ChemPubSoc Europe

DOI: 10.1002/cphc.201801066

CHEMPHYSCHEM Articles

Computational Prediction of ^{1}H and ^{13}C NMR Chemical Shifts for Protonated Alkylpyrroles: Electron Correlation and Not Solvation is the Salvation

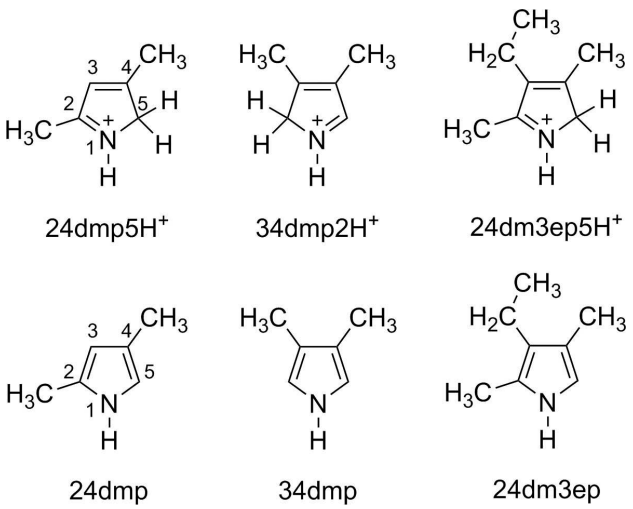
Evanildo G. Lacerda, Jr.,[a, c] Fadhil S. Kamounah,[a, b] Kaline Coutinho,[c] Stephan P. A. Sauer,*[a] Poul Erik Hansen,*[b] and Ole Hammerich*[a]

Prediction of chemical shifts in organic cations is known to be a challenge. In this article we meet this challenge for α-protonated alkylpyrroles, a class of compounds not yet studied in this context, and present a combined experimental and theoretical study of the ^{13}C and ^{1}H chemical shifts in three selected pyrroles. We have investigated the importance of the solvation model, basis set, and quantum chemical method with the goal of developing a simple computational protocol, which allows prediction of ^{13}C and ^{1}H chemical shifts with sufficient accuracy for identifying such compounds in mixtures. We find that density functional theory with the B3LYP functional is not sufficient for reproducing all ^{13}C chemical shifts, whereas already the simplest correlated wave function model, Møller–Plesset perturbation theory (MP2), leads to almost perfect agreement with the experimental data. Treatment of solvent effects generally improves the agreement with experiment to some extent and can in most cases be accomplished by a simple polarizable continuum model. The only exception is the NH proton, which requires inclusion of explicit solvent molecules in the calculation.

1. Introduction

The theoretical prediction of NMR spectra has progressed immensely in the last decade and a number of protocols useful for the organic chemist have been published.[1] However, common to the large benchmark data sets that have been used in these studies is that cations are almost absent. In addition, results from the studies dedicated to organic cations[2] have shown that accurate predictions of the chemical shifts pose a special challenge to the theoretical methods, sometimes demanding high-level *ab initio* methods such as coupled cluster theory.[2a–c,e] In this report we present the results of a computational study of the chemical shifts or nuclear magnetic shielding constants for three protonated alkylpyrroles. The aim is to develop a computational protocol that can assist in the identification of protonated alkylpyrroles in mixtures of their monomers and oligomers, for instance as they result from the oxidative coupling of pyrroles,[3] or from the classic acid-induced self-condensation reactions.[4]

The structures of the protonated and neutral species included in this study are shown in Scheme 1 together with the abbreviations to be used in the following and it is seen that the protonation of alkylpyrroles takes place in an unsubstituted α-position.[4c,5] The extent to which the results are affected by the choice of the solvation model, the presence of the counter ion, the basis set and different ways of treating electron correlation are addressed. The experimental data and, accordingly, the theoretical data, were all obtained in acetonitrile (CH_3CN) owing to the importance of this solvent in, for instance, the electrochemical oxidation of pyrroles.[3e–i,m] In addition, CH_3CN offers the advantage that the effects of ion-pairing with the counter-



Scheme 1. Structures and atom numbering for 2,4-dimethylpyrrole (24dmp), 3,4-dimethylpyrrole (34dmp), 2,4-dimethyl-3-ethylpyrrole (24dm3ep) and the corresponding protonated species.

[a] Dr. E. G. Lacerda, Jr., Dr. F. S. Kamounah, Prof. Dr. S. P. A. Sauer, Prof. Dr. O. Hammerich
Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen (Denmark)
E-mail: sauer@kiku.dk
o.hammerich@chem.ku.dk
[b] Dr. F. S. Kamounah, Prof. Dr. P. E. Hansen
Department of Science and Environment, Roskilde University, Universitetsvej 1, DK-4000 Roskilde (Denmark)
E-mail: poulerik@ruc.dk
[c] Dr. E. G. Lacerda, Jr., Prof. Dr. K. Coutinho
Instituto de Física da Universidade de São Paulo, CP 66318, 05314-970 São Paulo, SP (Brazil)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cphc.201801066

ion are expected to be small owing to the high dielectric constant (35.7[6]) of the solvent.

Methods

Computational Methods

Quantum Mechanical Calculations

The density functional theory (DFT) calculations, using the B3LYP[7] exchange-correlation functional as well as the second order Møller-Plesset perturbation theory (MP2)[8] calculations were carried out with Gaussian 09[9] implemented on standard work stations or available at the High Performance Computing Centre at the University of Copenhagen. For geometry optimizations we used the Dunning basis set cc-pVDZ, whereas for the shielding calculations we used both the standard energy optimized Pople basis sets 6-311++G(2d,p) as well as Jensen's (aug)-pcSseg-*n* ($n=1,2,3$) basis sets that are specially optimized for DFT calculations of shielding constants[1f] and can be downloaded from the basis set exchange database.[10] For the treatment of the solvent effects we employed both the polarizable continuum model (PCM)[11] as implemented in Gaussian as well as the average solvent electrostatic configuration (ASEC) approach,[12] where the solvent molecules are represented by explicit partial atomic charges, which are specified in the input files for the calculations with Gaussian. No vibrational corrections[13] have been applied in the calculations, as the question of combining solvent and vibrational effects is not settled[14] and the cost of performing vibrational correction calculations would anyway prevent them from being carried out routinely for larger organic molecules. Structure optimizations were carried out to the opt=tight level; true minima resulted in all cases as evidenced by the absence of imaginary frequencies in the frequency calculations.

Monte Carlo Simulations

The Monte Carlo simulations were carried out with the DICE program,[15] where the solute and the counter-ion were described by the OPLS force field parameters[16] for the van der Waals interactions and by partial atomic charges derived from QM calculations using the CHELPG scheme[17] obtained at the B3LYP/cc-pVDZ level. The solvent molecules were fully described by the OPLS force field.[18] We performed the MC simulation with the Metropolis sampling technique in the *NPT* ensemble at $T=300$ K and $P=1$ atm. A total of 500 solvent molecules were included in the simulations and the average density obtained for CH_3CN were 0.757 $g\,cm^{-3}$ in good agreement with the experimental value[19] of 0.776 $g\,cm^{-3}$ at the same thermodynamic conditions. All molecules were kept rigid during the simulation but were free to translate and rotate. The geometries of the solute ($24dmp5H^+$) and the counter-ion (CF_3COO^-) used in the MC simulation were optimized separately in PCM at the B3LYP/cc-pVDZ level and in the start of the MC simulation placed in random positions in the simulation box. The geometry of CH_3CN was optimized at the HF/6-311G(d,p) level. The MC simulation starts with a random configuration, where all molecules (solute, counter-ion and solvent, 500) were placed in random positions in the simulation box.

Experimental

Chemicals

2,4-dimethylpyrrole (Aldrich, 97%), 3,4-dimethylpyrrole (Synthon Chemicals, 98%), and 2,4-dimethyl-3-ethylpyrrole (Aldrich, 97%), trifluoroacetic acid (Fluka, 98%), trifluoromethanesulfonic acid (Aldrich, 99%) and acetonitrile-d_3 (Aldrich) were all used as received.

NMR Spectroscopy

The ^{1}H NMR spectra were recorded at 300 MHz and the ^{13}C NMR spectra at 75 MHz on a Varian Mercury 300 instrument at room temperature. The solvent was acetonitrile-d_3. An acquisition time of 4 s was used for the ^{1}H NMR spectra. The HSQC and HMBC spectra were recorded at 600 MHz at a Varian Inova or a Bruker 500 MHz instrument using standard recording parameters. Data and assignments are summarized in the SI.

Neutral pyrroles: A solution of 24dmp (36.2 mg; 0.38 mmol) or 34dmp (36.2 mg; 0.38 mmol) or 24dm3ep (46.9 mg; 0.38 mmol) in acetonitrile-d_3 (0.75 ml) were added at room temperature under a nitrogen atmosphere into an NMR tube.

Protonated pyrroles: A solution of 24dmp (36.2 mg; 0.38 mmol) or 34dmp (36.2 mg; 0.38 mmol) or 24dm3ep (46.9 mg; 0.38 mmol) in acetonitrile-d_3 (0.75 ml) were added at room temperature under a nitrogen atmosphere to trifluoroacetic acid (86.5 mg; 0.76 mmol, 2.0 equivalents), for 24dmp and 24dm3ep, or the stronger trifluoromethanesulfonic acid (114.05 mg; 0.76 mmol, 2.0 equivalents), for 34dmp owing to its lower basicity. The resulting homogeneous mixtures were stirred for 10 minutes and then charged into an NMR tube under a nitrogen atmosphere.

2. Results and Discussion

In the following, we will present first the results of our more extensive study of the ^{13}C and ^{1}H chemical shifts for one of the α-protonated alkylpyrroles, $24dmp5H^+$. Afterwards we will validate the developed computational protocol by applying it to the two other protonated species, $34dmp2H^+$ and $24dm3ep5H^+$, and then, finally, to the three neutral pyrroles, 24dmp, 34dmp and 24dm3ep.

Since it is not possible to compare directly the calculated absolute shielding constants, σ_{calc}, and experimental chemical shifts, δ_{exp}, (summarized in Tables S1–S6 in the Supporting Information) we should also calculate the corresponding absolute shielding constants of a reference molecule, e.g., tetramethylsilane. However, this is inconvenient for a study developing a computational protocol. We will therefore employ the often used approach of an internal reference[1k,2d,20] and compare the experimental chemical shifts, δ_{exp}, with calculated chemical shifts, δ_{calc}, which are obtained by least-squares fitting

of the calculated shielding constants to the experimental shifts according to Equation (1):

$$\delta = a\,\sigma_{calc} + b \qquad (1)$$

($\delta = \delta_{exp}$) with a and b being adjustable parameters. Consequently, the better the quality of the σ_{calc} values, the better the correlation with the δ_{exp} values. Our calculated chemical shifts, δ_{calc}, are then obtained from Equation (1) ($\delta = \delta_{calc}$) using the fitted parameters a and b and the calculated absolute shieldings σ_{calc}. Our goal then is to develop a computational protocol to get the coefficient of determination of the fits, R^2, as close as possible to unity and the standard deviation, *SD*, small. Furthermore, we will report the maximum deviation, *MaxDev*, for the C or H atom for which the largest difference between δ_{calc} and δ_{exp} is observed. The slope a is interpreted as a scaling factor,[1k,2d,20] but obtaining values close to -1 would allow direct comparison of the calculated shieldings and the experimental chemical shifts.

The focus of the discussion will be (i) on the effect of solvation by CH_3CN, where we investigate different implicit and explicit solvation models and the role of the counter-ion, trifluoroacetate (CF_3COO^-), (ii) on the choice of the basis set and (iii) on the treatment of electron correlation. The solvent models employed are the polarizable continuum model (PCM),[11] a sequential QM/MM method called the average solvent electrostatic configuration model (ASEC),[12] and the PCM with one or more explicit CH_3CN molecules or explicitly with the counter-ion, CF_3COO^-, in order to test for specific solvation of the protonated pyrrole, possible ion-pair effects or hydrogen bonding to the NH group. In the tables and figures to follow these calculations will be denoted as 'X+PCM', 'X+ASEC', 'X+1S+PCM' and 'X+A+PCM', where the solute molecule is referred to as 'X', the solvent molecules as 'S' and the counter ion as 'A'. The geometry of the X+S and X+A complexes were optimized at the same level as applied to the isolated 24dmp5H$^+$.

In addition, three sets of calculations were carried out where the positions of the solvent molecules were obtained from a Monte Carlo (MC) simulation. Two hundred statistically uncorrelated configurations from the MC simulation were selected and the geometry of the complex of 24dmp5H$^+$ with the nearest (X+1S), the five nearest (X+5S) or the 10 nearest (X+10S) solvent molecules were extracted from each of these configurations. For each of these three sets of 200 clusters a PCM calculation of the shielding constants was carried out and the 200 results for the shielding constants were finally averaged, leading to the 'MC X+1S+PCM', 'MC X+5S+PCM' and 'MC X+10S+PCM' results in the following tables.

The discussion will focus on the statistical data that are tabulated below. The calculated absolute shielding constants are tabulated in the SI.

2.1. 24dmp5H$^+$

2.1.1. Effects of the Choice of Solvation Model, Presence of a Counter Ion, and One or More Explicit Solvent Molecules

In the first step the carbon and hydrogen shielding constants σ_{calc} were calculated for 24dmp5H$^+$ both in vacuum and in CH_3CN using the PCM. The geometries of the isolated 24dmp5H$^+$ and of 24dmp5H$^+$ in CH_3CN were optimized correspondingly either in vacuum or in the PCM model for CH_3CN both at the B3LYP/cc-pVDZ level. For calculation of the shieldings, the B3LYP/6-311++G(2d,p) model was employed in this part of the study as recommended for ^{13}C[1g] with the slight modification that we added also diffuse functions for the hydrogen atoms.

The statistical results for ^{13}C are summarized in Table 1 and illustrated for the X+PCM model in Figure 1. A complete set of plots corresponding to all entries of Table 1 is available as Figures S1–S8. The calculated shielding constants are summarized in Table S7.

Looking at the statistics for all models we see a high correlation coefficient R^2 above 0.99. Secondly, the C-4 atom that carries a partial positive charge in the classical resonance structures has always the highest deviation with *MaxDev* values

Table 1. ^{13}C NMR statistical data for 24dmp5H$^+$ fitted to Equation (1) testing the effect of the solvation model, the presence of the counter ion and of explicit solvent molecules.[a]

Model[b]	R^2	*SD*	*MaxDev*	Slope
X	0.9970	3.8	6.5 (C-4)	−0.9311
X+PCM	0.9982	2.9	5.2 (C-4)	−0.9441
X+ASEC	0.9981	3.1	5.5 (C-4)	−0.9463
X+A+PCM	0.9988	2.4	4.5 (C-4)	−0.9650
X+1S+PCM	0.9986	2.6	4.6 (C-4)	−0.9496
MC X+1S+PCM	0.9984	2.8	5.0 (C-4)	−0.9469
MC X+5S+PCM	0.9986	2.6	4.7 (C-4)	−0.9576
MC X+10S+PCM	0.9984	2.8	4.9 (C-4)	−0.9674

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p). [b] The abbreviations are explained in the text.

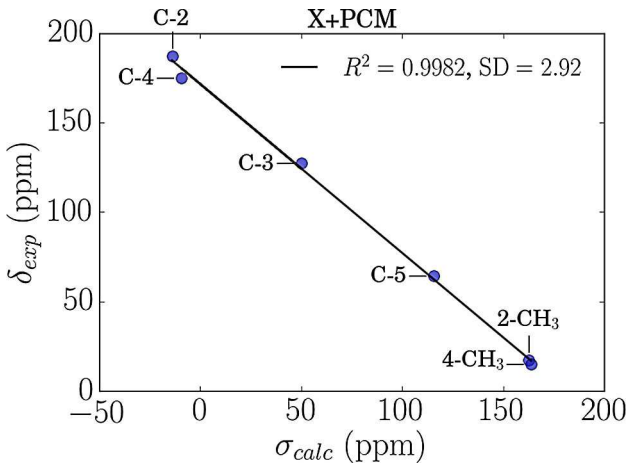


Figure 1. Experimental chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} for 24dmp5H$^+$ modeled as X+PCM.

between 4.5 and 6.5 ppm. The worst statistics are obtained for the calculations on the isolated solute X in the gas phase, with the lowest R^2 value of 0.9970, the largest *SD* of 3.8 ppm, the highest *MaxDev* (6.5 ppm) and a slope, −0.9311, that is significantly less than unity.

The inclusion of solvent effects as described by the PCM (see Figure 1) improves somewhat the agreement with the experimental values, since the overall correlation between the theoretical and experimental values is slightly improved, R^2 from 0.9970 (X) to 0.9982 (X + PCM) and *SD* from 3.8 ppm (X) ppm to 2.9 ppm (X + PCM). Similarly, also the deviation of the C-4 chemical shift is reduced (*MaxDev* = 5.2 ppm) but it is still rather large and stands out as can clearly be seen in Figure 1. Secondly, the slope, −0.9441, is also only slightly better than for the pure gas phase calculation. Introduction of the solvent via the PCM model clearly helps but is not sufficient yet.

Similar results as for PCM are obtained with the alternative electrostatic embedding model ASEC. Measured by the statistical parameters, X + PCM performs slightly better than X + ASEC with an *SD* (2.9 versus 3.1 ppm) and *MaxDev* (5.2 versus 5.5 ppm), although the slope is slightly better for X + ASEC. Possibly, as we are dealing with a positively charged solute, the PCM solvent model allows the continuous environment to polarize accordingly leading to improved results, while in ASEC the solvent point charges have previously fixed values generated for neutral environments. Nevertheless, the differences in the statistics between both models are small and none of them gives a significantly better agreement for C-4. Since our goal is to develop a protocol that is able to differentiate atoms that have close chemical shifts, as e.g. C-2 and C-4, the electrostatic embedding results are not quite satisfactory. However, both approaches do not take specific solute-solvent or solute-counter-ion interactions into account. For those one has to include explicit solvent (S) or counter-ion (A) molecules in the calculations. In the following only the PCM model will be considered further as ASEC cannot include explicit solvent molecules.

The next and natural step to improve the calculated chemical shifts is to consider the effect of the counter-ion CF_3COO^-. In order to find out, where one should place the counter-ion, we have performed a Monte Carlo (MC) simulation with 500 explicit acetonitrile molecules as the solvent and with CF_3COO^- as the counter-ion. The superposition of 100 statistically uncorrelated configurations taken from the MC simulation is shown in Figure 2. It becomes clear that the counter-ion does not stay close to the solute during the MC simulation.

Nevertheless, by including 'A', placed anyway close to the NH group and re-optimizing the geometry of the X + A + PCM system at the B3LYP/cc-pVDZ level of theory we obtain another improvement of the statistics. The *SD* value is now 2.4 ppm, the deviation of C-4 is reduced to 4.5 ppm and the slope is increased to −0.9650, while R^2 is practically unchanged compared to the model with no counter-ion. In the same way as the counter-ion an explicit solvent molecule (S) was added and the geometry of the X + 1S + PCM system was re-optimized. The ^{13}C chemical shifts are again better than for the X + PCM system but marginally worse than on explicit inclusion of the counter-ion.

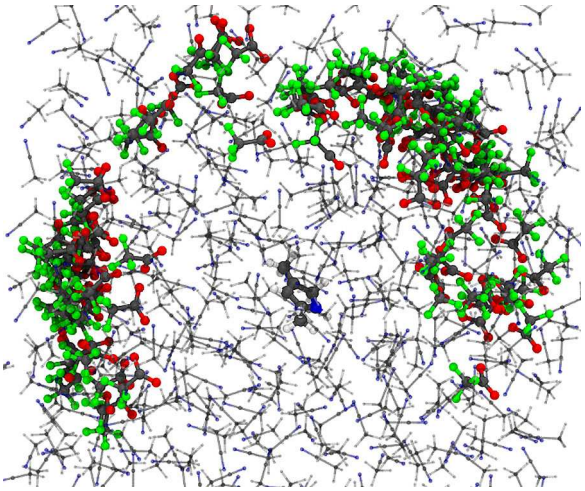


Figure 2. Superimposition of 100 statistically uncorrelated configurations generated by the MC simulations showing a perspective of the solute with the solvent molecules distribution and the positions of the counter-ion (CF_3COO^-) in the simulation box. The color scheme is grey = C, white = H, blue = N, red = O and green = F.

However, selecting only one solvent molecule from the MC simulation might be a somewhat biased approach. In addition, the solvent molecules remain in a solution obviously not fixed in the same position. We have therefore tested also the costlier approach of explicitly averaging over 200 uncorrelated conformations from the MC simulation as described previously. The statistics for these averaged results is shown in Table 1 (the full data in Table S7) in the rows MC X + 1S + PCM, MC X + 5S + PCM and MC X + 10S + PCM, respectively for 1, 5 or 10 explicit solvent molecules. The *SD* values are 2.8 ppm for one explicit solvent molecule, 2.6 for five and 2.8 for ten with *MaxDev* values of 5.0 ppm, 4.7 ppm and 4.9 ppm. These values are thus slightly lower than the values for the pure PCM environment X + PCM, indicating again the importance of including explicit solvent molecules in the calculations. However, all the MC X + nS + PCM results are, apart from the slope, worse than the results of the single conformation calculation X + 1S + PCM. This implies that averaging the shieldings over many configurations extracted from MC simulations does not improve the results compared with calculations from single optimized configurations X + A + PCM and X + 1S + PCM. Even the solvent model MC X + 10S + PCM, which includes 10 explicit solvent molecules in the calculation did not show any improvement compared to the relatively cheaper calculation X + 1S + PCM. The geometry optimization of the solute in the presence of the explicit molecule might explain the better results compared to the MC X + nS + PCM approach, where the geometry of the solute was optimized without the presence of the solvent or the counter-ion. One should also note that the structures and energies of X + 1S + PCM and MC X + 1S + PCM are not the same. The X + 1S + PCM system was optimized and has therefore the lower energy. Furthermore, it has a perfect linear hydrogen-bond

from NH in 24dmp5H$^+$ to N in CH_3CN, whereas this is not the case in any of the configurations of the MC X+1S+PCM calculation. Actually not all of these configurations have the solvent molecule at all hydrogen-bonded to the NH. Looking at the larger simulations, MC X+5S+PCM and MC X+10S+PCM, this idealized linear hydrogen-bond is not found either, but on the other hand there are several conformations with two CH_3CN molecules hydrogen-bonded to the NH in 24dmp5H$^+$. From the full set of data in Table S7 we can see that with increasing number of explicit solvent molecules the two carbons, C-2 and C-4, that carry a positive charge get more deshielded, while the other carbons, C-3, C-5, 2-CH_3 and 4-CH_3, get more shielded. Finally, analyzing the MC simulation in more detail we note that the counter-ion CF_3COO^- does not stay close to the solute during the MC simulation (see Figure 2). That could be the evidence that the solute and counter-ion do not interact closely in the experiments.

Overall, from Table 1 we can draw some partial conclusions. Including solvent effects is definitely necessary for somewhat improving the statistical results and the cheaper and effective way to include it is using PCM. Also, MC simulations seem not to be essential for getting better results, but inclusion of either an explicit solvent molecule or the counter-ion are important, where the last one so far looks most promising. However, it is statistically difficult to say, based only on the ^{13}C NMR data, whether the counter ion is close to the solute or not. Furthermore, there is clearly room for improvement of the *SD*, and in particular of the *MaxDev*, i.e. the too large deviation for C-4, and for obtaining a slope closer to −1. In order to get more insights, we will in the following analyze also the ^{1}H NMR chemical shift data for the same models.

The statistical results for ^{1}H are summarized in Table 2 and illustrated for the X+PCM model in Figure 3. A complete set of

Table 2. ^{1}H NMR statistical data for 24dmp5H$^+$ fitted to Equation (1) testing the effect of the solvation model, the presence of the counter ion and of explicit solvent molecules.[a]

Model[b]	NH[c]	R^2	*SD*	*MaxDev*	Slope
X	yes	0.8823	1.1	1.8 (H-3)	−1.5050
	no	0.9913	0.2	0.2 (H_2-5)	−1.0146
X+PCM	yes	0.9206	0.9	1.5 (H-3)	−1.4165
	no	0.9967	0.1	0.2 (H_2-5)	−0.9947
X+ASEC	yes	0.9233	0.9	1.5 (H-3)	−1.4187
	no	0.9968	0.1	0.1 (H_2-5)	−1.0000
X+A+PCM	yes	0.9504	0.7	1.3 (H-3)	−0.5540
	no	0.9985	0.1	0.1 (2-CH_3)	−1.0129
X+1S+PCM	yes	0.9989	0.1	0.2 (H_2-5)	−1.0022
	no	0.9956	0.1	0.2 (H_2-5)	−1.0194
MC X+1S+PCM	yes	0.9258	0.9	1.5 (H-3)	−1.4117
	no	0.9965	0.1	0.2 (H_2-5)	−0.9987
MC X+5S+PCM	yes	0.9331	0.9	1.4 (H-3)	−1.3750
	no	0.9964	0.1	0.2 (H_2-5)	−0.9836
MC X+10S+PCM	yes	0.9362	0.8	1.4 (H-3)	−1.3796
	no	0.9965	0.1	0.2 (H_2-5)	−0.9920

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p). [b] The abbreviations are explained in the text. [c] Including (yes) or excluding (no) the data point for the NH proton. See the text.

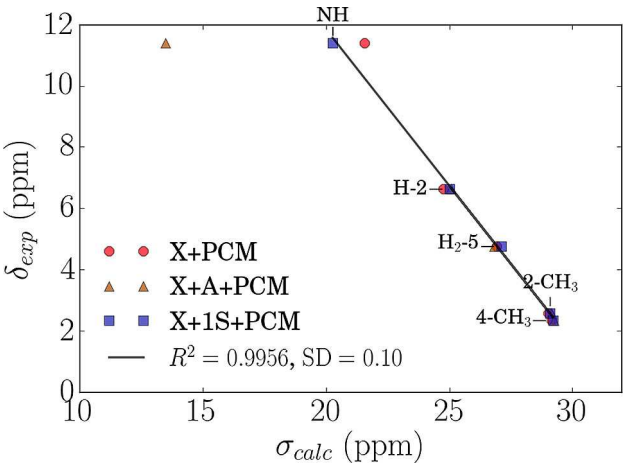


Figure 3. Experimental ^{1}H NMR chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} for the molecule at the models, X+PCM, X+A+PCM and X+1S+PCM. (from Figures S10, S12 and S13). The statistical data are for X+1S+PCM, omitting the data point for NH in the regression.

plots corresponding to all entries of Table 2 is available as Figures S9–S16. The calculated shielding constants are summarized in Table S8. On fitting the ^{1}H results to Equation (1) we noted that the fits could strongly be improved by excluding the NH proton from the data set for the fitting. We present therefore in Table 2 both the statistics for the fits with and without the NH proton.

For all the fits without the NH proton we thus obtain again a correlation coefficient R^2 above 0.99, standard deviations of maximal 0.2 ppm and *MaxDev* of 0.1 or 0.2 ppm mostly for the H_2-5 proton. Also the slopes of the fits deviate at most by 2% from unity. The differences between the different models for treating the solvation effects are small, but nevertheless the observed trend is similar as for the ^{13}C chemical shifts. Inclusion of any solvation model improves the statistics somewhat compared to the gas phase calculations. The sampling over many conformation calculations, MC X+nS+PCM, do not offer any advantage over the simpler X+PCM, X+ASEC or X+1S+PCM calculations. On the other hand, different from the ^{13}C data, the X+ASEC calculation performs in three out of the four statistical data better than the X+PCM model for the ^{1}H chemical shifts. It actually gives the perfect slope of unity on excluding the NH proton. Also including the counter-ion appears to give a better statistic than including one explicit solvent molecule.

However, including the NH proton in the fits gives for all but one calculation much worse statistics showing it not to be properly described by these models. This can clearly be seen in Figure 3 with the results for the ^{1}H NMR chemical shifts from the three models: X+PCM, X+A+PCM and X+1S+PCM.

The best fit is obtained, when we include an explicit solvent molecule hydrogen-bonded to the NH group, while the worst result is obtained, when we have the counter ion placed explicitly close to the NH group. Owing to the hydrogen-bonding donating properties of 24dmp5H$^+$ the X+1S+PCM model performs very good as observed also for other NH

containing molecules,[21] while the X+A+PCM model gives a too large change from the pure PCM model. The statistical data are only for the X+1S+PCM model similar on inclusion or exclusion of the NH proton in the fit and the X+1S+PCM model with the NH proton has actually the highest R^2 value and an almost perfect slope of −1.0022. It is interesting to note that the only significant difference in the shieldings between the three models in Figure 3 are for the NH proton, while all the others, H-2, H_2-5, 2-CH_3 and 4-CH_3, have similar shieldings for the three calculations X+PCM, X+A+PCM, X+1S+PCM. This indicates strongly that the presence of the counter-ion close to the NH group in the calculation is neither necessary nor important for describing the shieldings of protonated pyrroles in CH_3CN in contrast to a hydrogen-bonded solvent molecule. That might not be the case for a solvent with lower dielectric constant such as chloroform, where the ion-pair may not dissociate, but be important for describing the shieldings, or for cations which are worse hydrogen-bond donors than protonated *N*-aromatic compounds.[2d,i]

So far, all NMR calculations have been carried out with the same quantum chemical model, i.e. B3LYP/6-311++G(2d,p). In the following we will discuss whether the remaining disagreement with experiment, in particular the disagreement for C-4, is caused by our choice of quantum chemical model, i.e. whether we can obtain a better agreement by using better basis sets or a correlated wave function method.

2.1.2. Basis Set Effects

We will discuss first the effect of the basis set on the shieldings at the DFT/B3LYP level considering the three best solvation models: X+PCM, X+A+PCM and X+1S+PCM. The only family of basis sets especially optimized for the calculation of shieldings constants at the DFT level are, to our knowledge, the (aug-)pcSseg-n basis sets by Jensen.[1f] We tested therefore three basis sets in this series, pcSseg-1, pcSseg-2 and pcSseg-3, which correspond to polarized double-, triple- and quadruple-zeta basis sets. In addition, we investigated the effect of the additional diffuse functions in the aug-pcSseg-1 basis set. The structures are unchanged, i.e. were optimized at the B3LYP/cc-pVDZ level as in the previous sections. The results for the ^{13}C chemical shifts are shown in Tables 3 and S9 and illustrated in Figure 4. A complete set of plots is available in the SI as Figures S17–S28. The corresponding data for the ^{1}H chemical shifts are summarized in Tables 4 and S10 and illustrated in Figure 5. A complete set of plots is available in the SI (Figures S29–S40).

Analyzing first the ^{13}C shieldings in Table S9 one can observe that the absolute shieldings nicely converge within this series of basis sets, i.e. that the differences between the pcSseg-2 and pcSseg-3 results are all less than 1 ppm. Furthermore, the effect of augmenting the pcSseg-1 basis set with additional diffuse functions is also smaller than the change on going to pcSseg-2 and one can therefore expect that augmenting the larger basis sets will not lead to significant changes. However, looking at the statistics for the comparison

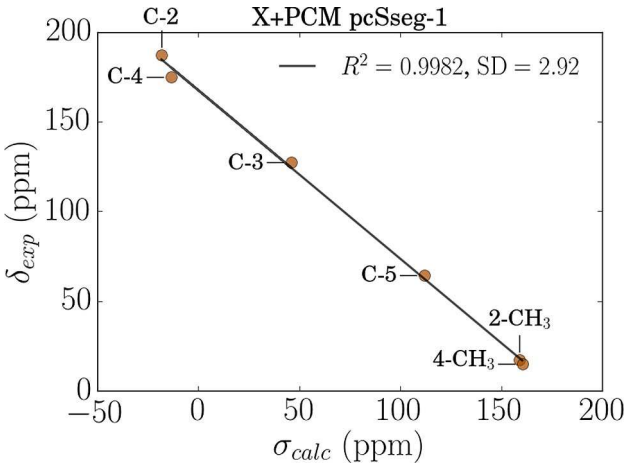


Figure 4. Experimental chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} for 24dmp5H$^+$ modeled as X+PCM with the pcSseg-1 Jensen basis set.

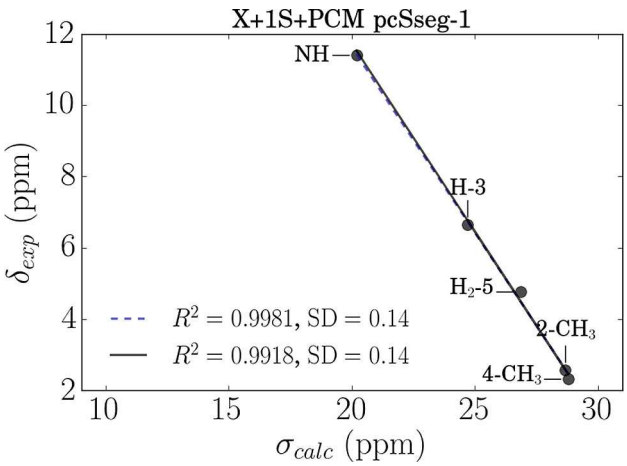


Figure 5. Experimental chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} for 24dmp5H$^+$ modeled as X+1S+PCM with the pcSseg-1 Jensen basis set excluding (full line) or including (dotted line) the data point for the NH proton.

with the experimental chemical shifts, Table 3, one observes on one hand the same convergence but unfortunately a convergence to somewhat larger values. Although the standard deviations of the fits and the deviation of C-4 are a bit better with the pcSseg-1 basis sets than with the 6-311++G(2d,p) basis set for the X+A+PCM and X+1S+PCM solvent models, they deteriorate slightly again for the large basis sets. This on the first look disappointing behavior has on the other hand a simple explanation. The B3LYP/pcSseg-3 results are certainly close to the basis set limit and thus show the true quality obtainable with the B3LYP functional without error cancellation due to a too small basis set.[22] In addition to the inherent error in the B3LYP functional there is also still the possibility of a remaining error due to the level of theory employed in the optimization of the geometry. Summarizing one has to say that the problem with the large deviation of C-4 is not solved by

Table 3. ^{13}C NMR statistical data for 24dmp5H$^+$ fitted to Equation (1) testing the effect of the basis set.[a]

Basis set and model[b]	R^2	*SD*	*MaxDev*	Slope
pcSseg-1				
X+PCM	0.9982	2.9	5.2 (C-4)	−0.9394
X+A+PCM	0.9990	2.2	3.9 (C-4)	−0.9610
X+1S+PCM	0.9988	2.4	4.3 (C-4)	−0.9450
aug-pcSseg-1				
X+PCM	0.9979	3.2	5.6 (C-4)	−0.9281
X+A+PCM	0.9989	2.3	4.1 (C-4)	−0.9458
X+1S+PCM	0.9985	2.7	4.6 (C-4)	−0.9320
pcSseg-2				
X+PCM	0.9978	3.3	5.7 (C-4)	−0.9080
X+A+PCM	0.9988	2.5	4.6 (C-4)	−0.9289
X+1S+PCM	0.9985	2.7	4.8 (C-4)	−0.9138
pcSseg-3				
X+PCM	0.9978	3.3	5.7 (C-4)	−0.9076
X+A+PCM	0.9988	2.4	4.5 (C-4)	−0.9278
X+1S+PCM	0.9984	2.8	4.9 (C-4)	−0.9127

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/see above. [b] The abbreviations are explained in the text.

using a better basis set as one can see from the *MaxDev* values in Table 3.

From Table S10 we note that the only ^1H absolute shielding, which is significantly influenced by the changes in basis sets, is the NH proton shielding. Concentrating therefore on the X+1S+PCM statistics in Table 4, we can draw the same conclusion as

Table 4. ^1H NMR statistical data for 24dmp5H$^+$ fitted to Equation (1) testing the effect of the basis set.[a]

Basis set and model[b]	NH[c]	R^2	*SD*	*MaxDev*	Slope
pcSseg-1					
X+PCM	yes	0.9256	0.9	1.5 (H-3)	−1.4724
	no	0.9953	0.1	0.2 (H_2-5)	−1.0415
X+A+PCM	yes	0.9476	0.8	1.3 (H-3)	−0.5563
	no	0.9986	0.1	0.1 (2-CH_3)	−1.0430
X+1S+PCM	yes	0.9981	0.1	0.3 (H_2-5)	−1.0396
	no	0.9918	0.1	0.3 (H_2-5)	−1.0535
aug-pcSseg-1					
X+PCM	yes	0.9426	0.8	1.3 (H-3)	−1.4177
	no	0.9969	0.1	0.2 (H_2-5)	−1.0309
X+A+PCM	yes	0.9451	0.8	1.4 (H-3)	−0.5407
	no	0.9976	0.1	0.1 (2-CH_3)	−1.0350
X+1S+PCM	yes	0.9972	0.2	0.3 (H_2-5)	−0.9713
	no	0.9936	0.1	0.2 (H_2-5)	−1.0449
pcSseg-2					
X+PCM	yes	0.9372	0.8	1.4 (H-3)	−1.3963
	no	0.9963	0.1	0.2 (H_2-5)	−1.0057
X+A+PCM	yes	0.9469	0.8	1.3 (H-3)	−0.5341
	no	0.9986	0.1	0.1 (2-CH_3)	−1.0072
X+1S+PCM	yes	0.9978	0.2	0.3 (H_2-5)	−0.9652
	no	0.9937	0.1	0.2 (H_2-5)	−1.0184
pcSseg-3					
X+PCM	yes	0.9414	0.8	1.4 (H-3)	−1.3908
	no	0.9965	0.1	0.2 (H_2-5)	−1.0092
X+A+PCM	yes	0.9452	0.8	1.3 (H-3)	−0.5289
	no	0.9987	0.1	0.1 (2-CH_3)	−1.0134
X+1S+PCM	yes	0.9972	0.2	0.3 (H_2-5)	−0.9518
	no	0.9933	0.1	0.2 (H_2-5)	−1.0212

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/see above. [b] The abbreviations are explained in the text. [c] Including (yes) or excluding (no) the data point for the NH proton. See the text.

for the ^{13}C chemical shifts, converging the absolute shieldings with respect to the basis set at the B3LYP level does not improve the agreement with experiment. And overall we have to conclude that the problem with the too large deviation of C-4 is not solved by improving the basis set. In the following we will therefore investigate the effect of replacing the most frequently employed DFT functional, B3LYP, by the most frequently employed correlated wave function method, second order Møller-Plesset perturbation theory (MP2).[8,23]

2.1.3. HF and MP2 Levels

The MP2 calculations have been carried out again at the B3LYP/cc-pVDZ optimized geometries in order to study solely the effect of electron correlation on the shielding calculations. As basis set we employed now the pcSseg-1 basis set. In addition to the MP2 calculations also uncorrelated Hartree-Fock (HF) calculations were carried out in order to get a measure of the correlation effect on the shieldings. The results for the ^{13}C chemical shifts are shown in the Tables 5 and S11 and illustrated in Figure 6. (A complete set of plots is available in

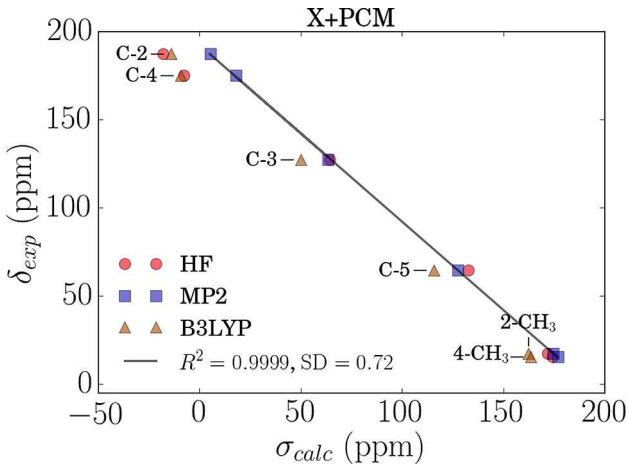


Figure 6. Experimental ^{13}C NMR chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} at the model 'X+PCM' for 24dmp5H$^+$. The regression line is for MP2. The data points for B3LYP are the same as in Figure 1 and are included for comparison.

the SI as Figures S41–S43), while the corresponding data for the ^1H chemical shifts are in Tables 6 and S12 and illustrated in Figure 7. (A complete set of plots is available in the SI as Figures S44–S46).

Comparison of the statistical data for MP2 in Table 5 with the corresponding B3LYP data in Table 3 shows great improvement on all parameters for the X+PCM and X+1S+PCM models. It is in particular gratifying to see, that using MP2 the C-4 carbon has no longer the largest deviation and that the maximum deviation is overall reduced to 1.4 ppm in the X+PCM model. Also the *SD* value changed for the X+PCM model

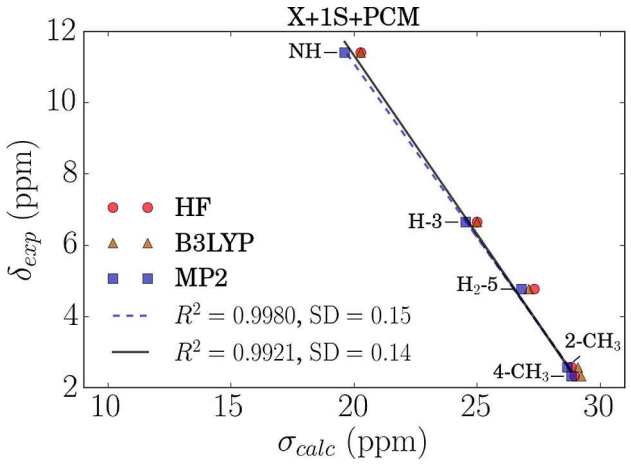


Figure 7. Experimental ^{1}H NMR chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} at the model 'X+1S+PCM' for 24dmp5H$^+$. The regression lines are for MP2 excluding (full line) or including (dotted line) the data point for the NH proton. The data points for B3LYP are the same as in Figure 1 and are included for comparison.

Table 5. ^{13}C NMR statistical data for 24dmp5H$^+$ fitted to Equation (1) testing the effect of using HF and MP2 calculations.[a]

Method and model[b]	R^2	*SD*	*MaxDev*	Slope
HF				
X+PCM	0.9915	6.4	10.2 (C-3)	−0.8809
X+A+PCM	0.9946	5.1	7.4 (C-3)	−0.9046
X+1S+PCM	0.9927	6.0	8.9 (C-3)	−0.8879
MP2				
X+PCM	0.9999	0.7	1.4 (C-3)	−1.0024
X+A+PCM	0.9987	2.5	4.8 (C-3)	−1.0319
X+1S+PCM	0.9996	1.4	3.0 (C-3)	−1.0108

[a] Opt: B3LYP/cc-pVDZ. NMR: HF/pcSseg-1 or MP2/pcSseg-1. [b] The abbreviations are explained in the text.

from 2.9 ppm at the B3LYP level to 0.7 ppm for the MP2 approach and the R^2 value and the slope are almost perfect.

It is interesting to compare the behavior of the individual carbon atoms in the three computational levels for the X+PCM model in Figure 6. While the MP2 results exhibit an almost perfect linear correlation with the experimental chemical shifts, the B3LYP results are in general shifted to lower absolute shieldings and suffer from C-4 being an outlier. The HF results are on one side quite close but below the MP2 results for the two methyl-groups and close but above the MP2 results for C-3 and C-5, but are on the other side close to the B3LYP results for the two carbon atoms carrying a partial positive charge: C-2 and C-4. One way to interpret this is to say, that the B3LYP calculation does not recover the relative large correlation effects for these two carbon atoms, while at the same time overestimating the correlation effect for the other carbon atoms. The same effect, i.e. prediction of a correlation effect by B3LYP calculations when there is almost none, has previously been observed, for example, for simple systems as the chemical shifts in noble gas dimers.[24]

Table 6. ^{1}H NMR statistical data for 24dmp5H$^+$ fitted to Equation (1) testing the effect of using HF and MP2 calculations.[a]

Method and model[b]	NH[c]	R^2	*SD*	*MaxDev*	Slope
HF					
X+PCM	yes	0.9289	0.9	1.6 (H-3)	−1.4943
	no	0.9835	0.2	0.4 (H_2-5)	−1.0641
X+A+PCM	yes	0.9401	0.8	1.4 (H-3)	−0.5325
	no	0.9993	<0.1	0.1 (2-CH_3)	−1.0685
X+1S+PCM	yes	0.9937	0.3	0.5 (H_2-5)	−1.0206
	no	0.9754	0.3	0.5 (H_2-5)	−1.0779
MP2					
X+PCM	yes	0.9477	0.8	1.3 (H-3)	−1.3508
	no	0.9950	0.1	0.2 (H_2-5)	−0.9928
X+A+PCM	yes	0.9482	0.8	1.3 (H-3)	−0.5347
	no	0.9994	<0.1	0.1 (2-CH_3)	−0.9986
X+1S+PCM	yes	0.9980	0.1	0.3 (H_2-5)	−0.9731
	no	0.9921	0.1	0.3 (H_2-5)	−1.0036

[a] Opt: B3LYP/cc-pVDZ. NMR: HF/pcSseg-1 or MP2/pcSseg-1. [b] The abbreviations are explained in the text. [c] Including (yes) or excluding (no) the data point for the NH proton. See the text.

For the ^{1}H chemical shifts the differences between the MP2 results and B3LYP results (absolute values in Tables S10 and S12, statistics in Tables 4 and 6) are marginal. Furthermore, it is also at the MP2 level only the X+1S+PCM model, which is able to reproduce the NH proton chemical shift, giving equally good statistics for fits with and without the NH proton, Figure 7.

2.1.4. Influence of the Geometry

One point, which has not been investigated yet, is the influence of the geometry on the calculated shieldings. In Figure 8 we present plots of the calculated ^{13}C shieldings versus the experimental chemical shifts calculated at the B3LYP (Figure 8 top) and MP2 (Figure 8 bottom) level for geometries optimized at the B3LYP/cc-pVDZ and MP2/cc-pVDZ level. One can see that there is virtually no influence of the geometry on these data. We will in the following therefore continue using the B3LYP/cc-pVDZ optimized geometries.

2.1.5. Conclusions Based on Results Obtained for 24dmp5H$^+$

So far, the conclusions from our calculations for the 24dmp5H$^+$ cation are the following: (1) For the ^{13}C NMR chemical shifts it might be enough to treat the effects of the solvent CH_3CN by the simple implicit solvent model PCM. (2) In order to reproduce the NH proton ^{1}H chemical shifts it is necessary to include one explicit solvent molecule hydrogen bonded to the NH group, the X+1S+PCM model. Including explicitly of the counter-ion deteriorates the results, which indicates that there is no strong interaction between the solute and the counter-ion in this solvent. (3) A consistent reproduction of all ^{13}C chemical shifts, including the partially positively charged C-4, is only possible at the MP2 level. The last conclusion confirms for our class of cations, the protonated alkylpyrroles, what previously has been found for smaller cations such as allyl,[25] vinyl,[2b,c]

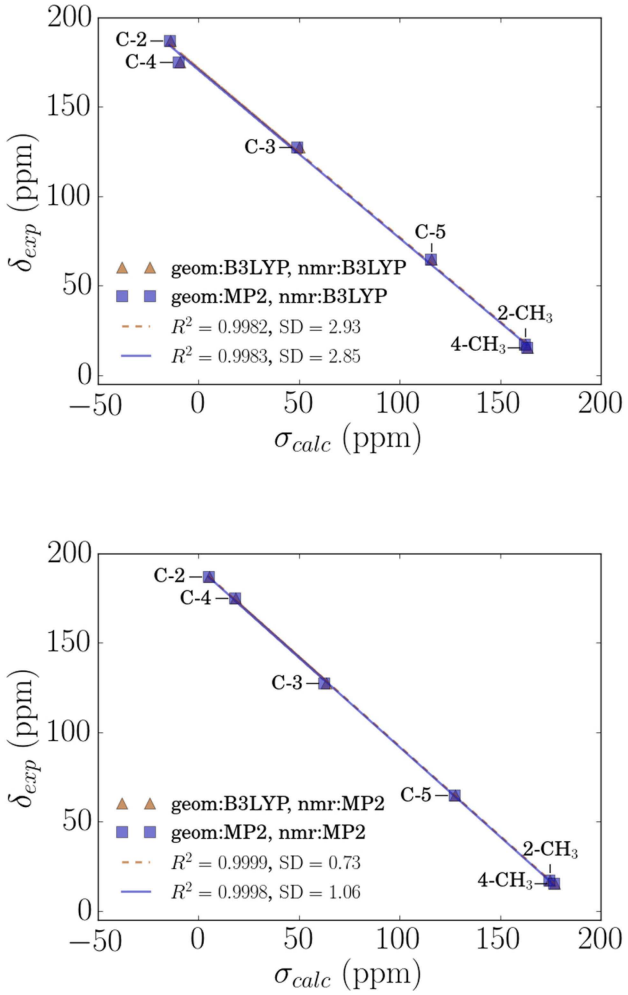


Figure 8. Experimental ^{13}C NMR chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} for 24dmp5H$^+$ calculated with the B3LYP/ 6-311++G(2d,p) (top) and MP2/pcSseg-1 (bottom) level at an B3LYP/cc-pVDZ and an MP2/cc-pVDZ optimized geometry and the X+PCM solvation model.

dienyl[2a,g] or even adamantly cations.[2e] HF and DFT at least with the B3LYP functional fail for unsaturated carbocations.[2h] In particular the positively charged carbon atoms are overly deshielded at these levels of theory due to an overestimation of the paramagnetic contribution. These non-systematic deviations of these carbon atoms compared to the remaining carbon atoms caused then problems with fitting approaches like Equation (1), as we have also observed for the protonated alkylpyrroles in Tables 1, 3, and 5. Systematic inclusion of electron correlation already at the low level of MP2, however, was shown to be able to overcome this problem and lead to good agreement with experiment for many unsaturated carbocations.[2h]

2.2. Applying the Best Methodology to 34dmp2H$^+$ and 24dm3ep5H$^+$

From the extensive study of 24dmp5H$^+$ we have concluded that the best approach is (1) to optimize the geometry at the B3LYP/cc-pVDZ level and (2) to calculate the shieldings at this optimized geometry at the MP2/pcSseg-1 level with the simple X+PCM model for the ^{13}C chemical shifts and with the X+1S+PCM model for the ^{1}H chemical shifts. In the following, we are going to refer to this approach as'MP2' and compare it with the more standard approach denoted 'B3LYP', which implies to calculate the shieldings at the B3LYP/6-311++G(2d,p) level. Both approaches are further tested for the other two protonated alkylpyrroles, 34dmp2H$^+$ and 24dm3ep5H$^+$.

The results for the ^{13}C chemical shifts are shown in Tables 7, 8 and S13, S14 and are illustrated in Figure 9 (a complete set of plots is available in the SI as Figures S47–S50), while the corresponding data for the ^{1}H chemical shifts are in Tables 9, 10 and S15, S16 and are illustrated in Figure 10 (a complete set of plots is available in the SI as Figures S51–S54).

Judged by the statistical data in Tables 7 and 8 it is clear that the MP2 approach performs significantly better than the B3LYP approach for the ^{13}C chemical shifts. Again as for 24dmp5H$^+$ the smallest deviations and best slope and R^2 values are for the simple X+PCM model. It is though interesting to note that the largest deviations are not always for the same carbon atoms. They vary from molecule to molecule and

Table 7. ^{13}C NMR statistical data for 34dmp2H$^+$ fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method and model[b]	R^2	*SD*	*MaxDev*	Slope
B3LYP				
X+PCM	0.9978	3.3	6.0 (C-3)	−0.9404
X+1S+PCM	0.9985	2.4	4.6 (C-3)	−0.9516
HF				
X+PCM	0.9902	6.8	11.4 (C-4)	−0.8799
X+1S+PCM	0.9916	6.3	10.1 (C-4)	−0.8922
MP2				
X+PCM	0.9996	1.3	2.3 (C-4)	−0.9991
X+1S+PCM	0.9993	1.9	3.6 (C-4)	−1.0126

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 and MP2/pcSseg-1. [b] The abbreviations are explained in the text.

Table 8. ^{13}C NMR statistical data for 24dm3ep5H$^+$ fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method and model[b]	R^2	*SD*	*MaxDev*	Slope
B3LYP				
X+PCM	0.9986	2.7	5.6 (C-4)	−0.9571
X+1S+PCM	0.9990	2.3	4.5 (C-4)	−0.9632
HF				
X+PCM	0.9911	6.7	12.6 (C-3)	−0.9046
X+1S+PCM	0.9922	6.3	11.3 (C-3)	−0.9123
MP2				
X+PCM	0.9998	1.1	1.9 (3-CH_2)	−1.0069
X+1S+PCM	0.9997	1.3	2.6 (C-3)	−1.0156

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. [b] The abbreviations are explained in the text.

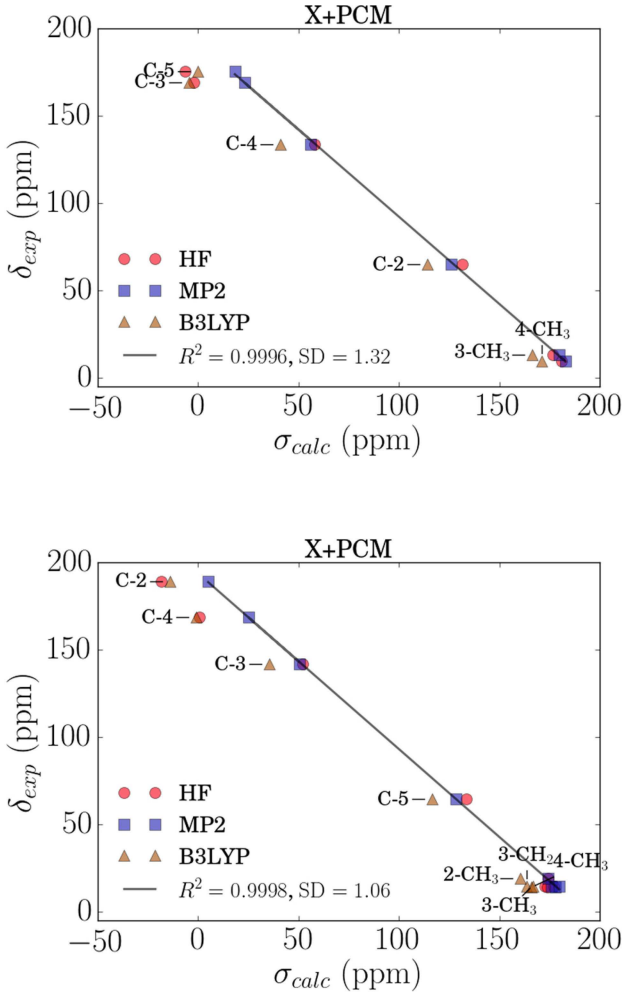


Figure 9. Experimental ^{13}C NMR chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} at the model 'X+PCM' for 34dmp2H$^+$.(top) and 24dm3ep5H$^+$ (bottom). The regression lines and statistical data are for MP2.

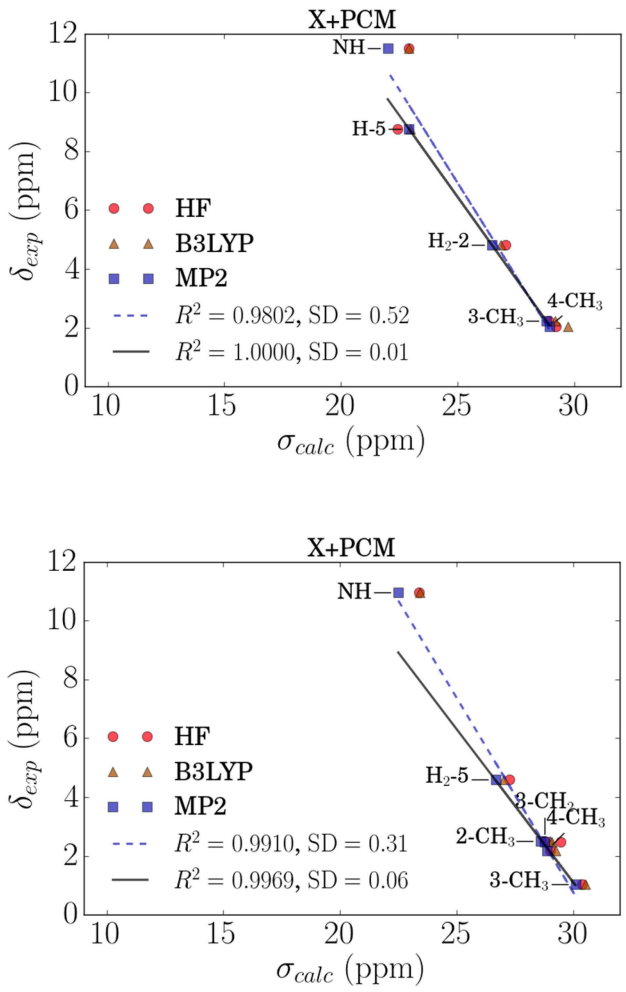


Figure 10. Experimental ^1H NMR chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} at the model 'X+PCM' for 34dmp5H$^+$ (top) and 24dm3ep5H$^+$ (bottom). The regression lines are for MP2 excluding (full line) or including (dotted line) the data point for the NH proton.

method to method. From Figure 9 and the absolute shielding values in Tables S13 and S14 we can however see a very systematic behavior. HF underestimates the shieldings of the two carbon atoms with partial positive charges, C-2 and C-4 in 5H$^+$ pyrroles and C-3 and C-5 in the 2H$^+$ pyrroles, by ~25 ppm while B3LYP also underestimates the shielding of the carbon next to the protonated carbon by ~25 ppm but the shielding of the carbon opposite to the protonated carbon by "only" ~20 ppm. For all the other carbons B3LYP underestimates the shieldings by 10 to 15 ppm, while the HF values differ by −5 to +5 ppm from the MP2 values. This confirms our previous conclusion that B3LYP for these systems is not able to deliver electron correlation, where there is need for a lot of it, and predicts electron correlation effects, where there is actually only little as judged by the difference between MP2 and Hartree-Fock.

The differences in the ^1H statistics between the different methods, in Tables 9 and 10, are again smaller, but Figure 10 and the corresponding Figures S52 and S54 for the X+1S+

Table 9. ^1H NMR statistical data for 34dmp2H$^+$ fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method and model[b]	NH[c]	R^2	*SD*	*MaxDev*	Slope
B3LYP					
X+PCM	yes	0.9439	0.9	1.4 (NH)	−1.2321
	no	0.9974	0.1	0.2 (3-CH_3)	−1.0178
X+1S+PCM	yes	0.9954	0.3	0.4 (H-5)	−0.9720
	no	0.9989	0.1	0.2 (H_2-2)	−1.0545
HF					
X+PCM	yes	0.9182	1.1	1.7 (NH)	−1.2272
	no	0.9882	0.3	0.5 (H_2-5)	−0.9937
X+1S+PCM	yes	0.9924	0.3	0.6 (H_2-5)	−0.9652
	no	0.9851	0.3	0.6 (H_2-5)	−1.0084
MP2					
X+PCM	yes	0.9802	0.5	0.8 (H-5)	−1.2651
	no	1.0000	<0.1	<0.1 (H_2-5)	−1.1077
X+1S+PCM	yes	0.9861	0.4	0.7 (H-5)	−0.9621
	no	0.9997	<0.1	0.1 (H_2-5)	−1.1313

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. [b] The abbreviations are explained in the text. [c] Including (yes) or excluding (no) the data point for the NH proton. See the text.

Table 10. ^{1}H NMR statistical data for 24dm3ep5H^{+} fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method and model[b]	NH[c]	R^2	*SD*	*MaxDev*	Slope
B3LYP					
X+PCM	yes	0.9808	0.5	0.8 (H_2-5)	−1.4179
	no	0.9978	0.1	0.1 (4-CH_3)	−1.0215
X+1S+PCM	yes	0.9986	0.1	0.3 (H_2-5)	−0.9663
	no	0.9974	0.1	0.1 (4-CH_3)	−1.0742
HF					
X+PCM	yes	0.9823	0.4	0.6 (3-CH_3)	−1.4357
	no	0.9386	0.3	0.4 (3-CH_3)	−1.0892
X+1S+PCM	yes	0.9921	0.3	0.4 (2-CH_3)	−0.9470
	no	0.9397	0.3	0.5 (2-CH_3)	−1.0911
MP2					
X+PCM	yes	0.9910	0.3	0.6 (H_2-5)	−1.3205
	no	0.9969	0.1	0.1 (3-CH_3)	−1.0403
X+1S+PCM	yes	0.9977	0.2	0.3 (H_2-5)	−0.9363
	no	0.9976	0.1	0.1 (4-CH_3)	−1.0865

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. [b] The abbreviations are explained in the text. [c] Including (yes) or excluding (no) the data point for the NH proton. See the text.

Table 11. ^{13}C NMR statistical data for 24dmp fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method	R^2	*SD*	*MaxDev*	Slope
B3LYP	0.9993	1.3	2.3 (C-4)	−0.9646
HF	0.9977	2.4	4.1 (C-3)	−0.9373
MP2	0.9996	1.0	1.8 (C-2)	−1.0181

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. Model: X+PCM.

Table 12. ^{13}C NMR statistical data for 34dmp fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method	R^2	*SD*	*MaxDev*	Slope
B3LYP	0.9988	1.7	2.2 (C-2)	−0.9668
HF	0.9987	1.8	2.3 (C-3)	−0.9470
MP2	0.9999	0.5	0.6 (C-2)	−1.0164

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. Model: X+PCM.

Table 13. ^{13}C NMR statistical data for 24dm3ep fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method	R^2	*SD*	*MaxDev*	Slope
B3LYP	0.9994	1.3	2.4 (C-5)	−0.9632
HF	0.9982	2.3	3.9 (C-3)	−0.9428
MP2	0.9996	1.0	1.5 (C-3)	−1.0152

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. Model: X+PCM.

PCM model show that the NH proton chemical shift is also for these two protonated pyrroles the problem.

Adding the explicit solvent molecule again improves drastically the results for B3LYP and MP2, while MP2 gives already with the X+PCM model more consistent values for this proton. For all three protonated alkylpyrroles we find thus with the X+1S+PCM solvent model B3LYP gives marginally better statistics than MP2, while the opposite is the case for the simpler X+PCM model. If the NH proton is, however, of no interest then both B3LYP and MP2 with the X+PCM model are the best and more or less equally good.

The conclusions for the protonated alkylpyrroles are thus quite clear: treatment of electron correlation at the MP2 level or better in combination with the simple PCM model is necessary for the ^{13}C chemical shifts, while for the ^{1}H chemical shifts B3LYP or MP2 will do. However, it is necessary to include an explicit solvent molecule hydrogen bonded to the NH proton, if one wants to reproduce also the NH ^{1}H chemical shift. The question that now remains is whether these conclusions are special for the protonated forms or whether they also hold for the neutral alkylpyrroles. This will be discussed in the following section.

2.3. Applying the Best Methodology to Neutral Pyrroles

The statistical data for the three neutral pyrroles are collected in Tables 11–13 for the ^{13}C shifts and the absolute shieldings are shown in Tables S17–S19 and illustrated for 24dm3ep in Figure 11. A full set of figures are reproduced in the Supporting Information (Figures S55–S56).

For the neutral pyrroles we have only used the simple PCM solvation model as the specific solvation effects in the previous sections turned out to be only really necessary for the NH ^{1}H chemical shift and might be expected to be less important for the neutral forms. The statistics clearly show that the MP2

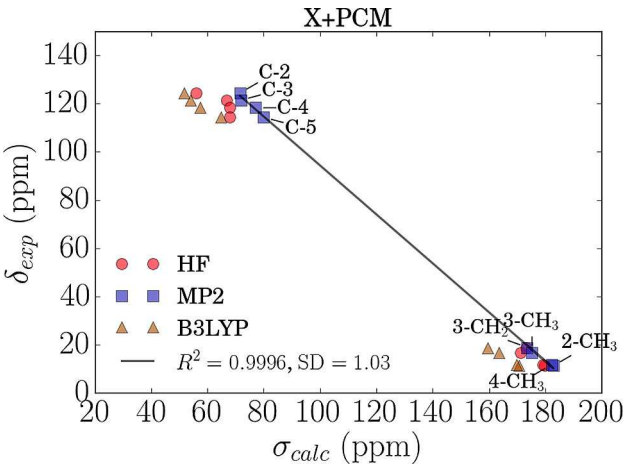


Figure 11. Experimental ^{13}C NMR chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} at the model 'X+PCM' for 24dm3ep. The regression lines and statistical data are for MP2.

approach is also for the neutral pyrroles performing better than B3LYP. Consulting the absolute shieldings in Tables S17–S19 one can see a reminiscence of the non-consistent difference between the B3LYP and MP2 results from the protonated pyrroles. For the C-2 and C-4 carbons in 24 dm and 24dm3ep and for C-3/C-4 in 34dmp the B3LYP results deviate from the MP2 results still by ~20 ppm while it is 12 to 15 ppm for the other carbon atoms.

In Table 14 and Figure 12 we compare then the calculated ^{13}C chemical shifts, Equation (1), for all molecules with the corresponding experimental values for the X+PCM/B3LYP/6-311++G(2d,p) and X+PCM/MP2/pcSseg-1 models. Additional plots in which the six compounds are separated in two groups, protonated and neutrals, have been included in the supplementary material as Figures S58–S59 (protonated) and S60–S61 (neutrals). All the figures clearly show again that B3LYP has in contrast to MP2 problems reproducing the deshielded carbon chemical shifts. In addition, it underestimates the MP2 chemical shifts on average by ~20 ppm. On the other hand, the slope for the MP2 results in Figure 12 is with −1.0056 almost perfect and would allow translating the MP2 absolute shieldings directly without scaling to chemical shifts by using 194 ppm as an internally consistent reference shielding for TMS.

In the previous sections we have several times hinted at a connection between the electron correlation contributions to the shielding of a carbon atom and its partial charge. In Figure 13 we have thus collected these data for all the pyrroles, where the electron correlation contribution is given as difference between MP2 and HF values and as charges we used the HF charges for the sake of simplicity. Although there is no linear correlation for all carbon atoms, there is quite clearly a trend for all the aromatic carbon atoms in the neutral pyrroles towards larger positive correlation contributions for more positive (or less negative) partial charges. For the protonated pyrroles one observes more a black and white picture with large positive correlation contributions for carbon atoms with positive HF charges and small negative correlation contributions for the carbon atoms with the negative HF charges confirming our previous statements.

Finally, the ^{1}H statistical data for the neutral pyrroles are presented in Tables 15–17. The absolute shieldings are summar-

Table 14. Statistical data for the ^{13}C NMR results for all molecules and cations studied comparing the effect of using B3LYP and MP2 calculations.[a]

	R^2	*SD*	*MaxDev*[b]	Slope
B3LYP	0.9982	2.6	8.0 (C-4)	−0.9525
MP2	0.9994	1.6	3.7 (C-5)	−1.0056

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), MP2/pcSseg-1. Model: X+PCM.[b] 34dmp2H$^+$.

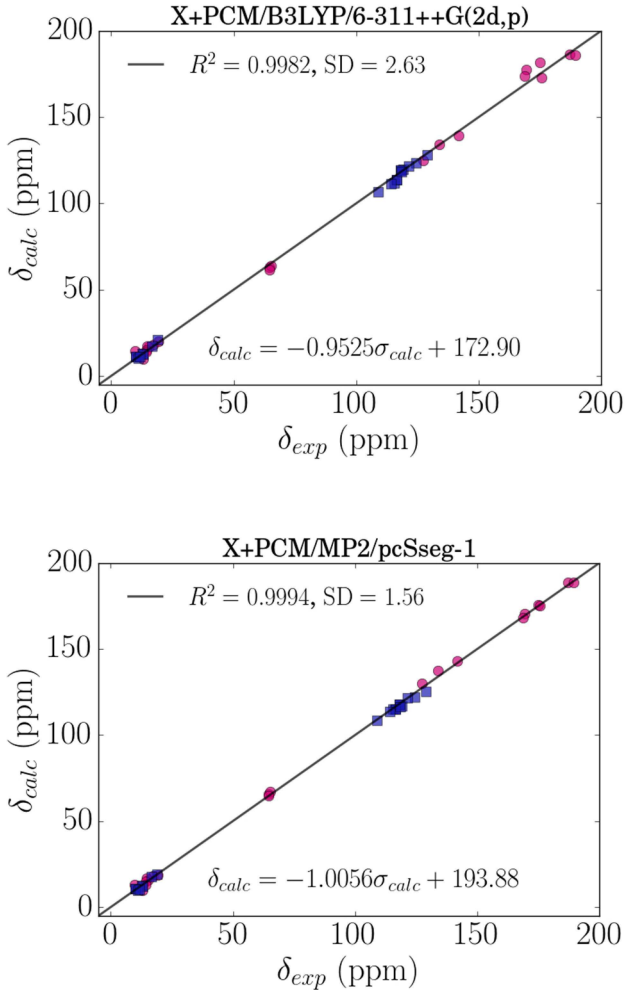


Figure 12. Calculated chemical shifts δ_{calc} versus the experimental ^{13}C NMR chemical shifts δ_{exp} for all molecules in this study with the X+PCM/B3LYP/6-311++G(2d,p) (top) and the X+PCM/MP2/pcSseg-1 (bottom) approaches. Red circles for the protonated pyrroles and blue squares for the neutral pyrroles.

Table 15. ^{1}H NMR statistical data for 24dmp fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method	NH[b]	R^2	*SD*	*MaxDev*	Slope
B3LYP	yes	0.9735	0.4	0.7 (NH)	−1.0876
	no	0.9990	0.1	0.1 (H-3)	−0.9599
HF	yes	0.9591	0.5	0.8 (NH)	−1.0692
	no	0.9998	<0.1	<0.1 (H-3)	−0.9242
MP2	yes	0.9808	0.3	0.5 (NH)	−1.0667
	no	0.9991	0.1	0.1 (H-3)	−0.9550

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. Model: X+PCM. [b] Including (yes) or excluding (no) the data point for the NH proton. See the text.

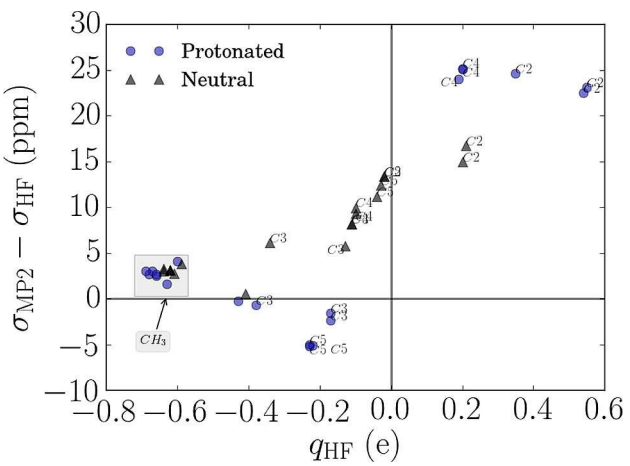


Figure 13. Correlation contributions at the MP2 level to the ^{13}C absolute shieldings for all molecules in this study versus the partial charges on the corresponding atoms calculated with the Natural Population Analysis approach at the Hartree-Fock/pcSseg-1 level.

Table 16. ^{1}H NMR statistical data for 34dmp fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method	NH	R^2	*SD*	*MaxDev*	Slope
B3LYP	yes	0.9794	0.4	0.6 (NH)	−1.0533
	no	1.0000	0	–	−0.9544
HF	yes	0.9549	0.6	1.0 (NH)	−1.0449
	no	1.0000	0	–	−0.9195
MP2	yes	0.9871	0.3	0.5 (NH)	−1.0759
	no	1.0000	0	-	−0.9910

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. Model: X+PCM. [b] Including (yes) or excluding (no) the data point for the NH proton. See the text.

Table 17. ^{1}H NMR statistical data for 24dm3ep fitted to Equation (1) comparing the effect of using B3LYP, HF and MP2 calculations.[a]

Method	NH	R^2	*SD*	*MaxDev*	Slope
B3LYP	yes	0.9993	0.1	0.1 (H-5)	−0.9550
	no	0.9999	<0.1	<0.1 (4-CH_3)	−0.9321
HF	yes	0.9954	0.2	0.3 (H-5)	−0.9513
	no	0.9975	0.1	0.2 (3-CH_2)	−0.9018
MP2	yes	0.9996	0.1	0.1 (3-CH_2)	−0.9531
	no	0.9995	<0.1	0.1 (3-CH_2)	−0.9670

[a] Opt: B3LYP/cc-pVDZ. NMR: B3LYP/6-311++G(2d,p), HF/pcSseg-1 or MP2/pcSseg-1. Model: X+PCM. [b] Including (yes) or excluding (no) the data point for the NH proton. See the text.

ized in Tables S20–S22 and illustrated in Figure 14 and in Figures S62–S64.

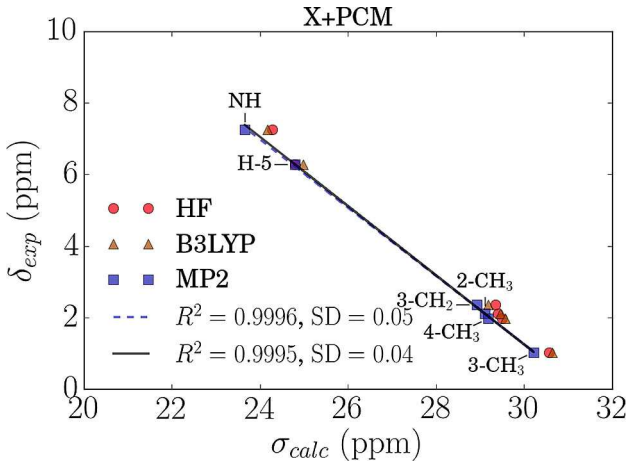


Figure 14. Experimental ^{1}H NMR chemical shifts δ_{exp} versus the calculated shieldings σ_{calc} at the model 'X+PCM' for 24dm3ep. The regression lines and statistical data are for MP2 excluding (full line) or including (dotted line) the data point for the NH proton.

Studying the Figures S62-S64 in more detail one observes that for 24dmp and 34dmp the NH proton is still a problem for the X+PCM model, while it fits nicely on the line for the MP2 results for 24dm3ep. Without the NH ^{1}H chemical shifts there is not much difference between the B3LYP and MP2 values, while there is a small advantage of the MP2 calculations on inclusion of the NH proton in the fitting.

3. Conclusions

For three α-protonated alkylpyrroles and their neutral precursors we have measured their chemical shifts and thoroughly investigated which computational protocol will be able to reproduce all the chemical shifts to an accuracy sufficient for identification of the compounds in mixtures. In particular, we have investigated several solvent models, the influence of the basis set and electron correlation effects.

For the ^{1}H chemical shifts we observed that both B3LYP with the 6-311++G(2d,p) basis set and MP2 with the specialized pcSseg-1 basis set in combination with the PCM solvation model will give good agreement with experimental values with the exception of the NH proton chemical shift. In order to reproduce also this, it is necessary to include an explicit solvent molecule hydrogen bonded to the NH group in the quantum chemical calculation. Inclusion of the counter-ion on the other hand deteriorates the agreement.

For the ^{13}C chemical shifts quite a different conclusion must be drawn. Treatment of solvation effects at the B3LYP level by PCM with or without inclusion of an explicit solvent molecule or the counter-ion improves somewhat the agreement with experiment, but is not sufficient for a consistently good agreement. We find that B3LYP calculations are not able to consistently reproduce the chemical shifts of all the carbon atoms in these molecules leading to worse fits and statistics compared to MP2. In particular, the carbon atoms carrying partial positive charges appear to be troublesome for B3LYP. Their shieldings are affected by large correlation effects in the order of 25 ppm, which B3LYP cannot reproduce. MP2, on the other hand, appears to recover enough electron correlation to give a consistent description of the shieldings of all the carbon atoms in these compounds. It leads thus to an almost perfect linear correlation with the experimental values. Our results confirm in this way, also for the protonated alkylpyrroles, that accurate calculations of ^{13}C chemical shifts in unsaturated carbocations require correlated wave function methods.[2h] Of course, MP2 calculations require more computer power than B3LYP calculations; still with standard computer facilities the MP2 approach is feasible for neutrals or cations including at least 20 carbon/nitrogens covering a large number of interesting species.

Acknowledgements

E.G.L. thanks CNPq and the Science without Borders program for a postdoc fellowship, 206886/2014-4. K.C. thanks CNPq for the grant 30437/2014-9 and CAPES for the grant 23038.004630/2014-35. S.P.A.S. thanks the Danish Agency for Science, Technology and Innovation for an International Network Programme grant 1370-00075B. Professor Frank Jensen, University of Aarhus, is acknowl-

edged for helpful discussions and preliminary calculations with his polarized consistent basis sets.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: ab initio calculations · density functional calculations · NMR spectroscopy · protonated alkylpyrroles · solvation

[1] a) R. Jain, T. Bally, P. R. Rablen, *J. Org. Chem.* **2009**, *74*, 4017–4023; b) K. W. Wiitala, T. R. Hoye, C. J. Cramer, *J. Chem. Theory Comput.* **2006**, *2*, 1085–1092; c) M. W. Lodewyk, M. R. Siebert, D. J. Tantillo, *Chem. Rev.* **2012**, *112*, 1839–1862; d) D. Xin, C. A. Sader, O. Chaudhary, P.-J. Jones, K. Wagner, C. S. Tautermann, Z. Yang, C. A. Busacca, R. A. Saraceno, K. R. Fandrick, N. C. Gonnella, K. Horspool, G. Hansen, C. H. Senanayake, *J. Org. Chem.* **2017**, *82*, 5135–5145; e) E. Toomsalu, P. Burk, *J. Mol. Model.* **2015**, *21*, 244; f) F. Jensen, *J. Chem. Theory Comput.* **2008**, *4*, 719–727; g) J. R. Cheeseman, G. W. Trucks, T. A. Keith, M. J. Frisch, *J. Chem. Phys.* **1996**, *104*, 5497–5509; h) D. Flaig, M. Maurer, M. Hanni, K. Braunger, L. Kick, M. Thubauville, C. Ochsenfeld, *J. Chem. Theory Comput.* **2014**, *10*, 572–578; i) A. M. Sarotti, S. C. Pellegrinet, *J. Org. Chem.* **2009**, *74*, 7254–7260; j) A. M. Sarotti, S. C. Pellegrinet, *J. Org. Chem.* **2012**, *77*, 6059–6065; k) M. A. Iron, *J. Chem. Theory Comput.* **2017**, *13*, 5798–5819; l) R. Suardiaz, C. Perez, R. Crespo-Otero d. l. V. J. M. Garcia, J. S. Fabian, *J. Chem. Theory Comput.* **2008**, *4*, 448–456.

[2] a) H.-U. Siehl, T. Mueller, J. Gauss, *J. Phys. Org. Chem.* **2003**, *16*, 577–581; b) H.-U. Siehl, T. Mueller, J. Gauss, P. Buzek, P. v. R. Schleyer, *J. Am. Chem. Soc.* **1994**, *116*, 6384–6387; c) J. F. Stanton, J. Gauss, H.-U. Siehl, *Chem. Phys. Lett.* **1996**, *262*, 183–186; d) D. S. Fadeev, I. P. Chuikov, V. I. Mamatyuk, *J. Fluorine Chem.* **2016**, *182*, 53–60; e) M. E. Harding, J. Gauss, P. v. R. Schleyer, *J. Phys. Chem. A* **2011**, *115*, 2340–2344; f) V. A. Semenov, D. O. Samultsev, L. B. Krivdin, *Magn. Reson. Chem.* **2018**, *56*, 727–739; g) H.-U. Siehl, S. Brixner, *J. Phys. Org. Chem.* **2004**, *17*, 1039–1045; h) H.-U. Siehl, *Adv. Phys. Org. Chem.* **2007**, *42*, 125–165; i) D. S. Fadeev, I. P. Chuikov, V. M. Karpov, V. I. Mamatyuk, *J. Fluorine Chem.* **2017**, *197*, 1–5.

[3] a) C. P. Andrieux, P. Audebert, P. Hapiot, J. M. Saveant, *J. Phys. Chem.* **1991**, *95*, 10158–10164; b) P. Audebert, P. Hapiot, *Synth. Met.* **1995**, *75*, 95–102; c) L. Guyard, P. Hapiot, P. Neta, *J. Phys. Chem. B* **1997**, *101*, 5698–5706; d) L. Guyard, P. Audebert, P. Hapiot, *J. Chim. Phys. Phys.-Chim. Biol.* **1998**, *95*, 1192–1195; e) M. Zhou, J. Heinze, *Electrochim. Acta* **1999**, *44*, 1733–1748; f) M. Zhou, J. Heinze, *J. Phys. Chem. B* **1999**, *103*, 8443–8450; g) M. Zhou, J. Heinze, *J. Phys. Chem. B* **1999**, *103*, 8451–8457; h) M. Zhou, V. Rang, J. Heinze, *Acta Chem. Scand.* **1999**, *53*, 1059–1062; i) M. Zhou, M. Pagels, B. Geschke, J. Heinze, *J. Phys. Chem. B* **2002**, *106*, 10065–10073; j) O. Hammerich, R. M. Henriksen, F. S. Kamounah, *Proc. Electrochem. Soc.* **2000**, *2000–15*, 9–12; k) G. Sabouraud, S. Sadki, N. Brodie, *Chem. Soc. Rev.* **2000**, *29*, 283–293; l) P. Audebert, F. Miomandre, S. G. Di Magno, V. V. Smirnov, P. Hapiot, *Chem. Mater.* **2000**, *12*, 2025–2030; m) G. H. Hansen, R. M. Henriksen, F. S. Kamounah, T. Lund, O. Hammerich, *Electrochim. Acta* **2005**, *50*, 4936–4955; n) C. P. Andrieux, P. Hapiot, P. Audebert, L. Guyard, M. N. D. An, L. Groenendaal, E. W. Meijer, *Chem. Mater.* **1997**, *9*, 723–729.

[4] a) G. F. Smith, *Adv. Heterocycl. Chem.* **1963**, *2*, 287–309; b) R. Bonnett, I. A. D. Gale, G. F. Stephenson, *J. Chem. Soc.* **1965**, 1518–1519; c) C. O. Bender, R. Bonnett, *J. Chem. Soc. C* **1968**, 2526–2528.

[5] a) R. J. Abraham, E. Bullock, S. S. Mitra, *Can. J. Chem.* **1959**, *37*, 1859–1869; b) Y. Chiang, E. B. Whipple, *J. Am. Chem. Soc.* **1963**, *85*, 2763–2767.

[6] C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry. 2nd Ed*, VCH Verlagsgesellschaft, **1988**.

[7] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[8] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618–622.

[9] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09, Rev. A.1*, Gaussian, Inc., Wallingford, CT, USA, **2009**.

[10] a) https://bse.pnl.gov; b) D. Feller, *J. Comput. Chem.* **1996**, *17*, 1571–1586; c) K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, T. L. Windus, *J. Chem. Inf. Model.* **2007**, *47*, 1045–1052.

[11] a) E. Cances, B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3032–3041; b) E. Cances, B. Mennucci, *J. Math. Chem.* **1998**, *23*, 309–326; c) G. Scalmani, M. J. Frisch, *J. Chem. Phys.* **2010**, *132*, 114110.

[12] K. Coutinho, H. C. Georg, T. L. Fonseca, V. Ludwig, S. Canuto, *Chem. Phys. Lett.* **2007**, *437*, 148–152.

[13] R. D. Wigglesworth, W. T. Raynes, S. Kirpekar, J. Oddershede, S. P. A. Sauer, *J. Chem. Phys.* **2000**, *112*, 736–746.

[14] R. Faber, J. Kaminsky, S. P. A. Sauer, in *Gas Phase NMR* (Eds.: K. Jackowski, M. Jaszuński), Royal Society of Chemistry, **2016**, pp. 218–266.

[15] K. Coutinho, S. Canuto, DICE: A Monte Carlo program for molecular liquid simulation, 2.9 ed., Universidade de São Paulo, São Paulo, **2003**.

[16] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

[17] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, *11*, 361–373.

[18] M. L. P. Price, D. Ostrovsky, W. L. Jorgensen, *J. Comput. Chem.* **2001**, *22*, 1340–1352.

[19] http://www.ddbst.com/en/EED/PCP/DEN_C3.php.

[20] L. Hemmingsen, L. Olsen, J. Antony, S. P. A. Sauer, *J. Biol. Inorg. Chem.* **2004**, *9*, 591–599.

[21] a) A. N. Troganis, E. Sicilia, K. Barbarossou, I. P. Gerothanassis, N. Russo, *J. Phys. Chem. A* **2005**, *109*, 11878–11884; b) D. Kubica, S. Molchanov, A. Gryff-Keller, *J. Phys. Chem. A* **2017**, *121*, 1841–1848; c) S. Molchanov, A. Gryff-Keller, *J. Phys. Chem. A* **2017**, *121*, 9645–9653; d) M. C. Caputo, P. F. Provasi, S. P. A. Sauer, *Theor. Chem. Acc.* **2018**, *137*:88.

[22] a) T. Kupka, M. Stachow, M. Nieradka, J. Kaminsky, T. Pluta, *J. Chem. Theory Comput.* **2010**, *6*, 1580–1589; b) F. Jensen, *J. Chem. Theory Comput.* **2018**, DOI: 10.1021/acs.jctc.8b00477.

[23] J. Gauss, *Chem. Phys. Lett.* **1992**, *191*, 614–620.

[24] M. Jankowska, T. Kupka, L. Stobinski, R. Faber, E. G. Lacerda, Jr., S. P. A. Sauer, *J. Comput. Chem.* **2016**, *37*, 395–403.

[25] P. Buzek, P. v. R. Schleyer, H. Vancik, Z. Mihalic, J. Gauss, *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 448–451.

Manuscript received: November 15, 2018
Accepted manuscript online: November 19, 2018
Version of record online: December 7, 2018