Substituent Effects on the Physical Properties and pK_a of Phenol

KEVIN C. GROSS, PAUL G. SEYBOLD

Department of Chemistry, Wright State University, Dayton, Ohio 45435

Received 25 February 2001; accepted 2 March 2001

ABSTRACT: Substituent effects on the physical properties and pK_a of phenol were studied using density functional theory [B3LYP/6-311G(d,p)] calculations. Substituents alter the physical properties of phenol such as the hydroxyl-group C—O and O—H bond lengths, the C—O—H bond angle, and the energy barrier to rotation about the C—O bond, and also influence the hydroxyl-group pK_a. Except for the rotational barrier, Hammett σ constants showed strong correlation with these property changes. Several quantum chemical parameters, including the natural charge on the phenolic hydrogen Q_n(H) and the natural charge on the phenoxide oxygen $Q_n(O^-)$, the HF/6-311G(d,p) HOMO energy E_{homo}, and the proton-transfer energy ΔE_{prot}, outperformed the empirical Hammett constants in modeling changes in the pK_a. All of these latter parameters yielded correlation coefficients $|r| > 0.94$ for the pK_a. Int J Quantum Chem 85: 569–579, 2001

Key words: Hammett constants; partial atomic charges; pK_a

Introduction

Understanding substituent effects on molecular properties has long been an important goal in chemistry. The first successful attempt at quantifying these effects came in the form of the Hammett constants (σ) [1, 2]. Introduced in the 1930s, these constants have enjoyed considerable success in relating changes in a number of physicochemical properties, such as NMR peaks, reaction rates, equilibrium positions, etc., to the natures of the attached substituents [3–7].

Correspondence to: P. G. Seybold; e-mail: paul.seybold@wright.edu.

Unfortunately, Hammett constants cannot be expected to work in all situations [4], and because of their empirical nature it is difficult to know exactly what chemical information about a substituent they encode [8]. Properly selected quantum chemical parameters may provide a more fundamental approach to understanding the observed property variations engendered by substituents. In earlier work with substituted anilines, we demonstrated that the amino-nitrogen natural charge was an excellent regression descriptor for several physical and chemical properties including the pK_a for 18 mono-substituted anilines [9]. In a subsequent study, multiply substituted anilines were examined, and it was concluded that several quantum chem-

ical parameters can in some cases outperform the Hammett constants [10].

Here we extend these studies to include a new class of compounds—the phenols. Phenol is a central molecule in organic chemistry. It is a precursor for many valuable products (e.g., the food preservative BHT and the insulator Bakelite) and displays a variety of biological activities [11, 12], particularly in the form of its chlorinated derivatives [13]. A number of experimental studies have examined substituent effects on the properties of phenol.

In this article, we examine changes in several calculated physicochemical properties—bond lengths and angles, the —OH group rotation energy barrier, and the pK_a—that accompany the introduction of substituents. We begin by studying phenol itself at several computational levels in order to establish a suitable level and basis set for subsequent study. Next, we evaluate the ability of the traditional Hammett constants to account for property variations induced by the substituents. Finally, we determine the ability of several quantum chemical parameters to model the above changes with particular attention to changes in the experimental pK_a values.

Methods

Following an examination of phenol itself at a number of levels of theory, the equilibrium geometries of phenol and 18 mono-substituted phenols, as well as their corresponding phenoxides, were determined at the B3LYP/6-311G(d,p) level. To establish the lowest energy conformations, AM1 geometry optimizations were first carried out for a systematically varied set of starting geometries. In some cases, several low-energy conformations were found. B3LYP/6-31G(d) energy calculations were then used to select the lowest energy conformer among these conformers. Transition state geometries corresponding to rotation about the C—O bond were also located at this level of theory for each of the phenols. Location of the transition states required two steps. Starting from the equilibrium geometry, the torsional angle between the —OH group and closest phenyl-ring carbons was first fixed at 90° and the remaining variables subjected to geometry optimization. Following this, the torsional angular constraint was relaxed and a transition state search was performed. Vibrational frequency analysis was performed to ensure minimum-energy conformations and first-order saddle points, as appropriate. Zero-point corrections to the energies were also determined.

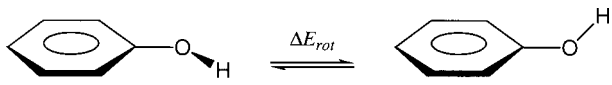
The equilibrium geometry B3LYP/6-311G(d,p) wavefunctions were used to determine the Mulliken, Merz–Sing–Kollman electrostatic [14, 15], and natural [16] charges (Q_M, Q_e, and Q_n, respectively). From the optimized geometries, single-point energy calculations at the HF/6-311G(d,p) level were used to obtain HOMO and LUMO energies in the context of Koopmans' theorem [17]. The barrier to rotation about the C—O bond, ΔE_{rot}, was defined as the difference between the energy of the transition state and that of the equilibrium geometry. Proton-transfer energies ΔE_{prot} were computed for the isodesmic reaction

Hammett constants [18] and pK_a values [19, 20] were taken from the literature. The Spartan program [21] was used in the initial conformational analysis, and the Gaussian 98 suite of programs [22] was employed for the final quantum chemical calculations. Statistical analyses were performed using Mathematica [23].

Results and Discussion

PHENOL

Phenol is a planar molecule as indicated by microwave [24] and far-infrared spectroscopy [25, 26], and also theoretical studies [27, 28]. Overlap between the phenyl π-system and oxygen's lone-pair electrons endows the planar form (**I**) with greater stability than the nonplanar form (**II**). Structure **II** represents a transition state for rotation about the C—O bond.



I. Parent geometry **II**. Transition state

The difference in energy between these two forms will be denoted ΔE_{rot} and represents the barrier to rotation of the —OH group.

Table I lists the calculated geometric parameters and ΔE_{rot} values for phenol determined using AM1, Hartree–Fock, second-order Møller–Plesset (MP2) and density functional (B3LYP) theories, each employed with several basis sets.

TABLE I
Calculated and experimental properties of phenol.

Level of Theory[a]	R(C—O) (Å)	R(O—H) (Å)	θ (deg)	ΔE_{rot}
AM1	1.377	0.968	107.9	2.44
		Hartree–Fock		
3-21G*	1.377	0.964	112.9	2.30
6-31G(d)	1.353	0.947	110.6	2.64
cc-pVDZ	1.351	0.945	110.2	2.77
6-311G(d,p)	1.351	0.940	110.6	2.56
6-311+G(2df,2pd)[b]	—	—	—	2.63
aug-ce-pVTZ[b]	—	—	—	2.71
		2nd-Order Møller–Plesset		
3-21G*	1.405	0.992	109.056	3.28
6-31G(d)	1.375	0.973	108.345	3.58
cc-pVDZ[c]	—	—	—	3.73
6-311G(d,p)[c]	—	—	—	3.37
		B3LYP Density Functional		
3-21G*	1.388	0.992	109.9	4.16
6-31G(d)	1.369	0.970	108.9	4.09
cc-pVDZ	1.367	0.969	108.6	4.10
6-311G(d,p)	1.367	0.963	109.1	3.84
6-311+G(2df,2pd)[d]	—	—	—	3.64
aug-cc-pVTZ[d]	—	—	—	3.24
experimental[e,f]	*1.381*	*0.958*	*106.4*	*3.36–3.56*

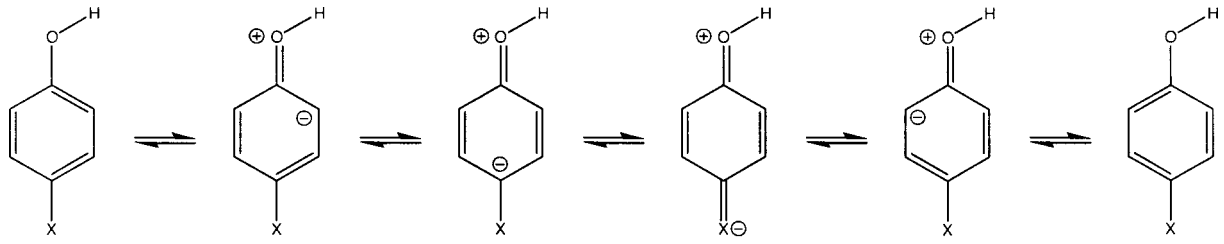
[a] Level of geometry optimization unless otherwise noted.
[b] Single-point calculation using HF/6-311G(d,p) geometry.
[c] Single-point calculation using MP2/6-31G(d) geometry.
[d] Single-point calculation using B3LYP/6-311G(d,p) geometry.
[e] Bond lengths and angle taken from [28].
[f] Several values for ΔE_{rot} (kcal/mol) have been reported: 3.36 [24], 3.47 [26], and 3.56 [25].

The semi-empirical AM1 level calculation [29] performs well in terms of its geometrical description of phenol as indicated by the close agreement with the experimental bond lengths and angles. Relative errors ($\frac{calc-expt}{expt}$) were in each case less than 1.5%. However, at this level ΔE_{rot} is substantially underestimated (30%). The AM1 method, like other semi-empirical methods, has been parameterized to generate good ground-state, equilibrium geometries and is not expected to give exceptionally good results when used to determine transition-state geometries and their corresponding energies [30].

At the Hartree–Fock (HF) level, the results were relatively poor apart from the auspicious (and presumably accidental) bond length values obtained with the 3-21G* basis set. With double- and triple-ζ quality basis sets, the HF geometries underestimated R(C—O) (~2%), R(O—H) (1–2%), and ΔE_{rot} (20–26%),* whereas they overestimated θ (~4%). As a result of the overlap between the oxygen lone-pair electrons and the phenyl π-system, the partial double-bond character existing in **I** must be broken to arrive at **II**. This double-bond character is evident in comparing ground- and transition-state values for R(C—O): at the HF/6-311G(d,p) level, the ground-state value is 1.351 Å, whereas the transition-state value is 1.369 Å. Since electron correlation must be accounted for in determining the energetics of bond-breaking processes, the rather large errors in ΔE_{rot} were not unexpected. The results at this level are consistent with similar calculations found elsewhere [27, 28].

MP2 and B3LYP calculations were also performed to account for some of the electron correla-

*Taking 3.47 kcal/mol as the true experimental value for ΔE_{rot}.



I II III IV V VI

tion neglected at the HF level. Most notable were the improvements in ΔE_{rot} (errors between 3 and 20%) with MP2 yielding better values. Ignoring results obtained with the 3-21G* basis set, geometries were better than, or on par with, the HF results. R(C—O) was underestimated by 0.4–1.0%, while R(O—H) and θ were too large by 0.5–1.6% and 1.8–2.6%, respectively.

Given its satisfactory results and its moderate computational demands, the B3LYP/6-311G(d,p) level of theory was selected for the remainder of this study.

USE OF HAMMETT CONSTANTS

It is well known that electron-withdrawing groups (EWGs), particularly those *para* to the hydroxy moiety, tend to decrease the C—O bond length in phenols on the basis of standard resonance arguments. Similar reasoning suggests that EWGs, which favor the quinoid-type resonance structures (especially **III** and **IV**), should increase the values of θ, ΔE_{rot}, and the pK_a. Likewise, electron-donating groups (EDGs) should produce opposite effects since the Kekulé forms (**I** and **VI**) will be more dominant.

The above arguments were explored using B3LYP/6-311G(d,p) calculations. Table II lists the calculated values for R(C—O), R(O—H), θ, and ΔE_{rot}, as well as the experimental pK_as and empirical Hammett (σ) constants for the compounds studied. A few available experimental values for ΔE_{rot} are also listed in Table II. The σ constants were first used to correlate the effects of substituents on the calculated parameters in Table II. Table III reports the resulting regression statistics and offers these broken down into *para* and *meta* subsets. As indicated in Figure 1, the fits obtained from the σ values were good to excellent with few outliers.

For each of the 18 substituted phenols, the —OH group was essentially co-planar with the phenyl ring in the equilibrium geometry. For the corresponding transition states, the —OH group was roughly perpendicular to the phenyl ring with some variation in this value for the *meta* substituted compounds.

The calculated C—O bond lengths showed a strong correlation with the substituent Hammett constants as indicated by the regression statistics:

$$R(\text{C—O})\ (\text{Å}) = -0.0114\ (\pm 0.0003) \times \sigma + 1.367\ (\pm 0.001)$$
$$n = 19, \qquad r^2 = 0.947, \qquad s = 0.001, \qquad F = 303.$$

Here, n is the number of compounds, r is the correlation coefficient, s is the standard deviation, and F is the Fisher statistic. In this study, *p*-nitrophenol was the only outlier (its regression and calculated bond lengths differed by more than twice the standard deviation). As one might expect, R(C—O) varied over a much wider range of values (1.356–1.374 Å) for the *para* subset than for the *meta* subset (1.361–1.368 Å).

Similar results, although not as good, were obtained for R(O—H):

$$R(\text{O—H})\ (\text{Å}) = 8.96\ (\pm 0.66) \times 10^{-4} \times \sigma + 0.9626\ (\pm 0.0001)$$
$$n = 19, \qquad r^2 = 0.916,$$
$$s = 1.1 \times 10^{-4}, \qquad F = 184.$$

There were no outliers. As indicated by the slope of the fit, R(O—H) is an order of magnitude less sensitive to changes in substituent σ than is R(C—O).

The C—O—H angle θ increased as substituents became more electron-withdrawing in nature. Comparing θ with σ, one finds

$$\theta(^\circ) = 0.762\ (\pm 0.086) \times \sigma + 109.18\ (\pm 0.04)$$
$$n = 19, \qquad r^2 = 0.823, \qquad s = 0.136, \qquad F = 78.9$$

with *m*-methoxyphenol as the sole outlier.

The correlation of σ with the barrier to rotation about the C—O bond ΔE_{rot} was relatively poor for this data set:

$$\Delta E_{rot}\ (\text{kcal/mol}) = 1.32\ (\pm 0.21) \times \sigma + 3.64\ (\pm 0.08)$$
$$n = 19, \qquad r^2 = 0.707, \qquad s = 0.328, \qquad F = 41.1.$$

TABLE II

Hammett σ values and B3LYP/6-311G(d,p) calculated bond lengths, angles, and rotational barriers for the hydroxy moiety, and experimental pK_as for a series of substituted phenols.

Substituent	σ[a]	R(C—O)[b]	R(O—H)[b]	θ[b]	ΔE_{rot}[c]	$\Delta E_{rot,zp}$[c]	$\Delta E_{rot,exp}$[d]	pK_a[e]
H (phenol)	0.00	1.3670	0.9627	109.12	3.84	3.25	3.56	9.98
m-amino	−0.16	1.3681	0.9625	108.82	4.18	3.61	—	9.87
m-bromo	0.39	1.3634	0.9629	109.54	3.91	3.30	—	9.03
m-chloro	0.37	1.3633	0.9629	109.58	3.93	3.32	—	9.02
m-cyano	0.56	1.3615	0.9631	109.76	4.10	3.45	—	8.61
m-fluoro	0.34	1.3636	0.9628	109.22	4.07	3.48	—	9.28
m-hydroxy	0.12	1.3653	0.9628	109.20	4.21	3.57	—	9.44
m-methoxy	0.12	1.3675	0.9626	108.95	4.11	3.52	—	9.65
m-methyl	−0.07	1.3674	0.9627	109.11	3.86	3.31	—	10.08
w-nitro	0.71	1.3607	0.9632	109.74	4.37	3.67	—	8.40
p-amino	−0.66	1.3737	0.9621	108.88	2.62	2.14	—	10.30
p-bromo	0.23	1.3644	0.9628	109.40	3.78	3.17	—	9.36
p-chloro	0.23	1.3647	0.9627	109.40	3.71	3.10	—	9.38
p-cyano	0.66	1.3582	0.9633	109.74	4.69	3.98	4.26	7.95
p-fluoro	0.06	1.3678	0.9625	109.30	3.26	2.67	2.96	9.95
p-hydroxy	−0.37	1.3717	0.9622	109.01	2.84	2.36	2.69	9.96
p-methoxy	−0.27	1.3714	0.9623	109.01	2.90	2.35	—	10.21
p-methyl	−0.17	1.3685	0.9626	109.06	3.60	3.01	3.24	10.14
p-nitro	0.78	1.3561	0.9635	109.83	5.01	4.28	4.54	7.15

[a] Hammett σ values taken from [18].
[b] Bond lengths in angstroms and bond angles in degrees.
[c] Electronic and zero-point corrected energies in hartrees.
[d] Experimental energies (kcal/mol) taken from [25].
[e] pK_a values taken from [19].

For this model, *m*-aminophenol was the only outlier. Inclusion of zero-point corrections to the energy worsened the results (to $r^2 = 0.670$). As indicated by the positive slope, electron-withdrawing groups tend to increase ΔE_{rot} and electron-donating groups decrease ΔE_{rot}. This is consistent with the shorter R(C—O) values associated with EWGs found above. Examining the *para* and *meta* compounds separately, a strong dependence of ΔE_{rot} on substituents *para* to the —OH group was found ($r^2 = 0.904$). No significant relationship was found for the *meta* compounds ($r^2 = 0.133$). Groups attached at the *meta* positions in phenol exert much less influence on the —OH group than do *para* substituents due to the lack of direct resonance.

We note that reasonably good agreement was found between the calculated values for ΔE_{rot} and the few available experimental values. The calculated rotational barriers underestimated the true values by 9% on average.

The Hammett constants are most naturally applied in the study of acidity. The Hammett σ parameters perform reasonably well in quantifying the influence of substituents on the acidity of phenol.

$$\text{p}K_a = -1.97\ (\pm 0.23) \times \sigma + 9.65\ (\pm 0.09)$$
$$n = 19, \qquad r^2 = 0.816, \qquad s = 0.362, \qquad F = 75.2.$$

Here, *p*-nitrophenol was the only outlier. Biggs and Robinson found a similar equation ($\text{p}K_a = -2.23 \times \sigma + 9.92$) using 13 substituted phenols [31], most of which are included in our data set. With the phenyl backbone and its associated resonance structures common to both the present system and the benzoic acids, one might have expected a stronger correlation between σ and pK_a. In fact, the Hammett constants described variations in the pK_as less accurately than they did for any of the geometric parameters described above.

pK_a PREDICTION VIA QUANTUM CHEMICAL DESCRIPTORS

In seeking to predict pK_as with better accuracy, we examined several quantum chemical parameters as structural descriptors. Table V lists the values for the most useful of these computed parameters:

TABLE III

Linear regression statistics (outliers inclusive) for several computed properties versus the Hammett σ values. Here, *n* is the number of compounds, *r* is the correlation coefficient, *s* the standard error, and *F* the Fisher ratio. Phenol is included in each subset.

	Subset	*n*	r^2	*s*	*F*
	All	19	0.947	1.0×10^{-3}	303
R(C—O)	*para*	10	0.972	9.6×10^{-4}	276
	meta	10	0.946	6.2×10^{-4}	139
	All	19	0.916	1.1×10^{-4}	184
R(O—H)	*para*	10	0.934	1.2×10^{-4}	114
	meta	10	0.879	7.5×10^{-5}	58
	All	19	0.823	0.136	79
θ	*para*	10	0.962	0.062	203
	meta	10	0.840	0.133	42
	All	19	0.707	0.328	41
ΔE_{rot}	*para*	10	0.904	0.240	75
	meta	10	0.133	0.160	1
	All	19	0.816	0.162	75
pK_a	*para*	10	0.815	0.456	35
	meta	10	0.943	0.137	133

the natural charges on the phenolic hydrogen Q_n(H) and the phenoxide oxygen $Q_n(O^-)$, the phenoxide HOMO energy E_{homo}, and the relative proton-transfer energy ΔE_{prot}. As indicated in Figure 2, each of these parameters proved useful in estimating pK_as. The regression statistics for the performance of these parameters are given in Table VI.

Atomic Charges

Although the charge on an atom in a molecule is not a physical observable, the atomic charge concept can offer insights into a variety of phenomena. A number of schemes for assigning charges to atoms within a molecule have been developed. Perhaps the most widely used of these is the Mulliken charge (Q_M), although the simple design of this measure brings with it potential problems [32, 33]. Several methods exist for fitting charges to the electrostatic potential and are called electrostatic charges (Q_e). The natural charge (Q_n) derived from natural population analysis (NPA) [33] is a theoretically cleaner approach which avoids many of the potential pitfalls associated with Q_M.

In previous studies [9, 10], we found that the natural charge on the amino nitrogen in a series of substituted anilines correlated strongly with their associated pK_as. It was also shown that the corresponding Mulliken and electrostatic charges performed poorly in this respect [9]. For purposes of comparison, we evaluated Q_M, Q_e, and Q_n as regression parameters for pK_a. To our surprise, the Mulliken charge performed extremely well for the present series of substituted compounds. Omitting the regression equations,

$$Q_M(\mathrm{H}): \quad n = 19, \quad r^2 = 0.913, \quad s = 0.248, \quad F = 179$$

$$Q_M(\mathrm{O}^-): \quad n = 19, \quad r^2 = 0.894, \quad s = 0.274, \quad F = 144.$$

The Merz–Sing–Kollman [14, 15] electrostatic charges performed quite poorly, although still better than expected given our results with the anilines:

$$Q_e(\mathrm{H}): \quad n = 19, \quad r^2 = 0.344, \quad s = 0.682, \quad F = 8.90$$

$$Q_e(\mathrm{Q}^-): \quad n = 19, \quad r^2 = 0.692, \quad s = 0.467, \quad F = 38.2.$$

Accordingly, NPA was put to the test for this series of phenol derivatives. Q_n(H) worked quite well as a regression descriptor of pK_a:

$$\mathrm{p}K_a = -265\ (\pm 23) \times Q_n(\mathrm{H}) + 132\ (\pm 11)$$
$$n = 19, \quad r^2 = 0.887, \quad s = 0.283, \quad F = 134.$$

Only *p*-nitrophenol was an outlier. Quite reasonably, a more positive hydrogen was associated with a lower pK_a.

The phenoxide oxygen charge was also useful in describing the changes in pK_a for this series:

$$\mathrm{p}K_a = -38.4\ (\pm 2.9) \times Q_n(\mathrm{O}^-) - 19.0\ (\pm 2.2)$$
$$n = 19, \quad r^2 = 0.911, \quad s = 0.252, \quad F = 173.$$

There were no outliers for this regression.

From a heuristic viewpoint, the charges on the acidic hydrogen and the deprotonated oxygen should serve as good measures of acidity. More acidic hydrogens have lower electron densities, and a reasonable measure of atomic charge should reflect this. In the same manner, delocalization of the phenoxide oxygen negative charge is expected to impart stability to the phenoxide, encouraging its formation and lowering the pK_a. Here, too, a good definition of atomic charge should account for this variation. The results above suggest that some definitions of charge (here Q_M and Q_n) are superior in

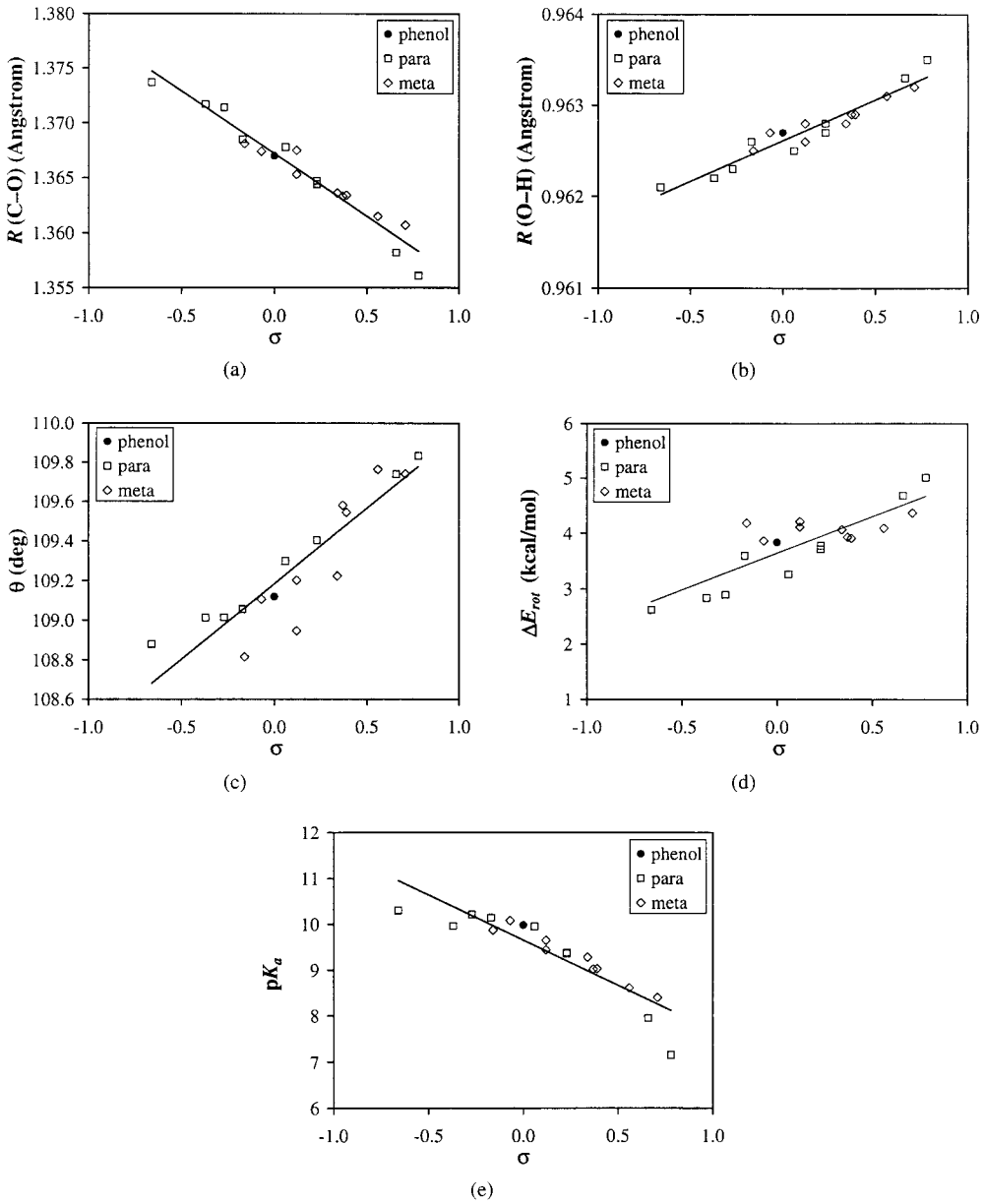


FIGURE 1. Illustrations of the relationships between the properties examined and the Hammet σ constants: (a) the hydroxy group C—O bond length, (b) the hydroxy group O—H bond, (c) the C—O—H angle, (d) the barrier to rotation about the C—O bond, and (e) the hydroxy group pK_a.

TABLE IV
Correlation coefficients (r) for linear fits between the studied properties and several regression parameters.

	σ	R(C—O)	R(O—H)	θ	ΔE_{rot}	pK_a
σ	1	−0.973	0.957	0.907	0.841	−0.903
Q_n(H)	0.978	−0.958	0.939	0.933	0.792	−0.942
Q_n(O$^-$)	0.824	−0.898	0.885	0.825	0.759	−0.954
E_{homo}	−0.915	0.942	−0.918	−0.913	−0.762	0.975
ΔE_{prot}	−0.927	0.948	−0.925	−0.951	−0.732	0.972

TABLE V

Calculated natural charges, phenoxide HOMO energies, and relative proton-transfer energies for a series of substituted phenols.

Substituent	Q_n(H)[a]	Q_n(O$^-$)[b]	E_{homo}[c]	ΔE_{prot}[d]	$\Delta E_{prot,zp}$[d]
H (phenol)	0.4617	−0.7521	−0.075	0.00	0.00
m-amino	0.4614	−0.7309	−0.078	2.17	2.32
m-bromo	0.4651	−0.7256	−0.098	−10.38	−10.25
m-chloro	0.4651	−0.7295	−0.097	−9.80	−9.67
m-cyano	0.4674	−0.7283	−0.105	−15.00	−14.74
m-fluoro	0.4646	−0.7397	−0.091	−6.11	−5.96
m-hydroxy	0.4628	−0.7485	−0.084	−1.87	−1.77
m-methoxy	0.4628	−0.7479	−0.081	−1.03	−1.04
m-methyl	0.4161	−0.7563	−0.076	0.32	0.34
m-nitro	0.4690	−0.7265	−0.109	−15.81	−15.58
p-amino	0.4588	−0.7532	−0.074	3.91	4.07
p-bromo	0.4642	−0.7368	−0.094	−8.99	−8.74
p-chloro	0.4641	−0.7393	−0.093	−8.40	−8.17
p-cyano	0.4677	−0.7027	−0.113	−20.09	−19.56
p-fluoro	0.4627	−0.7559	−0.081	−3.16	−3.00
p-hydroxy	0.4605	−0.7655	−0.067	2.30	2.19
p-methoxy	0.4601	−0.7519	−0.076	0.85	0.64
p-methyl	0.4610	−0.7524	−0.072	0.70	0.65
p-nitro	0.4689	−0.6767	−0.130	−25.36	−24.73

[a] Natural charge on the hydroxy-group hydrogen.
[b] Natural charge on the (deprotonated) phenoxide oxygen.
[c] HOMO energies in hartrees were taken from single-point HF/6-311G(d,p) calculations.
[d] Electronic and zero-point corrected energies in kcal/mol.

their ability to describe the acid-base chemistry of the hydroxyl group.

Orbital Energies

Koopmans' theorem [17] assigns a physical interpretation to the highest occupied and lowest unoccupied molecular orbital energies (E_{homo} and E_{lumo}, respectively):

$$E_{homo} \simeq -I, \qquad E_{lumo} \simeq -A.$$

Here, I is the molecular ionization potential and A is the molecular electron affinity. Two related quantities may be defined:

$$\chi = -(E_{homo} + E_{lumo})/2$$
$$\eta = +(E_{lumo} - E_{homo})/2.$$

To the extent that Koopmans' theorem holds, χ is the absolute electronegativity (equivalent to the Mulliken electronegativity) and η is the absolute hardness [34, 35]. It has been shown that η theoretically justifies and allows the quantification of the Hard-Soft-Acid-Base (HSAB) principle [34]. This principle states that hard acids react more readily with hard bases, and soft acids with soft bases.

Taking H^+ as a hard acid and the substituted phenoxides as bases of varying degrees of hardness, we examined η, χ, E_{homo}, and E_{lumo} as possible regression descriptors for the pK_a. None of these parameters correlated strongly with the pK_a at the B3LYP/6-311G(d,p) level (all displayed $r^2 < 0.45$). Accordingly,[†] we examined the same variables obtained by means of single-point HF/6-311G(d,p) calculations using the B3LYP/6-311G(d,p) geometries.[‡] Of the four quantities determined at the HF level, E_{homo} exhibited the strongest relationship with pK_a:

$$\mathrm{p}K_a = 49.5\ (\pm 2.7) \times E_{homo} + 13.8\ (\pm 0.2)$$
$$n = 19, \qquad r^2 = 0.951, \qquad s = 0.187, \qquad F = 328.$$

[†]While Koopmans' theorem does not hold within density functional theory [36], some interpretation of the Kohn–Sham orbital eigenvalues is possible [36–38]. If true, albeit thus far unknown, energy functional is used, the magnitude of the highest occupied orbital energy will equal the ionization potential [37].

[‡]We later verified that geometries obtained at either the B3LYP/6-311G(d,p) or HF/6-311G(d,p) level made no significant difference in the regression statistics. This may justify "stepping down" to single-point HF calculations for many applications.

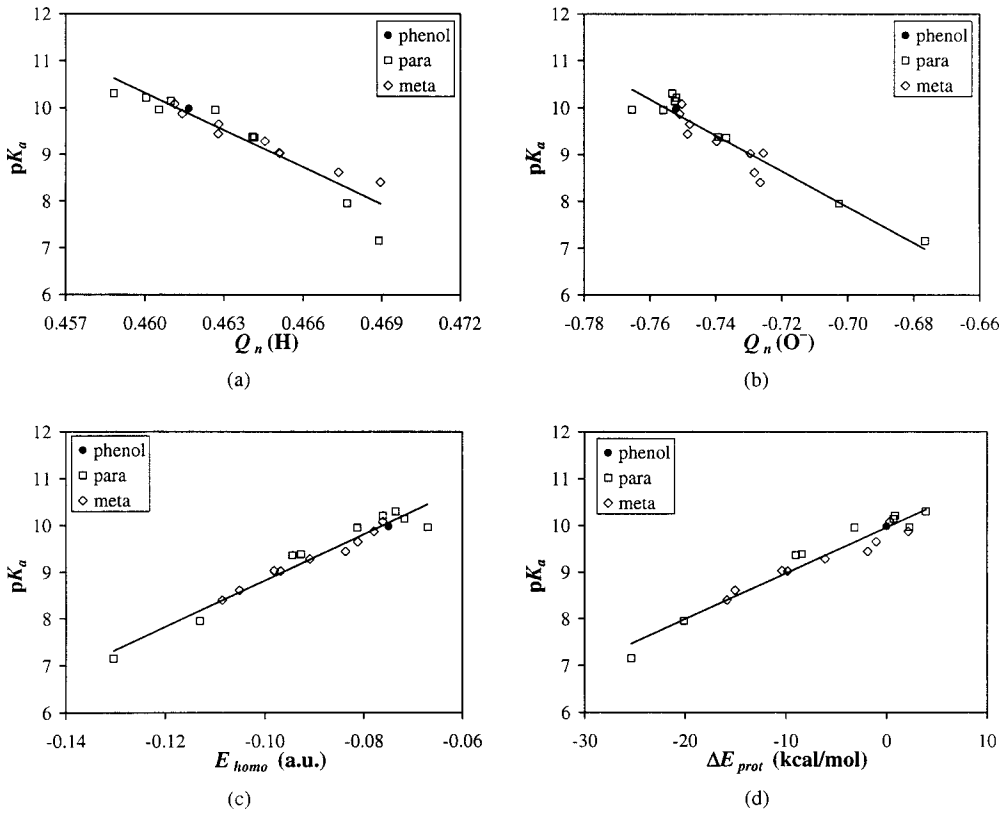


FIGURE 2. Plots of several descriptors versus the pK_a for a series of substituted phenols: (a) the natural charge on the substituted phenol's acidic hydrogen, (b) the natural charge on the substituted phenoxide's oxygen, (c) the HF/6-311G(d,p) HOMO energy of the substituted phenoxide, and (d) the relative proton-transfer energy.

The other parameters were less effective in describing the pK_a. Omitting the regression equations,

$$\chi: \quad n = 19, \quad r^2 = 0.721, \quad s = 0.445, \quad F = 43.9$$

$$E_{lumo}: \quad n = 19, \quad r^2 = 0.429, \quad s = 0.636, \quad F = 12.8$$

$$\eta: \quad n = 19, \quad r^2 = 0.008, \quad s = 0.839, \quad F = 0.14.$$

As computed, η represents the global, average hardness, yet the hardness can take on different, local values within the molecular framework. Thus, in trying to compare η to the site-specific pK_a of the hydroxyl group, these poor results might have been anticipated. We also note that E_{lumo} is a more questionable approximation to A than E_{homo} is to I, and E_{lumo} often varies considerably with the basis set used [39]. This may explain why E_{lumo} and χ are less effective at quantifying changes in pK_a than is E_{homo}.

Relative Proton-Transfer Energy

A more direct measure of changes in acidity can be determined by considering the relative proton-transfer energy between the substituted phenoxides and phenol as depicted earlier. If it is assumed that the zero-point energies, thermal corrections, and entropy contributions are similar across the series of phenols, ΔE_{prot} should indicate with reasonable accuracy the thermodynamic tendency (ΔG) of this reaction. Thus, a positive value of ΔE_{prot} indicates that the substituted phenol is less acidic than phenol itself, and a negative ΔE_{prot} suggests the substituted phenol is more acidic.

As a regression descriptor, ΔE_{prot} performed well:

$$\text{p}K_a = 0.098\ (\pm 0.06) \times \Delta E_{prot} + 9.95\ (\pm 0.06)$$
$$n = 19, \quad r^2 = 0.944, \quad s = 0.199, \quad F = 287.$$

The results were essentially unchanged when zero-point corrections to the energy were applied.

While its correlation with the pK_a is strong, ΔE_{prot} did not accurately rank the compounds in order of their observed solution acidities. The most

TABLE VI

Linear regression statistics (outliers inclusive) for different descriptors' abilities to estimate substituted phenol pK_as. Here, *n* is the number of compounds, *r* is the correlation coefficient, *s* the standard error, and *F* the Fisher ratio. Phenol is a member of each subset.

	Subset	*n*	r^2	*s*	*F*
Q_n(H)	All	19	0.887	0.283	134
	para	10	0.919	0.303	90
	meta	10	0.969	0.102	246
$Q_n(O^-)$	All	19	0.911	0.252	173
	para	10	0.952	0.233	157
	meta	10	0.845	0.226	44
E_{homo}[a]	All	19	0.951	0.187	328
	para	10	0.945	0.248	138
	meta	10	0.980	0.082	387
ΔE_{prot}	All	19	0.944	0.199	287
	para	10	0.964	0.202	213
	meta	10	0.942	0.139	129

[a] HOMO energies taken from single-point HF/6-311G(d,p) calculations.

notable exceptions were *m*-aminophenol and *p*-hydroxyphenol. Solvent effects, which are not accounted for in the present calculation of ΔE_{prot}, are a likely cause. The two noted compounds have substituents that are moderately electron-donating (*p*-hydroxy, $\sigma_p = -0.37$; *m*-amino, $\sigma_m = -0.16$) and readily form hydrogen bonds in aqueous solvent. Such hydrogen bonding is expected to alter the influence of these substituents.

The *r* values listed in Table IV illustrate the strength of correlation between the most useful regression parameters and each of the studied properties. It is apparent that the general-purpose Hammett constants correlate strongly with the calculated properties considered in this report and yield a reasonable representation for the pK_a. It is also clear that the quantum chemical parameters Q_n(H), $Q_n(O^-)$, E_{homo}, and ΔE_{prot} all yield superior regression models for the pK_a.

Conclusions

Our earlier work demonstrated the ability of quantum chemical parameters to account for substituent effects in a series of anilines, particularly as applied to estimating changes in the amino group pK_a. We have now extended this study to a series of substituted phenols. The present results suggest that parameters such as Q_n, E_{homo}, and ΔE_{prot} should have a general applicability in such studies.

References

1. Hammett, L. P. J Am Chem Soc 1937, 59, 96–103.
2. Hammett, L. P. Trans Faraday Soc 1938, 34, 156–165.
3. Jaffé, H. H. Chem Rev 1953, 53, 191–261.
4. Lowry, T. H.; Richardson, K. S. Mechanism and Theory in Organic Chemistry, 3rd ed.; Harper Collins: New York, 1987.
5. Shorter, J. Correlation Analysis in Organic Chemistry: An Introduction to Linear Free-Energy Relationships; Clarendon Press: Oxford, 1973.
6. Shorter, J. Correlation Analysis of Organic Reactivity, with Particular Reference to Multiple Regression; Research Studies Press: New York, 1982.
7. Exner, O. Correlation Analysis of Chemical Data; Plenum Press: New York, 1988.
8. Schüürmann, G. Sci Total Environ 1991, 221–235.
9. Gross, K. C.; Seybold, P. G. Int J Quantum Chem 2000, 80, 1107–1115.
10. Gross, K. C.; Seybold, P. G.; Peralta-Inga, Z.; Murray, J. S.; Politzer, P. J Org Chem (in press).
11. Ruiz, J.; Pérez, A.; Pouplana, R. Quant Struct Act Relat 1996, 15, 219–223.
12. Selassie, C. D.; DeSoyza, T. V.; Rosario, M.; Gao, H.; Hansch, C. Chemico-Biological Interation 1998, 113, 175–190.
13. Schüürmann, G. Quantum Struct Act Relat 1996, 15, 121–132.
14. Singh, U. C.; Kollman, P. A. J Comput Chem 1984, 5, 129.
15. Besler, B. H.; Merz, K. M. J.; Kollman, P. A. J Comput Chem 1990, 11, 431.
16. Reed, A. E.; Weinstock, R. B.; Weinhold, F. J Chem Phys 1985, 83, 735–746.
17. Koopmans, T. A. Physica 1933, 1, 104.
18. Hansch, C.; Leo, A.; Taft, R. W. Chem Rev 1991, 91, 165–195.
19. Albert, A.; Serjeant, E. P. Ionization Constants of Acids and Bases; Methuen: London, 1962.
20. Perrin, D. D., Ed. Dissociation Constants of Organic Bases in Aqueous Solution; Butterworth: London, 1965.
21. Wavefunction, Spartan 5.1, Irvine, CA 92612.
22. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laharn, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A.

Gaussian 98, Revision A.9; Gaussian, Inc.: Pittsburg, PA, 1998.

23. Wolfram Research, Inc., Mathematica 4.0, Champaign, IL 61820, 1999.
24. Pedersen, T.; Larsen, N. W.; Nygaard, L. J Mol Struct 1969, 4, 59–77.
25. Radom, L.; Hehre, W. J.; Popk, J. A.; Carlson, G. L.; Fateley, W. G. J C S Chem Comm 1972, 308–309.
26. Larsen, N. W.; Nicolaisen, F. M. J Mol Struct 1974, 22, 29–43.
27. Head-Gordon, M.; Pople, J. A. J Phys Chem 1993, 97, 1147–1151.
28. Lampert, H.; Mikenda, W.; Karpfen, A. J Phys Chem A 1997, 101, 2254–2263.
29. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. J Am Chem Soc 1985, 107, 3902–3909.
30. Hehre, W. J. Wavefunction: Irvine, 1995.
31. Biggs, A. I.; Robinson, R. A. J Chem Soc 1961, 388–393.
32. Bachrach, S. M. Population analysis and electron densities from quantum mechanics. In Reviews in Computational Chemistry, Vol. 5; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH: New York, 1994.
33. Reed, A. E.; Weinstock, R. B.; Weinhold, F. A. J Chem Phys 1985, 83, 735–746.
34. Parr, R. G.; Pearson, R. G. J Am Chem Soc 1983, 105, 7512–7516.
35. Pearson, R. G. J Org Chem 1989, 54, 1423–1430.
36. Politzer, P.; Abu-Awwad, F. Theor Chem Acc 1988, 99, 83–87.
37. Savin, A.; Umrigar, C. J.; Gonze, X. Chem Phys Lett 1998, 288, 391–395.
38. Stowasser, R.; Hoffmann, R. J Am Chem Soc 1999, 121, 3414–3420.
39. Jensen, F. Introduction to Computational Chemistry; Wiley: New York, 1999.