Rationalizing the regioselectivity of substituted phenols from the FERMO concept: stereoelectronic effects on protonation and functionalization

Letícia Santos Braga, Daniel Henriques Soares Leal, Teodorico Castro Ramalho

\textsuperscript{a}Department of Chemistry, Federal University of Lavras, Lavras, Brazil; \textsuperscript{b}Institute of Physics and Chemistry, Federal University of Itajubá, Itajubá, Brazil; \textsuperscript{c}Department of Chemistry, Faculty of Science, University of Hradec Královen, Hradec Králové, Czech Republic.

\textsuperscript{a}bragaleticia08@gmail.com; \textsuperscript{b}dhsleal@unifei.edu.br; \textsuperscript{c}teo@ufla.br
Rationalizing the regioselectivity of substituted phenols from the FERMO concept: stereoelectronic effects on protonation and functionalization

The relative extent of protonation in oxygen and carbon atoms and the position of protonation in carbons depend on several factors. We seek to locate the frontier molecular orbitals involved in the protonation reactions of substituted phenols using the FERMO concept through the MOLPROJ software, to compare the computational results with experimental NMR data obtained in the literature. We evaluate computationally the stereo-electronic effects that govern reactions of aromatic electrophilic substitution using an experimental study as an example. The MOLPROJ returned a percentage of correct answers of approximately 86% in the protonation sites. The experimental results on the protonation sites were rationalized in terms of stereoelectronic effects.

Keywords: Phenols, Protonation, molecular orbital, FERMO, MOLPROJ.

Introduction

Phenols are compounds with a hydroxyl functional group directly attached to a sp\(^2\)-hybridized carbon atom of a phenyl group. Namely, they are compounds containing one or more hydroxyl groups attached to benzene or another arene ring [1].

Many manufactured products such as dyes and oils, among others, have in their composition phenolic compounds that are of natural origin [2]. They are present in vegetables in free form or linked to sugars (glycosides) and proteins [3]. These compounds also lead to the polymerization of natural polyphenols such as lignin and melanin, besides forming an important class of antioxidants, which inhibit the oxidative degradation of organic or bio-organic molecules [2,4,5]. The antioxidant activity of phenols against free radicals comes from their elimination role, which is related to their ability to react with radicals much more quickly than other organic substrates. For example, tocopherol (vitamin E) is mentioned as an efficient capture agent that can eliminate harmful peroxide radicals in the blood plasma [5,6].

Protonation and deprotonation processes in aromatic molecules such as phenols are important in organic chemistry and biochemistry [7]. Through the deprotonation of phenol, the phenoxide or phenolate ion is formed, stabilized by resonance. In comparison with phenol, the phenoxide ion is more stable because of the high displacement of the
negative charge on the aromatic ring. The resonance structures of phenol involve separation of negative and positive charges. Thus, phenol is more likely to form phenoxide, releasing the proton [6].

The preferred protonation sites for phenols have been discussed from an experimental point of view [8–12] and theoretical [13–16]. In super-acidic solutions, phenols are protonated in oxygen atoms, to form oxon ions (O-PhH$^+$), or carbon in the aromatic ring in the ortho (o-PhH$^+$), meta (m-PhH$^+$) or para (p-PhH$^+$)[7,17–20].

The relative extent of protonation in oxygen and carbon atoms and the position of protonation in carbons depend on several factors, such as electronic structure of the base, acidity and solvation properties of the medium and temperature [21]. Theoretical studies have been dedicated mainly to the exploration of the intrinsic basicities in the different positions in phenol molecules [15] and in the relationship between protonic affinities and ionization energies [18].

Characteristics such as spatial and electronic structure and physico-chemical and biochemical properties of phenols can be influenced by substituting groups [20]. A substituent is a minor structural subunit (atom or functional group) that disturbs the properties of a molecular system in a quantitative sense, maintaining its general character. That is, the substituent can mischaracterize the molecule in a measurable, but not dramatic, way by modifying its spatial and electronic structure and, consequently, its behavior [16,22].

The acidity of phenols can be influenced by other groups linked to the ring [21–24]. The presence of any substituent on the aromatic ring that can stabilize the phenoxide ion will tend to increase the acidity of the phenol. On the other hand, any substituent that destabilizes the phenoxide ion, increasing its negative charge, will decrease the acidic nature of the phenol. In other words, the presence of electron withdrawing groups in phenols will increase their acidity, while electron density donor groups reduce their acidity [20,25].

To measure the basicity of organic compounds in the gas phase, the proton affinity (PA) can be used. The proton affinity is defined as the negative of the enthalpy change for the reaction $M(g) + H^+(g) \rightarrow MH^+(g)$, being M a chemical species: molecule, radical or atom [23,25]. The higher the PA, the stronger the conjugate base and the weaker the conjugated acid in the gas phase. The PA value also illustrates the role of hydration in Brønsted acidity in the aqueous phase [26–28]. The relationships between the energies of
the highest molecular orbitals (HOMO and other frontier orbitals; HOMO: Highest Occupied Molecular Orbital) and PA are often present for families of compounds, such as phenols. However, for a large number of compounds, the energies of their HOMOs do not present a good correlation with PA values [27]. Accordingly, the question arose: why are HOMO energies good acid-base descriptors for some compounds and not for others?

Faced with such limitations, another approach has emerged to understand chemical reactivity: the concept of FERMO (Frontier Effective for Reaction Molecular Orbital), proposed by Silva and Ramalho [29–31].

This concept is based on simultaneous analysis of composition and shapes of frontier molecular orbitals to determine the real molecular orbital governing a reaction. According to this approach, a frontier molecular orbital may even correspond to the FERMO for a given reaction and not for another. The HOMO itself would only be the orbital that rules a reaction if it fulfills the requirements to be the FERMO [28,29].

A strategy to quantify the location of the FERMO was developed, leading to the construction of the MOLPROJ software, based on the use of projection operators to build the molecular orbitals (MOs) by linear combination of atomic orbitals (AOs; the LCAO approach). In this same study [31], it was possible to determine the reaction site of a series of amines and describe their acid-base behavior. Thus, the location of the FERMO would indicate the orbital in which the reaction occurs and, consequently, would point to the most favorable location for protonation.

Other computational studies also look for different approaches to rationalize reactivity properties. One of these approaches is the study of the atomic charge distribution in molecules to quantify regioselectivity [42-46]. Liu and coworkers [42] proposed a method to simultaneously quantify electrophilicity and nucleophilicity using the Hirshfeld charge. This quantification is based on the Information Conservation Principle, which states that information must be conserved before and after a molecular system is formed.

It decides both where electrophilic and nucleophilic attacks will preferentially occur, but also dictates the amount of Hirshfeld charge distribution, which correlates with experimental scales of electrophilicity and nucleophilicity. In order to have the information conservation, the system that was formed is adjusted in order to have each one of the components loaded according to the contribution of its "stock" of electronic
density. Therefore, the Hirshfeld charge should be a good descriptor for both electrophilicity and nucleophilicity [42-46].

Another very useful method for modeling molecule-molecule interactions are derived from a least squares fit to the electrostatic potential (ESP). ESP is one of the useful properties to acquire partial atomic charges suitable for modeling short- and long-range molecule-molecule interactions [46]. The grid-oriented CHELP (Charges from Electrostatic potentials method) was the first method of its kind to be developed, being modified to CHELPG (Charges from Electrostatic Potentials using a Grid based method) by Breneman and Wiberg. The CHELPG method is less dependent upon molecular orientation than the original CHELP method in which partial atomic charges are fitted to reproduce the molecular ESP at a number of points around the molecule [46,48,49,50].

In addition to computational studies, experimental liquid-phase spectroscopy studies, mainly NMR and IR spectroscopy, were widely used to locate the preferred protonation site(s) of the phenol and benzene molecules substituted as a function of solvent and temperature [17,18,32-34]. To produce information about the competitive protonation in the substituents in substituted aromatic compounds, in addition to the protonation in the ring itself, it is necessary to separate the effects of the solvent from the molecular electronic properties, whose investigations must be made in the gas phase [32].

The spectroscopic data, particularly those from NMR, are sufficiently accurate to unambiguously identify the protonation sites in aromatic compounds. Experimental information on protonation in the gas phase also comes from mass spectrometry studies involving proton transfer reactions [5,7,11,12,14,23,32,35]. Even though these studies have shown the existence of several protonated isomers of many aromatic ions, and in some cases also the specific locations of the protons, the details of their structures constantly remained confused. For these cases, the structural attribution of the various isomers sometimes depended on computational chemical calculations [14,15,36,37].

In this article, we seek to locate the frontier molecular orbitals involved in the protonation reactions of substituted phenols using the FERMO concept, quantifying the orbital coefficients of the carbon and oxygen atoms involved. Once these orbitals are identified, we seek to compare the computational results with experimental NMR data obtained in the literature. In addition, we also search to computationally evaluate these stereo-electronic effects that govern reactions of aromatic electrophilic substitution using an experimental study as an example.
Methodology

Computational details

All calculations were performed with the GAMESS software for Linux, version 30 SEPT 2017 (R2). Furthermore, all studied compounds were fully optimized using the DFT method with the functional B3LYP and the basis set 6-31G(d,p) in water solution. The description of the solvent model was carried out with the Polarizable Continuum Model (PCM) using the integral equation formalism variant (IEFPCM). No symmetry restrictions were imposed during the optimization process. No imaginary frequencies were found for the optimized geometries, which were used in all subsequent calculations and had their single point energies determined with the same functional and basis set. MOs figures were prepared using Avogadro software [31] with a contour value of 0.010.

The charge calculations have been performed with Gaussian 09 [47]. The Hirshfeld and CHELPG (Charges from Electrostatic Potentials using a Grid based method) charges are obtained from the population analysis with the keyword pop=hirshfeld and pop=chelpg, respectively.

Criteria for choosing the orbitals

Considering the Brönsted-Lowry acid-base concept, the MO that drives the protonation reactions of the investigated phenols must be centered on the atoms that bind to the proton. Using the software developed in our group, MOLPROJ [31], the investigation of these orbitals was carried out in a quantitative way from the output data generated by GAMESS.

As will be said below, MOLPROJ uses projection operators to quantify the location of FERMO in a reaction using Equation 1, in which MOs are built via LCAO [3,31]:

$$\theta_\mu = \sum C_{i\mu} \zeta_i(x)$$

(Eq. 1)

being $C_{i\mu}$ the matrix of the coefficients of the molecular orbitals and the Gaussian $\zeta_i(x)$ AOs. These are created as an orthogonal set of vectors, which have components in
each atomic orbital of a given molecule. The $S_{ij}$ overlap matrix is then defined by the set of AOs $\zeta_i(x)$, which forms a set of non-orthogonal bases [29–31]:

$$S_{ij} = \int \zeta_i(x)\zeta_j(x)dx \quad \text{(Eq. 2)}$$

Through the overlap matrix $S_{ij}$, it is possible to build a projection operator for $P$ a set of AOs $G$ [29-31]:

$$P_G = \sum_{i \in G} \sum_{j \in G} \zeta_i \zeta_j^{-1} \zeta_j \quad \text{(Eq. 3)}$$

The projector can be understood as the projection of a “shadow” of a selected MO in the subspace of a set of AOs [3,31], thus making it a quantitative characterization of the shape of MO in a set of AOs and, consequently, a set of atoms.

Thus, it is possible to define the degree of localization $\Gamma_{FERMO}$: it is the norm of an MO projected on the expected set of atomic orbitals, which are important and that participate in the reaction of a certain compound, as follows in Equation 4 [3, 29-31]:

$$\Gamma_{FERMO} = \sqrt{\sum_i \sum_{j \in G} \sum_k \sum_{\mu} C_l^\mu S_{li} S_{ij}^{-1} S_{jk} C_{k\mu}} \quad \text{(Eq. 4)}$$

**Phenols**

Figure 1 shows the 13 phenols used in this work. The choice was made based on phenols that had experimentally determined protonation sites [8-10, 12].

01-13

Figure 1. Structures of the 13 phenols with protonation sites experimentally determined by NMR chosen for the work.
Results and discussion

Theory versus Experiment

Using the MOLPROJ software, optimized structures of the selected compounds were analyzed using their MOs eigenvectors and overlap matrices, both extracted from output files of single point energy calculations with GAMESS. The degree of localization for each MO and their respective energies were calculated and analyzed. For this, the $P_G$ projection operators were applied to all $2p_x, 2p_y$ orbitals of aromatic carbon atoms as for the hydroxyl oxygen atom. Subsequently, we calculated the coefficients of degree of localization for the HOMO and FERMO molecular orbitals ($\Gamma_{\text{HOMO}}$ and $\Gamma_{\text{FERMO}}$) for comparison. These coefficients are shown in Table I, together with the corresponding MO energies.

Table I: results from the degree of localization of phenols MOs from compounds 01 to 13 as returned by MOLPROJ along with protonation experimental data.

<table>
<thead>
<tr>
<th>Phenol</th>
<th>HOMO</th>
<th>Shape-based FERMO for protonation</th>
<th>Experimental protonation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Atom</td>
<td>$\Gamma_{\text{HOMO}}$</td>
<td>Energy (eV)</td>
</tr>
<tr>
<td>01</td>
<td>C4</td>
<td>0.500961</td>
<td>-159.704280</td>
</tr>
<tr>
<td>02</td>
<td>C4</td>
<td>0.489308</td>
<td>-153.633000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03</td>
<td>C4</td>
<td>0.454808</td>
<td>-159.482160</td>
</tr>
<tr>
<td>04</td>
<td>C4</td>
<td>0.500184</td>
<td>-153.781080</td>
</tr>
<tr>
<td>05</td>
<td>C4</td>
<td>0.480443</td>
<td>-152.892600</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06</td>
<td>C4</td>
<td>0.500974</td>
<td>-151.041600</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07</td>
<td>C4</td>
<td>0.526782</td>
<td>-154.225320</td>
</tr>
<tr>
<td>08</td>
<td>C4</td>
<td>0.490162</td>
<td>-150.301200</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09</td>
<td>C4</td>
<td>0.487159</td>
<td>-148.080000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>C4</td>
<td>0.473252</td>
<td>-147.561720</td>
</tr>
</tbody>
</table>
From these data, two doubts emerged: (1) is the reaction site pointed by the software in accordance with experimental results found in the literature, based on NMR? (2) can the software be applied to a study that has only experimental results?

We know that even simple molecules have favorable sites for protonation and that the preferred site depends significantly on its chemical environment. Much of the experimental information on protonation processes in gas phase came from mass spectrometry studies involving proton transfer reactions \[5,7,11,12,14,23,32,35\]. These studies revealed the existence of several isomers of MH\(^+\) ions, and, in some cases, specific protonation sites. As already mentioned, this protonation can occur in the oxygen atom to form an oxonium ion (O-PhH\(^+\) or PhOH\(^2+\)) or in carbon atoms in ortho (o-PhH\(^+\)), meta (m-PhH\(^+\)) or para (p-PhH\(^+\)) positions in the aromatic ring \[7,17–20\], respectively C2, C3 and C4 (Figure 2).

Using data from literature about protonation of phenols based on NMR techniques, as can be seen in Table I, we performed an analysis of the protonation sites as pointed by NMR technique and by the MOLPROJ software (Table I). Literature data for

<table>
<thead>
<tr>
<th>Phenol</th>
<th>HOMO</th>
<th>Shape-based FERMO for protonation</th>
<th>Experimental protonation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Atom</td>
<td>Energy (eV)</td>
<td>Atom</td>
</tr>
<tr>
<td>11</td>
<td>C4</td>
<td>0.516664</td>
<td>C4</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>0.495021</td>
<td>HOMO-1</td>
</tr>
<tr>
<td>12</td>
<td>C4</td>
<td>0.498601</td>
<td>C4</td>
</tr>
<tr>
<td></td>
<td>Solvent A, C, D; C4: 100%;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>C4</td>
<td>0.496808</td>
<td>C4</td>
</tr>
<tr>
<td></td>
<td>Solvents A, C, D; C4: 100%;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A: HSO\(_3\)-/SbF\(_5\)\[12\]; B: \(H_2\)SO\(_4\)\[10\]; C: HSO\(_3\)-F\[8\]; D: CF\(_3\)SO\(_3\)-H \[9\]

Figure 2. Nomenclature of the protonated phenols adopted in this work. Resonance structures for o-PhH\(^+\), m-PhH\(^+\) and p-PhH\(^+\) are not represented. The numbers C1-C6 refers to the notation used in Table I.
The protonation of phenols through NMR technique was indeed rare and, therefore, there is not a large number of samples.

As it can be seen, in some cases, and as previously discussed, the protonation site changes according to the chemical environment in which the compound is present. In our calculations, we used the water as implicit solvent and at the same temperature. Due to the scarcity of data in the literature for computational reproduction of the experimental solvation environment (i.e., H₂SO₄, HSO₃F or CF₃SO₃H in aqueous solution), we chose to describe the system only in implicit water, considering that, experimentally, water would already be the more abundant species. As there could be more than one protonation site, we compared the experimental reaction sites with the values of the location coefficients $\Gamma$ generated by MOLPROJ for the frontier orbitals related to the atoms corresponding to these sites, checking which ones had the highest $\Gamma$ values (above 0.4) between the possible protonation sites. It was done for all compounds. We start from the premise that atoms with higher location coefficients, as returned by MOLPROJ, contribute better to the protonation sites and to the construction of the involved boundary molecular orbitals.

Pictures of the highest MOs for each compound are shown in Figure 3. The $\Gamma$ coefficients from HOMO to HOMO-3 for O and C1-C6 atoms of all compounds can be find in Supplementary Material.

<table>
<thead>
<tr>
<th>Compounds</th>
<th>HOMO</th>
<th>HOMO-1</th>
<th>HOMO-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="HOMO" /></td>
<td><img src="image2.png" alt="HOMO-1" /></td>
<td><img src="image3.png" alt="HOMO-2" /></td>
</tr>
<tr>
<td>2</td>
<td><img src="image4.png" alt="HOMO" /></td>
<td><img src="image5.png" alt="HOMO-1" /></td>
<td><img src="image6.png" alt="HOMO-2" /></td>
</tr>
</tbody>
</table>
Figure 3. Shapes of the four highest MOs for the protonated phenols studied in this work.

For the simplest compound, phenol (01), protonation in H₂SO₄ medium occurs only in the oxygen atom. An analysis performed with the MOLPROJ software returned a slightly higher $\Gamma$ coefficient value for the HOMO-2 (0.505728), located at the oxygen atom, in relation to the HOMO orbital (0.500961), which is mainly located at the C4 carbon atom (the para-hydroxy carbon; Figure 3). With that, it can be understood that protonation would be more favored in this oxygen atom, which agrees with the experimental results in H₂SO₄ solvent, indicating that, in this case, HOMO-2 is the FERMO of the reaction. Analyzing the figure of the orbitals for compound 01, we can see that HOMO-2 describes the protonation site better than the HOMO-1 orbital.

The $p$-fluoro compound 3 shares the same 100% experimental protonation in the oxygen atom in H₂SO₄. For this compound, the MOLPROJ software returned the HOMO-2 as the FERMO of the reaction (centered at the oxygen atom). As with compound 01, HOMO-2 has a high localization coefficient $\Gamma$ value in the oxygen atom (0.505810 over 0.454808 for the HOMO at C4). Thus, we understand that HOMO-2 would be the FERMO of this reaction as well. In the same way as compound 01, based on the analysis
of the shape of the orbitals in Figure 3, it is possible to verify that HOMO-2 has, in fact, the characteristics of a favorable location for protonation.

For compound 10 (2,4,6-trimethylphenol), which is experimentally fully protonated at C3, our theoretical findings point out protonation at C5 (the main location of HOMO-1/FERMO). But they are chemically equivalent and can react the same way. The difference in the coefficients can only be attributed to the fact that the conformation shows the H of the hydroxyl positioned on one side or the other. In this case, an experimental protonation was conducted on HSO$_3$F / SbF$_5$ instead of H$_2$SO$_4$.

For the compounds 07, 12 and 13, experimentally protonated only at the C4 carbon atom in HSO$_3$F (with/without SbF$_5$) or CF$_3$SO$_2$H, the MOLPROJ software returned sufficiently high (higher than 0.4) Γ coefficient values only for the HOMO, mainly located at the C4 carbon atom (Figure 3). In these compounds, the HOMO is the FERMO of the reaction.

The compound 11, also with a high Γ coefficient value for its HOMO at C4 carbon and for HOMO-1 at C2 (Figure 3), was experimentally protonated at C2 and, in less extension, at C4. For the protonation of 2,4,5-trimethylphenol (09), the experimentally preferred site in the presence of HSO$_3$F was the C4 carbon (85% protonation), accompanied by about 15% protonation in the C6 carbon. Our MOLPROJ software returned that the three largest location coefficients were found at carbon C4, C3 and carbon C6 (0.487159, 0.492505 and 0.539112, respectively). It can be seen that, when the protonation occurs at C4 (para), the reaction orbital is the HOMO. When it occurs at C6 (ortho), the orbital that is describing the reaction is the HOMO-1. Therefore, both orbitals are the FERMOs of the reaction (Figure 3).

In the case of the compound 06 (3,4-dimethylphenol), the experimentally preferred sites in the presence of the same solvent (HSO$_3$F) were the C2 (90% protonation and a Γ coefficient of 0.531917) and C4 carbon atoms (10% protonation and a Γ coefficient of 0.502416).

Finally, for the compounds 05 and 08, it was observed that some protonation sites were different from those experimentally found. In 05, the C4 and C5 carbon atoms would be protonated according to MOLPROJ but, however, it was experimentally observed that protonation occurred at C4 and C6. We attribute this to the fact that the MOLPROJ analysis is based only on electronic effects and the fact that there is a mixture of solvent, which makes computational analysis difficult. Also because the studies used as reference
are relatively old and the parameterization of implicit solvation calculations containing solvent mixtures is extremely difficult. We mean that, based on the analysis of MOs, the software helps to determine which AOs are relevant for the reaction center of a given compound.

An alternative to such approach is to apply electron density localizations at the respective atomic sites as the measure of reactivity, specifically for the protonation reactions considered. According to Liu (2014) [42], the Hirshfeld charge can determine regioselectivity. As the Hirshfeld charge is derived from the Conservation of Information Principle, which requires that atomic identity be maintained as much as possible in molecules. So, this charge can reflect the electronegativity or the electropositivity nature of atoms. Therefore, the regioselectivity can be evaluated by using Hirshfeld charges [42-46].

Considering of comparing this method and the CHELPG charge with the MOLRPOJ, we calculated the two types of charges of the 13 compounds (charges results in the supplementary material, Table SIII).

When the direct analysis of the values of the Hirshfeld charges was performed, that is, direct analysis of the charges on the atoms involved in the protonation reaction, it was observed that the most negative values were present in the hydroxyl oxygen and in the methyl substituents and in the fluorine. Considering only the carbons of the benzene ring and oxygen and that there was at least one hit in the protonation site when compared to the experimental protonation site, 69.23% of hits were obtained (01, 02, 03, 05, 06, 08, 09, 10 and 11 compounds).

When the charge difference analysis was performed [45], with phenol being the standard of Hirshfeld charge ($\Delta_{\text{Hirsh}}$ ) values compared with the insertion of substituents in phenol, 46.15% (01, 06, 08, 10, 11 and 12 compounds) of correct answers were obtained. The analysis of the CHELPG charges was done in the same way as the Hirshfeld charges, in a direct way. Reaching 61.53% hits (01, 02, 03, 06, 09, 10, 11 and 12 compounds) with a correct answer.

As the MOLPROJ it is still a code under development, we will continue to seek in future works to evaluate better the effects of solvents on molecular orbitals. However, even considering that these compounds were computationally studied in an aqueous medium, without the appropriate experimentally applied solvents and compared to experimental results in different phases, we obtained 86% of correctness of the
protonation sites using the developed software. We mean that 11 compounds (01, 02, 03, 04, 06, 07, 09, 10, 11, 12, 13) presented exact results when compared to the experimental results and the Hirshfeld and CHELPG charge methods.

Therefore, from our findings, we got a much more representative answer with the FERMO approach. In fact, from the MOLPRJ software, 86% of correctness of the protonation sites using the developed software was obtained.

**Stereo-electronic effects on protonation and functionalization of substituted phenols**

In a second step, we validated, through a computational study using MOLPRJ, an experimental approach regarding the ortho-cyanation reaction of phenols to obtain aromatic nitriles by means of Lewis acids [38]. The importance of aromatic nitriles comes from the fact that they can be easily converted into a variety of valuable sytones, such as ketones, aldehydes, amines, among others, conferring them an important role in synthetic chemistry [38–41].

According to Zhang, Yang and Zhao [38], authors of the study under consideration, the control of regioselectivity with 3,4-disubstituted phenols is still a significant and challenging work, which remains unsolved. So, they proposed to develop a selective method for C–H ortho-cyanation promoted by Lewis acid in 3-substituted and 3,4-disubstituted phenols in two different ortho positions, as can be seen in Figure 4. After an initial step for optimizing reaction conditions, in which it was established that the reaction would be best conducted under a combination of both AlCl$_3$ and BF$_3 \cdot$ OEt$_2$ as Lewis acid catalysts and CH$_3$SCN as cyan donor, they explored the scope of its reaction through the cyanation of 3-substituted and 3,4-disubstituted phenols to afford a wide range of the corresponding 2-hydroxybenzonitriles in good or excellent yields and with enhanced regioselectivity.
Figure 4. General structure of the two possible ortho-cyan phenols obtained through electrophilic cyanation of 3-substituted or 3,4-disubstituted phenols from Zhang, Yang and Zhao [38].

In order to explain their results, the authors proposed a mechanism in which the intermediate I (Figure 5) is generated from the 3-substituted phenol and BF$_3$·OEt$_2$ (activated by AlCl$_3$). Then, the methyl thiocyanate couples with the intermediate I through the transition state TS$_1$ to form the intermediate II. Another possible transition state in this step would be TS$_2$, but its formation would be unfavored because of steric hindrance due to the meta substituent. The intermediate II, through tautomerization, gives the key intermediate III which, after treating with aqueous NaOH solution (leading to IV) followed by acid treatment, leads to the desired 4-substituted 2-hydroxybenzonitriles. These steps are shown in Figure 5 [38]. To validate the mechanism proposed by the authors, we again use MOLPROJ to find the cyanation site in the model molecule I.
In order to find out what carbon atom would be the site for the cyanation reaction, we started from two different conformations of the first intermediate \( \text{I} \) which could lead the two possible transition state structures \( \text{TS1} \) and \( \text{TS2} \), respectively named as \( \text{A} \) and \( \text{B} \) as it can be seen in Figure 6. For this step, the compound meta-cresol was chosen as model. These two structures were generated through optimization with B3LYP/6-31G(d,p). According to the authors, C2 and C6 would be the ring carbons favorable for cyanation but the reaction in C6 would happen in a major way. Their explanation for the preference of \( \text{TS1} \) formation is related to the steric hindrance at the C3 meta-substituent, therefore unfavoring the formation of \( \text{TS2} \). Based on this assumption, we obtained the values of the coefficients \( \Gamma \) over C2 and C6 for both structures \( \text{A} \) and \( \text{B} \) in Figure 6. The data returned by the MOLPROJ software and the experimental cyanation site are shown in Table II.
Figure 6. Structures of the two optimized conformations of intermediate I chosen for prediction of cyanation sites. **A**: structure of I with methyl thiocyanate close to C6, related to **TS1**. **B**: structure of I with methyl thiocyanate close to C2, related to **TS2**. Only the C2 and C6 carbon atoms are highlighted in both structures. Conformations were obtained through B3LYP/6-31G(d,p) calculations.

Table II: results for the degree of localization of MOs from structures A and B (derived from *meta*-cresol) as returned by MOLPROJ.

<table>
<thead>
<tr>
<th>Structure</th>
<th>HOMO</th>
<th>Shape-based FERMO</th>
<th>Experimental cyanation site</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Átomo</td>
<td>$\Gamma_{HOMO}$</td>
<td>Energy (eV)</td>
</tr>
<tr>
<td>A</td>
<td>C2</td>
<td>0.247323</td>
<td>-169.847760</td>
</tr>
<tr>
<td></td>
<td>C6</td>
<td>0.375934</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>C2</td>
<td>0.196610</td>
<td>-164.813040</td>
</tr>
<tr>
<td></td>
<td>C6</td>
<td>0.406361</td>
<td></td>
</tr>
</tbody>
</table>

As it can be seen in Table II, the site for *ortho*-cyanation reaction indicated by MOLPROJ corresponds to C2, which is does not correspond to the experimental results, pointing to C6 as the experimental site. The MOs analysis identified the HOMO-1 as the FERMO for the *ortho*-cyanation reaction instead of HOMO. The $\Gamma_{HOMO-1}$ coefficients calculated by MOLPROJ showed higher values for C2 than for C6 for both A and B. In Figure 7, the shapes of their corresponding MOs are shown.
From these results, another question arose: would this reaction be coordinated by steric or electronic effects (or even by both)? To answer this question, we decided to replace the 3-methyl group by an isoelectronic and smaller fluorine atom in the new structures A’ and B’ and running only single point calculations with the same base. We did this because 3-fluorophenol was one of the 3-substituted phenols examined in the scope of the selective ortho-cyanation reaction substrate in the article [38]. According to the results returned by MOLPROJ, the main reaction site, when a meta-fluorine atom is present, remains at the C2 atom in the aromatic ring, as can be seen below in Table III.

Table III: results for the degree of localization of MOs from fluorinated structures A’ and B’ (derived from 3-fluorophenol) as returned by MOLPROJ.

<table>
<thead>
<tr>
<th>Structure</th>
<th>HOMO</th>
<th>Shape-based FERMO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>átomo</td>
<td>Energy (eV)</td>
</tr>
<tr>
<td>A’</td>
<td>C2</td>
<td>0.128708</td>
</tr>
<tr>
<td></td>
<td>C6</td>
<td>0.459591</td>
</tr>
<tr>
<td>B’</td>
<td>C2</td>
<td>0.129347</td>
</tr>
<tr>
<td></td>
<td>C6</td>
<td>0.461176</td>
</tr>
</tbody>
</table>

Looking at the results, we can see that the reaction path depends on two factors: the size of the substituent and its position. This can be explained by the fact that, in the presence of substituents with heavy atoms, the steric effect will drive the reaction path. Knowing that the volume of an atom or any substituting group is practically expressed by
its electronic density, thus, the steric effect can be interpreted as an electronic repulsion. This electronic repulsion then makes it difficult to approach the substituent, making the electrophilic substitution reaction unfeasible, as occurs in the studied reaction. Now thinking about lower size groups, as in the case of fluorine, we can think about non-classical models. In other words, we are analyzing the electronic density of FERMO to correctly predict which carbon atoms would be the reaction site of this reaction. Therefore, the attractive interactions between the frontier orbitals will be responsible for the reaction site in the cyanation reaction in cases where the substituent is not bulky.

**Conclusion**

The MOLPROJ software, based on the use of projection operators for quantifying the FERMO localization, was employed again successfully now for phenol protonation. It is possible to obtain different values for the localization degrees of MOs according to any combination of AOs. It returns as results not only the localization degree of the MOs in the subspace of one or more AOs as well as the energy of each of them. The code allows identifying the FERMOs based on the joint analysis of the localization coefficients and orbital energies.

The FERMO concept can be applied to find the protonation site for phenols. It was possible, even with different environments in which the experimental compounds were inserted, to have an idea of the preferred protonation site of these molecules. We obtained approximately 86% of correctness of the protonation sites using our theoretical methodology. It was also possible to rationalize a reaction with only experimental results that the reaction path depended on two factors: the size of the substituent and its position. In the presence of bulky substituents, the steric effect will direct the path of the reaction. In lower size groups, the attractive interactions between the frontier orbitals will be responsible for the reaction site in the cyanation reaction.

**Acknowledgements**

We thank the Brazilian agencies FAPEMIG, CNPq, and CAPES for the financial support. This work was also supported by University of Hradec Kralove (Faculty of Science, VT2019-2021).

**References**


Toledo EJL, Ramalho TC. Controversies about hydrogen bonds in water molecules on the influence of high magnetic fields: implications on structural and electronic parameters. Mol. Sim. 2021, 47, 1159-1167.