxylic acids XCOOH in different solvents

R	Acetonitrile	Propylene carbonate	Acetone	*N*,*N*-Dimethylformamide	Dimethyl sulfoxide	Nitromethane
1 Cl_3C						212.312
2 Cl_2CH	213.201	213.258	213.094	213.199	213.227	213.195
3 $CNCH_2$						213.640
4 $ClCH_2$	214.213	214.271	214.104	214.211	214.240	214.207
5 $PhCH_2$	215.502	215.584	215.350	215.500	215.540	215.494
6 CH_3	215.424	215.485	215.308	215.422	215.452	
7 CH_3CH_2	215.409	215.481	215.276	215.407	215.442	
8 $(CH_3)_2CH$	215.216	215.286	215.088	215.214	215.248	
9 $(CH_3)_3C$	214.668	214.739	214.537	214.666	214.701	

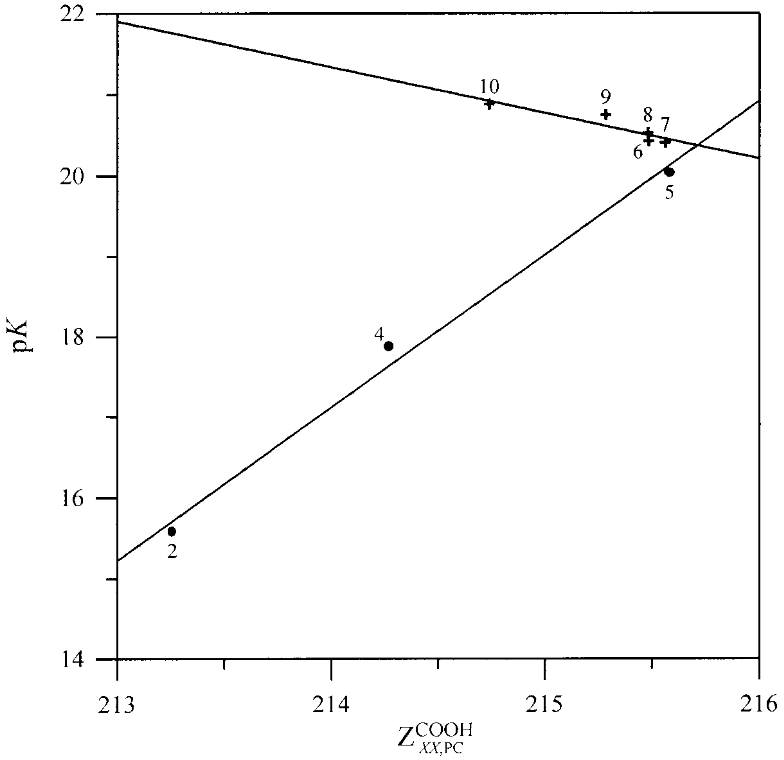

Figure 3. Dependence of calculated MQS-SM $Z^{COOH}_{XX,PC}$ for a series of carboxylic acids on p*K* values in propylene carbonate (PC)

Table 5. Calculated statistical parameters of the rectified linear relationships between MQS-SM $Z_{XX,\Sigma}^{\mathrm{COOH}}$ and p*K* values for the series of carboxylic acids in different solvents

Solvent	Type of alkyl-substituted acid	Linear equation	*n*	r^2
Water	'Polar'	$pK = 1.167Z_{XX,\Sigma}^{\mathrm{COOH}} - 247.121$	13	0.968
Acetonitrile		$pK = 1.169Z_{XX,\Sigma}^{\mathrm{COOH}} - 231.227$	3	0.990
Propylene carbonate		$pK = 1.900Z_{XX,\Sigma}^{\mathrm{COOH}} - 389.524$	3	0.992
Acetone		$pK = 2.265Z_{XX,\Sigma}^{\mathrm{COOH}} - 470.448$	3	0.984
N,*N*-Dimethylformamide		$pK = 2.106Z_{XX,\Sigma}^{\mathrm{COOH}} - 440.815$	3	0.998
Dimethyl sulfoxide		$pK = 2.472Z_{XX,\Sigma}^{\mathrm{COOH}} - 521.016$	3	0.994
Nitromethane		$pK = 2.093Z_{XX,\Sigma}^{\mathrm{COOH}} - 436.928$	5	0.979
Water	'Non-Polar'	$pK = -0.280Z_{XX,\Sigma}^{\mathrm{COOH}} + 65.067$	5	0.857
Acetonitrile		$pK = -0.618Z_{XX,\Sigma}^{\mathrm{COOH}} + 154.91$	4	0.473
Propylene carbonate		$pK = -0.521Z_{XX,\Sigma}^{\mathrm{COOH}} + 132.84$	4	0.802
Acetone		$pK = -0.610Z_{XX,\Sigma}^{\mathrm{COOH}} + 149.89$	4	0.544
N,*N*-Dimethylformamide		$pK = -0.635Z_{XX,\Sigma}^{\mathrm{COOH}} + 150.53$	4	0.712
Dimethyl sulfoxide		$pK = -0.546Z_{XX,\Sigma}^{\mathrm{COOH}} + 130.12$	4	0.821

slope for 'polar' alkyl-substituted acids and another with a negative slope for 'non-polar' alkyl-substituted acids.

The parameters of the resulting regression lines are summarized in Table 5. The most interesting information that can be extracted from this table concerns the values of the slopes, which characterize the sensitivity of the dissociation constant to substitution for each particular solvent. These values represent within the present approach the theoretical counterpart of experimental ρ constants derived from Taft-like regression analysis[14] and, therefore, it is interesting to compare both sets of data. An example of such a correlation is shown in Fig. 4 and, as can be seen, there is a fairly good linear relationship between experimental and 'theoretical' ρ

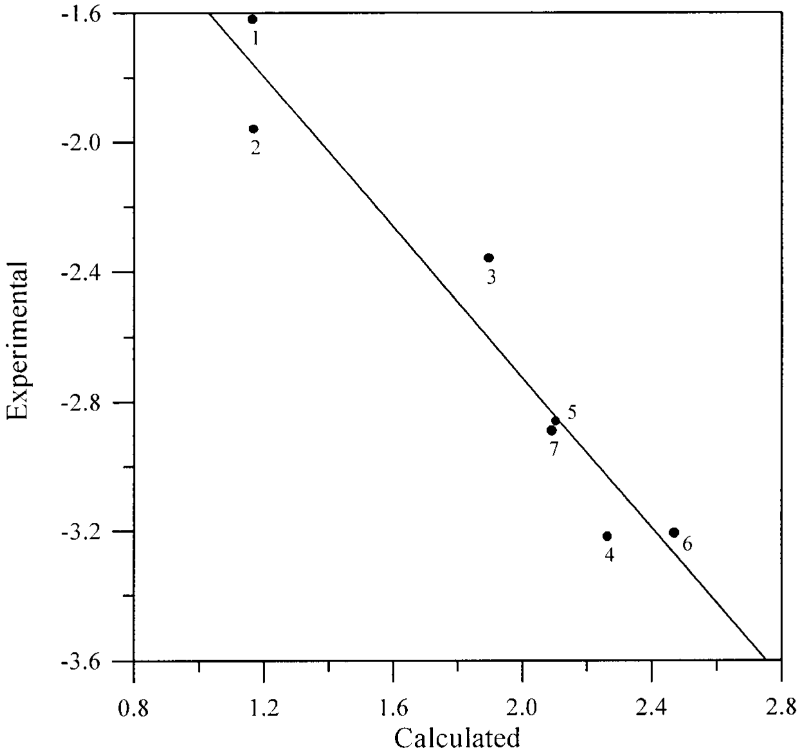

Figure 4. Correlation of calculated and experimental ρ constants from Eqns (9) and (10) for a series of 'polar' alkyl-substituted carboxylic acids in seven different solvents: (1) water; (2) acetonitrile; (3) propylene carbonate; (4) acetone; (5) *N*,*N*-dimethylformamide; (6) dimethyl sulfoxide; (7) nitromethane

constants, at least for ‘polar’ alkyl-substituted acids. The correlation for ‘non-polar’ alkyl-substituted acids is slightly poorer, but this can probably be attributed to the lower reliability of the regression parameters due to the lower correlation coefficients. This result is very interesting for at least two reasons. The first is that it provides a justification for the present use of the polarizable continuum model for the description of solvent effects. The other is of direct importance for the applicability of the similarity approach, since the reported correlations of pK with self-similarity measures clearly suggest that in systems free of complicating steric interactions, these self-similarity indices do indeed represent an appropriate theoretical parameter, allowing the description of both substituent and solvent effects even in aliphatic systems.

CONCLUSIONS

Several recent advances in MQSM have provided powerful tools for the understanding of physico-chemical properties. In this work, a quantum mechanical description using self-similarity measures applied to the prediction of substituent and solvent effects of aliphatic carboxylic acids in pK values has been successfully described. The QSPR found here are associated with the COOH group QSM fragment density. In fact, this is another application and indirect proof of the Mezey holographic density theorem,[26] which has been already used to characterize molecular chirality[27].

Acknowledgments

This work was performed during the stay of one of us (R. Ponec) at the University of Girona. This research was partially supported by a CICYT grant, SAF 96-0158, by EC grant, ENV4-CT97-0508, and by the Fundació Maria Francisca de Roviralta. The authors gratefully acknowledge all these sources of support.

REFERENCES

1. D. H. Rouvray, *Top. Curr. Chem.* **173**, 1–30 (1995).
2. R. Carbó (Ed.), *Molecular Similarity and Reactivity. From Quantum Chemical to Phenomenological Approaches*. Kluwer, Dordrecht (1995).
3. R. I. Zalewski, T. M. Krygowski and J. Shorter (Eds), *Similarity Models in Organic Chemistry, Biochemistry and Related Fields*. Elsevier, Amsterdam (1991).
4. M. A. Johnson and G. M. Maggiora (Eds), *Concepts and Applications of Molecular Similarity*. Wiley, New York (1990).
5. R. Carbó-Dorca and P. G. Mezey (Eds), *Advances in Molecular Similarity*, Vol. 1. JAI Press, Greenwich, CT (1996).
6. R. Carbó, L. Leyda and M. Arnau, *Int. J. Quantum Chem.* **17**, 1185–1189 (1980); see also E. Besalú, R. Carbó, J. Mestres and M. Solà, *Top. Curr. Chem.*, **173**, 31–62 (1995); P. G. Mezey, *Top. Curr. Chem.* **173**, 63–83 (1995); N. L. Allan and D. L. Cooper, *Top. Curr. Chem.* **173**, 85–111 (1995); R. Ponec, *Top. Curr. Chem.* **174**, 1–26 (1995).
7. R. Ponec, L. Amat and R. Carbó-Dorca, *J. Comput.-Aided Mol. Design* **13**, 259–270 (1999).
8. L. Amat, R. Carbó-Dorca and R. Ponec, *J. Comput. Chem.* **19**, 1575–1583 (1998).
9. P. Constans and R. Carbó, *J. Chem. Inf. Comput. Sci.* **35**, 1046–1053 (1995).
10. L. Amat and R. Carbó-Dorca, *J. Comput. Chem.* **18**, 2023–2039 (1997).
11. ASA coefficients and exponents can be downloaded from the WWW site http://iqc.udg.es/cat/similarity/ASA/funcset.html.
12. R. Carbó, E. Besalú, L. Amat and X. Fradera, *J. Math. Chem.* **18**, 237–246 (1995).
13. D. R. Lide (Ed.), *Handbook of Chemisty and Physics* 76th ed. CRC Press, Boca Raton, FL (1995–96).
14. H. Bartnicka, I. Bojanowska and M. K. Kalinowski, *Aust. J. Chem.* **44**, 1077–1084 (1991).
15. A. G. Kozatchenko, E. I. Matrosov and N. I. Kabatchnik, *Izv. Akad. Nauk SSSR*, *Ser. Khim.* 1476 (1976).
16. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez and J. A. Pople, *Gaussian-94 (Revision E.2)*. Gaussian, Pittsburgh, PA (1995).
17. S. Miertus, E. Scrocco and J. Tomasi, *Chem. Phys.* **55**, 117–129 (1981).
18. S. Miertus and J. Tomasi, *Chem. Phys.* **65**, 239–245 (1982).
19. G. P. Ford and B. Wang, *J. Mol. Struct. (Theochem)* **283**, 49–55 (1993).
20. J. Andzelm, Ch. Kölmel and A. Klamt, *J. Chem. Phys.* **103**, 9312–9320 (1995).
21. G. Schüürmann, M. Cossi, V. Barone and J. Tomasi, *J. Phys. Chem. A.* **102**, 6706–6712 (1998).
22. G. Schüürmann, *Quant. Struct.–Act. Relat.* **15**, 121 (1996).
23. R. Yamdagni and P. Kebarle, *J. Am. Chem. Soc.* **95**, 4050–4052 (1973).
24. K. Hiraoka, R. Yamdagni and P. Kebarle, *J. Am. Chem. Soc.* **95**, 6833–6835 (1973).
25. R. Yamdagni and P. Kebarle, *Can. J. Chem.* **52**, 861–863 (1974).
26. P. G. Mezey, *Mol. Phys.* **96**, 169–178 (1999).
27. P. G. Mezey, R. Ponec, L. Amat and R. Carbó-Dorca, *Enantiomers*, in press.