

## Differential reactivity of thiophene-2-carboxylic and thiophene-3-carboxylic acids Results from DFT and Hartree–Fock theory

Salvatore Profeta Jr.<sup>a,\*</sup>, V.S. Senthil Kumar<sup>a</sup>, Richard Austin<sup>b</sup>, S. Stanley Young<sup>c</sup>

<sup>a</sup>Department of Basic Pharmaceutical Sciences, South Carolina College of Pharmacy, University of South Carolina, Columbia, SC 29208, United States

<sup>b</sup>Sanofi-Aventis Pharmaceuticals, Tucson, AZ 85737, United States

<sup>c</sup>National Institute of Statistical Sciences, Research Triangle Park, NC 27709, United States

### ARTICLE INFO

#### Article history:

Received 4 April 2009

Received in revised form 25 November 2009

Accepted 30 November 2009

Available online 4 December 2009

Dedicated to Professor N.L. Allinger in celebration of his 60+ years in Organic Chemistry!.

#### Keywords:

Thiophene carboxylic acid

Reactivity

Nucleophile

Electrophile

Hartree–Fock theory

Density functional theory-conformational influence

### ABSTRACT

In the course of investigating the propensity of aromatic acids to react with selected nucleophiles, we came across an interesting difference in yields for two structurally similar thiophene carboxylic acids. Given that these yields were consistent across more than 40 repetitions for each structure, we felt that the difference was real, and worth exploration. To extract the potential steric and electronic origins of such differences, we employed both DFT B3LYP/6-31G\* and HF/6-311+G\*\* levels of theory to evaluate structures and energetics. Two somewhat different pictures emerge of the origin of the differences between the carboxylic acids and their individual conformations. In particular, Hartree–Fock calculations with the larger 6-311+G\*\* basis set, which includes multiple diffuse functions, show a profound difference in the delocalization of the LUMO, relative to the DFT method which, expectedly, provides a more localized picture of the contributions to the LUMO. For each of the conformers, the molecular electrostatic potential and the ionization potential (ESP and IP, respectively), together with charge distributions, dipole moments and orbital energies have been explored. A potential explanation for the somewhat more reactive character of the thiophene-2-carboxylic acid appears to arise from the presence of a conformer which has an internal hydrogen bond to the thiophene sulfur, which, in turn, polarizes the acid function significantly relative to other conformers, and optimizes the angle of attack of the nucleophile.

© 2009 Elsevier Inc. All rights reserved.

## 1. Introduction

Over the past 15 years, the role of combinatorial chemistry has grown substantially in the pharmaceutical industrial sector as a means to access both chemically diverse molecules and potential pharmaceutical leads [1]. One area of research in computational chemistry that has recently begun to grow in response to the needs of combinatorial chemistry is reaction/yield prediction and analysis [2]. This effort arises from difficulties encountered optimizing the number of reaction variables that are involved in the development of combinatorial methods that can be scaled effectively to provide multi-gram quantities of desired compounds. The amount of work required to establish the optimal conditions for a combinatorial “campaign” can be substantial, and,

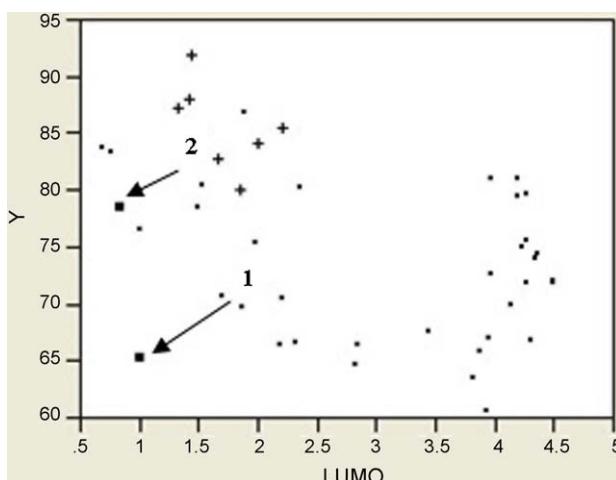
at times, is the most time-consuming step in providing compounds to drug discovery teams.

Whereas, in the past, the computational analysis of sets of compounds might have been the rate-limiting step in a discovery program, this is no longer necessarily the case. With faster computers, more highly optimized software tools, and the use of computer farms, molecular modelers can now generate hundreds of chemical descriptors for structure datasets, and run them through multiple statistical modeling tools in a fraction of the time required 10 years ago. This can be accomplished for datasets ranging from hundreds to millions of structures in a matter of days to weeks. This timeline is often faster than the timeline required to optimize a combinatorial synthesis protocol [3]. Thus, it is now possible to use computational tools to characterize the reactivity of the pairs of components involved in a combinatorial library, and to prioritize which of these candidates will give the best yields. This has the effect of increasing the efficiency of the preliminary library work, and reducing cycle time.

While we can now readily examine large sets of molecules, it is still frequently the case that a smaller subset of structures pose

\* Corresponding author. Present address: Department of Chemistry & Biochemistry, Florida State University, 95 Chieftan Way, 614 Dittmer Hall, Tallahassee, FL 32306, United States. Tel.: +1 850 644 8278; fax: +1 850 644 8281.

E-mail address: [profeta@chem.fsu.edu](mailto:profeta@chem.fsu.edu) (S. Profeta Jr.).

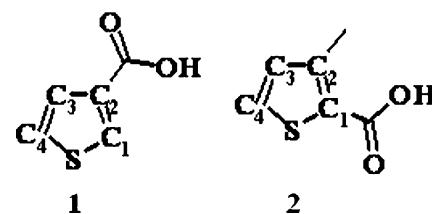


**Fig. 1.** A distribution of product yields for a variety of organic acids as a function of the CNDO/1 computed LUMO (lowest unoccupied molecular orbital) of the carboxylic acid reactants. Note the bolded squares which represent molecules **1** and **2**.

interesting scientific questions, and the study herein is a case example.

During a synthesis methods optimization, two structurally similar thiophene carboxylic acids were found to have slightly, but statistically significant, different yield trends toward nucleophilic substitution in a condensation reaction, 65% vs 79%. At first, it was assumed that this might be an experimental fluctuation; however, these values are the results of 40 or more individual experiments each. Under such circumstances, one must assume that either the experimental protocol is biasing the results (possible), or, that a small, but real intrinsic difference in reactivity between the two molecules exists. We assumed the latter and executed a brief examination [4] of the electronic structure of the two carboxylic acids and their conformations.

In Fig. 1, we show a plot of the correlation between the LUMO energy (computed using the CNDO/2 method) and the yields of a series of carboxylic acids used in developing a library of condensation products with nucleophilic amines. The plot itself is part of a comprehensive analysis of the potential relationship between computed physical chemical descriptors and observed yields; we include it here to provide context, and to show that the



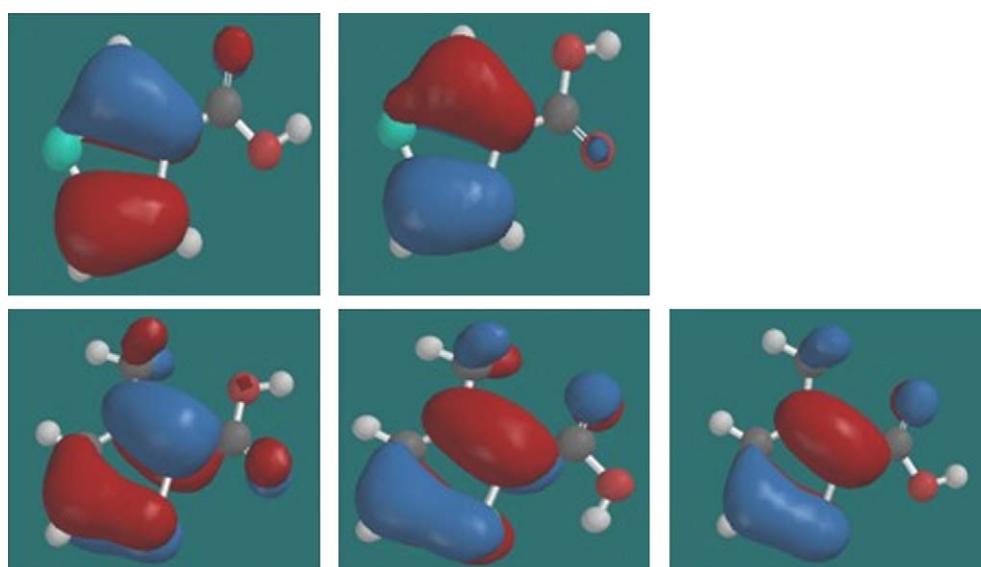
**Fig. 2.** The structures and numbering of thiophene-3-carboxylic acid (**1**) and 3-methylthiophene-2-carboxylic acid (**2**).

two thiophene carboxylic acids in question have a measurable (and reproducible) difference in their yields of condensation products. This difference prompted us to investigate the nature of the acids in more detail.

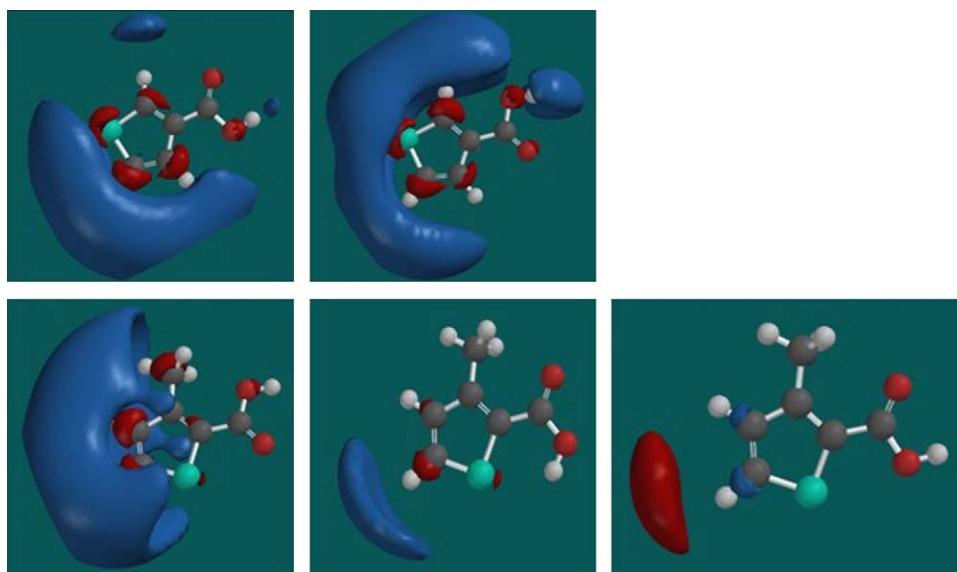
## 2. Methods

The geometry of the thiophene carboxylic acids **1** and **2** (Fig. 2) were derived from the *ab initio* Hartree–Fock (HF) theory using the SPARTAN'04[5] program (Mac v. 1.0.1) on an Apple Macintosh G5 running at 1.8 GHz with 2 GB of memory. We used the 6-311+G\*\* basis set [6] which, while originally designed for use with correlated wavefunctions, can be effective in the analysis of third row atoms, such as sulfur, even in the divalent state. The 311 “triple” split has been developed to give more overall flexibility in the representation, and improves the description of the outer valence region, which can be useful in sulfur-containing systems. The d-type Gaussians added use 5 instead of a full set of 6 functions, with the loss of only an s-type Gaussian function. To this set, a single set of diffuse type s- and p-type functions that provide additional coverage of the long-range interactions, and a more complete picture of weakly bonded atomic arrangements, such as sulfur–hydrogen hydrogen bonding [6]. DFT calculations were executed using the B3LYP corrected method together with the 6-31G\* basis set, as described in Ref. [7], and implemented in SPARTAN'04 [5] (Figs. 7 and 8).

Different conformers of **1** (**1a–b**) and **2** (**2a–c**) (see Figs. 5 and 6 for the structures) were generated by either a single C–C or C–O bond rotation, or, by simultaneous rotations around the central C–C as well as C–O bonds. Geometry optimizations of these rotamers were performed, and the ESP, IP and the molecular orbitals (HOMO and LUMO) for these conformers were generated and examined



**Fig. 3.** Top row: HOMOs for structures **1a** and **1b**. Bottom row: HOMOs for structures **2a–c**. Computed at HF/6-311+G\*\*//6-311+G\*\*.



**Fig. 4.** Top: The LUMO for **1a** and **1b**. Bottom: The LUMO for **2a–c**. Computed at HF/6-311+G\*\*//6-311+G\*\*. Note the relative lack of participation of the acid functions in both sets.

graphically. All charges noted in Table 1 and in the text are ESP-derived values, as available in SPARTAN [5,8]. The ESP,  $V(\mathbf{r})$ , at a point  $\mathbf{r}$  due to a molecular system with nuclear charges  $\{Z_A\}$  located at  $\{\mathbf{R}_A\}$  and electron density  $\rho(\mathbf{r})$  is given by

$$V(\mathbf{r}) = \sum_{A=1}^N \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} - \int \frac{\rho(\mathbf{r}') d^3 \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}$$

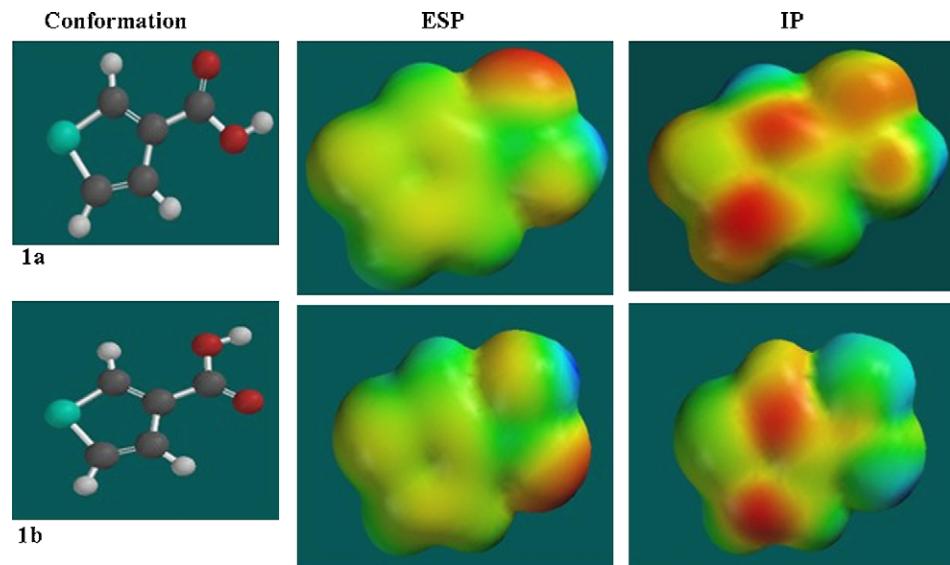
The graphical representation of the molecular electrostatic potential surface (MEP or ESP), as described by Kollman and

Singh [8] is a series of values representing the evaluation of the interaction energy between a positively charged (proton) probe and points on a solvent accessible surface as defined by Connolly [9]. As implemented within the SPARTAN program, areas of high electron density, representing a strong attraction between the proton and the points on the molecular surface, have the brightest red color and areas of lowest electron density, have deep blue to indigo color, indicating the regions of maximum repulsion.

**Table 1**

Total energies, relative conformational energies, solvation energies, dipole moments and partial charges for the conformers of **1** and **2** at HF/6-311+G\*\*//6-311+G\*\* and B3LYP/6-31G\*//6-31G\*.

| Conformer                  | Energy (Hartrees) |             | Rel. E (kcal/mol) |        | Dipole (D) |        | E LUMO (eV)         |          |
|----------------------------|-------------------|-------------|-------------------|--------|------------|--------|---------------------|----------|
|                            | HF                | B3LYP       | HF                | B3LYP  | HF         | B3LYP  | HF                  | B3LYP    |
| Acid <b>1a</b>             | -738.91243        | -741.576188 | 0.34              | 0.29   | 2.024      | 1.842  | 1.61932             | -1.20231 |
| Acid <b>1b</b>             | -738.91286        | -741.576651 | 0.00              | 0.00   | 1.614      | 1.447  | 1.63558             | -1.22427 |
| Acid <b>2a</b>             | -777.94764        | -780.894085 | 0.89              | 0.91   | 3.364      | 2.807  | 1.62580             | -1.45081 |
| Acid <b>2b</b>             | -777.93550        | -780.883658 | 8.02              | 7.46   | 4.931      | 4.477  | 1.43687             | -1.60629 |
| Acid <b>2c</b>             | -777.94908        | -780.895539 | 0.00              | 0.00   | 1.813      | 1.623  | 1.61831             | -1.43515 |
| Conformer                  | E HOMO (eV)       |             |                   |        | GAP        |        |                     |          |
|                            | HF                |             | B3LYP             |        | HF         |        | B3LYP               |          |
| Acid <b>1a</b>             | -9.5591148        |             | -6.79042544       |        | 11.178     |        | 5.5881              |          |
| Acid <b>1b</b>             | -9.5250402        |             | -6.81905033       |        | 11.161     |        | 5.5948              |          |
| Acid <b>2a</b>             | -9.3611228        |             | -6.63537338       |        | 10.987     |        | 5.1846              |          |
| Acid <b>2b</b>             | -9.6078129        |             | -6.84576615       |        | 11.045     |        | 5.2395              |          |
| Acid <b>2c</b>             | -9.3341143        |             | -6.62204003       |        | 10.952     |        | 5.1869              |          |
| ESP-derived atomic charges |                   |             |                   |        |            |        |                     |          |
| Conformer                  | C1                |             | C2                |        | C3         |        | C4                  |          |
|                            | HF                | B3LYP       | HF                | B3LYP  | HF         | B3LYP  | HF                  | B3LYP    |
| Acid <b>1a</b>             | -0.239            | -0.229      | 0.046             | 0.062  | -0.239     | -0.209 | -0.192              | -0.179   |
| Acid <b>1b</b>             | -0.220            | -0.216      | 0.073             | 0.078  | -0.283     | -0.242 | -0.152              | -0.143   |
| Acid <b>2a</b>             | -0.284            | -0.194      | 0.398             | 0.321  | -0.372     | -0.321 | -0.157              | -0.146   |
| Acid <b>2b</b>             | -0.542            | -0.362      | 0.573             | 0.432  | -0.440     | -0.350 | -0.023              | -0.060   |
| Acid <b>2c</b>             | -0.326            | -0.238      | 0.418             | 0.332  | -0.368     | -0.302 | -0.154              | -0.170   |
| Conformer                  | O(C=O)            |             | O(OH)             |        | H(OH)      |        | C(CH <sub>3</sub> ) |          |
|                            | HF                | B3LYP       | HF                | B3LYP  | HF         | B3LYP  | HF                  | B3LYP    |
| Acid <b>1a</b>             | -0.631            | -0.515      | -0.657            | -0.563 | 0.459      | 0.419  |                     |          |
| Acid <b>1b</b>             | -0.610            | -0.502      | -0.718            | -0.594 | 0.480      | 0.430  |                     |          |
| Acid <b>2a</b>             | -0.608            | -0.502      | -0.661            | -0.561 | 0.442      | 0.406  | -0.532              | -0.462   |
| Acid <b>2b</b>             | -0.600            | -0.475      | -0.702            | -0.586 | 0.468      | 0.423  | -0.609              | -0.570   |
| Acid <b>2c</b>             | -0.634            | -0.516      | -0.629            | -0.536 | 0.433      | 0.398  | -0.555              | -0.482   |

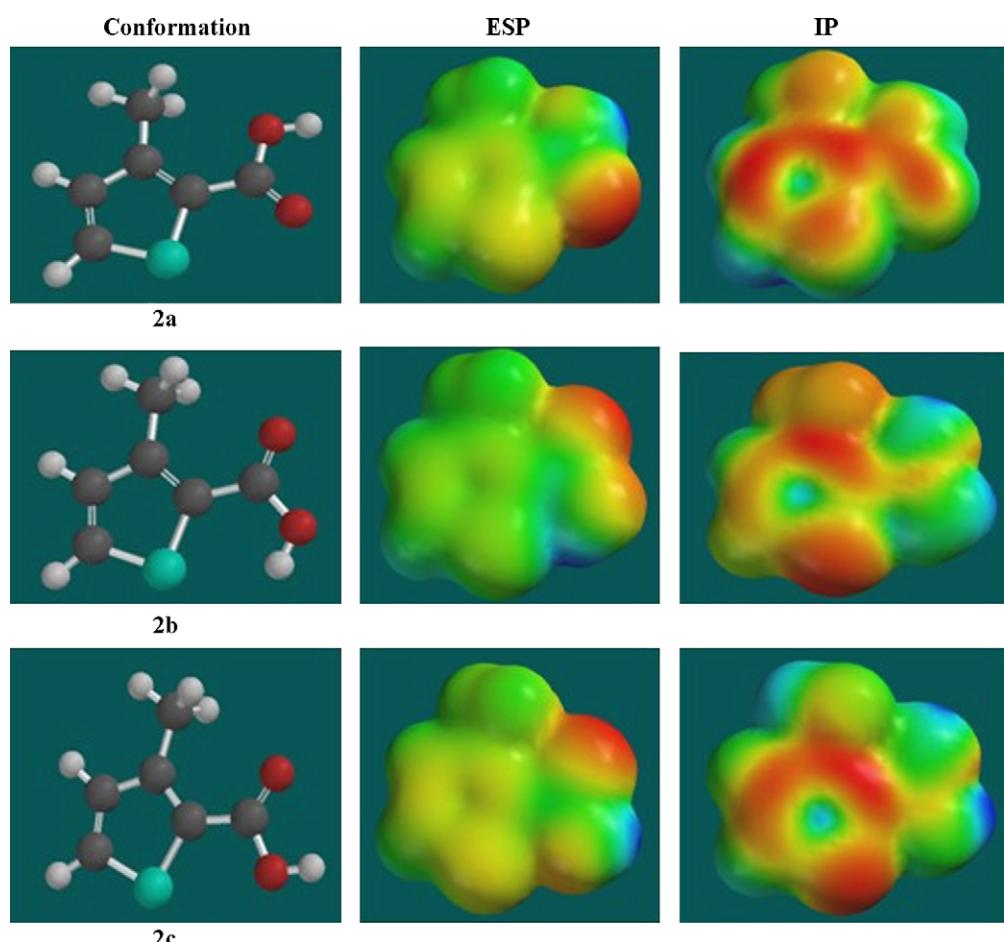


**Fig. 5.** The ESP and IP surfaces for **1a–b**, computed at HF/6-311+G\*\*//6-311+G\*\*.

Another quantity of utility in the analysis of intrinsic reactivity is the local ionization potential,  $I(\mathbf{r})$ , IP. This is defined as the sum over orbital electron densities,  $\rho_i(\mathbf{r})$  times the absolute orbital energies,  $|\varepsilon_i|$ , and divided by the total electron density,  $\rho(\mathbf{r})$ .

$$I(\mathbf{r}) = \sum_i^{\text{occupied molecular orbitals}} \rho_i(\mathbf{r}) |\varepsilon_i| / \rho(\mathbf{r})$$

The local IP is intended to reflect the relative ease of electron removal at any location around a molecule [10]. Thus, as we use it here, it is a measure of the susceptibility of a molecular region to electrophilic attack (reactivity). The surface mapped values of the IP, as derived from SPARTAN'04 and shown in Figs. 5–8, used the default values of the program



**Fig. 6.** The ESP and IP surfaces for **2a–c**, computed at HF/6-311+G\*\*//6-311+G\*\*.

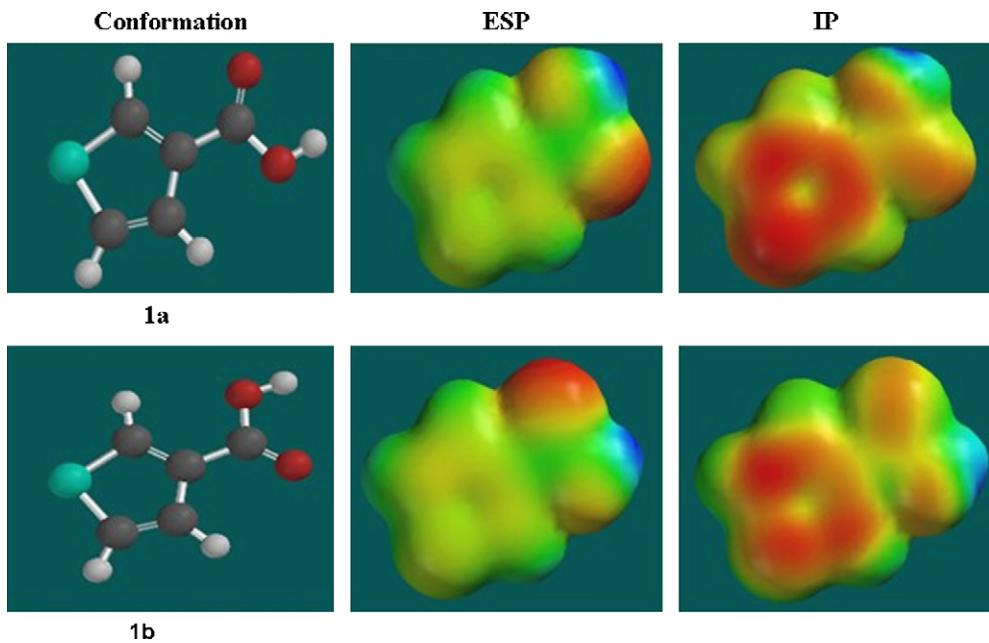


Fig. 7. The ESP and IP surfaces for **1a–b**, computed at B3LYP/6-31G\*//6-31G\*.

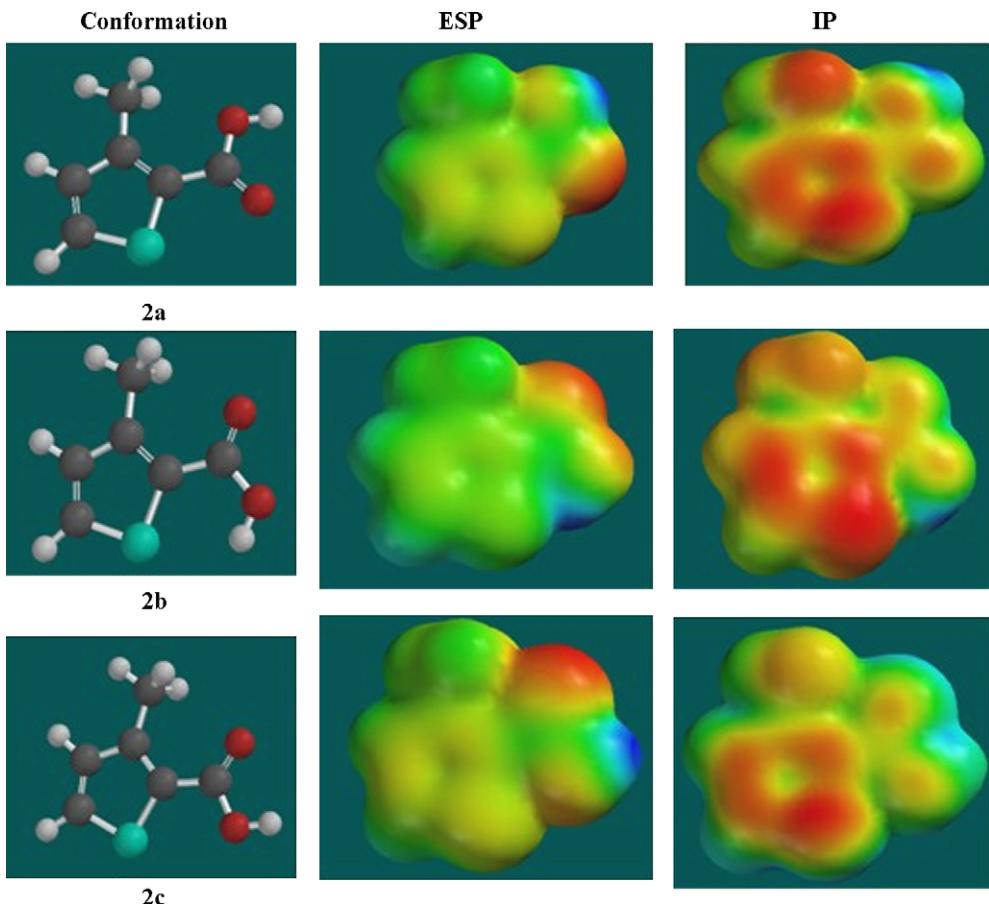


Fig. 8. The ESP and IP surfaces for **2a–c**, computed at B3LYP/6-31G\*//6-31G\*.

throughout (Fixed Value: (0.032e); Medium Resolution). While evaluation of the IP may be of value in quantifying reactivity, our purpose is primarily to convey a qualitative graphical representation of the differences in potential reactivity between constitutional isomeric systems and conformations within those

systems [11]. The graphical convention used (Figs. 5–8) indicates a range of relative values, with red representing regions of highest likelihood of electrophilic attack, and blue, at the other extreme, areas of lowest likelihood (most nucleophilic regions).

### 3. Discussion

As a part of ongoing studies in predicting reaction yields for combinatorial reactions, quantum mechanical calculations were carried out to study the molecular geometry and electronic structure and reactivity of the different conformations of two thiophene carboxylic acids (**1(a–b)** and **2(a–c)**: see Figs. 5 and 6). There are two conformations for thiophene-3-carboxylic acid (**1a** and **1b**) and three conformations possible for 3-methyl thiophene-2-carboxylic acid. They differ in the rotation about the C–C bond between thiophene ring and the carboxylic acid group. We did not consider non-hydrogen bonded conformations of **1**, as there is no equivalent conformation to **2b**. The charge density, electrostatic potential, ionization potential and the spatial distribution of HOMO and LUMO potentially provide insights regarding the reactivity towards nucleophilic attack.

#### 3.1. HOMO/LUMO behavior at the HF/6-311+G\*\* level

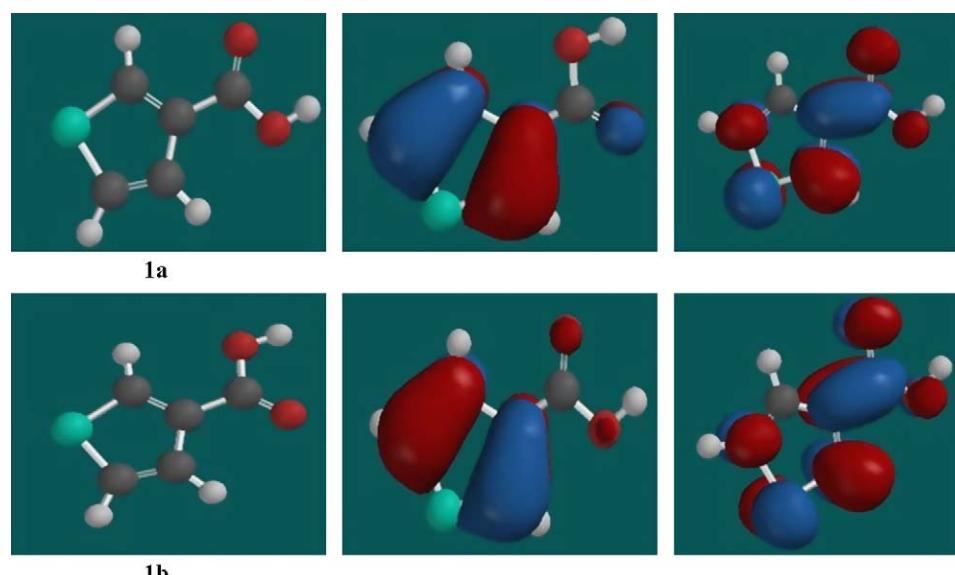
Historically, HOMO/LUMO and gap values as calculated by simple Hückel Molecular Orbital theory (SHMO) have provided insights into the interactions of molecular orbitals and an understanding the reactivity of unsaturated and aromatic hydrocarbons [12a,12c], nucleic acids and carbonyl containing compounds [12c]. More recent exploration of the role of gap values can be seen in the study of pigmentary macromolecules such as melanins [13]. In a representative study [13], the authors have used the Perdew–Burke–Ernzerhof (PBE) exchange correlation functional method to evaluate the HOMO–LUMO gap and correlation to observed optical properties in a series of melanin precursor quinones. Their DFT results are consistent with the threshold for optical absorption found by semiempirical methods. The authors [13] conclude: “The large difference in the HOMO–LUMO gap between HQ (hydroquinone) and the other molecules may be important for “molecular engineering” applications and for explaining the broadband optical absorption of eumelanin. In particular, if the model of the eumelanins as amorphous semiconductors is correct it may allow the chemical tuning of the semiconducting bandgap.” Thus, it appears that HOMO/LUMO and gap values have again come to the forefront in the chemist’s toolkit to assist in both understanding electronic transition, relative reactivity, and in designing new and relevant molecules.

The HOMO/LUMO and gap values for the two conformations of **1** and the three conformations of **2** are given in Table 1. The structure with the lowest absolute energy of both the HOMO and LUMO is **2b**. In general, the MOs for structure **2** are lower than in **1** by ~0.165 eV. While the gap for the conformations of **1** is small (0.018 eV), the gap for the conformations of **2**, are somewhat larger ( $|2a-2b| = 0.058$  eV;  $|2a-2c| = 0.093$  eV).

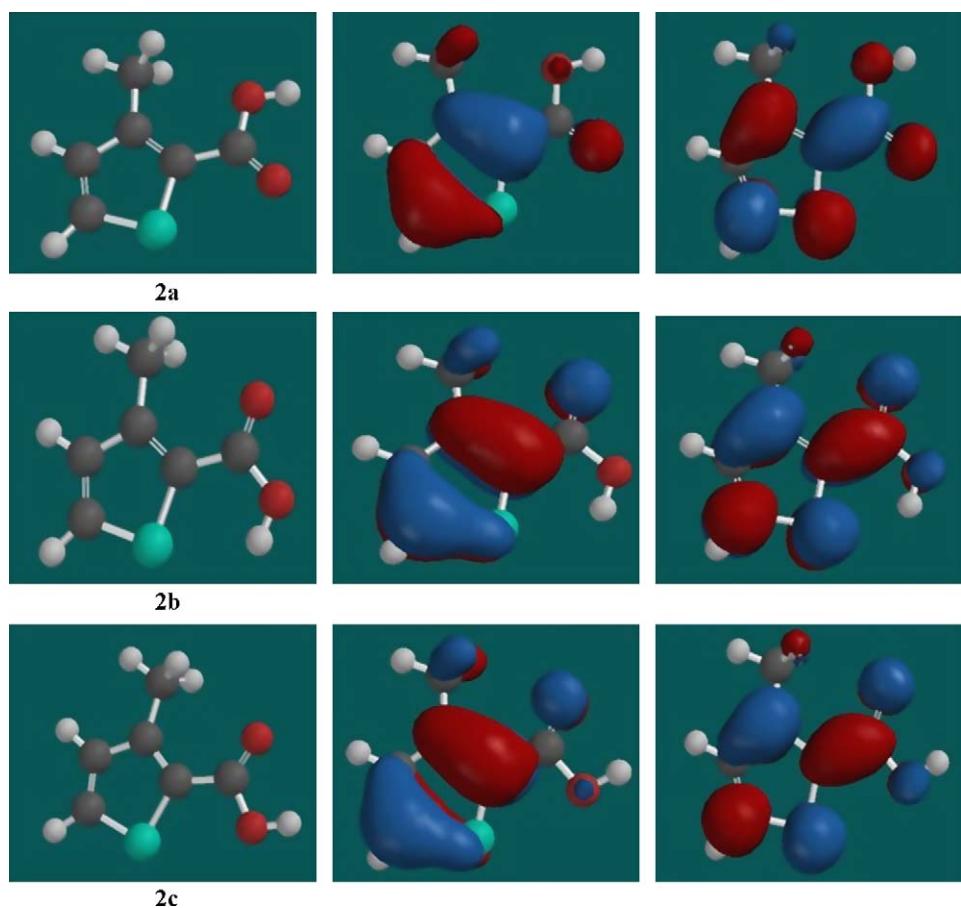
In the case of **1**, the thiophene-3-carboxylic acid conformations, neither the sulfur atom nor the carboxylic acid functionality contribute significantly to the HOMO. This is somewhat perplexing given that the sulfur does donate charge into the ring, and increases the charge at C1 in the ring (see Figs. 3 and 9, Table 1). In **2**, the sulfur atom and the carboxylic acid group, appear to contribute significantly to the HOMO, regardless of conformation. Conversely, the carboxylic acid group and the sulfur atom in **1a** and **1b** contribute to the LUMO while this is not observed in the conformations in **2a–c** (Fig. 4). Thus, an overall interpretation of the data suggests that conformations **2a–c** with larger HOMO–LUMO gaps are more likely to involve reactions with a nucleophile. In particular, conformation **2b**, with a LUMO ~0.19 eV lower than any other structure, potentially provides a more attractive target to an incoming electron pair (Fig. 10).

Representative plots of the local ionization potentials show interesting trends in the regions more susceptible to electrophilic and nucleophilic attack. Examination of the structures for **1**, Fig. 6, shows that the acid group conformation **1a** is more electron rich than in **1b**. From this data, we conclude that **1b** is the conformation that undergoes the reaction, and happens to be the lower energy conformation. This is further substantiated by viewing the probable approach of the nucleophile along a vector coincident with the C=O bond, but displaced above and at an angle of about 100° (approximately perpendicular) to the plane of the entire molecule. See Fig. 11. Examination of the IP surface suggests that the ideal charge distribution for attack by a nucleophile on the surface of **1a** or **1b** is best at **1b** immediately in the vicinity of the carbonyl carbon, which is shown as greenish-blue on the surface. Conformation **1a**, in fact, appears more susceptible to electrophilic attack at the acid function from its IP distribution (reddish-orange).

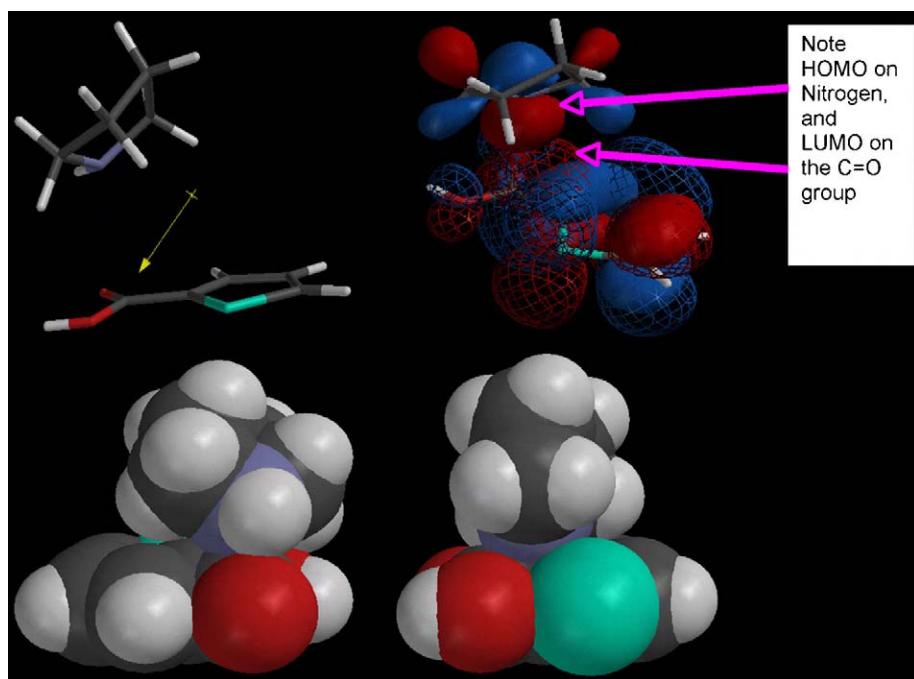
The situation in **2** is a somewhat similar. Conformation **2c** is the lowest in energy. However, conformation **2a** is not much higher in energy, being 0.89 kcal/mol higher at HF/6-311+G\*\*, and 0.91 kcal/mol higher at B3LYP/6-31G\*. However, **2a** does not appear to have



**Fig. 9.** From left to right, the structures, HOMO and LUMO representations of **1** from B3LYP/6-31G\*/6-31G\* calculations.



**Fig. 10.** From left to right, the structures, HOMO and LUMO representations of **2** from B3LYP/6-31G\*//6-31G\* calculations.



**Fig. 11.** Pyrrolidine-thiophene-2-carboxylic acid van der Waals complex at HF/6-31G\*//6-31G\*. Stick figure (top left) showing the structure and dipole moment, HOMO (solid surface)-LUMO (mesh) interactions (top right), and CPK representations of a model of nucleophilic attack by pyrrolidine on the acid group in thiophene-2-carboxylic acid.

a favorable charge distribution from the IP analysis at the acid group, with that surface being dominated by a red-orange color, indicating a good region for electrophilic, rather than nucleophilic attack. Conversely, both **2c** and **2b** appear more susceptible to nucleophilic attack at the acid function, given the green-blue color distribution, with **2b** looking, perhaps, the more reactive of the two to such attack. If a nucleophile were to approach **2b** or **2c**, it would have a clear “path of approach” which would be polarized in favor of a nucleophile approaching the acid function more readily, with a minimum of charge repulsion presented by the sulfur or other ring atoms. From this analysis we conclude that **2c** is the primary source of reactivity for 2 from a statistical perspective, and despite the higher relative conformational energy of **2b**, it is a probably also a contributor to the differential in reactivity seen in the synthesis efforts.

### 3.2. Pyrrolidine-thiophene carboxylic acid model calculations

To assess the conclusions elaborated above, we chose a simple model of the attack of pyrrolidine on thiophene-2-carboxylic acid as a way to verify our hypothesis. Starting with approximately (i.e., manually) aligned structures for the two components, we performed HF/3-21G\* followed by HF/6-31G\* geometry optimizations within SPARTAN. The resulting geometries of the van der Waals complexes were virtually the same as those derived manually. The dipole moments calculated were 1.92D and 1.80D, at HF/3-21G\* and HF/6-31G\*, respectively. The HF/6-31G\* complex with the dipole aligned is shown in Fig. 11. Also shown in Fig. 11 are CPK renderings showing the tight fit of the pyrrolidine to the region of positive charge at the carbonyl carbon. The figure also shows the alignment of the HOMO on nitrogen and its impingement on the LUMO of the carbonyl carbon. The “angle of attack” appears to be directly as proposed herein and first suggested by Burgi and Dunitz [14].

The optimized distance between the atomic center at nitrogen and the carbonyl carbon is 2.897 Å, well within the sum of the van der Waals radii, ~3.20 Å. This distance does not explicitly account for the lone pair on nitrogen, which is estimated at about 0.5–0.6 Å beyond the atomic center of nitrogen [15]. The parallel alignment of the C=O and N-H bonds makes sense in terms of both the dipole alignment and the orientation of the attacking species. Given that the thiophene group itself has a significant dipole (0.88D at HF/6-31G\*/6-31G\*) which is pointed directly at the sulfur in the plane from the bisection of the C2-C3 bond, the net moment of the complex appears dominated by the N-lone pair  $\Rightarrow$  C=O interaction.

## 4. Summary

While the relative yields of the two thiophene carboxylic acids may be only different by about 15%, it is now clear that there is an intrinsic reactivity difference between 1 and 2, such that this difference has chemical and statistical meaning. The relative energies and conformational distributions contribute to the difference as well, although for 2, it appears that a high energy/low population conformer may have an unanticipated large influence on the overall reactivity due to a built-in enhancement toward the particular reaction type. These factors suggest that when pursuing reaction optimization studies, that a more careful evaluation of intrinsic conformational and electronic properties may be appropriate as part of the design of experiments to improve yields. Not all systems will require such scrutiny; however, in cases such as this, such scrutiny provides an insight into the fundamental reactivity of the species involved, and that variation of reaction conditions, such as solvent polarity (which can have a dramatic

influence on conformational energetics), may have a significant influence on the ultimate yields achieved.

## Acknowledgement

We wish to thank the University of South Carolina College of Pharmacy for computational support and Aventis Pharmaceuticals for financial support to SP.

## References

- [1] (a) D.A. Horton, G.T. Bourne, M.L. Smythe, The combinatorial synthesis of bicyclic privileged structures or privileged substructures, *Chem. Rev.* 103 (2003) 893–930;  
 (b) E.M. Gordon, J.F. Kerwin, Jr. (Eds.), *Combinatorial Chemistry and Molecular Diversity in Drug Discovery*, Wiley-Liss, 1998;  
 (c) W. Bannwarth, B. Hinzen (Eds.), *Combinatorial Chemistry: From Theory to Application*, Wiley-VCH, 2006.
- [2] S. Profeta Jr., S.S. Young, E.M. Collantes, P.D. Mosier, QSAR and QSRA (quantitative structure-reactivity analysis): partners in combinatorial library design and optimization, in: 215th ACS Meeting, Dallas, TX, Paper #15, AGFD, Spring 1998;  
 J. Gasteiger, M. Pförtner, M. Sitzmann, R. Höllering, O. Sacher, T. Kostka, N. Karg, Computer-assisted synthesis and reaction planning in combinatorial chemistry, *Perspect. Drug Design Discov.* 20 (2000) 245–264.
- [3] C.J. Manly, S. Louise-May, J.D. Hammer, The impact of informatics and computational chemistry on synthesis and screening, *Drug Discov. Today* 6 (2001) 1101–1110.
- [4] A reviewer made two good points that we wish to address: (1) Why not perform reaction pathway trajectory calculations to assess the reaction coordinates? and, (2) Would not the water generated by the condensation reaction preclude hydrogen bonding between the ring sulfur (as we suggest) and the hydroxyl hydrogen? It will assist the reader to have additional context to understand our approach: regarding question 1, Pathway calculations would be the ideal scenario, should we have had the time to do so. This was a very brief part of a much larger effort to develop physical chemical descriptors of large sets (thousands) of acids and amines. In addition to reaction pathway analysis, we would have needed to include dynamics to capture conformational behavior. Pathway analysis does not assure that the conformational dynamics are sampled. Regarding question 2, the reactions in question were run on 384 well plates. Product yields and residual reactants were quantified by NMR and GC/MS. The scale of these reactions is such that while the water may be equimolar to the reactants, it's likely insufficient to cause a significant interruption to the conformational dynamics. Recall that, as a rule of thumb, the rate of hydrogen bonding for intramolecular interactions is about 10<sup>3</sup> greater than the rate for intermolecular H-bonding.
- [5] SPARTAN'04 for Windows and Mac, Wavefunction, Inc., Irvine, CA 92612. See: W.J. Hehre, *A Guide to Molecular Mechanics and Quantum Chemical Calculations*, Wavefunction, Irvine, 2003.
- [6] W.J. Hehre, L. Radom, P.v.R. Schleyer, J.A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley-Interscience, New York, 1986, pp. 63–100.
- [7] W. Kohn, A.D. Becke, R.G. Parr, Density functional theory of electronic structure, *J. Phys. Chem.* 100 (1996) 12974–12980.
- [8] U.C. Singh, P.A. Kollman, An approach to computing electrostatic charges for molecules, *J. Comput. Chem.* 5 (1984) 129–145.
- [9] M.L. Connolly, Solvent-accessible surfaces of proteins and nucleic acids, *Science* 221 (1983) 709–713.
- [10] (a) P. Sjöberg, J.S. Murray, T. Brinck, P. Politzer, *Can. J. Chem.* 68 (1990) 1440;  
 (a) For a review, see: J.S. Murray, P. Politzer, in: C. Párkányi (Ed.), *Theoretical Organic Chemistry. Theoretical and Computational Chemistry*, vol. 5, Elsevier, Amsterdam, 1998, p. 189;  
 (c) P. Politzer, J.S. Murray, M.C. Concha, *Int. J. Quant. Chem.* 88 (2002) 19.
- [11] Hoffmann and Stowasser have examined the relationship between IP values derived from HF and DFT methods. For the Kohn-Sham eigenvalues they find an approximately linear dependency of  $|e_i^{\text{KS}} - e_i^{\text{HF}}|$  vs  $e_i^{\text{HF}}$  ( $\approx$ IP) for the occupied as well as for the unoccupied orbital eigenvalues. They do suggest that a linear scaling is probably justified, however, as we noted, we are not so interested in the quantitation of the IP values as the conformational trends. See: R. Stowasser, R. Hoffmann, What do the Kohn-Sham orbitals and eigenvalues mean? *J. Am. Chem. Soc.* 121 (1999) 3414–3420.
- [12] (a) A. Rauk, *Orbital Interaction Theory of Organic Chemistry*, 2nd ed., Wiley-Interscience, New York, 2001, pp. 86–136;  
 (b) B. Pullman, A. Pullman, Electronic aspects of biochemistry, *Rev. Mod. Phys.* 32 (1960) 428–436;  
 B. Pullman, A. Pullman, *Quantum Biochemistry*, Interscience, New York, 1963;  
 (c) A. Streitwieser Jr., *Molecular Orbital Theory for Organic Chemists*, Wiley, New York, 1961.
- [13] B.J. Powell, T. Baruah, N. Bernstein, K. Brake, R.H. McKenzie, P. Meredith, M.R. Pederson, A first-principles density-functional calculation of the electronic and vibrational structure of key melanin monomers, *J. Chem. Phys.* 120 (2004) 8608–8615.
- [14] H.-B. Bürgi, J.D. Dunitz, J.M. Lehn, G. Wipff, Stereochemistry of reaction paths at carbonyl centres, *Tetrahedron* 30 (1974) 1563–1572.
- [15] S. Profeta Jr., N.L. Allinger, Molecular mechanics calculations on aliphatic amines, *J. Am. Chem. Soc.* 107 (1985) 1907–1918.