DOI: 10.1002/chem.200801272

# Computational NMR Spectroscopy of Organoarsenicals and the Natural Polyarsenic Compound Arsenicin A

**Petri Tähtinen,[a, b] Giacomo Saielli,[c] Graziano Guella,[d] Ines Mancini,[d] and Alessandro Bagno*[a]**

**Abstract:** The $^{1}H$ and $^{13}C$ NMR chemical shifts and coupling constants of a series of organoarsenic compounds were calculated with DFT methods and compared with available experimental spectra. We show that non-relativistic methods successfully model the NMR spectra of these molecules; relativistic spin–orbit effects are small but appreciable for $^{13}C$ shifts, and their inclusion is beneficial. Application of the same methods of calculation to the intriguing natural polyarsenic compound arsenicin A allowed several viable alternative structures to be ruled out and thereby confirmed the previously suggested adamantane-like structure of arsenicin A. These results not only reinforce the known predictive power of DFT NMR calculations, but also open the way for the investigation of other naturally occurring molecules with unusual structures outside the scope of empirical methods.

**Keywords:** arsenic · density functional calculations · natural products · NMR spectroscopy · structure elucidation

## Introduction

Arsenic is the epitome of toxicity and, indeed, all arsenic compounds are toxic to some degree. Nevertheless, some arsenicals have been used since more than a century for the treatment of certain infections such as syphilis, in the form of the drug Salvarsan, which is well known although poorly understood at the structural level.[1] In the antibiotics era, arsenicals have only played a specialized role, for instance, as antiparasitic agents in the treatment of sleeping sickness.[2] We are now witnessing a revival of arsenicals in the chemotherapy of cancer,[3] for example, the recent approval of $As_2O_3$ (arsenic trioxide, ATO) for the treatment of promyelocytic leukaemia[4] and demonstration of its inhibitory activity towards thioredoxin reductase.[5]

Less well known by the public are the naturally occurring organoarsenic compounds. They are also found in various foodstuffs, particularly seafood.[6,7] Research in the last 30 years has identified over 100 naturally occuring organoarsenicals, and has therefore raised human health concerns about the safety of such foods. Increasingly, research on such issues involves metabolic investigations, because the original, perhaps non-toxic, organoarsenical(s) present in the food may be metabolized to other arsenicals with currently unknown toxic effects.[8] Inorganic arsenic compounds are alkylated by marine organisms to give mono-arsenic compounds, which are present in brown algae, molluscs, arthropods, and vertebrates.[9] Besides volatile compounds like alkyl arsines,[9e] more commonly found are non-volatile methyl arsine oxides, methylarsonic acid, and dimethylarsinic acid (cacodylic acid), alongside water-soluble betaines, cholines, and derivatives of carbohydrates, lipids, and amino acids.[9–11] Sulfur derivatives such as thio-organoarsenates are rarely encountered,[12] while thioarsenic acid is known to be derived from the metabolism of cod-liver arsenolipids in humans.[13] The presence of arsenic in drinking water is a major problem in certain countries, and efforts are directed at removing it, especially $As^{III}$, with chelating agents.[14]

[a] Dr. P. Tähtinen, Prof. A. Bagno
Dipartimento di Scienze Chimiche
Università di Padova
via Marzolo 1, 35131 Padova (Italy)
Fax: (+39) 0498275239
E-mail: alessandro.bagno@unipd.it

[b] Dr. P. Tähtinen
Department of Chemistry, University of Turku
Vatselankatu 2, 20014 Turku (Finland)

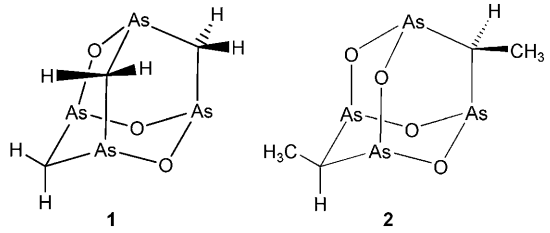
[c] Dr. G. Saielli
Istituto CNR per la Tecnologia delle Membrane, Sezione di Padova
via Marzolo 1, 35131 Padova (Italy)

[d] Prof. G. Guella, Prof. I. Mancini
Laboratorio di Chimica Bioorganica
Dipartimento di Fisica, Università di Trento
via Sommarive 14, 38050 Povo-Trento (Italy)

Recently, some of us isolated the first natural polyarsenical, that is, arsenicin A from the New Caledonian sponge *Echinochalina bargibanti*.[15] Elucidating the structure of this substance with empirical formula $C_3H_6As_4O_3$, isolated in very small quantities, met with substantial difficulties owing to a combination of factors: 1) a deceptive fragmentation pattern indicating loss of $H_2CO$ and formation of a $[As_4]^+$ ion-fragment at $m/z$ 300 in the electron impact (EI) mass spectrum; 2) extremely simple and uninformative $^1H$ and $^{13}C$ NMR spectra; 3) a lack of reference data for the assignment of such spectra; 4) NMR measurements on the most obvious probe, the $^{75}As$ nucleus ($I=3/2$, 100% natural abundance), are not feasible because three-coordinate $As^{III}$ compounds have a large electric field gradient at the arsenic center, which couples with the high quadrupole moment of $^{75}As$ and leads to extremely broad lines.[16]

Nevertheless, arsenicin A was proposed to have the intriguing adamantane-like structure **1** by comparison of its experimental and calculated vibrational spectra with those of the known tetraarsenic compound **2**, whose synthesis and crystal structure were available[17] (Scheme 1). However, although all the $^1H$ and $^{13}C$ NMR resonances were reported,[15] a thorough understanding of its NMR spectra could not be achieved.




Scheme 1. Reported structures of arsenicin A (**1**) and synthetic tetraarsenic compound **2**.

Many empirical approaches have been developed for the prediction of NMR spectra.[18] Such methods are based on database lookup, or on other knowledge-based methods such as neural networks, and usually perform well in the case of naturally occurring molecules with common functional groups. However, none of these can be expected to yield reliable results for molecules which fall outside the scope of the database used for prediction or training. Instances in which this may happen include, for example, strained or sterically hindered molecules, and molecules with unusual functionalities.

One such prominent case is indeed provided by organoarsenicals, for which a relatively small knowledge base is available. Thus, one can hardly turn to empirical estimates for molecules like **1** and **2**, which do not bear any significant functional groups analogous with typical organic molecules.

On the other hand, the prediction of NMR parameters by quantum chemical, and especially DFT, methods has been established as a reliable tool to obtain accurate estimates of all relevant molecular properties (nuclear shieldings and coupling constants).[19] Thus, DFT methods have proved effective at predicting the NMR spectral patterns of a wide variety of organic, organometallic, and inorganic species, ranging from natural substances with intricate connectivities based on a carbon framework,[20] to inorganic complexes, in which connectivities are based on a central metal atom,[21] and the complicated networks of polyoxotungstates,[22] whose NMR properties are strongly affected by relativistic effects. Current capabilities for predicting NMR spectra therefore virtually encompass the whole Periodic Table. Computations on arsenic clusters and rings have revealed a variety of fascinating motifs.[23,24]

Hence, we believe the computational elucidation of the structure of arsenicin A (**1**) to be a valuable contribution to understanding the chemistry and NMR spectroscopy of an emerging class of intriguing natural substances such as organoarsenicals. Arsenicin A is also interesting from a computational viewpoint, since it is a relatively small and nonpolar molecule but contains a set of fairly heavy As atoms. Thus, the NMR spectra of all nuclei may be affected by relativistic spin–orbit effects.

To this end, we have computed the structure and NMR parameters (chemical shifts and coupling constants) of a series of arsenicals, ranging from simple models to the target arsenicin A, by a variety of DFT methods.

## Results and Discussion

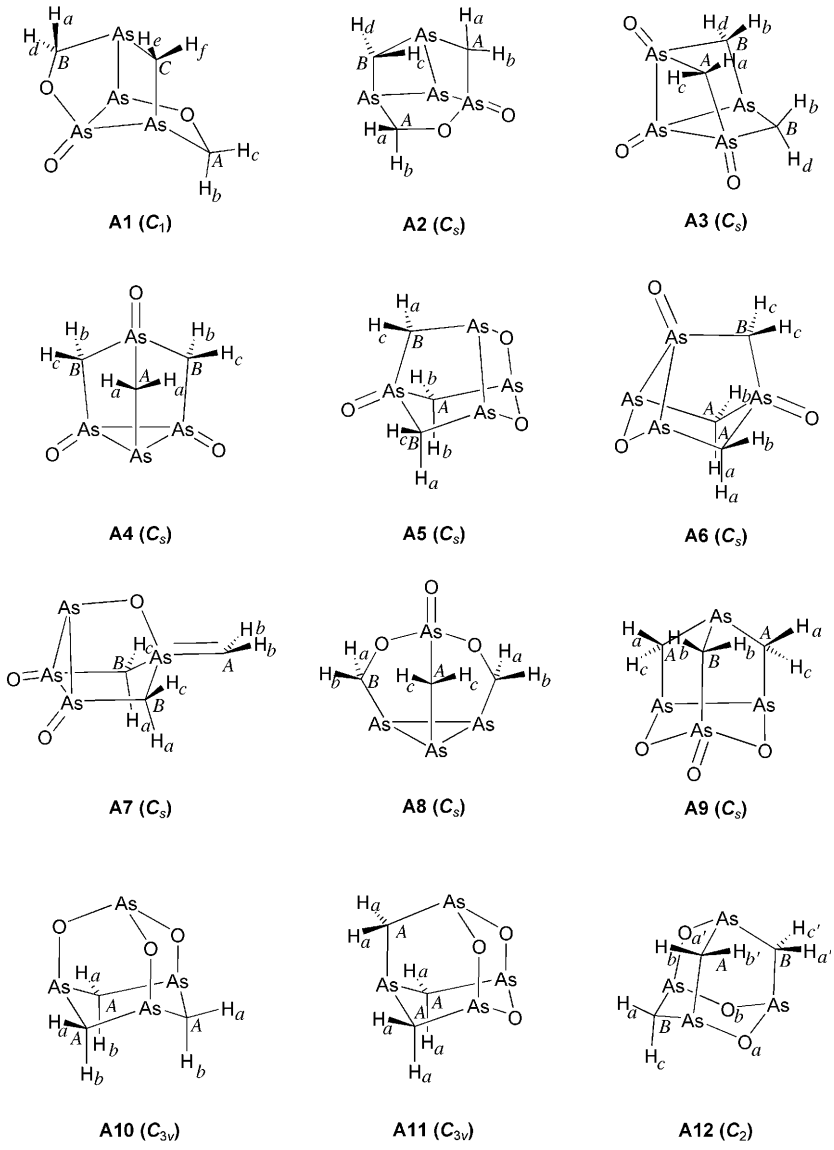
Apart from the arsenicals used for calibration below, we investigated synthetic tetraarsenic compound **2**[17] and various candidate structures for arsenicin A (Scheme 2).

**Calibration**: We have first computed the structure and NMR parameters of a series of organoarsenicals for which $^1H$ and $^{13}C$ NMR data are available or were thought to be relevant: 2-(dimethylarsino)trimethylarsine, $(CH_3)_2AsCH_2As(CH_3)_2$ (**3**);[25] trimethylarsine ylide, $(CH_3)_3As{=}CH_2$ (**4**);[26,27] trimethylarsine, $(CH_3)_3As$ (**5**);[28] triphenylarsine, $Ph_3As$ (**6**);[28] arsabenzene, $C_5H_5As$ (**7**);[29] dimethylarsino dimethyldithioarsinate, $(CH_3)_2As(S){-}S{-}As(CH_3)_2$ (**8**).[30]

Basis-set convergence was tested on **8**, to establish the effects of increased valence splitting and additional polarization and diffuse functions. Increasing the valence splitting from double-zeta to triple-zeta and the addition of extra d and p functions on heavy atoms and hydrogen atoms, respectively, influenced bond lengths significantly, whereas the influence of additional diffuse functions was minimal. Thus, the 6-311G(2d,2p) basis set was found to be adequate for geometry optimization and used thereafter.

Energetically favorable conformers of **3** and **8** were searched by systematically changing their skeletal torsions. It was found that **3** occupies two degenerate global minima in the potential-energy surface which are enantiomeric rotamers of each other and are 3.7 kcal mol$^{-1}$ more stable than the next lowest minimum-energy conformation. For **8**, four minima were found, of which the two lowest are degenerate




Scheme 2. Candidate structures for arsenicin A (see also Scheme 1). Structures **A4**, **A8, A9, A10**, and **A11** correspond respectively to **2**, **1**, **3**, **5**, and **6** in ref. [15].

enantiomeric rotamers (92.2%) that are 1.3 kcal mol$^{-1}$ (4.7%) and 1.6 kcal mol$^{-1}$ (3.1%) below the two next lowest minima. For each conformer of **8** the chemical shifts were calculated and their averages were weighted according to the mole fractions estimated from Boltzmann populations.

For the optimized geometries of the selected compounds, the chemical shifts were calculated at the non-relativistic B3LYP/cc-pVTZ (Method A) and relativistic BP86-ZSO/TZ2P (Method C) levels. The results are presented in Table 1 and plotted in Figure 1. The $^{1}$H and $^{13}$C shifts are well predicted by DFT methods. For the carbon atoms directly bound to arsenic, spin–orbit shieldings $\sigma_{SO}$ are small but appreciable (up to 8 ppm); nevertheless, the non-relativistic and relativistic methods perform equally well.

**Synthetic tetraarsenic compound 2 ($C_4H_8As_4O_4$)**: For further calibration we considered **2** (Scheme 1), which can be obtained by reaction of $As_2O_3$ with propionic acid/anhydride.[17] Its $C_s$ symmetry results in very simple NMR spectra: two $^{13}$C and two $^{1}$H signals are observed, with a single $^{1}$H,$^{1}$H coupling constant $^{3}J$(H,CH$_3$) of 7.9 Hz (Table 2). Both non-relativistic and relativistic calculations provide the correct ordering of $^{13}$C and $^{1}$H signals (Figure 2), with rather similar mean absolute errors (MAE). Similar performance is expected in view of the very small spin–orbit shieldings, which are less than 1 ppm also for $^{13}$C. Moreover, the calculated coupling constant $^{3}J$(H,CH$_3$) is in good agreement with the experimental value for all methods.

Finally, the electric field gradient at As (computed at the ZORA SO level) results in a nuclear quadrupolar coupling constant (NQCC) of $\chi(^{75}\mathrm{As}) = -240$ MHz with asymmetry parameter $\eta = 0.57$. If these values are inserted into the equation for quadrupolar relaxation,[31] even assuming a relatively fast correlation time of 1–10 ps, a line width of about 80–800 kHz is obtained, which renders any direct $^{75}$As NMR measurement impossible.

It can be concluded that both non-relativistic and relativistic DFT methods provide an accurate framework for predicting the NMR spectra of this unusual arsenic compound.

**Arsenicin A ($C_3H_6As_4O_3$)**: The relatively simple NMR spectra of arsenicin A (first entry in Table 3) suggest the presence of three methylene groups, two of which are magnetically equivalent, and symmetry elements, so that strong constraints are imposed on the structures that can be proposed. However, we relaxed these requirements somewhat in our investigation, to take into account the possibility that signals which are distinct by symmetry may be accidentally isochronous.

Thus, we investigated a series of candidate structures (Scheme 2), many of which contain an $As_4$, $As_3$ or $As_2$ groups, as seemingly indicated by MS results.

Table 1. Experimental and calculated $^{13}C$ and $^{1}H$ chemical shifts for model organoarsenicals.[a]

| | $^{13}C$ | | | | $^{1}H$ | | | |
|---|---|---|---|---|---|---|---|---|
| | Exptl | Method A | Method C | $\sigma_{SO}$ | Exptl | Method A | Method C | $\sigma_{SO}$ |
| $Me_2AsCH_2AsMe_2$ (**3**) | | | | | | | | |
| $CH_3$ | – | 17.8 | 16.3 | 4.0 | 1.1 | 0.86 | 1.02 | −0.08 |
| $CH_2$ | – | 42.9 | 41.1 | 7.6 | 1.6 | 1.14 | 1.69 | −0.21 |
| $Me_3As{=}CH_2$ (**4**) | | | | | | | | |
| $CH_3$ | 15.6 | 19.4 | 20.7 | 2.6 | 0.82 | 1.08 | 1.23 | −0.05 |
| $CH_2$ | 7.6 | 8.4 | 14.4 | 2.1 | -0.19 | 0.06 | 0.58 | −0.06 |
| $Me_3As$ (**5**) | | | | | | | | |
| $CH_3$ | 11.2 | 17.3 | 15.6 | 4.1 | 0.95 | 0.78 | 0.93 | −0.08 |
| $Ph_3As$ (**6**) | | | | | | | | |
| *ipso* | 139.6 | 152.2 | 151.1 | 2.6 | – | – | – | |
| *ortho* | 133.7 | 139.9 | 139.3 | 0.4 | 7.32 | 7.47 | 7.60 | −0.04 |
| *meta* | 128.6 | 133.7 | 133.3 | 0.6 | 7.32 | 7.55 | 7.50 | −0.03 |
| *para* | 128.4 | 133.5 | 132.8 | 0.7 | 7.32 | 7.55 | 7.55 | −0.04 |
| $C_5H_5As$ (**7**) | | | | | | | | |
| *ortho* | 167.7 | 182.8 | 178.1 | 1.5 | 9.68 | 10.16 | 10.31 | −0.17 |
| *meta* | 133.2 | 137.6 | 136.8 | 0.4 | 7.83 | 8.28 | 8.25 | −0.03 |
| *para* | 128.2 | 132.0 | 133.1 | −1.1 | 7.52 | 7.71 | 7.70 | −0.10 |
| $Me_2As(S){-}S{-}AsMe_2$ (**8**) | | | | | | | | |
| $(CH_3)_2AsS_2$ | 27.7 | 31.8 | 34.8 | 2.5 | 2.16 | 1.70 | 1.93 | −0.11 |
| $(CH_3)_2AsS$ | 15.7 | 21.1 | 20.0 | 3.7 | 1.40 | 1.25 | 1.48 | −0.15 |

[a] All shifts in ppm; see text for definition of computational methods. The value of $\sigma_{SO}$ is given only for method C.

Some of these had previously been investigated with regard to their vibrational spectra.[15] Structures **A1**–**A9** contain $CH_2$ groups bound to As, O or As(O) in various patterns; **A7** also features an arsine ylide group. The $C_s$ symmetry of **A2**–**A9** would result in the required number of $^{13}C$ and $^{1}H$ NMR signals; at least one arsine oxide (As=O) functionality is required to accommodate all oxygen atoms. Compound **A1** is chiral ($C_1$); however, no attention was paid to stereochemistry in this exploratory part.

The last three structures (**A10**–**A12**) do not contain any As–As bond. Structures **A10** and **A11** were previously considered and rejected,[15] owing

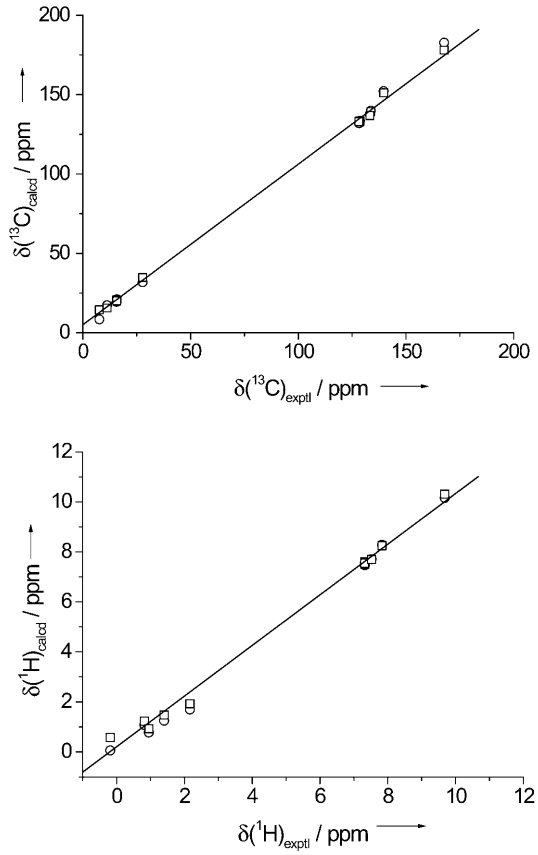



Figure 1. Correlation between experimental and calculated $^{13}C$ and $^{1}H$ chemical shifts for model compounds **3**–**8**. Circles: non-relativistic (method A); squares: relativistic (method C); The linear fits of $\delta_{calcd} = a\delta_{exptl} + b$ have the following parameters: $^{13}C$, Method A: $a = 1.04$, $b = 3.02$ ppm ($R = 0.9988$); $^{13}C$, Method C: $a = 1.01$, $b = 5.04$ ppm ($R = 0.9993$); $^{1}H$, Method A: $a = 1.06$, $b = -0.18$ ppm ($R = 0.9979$); $^{1}H$, Method C: $a = 1.02$, $b = 0.14$ ($R = 0.9973$).

Table 2. Experimental and calculated $^{13}C$ and $^{1}H$ NMR parameters for **2**.[a]

| | Exptl | Method A | Method B | Method C | $\sigma_{SO}$ |
|---|---|---|---|---|---|
| *C*H | 32.32 | 41.64 | 42.90 | 43.76 | 0.58 |
| *C*$H_3$ | 10.29 | 11.48 | 12.07 | 12.88 | 0.31 |
| C*H* | 0.54 | 0.80 | 0.81 | 1.05 | −0.18 |
| C$H_3$ (av) | 1.08 | 1.50 | 1.52 | 1.61 | 0.04 |
| MAE ($^{13}C$) | | 5.25 | 6.18 | 7.01 | |
| MAE ($^{1}H$) | | 0.34 | 0.35 | 0.52 | |
| MAE (all) | | 2.79 | 3.27 | 3.76 | |
| $^{3}J$(H,$CH_3$) | 7.9 | 8.6 | 10.2 | 7.9 | |

[a] Chemical shifts in ppm, coupling constants in Hz. Experimental values from ref. [17]. $MAE = \Sigma_i |\delta_i^{exptl} - \delta_i^{calcd}| / n$. See text for definition of methods; the value of $\sigma_{SO}$ is given only for method C.

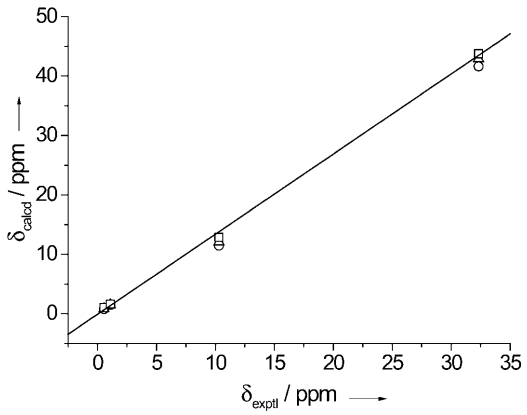



Figure 2. Correlation between experimental and calculated $^{1}H$ and $^{13}C$ chemical shifts for **2**. Circles: non-relativistic (method A); triangles: non-relativistic (method B); squares: relativistic (method C). The linear fits of $\delta_{calcd} = a\delta_{exptl} + b$ have the following parameters: Method A: $a = 1.29$, $b = -0.38$ ppm, $R = 0.9988$; Method B: $a = 1.33$, $b = -0.35$, $R = 0.9991$; Method C: $a = 1.35$, $b = -0.09$, $R = 0.9995$.

Table 3. Experimental and calculated $^{1}H$ and $^{13}C$ chemical shifts for trial structures **A1**–**A12** of arsenicin A, and relative energies.[a]

| | $\delta E$[b] | $H_a$ | $H_b$ | $H_c$ | $H_d$ | $H_e$ | $H_f$ | $C_A$ | $C_B$ | $C_C$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Exptl | | 1.37 | 2.23 | 2.42 | | | | 17.03 | 23.05 | |
| **A1** | 52.1 | 5.21 | 5.13 | 5.05 | 4.66 | 2.24 | 1.73 | 87.4 | 75.8 | 17.8 |
| **A2** | 45.4 | 5.11 | 4.37 | 3.61 | 2.07 | | | 78.6 | 3.94 | |
| **A3** | 99.4 | 5.47 | 4.73 | 4.14 | 3.79 | | | 91.7 | 64.8 | |
| **A4** | 103.8 | 3.26 | 3.10 | 2.89 | | | | 46.1 | 44.9 | |
| **A5** | 22.8 | 2.99 | 1.65 | 1.49 | | | | 49.7 | 44.1 | |
| **A6** | 65.2 | 4.00 | 2.86 | 2.08 | | | | 57.6 | 28.9 | |
| **A7** | 126.2 | 2.92 | 2.83 | 2.68 | | | | 43.6 | 25.3 | |
| **A8** | 53.8 | 6.11 | 5.96 | 0.88 | | | | 80.2 | −0.6 | |
| **A9** | 25.0 | 3.55 | 2.02 | 1.14 | | | | 43.9 | 29.3 | |
| **A10** | (0.0) | 2.99 | 1.33 | | | | | 27.4 | | |
| **A11** | 0.7 | 1.67 | | | | | | 23.4 | | |
| **A12** | 3.0 | 1.04 | 2.08 | 2.19 | | | | 27.0 | 31.0 | |

[a] In ppm by method C. No signal assignment is implied. See text for definition of methods. [b] In kcal mol$^{-1}$ at the B3LYP/6–311G(2d,2p) level, relative to the most stable isomer **A10**.

to their higher $C_{3v}$ symmetry, which would result in too few NMR signals. We included them in our survey, because 1) we can thus better characterize the $C_3H_6As_4O_3$ potential-energy surface and therefore the relative stability of all isomers, and 2) NMR predictions can be made for the organic moiety in these chemical environments, which are closely reminiscent of **1**, **2** (and $As_4O_6$).

Finally, structure **A12** is that proposed for arsenicin A; its $C_2$ symmetry can result in the correct number of NMR signals, and its vibrational spectrum was found to be fairly consistent with experiment. However, it also does not contain any As−As bond.

The structures and energies of **A1**–**A12** were optimized at the B3LYP/6-311G(2d,2p) level, and chemical shifts calculated with method C; the results are reported in Table 3.

A first screening of the candidate structures is provided by relative energies: the most stable isomer is **A10**; **A1**–**A9** lie more than 23 kcal mol$^{-1}$ higher in energy. The arsine ylide **A7** is the least stable, by 126 kcal mol$^{-1}$. Similarly, **A3** and **A4**, which contain strained three- and four-membered rings, are also quite unstable. Adamantane-like structures **A10**–**A12** are the most stable, and lie within 3 kcal mol$^{-1}$ of each other.

Calculated NMR data are presented in Table 3; since no consistent assignment can be made, MAEs are not meaningful and are not given. The results for **A1**–**A9** reinforce the previous conclusion, since a consistently poor match of $^{1}H$ and $^{13}C$ data is found; in most cases the predicted $^{13}C$ chemical shifts ($\delta = 50$–$70$ ppm) do not even lie in the correct range ($\delta \approx 20$ ppm). In conclusion, none of the structures **A1**–**A9** (containing O-$CH_2$-As, As-$CH_2$-As(O) or (O)As-$CH_2$-As(O) moieties) is a good candidate for arsenicin A.

Conversely, for **A10**–**A12** all chemical shifts at least lie in the expected range, which strongly suggests that their chemical environment is similar to that of arsenicin A. All three structures have an $AsCH_2As$ group like in **2** and **3**; however, for the latter, no experimental $^{13}C$ NMR data are available.

These arguments allow the possibilities to be narrowed down to **A10**–**A12**. However, **A10** and **A11** must be ruled out owing to their wrong symmetry. We are then left with $C_2$-symmetric structure **A12** as the sole viable candidate. Its thermodynamic stability is somewhat lower than that of **A10** or **A11**. To further probe this issue we carried out vibrational analyses and ab initio (MP2/cc-pVTZ) calculations for **A10**–**A12**. Inclusion of thermochemical corrections reduces the gap to $\Delta G = 2.2$ kcal mol$^{-1}$ for **A12**, while **A11** becomes slightly more stable than **A10** ($\Delta G = -0.1$ kcal mol$^{-1}$). The MP2 relative energies are rather similar to the DFT values ($\Delta E$ vs. **A10** of 0.5 and 4.2 kcal mol$^{-1}$ for **A11** and **A12**, respectively).

This small energy gap is definitely compatible with a kinetically controlled, irreversible biochemical path leading to a somewhat less stable isomer; for example, unsaturated fatty acids, which occur in Nature only as the less stable *cis* isomers with stability difference of about 1 kcal mol$^{-1}$ per double bond;[33] many other examples of highly strained natural molecules are known.

We then focussed on a more detailed computational characterization of structure **A12**. The calculated chemical shifts for **A12** are in the correct range and ordering with non-relativistic and relativistic methods, and overall MAEs do not exceed 4 ppm (Table 4, Figure 3). The $\sigma_{SO}$ term is small (3 ppm at most) but sizable for $^{13}C$, and including its contribution leads to improved agreement. The quality of the correlation is somewhat lower than for **2**, partly owing to the narrower range spanned. Interestingly, the signal that deviates most is that of $C_A$, which also has the largest spin–orbit shielding (3 ppm). It thus seems that the electronic structure of arsenicin A is rather different from that of **2**, in which $\sigma_{SO}$ does not exceed 0.6 ppm. Similar effects were noticed for the halogen-bound carbon atoms in halobenzenes.[34]

Table 4. Experimental and calculated $^{13}C$ and $^{1}H$ chemical shifts [ppm] for structure **A12**.

| | Exptl | Method A | Method B | Method C | $\sigma_{SO}$ |
|---|---|---|---|---|---|
| $C_A$ | 17.03 | 27.32 | 28.19 | 27.00 | 3.18 |
| $C_B$ | 23.05 | 30.25 | 31.07 | 31.05 | 0.91 |
| $H_a$ | 1.37 | 0.76 | 0.81 | 1.04 | −0.21 |
| $H_b$ | 2.23 | 1.73 | 1.75 | 2.08 | −0.27 |
| $H_c$ | 2.42 | 1.99 | 2.02 | 2.19 | −0.21 |
| MAE ($^{13}C$) | | 8.75 | 9.59 | 8.98 | |
| MAE ($^{1}H$) | | 0.51 | 0.48 | 0.24 | |
| MAE (all) | | 3.81 | 4.13 | 3.74 | |

The $^{75}As$ NQCCs are again very large (between −230 and +200 MHz, $\eta = 0.54$–$0.97$; method C), which definitely rules out $^{75}As$ as a viable NMR probe.

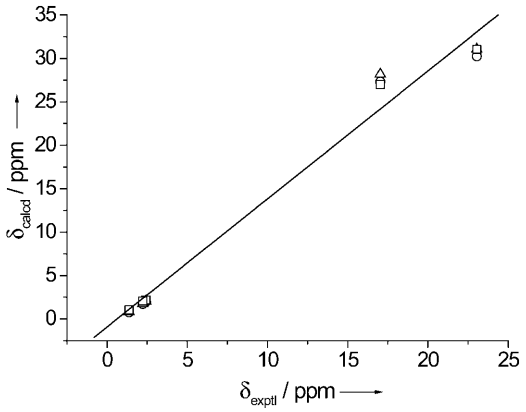



Figure 3. Correlation between experimental and calculated $^{1}H$ and $^{13}C$ chemical shifts for structure **A12** of arsenicin A. The linear fits of $\delta_{calcd} = a\delta_{exptl} + b$ have the following parameters: Method A: $a = 1.47$, $b = -1.12$ ppm, $R = 0.98976$; Method B: $a = 1.51$, $b = -1.16$, $R = 0.98929$; Method C: $a = 1.47$, $b = -0.91$, $R = 0.99317$.

For the sake of completeness we also examined $^{17}O$ spectra: two non-equivalent $^{17}O$ signals are expected by method C at 74 and 123 ppm for $O_a$ and $O_b$, respectively (Scheme 2). Their NQCCs are moderate (ca. −8 MHz, $\eta = 0.5$). Setting the correlation time to 1–10 ps[31] their line widths can be expected to be around 20–200 Hz, which is a viable range. However, the very small quantities isolated place a severe limitation on this technique.

**Couplings in arsenicin A**: Coupling constants for structure **A12** were also calculated. The computations yielded large $^{2}J$(H,H) values for the geminal protons, as expected (Table 5). Calculations with methods A and C yield $^{2}J$ values that are significantly smaller than experiment; conversely, the $^{2}J$ values from method B (based on the accurate pcJ-2 basis set) are larger and in better agreement with the experimental results, which further illustrates the better performance of basis sets constructed according to established criteria (decontraction of core functions etc.). Additionally, several smaller long-range couplings were predicted. Notably, the calculations also yield a fairly large $^{4}J(H_a,H_b)$ value of 1.5 Hz, which was not observed. On the other hand, sizable values of $^{4}J(H_c,H_b)$ and $^{4}J(H_c,H_{b'})$ are predicted by the computations and are in agreement with the experimental results. All calculations predict observable long-range couplings between almost all protons, but in general their magnitude is below the computational accuracy (and below experimental detection limits).

Table 5. Experimental and calculated $J(^{1}H,^{1}H)$ values [Hz] for arsenicin A.[a]

| | Exptl | Method A | Method B | Method C |
|---|---|---|---|---|
| $^{2}J(H_a,H_c)$ | 13.8 | −9.7 | −12.6 | −7.4 |
| $^{2}J(H_b,H_{b'})$ | n.d. | −10.1 | −13.1 | −7.7 |
| $^{4}J(H_a,H_b)$ | n.d. | 1.5 | 1.5 | 0.9 |
| $^{4}J(H_a,H_{b'})$ | n.d. | −0.6 | −0.5 | −0.6 |
| $^{4}J(H_c,H_b)$ | 0.9 | −0.7 | −0.6 | −0.7 |
| $^{4}J(H_c,H_{b'})$ | 0.9 | 0.1 | 0.6 | 0.5 |
| $^{6}J(H_a,H_{c'})$ | n.d. | −0.8 | −0.7 | −0.5 |
| $^{6}J(H_a,H_{a'})$ | n.d. | −0.8 | −0.7 | −0.5 |
| $^{6}J(H_c,H_{c'})$ | n.d. | −0.8 | −0.7 | −0.5 |
| $^{1}J(C,H_a)$ | 132 | 122.7 | 141.3 | 143.9 |
| $^{1}J(C,H_c)$ | 132 | 126.5 | 145.3 | 148.1 |
| $^{1}J(C,H_b)$ | 132 | 124.1 | 142.9 | 145.5 |
| MAE $(J_{H,H})$[b] | | 1.7 | 0.6 | 2.4 |
| MAE $(J_{C,H})$ | | 7.6 | 11.2 | 13.8 |

[a] Experimental data are absolute values; n.d. = not determined. [b] Including only $^{2}J(H_a,H_c)$, $^{4}J(H_c,H_b)$ and $^{4}J(H_a,H_{b'})$.

The overall performance of computed NMR parameters is best appreciated by comparison of $^{1}H$ spectra simulated at 400 MHz with experimental and calculated chemical shifts and couplings (Figure 4). In the simulation we employed the most accurate parameters available, that is, chemical shifts from method C and coupling constants from method B. Apart from a systematic deshielding (ca. 0.4 ppm), the spectrum predicted by DFT matches the experimental one with regard to the relative spacing and spin–spin splitting of the signals.

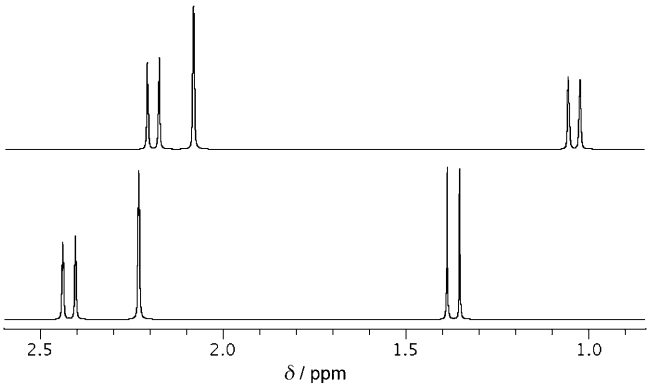



Figure 4. Experimental (bottom) and calculated (top) $^{1}H$ spectra of arsenicin A (**1**) and structure **A12**. Chemical shifts and couplings computed with methods C and B, respectively (see text). Simulations at 400 MHz.

## Conclusion

Computational NMR has confirmed the structure of the unusual marine natural product arsenicin A, structure elucidation of which was hampered by several factors: 1) the simple spin system and spatial arrangement severely limit the connectivity information that can be extracted from its experimental NMR spectra; 2) chemical shifts can hardly be compared to known databases, owing to the lack of such reference data; 3) NMR measurements on the $^{75}As$ nucleus cannot be carried out since its line widths are beyond detection. On the other hand, such systems are ideally suited for the application of DFT calculations, since they are nonpolar and rigid. Thus, among twelve isomers, **A12** (**1**) proved be the only structure with computed NMR parameters fully consistent with experiment. Heavy-atom spin–orbit effects on $^{13}C$ shieldings are relatively small but appreciable (a few ppm) and argue for the use of relativistic DFT methods.

We are convinced that this approach will be very useful for elucidating the structure of other naturally occurring organoarsenicals, which are coming to the forefront as biologically active compounds or environmental hazards.

## Computational Details

The structure of each species was optimized at the non-relativistic B3LYP/6-311G(2d,2p) level, which was found to be appropriate for the studied compounds (see above). For the optimized structures, the nuclear shieldings $\sigma$ and coupling constants $J$ were calculated with three methods. Methods A and B are non-relativistic and employed the hybrid B3LYP functional. They differ in adopting either the cc-pVTZ basis set (A) or a mixed basis set comprising the pcJ-2 basis set[35] for S, O, C and H, and cc-pVTZ for As. This choice was dictated by the good performance of the pcJ-*n* basis sets in the calculation of coupling constants;[35,36] however, this basis set is not available for As. Method C includes relativistic corrections by means of the zero-order regular approximation (ZORA) up to spin–orbit coupling (ZSO) with the Becke88–Perdew86 GGA functional (BP) and a double-zeta, twice-polarized Slater basis set (TZ2P). Geometry optimizations and non-relativistic calculations were performed with Gaussian 03,[37] and relativistic NMR parameter calculations with ADF 2007[38] and associated *nmr*[39] and *cpl*[40] modules, which allow for calculation of NMR properties within ZORA. Chemical shifts were calculated from $\delta = \sigma_{ref} - \sigma$. TMS was used as the reference compound, with the following $\sigma_{ref}$ values: $^{13}$C: 184.5978 (A), 181.1052 (B), 186.88 (C); $^{1}$H: 31.7660 (A), 31.6927 (B), 31.64 (C). The reference shielding for $^{17}$O ($H_2O$) was 331.8 (C).

## Acknowledgement

The Academy of Finland is acknowledged for financial support, grant no. 119979 (P.T.). Computational resources were provided by the Laboratorio Interdipartimentale di Chimica Computazionale (LICC) at the Department of Chemistry of the University of Padova and the Finnish IT Center for Science (CSC).

[1] N. C. Lloyd, H. W. Morgan, B. K. Nicholson, R. S. Ronimus, *Angew. Chem.* **2005**, *117*, 963–966; *Angew. Chem. Int. Ed.* **2005**, *44*, 941–944.
[2] S. V. Barrett, M. P. Barrett, *Parasitol. Today* **2000**, *16*, 7–9.
[3] S. Waxman, K. C. Anderson, *Oncologist* **2001**, *6*, 3–10.
[4] a) W. H. Miller, H. M. Schipper, J. S. Lee, J. Singer, S. Waxman, *Cancer Res.* **2002**, *62*, 3893–3903; b) H. Chen, R. C. MacDonald, S. Li, N. L. Krett, S. T. Rosen, T. V. O'Halloran, *J. Am. Chem. Soc.* **2006**, *128*, 13348–13349.
[5] J. Lu, E.-C. Chew, A. Holmgren, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12288–12293.
[6] K. A. Francesconi, D. Kuehnelt in *Environmental Chemistry of Arsenic* (Ed.: W. T. Frankenberger), Marcel Dekker, New York, **2002**, pp. 51–94.
[7] Y. Nakamura, T. Narukawa, J. Yoshinaga, *J. Agric. Food Chem.* **2008**, *56*, 2536–2540.
[8] a) R. M. Brown, D. Newton, C. J. Pickford, J. C. Sherlock, *Hum. Exp. Toxicol.* **1990**, *9*, 41–46; b) R. Raml, W. Goessler, P. Traar, T. Ochi, K. A. Francesconi, *Chem. Res. Toxicol.* **2005**, *18*, 1444–1450.
[9] a) R. Bentley, T. G. Chasteen, *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 250–271; b) J. S. Edmonds, K. A. Francesconi, R. V. Stick, *Nat. Prod. Rep.* **1993**, *10*, 421–428; c) K. A. Francesconi, J. S. Edmonds, *Croat. Chem. Acta* **1998**, *71*, 343–359; d) V. M. Dembitsky, T. Rezanka, *Plant Sci.* **2003**, *165*, 1177–1192; e) T. Kaise, T. S. Fukui, *Appl. Organomet. Chem.* **1992**, *6*, 155–160; f) V. M. Dembitsky, D. O. Levitsky, *Prog. Lipid Res.* **2004**, *43*, 403–448; g) P. A. O'Day, *Elements* **2006**, *2*, 77–83.
[10] a) W. R. Cullen, K. J. Reimer, *Chem. Rev.* **1989**, *89*, 713–764; b) J. S. Edmonds, K. A. Francesconi, *Nature* **1981**, *289*, 602–604; c) J. S. Edmonds, K. A. Francesconi, *J. Chem. Soc. Perkin Trans. 1* **1983**, 2375–2382; d) K. A. Francesconi, S. S. Khokiattiwong, W. Goessler, S. N. Pedersen, M. Pavkov, *Chem. Commun.* **2000**, 1083–1084; e) I. Haiduc, J. J. Zuckerman in *Basic Organometallic Chemistry*, W. de Gruyter, Berlin, **1985**, pp. 183–205; f) M. Baudler, J. Hellman, P. Bachmann, K. F. Tebbe, R. Fröhlich, M. Feher, *Angew. Chem.* **1981**, *93*, 415–417; *Angew. Chem. Int. Ed. Engl.* **1981**, *20*, 406–408; g) Y. Yamaoka, M. L. Carmona, J. M. Oclarit, K. Z. Jin, Y. Shibata, *Appl. Organomet. Chem.* **2001**, *15*, 261–265; h) K. Shiomi, *Comp. Biochem. Physiol. Part C* **1988**, *90*, 361–365; i) A. Rumpler, J. S. Edmonds, M. Katsu, K. B. Jensen, W. Goessler, G. Raber, H. Gunnlaugsdottir, K. A. Francesconi, *Angew. Chem.* **2008**, *120*, 2705–2707; *Angew. Chem. Int. Ed.* **2008**, *47*, 2665–2667.
[11] A. Raab, S. H. Wright, M. Jaspars, A. A. Meharg, J. Feldmann, *Angew. Chem.* **2007**, *119*, 2648–2651; *Angew. Chem. Int. Ed.* **2007**, *46*, 2594–2597.
[12] H. R. Hansen, R. Pickford, J. Thomas-Oates, M. Jaspars, J. Feldmann, *Angew. Chem.* **2004**, *116*, 341–344; *Angew. Chem. Int. Ed.* **2004**, *43*, 337–340.
[13] E. Schmeisser, A. Rumpler, M. Kollroser, G. Rechberger, W. Goessler, K. A. Francesconi, *Angew. Chem.* **2006**, *118*, 157–160; *Angew. Chem. Int. Ed.* **2006**, *45*, 150–154.
[14] A. Kolozsi, A. Lakatos, G. Galbács, A. O. Madsen, E. Larsen, B. Gyurcsik, *Inorg. Chem.* **2008**, *47*, 3832–3840, and references therein.
[15] I. Mancini, G. Guella, M. Frostin, E. Hnawia, D. Laurent, C. Debitus, F. Pietra, *Chem. Eur. J.* **2006**, *12*, 8989–8994.
[16] K. R. Dixon in *Multinuclear NMR* (Ed.: J. Mason), Plenum, New York, **1987**.
[17] M. B. L. Marx, H. Pritzkow, B. K. Keppler, *Z. Anorg. Allg. Chem.* **1996**, 622, 1097–1100.
[18] M. E. Elyashberg, A. J. Williams, G. E. Martin, *Prog. Nucl. Magn. Reson. Spectrosc.* **2008**, *53*, 1–104.
[19] *Calculation of NMR and EPR Parameters* (Eds.: M. Kaupp, M. Bühl, V. G. Malkin), Wiley-VCH, Weinheim, **2004**.
[20] A. Bagno, F. Rastrelli, G. Saielli, *Chem. Eur. J.* **2006**, *12*, 5514–5525.
[21] A. Bagno, G. Saielli, *Theor. Chem. Acc.* **2007**, *117*, 603–619.
[22] A. Bagno, M. Bonchio, J. Autschbach, *Chem. Eur. J.* **2006**, *12*, 8460–8471.
[23] A. J. Karttunen, M. Linnolahti, T. A. Pakkanen, *ChemPhysChem* **2007**, *8*, 2373–2378.
[24] P. Nava, R. Ahlrichs, *Chem. Eur. J.* **2008**, *14*, 4039–4045.
[25] A. F. Chiffey, J. Evans, W. Levason, M. Webster, *Polyhedron* **1996**, *15*, 591–596.
[26] H. Schmidbaur, W. Tronich, *Inorg. Chem.* **1968**, *7*, 168–169.
[27] H. Schmidbaur, W. Richter, W. Wolf, F. H. Köhler, *Chem. Ber.* **1975**, *108*, 2649–2655.
[28] C. L. Watkins, L. K. Krannich, R. K. Kanjolia, D. K. Srivastava, *Magn. Reson. Chem.* **1989**, *27*, 616–623.
[29] A. J. Ashe III, R. R. Sharp, J. W. Tolan, *J. Am. Chem. Soc.* **1976**, *98*, 5451–5456.
[30] M. W. Fricke, M. Zeller, H. Sun, V. W.-M. Lai, W. R. Cullen, J. A. Shoemaker, M. R. Witkowski, J. T. Creed, *Chem. Res. Toxicol.* **2005**, *18*, 1821–1829.
[31] The transverse and longitudinal relaxation rates for quadrupolar relaxation of a spin-3/2 nucleus are given by $1/T_1 = 1/T_2 = (2/5)\pi^2\chi^2(1+\eta^2/3)\tau_c$, where $\chi$ is the nuclear quadrupolar coupling constant (NQCC), $\eta$ its asymmetry parameter, and $\tau_c$ the rotational correlation time.[32] A useful rule of thumb for estimating $\tau_c$ is to assign it a value (in picoseconds) numerically equal to the molecular weight. Spherically symmetric molecules such as these may tumble at a substantially faster rate; hence (lacking further data) we can only estimate a range of 1–10 ps.
[32] J. H. Noggle, R. E. Schirmer, *The Nuclear Overhauser Effect: Chemical Applications*, Academic Press, New York, **1971**.

[33] D. W. Rogers, O. P. A. Hoyte, R. K. C. Ho, *J. Chem. Soc. Faraday Trans. 1* **1978**, 46–52.
[34] A. Bagno, F. Rastrelli, G. Saielli, *J. Phys. Chem. A* **2003**, *107*, 9964–9973.
[35] F. Jensen, *J. Chem. Theory Comput.* **2006**, *2*, 1360–1369.
[36] M. U. Roslund, P. Tähtinen, M. Niemitz, R. Sjöholm, *Carbohydr. Res.* **2008**, *343*, 101–112.
[37] Gaussian 03 (Revision A.1), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, Gaussian, Inc., Wallingford, CT, 2004.
[38] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, *22*, 931–967.
[39] a) G. Schreckenbach, T. Ziegler, *J. Phys. Chem.* **1995**, *99*, 606–611; b) G. Schreckenbach, T. Ziegler, *Int. J. Quantum Chem.* **1997**, *61*, 899–918; c) S. K. Wolff, T. Ziegler, *J. Chem. Phys.* **1998**, *109*, 895–905; d) S. K. Wolff, T. Ziegler, E. van Lenthe, E. J. Baerends, *J. Chem. Phys.* **1999**, *110*, 7689–7698.
[40] a) J. Autschbach, T. Ziegler, *J. Chem. Phys.* **2000**, *113*, 936–947.b) J. Autschbach, T. Ziegler, *J. Chem. Phys.* **2000**, *113*, 9410–9418.

Received: June 26, 2008
Published online: October 8, 2008