DOI: 10.1002/ejoc.200800652

Theoretical Description of Substituent Effects in Electrophilic Aromatic Substitution Reactions

Tobias Schwabe[a] and Stefan Grimme*[a]

Keywords: Phenol / Proton affinity / Electrophilic substitution / Ab initio and DFT calculations / Double-hybrid density functionals

The ability of the current Kohn–Sham density functional theory (DFT) to compute the change of the proton affinity (PA) of phenol derivatives due to substitution is investigated. These systems can be used as models to predict reactivities in electrophilic aromatic substitution reactions. The complexity of the problem is increased systematically by introducing successively up to four substituents in five typical cases (methyl, cyano, fluorine, chlorine, and bromine). Our investigation can be regarded as representative for an important class of problems consistently encountered in the DFT modeling of organic reactions. High-level theoretical reference data from CCSD(T) and SCS-MP2 wave-function calculations are presented, and the PAs are compared to those obtained by a series of density functionals (DFs). It is shown that not all DFs are capable of quantitatively reproducing the substituent effects. These can be simply linear in the number of substituents or show more complicated patterns. Especially for halogens, some DFs even fail completely. In these cases, linearly increasing errors with the number of groups are observed. Reliable results are obtained with hybrid DFs or the even more accurate double-hybrid DF approach. The errors are attributed to the common self-interaction (over-delocalization) error in part of the DFs. Comparison with Hartree–Fock results shows that a reliable account of electron correlation is necessary to compute the PA of unsaturated and highly substituted molecules with chemical accuracy.

Introduction

The special properties of functional groups lay the foundations for the richness and complexity of organic chemistry. Their diversity and complex traits allow us to modify and tune the electronic properties of compounds for a desired target or function. Therefore it is of great interest to predict and quantify by theoretical means the effects of substitution. This does not only lead to a better understanding, but also allows us to develop new compounds "from scratch" – the goal of theoretical rational design.[1–3]

The advent and success of the Kohn–Sham density functional theory (DFT) yields promising access to the theoretical investigation of substitution effects because of its very good relation between accuracy and computational effort,[4,5] which enables the treatment of systems up to a few hundred atoms routinely.

DFT is the main method for computational investigations in various fields of organic chemistry. Good results have been reported for basic thermodynamic data,[6,7] isomerization reactions,[8] and even reaction barriers.[9,10] New density functionals (DFs), like the meta-hybrid-generalized gradient-approximated (GGA) functionals (BMK,[11] M06 family[12,13]), long-range-corrected functionals (ωB97[14]), and the double-hybrid approach (B2PLYP[15,16]), often yield results within the so-called chemical accuracy (about 1 kcal mol^{-1}). The problems in describing London-dispersion (van der Waals) effects have also been solved empirically, such that large biological or supramolecular systems can be investigated with similar accuracy.[17–19]

Recently, some warnings[20] and important problems[21–24] concerning the uncritical usage of DFT in its current form for organic reactions have been reported. The seemingly simple case of alkane isomerization[24] has attracted much attention beyond the theoretical community and has led to further intensive research on the general accuracy of widely used functionals in particular for organic chemistry.[25–27] Thus, it seems worthwhile to ask how accurately common substitution effects are described on a standard DFT level. This is the main topic of the present work.

We carried out a thorough investigation of a model reaction, namely the protonation of a multi-substituted phenol, and report how well the effects of substitution on the proton affinity (PA) are described. The protonated species can be regarded as a prototype for the π-complex in electrophilic aromatic substitution reactions (S_EAr). Substituent effects are typically interpreted by using Hammett con-

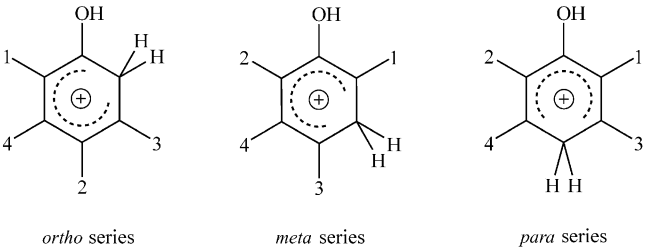
---

[a] Organisch-Chemisches Institut, Universität Münster,
Corrensstraße 40, 48149 Münster
Fax: +49-251-83-36515
E-mail: grimmes@uni-muenster.de

Supporting information for this article is available on the WWW under http://www.eurjoc.org or from the author.

stants[28] which are, however, not generally available for multiple substitutions. Incremental schemes may be a solution, but accurate and reliable ones are limited to special systems. Theoretical methods may be helpful in complicated cases, but the final predictions are, of course, sensitive to small (relative) energy changes. Our investigation should provide insight into which level of theory is necessary to achieve satisfactory results for such a task. Note that because all our treatments refer to single-molecule (gas-phase) conditions, no direct comparison with experimental results, which are routinely obtained in solution, is possible (although achieving comparability is straightforward in principle by applying appropriate theoretical solvation models).

Starting from phenol (X = H), substituents are introduced into the system in a successive way. This choice was triggered by unpublished experimental studies on aromatic ethers for which phenol is a good model. The following functional groups are considered: methyl, cyano, fluorine, chlorine, and bromine. The three possible substitution series are depicted in Scheme 1. The labeling refers to the protonation position relative to the hydroxy group.



Scheme 1. General series for *ortho*, *meta*, and *para* protonation of substituted phenols. The substituents X = CH_3, CN, F, Cl, Br are subsequently placed at the positions 1–4.

Variation of X and the number of substituents should mainly reflect the electronic properties of X when it is interacting with the (charged) π-system. Because the addition of protons (without counterions) to stronger bases is usually barrierless,[29] it can be expected that a larger PA corresponds to an increased reactivity (substitution rate) of the substituted compound compared to phenol. The π-complex that can be relevant as an intermediate of real systems in solution[29] is not considered here.

Although substituent effects in electrophilic aromatic substitution reactions are textbook knowledge, not many systematic theoretical studies have been carried out on this topic. A related set of systems has been investigated theoretically by Tishchenko et al. but with emphasis on the protonation site[30] also including the hydroxy group. Furthermore, only monosubstituted phenols have been considered. One of the main aspects of the present work is the question of how to perform accurate computations of multiple substituent effects. These were considered previously for arene protonation at the MP2 level by Eckert-Maksic et al.[31] These authors showed that an additivity model holds for the PA of polysubstituted (X = CH_3, CN, F) benzenes.

When dealing with π-systems on the DFT level of theory, it is well known that the description might suffer from the so-called self-interaction error (SIE).[22,32–34] This term describes the observation that most of the common exchange-correlation functionals do not cancel out the unphysical Coulomb interaction of one electron with itself. As a result, they predict delocalized radicals and transition states of many radical reactions to be too stable. A good example is arene radical cations and their aggregates, which represent a "nightmare" for DFT.[14] Consideration of the SIE is a very active field of theoretical research, and important progress in its understanding has been achieved.[35] Although in closed-shell systems the SIE is smaller (or cancels out for relative energies), related effects are still present but may be hidden by error cancellations with effects of static electron correlation. In very delocalized π-systems, the errors may not cancel each other out, leading to non-negligible errors. In ref.[35], SIE effects in the many-electron case (for which it originally has not been defined) have been dubbed delocalization error (DLE). This term is used also in the following to discuss our observed errors for the PA of phenols. In a related study, the PA of a series of alkenes of increasing size has been used to benchmark a large set of DFs.[22] When the SIE (or DLE) is a major problem, it can be expected that our methodological conclusions are also transferable to the DFT description of the attacks of radicals on unsaturated molecules.

A variety of different but representative DFs has been tested in our work (GGA: BP86[36,37] PBE,[38] meta-GGA: TPSS,[39] hybrid-GGA: B3LYP[40,41] PBE0,[42] double hybrid: B2PLYP,[15] B2PLYP-D[19]). The DFT data are also compared to uncorrelated but SIE-free Hartree–Fock (HF) results in order to rationalize the errors. As theoretical reference data, we employ the (SIE- and DLE-free) post-HF methods single and double excitation coupled-cluster with approximative triple excitations [CCSD(T)[43–45]] and the spin-component-scaled second-order Møller-Plesset perturbation theory (SCS-MP2)[46] for the larger systems. SCS-MP2 yields results very close to CCSD(T) for many main-group molecules at a fraction of computational cost (for a recent overview, see ref.[7]).

Methods and Technical Details

For all compounds, the proton affinity (PA) was computed as the differences between the total energies of the neutral reactand and the cation without vibrational or thermal contributions.

All molecular structures were obtained at the B97-D/TZVPP level of theory.[47] Subsequent single-point energies were also calculated by using the large TZVPP basis set[48] and the functionals BP86,[36,37] BLYP,[36,49] PBE,[38] TPSS,[39] B3LYP,[40,41] PBE0,[42] B2PLYP,[15] and B2PLYP-D.[19] For all semi-local (non-hybrid) functionals, the resolution-of-identity (RI) approximation (also called density fitting) for the two-electron Coulomb interactions has been applied. The corresponding optimized auxiliary basis sets[50] have been used. The reference single-point energies were obtained at the SCS-MP2/TZVPP[46] level, by also using the RI approximation with the corresponding auxiliary basis sets.[51]

The accuracy of the SCS-MP2 method has been tested by computation of PAs for phenol on the CCSD(T) level. Extrapolation to the complete-basis-set (CBS) limit was done with a standard two-point scheme[52] based on cc-pVDZ and cc-pVTZ basis-set[53] results. Because of its high computational demand, this procedure (which has an estimated error of ±0.5 kcal mol^{-1} compared to experimental data) is not feasible for most of the substituted species. Basis-set effects are also considered only for phenol by comparison of SCS-MP2/CBS and SCSMP2/TZVPP results.

The zero-point vibrational-energy contributions to the PA have been computed in an exemplary manner for phenol, 2-fluorophenol and 2,3,4,5-tetrafluorophenol on the B3LYP/TZVPP level. The contributions are almost constant, independent of the number of substituents, with values between 6.8 and 7.3 kcal mol^{-1} (i.e., lowering the electronic PA). Thus, these corrections are not expected to play a significant role for our conclusions, and are therefore not included.

Most of the computations have been carried out with a locally modified version of TURBOMOLE 5.7.[54–56] For the coupled-cluster calculations, MOLPRO[57] has been used.

Supporting Information is available with theoretical PAs on the SCS-MP2/TZVPP level of theory for all substituted phenols (Tables S1–S5), including the deviations for HF and the functionals BP86, BLYP, PBE, TPSS, PBE0, B3LYP, B2PLYP and B2PLYP-D.

Results and Discussion

In the following discussion, only DFT results for the functionals PBE, PBE0, the popular B3LYP, and B2PLYP-D are shown. These functionals represent prototypes for the different functional classes: semi-local without HF exchange (represented by PBE), hybrid functionals (represented by PBE0 and B3LYP), and double-hybrid functionals also including a perturbation correction for non-local correlation effects (represented by B2PLYP). The latter DF is also augmented by a semi-empirical correction for long-range London dispersion effects (DFT-D[47]). The DFT-D correction is not of utmost importance for the reactions considered here, but the B2PLYP-D approach is one of the most consistent and accurate state-of-the-art methods for general thermochemistry.[7,19] Results for all other functionals mentioned above can be found in the Supporting Information.

Before the quality of the DFT results is considered, we first validated our reference method. This was done by comparing SCS-MP2 results for the PA in *ortho*, *meta* and *para* positions of phenol with high-level CCSD(T) data. The deviations of SCS-MP2/CBS and SCS-MP2/TZVPP with respect to CCSD(T)/CBS, along with HF and DFT results, are given in Table 1.

As can be seen from these data, SCS-MP2 is very accurate for our problem. At the CBS level, the deviations are small (about 1 kcal mol^{-1}) and very systematic, i.e., almost the same for the PAs in all three positions (which differ by up to 17 kcal mol^{-1}). The PAs with the TZVPP set are of course not fully converged with respect to basis set size, but generally within 1 kcal mol^{-1} of the CBS data. We estimate overall SCS-MP2/TZVPP errors compared to the "true" electronic PA of about ±2 kcal mol^{-1} (about 1%). The regioselectivity should be more accurate to about ±1 kcal mol^{-1}. This should be suffcient for our goal of testing the other methods that (except for B2PLYP-D) have errors in the 2–10 kcal mol^{-1} range. Note that, although B2PLYP-D has a slightly larger mean absolute error (MAE) than SCS-MP2, the regioselectivity predictions are more accurate. One of the interesting questions at this point is if the DLE of the (not completely SIE-free) B2PLYP-D functional increases in particular for highly substituted phenols.

Tables 2, 3, 4, 5, and 6 list the computed PAs for the substituents fluorine, chlorine, bromine, cyano, and methyl and the deviations (errors) of the considered methods for each of the three substitution series according to Scheme 1. All these results refer to the TZVPP basis set and employ SCS-MP2 as reference. Note that the additional substituents in the *ortho*, *meta*, and *para* series enter in various positions relative to the protonation site. For example in the *meta* series, the first two additional substituents are placed in *ortho* and *para* position relative to the protonation site, whereas this relation becomes *meta* and *meta*, respectively, in the *ortho* series.

Before discussing the performance of the DFT methods, the magnitude and sign of the substituent effects on the PA are considered. Therefore, the SCS-MP2/TZVPP reference data are also shown graphically in Figure 1. Clearly, the phenols are more easily protonated in the *ortho* and *para* positions than in the *meta* position, in agreement with common textbook knowledge. The *ortho* and *para* PAs are in general very similar. The only exceptions are the di- and

Table 1. Computed proton affinities on the CCSD(T)/CBS level of theory (PA$^{\text{ref}}$) for phenol and deviations for different methods. All values are in kcal mol^{-1}.

Protonation site	PA$^{\text{ref}}$	Deviation[a]					
		SCS-MP2/CBS	SCS-MP2	HF	PBE	PBE0	B2PLYP-D
ortho	197.9	1.4	0.6	–9.5	–3.1	–3.0	–1.6
meta	184.5	1.2	1.4	–3.8	–4.0	–2.4	–1.2
para	201.2	1.2	0.4	–9.9	–2.3	–2.6	–1.2

[a] Deviation = PA(reference) – PA(method). Based on TZVPP results if not noted otherwise.

trisubstituted fluorophenols, for which the *ortho* and *meta* values are almost equal.

The effects of methyl substitution on the cation are easily rationalized. Each additional substituent stabilizes the positive charge by roughly 4 kcal mol^{-1} because of its +I effect, revealing an almost linear behavior. For the CN substituent, the –I effect results in a decrease of the PA by about 10 kcal mol^{-1} for each successive substitution, also showing a linear dependency. This is in agreement with previous results of Eckert-Maksic et al.[31] No mesomeric stabilization by CN is observed. The effects are similar in magnitude for all three positions, although the *meta* series qualitatively gives slightly different curves than the *ortho* and *para* ones.

In contrast, the curves for the halogens are somewhat different between the protonation sites, although they are alike among each other. In the *ortho* and *para* series, the PAs decrease to a minimum for the disubstituted chlorine and bromine species and slightly increase for the third and fourth substitutions. For fluorine, the PAs become more or less constant after the second substitution. In the *meta* series, all halogens induce little variation of the PA.

We interpret this by invoking the dual electronic character of the halogens. On the one hand they destabilize the protonated species because of their –I effect, whereas on the other hand the +M effect allows better delocalization of the positive charge if appropriate mesomeric structures are possible. Thus, the number of substituents is as important

Table 2. Computed proton affinities on the SCS-MP2/TZVPP level of theory (PA$^{\text{ref}}$) for methyl-substituted phenols and the deviations for different methods (TZVPP basis set). All values are in kcal mol^{-1}.

Protonation site	*n*[b]	PA$^{\text{ref}}$	Deviation[a]				
			HF	PBE	PBE0	B3LYP	B2PLYP-D
ortho	0	197.3	–10.2	–3.7	–3.7	–4.2	–2.3
	1	200.4	–9.3	–4.1	–3.7	–4.2	–2.3
	2	202.5	–8.7	–4.7	–4.0	–4.5	–2.5
	3	207.6	–9.4	–5.5	–4.8	–5.2	–3.2
	4	213.1	–10.6	–5.7	–5.3	–5.5	–3.5
meta	0	183.1	–5.2	–5.5	–3.8	–4.7	–2.6
	1	189.8	–6.6	–5.8	–4.3	–5.1	–3.0
	2	197.1	–8.4	–6.5	–5.3	–5.9	–3.6
	3	202.2	–9.6	–7.2	–6.2	–6.6	–4.2
	4	205.0	–8.6	–7.2	–6.0	–6.3	–4.1
para	0	200.8	–10.3	–2.7	–3.0	–3.3	–1.6
	1	204.2	–10.0	–3.1	–3.2	–3.5	–1.7
	2	207.0	–9.5	–3.7	–3.5	–3.7	–1.9
	3	211.7	–10.0	–4.3	–4.2	–4.3	–2.4
	4	217.0	–10.9	–4.6	–4.6	–4.6	–2.8

[a] Deviation = PA(reference) – PA(method). [b] Number of substituents.

Table 3. Computed proton affinities at the SCS-MP2/TZVPP level of theory (PA$^{\text{ref}}$) for cyano-substituted phenols and the deviations for different methods (TZVPP basis set). All values are in kcal mol^{-1}.

Protonation site	*n*[b]	PA$^{\text{ref}}$	Deviation[a]				
			HF	PBE	PBE0	B3LYP	B2PLYP-D
ortho	0	197.3	–10.2	–3.7	–3.7	–4.2	–2.3
	1	183.2	–9.5	–4.0	–3.5	–4.1	–2.3
	2	170.2	–7.9	–3.8	–2.7	–3.5	–1.9
	3	162.1	–4.8	–5.5	–3.0	–4.0	–2.4
	4	154.1	–2.4	–7.4	–3.6	–4.9	–3.1
meta	0	183.1	–5.2	–5.5	–3.8	–4.7	–2.6
	1	172.3	–1.9	–5.6	–2.7	–4.0	–2.2
	2	161.6	3.4	–6.8	–2.2	–3.8	–2.0
	3	152.5	8.8	–7.8	–1.5	–3.5	–1.8
	4	142.6	8.1	–9.6	–2.7	–4.8	–3.0
para	0	200.8	–10.3	–2.7	–3.0	–3.3	–1.6
	1	186.9	–9.0	–2.5	–2.3	–2.8	–1.2
	2	174.2	–7.6	–2.7	–1.9	–2.5	–1.0
	3	165.1	–5.1	–4.3	–2.2	–3.1	–1.6
	4	156.8	–2.3	–5.9	–2.4	–3.6	–2.1

[a] Deviation = PA(reference) – PA(method). [b] Number of substituents.

Table 4. Computed proton affinities on the SCS-MP2/TZVPP level of theory (PA$^{\text{ref}}$) for fluorine-substituted phenols and the deviations for different methods (TZVPP basis set). All values are in kcal mol^{-1}.

Protonation site	*n*[b]	PA$^{\text{ref}}$	Deviation[a]				
			HF	PBE	PBE0	B3LYP	B2PLYP-D
ortho	0	197.3	–10.2	–3.7	–3.7	–4.2	–2.3
	1	190.3	–8.3	–5.1	–4.2	–4.5	–2.6
	2	184.4	–6.8	–6.1	–4.5	–4.5	–2.6
	3	182.2	–6.0	–7.4	–5.1	–4.9	–3.1
	4	184.7	–5.7	–7.0	–4.8	–4.0	–2.6
meta	0	183.1	–5.2	–5.5	–3.8	–4.7	–2.6
	1	182.7	–5.7	–5.6	–3.9	–4.3	–2.5
	2	185.5	–5.9	–5.5	–3.8	–3.6	–2.1
	3	184.8	–5.7	–5.9	–4.0	–3.3	–2.1
	4	178.8	–3.5	–7.3	–4.4	–3.5	–2.3
para	0	200.8	–10.3	–2.7	–3.0	–3.3	–1.6
	1	195.0	–9.0	–3.7	–3.3	–3.4	–1.7
	2	188.7	–7.2	–5.0	–3.8	–3.7	–2.0
	3	188.4	–6.5	–5.3	–3.7	–3.2	–1.8
	4	187.0	–5.7	–5.7	–3.8	–3.0	–1.7

[a] Deviation = PA(reference) – PA(method). [b] Number of substituents.

Table 5. Computed proton affinities on the SCS-MP2/TZVPP level of theory (PA$^{\text{ref}}$) for chlorine-substituted phenols and the deviations for different methods (TZVPP basis set). All values are in kcal mol^{-1}.

Protonation site	*n*[b]	PA$^{\text{ref}}$	Deviation[a]				
			HF	PBE	PBE0	B3LYP	B2PLYP-D
ortho	0	197.3	–10.2	–3.7	–3.7	–4.2	–2.3
	1	191.5	–8.7	–4.7	–3.9	–4.4	–2.5
	2	186.2	–7.5	–5.6	–4.2	–4.6	–2.7
	3	186.7	–5.8	–6.9	–4.7	–5.1	–3.3
	4	189.3	–4.3	–7.4	–4.8	–5.0	–3.2
meta	0	183.1	–5.2	–5.5	–3.8	–4.7	–2.6
	1	183.1	–4.0	–5.7	–3.6	–4.3	–2.6
	2	185.8	–2.4	–6.4	–3.7	–4.2	–2.5
	3	185.8	–1.2	–7.2	–4.0	–4.4	–2.8
	4	182.9	0.5	–7.9	–4.1	–4.5	–3.0
para	0	200.8	–10.3	–2.7	–3.0	–3.3	–1.6
	1	195.6	–8.9	–3.5	–3.1	–3.5	–1.7
	2	190.3	–7.6	–4.6	–3.5	–3.8	–2.0
	3	191.3	–5.9	–5.3	–3.6	–3.8	–2.2
	4	192.2	–4.6	–5.9	–3.7	–3.8	–2.4

[a] Deviation = PA(reference) – PA(method). [b] Number of substituents.

Table 6. Computed proton affinities at the SCS-MP$_2$/TZVPP level of theory (PA^{ref}) for bromine-substituted phenol and the deviations for different methods (TZVPP basis set). All values are in kcal mol^{-1}.

Protonation site	n[b]	PA^{ref}	Deviation[a]				
			HF	PBE	PBE0	B3LYP	B2PLYP-D
ortho	0	197.3	–10.2	–3.7	–3.7	–4.2	–2.3
	1	192.3	–8.7	–4.6	–3.8	–4.4	–2.4
	2	187.3	–7.6	–5.4	–4.0	–4.5	–2.6
	3	188.4	–5.7	–6.7	–4.4	–5.0	–3.2
	4	191.1	–3.7	–7.5	–4.7	–5.0	–3.3
meta	0	183.1	–5.2	–5.5	–3.8	–4.7	–2.6
	1	183.5	–3.5	–5.6	–3.3	–4.2	–2.5
	2	186.4	–1.0	–6.5	–3.4	–4.1	–2.5
	3	186.8	–0.7	–7.4	–3.6	–4.2	–2.7
	4	184.5	2.1	–8.0	–3.7	–4.3	–3.0
para	0	200.8	–10.3	–2.7	–3.0	–3.3	–1.6
	1	196.2	–8.8	–3.5	–3.0	–3.4	–1.6
	2	191.4	–7.6	–4.5	–3.4	–3.8	–1.8
	3	192.7	–5.6	–5.3	–3.5	–3.8	–2.1
	4	194.0	–4.2	–5.9	–3.5	–3.8	–2.4

[a] Deviation = PA(reference) – PA(method). [b] Number of substituents.

as the position in which they are placed. One should keep in mind that the form of the curves in Figure 1 is somewhat arbitrary because they are also influenced by the (arbitrary) substitution order depicted in Scheme 1.

All these mentioned aspects have to be reproduced by a theoretical method which claims to be able to quantify substitution effects. To investigate the DFT methods with respect to this requirement, the DFT errors are listed in Tables 2, 3, 4, 5, and 6. For more clarity, they are also shown in Figure 2, taking chlorine as representative for the halogens. Note that approximately flat curves in these plots (constant error with respect to the number of substituents)

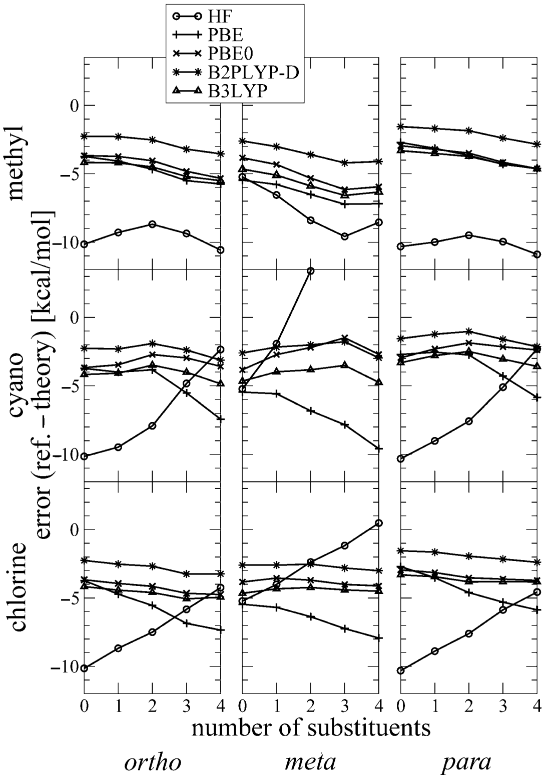


Figure 2. Errors (with respect to SCS-MP2/TZVPP) of HF and various DFT methods for the PAs in the series of substituted phenols.

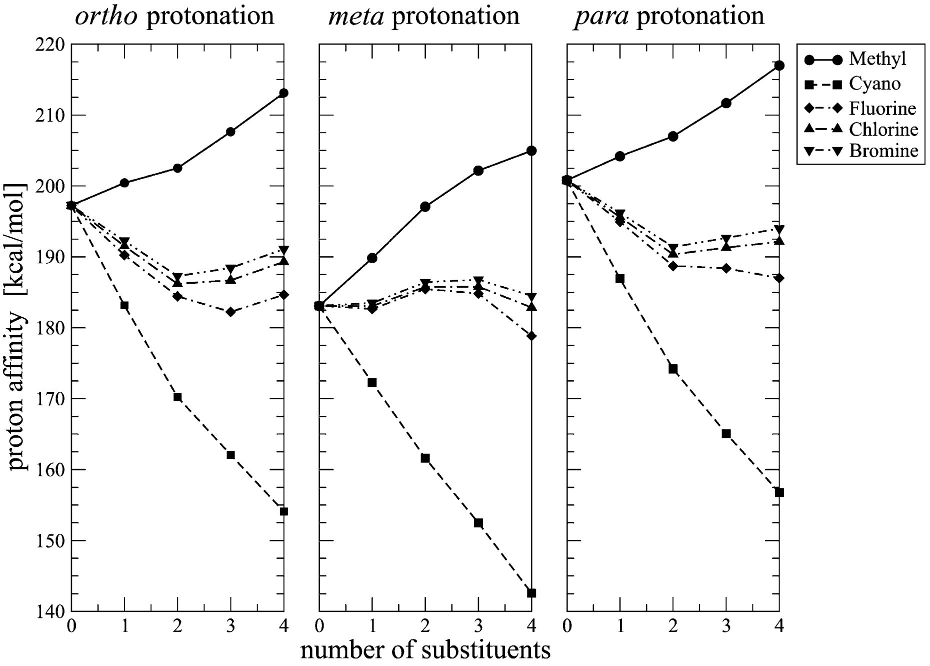


Figure 1. Proton affinities for substituted phenols based on SCS-MP2/TZVPP calculations. The labels *ortho*, *meta*, and *para* refer to the protonation site relative to the hydroxy group.

mean that the respective method solely provides an absolute error for the PA but describes the substituent effect correctly. For an appropriate isodesmic reaction (e.g. proton exchange between phenol and substituted phenol) the corresponding error would be small.

For the effect of the methyl group, there seems to be at least qualitatively no problem. All methods yield an almost constant error, independent of the number of groups. The absolute errors are largest for HF (about 10 kcal mol^{-1}), they are about halved by the PBE, PBE0 and B3LYP approaches, and further halved for B2PLYP-D. The errors are all negative, meaning that the PA is overestimated, which generally holds for most of the computed values. For DFT, this can be explained by an overstabilization due to overdelocalization of the protonated (stronger π-delocalized) species (see Figure 3). Consistent with this picture is that the error is largest for (uncorrelated) HF. This can be attributed to an absolutely larger correlation energy (that is neglected in HF theory) in the neutral than in the protonated form (electron correlation typically favors more localized electronic structures). The implicit correlation treatment in the DFs at least qualitatively corrects in the right direction.

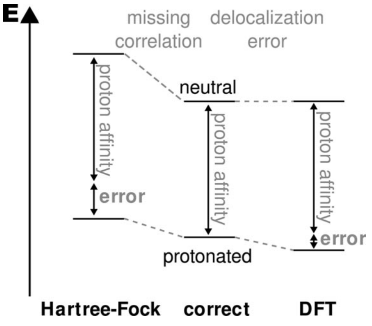


Figure 3. Schematic explanation of the observed errors for PAs of phenols calculated with HF and KS-DFT based methods. In DFT, the correlation energy is computed rather accurately for the neutral and cationic species, but the DLE overstabilizes the cation, which leads to a too large PA. In the HF method, the sign of the error is the same but arises from the neglect of electron correlation, which is larger in the neutral electronically more localized system.

The different functionals form a hierarchy regarding their accuracy (B2PLYP-D > PBE0 ≈ B3LYP > PBE), which correlates nicely with their theoretical complexity. The PBE0 hybrid functional with 25% HF exchange performs slightly better than B3LYP, which contains only 20% HF exchange. The very small errors of B2PLYP-D (2–3 kcal mol^{-1}) can be explained by the increased portion of exact exchange (53%). In fact, the B2PLYP-D-computed PAs are close to those of the SCS-MP2 reference method, and at this point it seems diffcult to judge which method performs better.

The CN and chlorine substituents are more diffcult to describe, in particular by HF and the PBE non-hybrid functional. The errors of HF for both CN and Cl substitutions show (opposed to the case of methyl) a large positive slope with increasing number of substituents. Although the absolute HF errors initially become smaller with increasing substitution, this is due to error compensation, and the description is totally unbalanced for the differently sized systems. This error compensation is especially evident for the chlorine *meta* series, where a simple statistical analysis of the methodological performance for the absolute error of the PA would rank HF even better than most of the DFs, although the description of the substituent effects is rather inconsistent. The positive slope can be interpreted such that the increasing instability of the cation (the –I effect) is underestimated. Especially in the *meta* series of the cyanophenols, the HF-calculated PAs of multiply substituted systems have large errors (not visible on the chosen scale). A similar, but less pronounced, behavior is seen for the other halogen-substituted systems.

At first glance, DFT seems to behave much more favorably than the other methods, which one would expect because the error of the HF method is likely caused by the neglect of correlation effects, which are implicitly incorporated within DFT. However, a closer look reveals that PBE also suffers from an increasing error with the number of substituents but in the opposite direction to HF. This again indicates an overstabilization of the protonated systems with increasing number of substituents, a behavior that is also observed for all other tested GGAs (see Supporting Information); meta-GGAs like the investigated TPSS functional offer no significant advantage.

The hybrid and double-hybrid approaches suffer less from this effect and provide results closer to the reference values. Also for the Cl and CN substituents, the B2PLYP-D error is smallest (about 2 kcal mol^{-1}) and furthermore only slightly dependent on the number of substituents. As already noted for the methyl case, the PBE0 functional performs a bit better than B3LYP, which is most clearly visible for the *meta* series of the cyanophenols.

Conclusions

We have investigated the ability of HF and several density functionals to describe substituent effects on the protonation of multisubstituted phenols. The typical groups methyl, cyano, fluorine, chlorine, and bromine have been considered. As reference we have used the SCS-MP2/TZVPP approach, which in turn has been proven to yield very small errors compared to estimated CCSD(T)/CBS data.

It has been shown that the effects on the PA depend on kind, placement and number of substituents. The HF method is not able to give a qualitatively correct description of these effects for most of the substituents. Similarly, (meta-)GGA-type functionals fail for this purpose, but the reason is different. The analysis of the errors for different systems leads to the conclusion that the overdelocalization error is the major problem of common density functionals. In the HF method, a similar overestimation of the PA as by (meta-)GGAs is observed, but could be traced back to a different reason (neglect of electron correlation). For both levels of theory an almost linear dependence of the error on the number of substituents has been found.

By using hybrid and double-hybrid methods, qualitatively correct results are obtained. The reduced self-interaction (delocalization) error is certainly the main reason for this. With the best DFT approach (B2PLYP-D), the absolute PA values for multiply substituted phenols are accurate to within 2–3 kcal mol^{-1} (almost independent of the number of substituents). Errors for relative values and regioselectivity predictions are typically <1 kcal mol^{-1}. This seems to be suffcient for most purposes (and also within the accuracy of the reference method).

Our study shows once more for a typical example occurring in organic chemistry, that DFT can be a valuable tool for the study of substituent effects. It also suggests that thorough validation is still needed. Only the use of accurate reference data for the whole substitution series allowed us to identify the problems of (meta-)GGA functionals and to trace their failure to the overdelocalization error. It may be dangerous to rely on general validation studies when choosing a computational method for a problem at hand. Chemistry is too diverse, and current DFT is sometimes not predictive enough (as other recent studies have also proven). Therefore, we want to encourage the use of more sophisticated functionals in future reactivity studies based on DFT.

Supporting Information (see footnote on the first page of this article): Theoretical PAs at the SCS-MP2/TZVPP level of theory for all substituted phenols (Tables S1–S5) including also the deviations for HF and the functionals BP86, BLYP, PBE, TPSS, PBE0, B3LYP, B2PLYP and B2PLYP-D.

Acknowledgments

We appreciate financial support by the Deutsche Forschungsgemeinschaft in the framework of the SFB "Molecular Orientation and its Function in Chemical Systems".

---

[1] C. Lipinski, A. Hopkins, *Nature* **2004**, *432*, 855.
[2] M. Wang, X. Hu, D. N. Beratan, W. Yang, *J. Am. Chem. Soc.* **2006**, *128*, 3228–3232.
[3] M. C. Durrant, *Chem. Eur. J.* **2007**, *13*, 3406–3413.
[4] W. Koch, C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, Weinheim, **2000**.
[5] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford, **1989**.
[6] P. Mori-Sánchez, A. J. Cohen, W. Yang, *J. Chem. Phys.* **2006**, *124*, 091102.
[7] S. Grimme, T. Schwabe, *Acc. Chem. Res.* **2008**, *41*, 569–579.
[8] S. Grimme, M. Steinmetz, M. Korth, *J. Org. Chem.* **2007**, *72*, 2118–2126.
[9] B. Lynch, P. Fast, M. Harris, D. G. Truhlar, *J. Phys. Chem. A* **2000**, *104*, 4811–4815.
[10] S. Grimme, C. Mück-Lichtenfeld, E.-U. Würthwein, A. W. Ehlers, T. P. M. Goumans, K. Lammertsma, *J. Phys. Chem. A* **2006**, *110*, 2583–2586.
[11] A. D. Boese, J. M. L. Martin, *J. Chem. Phys.* **2004**, *121*, 3405–3416.
[12] Y. Zhao, N. E. Schultz, D. G. Truhlar, *J. Chem. Theory Comput.* **2006**, *2*, 364–382.
[13] Y. Zhao, D. G. Truhlar, *Acc. Chem. Res.* **2008**, *41*, 157–167.
[14] J. D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2008**, *128*, 084106.
[15] S. Grimme, *J. Chem. Phys.* **2006**, *124*, 034108.
[16] T. Schwabe, S. Grimme, *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.
[17] P. Jurečka, J. Černý, P. Hobza, D. R. Sahlahub, *J. Comput. Chem.* **2007**, *28*, 555–569.
[18] S. Grimme, J. Antony, T. Schwabe, C. Mück-Lichtenfeld, *Org. Biomol. Chem.* **2007**, *5*, 741–758.
[19] T. Schwabe, S. Grimme, *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
[20] P. R. Schreiner, *Angew. Chem. Int. Ed.* **2007**, *46*, 4217–4219.
[21] H. L. Woodcock, H. F. Schaefer III, P. R. Schreiner, *J. Phys. Chem. A* **2002**, *106*, 11923–11931.
[22] Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2006**, *110*, 10478–10486.
[23] P. R. Schreiner, A. A. Folkin, R. A. Pascal Jr, A. de Meijere, *Org. Lett.* **2006**, *8*, 3635–3638.
[24] S. Grimme, *Angew. Chem. Int. Ed.* **2006**, *45*, 4460–4464.
[25] S. Grimme, M. Steinmetz, M. Korth, *J. Chem. Theory Comput.* **2007**, *3*, 42–45.
[26] Y. Zhao, D. G. Truhlar, *Org. Lett.* **2006**, *8*, 5753–5755.
[27] G. I. Csonka, A. Ruzsinszky, J. P. Perdew, S. Grimme, *J. Chem. Theor. Comput.* **2008**, *4*, 888–891.
[28] E. V. Anslyn, D. A. Dougherty, *Modern Physical Organic Chemistry*, University Science Books, Sausalito, **2006**.
[29] D. Heidrich, *Phys. Chem. Chem. Phys.* **1999**, *1*, 2209–2211.
[30] O. Tishchenko, N.-N. Pham-Tran, E. S. Kryachko, M. T. Nguyen, *J. Phys. Chem. A* **2001**, *105*, 8709–8717.
[31] M. Eckert-Maksić, M. Klessinger, Z. B. Maksić, *Chem. Eur. J.* **1996**, *2*, 1251–1257.
[32] T. Bally, G. N. Sastry, *J. Phys. Chem. A* **1997**, *101*, 7923–7925.
[33] Y. Zhang, W. Yang, *J. Chem. Phys.* **1998**, *109*, 2604–2608.
[34] O. V. Gritsenko, B. Ensing, P. R. T. Schipper, E. J. Baerends, *J. Phys. Chem. A* **2000**, *104*, 8558–8565.
[35] P. Mori-Sánchez, A. J. Cohen, W. Yang, *J. Chem. Phys.* **2006**, *125*, 201102.
[36] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.
[37] J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822–8824.
[38] J. P. Perdew, K. Burke, M. Enzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
[39] J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, *Phys. Rev. Lett.* **2003**, *91*, 146401.
[40] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.
[41] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.
[42] C. Adamo, V. J. Barone, *J. Chem. Phys.* **1999**, *110*, 6158–6170.
[43] T. D. Crawford, H. F. Schaefer III, *Rev. Comput. Chem.* **2000**, *14*, 33–136.
[44] J. A. Pople, M. Head-Gordon, K. Raghavachari, *J. Chem. Phys.* **1987**, *87*, 5968–5975.
[45] T. J. Lee, A. P. Rendell, P. R. Taylor, *J. Phys. Chem.* **1990**, *94*, 5463–5468.
[46] S. Grimme, *J. Chem. Phys.* **2003**, *118*, 9095–9102.
[47] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787–1799.
[48] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* **1994**, *100*, 5829–5835.
[49] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
[50] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, *Theor. Chem. Acc.* **1997**, *97*, 119–124.
[51] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, *Chem. Phys. Lett.* **1998**, *294*, 143–152.
[52] T. Helgaker, W. Klopper, H. Koch, J. Noga, *J. Chem. Phys.* **1997**, *106*, 9639–9646.
[53] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007–1023.
[54] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162*, 165–169.
[55] M. Häser, R. Ahlrichs, *J. Comput. Chem.* **1989**, *10*, 104–111.
[56] R. Ahlrichs, M. Bär, H.-P. Baron, R. Bauernschmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliott, F. Furche, F. Haase, M. Häer, H. Horn, C. Huber, U. Huniar, M. Kattannek, C. Kölmel, M. Kollwitz, K. May, C. Ochsenfeld, H. Öhm, A. Schäfer, U. Schneider, O. Treutler, M. von Arnim, F. Weigend, P. Weis, H. Weiss, *TURBOMOLE*, version 5.7, University of Karlsruhe, **2005**, http://www.turbomole.com.

[57] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, *MOLPRO, a package of ab initio programs*, version 2006.1, http://www.molpro.net.

Received: July 2, 2008