

Comparison of Different Atomic Charge Schemes for Predicting pK_a Variations in Substituted Anilines and Phenols*

KEVIN C. GROSS,¹ PAUL G. SEYBOLD,¹ CHRISTOPHER M. HADAD²

¹Department of Chemistry, Wright State University, Dayton, Ohio 45435

²Department of Chemistry, The Ohio State University, Columbus, Ohio 43210

Received 31 August 2001; accepted 21 November 2001

Published online 22 January 2002 in Wiley InterScience (www.interscience.com)

DOI 10.1002/qua.10108

ABSTRACT: A number of different methods have been proposed for assigning partial charges to the atoms of a molecule, including both quantum chemical and empirical schemes. A reasonable expectation for any successful calculational scheme is that the atomic charges it produces should vary in a manner consistent with chemical intuition and, more specifically, that these variations should be correlated in a sensible way with experimental observations. Seven of the most popular atomic charge schemes (Bader's AIM charges Q_{AIM} , electrostatic potential charges Q_{ESP} , GAPT charges Q_{GAPT} , Gasteiger π charges Q_{Gast} , Löwdin charges $Q_{Löw}$, Mulliken charges Q_{Mull} , and charges derived from natural population analysis Q_{NPA}) were tested for their ability to represent variations in the pK_a 's of 19 monosubstituted anilines and 19 monosubstituted phenols. In most cases the calculations were performed at the B3LYP/6-311G** level of theory. For the substituted anilines, the amino nitrogen, anilinium proton, and total amino group charges were taken as representative regression parameters, and for the phenols, the phenolic hydrogen, phenoxide oxygen, and hydroxyl group charges were employed. Overall, Q_{AIM} , $Q_{Löw}$, and Q_{NPA} yielded the most successful correlations with the pK_a 's of these compounds, although for the phenol series, Q_{GAPT} and Q_{Mull} also yielded good results. © 2002 Wiley Periodicals, Inc. *Int J Quantum Chem* 90: 445–458, 2002

Key words: atomic charges; AIM charges; GAPT charges; Löwdin charges; Mulliken charges

*Dedicated to the memory of Prof. Per-Olov Löwdin (1916–2000), quantum chemistry pioneer, visionary, and mentor to generations of young scientists.

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

Contract grant sponsor: National Science Foundation.

Contract grant number: CHE-9733457.

Introduction

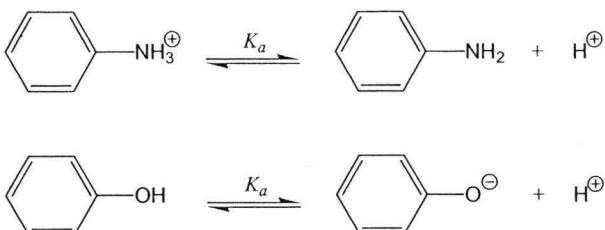
Although the notion of an "atomic charge" in a molecule is not a proper quantum chemical observable, the idea of assigning charges to atoms has proved an immensely valuable heuristic tool for chemists. Atomic charges have been a key concept in the understanding of many types of chemical reactions, and they are central to the interpretations of a number of other phenomena, such as dipole moments and nuclear magnetic resonance (NMR) chemical shifts. They also are important parameters in molecular structure–property and structure–activity relations. Consequently, a great many schemes, both quantum chemical [1–3] and empirical, have been proposed for partitioning the electronic density distribution among the atoms of a molecule or otherwise assigning charge to these atoms.

Traditionally, the Mulliken [4, 5] and electrostatic potential [1, 6] methods have been the most widely used for determining atomic charges, although the weaknesses of both approaches are well known. Mulliken analysis divides overlap populations equally between the two atoms of a bond and thus has the advantage of simplicity. However, its results tend to vary with the basis set employed and it yields unnatural values in some cases [7, 8]. (Many of these weaknesses were originally pointed out by Mulliken himself.) The electrostatic potential approach assigns point charges to the atomic centers in order to best reproduce the electrostatic potential. Here the main weakness is that the molecular electrostatic potential arises from a far more subtle and nonisotropic distribution of polarized and distributed electronic densities than can be accounted for by a simple set of atom-centered point charges. Part of the weakness in the Mulliken approach arises from the fact that it employs a nonorthogonal basis set, and this problem was overcome by Löwdin, who transformed the atomic orbitals to an orthogonal basis set to achieve an improved population analysis (Löwdin charges) [9, 10]. Weinhold and co-workers have extended the idea of "natural orbitals," originally proposed by Löwdin [11], to yield natural population analysis (NPA) charges that are relatively insensitive to basis set changes and have other advantages [12, 13]. These approaches are based on differing analyses of the atomic and molecular orbitals. A very different approach, in which electron densities are integrated over the volumes (basins) assigned to the individual

atoms, has been proposed by Bader in his "atoms-in-molecules" (AIM) method [14]. The volumes are assigned in terms of "zero flux surfaces" defined by the gradients of the electronic density function ρ . Cioslowski has also proposed a different approach, in this case based on the generalized atomic polar tensor (GAPT) [15]. The resulting GAPT charges do not require any direct reference to the basis set employed. Truhlar, Cramer, and their colleagues have also devised a formalism for determining charges based on a parameterized mapping procedure [16].

In addition to these quantum chemical approaches, a number of empirical schemes, mostly based on Sanderson's concept of electronegativity equalization [17], have been proposed. These schemes require only information on the molecular topology and are much simpler to compute. Schemes of this type have been proposed by Gasteiger et al. [18–20], Mullay [21], Abraham and co-workers [22, 23], Scheraga and co-workers [24, 25], Rappé and Goddard [26], and others. Dixon and Jurs have modified Abraham's approach for use in quantitative structure–property relationships [27, 28].

Any definition of atomic charge in a molecule is of necessity an arbitrary distinction, and its *usefulness* (one can hardly speak of validity in this context) must be ascertained from its ability to correlate with experimental properties. One important experimental property is the pK_a of a compound. A good correlation between aqueous acidity and easily calculated gas-phase properties is especially convenient if the atomic charge method is practical for rapid evaluation. In recent studies, we have examined the effects of substitution on the electronic and structural properties of anilines and phenols, including the pK_a 's of these compounds [29–31]. In this work, it was found that a variety of energetic and structural variables were correlated with the variations in the pK_a 's. Most relevant to the present investigation, the atomic charges on the dissociating functional groups of these compounds obtained from NPA (Scheme 1) were strongly correlated with



SCHEME 1.

the variations observed in the pK_a's. In the following we examine the abilities of other charge schemes to yield correlation with the pK_a's of this same set of substituted anilines and phenols.

Methods

For each aniline, anilinium cation, phenol, and phenoxide anion, a series of *para*- and *meta*-substituted derivatives were studied using the B3LYP [32] hybrid density functional theory [33] method. Nine substituents were used for a total of 19 compounds in each series (parent compound included). Geometry optimizations were carried out at the B3LYP/6-311G(d,p) level for each compound. The absence of imaginary frequencies ensured that each geometry corresponded to an equilibrium conformation. The 6-311G(d,p) basis set with six Cartesian *D* functions was used.

Several of the compounds exhibited multiple low-energy conformations. For these compounds AM1 geometry optimizations were performed on systematically varied starting geometries. When necessary, B3LYP/6-31G(d) single-point energy calculations distinguished the minimum-energy conformation.

Using the B3LYP/6-311G(d,p) wave functions, five quantum chemical charges were computed—Bader's atoms-in-molecules charges Q_{AIM} , Merz-Kollman charges fit to the electrostatic potential Q_{ESP} , charges derived from the atomic polar tensor Q_{GAPT} , Mulliken charges Q_{Mull} , and charges derived from natural population analysis Q_{NPA} . Löwdin charges $Q_{Löw}$ were calculated at the HF/6-311G(d,p) level using the B3LYP/6-311G(d,p) geometries. Topological Gasteiger charges Q_{Gast} were also computed. For the amino nitrogen and phenoxide oxygen, charges $Q(N)_{Gast}^{\pi}$ and $Q(O^-)_{Gast}^{\pi}$ were computed; for the acidic hydrogens in both systems, the charges $Q(H^+)_{Gast}^{\sigma}$ and $Q(H)_{Gast}^{\sigma}$ were computed.

Spartan [34] was used for initial determinations of the minimum-energy conformations. The Gaussian98 [35] suite of programs was used for the final geometry optimizations and the quantum chemical population analyses, except where noted. Using wave function files generated by Gaussian98, the AIMALL collection of programs generated the AIM charges [36]. Gasteiger charges were calculated using Petra via the interactive website <http://www2.chemie.uni-erlangen.de>. Löwdin charges were calculated using the Gamess-US [37]

package. Since bromine is not included in the Gamess-US 6-311G(d,p) basis set, the bromo compounds were excluded from Löwdin analysis. The experimental pK_a's were taken from the compilation of Albert and Serjeant [38]. For the anilines the pK_a values refer to dissociations of the anilinium ions. A variety of atomic and functional group charges were examined, as will be described.

Results and Discussion

In developing pK_a models it is reasonable to expect that variations in the pK_a should be correlated in some way with electronic changes at the dissociating atomic positions. For substituted anilines and phenols, we have found that atomic charges on the functional group heavy atoms (e.g., the N in $-\text{NH}_2$ or the O of the $-\text{OH}$ or $-\text{O}^-$) and on the acidic hydrogens (e.g., in $-\text{NH}_3^+$ or $-\text{OH}$) can serve as good regression parameters [29–31]. In the phenol series, for example, the $-\text{OH}$ group pK_a varies with substitution, increasing with electron-donating groups (EDGs) and decreasing with electron-withdrawing groups (EWGs). EWGs increase, and EDGs decrease, delocalization of the negative charge of the phenoxide ion formed upon proton loss. Appropriately defined, the charge on the phenoxide oxygen should serve as an indicator of this delocalization, and in turn should correlate with the $-\text{OH}$ group's pK_a. When the $-\text{OH}$ hydrogen is involved in a more polarized bond, it is generally more acidic. Hydrogen's atomic charge can characterize this degree of polarization and can also serve as a pK_a regression parameter. Similar reasoning suggests that the amino nitrogen charge in the anilines and the hydrogen charge in the anilinium ions should correlate with the amino group's pK_as. It is important to emphasize again that for the anilines the pK_a refers to that of the anilinium ion.

Seven population analysis schemes were examined for their ability to generate charges that correlate with the experimental pK_a's for a series of 19 monosubstituted anilines and 19 monosubstituted phenols. For the anilines, the charges investigated were those for the amino nitrogen $Q(N)$ and anilinium hydrogen $Q(H^+)$, and the total charges on the $-\text{NH}_2$ and $-\text{NH}_3^+$ functional groups. The phenolic hydrogen and phenoxide oxygen charges, $Q(H)$ and $Q(O^-)$, respectively, and the total charge on the $-\text{OH}$ group were tabulated for the phenols. The respective atomic charges and pK_a's are

TABLE I

Calculated atomic charges for the amino-group nitrogen in a series of substituted anilines.^a

Substituent	$Q(N)_{AIM}$	$Q(N)_{ESP}$	$Q(N)_{GAPT}$	$Q(N)_{Gast}^{\pi}$	$Q(N)_{Löw}$	$Q(N)_{Mul}$	$Q(N)_{NPA}$
H (aniline)	-1.0356	-0.7689	-0.6488	0.0852	-0.1089	-0.4722	-0.7758
<i>m</i> -Amino	-1.0326	-0.7991	-0.6528	0.0848	-0.1031	-0.4741	-0.7763
<i>m</i> -Bromo	-1.0393	-0.7367	-0.6809	0.0875	<i>n/a</i>	-0.47	-0.7702
<i>m</i> -Chloro	-1.0395	-0.7461	-0.6766	0.0888	-0.0967	-0.47	-0.7702
<i>m</i> -Cyano	-1.0434	-0.7499	-0.6736	0.0896	-0.0952	-0.4744	-0.7682
<i>m</i> -Fluoro	-1.0393	-0.7513	-0.6649	0.0897	-0.0972	-0.4707	-0.7714
<i>m</i> -Hydroxy	-1.0346	-0.7502	-0.654	0.0881	-0.1008	-0.4727	-0.7746
<i>m</i> -Methoxy	-1.0334	-0.7529	-0.6536	0.0881	-0.1013	-0.4716	-0.7755
<i>m</i> -Methyl	-1.0348	-0.7661	-0.6546	0.0856	-0.1044	-0.474	-0.7759
<i>m</i> -Nitro	-1.0443	-0.7487	-0.6694	0.2361	-0.094	-0.4706	-0.7675
<i>p</i> -Amino	-1.0152	-0.8099	-0.638	0.0854	-0.1193	-0.4761	-0.7841
<i>p</i> -Bromo	-1.0379	-0.7649	-0.7007	0.0885	<i>n/a</i>	-0.4721	-0.772
<i>p</i> -Chloro	-1.0368	-0.7588	-0.6885	0.0891	-0.105	-0.4721	-0.7726
<i>p</i> -Cyano	-1.0584	-0.7732	-0.7618	0.1306	-0.0912	-0.4718	-0.76
<i>p</i> -Fluoro	-1.0275	-0.7431	-0.6342	0.0907	-0.1105	-0.4717	-0.7775
<i>p</i> -Hydroxy	-1.0185	-0.7311	-0.6292	0.0875	-0.1163	-0.4739	-0.7819
<i>p</i> -Methoxy	-1.0201	-0.7335	-0.6492	0.0877	-0.116	-0.4741	-0.7817
<i>p</i> -Methyl	-1.0309	-0.7648	-0.6655	0.0849	-0.1116	-0.4742	-0.7777
<i>p</i> -Nitro	-1.0667	-0.7692	-0.7812	0.2294	-0.0854	-0.4703	-0.7539

^a Seven charge schemes were employed, namely, the atoms-in-molecules charge $Q(N)_{AIM}$, the electrostatic-potential determined charge $Q(N)_{ESP}$, the atomic-polar-tensor charge $Q(N)_{GAPT}$, the Gasteiger π charge $Q(N)_{Gast}^{\pi}$, the Löwdin charge $Q(N)_{Löw}$, the Mulliken charge $Q(N)_{Mul}$, and the natural charge $Q(N)_{NPA}$.

TABLE II

Calculated atomic charges for the amino-group hydrogen in a series of substituted aniliniums and the corresponding experimental pK_a 's.

Substituent	$Q(H^+)_{AIM}$	$Q(H^+)_{ESP}$	$Q(H^+)_{GAPT}$	$Q(H^+)_{Gast}^{\sigma}$	$Q(H^+)_{Löw}$	$Q(H^+)_{Mul}$	$Q(H^+)_{NPA}$	pK_a
H (anilinium)	0.4525	0.4303	0.2443	0.2264	0.1723	0.3099	0.4411	4.58
<i>m</i> -Amino	0.4492	0.4115	0.2401	0.2264	0.1714	0.3068	0.4387	4.88
<i>m</i> -Bromo	0.4552	0.4182	0.2479	0.2264	<i>n/a</i>	0.3121	0.443	3.51
<i>m</i> -Chloro	0.455	0.4131	0.2488	0.2264	0.1739	0.312	0.4429	3.34
<i>m</i> -Cyano	0.4574	0.4335	0.2521	0.2264	0.1751	0.3142	0.4446	2.76
<i>m</i> -Fluoro	0.4554	0.4122	0.2488	0.2264	0.1748	0.3123	0.4431	3.59
<i>m</i> -Hydroxy	0.4527	0.4164	0.2455	0.2264	0.1732	0.3099	0.4412	4.17
<i>m</i> -Methoxy	0.4515	0.3903	0.2442	0.2264	0.1722	0.3089	0.4403	4.2
<i>m</i> -Methyl	0.4509	0.4135	0.2425	0.2264	0.1715	0.3085	0.44	4.69
<i>m</i> -Nitro	0.4583	0.4226	0.2532	0.2285	0.1759	0.315	0.4453	2.5
<i>p</i> -Amino	0.4446	0.3955	0.2266	0.2264	0.1682	0.3048	0.4347	6.08
<i>p</i> -Bromo	0.4547	0.4265	0.2452	0.2264	<i>n/a</i>	0.3115	0.4425	3.91
<i>p</i> -Chloro	0.4535	0.4196	0.2457	0.2264	0.1734	0.311	0.4418	3.98
<i>p</i> -Cyano	0.4568	0.4219	0.2515	0.2264	0.1757	0.3138	0.4443	1.74
<i>p</i> -Fluoro	0.4533	0.4122	0.2451	0.2264	0.1732	0.3109	0.4416	4.65
<i>p</i> -Hydroxy	0.4494	0.4053	0.2376	0.2264	0.1708	0.3077	0.4388	5.5
<i>p</i> -Methoxy	0.4478	0.3932	0.2349	0.2264	0.1698	0.3064	0.4377	5.29
<i>p</i> -Methyl	0.4504	0.4148	0.2408	0.2264	0.1711	0.3082	0.4396	5.12
<i>p</i> -Nitro	0.4581	0.4232	0.2538	0.2285	0.1768	0.3149	0.4452	1.02

TABLE III

Calculated atomic charges for the phenolic hydrogen in a series of substituted phenols and the corresponding experimental pK_a's.

Substituent	Q(H) _{AIM}	Q(H) _{ESP}	Q(H) _{GAPT}	Q(H) _{Gast} ^σ	Q(H) _{Löw}	Q(H) _{Mul}	Q(H) _{NPA}	pK _a
H (phenol)	0.5566	0.4088	0.2823	0.2177	0.1108	0.245	0.4617	9.98
<i>m</i> -Amino	0.556	0.412	0.2844	0.2177	0.1118	0.2446	0.4614	9.87
<i>m</i> -Bromo	0.5618	0.4211	0.2919	0.2177	<i>n/a</i>	0.2496	0.4651	9.03
<i>m</i> -Chloro	0.5619	0.4231	0.2912	0.2177	0.1147	0.2495	0.4651	9.02
<i>m</i> -Cyano	0.5649	0.4175	0.2947	0.2177	0.1167	0.2522	0.4674	8.61
<i>m</i> -Fluoro	<i>n/a</i>	0.4121	0.2903	0.2177	0.1151	0.2484	0.4646	9.28
<i>m</i> -Hydroxy	0.5583	0.412	0.2835	0.2177	0.1125	0.2468	0.4628	9.44
<i>m</i> -Methoxy	0.5577	0.4179	0.2844	0.2177	0.1122	0.2462	0.4628	9.65
<i>m</i> -Methyl	0.5557	0.4133	0.2801	0.2177	0.1101	0.244	0.4611	10.08
<i>m</i> -Nitro	0.5672	0.4328	0.2967	0.2968	0.1182	0.2543	0.469	8.4
<i>p</i> -Amino	0.5528	0.4126	0.279	0.2177	0.1076	0.2413	0.4588	10.3
<i>p</i> -Bromo	0.5605	0.4158	0.2935	0.2177	<i>n/a</i>	0.2481	0.4642	9.36
<i>p</i> -Chloro	0.5604	0.413	0.2919	0.2177	0.1133	0.2481	0.4641	9.38
<i>p</i> -Cyano	0.5655	0.4164	0.302	0.2177	0.1176	0.2525	0.4677	7.95
<i>p</i> -Fluoro	0.5584	0.411	0.2844	0.2177	0.1118	0.2465	0.4627	9.95
<i>p</i> -Hydroxy	0.5553	0.4149	0.2814	0.2177	0.1092	0.2437	0.4605	9.96
<i>p</i> -Methoxy	0.5545	0.4123	0.2803	0.2177	0.1087	0.243	0.4601	10.21
<i>p</i> -Methyl	0.5557	0.4102	0.2828	0.2177	0.11	0.244	0.461	10.14
<i>p</i> -Nitro	0.5673	0.4185	0.3048	0.2968	0.1192	0.2545	0.4689	7.15

found in Tables I-IV. (The group charge values have been omitted for reasons of economy but can be obtained from the authors.) It is evident that the different charges obtained vary considerably in

magnitude. For the most part, Q_{AIM} , Q_{NPA} , and Q_{ESP} tend to be large compared to $Q_{Löw}$ and Q_{Mul} . We note that trends in the charges are more important than the absolute values. One might falsely con-

TABLE IV

Calculated atomic charges for the phenoxide oxygen in a series of substituted phenoxides.

Substituent	Q(O ⁻) _{AIM}	Q(O ⁻) _{ESP}	Q(O ⁻) _{GAPT}	Q(O ⁻) _{Gast} ^π	Q(O ⁻) _{Löw}	Q(O ⁻) _{Mul}	Q(O ⁻) _{NPA}
H (phenoxide)	-1.2063	-0.7358	-0.9705	-0.1374	-0.4216	-0.5461	-0.7521
<i>m</i> -Amino	-1.2057	-0.7297	-0.9859	0.0327	-0.4187	-0.5462	-0.7509
<i>m</i> -Bromo	-1.1899	-0.694	-1.0042	0.046	<i>n/a</i>	-0.5227	-0.7256
<i>m</i> -Chloro	-1.1923	-0.7068	-0.9888	0.047	-0.402	-0.5261	-0.7295
<i>m</i> -Cyano	-1.1898	-0.7123	-0.9462	0.0559	-0.4018	-0.5232	-0.7283
<i>m</i> -Fluoro	-1.1993	-0.7214	-0.9636	0.0462	-0.4069	-0.5378	-0.7397
<i>m</i> -Hydroxy	-1.2046	-0.7272	-0.9712	0.0414	-0.415	-0.5442	-0.7485
<i>m</i> -Methoxy	-1.2037	-0.7322	-0.9748	0.0413	-0.416	-0.5384	-0.7479
<i>m</i> -Methyl	-1.2047	-0.7358	-0.9791	0.0766	-0.4213	-0.5442	-0.7503
<i>m</i> -Nitro	-1.1891	-0.71	-0.9202	0.1862	-0.3988	-0.5209	-0.7265
<i>p</i> -Amino	-1.2059	-0.7077	-0.9894	0.0754	-0.4246	-0.5464	-0.7532
<i>p</i> -Bromo	-1.1959	-0.7181	-0.9998	0.0776	<i>n/a</i>	-0.5308	-0.7368
<i>p</i> -Chloro	-1.1973	-0.7224	-0.9891	0.0782	-0.4077	-0.5331	-0.7393
<i>p</i> -Cyano	-1.1766	-0.6925	-1.0079	0.0795	-0.3666	-0.5016	-0.7027
<i>p</i> -Fluoro	-1.2059	-0.7214	-0.9561	0.0787	-0.4263	-0.5484	-0.7559
<i>p</i> -Hydroxy	-1.2102	-0.7217	-0.9784	0.0779	-0.4387	-0.5571	-0.7655
<i>p</i> -Methoxy	-1.2045	-0.7293	-0.9986	0.0779	-0.4248	-0.545	-0.7519
<i>p</i> -Methyl	-1.2053	-0.7279	-1.006	0.0769	-0.4239	-0.5459	-0.7524
<i>p</i> -Nitro	-1.1591	-0.6614	-0.9867	0.2342	-0.3405	-0.4811	-0.6767

TABLE V

Linear regressions and statistics for pK_a vs. charge for a series of substituted anilines.^a

	Aniline N charge						
	$Q(N)_{AIM}$	$Q(N)_{ESP}$	$Q(N)_{GAPT}$	$Q(N)_{Gast}^\pi$	$Q(N)_{Löw}$	$Q(N)_{Mul}$	$Q(N)_{NPA}$
Slope	101	-7.63	27.7	-19.6	-136	-499	-174
Intercept	109	-1.82	22.6	6.03	-10.1	-232	-131
r	0.961	-0.12	0.846	-0.697	-0.948	-0.672	-0.983
s	0.36	1.29	0.691	0.928	0.437	0.959	0.236
F	203	0.247	42.7	16.1	132	14	493
	Anilinium H^+ charge						
	$Q(H^+)_{AIM}$	$Q(H^+)_{ESP}$	$Q(H^+)_{GAPT}$	$Q(H^+)_{Gast}^\pi$	$Q(H^+)_{Löw}$	$Q(H^+)_{Mul}$	$Q(H^+)_{NPA}$
Slope	-322	-64	-173	-1150	-559	-406	-430
Intercept	150	30.5	46.3	265	101	130	194
r	-0.915	-0.582	-0.9	-0.603	-0.947	-0.924	-0.912
s	0.523	1.05	0.566	1.03	0.441	0.497	0.531
F	87.1	8.71	72.1	9.69	129	98.6	84.1
	Aniline NH_2 group charge						
	$Q(NH_2)_{AIM}$	$Q(NH_2)_{ESP}$	$Q(NH_2)_{GAPT}$	$Q(NH_2)_{Gast}^\pi$	$Q(NH_2)_{Löw}$	$Q(NH_2)_{Mul}$	$Q(NH_2)_{NPA}$
Slope	-172	-51.5	28.3	n/a	-78.6	-99.8	-73.9
Intercept	-50.7	-0.852	12.5	n/a	11.3	-2.31	0.887
r	-0.783	-0.857	0.499	n/a	-0.987	-0.985	-0.993
s	0.806	0.668	1.123	n/a	0.216	0.221	0.157
F	26.9	46.82	5.6	n/a	584	569	1133
	Anilinium NH_3^+ group charge						
	$Q(NH_3^+)_{AIM}$	$Q(NH_3^+)_{ESP}$	$Q(NH_3^+)_{GAPT}$	$Q(NH_3^+)_{Gast}^\pi$	$Q(NH_3^+)_{Löw}$	$Q(NH_3^+)_{Mul}$	$Q(NH_3^+)_{NPA}$
Slope	-124	-29.9	-8.81	n/a	-136	-123	-176
Intercept	53.6	20.6	8.77	n/a	108	74.4	120
r	-0.914	-0.264	-0.054	n/a	-0.894	-0.889	-0.897
s	0.525	1.249	1.293	n/a	0.612	0.593	0.572
F	86.3	1.27	0.0493	n/a	60	64	70.1

^a Here, r is the correlation coefficient, s the standard error, and F the Fisher statistic. Note that $n = 17$ for $Q_{Löw}$.

clude that the acidic hydrogen in phenol is more acidic than in anilinium by comparing $Q(H)_{AIM}$ with $Q(H^+)_{AIM}$, $Q(H)_{GAPT}$ with $Q(H^+)_{GAPT}$, or $Q(H)_{NPA}$ with $Q(H^+)_{NPA}$, since in each case listed, $Q(H) > Q(H^+)$.

Tables V and VI collect the regression statistics associated with the pK_a vs. atomic charge equations for the anilines and phenols, respectively. Figures 1-4 illustrate relationships between the different quantum chemical atomic charges and the experimental pK_a 's. We now examine each charge

scheme regarding its ability to estimate variations in the pK_a 's of the anilines and phenols.

AIM CHARGE Q_{AIM}

The AIM charges serve as excellent pK_a regression parameters. For the anilines, $Q(N)_{AIM}$ correlated more strongly with pK_a than did $Q(H^+)_{AIM}$. Interestingly, however, $Q(N)_{AIM}$ varied with pK_a counterintuitively: The nitrogen charge *became more positive* with increasing pK_a , suggesting that a nitrogen with less associated electron density more

TABLE VI

Linear regressions and statistics for pK_a vs. charge for a series of substituted phenols.^a

	Phenol H charge						
	Q(H) _{AIM}	Q(H) _{ESP}	Q(H) _{GAPT}	Q(H) _{Gast} ^π	Q(H) _{Löw}	Q(H) _{Mul}	Q(H) _{NPA}
Slope	-182.5	-87.9	-106	-22.3	-241	-207	-265
Intercept	111.5	45.9	40	14.4	36.6	60.5	132
r	-0.944	-0.586	-0.958	-0.662	-0.946	-0.946	-0.942
s	0.285	0.68	0.24	0.631	0.287	0.274	0.283
F	131	8.91	192	13.2	129	144	134
	Phenoxyde O ⁻ charge						
	Q(O ⁻) _{AIM}	Q(O ⁻) _{ESP}	Q(O ⁻) _{GAPT}	Q(O ⁻) _{Gast} ^π	Q(O ⁻) _{Löw}	Q(O ⁻) _{Mul}	Q(O ⁻) _{NPA}
Slope	-64.4	-38.3	-4.88	-6.59	-35.3	-43.9	-38.4
Intercept	-67.7	-18	4.57	9.8	-5.06	-14.1	-19
r	-0.958	-0.832	-0.129	-0.55	-0.946	-0.956	-0.954
s	0.241	0.467	0.835	0.703	0.288	0.248	0.252
F	190	38.2	0.286	7.38	128	179	173
	Phenol OH group charge						
	Q(OH) _{AIM}	Q(OH) _{ESP}	Q(OH) _{GAPT}	Q(OH) _{Gast} ^π	Q(OH) _{Löw}	Q(OH) _{Mul}	Q(OH) _{NPA}
Slope	-151	-52.7	15.8	n/a	-65	-67.9	-83.4
Intercept	-67.9	1.44	16	n/a	12.4	0.848	-7.02
r	-0.937	-0.721	0.36	n/a	-0.929	-0.959	-0.963
s	0.303	0.583	0.786	n/a	0.328	0.239	0.227
F	115	18.5	2.53	n/a	95.2	194	218
	Phenol O charge						
	Q(O) _{AIM}	Q(O) _{ESP}	Q(O) _{GAPT}	Q(O) _{Gast} ^π	Q(O) _{Löw}	Q(O) _{Mul}	Q(O) _{NPA}
Slope	-444	-29.8	22.2	n/a	-87.4	-98.2	-118
Intercept	-465	-7.5	25.1	n/a	3.6	-27.2	-68.8
r	-0.637	-0.422	0.611	n/a	-0.915	-0.951	-0.959
s	0.668	0.763	0.666	n/a	0.359	0.259	0.238
F	10.95	3.68	10	n/a	77	162	196

^a Note that n = 17 for Q_{Löw} and n = 18 for Q(H)_{AIM}, Q(O)_{AIM}, and Q(OH)_{AIM}.

readily accepts a proton. Compared to the amino nitrogen in aniline [Q(N)_{AIM} = -1.0356, pK_a = 4.58], Q(N)_{AIM} was more negative in *p*-nitroaniline [Q(N)_{AIM} = -1.0667, pK_a = 1.02], and Q(N)_{AIM} was more positive in *p*-aminoaniline [Q(N)_{AIM} = -1.0152, pK_a = 6.08]. In the phenol series, Q(O⁻)_{AIM} better estimated pK_a variations than did Q(H)_{AIM}. (Numerical problems were encountered with *m*-fluorophenol and its charges were omitted from the analysis.) For the aniline series the change in the acidic nitrogen AIM charges, Q_{AIM}(H⁺) - Q_{AIM}(H), also gave good results, with r = 0.962, s =

0.352, F = 213. Here, r is the correlation coefficient, s the standard error, and F the Fisher statistic. In the phenol series, the -OH group charge Q(OH)_{AIM} produced a good correlation with pK_a (r = -0.937), but the O charge Q(O)_{AIM} did not (r = -0.637).

ELECTROSTATIC CHARGE Q_{ESP}

The Besler-Merz-Kollman electrostatic charge [39] was not particularly useful in predicting pK_a shifts accompanied by changes in substituent. Only Q(O⁻)_{ESP} produced meaningful results, with

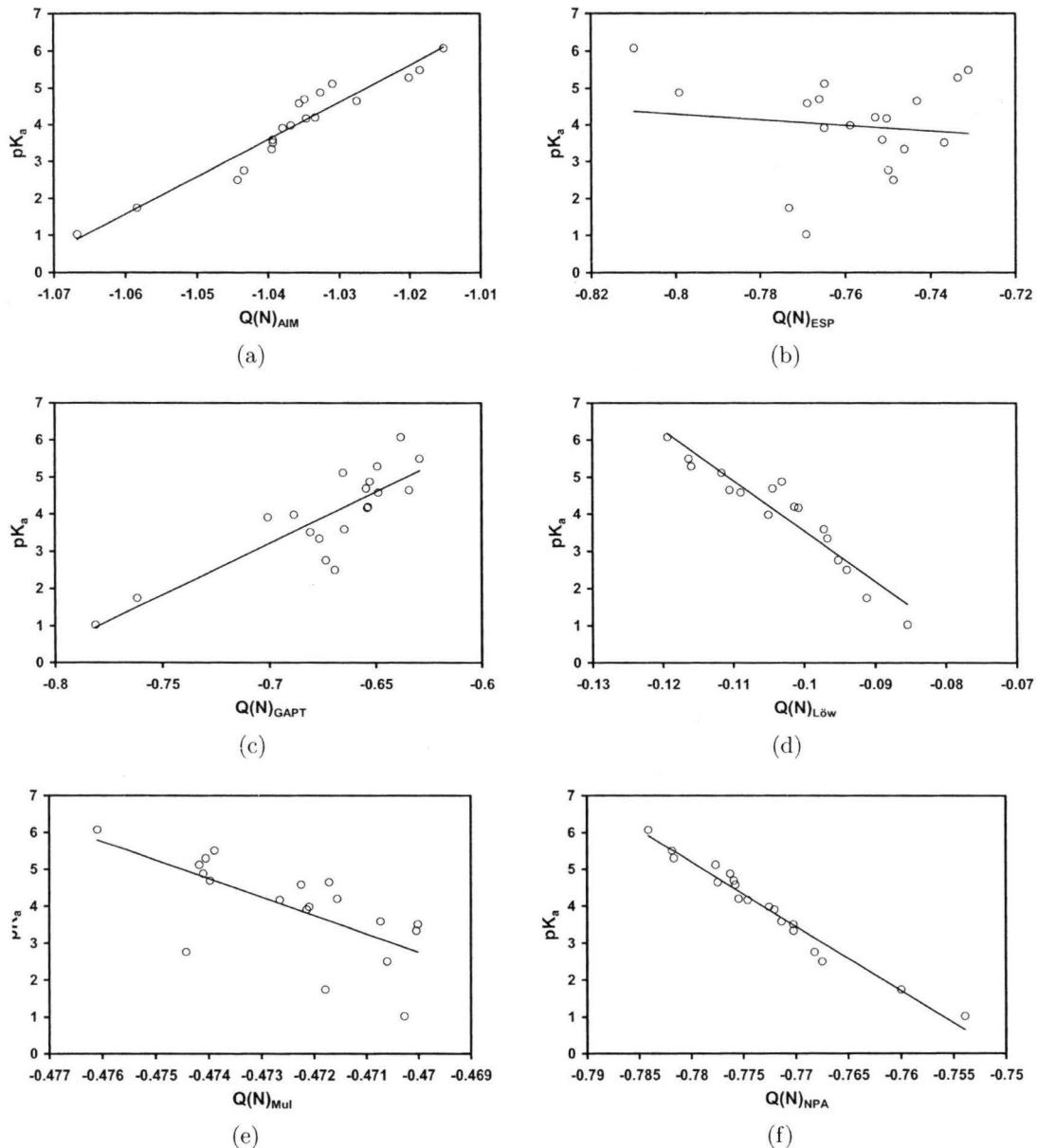


FIGURE 1. Plots of the amino-group nitrogen pK_a vs. charge using several charge definitions for a series of monosubstituted anilines: (a) the atoms-in-molecules charge $Q(N)_{AIM}$, (b) the electrostatic-potential derived charge $Q(N)_{ESP}$, (c) the atomic-polar-tensor charge $Q(N)_{GAPT}$, (d) the Löwdin charge $Q(N)_{Löw}$, (e) the Mulliken charge $Q(N)_{Mul}$, and (f) the natural charge $Q(N)_{NPA}$.

$r = -0.832$, $s = 0.467$, and $F = 38.2$. Various procedures exist for fitting atomic charges to the electrostatic potential. Another such method, CHelpG [40], was evaluated and gave similarly poor results, again except for $Q(O^-)_{ESP}$, which had regression statistics of $r = -0.921$, $s = 0.347$, and $F = 83.7$. The brominated compounds were omitted since the CHelpG method as implemented in Gaussian98 cannot treat them.

GAPT CHARGE Q_{GAPT}

The generalized atomic polar tensor charge of Cioslowski [15] proved somewhat useful as a pK_a descriptor, with the exception of $Q(O^-)_{GAPT}$. As with $Q(N)_{AIM}$, though, $Q(N)_{GAPT}$ varied counterintuitively with pK_a . More positive amino-nitrogen charges were associated with higher pK_a 's while

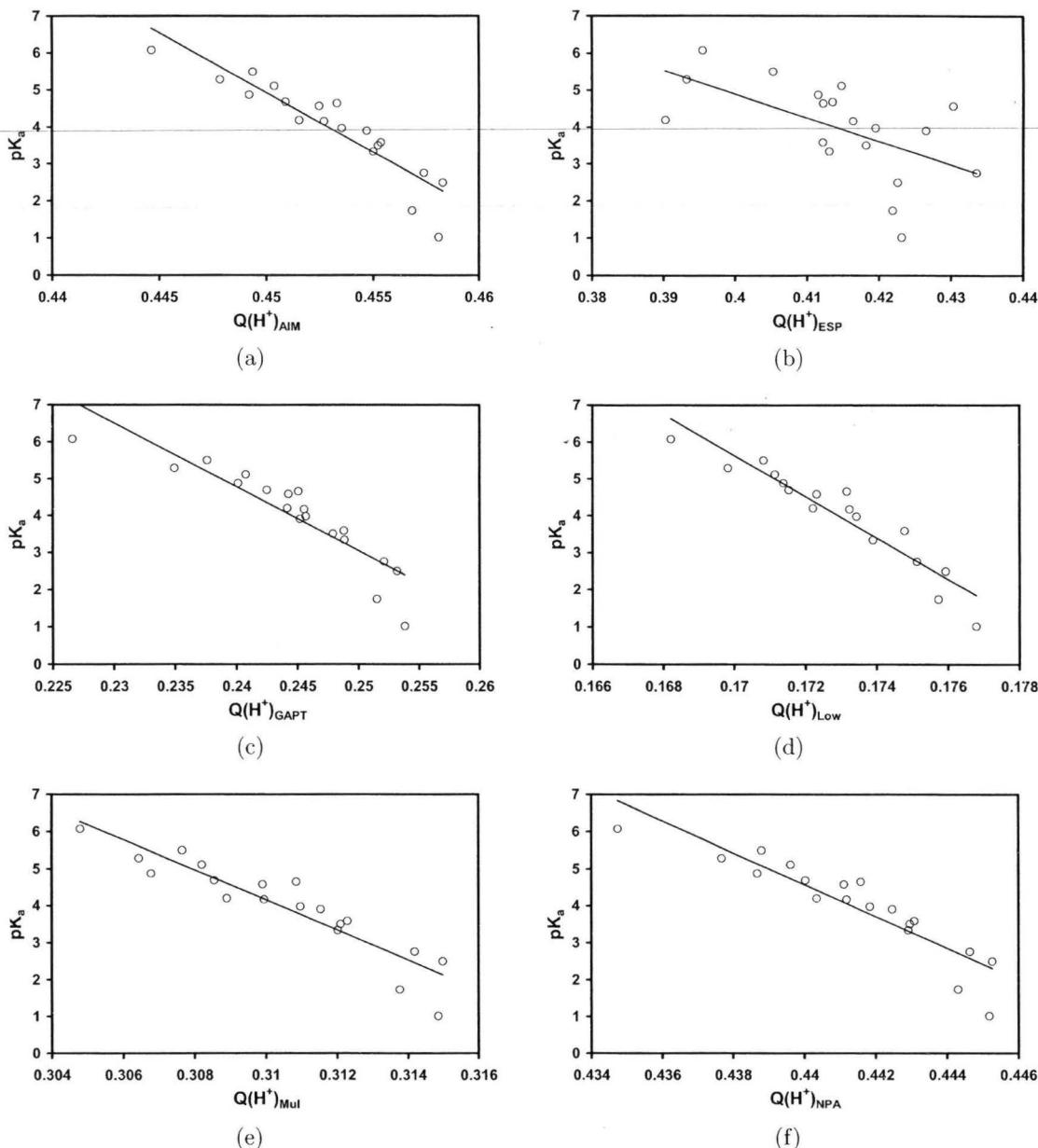


FIGURE 2. Plots of the averaged amoni-group hydrogen pK_a vs. charge using several charge definitions for a series of monosubstituted aniliniums: (a) $Q(H^+)_{AIM}$, (b) $Q(H^+)_{ESP}$, (c) $Q(H^+)_{GAPT}$, (d) $Q(H^+)_{Low}$, (e) $Q(H^+)_{Mul}$, and (f) $Q(H^+)_{NPA}$.

more negative charges were associated with lower pK_a's.

GASTEIGER CHARGE Q_{Gast}

Q_{Gast} , the only “topological charge” examined, performed poorly in correlating the aniline and phenol pK_a's. This was somewhat surprising since in other studies, this method was successful in estimating variations in pK_a's, dipole moments,

and NMR shifts for collections of diverse compounds [18–20, 41, 42]. For the substituent-induced changes in the present compounds, $Q(H^+)_{Gast}^\sigma$ and $Q(H)_{Gast}^\sigma$ in particular displayed no change with differing substituents, except when the nitro group was used. $Q(N)_{Gast}^\pi$ and $Q(O^-)_{Gast}^\pi$ displayed more variation with different substituents, but still did not produce meaningful regression models.

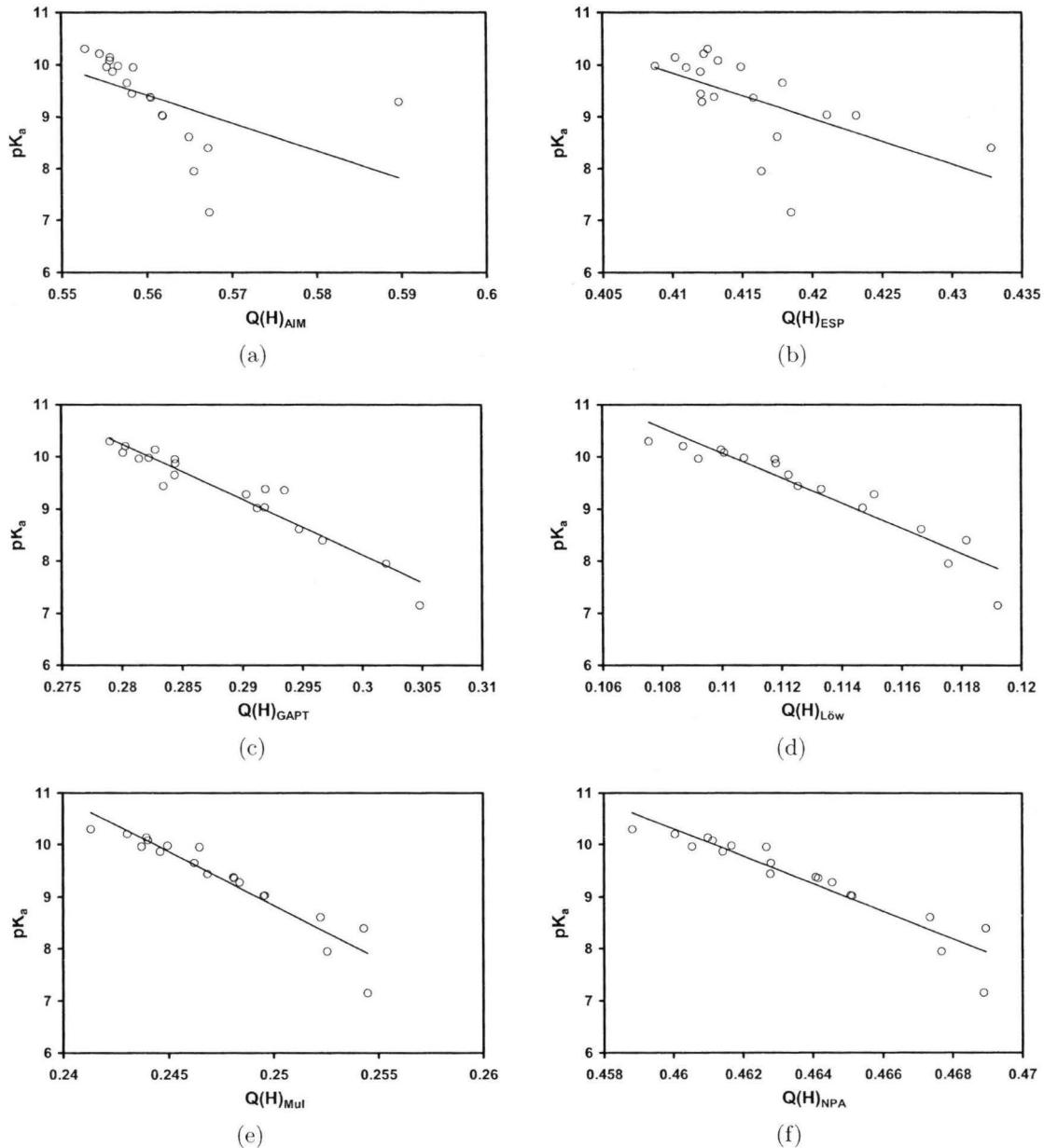


FIGURE 3. Plots of the phenolic hydrogen pK_a vs. charge using several charge definitions for a series of monosubstituted phenols: (a) $Q(H)_{AIM}$, (b) $Q(H)_{ESP}$, (c) $Q(H)_{GAPT}$, (d) $Q(H)_{Löw}$, (e) $Q(H)_{Mul}$, and (f) $Q(H)_{NPA}$.

LÖWDIN CHARGE $Q_{Löw}$

The Löwdin charges served as excellent regression parameters for the aniline and phenol pK_a 's. $Q_{Löw}$ also produced the most consistent correlations, e.g., $Q(N)_{Löw}, r = -0.948$; $Q(H^+)_{Löw}, r = -0.947$; $Q(H)_{Löw}, r = -0.946$; $Q(O^-)_{Löw}, r = -0.946$. For the anilines the total Löwdin charge on the $-NH_2$ group gave an excellent correlation with the pK_a : $r = -0.987, s = 0.216$, and $F = 584$. While simi-

lar to the Mulliken scheme, this population analysis differs by, and clearly benefits from, working with an orthogonal basis set.

MULLIKEN CHARGE Q_{Mul}

In spite of their shortcomings, the Mulliken charges accounted for pK_a variations quite well, except for the amino-nitrogen charge in the aniline series. The correlation coefficient magnitudes were

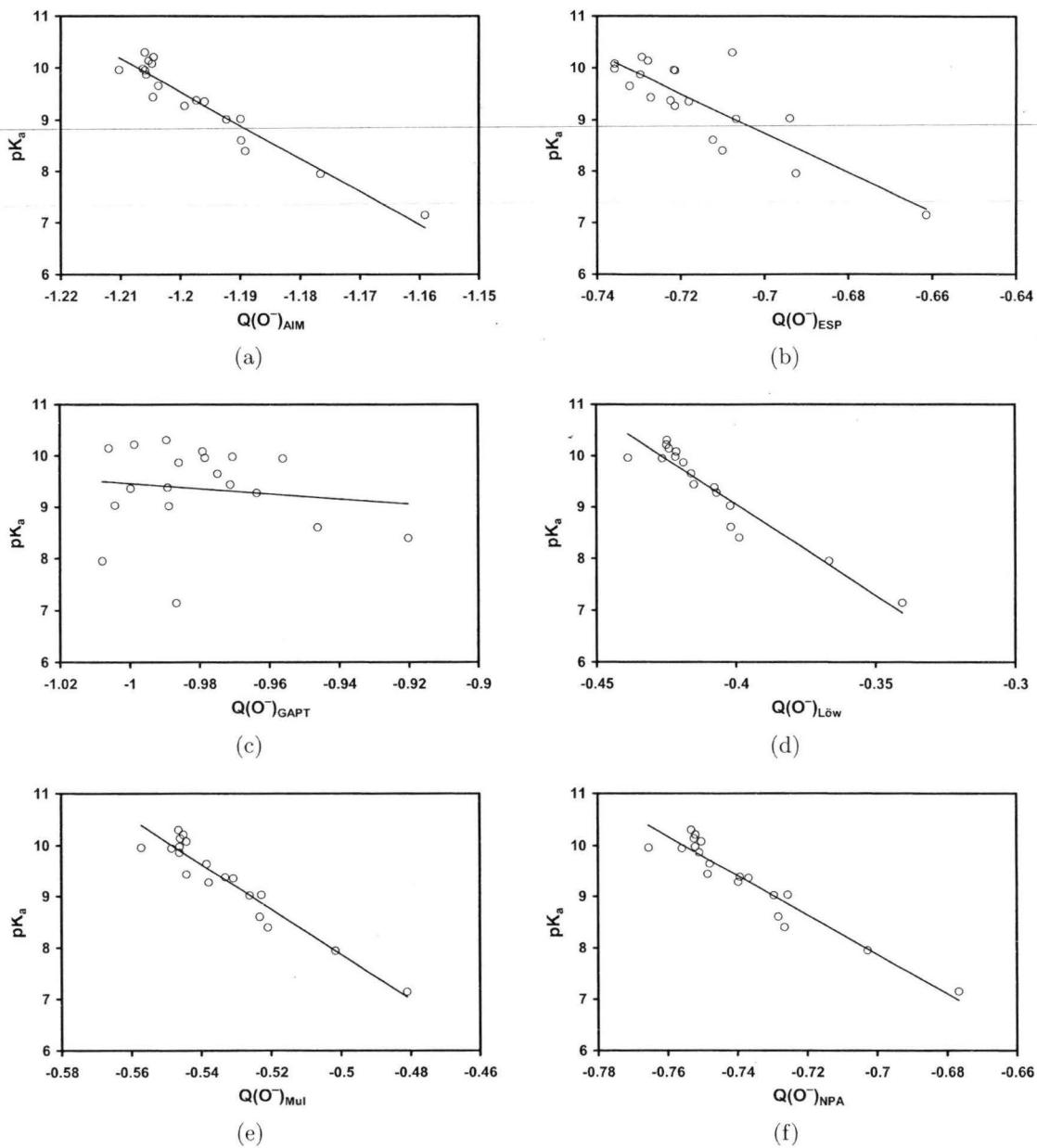


FIGURE 4. Plots of the phenoxide oxygen pK_a vs. charge using several charge definitions for a series of monosubstituted phenoxides: (a) Q(O⁻)_{AIM}, (b) Q(O⁻)_{ESP}, (c) Q(O⁻)_{GAPT}, (d) Q(O⁻)_{Loew}, (e) Q(O⁻)_{Mul}, and (f) Q(O⁻)_{NPA}.

all greater than 0.924, excluding Q(N)_{Mul}. Q(O⁻)_{Mul} was slightly better at estimating phenol pK_a variations than was Q(H)_{Mul}.

NATURAL CHARGE Q_{NPA}

The natural charges functioned as excellent pK_a regression parameters and, like the Löwdin charges, behaved consistently across all four data sets. For

example, Q(N)_{NPA}, $r = -0.983$; Q(H⁺)_{NPA}, $r = -0.912$; Q(H)_{NPA}, $r = -0.942$; and Q(O⁻)_{NPA}, $r = -0.954$. For estimating both the aniline and phenol pK_a's, NPA charges on the heavy atoms—Q(N)_{NPA} and Q(O⁻)_{NPA}, respectively—worked better than the corresponding charges on the acidic hydrogens. The group charges were especially effective, with $r = -0.993$ for Q(NH₂)_{NPA} and $r = -0.963$ for Q(OH)_{NPA}. The changes in the atomic NPA charges

TABLE VII

Cross correlations (r) between each of the atomic charges for the aniline and anilinium data sets.^a

Q(N) Cross-correlation matrix						
$Q(N)_{AIM}$	$Q(N)_{ESP}$	$Q(N)_{GAPT}$	$Q(N)_{Gast}^{\pi}$	$Q(N)_{Löw}$	$Q(N)_{Mul}$	$Q(N)_{NPA}$
1	0.037	0.891	-0.642	-0.938	-0.621	-0.988
$Q(N)_{ESP}$	1	0.147	-0.012	0.031	0.424	0.011
$Q(N)_{GAPT}$		1	-0.578	-0.759	-0.46	-0.907
$Q(N)_{Gast}^{\pi}$			1	0.615	0.44	0.683
$Q(N)_{Löw}$				1	0.686	0.935
$Q(N)_{Mul}$					1	0.64
$Q(N)_{NPA}$						1

Q(H ⁺) Cross-correlation matrix						
$Q(O^-)_{AIM}$	$Q(O^-)_{ESP}$	$Q(O^-)_{GAPT}$	$Q(O^-)_{Gast}^{\pi}$	$Q(O^-)_{Löw}$	$Q(O^-)_{Mul}$	$Q(O^-)_{NPA}$
1	0.903	-0.063	0.584	0.984	0.995	0.996
$Q(O^-)_{ESP}$	1	-0.152	0.613	0.874	0.876	0.885
$Q(O^-)_{GAPT}$		1	0.106	-0.073	-0.069	-0.084
$Q(O^-)_{Gast}^{\pi}$			1	0.533	0.554	0.542
$Q(O^-)_{Löw}$				1	0.988	0.994
$Q(O^-)_{Mul}$					1	0.997
$Q(O^-)_{NPA}$						1

^a Note that $n = 17$ for $Q_{Löw}$.

for the dissociation process also were successful in this respect, with $Q(N)_{NPA} = Q(N^+)_{NPA} - Q(N)_{NPA}$ yielding a correlation $r = 0.937$ for the aniline pK_a 's, and $Q(O)_{NPA} = Q(O)_{NPA} - Q(O^-)_{NPA}$ yielding $r = 0.904$ for the phenol series.

INTERCORRELATIONS

Substituents cause changes in the electron density at the dissociating functional groups in aniline and phenol. Tables VII and VIII show the correlations among the atomic charge models for the aniline N and H⁺ and the phenol O and H cases, respectively. As expected, the charges that correlated strongly with pK_a also tended to correlate strongly between one another. In general the AIM, Löwdin, and NPA charges were strongly correlated in their variations.

Conclusions

A variety of methods exist for calculating partial atomic charges. To gauge their usefulness, the abilities of seven popular schemes to correlate with substituent-induced pK_a shifts for a series of ani-

lines and phenols have been examined. On this basis, the AIM, Löwdin, and NPA charges are observed to be quite useful and effective measures. The GAPT and Mulliken charges were also fairly successful in correlating with the observed pK_a shifts of the examined series, but both schemes exhibited difficulties in at least one data set. Although the charge measures $Q(N)_{AIM}$ and $Q(N)_{GAPT}$ both produced good regression results, their variations appeared to be counterintuitive in the sense that they suggested that an increase in the aniline pK_a was accompanied by a decrease in the negative charge on the nitrogen atom. By the admittedly limited standard of usefulness of correlating with the substituent-induced pK_a changes in the anilines and phenols, the electrostatic and Gasteiger charges were not successful, although these charges have produced informative results in other contexts. Judging from the qualities of the regression results, the properties of the functional groups of the bases, i.e., the substituted anilines and phenoxide ions, as represented by the charges $Q(N)$, $Q(NH_2)$, and $Q(O^-)$, are especially effective indices in relation to the pK_a variations observed experimentally.

TABLE VIIICross correlations (*r*) between each of the atomic charges for the phenol and phenoxide data sets.^a

Q(H) Cross-correlation matrix							
	Q(H) _{AIM}	Q(H) _{ESP}	Q(H) _{GAPT}	Q(H) _{Gast} ^σ	Q(H) _{Löw}	Q(H) _{Mul}	Q(H) _{NPA}
Q(H) _{AIM}	1	0.699	0.952	0.63	0.988	0.999	0.999
Q(H) _{ESP}		1	0.561	0.634	0.627	0.686	0.684
Q(H) _{GAPT}			1	0.573	0.955	0.945	0.946
Q(H) _{Gast} ^σ				1	0.623	0.63	0.624
Q(H) _{Löw}					1	0.987	0.989
Q(H) _{Mul}						1	0.999
Q(H) _{NPA}							1

Q(O ⁻) Cross-correlation matrix							
	Q(O ⁻) _{AIM}	Q(O ⁻) _{ESP}	Q(O ⁻) _{GAPT}	Q(O ⁻) _{Gast} ^π	Q(O ⁻) _{Löw}	Q(O ⁻) _{Mul}	Q(O ⁻) _{NPA}
Q(O ⁻) _{AIM}	1	0.903	-0.063	0.584	0.984	0.995	0.996
Q(O ⁻) _{ESP}		1	-0.152	0.613	0.874	0.876	0.885
Q(O ⁻) _{GAPT}			1	0.106	-0.073	-0.069	-0.084
Q(O ⁻) _{Gast} ^π				1	0.533	0.554	0.542
Q(O ⁻) _{Löw}					1	0.988	0.994
Q(O ⁻) _{Mul}						1	0.997
Q(O ⁻) _{NPA}							1

^a Note that *n* = 17 for Q_{Löw} and *n* = 18 for Q(H)_{AIM}.**ACKNOWLEDGMENT**

P.G.S. thanks Prof. Johann Gasteiger for helpful conversations related to this work. C.M.H. acknowledges support from the National Science Foundation (CHE-9733457) and the Ohio Supercomputer Center.

References

- Williams, D. E. In *Reviews of Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; 1991; Vol. II, pp. 219–271.
- Bachrach, S. M. In *Reviews of Computational Chemistry*; Lipkowitz, K. B.; Boyd, D. B., Eds.; 1994; Vol. V, pp. 171–227.
- (a) Wiberg, K. B.; Hadad, C. M.; Breneman, C. M.; Laidig, K. E.; Murcko, M. A.; LePage, T. J. *Sci* 1991, 252, 1266–1272.
(b) Wiberg, K. B.; Rablen, P. R. *J Comput Chem* 1993, 14, 1504.
- Mulliken, R. S. *J Chem Phys* 1955, 23, 1833, 1841, 2338, and 2343.
- Mulliken, R. S. *J Chem Phys* 1966, 36, 3428.
- Wang, B.; Ford, G. P. *J Comput Chem* 1994, 15, 200–207.
- Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J Chem Phys* 1985, 83, 735–746.
- Weinhold, F. In *Encyclopedia of Computational Chemistry*; Wiley: New York, XXXX; Vol. 3, pp. 1792–1811.
- Löwdin, P.-O. *Adv Quantum Chem* 1970, 5, 185–199.
- Cusachs, I. C.; Politzer, P. *Chem Phys Lett* 1968, 1, 529.
- Löwdin, P.-O. *Phys Rev* 1955, 97, 1474–1489.
- Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J Chem Phys* 1985, 83, 735–746.
- Weinhold, F. In *Encyclopedia of Computational Chemistry*; Wiley: New York, XXXX; Vol. 3, pp. 1792–1811.
- Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, 1990.
- Cioslowski, J. *J Am Chem Soc* 1989, 111, 8333–8336.
- Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J Phys Chem A* 1998, 102, 1820–1831.
- Sanderson, R. T. *Chemical Periodicity*; Reinhold: New York, 1960.
- Gasteiger, J.; Marsili, M. *Tetrahedron* 1980, 36, 3219–3228.
- Mortier, W. J.; Van Geenchten, K.; Gasteiger, J. *J Am Chem Soc* 1985, 107, 829–835.
- Gasteiger, J.; Marsili, M. *Acta* 1980, 53, 601–614.
- Mullay, J. *J Am Chem Soc* 1986, 108, 1770–1775.
- Abraham, R. J.; Griffiths, L.; Loftus, P. *J Comput Chem* 1982, 3, 407–416.
- Abraham, R. J.; Smith, P. E. *J Comput Chem* 1987, 9, 288–297.
- No, K. T.; Grant, J. A.; Scheraga, H. A. *J Phys Chem* 1990, 94, 4732–4739.
- Park, J. M.; No, K. T.; Jhon, M. S.; Scheraga, H. A. *J Comput Chem* 1993, 14, 1482–1490.
- Rappé, A. K.; Goddard, W. A. III. *J Phys Chem* 1991, 95, 3358–3363.

27. Dixon, S. L.; Jurs, P. C. *J Comput Chem* 1992, 13, 492–504.
28. Dixon, S. L.; Jurs, P. C. *J Comput Chem* 1993, 14, 1460–1467.
29. Gross, K. C.; Seybold, P. G. *Int J Quantum Chem* 2000, 80, 1107–1115.
30. Gross, K. C.; Seybold, P. G. *Int J Quantum Chem* 2001, 85, 569–579.
31. Gross, K. C.; Seybold, P. G.; Peralta-Inga, Z.; Murray, J. S.; Politzer, P. *J Org Chem* 2001, 66, 6919–6925.
32. (a) Becke, A. D. *Phys Rev A* 1988, 38, 3098. (b) Becke, A. D. *J Chem Phys* 1993, 98, 5648. (c) Lee, C.; Yang, W.; Parr, R. G. *Phys Rev B* 1988, 37, 785.
33. (a) Labanowski, J. W.; Andzelm, J. *Density Functional Methods in Chemistry*; Springer: New York, 1991. (b) Parr, R. G.; Yang, W. *Density Functional Theory in Atoms and Molecules*; Oxford University Press: New York, 1989.
34. Spartan 5.1, Wavefunction, Inc., 18401 Von Karman, Suite 370 Irvine, CA 92612.
35. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrewski, V. G.; Montgomery, Jr., J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farhas, 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-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian98, Revision A.9*; Gaussian: Pittsburgh, 1998.
36. AIMALL, Todd Keith, Yale University.
37. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J Comput Chem* 1993, 14, 1347–1363.
38. Albert, A.; Serjeant, E. P. *Ionization Constants of Acids and Bases*; Meuthen: London, 1962.
39. Besler, B. H.; Merz, K. M.; Kollman, P. A. *J Comput Chem* 1990, 11, 431.
40. (a) Chirlian, L. E.; Franci, M. M. *J Comput Chem* 1987, 8, 894. (b) Breneman, C. M.; Wiberg, K. B. *J Comput Chem* 1990, 11, 361.
41. Gasteiger, J.; Marsili, M. *Org Magn Reson* 1980, 15, 353–360.
42. Gasteiger, J.; Guillen, M. D. *J Chem Res (S)* 1983, 304–305.