J. Org. Chem. **1994,** *59,* 8040–8046

Analysis of Solute/Solvent Interactions for the Acidity of Acetic Acids by Theoretical Descriptors

Allan D. Headley,[*,†] Stephen D. Starnes,[†] Leland Y. Wilson,[‡] and George R. Famini[§]

Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, Texas 79409-1061, Department of Chemistry, La Sierra University, Riverside, California 92515, and U.S. Army Edgewood Research, Development and Engineering Center, Aberdeen Proving Ground, Maryland 21010

Received May 23, 1994[⊗]

Quantitative structure–activity relationships (QSAR) are used now routinely to relate the biological activities and chemical properties of molecules to their molecular structural features. When theoretically determined descriptors are used in QSAR, quick and good *a priori* predictions of molecular properties can be accomplished. In this study, theoretical linear solvation energy relationship descriptors are used to model the solute/solvent interactions that influence the acidity of substituted acetic acids in different solvents. A multilinear regression (MLR) analysis approach is used to generate equations, and owing to the goodness of the fit of the different MLR equations, these descriptors account fairly well for the significant solute/solvent interactions that affect the acidity. Good agreement is obtained between the acidities predicted by this method and the experimental acidities, which are determined from potentiometric titrations, for twenty-five substituted acetic acids in seven solvents. For the acidity of substituted acetic acids where good acidity predictions were not achieved by this method, a detailed discussion is presented.

Introduction

The acidity of compounds in a particular medium is influenced not only by the structure of the molecules but also by the nature of significant solute/solvent interactions that exist. Such relationships have been used to correlate the molecular structural properties of compounds with known biological, chemical, and physical properties.[1] These relationships are often referred to as quantitative structure–activity relationships (QSAR). In cases where a specific property is examined, as in the case of this study, the term that is often used is quantitative structure–property relationships (QSPR). The success of QSAR depends on the assumption that quantitative relationships exist between microscopic features and macroscopic properties of molecules. Once a property relationship has been determined for a series of compounds, it can be used to predict that property of any compound with similar molecular features as the other compounds of the series.[2] The Hammett equation is one such relationship and it is based on the assumption that linear free energy relationships (LFER) exist between the structure and the property of compounds.[3] The ability of QSAR to make successful predictions of macroscopic properties depends strongly on the accurate quantification of microscopic molecular features. Quantified molecular features are often referred to as descriptors. Over the years, a number of descriptors have been developed,[4] but those developed by Taft and co-workers are used now routinely in most LFER.[5] The family of carboxylic acids represents one of the many families of compounds for which LFER have been used to analyze the effects of substituents on acidity variations in the gas phase and water.[6]

Of practical importance, however, is the ability to predict property variations of different molecules in various environments. To accomplish this task effectively, solute/solvent interactions must be considered. Kamlet, Taft, and co-workers have examined such interactions and have described the important ones as shown in eq 1.[7] The bulk/cavity term is a measure of the energy

$$\begin{aligned}\text{Property} = {} & \text{bulk/cavity term} + \\ & \text{dipolarity/polarizability term(s)} + \\ & \text{hydrogen bonding term(s)} + \text{constant} \qquad (1)\end{aligned}$$

that is needed to overcome the cohesive solvent/solvent interactions to form a cavity for the solute molecule. The dipolarity/polarizability terms are measures of the energies of solute/solvent dipole and induced dipole interactions that contribute to solvation. Hydrogen bonding terms measure specific interactions between solvent and solute. They reflect the ability of the solvent to accept a hydrogen bond(s) from the solute which is described as the hydrogen bond acceptor basicity (HBAB)

[†] Texas Tech University.

[‡] La Sierra University.

[§] U.S. Army Edgewood Research, Development and Engineering Center.

[⊗] Abstract published in *Advance ACS Abstracts,* November 15, 1994.

(1) Olson, E. C.; Christoffersen, R. E. *Computer-Assisted Drug Design;* ACS Symposium Series 112: American Chemical Society, Washington, DC, 1979.

(2) Gupta, S. *Chem. Rev.* **1987**, *87*, 1183.

(3) (a) Burkhart, G. N. *Nature (London)* **1935**, *17*, 684. (b) Hammett, L. P. *Chem. Rev.* **1935**, *17*, 125. (c) Hammett, L. P. *J. Am. Chem. Soc.* **1937**, *59*, 125. (d) Johnson, C. D. *The Hammett Equation*; Cambridge University Press: England, 1973. (e) Exner, O. In *Correlation Analysis of Chemical Data*, Shorter, J., Ed.; Plenum Press: New York, 1988.

(4) (a) Hansch, C. *Acc. Chem. Res.* **1962**, *B2*, 232. (b) Koppel, I. A.; Palm, V. A. In *Advances in Linear Free Energy Relationships*; Chapman, N. B., Shorter, J. Eds.; Plenum Press: London, 1972. (c) Charton, M. *Prog. Phys. Org. Chem.* **1981**, *13*, 119. (d) Reynolds, W. F. *Prog. Phys. Org. Chem.* **1983**, *14*, 165.

(5) (a) Taft, R. W. In *Steric Effects in Organic Chemistry*; Newman, M. S., Ed.; Wiley and Sons: New York, 1956. (b) Fujita, T.; Takayama, C.; Nakajima, M. *J. Org. Chem.* **1973**, *38*, 1623. (c) Kamlet, M. J.; Taft, R. W.; Abboud, J.-L. M. *J. Am. Chem. Soc.* **1977**, *92*, 8325. (d) Taft, R. W.; Abboud, J.-M. L.; Anvia, F.; Berthelot, M.; Fujio, M.; Gal, J.-F.; Headley, A. D.; Henderson, W. G.; Koppel, I.; Qian, J. H.; Mishima, M.; Taagepera, M.; Uiji, S. *J. Am. Chem. Soc.* **1988**, *110*, 1979.

(6) (a) Jinfeng, C.; Topsom, R. D. *J. Mol. Struct. (Theochem.)* **1989**, *188*, 45. (b) Topsom, R. D. *Prog. Phys. Org. Chem.* **1990**, *17*, 107. (c) Cai, J.; Topsom, R. D.; Headley, A. D.; Koppel, I.; Mishima, M.; Taft, R. W.; Ueji, S. *J. Mol. Struct. (Theochem.)* **1988**, *168*, 141.

(7) (a) Kamlet, M. J.; Doherty, R. M.; Abraham, M. H.; Taft, R. W. *Quant. Struct.-Act. Relat.* **1988**, *7*, 71. (b) Mishima, M.; McIver, R. T., Jr.; Taft, R. W.; Bordwell, F. G.; Olstead, W. N. *J. Am. Chem. Soc.* **1984**, *106*, 2717.

0022-3263/94/1959-8040$04.50/0

and the ability of the solvent to donate a hydrogen bond(s) to the solute which is described as the hydrogen bond donor acidity (HBDA). The first term in eq 1 is endoergic, whereas the last two terms are exoergic. For eq 1, linear statistical fitting between the variables is used. This approach is not the only form of statistical fitting that can be used, but numerous QSAR analyses have been carried out successfully by this approach. Descriptors[8] developed from thermodynamic and spectroscopic data for solvents and solutes are used commonly in eq 1, and they have been used successfully to correlate the chemical, physical, and biological properties of more than 200 compounds by this approach.[9] For multilinear relationships (MLR) where solute/solvent interactions are considered, the term that is often used to describe the relationships is linear solvation energy relationships (LSER).

One disadvantage in the use of LSER and other classical approaches for the analysis of solute/solvent interactions is that the descriptors used to analyze property variations are empirically developed. Thus, unusual interactions may be misinterpreted. For example, the ability of solvent molecules to access adequately the region between dipoles or charges that are in close proximity to each other in a molecule is not reflected by the solvent solvatochromic parameters.[10] One way to circumvent this problem is to use descriptors in QSAR that are developed from theoretical chemistry. Molecular orbital calculations have been used for the development of molecular descriptors.[11] Statistically-based interaction indices derived from molecular surface electrostatic potentials have been used also to predict the properties of molecules.[12] In addition, descriptors that are obtained by computational methods are reliable and they are obtained quickly.[13] Two of us (G.R.F. and L.Y.W.) have developed a set of six descriptors solely from molecular orbital calculations. These theoretical linear solvation energy relationship (TLSER) descriptors have been developed for a wide variety of compounds,[14] and they have been used successfully to correlate the properties of the following: five nonspecific toxins,[15] the activity of some local anesthetics and their molecular transform,[16] opiate activity of some fentanyl-like compounds,[17] and six physicochemical properties (charcoal on absorption, HPLC retention index, octanol−water partition coefficients, phosphonothiolate hydrolysis rate constants, aqueous equilibrium constants, and electronic absorption of some ylides).[18] They have been used successfully also to describe the effects of structural variations on the acidity of carboxylic acids, alcohols, silanols, anilines, hydrocarbons, and oximes in the gas phase.[19] Since these TLSER parameters are determined solely from computational methods, *a priori* predictions of the properties of compounds are possible. The TLSER descriptors were developed to correlate closely with LSER descriptors, to give MLR equations with correlation coefficients, *R*, and standard deviations (SD) close to those for LSER, and to be as widely applicable to solute/solvent interactions as the LSER set. The TLSER descriptors are developed to represent specific electronic properties of molecules in the gas phase. Since the electronic interactions between molecules depend only on the nature of the molecules, these descriptors also depict the interactions between two molecules—a solute and a solvent. Tables 1 and 2 give the summary of the TLSER descriptors for the compounds that are used in this paper.

Table 1. Theoretical Descriptors for Substituted Acetic Acids XCH_2COOH

no.	X	V_{mc}	π_i	ϵ_B	q_-	ϵ_A	q_+
1	H	50.15	0.1000	0.1298	0.3651	0.1696	0.2161
3	CH_3	65.83	0.1050	0.1306	0.3661	0.1691	0.2160
3	C_2H_5	82.10	0.1073	0.1307	0.3662	0.1691	0.2158
4	C_3H_7	101.82	0.1051	0.1308	0.3662	0.1691	0.2158
5	*i*-C_3H_7	102.31	0.1041	0.1308	0.3668	0.1691	0.2157
6	C_6H_{11}	140.74	0.1110	0.1312	0.3666	0.1691	0.2157
7	*t*-C_4H_9	117.97	0.1058	0.1310	0.3663	0.1692	0.2158
8	$Si(CH_3)_3$	125.48	0.1024	0.1323	0.3778	0.1697	0.2148
9	Ada	189.50	0.1145	0.1327	0.3667	0.1690	0.2155
10	SCH_3	84.23	0.1016	0.1462	0.3519	0.1731	0.2190
11	OCH_3	73.42	0.1054	0.1347	0.3645	0.1700	0.2170
12	C_6F_5	120.66	0.1233	0.1499	0.3668	0.1785	0.2155
13	Cl	64.98	0.1068	0.1269	0.3396	0.1801	0.2231
14	$N(CH_3)_2$	96.01	0.1080	0.1466	0.4480	0.1681	0.2143
15	$Cl(CH_2)_3$	100.65	0.1069	0.1284	0.3591	0.1729	0.2188
16	CH_3CHCl	115.55	0.1093	0.1293	0.3650	0.1730	0.2174
17	C_6H_5	119.39	0.1246	0.1499	0.3668	0.1785	0.2155
18	$Cl(CH_2)_2$	81.96	0.1080	0.1270	0.3635	0.1730	0.2206
19	F	52.46	0.0975	0.1272	0.3547	0.1736	0.2212
20	CF_3	78.17	0.0888	0.1229	0.3308	0.1781	0.2278
21	Br	71.16	0.1141	0.1287	0.2944	0.1820	0.2219
22	CN	63.89	0.1073	0.1260	0.3452	0.1743	0.2233
23	$CH{=}CH_2$	77.17	0.1098	0.1422	0.3672	0.1715	0.2154
24	SC_6H_5	137.41	0.1226	0.1495	0.3513	0.1813	0.2187
25	OC_6H_5	124.58	0.1287	0.1507	0.3555	0.1791	0.2175

Table 2. Hildebrand Parameters and the Theoretical Descriptors for Selected Solvents

no.	X	δ_H	π_i	ϵ_B	q_-	ϵ_A	q_+
26	H_2O	40.4	0.0670	0.1237	0.3256	0.1237	0.1628
27	MeOH	24.0	0.0860	0.1314	0.3291	0.1402	0.1803
28	EtOH	20.0	0.0927	0.1326	0.3235	0.1429	0.1800
29	2-PrOH	16.0	0.0962	0.1335	0.2354	0.1461	0.1802
30	DME	6.8	0.1076	0.1358	0.3550	0.1478	0.0124
31	EG	6.8	0.0953	0.1331	0.3306	0.1444	0.1842
32	*t*-BuOH	10.2	0.1034	0.1334	0.3180	0.1442	0.1764

The generalized TLSER equation for solutes in a given solvent is shown in eq 2:

---

(8) (a) Taft, R. W.; Abboud, J.-M. L.; Kamlet, M. J.; Abraham, M. H. *J. Solution Chem.* **1985**, *14*, 153. (b) Abrahams, M. H.; Grellier, P. L.; Prior, D. V.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans. 2* **1990**, 521.

(9) Kamlet, M. J.; Taft, R. W.; Famini, G. R.; Doherty, R. M. *Acta Chem. Scand.* **1987**, *41*, 589.

(10) (a) Headley, A. D. *J. Chem. Soc., Perkin Trans. 2* **1989**, 457. (b) Headley, A. D.; McMurry, M. E. *J. Phys. Org. Chem.* **1994**, *7*, 63.

(11) (a) Loew, G. H.; Poulsen, M.; Kirkjian, E.; Ferreil, J.; Sudhindra, B. S.; Rebagliati, M. *Environ. Health Perspect.* **1985**, *61*, 69. (b) Pedersen, L. *Environ. Heath Perspect.* **1985**, *61*, 185. (c) Chastrette, M.; Raizmann, M.; Chanon, M.; Purcell, K. F. *J. Am. Chem. Soc.* **1985**, *107*, 1. (d) Lewis, D. F. V. In *Progress in Drug Metabolism*; Bridges, J. W., Chasseaud, L. F., Eds.; John Wiley: London, 1990; p 205. (e) Kier, L. B.; Hall, L. *Molecular Connectivity in Structure Activity Analyses*, Research Studies Press: Letchworth, 1986.

(12) (a) Politzer, P.; Daiker, K. C. In *The Force Concept in Chemistry*; Deb, B. M., Ed.; Van Nostrand Reinhold Co: New York, 1981; Chapter 6. (b) Sjoberg, P.; Murray, J. S.; Brinck, T.; Potlizer, *Can J. Chem.* **1990**, *68*, 1440. (c) Potlizer, P.; Murray, J. S. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1991, Chapter 7. (d) Brinck, T.; Murray, J. S.; Potlizer, P. *J. Org. Chem.* **1991**, *56*, 5012.

(13) Ford, M. G.; Livingstone, D. J. L. *Quant. Struct.-Act. Relat.* **1990**, *9*, 107.

(14) Famini, G. R. *Using Theoretical Descriptors in Quantitative Structure Activity Relationships*; V. CRDEC-TR-085, U.S. Army Chemical Research, Development and Engineering Center, Aberdeen Proving Ground, MD, 1989.

(15) Wilson, L. Y.; Famini, G. R. *J. Med. Chem.* **1991**, *34*, 1668.

(16) Famini, G. R.; Kassel, R. J.; King, J. W.; Wilson, L. Y. *Quant. Struct.-Act. Relat.* **1991**, *10*, 344.

(17) Famini, G. R.; Famini, W. P.; Ashman, A. P.; Mickiewicz, A. P.; Wilson, L. Y. *Quant. Struct.-Act. Relat.* **1992**, *5*, 395.

(18) Famini, G. R.; Penski, C. E.; Wilson, L. Y. *J. Phys, Org. Chem.* **1992**, *5*, 395.

(19) Famini, G. R.; Marquez, B. C.; Wilson, L. Y. *J. Chem. Soc. Perkin Trans. 2* **1993**, 773.

$$\mathrm{SSP} = aV_{\mathrm{mc}} + b\pi_{\mathrm{i}} + c\epsilon_{\mathrm{B}} + dq_{-} + e\epsilon_{\mathrm{A}} + fq_{+} + \mathrm{SSP}_0 \tag{2}$$

In this study, SSP represents the solute/solvent interactions that cause acidity variations for different substituted acetic acids. V_{mc} describes the molecular van der Waals volume (in units of cubic angstroms, Å^3). π_{i} describes the dipolarity/polarizability contribution and is obtained from the division of the polarizability volume by the molecular volume to produce a unitless, size independent quantity which indicates the ease with which the electron cloud of a solute may be moved or polarized. ϵ_{B} is part of the hydrogen bond acceptor basicity (HBAB) contribution and is the energy difference between the highest occupied molecular orbital (HOMO) of the solute and the lowest unoccupied molecular orbital of (LUMO) water. Water was chosen as the reference because it is the most common solvent. The electrostatic term (q_{-}) is the largest negative formal charge on an atom of the solute; the units are atomic charge units (acu). ϵ_{A} describes the covalent acidity, which reflects the ability of a solute to act as a Lewis acid. These descriptors are obtained from the difference between the energies of the LUMO of the solute and the HOMO of water. q_{+} is the electrostatic acidity term and is the largest positive formal charge on an atom of the solute; the units are in acu. SSP_0 is the intercept, and for the acidity of substituted acetic acids; it represents the acidity of the unsubstituted acid (acetic acid) if all solute/solvent interactions were absent. The coefficients of the MLR equations, *a*, *b*, *c*, *d*, *e*, and *f*, indicate the significance of the different solute/solvent interactions to the property being analyzed. In this case, contributions that increase the acidity result in negative coefficients, whereas contributions that decrease the acidity result in positive coefficients. Once eq 2 has been determined from an appropriate data set, the property of any compound, that is similar to those used to generate the data set, in any solvent can be predicted. A major advantage of this method for the determination of properties of compounds is the low cost. The EPA estimates that the cost to fully understand, through experimentation, most properties of a single molecule exceeds \$100,000.[20] Owing to the relatively low cost to carry out computations, accurate estimates of the properties of most compounds can be achieved by this method at a fraction of the cost of conventional methods.

For the acidity of substituted acetic acids, various solute/solvent interactions contribute to the acidity. Thus, in order to predict effectively the acidity of acetic acids in different media, the identification and contribution that significant solute/solvent interactions make to acidity variations must be accomplished. In this paper, equations of the format shown in eq 2 are generated, in which TLSER descriptors are used, to model the solute/solvent interactions that affect the acidity variations of a wide variety of substituted acetic acids in different solvents. The quality of the model is determined by the statistical fit of the MLR equations. On the basis of the magnitude of the coefficients of the equations, the significant solute/solvent interactions that contribute to the acidity variation in each solvent are identified. The source and nature of these interactions are discussed.

Experimental Section

Experimental acidities were determined from potentiometric titrations in solutions of less that 10^{-3} M in 0.1 M tetrabutylammonium bromide in order to maintain a constant ionic strength. The procedure for the determinations is described elsewhere.[21] pK_{a} values shown in Table 2 were determined to within ±0.1 p*K* unit. The TLSER descriptors shown in Tables 1 and 2 were computed using the MNDO algorithm contained in MOPAC, version 6.0.[22] The molecular volume of the optimized geometry was determined using the algorithim of Hopfinger.[23] ϵ_{A} and ϵ_{B} values that are shown in Table 1 and are used for the MLR analysis are relative to water. The Minitab multilinear regression analysis computer program was used to obtain the correlation equations.[24]

Results and Discussion

The TLSER descriptors that are used in this study are shown in Table 1 (solutes) and Table 2 (solvents). Table 3 shows the experimental acidity (along with predicted acidity) of the various substituted acetic acids in the gas phase[25] and in solution. The predicted acidities are obtained from MLR correlation equations of the format shown in eq 2. Table 4 shows the coefficients and the statistics of the TLSER correlation equations that are used to predict the acidities of the substituted acetic acids in the different media shown in Table 3. The gas phase acidity values used for this correlation are free energy terms and are in kcal/mol. The absolute acidity of acetic acid in methanol was used to establish the relative acidity scale for other acetic acids in dimethoxyethane (DME) and ethylene glycol (EG). Acidity measurements in 2-methyl-2-propanol were carried out at a different temperature (303 K) than the temperature of the acidity in the other solvents (298 K). For the TLSER equations, good correlation coefficients are obtained ($0.960 < R < 0.995$). Not all terms of the equations are significant—the terms that are retained are at the 0.95 level or higher. Each coefficient is accompanied by its standard error (±), *t*-statistic (*t*-stat), and the variance inflation factor (VIF) to indicate the quality of the "fit" and the degree of cross correlation of the independent variables. For each regression (solution phase), methoxyacetic acid (compound **11**) was found to be an outlier and its acidity value was therefore omitted from the entries shown in Table 4. A total of nine regression equations are developed, one for each medium.

Gas Phase Acidity. The trend of the gas phase acidity data shown in Table 3 implies that the acidity of substituted acetic acids increases as the size of the substituents increases, i.e., *tert*-butylacetic acid is a stronger acid than acetic acid. The accepted rationale for this trend is that in the gas phase, charge-induced stabilization[26] of the conjugate base is gained from nearby substituents and as a result, acids with bulky alkyl substituents are more acidic than those with less bulky substituents. Similar observations are made for the

---

[20] Borman, S. *Chem. Eng. News* **1990**, *68*(No. 8), 20.

[21] Headley, A. D.; McMurry, M. E.; Starnes, S. D. *J. Org. Chem.* **1994**, *59*, 1863.

[22] (a) Dewer, M. J. K.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899. (b) Stewart, J. J. P. *Mopac Manual*, FJSRL-TR-88-PP7. (c) Seiler, F. J. Research Laboratory, U.S. Air Force Academy, Colorado Springs, CO, 1988.

[23] Hopfinger, A. J. *J. Am. Chem. Soc.* **1980**, *102*, 7126.

[24] The program is marketed by Minital, State College, PA.

[25] Gas phase data were obtained from Professor J. Bartmess, Department of Chemistry, University of Tennessee, Knoxville, TN 37996-1600.

[26] Hehre, W. J.; Pau, C.-F.; Headley, A. D.; Taft, R. W. *J. Am. Chem. Soc.* **1986**, *108*, 1711.

Table 3. Experimental and Predicted (in Parentheses) pK_a[a] of Substituted Acetic Acids (XCH_2COOH) in Various Solvents at 298 K

no.	X	gas	H_2O	MeOH	EtOH	2-PrOH	DME	EG	*t*-BuOH[b]
1	H	341.1 (340.9)	4.75 (4.72)	9.7 (9.8)	10.3 (10.4)	11.3 (11.6)	9.7 (9.6)	9.7 (9.7)	14.2 (14.6)
2	CH_3	340.3 (339.9)	4.87 (4.76)	10.1 (10.0)	10.6 (10.6)	12.2 (11.9)	9.9 (9.9)	10.0 (9.9)	15.2 (14.9)
3	C_2H_5	339.5 (339.8)	4.82 (4.76)	10.2 (10.0)	10.7 (10.7)	12.2 (12.0)	10.2 (10.0)	10.1 (10.0)	15.2 (15.2)
4	C_3H_7	339.3 (339.4)	4.86 (4.76)	10.1 (10.0)	10.7 (10.8)	12.3 (12.2)	10.3 (10.2)	10.1 (10.1)	15.6 (15.5)
5	*i*-C_3H_7		4.78 (4.80)	10.2 (10.1)	10.8 (10.8)	12.3 (12.2)	10.3 (10.2)	10.3 (10.3)	15.6 (15.5)
6	C_6H_{11}		4.82 (4.78)	10.4 (10.2)	10.9 (11.0)	12.5 (12.6)	10.3 (10.6)	10.3 (10.2)	15.6 (16.1)
7	*t*-C_4H_9	338.1 (338.6)	5.04 (4.76)	10.4 (10.3)	11.2 (10.9)	12.6 (12.3)	10.7 (10.3)	10.3 (10.2)	16.5 (15.7)
8	$Si(CH_3)_3$		5.22 (5.51)	10.5 (10.4)	11.2 (11.0)		10.5 (10.3)		16.6 (16.3)
9	ADA			10.3 (10.5)	11.2 (10.9)	13.0 (13.1)	10.8 (11.0)		16.9 (16.9)
10	SCH_3		3.66 (3.32)	8.9 (9.2)	9.5 (9.7)	10.6 (11.0)	8.8 (9.1)	9.0 (9.1)	13.2 (13.8)
11	OCH_3	335.0 (335.3)	3.5	8.6	9.2	10.2	8.3	8.5	12.9
12	C_6F_5			8.6 (8.9)	8.8 (9.0)	9.4 (10.1)	7.8 (8.4)	8.3 (8.9)	11.9
13	Cl	328.8 (329.4)	2.77 (3.05)	7.8 (7.8)	8.2 (8.0)	9.2 (9.0)	7.3 (7.4)	7.7 (7.7)	12.2 (11.5)
14	$N(CH_3)_2$		9.88 (9.90)	12.1 (12.2)	12.1 (12.1)	12.8 (12.8)			
17	C_6H_5	334.1 (334.0)	4.28 (4.22)	9.5 (8.9)	10.2 (10.2)	10.8 (10.1)		9.4 (8.9)	14.0 (13.9)
23	$CH{=}CH_2$		4.34 (4.49)	9.5 (9.7)	10.2 (10.2)	11.2 (11.4)		9.7 (9.7)	14.2 (14.7)
24	SC_6H_5					9.9 (9.5)	8.6 (7.9)	8.4 (8.2)	12.8 (12.9)
25	OC_6H_5		3.17 (3.42)			9.5 (9.9)	8.3 (8.3)	8.3 (8.9)	12.4 (13.2)

[a] DME and EG are relative values, the same pK_a scale as methanol was used for acidity determination. [b] Experimental determinations of pK_a were done at 30 °C.

Table 4. Coefficients and Statistics for the MLR for the Individual Sets. $pK_a = aV_{mc} + b\pi_i + c\epsilon_B + dq_- + e\epsilon_A + fq_+ + g$

coeff. ± *t*-stat. VIF		*a*	*b*	*c*	*d*	*e*	*f*	*g*	*N*	*R*	*F*	*s*
3	gas	n/s	−273.54 79.35 3.84 1.3	n/s	n/s	n/s	−1504.1 161.90 9.29 1.3	665.33	14	0.960	65	1.434
4	H_2O	n/s	n/s	−30.221 7.05 4.28 1.3	68.693 2.48 27.66 1.3	n/s	n/s	−16.441	14	0.993	389	0.209
5	MeOH	0.446 0.20 2.26 1.0	n/s	n/s	23.636 3.27 7.24 1.0	−135.08 18.56 7.28 1.2	n/s	23.857	15	0.981	64	0.258
6	EtOH	0.509 0.14 3.62 1.0	n/s	n/s	13.661 2.32 5.90 1.3	−204.98 15.14 13.54 1.3	n/s	39.944	15	0.988	138	0.180
7	2-PrOH	0.975 0.29 3.33 1.0	n/s	n/s	4.848 5.00 0.97 1.3	−244.13 23.60 10.35 1.3	n/s	50.731	16	0.963	51	0.393
8	EG	0.569 0.30 1.91 1.4	n/s	n/s	23.890 12.33 1.94 2.1	−143.26 22.37 6.4 2.3	n/s	25.030	14	0.964	43	0.263
9	*t*-BuOH	1.151 0.37 4.09 1.2	n/s	n/s	44.25 21.31 2.08 2.5	−203.60 40.71 5.00 2.3	n/s	32.18	15	0.960	43	0.345
10	DME	0.921 0.26 3.55 1.0 1.0	n/s	n/s	n/s	−216.53 19.53 11.09 1.0 1.0	n/s	45.900	14	0.961	67	0.345

acidities of other compounds in the gas phase.[27] For nonalkyl acetic acids, LFER show that both inductive and polarizability effects dictate acidity variations.[28] An alternate approach for the analysis of how structural variations affect the acidity of acetic acids is by the use of TLSER descriptors. Entry 3 of Table 4 shows the coefficients and statistics of the MLR equation for the acidity variations in the gas phase.[29] On the basis of the standard deviation, *R*, and *F*, the TLSER descriptors do not model the acidity variations as well as desired, but they are good enough to gain useful interpretation about how structural variations affect the acidity of acetic acids.

From entry 3 of Table 4, the two terms that are significant to acidity variations are the dipolarity/polarizability and electrostatic acidity terms. The dipolarity/polarizability effect describes how well the electron cloud

---

[27] (a) Bartmess, J. F.; Scott, J. A.; McIver, R. T., Jr. *J. Am. Chem. Soc.* **1979**, *101*, 6046. (b) Brauman, J. I.; Blair, L. K. *J. Am. Chem. Soc.* **1968**, *90*, 5636. (c) McIver, R. T., Jr.; Silvers, J. H. *J. Am. Chem. Soc.* **1973**, *95*, 8462. (d) Graul, S. T.; Schnute, M. E.; Squires, R. R. *Int. J. Mass. Spec., Ion Processes* **1990**, *96*, 181. (e) Aue, D. H.; Bowers, M. J. *Gas Phase Chemistry*; Academic Press: New York, 1979.

[28] Taft, R. W.; Topsom, R. D. *Prog. Phys. Org. Chem.* **1987**, *16*, 1.

[29] The compounds used for this correlation are **1**−**7**, **11**, **13**, **17**, **18**, **19**, and **20**.

of the acid can be moved. For acids that are large and polarizable, such ease in electron movement serves to stabilize the most negative region of the acid. The sign for this coefficient is negative, which means that this contribution favors the dissociation of the acid. This observation is consistent with that gained from the LFER analyses.[6] The electrostatic acidity term (q_+) is the other significant contribution to acidity. This term describes the most positive region on an acetic acid molecule, and owing the to polarity of the O−H bond, the acidic hydrogen of the carboxylic acid functionality is the most positive atom. Electronegative substituents, such as the Cl atom, increase the positive character of this hydrogen and as a result increase the acidity of the molecule. An increase in the partial positive character of this hydrogen will favor deprotonation. As a result, the coefficient for this contribution is negative. From the TLSER analysis, these two contributions are important to acidity variations, but entries 4 through 10 of Table 4 indicate that in the condensed phase they are overwhelmed by different solute/solvent interactions.

Aqueous Acidity. For the LSER approach in the analysis of the structural effects on the acidity of acetic acids in different solvents, the solvation effects on the equilibrium species are examined.[21] An alternate approach, as is done in this study, is the examination of the solute/solvent interactions on the undissociated acetic acid. Such interactions will give an indication of the acidity of the acetic acid. Thus, the acidity of different substituted acetic acids in water can be predicted from significant solute/solvent interactions of the neutral acetic acid. In water, the nature of the interactions depend on the properties of water. Water has an exceptional ability to form hydrogen bonds with solutes possessing at least one unshared pair of electrons and with solutes possessing acidic hydrogen(s). The solvatochromic parameters,[30] the Lewis acidity and basicity parameters,[31] and the TLSER descriptors (Table 2) have been used to quantify these abilities. Water also has a fairly large dielectric constant,[32] and as a result, charged or dipolar solutes in water exist as stable solvated species. On the basis of the quality of the fit of entry 4 of Table 4, the TLSER descriptors model very well the solute/aqueous interactions and their effects on the ability of acetic acids to dissociate. From the magnitude of the coefficients of this equation, the important solute/solvent contributions to acidity variations come from the solute/solvent interaction caused by the covalent basicity and the electrostatic basicity of the acids.

The TLSER covalent basicity descriptor reflects the ability of solutes to donate a pair of electrons to their environment. For compounds that have the ability to donate at least one pair of electrons to their environment, the potential of those compounds (in such environments) to lose a positive species, such as a proton, is increased. Thus, for an acetic acid (a neutral molecule), if the electron density is shifted to water, the loss of the most positive species, here a proton, is likely to occur in order to maintain neutrality. This interaction will dictate the acidity since it dictates the ability to lose a proton. The increase in acidity by this contribution is reflected by the negative sign of the coefficient *c* in entry 4 of Table 4. Compared to another solvent for which the energy difference between the HOMO (of the solvent) and the LUMO (of the water) is greater than that of water, the potential for an acetic acid to lose a proton to the solvent is not as great as it is in water. As will be shown in the next sections, this contribution to acidity variations is not significant in the other solvents.

The other important solute/solvent interaction that affects the ability of acetic acid to dissociate is caused by the electrostatic basicity of the acetic acids, q_-. This descriptor represents the most negative atom of each acetic acid. For alkyl-substituted acetic acids, the oxygens of the carboxylic acid functionality are the most negative atoms. From Table 1, the magnitude of the values for q_- increases as the bulkiness of the group increases. This trend demonstrates that alkyl groups are indeed electron releasing. Thus, the partial negative charge that resides on the oxygen atom of the polarized O−H bond is increased as the bulkiness of the acid increases. As a result, the O−H bond is a stronger bond for bulky alkyl acids compared to less bulky acids. For acids with polar substituents, such as $ClCH_2COOH$, the partial negative charge is dispersed into the electronegative regions of the acid. As a result, such acids are more acidic compared to alkylacetic acids. In water, any structural variations that cause an increase in the negative partial charge on the oxygen atoms of acetic acids (i.e. larger q_- values) will decrease the acidity, and as a result, the sign for this coefficient is positive.

For the derivation of the correlation equation (coefficients are shown in entry 4 of Table 4), methoxyacetic acid was not included. The "fit" shown for that equation was not as good if methoxyacetic acid were included. *R*, *F*, and *s* are 0.980, 143, and 0.349, respectively. One implication is that compared to the other compounds of that data set, the outlier has an unusual feature. As discussed above, a significant contribution to the acidity comes from the covalent basicity interaction of the acids with the solvent. For methoxyacetic acid, the additional unshared pairs of electrons of the methoxy group introduce additional sites for interaction with the solvent, which in turn will affect the acidity of the molecule. For each acid, there is only a single descriptor for the Lewis basicity, which accounts for the overall Lewis basicity—the methoxy group and the carboxylic acid functionality for methoxyacetic acid. Since methoxyacetic acid is an outlier, the separate interaction of both sites with the solvent cannot be accounted for by a single parameter (ϵ_B). Methylthioacetic acid, which has a similar structural feature as methoxyacetic acid, was used for the derivation of the correlation equation, and as shown in Table 3, the difference between predicted and experimental values is fairly large. This deviation signifies that a similar solute/solvent interaction exists for thioacetic acid. In addition, the magnitude of the deviation indicates the importance of the substituent/solvent interaction of the thio substituent relative to the other substituents. A large deviation suggests an important interaction, whereas small deviations suggest minor substituent/solvent interactions by this mode. The predicted acidity of phenoxyacetic acid correlates fairly well with the experimental acidity. A major difference between the pairs of electrons of phenoxyacetic acid and methoxyacetic acid is that the electron pairs of the oxygen of phenoxyacetic acid are delocalized into the phenyl ring

---

(30) Kamlet, M. L.; Abboud, J.-K. M.; Abraham, M. H.; Taft, R. W. *J. Org. Chem.* **1983**, *48*, 2877.

(31) Shorter, J. *Correlations Analysis of Organic Reactivity;* Research Studies, Press: New York, 1982.

(32) Ridick, J. A.; Bunger, W. B. Organic Solvents, 3rd ed. In *Techniques of Chemistry;* Weissberger, A., Ed.; Wiley-Interscience: New York, **1970**; Vol. 11.

and not as available for interaction with water as those of methoxyacetic acid.

N,N-Dimethylglycine is expected to destroy the fit for the MLR equation. However, the predicted acidity is very close to the experimental acidity. In order to understand this agreement, the experimental acidity must be discussed. The acidity of dimethylamino acid is dictated by the magnitude of the tautomeric equilibrium.[33] This magnitude is medium dependent. In the gas phase, amino acids are known to exist as neutral molecules,[34] whereas glycine and other soluble amino acids exist as zwitterions in aqueous solution and in the solid crystalline state.[35] Thus, the experimental acidity of *N,N*-dimethylglycine shown in Table 2 is a measure of the acidity of the dimethylammonium functionality and not the acidity of the carboxylic acidity. The predicted acidity of the neutral *N,N*-dimethylglycine in water is 5.19.[21] The experimental acidity is much lower than the expected acidity. Thus, neutral *N,N*-dimethylglycine, like methoxyacetic acid, is indeed an outlier for the acidity of substituted acetic acids. The acidity of neutral *N,N*-dimethylglycine would be an outlier since the *N,N*-dimethyl group is a very basic group. Inspection of Table 1 shows that q_- for the dimethylglycine acid is relatively large compared to the other q_- values, and thus interaction with the solvent by this mode will contribute greatly to the acidity of this acid. This factor, as pointed out in the previous section, decreases the acidity of the acetic acid. The other factor is the covalent basicity (ϵ_B) which relates the ability of the relatively basic nitrogen to interact with the solvent. This interaction increases the acidity, whereas the electrostatic interaction decreases the acidity. However, owing to the magnitude of q_-, the acidity is drastically reduced in water. Thus, the closeness of the experimental and predicted values appears to be coincidental.

Nonaqueous Acidity. The nonaqueous solvents that are used in this study are all alcohols, except one—dimethoxyethane (DME). From the MLR equations, of which the coefficients are shown in entries 5, 6, 7, 8, and 9 (Table 4), the important contributions to acidity variations are the volume, electrostatic basicity, and covalent acidity terms. In DME (entry 10), only the volume and covalent acidity contributions are significant. The volume contribution indicates the importance of the creation of a cavity in the solvent in order to force the initial solvation of the acetic acid before actual dissociation takes place. This process requires energy since the intermolecular attractive forces of the solvent molecules must be broken. On the other hand, energy is gained because of the favorable solute/solvent interactions that result. In water, the cavity contribution is not significant. Apparently, the energy required to disrupt the solvent/solvent interactions just about equals the energy gained from the hydration of the acids. For the other solvents, however, this contribution is significant and the sign is positive. A positive coefficient means that the energy gained by solvation of the solute is less than the energy lost by the creation of the cavity for the solute. This difference in energies is reflected in the magnitudes of the coefficients. Large coefficients result if large bulky solvent molecules cannot effectively solvate the solute, and especially if solute molecules have polar functionalities that are in close proximity to each other. The trend in coefficients is in agreement with the Hildebrand parameters, which is a measure of the energy required to separate solvent molecules from one another and as a result is a good measure of the solvent/solvent interactions that are disrupted for the creation of a suitable size cavity for the solutes.[36] For the solvents used in this study, the Hildebrand parameters are shown in Table 2. The magnitude of the coefficient for pK_a of entry 9 (Table 4) for 2-methyl-2-propanol cannot be used for comparison with the other coefficients, since the temperature of the pK_a values used to generate the correlation equation is not the same as that for the other pK_a values.

The next contribution that is of significance to acidity variations and that is common to the nonaqueous solvents is the interaction of the solute and the solvent by virtue of the covalent acidity (ϵ_A) of the solute. The descriptor of this contribution depicts the ability of each acid to act as a Lewis acid in an acid–base interaction with the solvent. Since the O–H bond is the most polar covalent bond for the acid, the acceptance of electrons from the solvents would result in the formation of a solvated proton and this effect would increase the acidity of each acetic acid. As a result, the coefficients for this contribution to the potential acidity of the acetic acids is negative. The ability of the solvent molecule to donate a pair of electrons to each acid is reflected by the covalent basicity terms, which are shown in Table 2. Large values mean that the solvent molecule will readily donate a pair of electrons to the solute, here the acid, causing an increase in acidity. Compared to the hydrogens of water, alkyl groups have been shown to increase the basicity of alcohols and ethers. The solvatochromic hydrogen acceptor parameter β is larger for bulky alcohols compared to methanol.[37] Since the ability of water to donate a pair of electrons to solute molecules is much lower than that of alcohols and ethers, this contribution is not of significance in water. The electrostatic basicity contribution is significant, in water and all the nonaqueous solvents, except in DME, for reasons shown in the previous section. Since DME has no acidic hydrogens, its electrostatic acidity term is not significant. From Table 2, the magnitude of the q_+ descriptor for this solvent is the smallest of all the solvents used.

Conclusions. On the basis of the goodness of the "fit" of the relationships shown in Table 4, the TLSER descriptors model fairly well the solute/solvent interactions that affect the acidity of acetic acids. As a result, characterization of significant solute/solvent interactions can be determined from the coefficients of the MLR equations. Owing to the good agreement that is obtained between experimental and predicted acidities for 25-substituted acetic acids in seven solvents, this model is very useful for the prediction of the acidity of other substituted acetic acids in these solvents. Even though the coefficients for the equations in which 2-methyl-2-propanol was considered cannot be compared directly

---

(33) (a) Edsall, J. T.; Blanchard, M. H. *J. Am. Chem. Soc.* **1933**, *55*, 2337. (b) Borrow, G. M. *J. Am. Chem. Soc.* **1958**, *80*, 86. (c) Wada, G.; Tamura, E. Okina, M.; Nakamura, M. *Bull. Soc. Chem. Jpn.* **1982**, *55*, 3064. (d) Hughes, D. L.; Bergan, J. J.; Grabowski, J. J. *J. Org. Chem.* **1986**, *51*, 2579.

(34) Gorman, G.; Speir, J. P.; Turner, C. A. Amster, I. J. *J. Am. Chem. Soc.* **1992**, *114*, 3986.

(35) (a) Klotz, I. M.; Green. D. M. *J. Phys. Chem.* **1948**, *52*, 961. (b) Haberfield, P. *J. Chem. Ed.* **1980**, *57*, 346. (c) Butler, J. N. *Ionic Equilibrium: A Mathemitical Approach*; Addison-Wesley Publishing Co.: Reading, MA, 1964; pp 231–237.

(36) (a) Herbrandson, H. F.; Neufeld, F. R. *J. Org. Chem.* **1966**, *31*, 1140. (b) Barton, A. F. M. *Chem. Rev.* **1975**, *75*, 731.

(37) Abrahams, M. H.; Grellier, P. L.; Prior, D. V.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans. 2* **1990**, 521.

with the other solvents owing the difference in temperatures, the goodness of the fit does indicate the applicability of the TLSER descriptors for the acidity of acetic acids in that solvent. For acetic acids that have localized electron pair(s), however, this model fails to give accurate acidity predictions but works fairly well for acetic acids that have substituents with delocalized electrons or alkyl substituents. On the basis of the analysis of the solute/solvent interactions for the acidity of acetic acids by TLSER, the acidity variations in water are dictated primarily by the acidic property of water. The hydrogen bonds that are formed from water to the acetic acids and any structural variations that cause an increase in the negative partial charge on the oxygen atoms of carboxylic functionality dictate the acidity. Since alcoholic solvents are less acidic and bulkier than water, the ability to form a cavity in the solvent for which the solute must be placed and the solvation of the most positive region of the acid are important contributions for acidity determination. For nonacidic solvents, such as DME, only the interaction of the most positive region of the acid with the solvent is important for acidity variations.

Acknowledgment. A.D.H. is pleased to acknowledge support from a grant from The National Science Foundation (RII-9019759).