CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200501583

Toward the Complete Prediction of the ^{1}H and ^{13}C NMR Spectra of Complex Organic Molecules by DFT Methods: Application to Natural Substances

Alessandro Bagno,*[a] Federico Rastrelli,[a] and Giacomo Saielli[b]

Abstract: The NMR parameters (^{1}H and ^{13}C chemical shifts and coupling constants) for a series of naturally occurring molecules have been calculated mostly with DFT methods, and their spectra compared with available experimental ones. The comparison includes strychnine as a test case, as well as some examples of recently isolated natural products (corianlactone, daphnipaxinin, boletunone B) featuring unusual and/or crowded structures and, in the case of boletunone B, being the subject of a recent revision. Whenever experimental spectra were obtained in polar solvents, the calculation of NMR parameters was also carried out with the Integral Equation-Formalism Polarizable Continuum Model (IEF-PCM) continuum method. The computed results generally show a good agreement with experiment, as judged not only by statistical parameters but also by visual comparison of line spectra. The origin of the remaining discrepancies is attributed to the incomplete modeling of conformational and specific solvent effects.

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

Introduction

There is hardly any doubt that NMR spectroscopy is an essential tool for all branches of chemistry. The scope of NMR spectroscopy encompasses both structure and dynamics, but the most widespread use concerns the determination of the structure of both organic and inorganic species. The characterization of naturally occurring substances through ^{1}H and ^{13}C NMR spectroscopy is one of the major applications due to the richness of information that can be obtained, including information on through-bond and through-space connectivities as well as chemical shifts and coupling constants. Very often, application of available pulse sequences yields most C–H connectivities, leading to cogent indication of the molecular structure (often complemented by NOE-derived data). Even though we see a steady advancement in NMR techniques, this information may not lead to an unambiguous structure: among many possible causes there are difficulties in resolving crowded spectral regions, in determining small long-range couplings, or in assigning them. As a result, the confirmation of the proposed structure through X-ray analysis or total synthesis is required, but these avenues are not always available. The scope of the problem is best appreciated by the recent review by Nicolaou and Snyder, devoted to structural misassignments in natural-product chemistry.[1] Indeed, even the spectrum of a molecule as simple and well-known as α-pinene has been revised.[2]

While automated routines, based on database queries or other predictor systems,[3–5] have become increasingly popular and successful, their scope is limited to structures similar to those for which the system has been validated, but "unusual" molecules (often a valuable asset in natural-product chemistry) may not be properly treated by using these approaches.

On the other hand, there is unmatched progress in the computational modeling of NMR parameters.[6–10] Such calculations, most often by DFT methods, have proved very effective at predicting nuclear shieldings and coupling constants of essentially all NMR-active nuclei. Quite naturally, organic biomolecules such as proteins, nucleic acids,[11] and carbohydrates[12] have been the object of such investigations. Particular emphasis was placed by ourselves[13–15] and by Bifulco and co-workers,[16–20] as well as others,[21] on the modeling of organic and naturally occurring molecules through the calculation of ^{1}H and ^{13}C NMR spectra.

For natural products with challenging NMR spectra we face four major problems: a) These molecules possess a

[a] Prof. A. Bagno, Dr. F. Rastrelli
Dipartimento di Scienze Chimiche, Università di Padova (Italy)
Fax: (+39)-049-8275239
E-mail: alessandro.bagno@unipd.it

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

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

large and complex molecular backbone, often with unprecedented features; b) flexible molecules may exist in many conformations, thus greatly increasing the effort needed; c) their spectral assignment relies largely on ^{1}H and ^{13}C NMR spectra with most signals lying within very crowded regions for each functionality; d) ^{1}H chemical shifts are generally prone to solvent effects.

The question is whether the accuracy of quantum chemistry is adequate for this task. With regard to the size problem, important advancements have been made by Helgaker[22] for couplings and by Ochsenfeld[23] for chemical shifts, but a thorough validation is still needed. Most of the effort has, thus, been devoted to the evaluation of modern DFT methods for predicting NMR spectra containing several closely spaced signals and that are influenced by solvent effects.

A crucial issue to be addressed is the way in which the accuracy of computed chemical shifts and coupling constants is assessed. Very often, in the literature, a plot of calculated versus experimental data is presented, and the accuracy judged according to whether: a) there is a linear correlation between the two; b) a statistical parameter such as the correlation coefficient is sufficiently high; c) the mean absolute error is minimal; d) the slope is close to unity and the intercept is close to zero. This approach is certainly a valid one when only the major features are sought, for example, when the whole chemical shift range of a given nucleus is being studied. In such cases (especially with heteronuclei), any given molecule contains only one or very few nuclei of interest and the spread of chemical shifts is so large that statistical errors are likely to be smaller than the $\Delta\delta$ between chemical shifts of individual molecules. This is, however, not necessarily a good guide for assessing the accuracy of calculated ^{13}C and especially ^{1}H shifts as these nuclei resonate in a relatively small range, and particularly as organic molecules contain dozens of such atoms in very similar chemical environments. Thus, no experimental NMR spectroscopist would be interested in sorting out an olefinic and an aromatic carbon or hydrogen atom by calculation, except in special instances. The crucial concern of structural or synthetic chemists is the assignment of carbon and proton signals belonging to similar functional groups the chemical shift of which is influenced by subtle differences in structure and environment. If quantum-chemical calculations are to be of any use in structure elucidation of complex organic molecules, they should be able to provide computed parameters affected by a smaller error than the smallest $\Delta\delta$ being looked for (often <0.1 ppm for ^{1}H). As a consequence, statistical arguments may overlook or hide the fact that a successful calculation must provide chemical shifts in the correct order, in addition to the value of the other indicators. Some of these issues have been discussed in detail by Baldridge and Siegel.[24] For this reason, we have devoted great attention to the evaluation problem. Since, as just mentioned, the ordering of a given pair of signals may be reversed by extremely small inaccuracies, this is the challenge that computational chemistry has to face.

Thus, we have proceeded to calculate the NMR parameters of a series of natural substances to establish whether computational methods can provide a complementary approach to spectroscopic structure determination. In this work we have restricted our study to relatively rigid molecules, for which a single conformation can be reasonably assumed to represent the dynamic structure in solution.

Experimental and Computational Section

Unless otherwise stated, the structure of the molecules investigated was firstly preoptimized at the AM1 level and further optimized at the B3LYP/6-31G(d,p) level. In almost all cases, the final steps of the optimization had to be run by using the GDIIS algorithm.[25] When more than one conformer was conceivable, the choice was based on the experimentally observed NOESY correlations reported in the literature. The optimized structure was then used to calculate NMR properties (isotropic shielding constants σ and spin–spin coupling constants *J*). Most of these calculations have been performed at the B3LYP/cc-pVTZ level, as in our previous works.[13–15] In addition to this, other levels of theory have been tested; when a different method was used to calculate the NMR properties of the molecule, the shielding constant of the reference tetramethylsilane (TMS) was also recalculated with the same method by using the same geometry. Chemical shifts were calculated as $\delta=\sigma_{ref}-\sigma$, where σ_{ref} is the shielding constant of ^{1}H or ^{13}C in the reference compound TMS. Calculated coupling constants were restricted to the Fermi contact term,[13] calculated by the finite-perturbation method. In order to account for solvent effects we adopted the Integral Equation-Formalism Polarizable Continuum Model (IEF-PCM) method.[26] Calculations with Gaussian 03[27] employed the B3LYP[28] and the PBE1PBE[29] hybrid density functionals or the MP2 ab initio method, with the 6-31G(d,p) and cc-pVTZ Gaussian basis sets. Calculations with ADF[30] employed the Becke 88–Perdew 86 (BP)[31] pure functional with the triple-zeta, polarized TZP and TZ2P Slater basis sets as defined in the package.

For experimental NMR studies on strychnine the sample had a concentration of approximately 100 mM in $CDCl_3$. NMR measurements were carried out at 298 K on a Bruker Avance DMX 600 spectrometer equipped with a 5 mm TXI *xyz*-gradient inverse probe. *J*(C,H) coupling constants were measured through heteronuclear *J*-resolved and non-decoupled HSQC experiments. We measured the *J*(H,H) couplings by means of standard *J*-resolved spectroscopy; when assignment was ambiguous, we employed a doubly-selective *J*-resolved pulse sequence based on a modification of the scheme reported in the literature[32] (see Supporting Information for details).

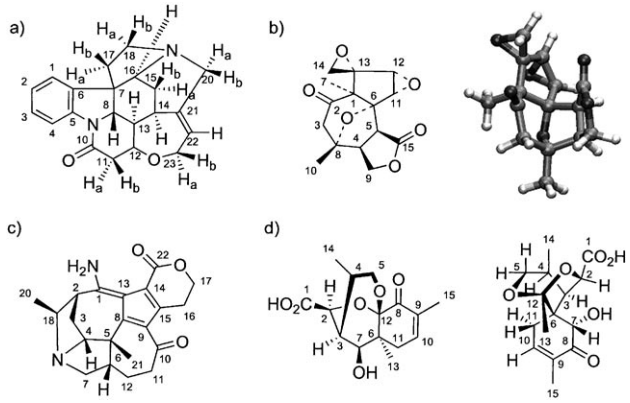
Results and Discussion

As mentioned above, several statistical parameters are available to judge the quality of a calculation but none of them, if taken alone, is fully satisfactory. As will be seen later, most calculations concern chemical shifts rather than couplings, and these have been analyzed in greater detail. For each system we present the parameters *a* and *b* of the linear regression $\delta_{calcd}=a+b\delta_{exptl}$; the correlation coefficient, r^2; the mean absolute error (MAE) defined as $\Sigma_n|\delta_{calcd}-\delta_{exptl}|/n$; the corrected mean absolute error, CMAE,[18] defined as $\Sigma_n|\delta_{corr}-\delta_{exptl}|/n$, where $\delta_{corr}=(\delta_{calcd}-a)/b$ and therefore corrects for systematic errors. For coupling constants the above formulae may be used by replacing δ with *J*.

In addition, we also compare calculated and experimental chemical shifts by means of bar-graph spectra, connected by

dotted lines to visualize any inversion in the order of the resonances, thus providing an intuitive guide to appreciating the quality of calculated data. All tables with numerical values of calculated and experimental NMR parameters are collected in the Supporting Information.

The structures of all substances investigated are presented in Scheme 1.



Scheme 1. Structures and numbering of natural products investigated: a) Strychnine; b) corianlactone with three-dimensional structure; c) daphnipaxinin; d) boletunone B: originally proposed structure (left); revised structure (right).

Strychnine: Strychnine, a well-characterized alkaloid of the *Strychnos* genus, was chosen as a test example to investigate the applicability of the protocol because its NMR parameters are well understood (it is often used to test new pulse sequences) and is also a fairly rigid molecule.

The ^{1}H NMR spectrum in $CDCl_3$ was assigned by Carter and co-workers[33] and was later corrected by Chazin et al.[34] For ^{13}C NMR chemical shifts for strychnine see for example ref. [35]. We have calculated the NMR properties at the following levels of theory: A) B3LYP/cc-pVTZ; B) PBE1PBE/cc-pVTZ; C) BP/TZP; D) BP/TZ2P (Table 1).

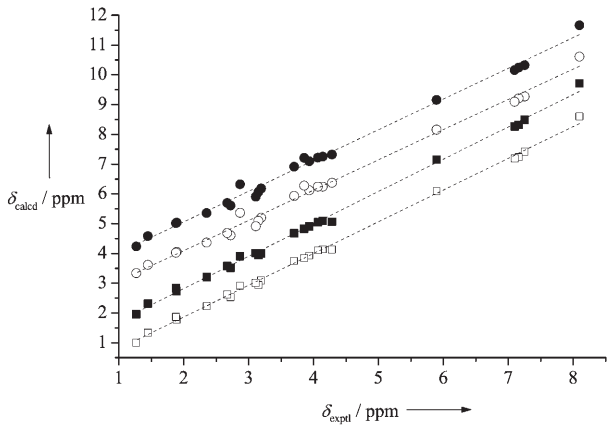


Figure 1. Correlation between calculated and experimental ^{1}H chemical shifts of strychnine. Methods: A (□); B (■); C (○); D (●). For clarity, results obtained with methods B, C, and D are shifted along the *y* axis by 1, 2, and 3 ppm, respectively. For each set of data the linear fitting is also reported as a dashed line.

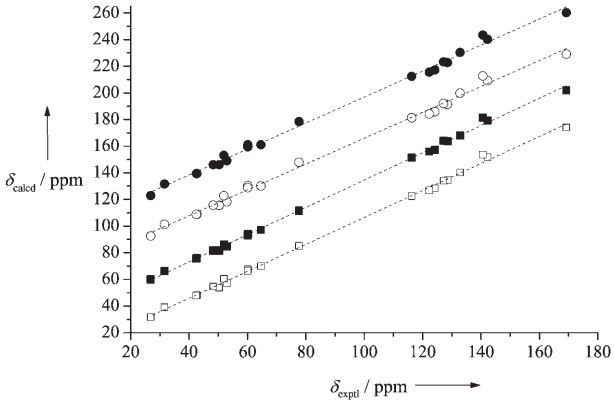


Figure 2. Correlation between calculated and experimental ^{13}C chemical shifts of strychnine. Methods: A (□); B (■); C (○); D (●). For clarity, results obtained with methods B, C, and D are shifted along the *y* axis by 30, 60, and 90 ppm, respectively. For each set of data the linear fitting is also reported as a dashed line.

Table 1. Correlation and fitting parameters of calculated NMR properties of strychnine.[a]

Method		*a*	*b*	r^2	MAE[b]	CMAE[c]
$\delta(^1H)$						
B3LYP/cc-pVTZ (A)		−0.27±0.05	1.07±0.01	0.9977	0.12	0.07
PBE1PBE/cc-pVTZ (B)		−0.36±0.05	1.09±0.01	0.9974	0.16	0.08
BP/TZP (C)		0.06±0.09	1.02±0.02	0.9923	0.16	0.14
BP/TZ2P (D)		−0.01±0.09	1.03±0.02	0.9933	0.16	0.13
$\delta(^{13}C)$						
A		5±1	1.01±0.01	0.9979	5.7	1.4
B		2±1	1.02±0.01	0.9985	3.2	0.9
C		9±1	0.97±0.01	0.9953	7.1	1.9
D		9±1	0.98±0.01	0.9960	7.6	1.8
A	^nJ(H,H)	0.17±0.09	0.93±0.01	0.9940	0.6	0.4
A	^1J(C,H)	−29±6	1.12±0.04	0.9741	12.5	1.4
A	^nJ(C,H)[d]	0.5±0.2	0.82±0.03	0.9190	2.6	2.4

[a] Linear fitting parameters refer to $\delta_{calcd}=a+b\delta_{exptl}$. [b] Mean average error: $MAE=\Sigma_n|\delta_{calcd}-\delta_{exptl}|/n$. [c] Corrected mean average error: $CMAE=\Sigma_n|\delta_{corr}-\delta_{exptl}|/n$ (see text). For coupling constants replace δ with *J*. [d] Experimental values from ref. [41], method I.

The PBE1PBE functional slightly increases the agreement for ^{13}C chemical shifts but does not significantly affect the results for the ^{1}H chemical shifts.[36] Pure functionals combined with the Slater basis sets (methods C and D above) yield slightly worse results. The correlation between calculated and experimental chemical shifts is shown in Figure 1. In this Figure and those that follow, it is important to point out two facts: a) The general correlation is very satisfactory, in the sense that all signals originating from

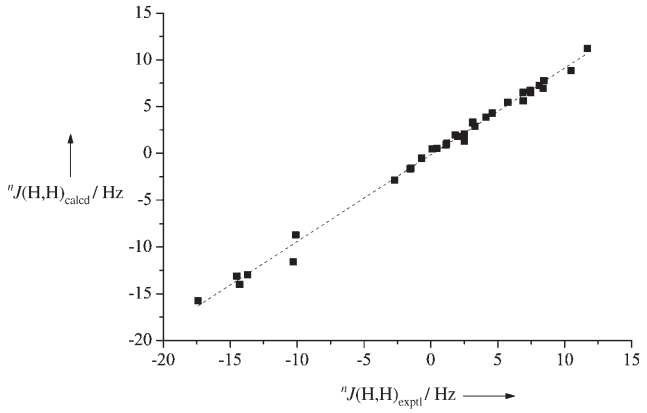


Figure 3. Calculated versus experimental J(H,H) coupling constants in strychnine. Linear fitting is shown as a dashed line.

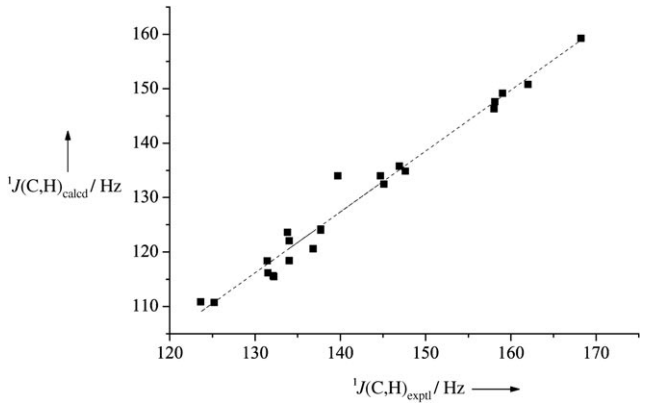


Figure 4. Calculated versus experimental direct ^{1}J(C,H) coupling constants in strychnine. Linear fitting is shown as a dashed line.

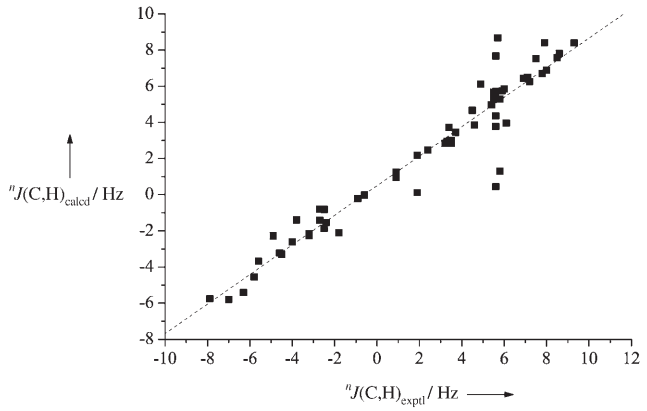


Figure 5. Calculated versus experimental long-range ^{n}J(C,H) coupling constants in strychnine. Linear fitting is shown as a dashed line.

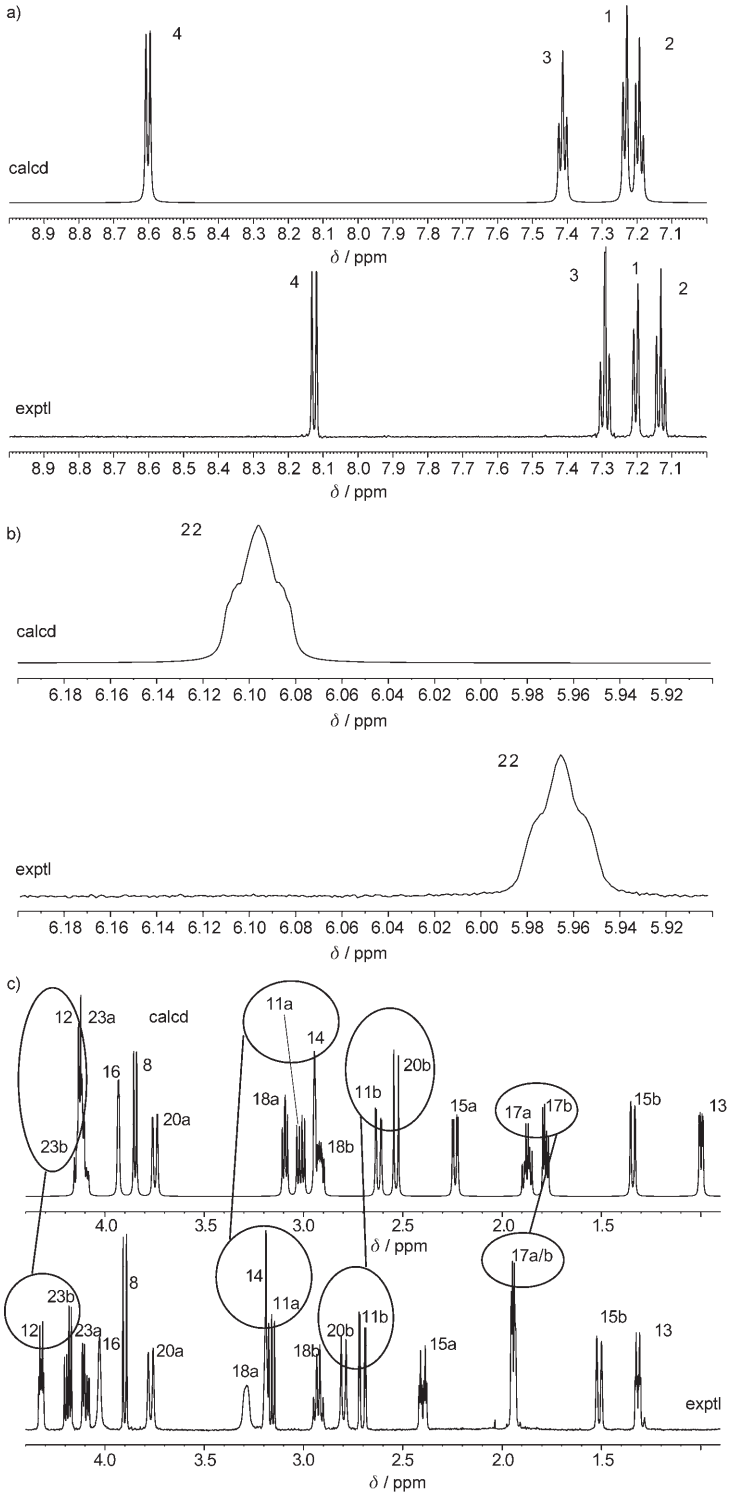


Figure 6. Calculated (method A) and experimental ^{1}H spectrum of strychnine: a) Aromatic region; b) olefinic region; c) aliphatic region. Inversions between calculated and experimental resonances are highlighted.

different functional groups fall into their own distinctive regions. Thus, the overall reliability of these calculations is borne out. b) Whenever two data points can be connected by a segment with negative slope, those signals appear in reverse order than they would in the experimental spectrum, that is, a misassignment would be made. We note that a few

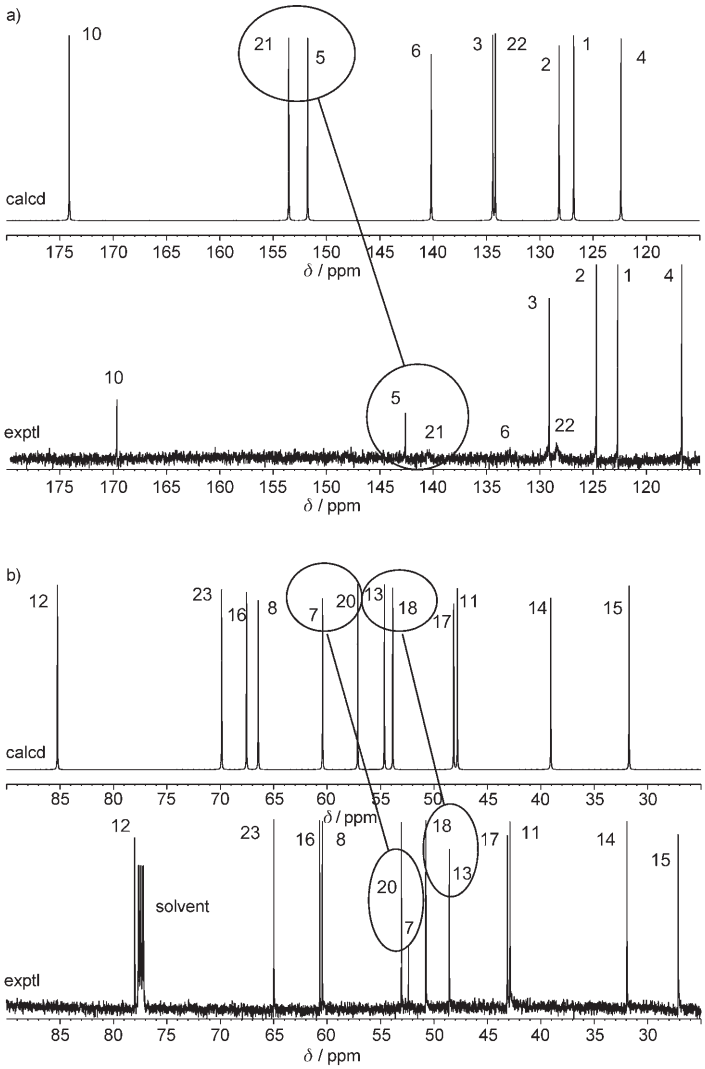


Figure 7. Calculated (method A) and experimental ^{13}C spectrum of strychnine: a) Low-field region; b) high-field region. Inversions between calculated and experimental resonances are highlighted.

of these cases do occur, however, only for closely spaced points, hence presumably originating from nuclei in the same functional group.

With method A the order is correct for 14 (out of 22) ^{1}H NMR signals, but four pairs are interchanged with respect to the experimental ordering: H-12,H-23b ($\Delta\delta = 0.14$ ppm); H-14,H-11a ($\Delta\delta = 0.04$ ppm); H-20b,H-11b ($\Delta\delta = 0.05$ ppm); H-17b,H-17a ($\Delta\delta = 0.01$ ppm). The first signal of each pair has a calculated value that is too much deshielded, while the second one is in good agreement with the experimental data.

Most of the 21 ^{13}C NMR signals (Figure 2) are also in agreement with literature data,[35] except for the pairs C-5,C-21; C-20,C-7; and C-18,C-13 that are calculated in reverse order. Again we note that the two signals of these pairs are quite close in the experimental spectrum: δ 1.6, 1.0, and 2.1 ppm, respectively. Also, the wrong prediction for C-13 (CH) and C-18 (CH_2) would be very easy to correct.

We have also calculated (method A) all J(H,H) and J(C,H) coupling constants (Figures 3–5). Most experimental J(H,H) couplings[33,34,37] have been recently corrected and agree better with our calculated values: thus, ^{2}J(H-17a,H-17b) was reported[34] to be 15.5 Hz and was revised to 10.3 Hz;[38] (calcd −11.6 Hz). Similarly, ^{2}J(H-20a,H-20b) was revised to 14.3 Hz[37] against 14.7 (calcd −14.0 Hz); ^{2}J(H-23a,H-23b) was revised to 13.7 Hz[39] against 14.2 Hz (calcd −13.0 Hz). Some new J(H,H) couplings of strychnine have been recently determined by an indirect deconvolution (post-processing) method, including a remarkable ^{5}J(H-20a,H-23a) of 1.80 Hz (calcd 2.04 Hz).[40] We have measured the same previously unreported proton couplings of strychnine with our new pulse sequence obtaining similar but slightly higher values (see Supporting Information) together with an additional coupling constant, ^{5}J(H-15b,H-22), close to our experimental resolution of 1.0 Hz (calcd 0.7 Hz).

As mentioned in the introduction, only the Fermi-contact term was calculated. In this context, we note that diamagnetic and paramagnetic spin–orbit contributions to long-range C–H coupling constants might be significant for aromatic spin systems, but they are negligible in aliphatic spin systems.[14] Our own measured ^{1}J(C,H) coupling constants appear to be systematically underestimated by approximately 15–20 Hz. Nevertheless the correlation shown in Figure 4 is satisfactory. Recently, long-range ^{n}J(C,H) coupling constants of strychnine have been measured with two new methods.[41] There is a good correlation between the calculated values and both experimentally derived values except for a few cases: ^{3}J(C-16,H-17a) is found

Table 2. Correlations and fitting parameters of ^{1}H and ^{13}C chemical shifts of corianlactone.[a]

Geometry[b]	a	b	r^2	MAE	CMAE
$\delta(^{1}H)$					
X-ray (A)	−0.3 ± 0.3	0.93 ± 0.08	0.9268	0.50	0.23
X-ray, H opt (B)	−0.1 ± 0.2	0.91 ± 0.06	0.9521	0.40	0.16
X-ray, H opt + PCM (B′)	−0.1 ± 0.2	1.00 ± 0.06	0.9631	0.16	0.13
B3LYP/6-31G(d,p) (C)	−0.05 ± 0.2	0.90 ± 0.06	0.9513	0.41	0.19
B3LYP/6-31G(d,p) + PCM (C′)	−0.1 ± 0.2	0.98 ± 0.04	0.9805	0.16	0.11
MP2/cc-pVDZ (E)	0.2 ± 0.2	0.87 ± 0.06	0.9370	0.26	0.21
MP2/cc-pVDZ + PCM (E′)	0.2 ± 0.1	0.95 ± 0.03	0.9898	0.10	0.09
$\delta(^{13}C)$					
A	6 ± 2	0.99 ± 0.02	0.9942	5.9	3.1
B	7 ± 2	0.98 ± 0.02	0.9956	6.2	2.7
B′	6 ± 1	1.02 ± 0.02	0.9964	7.2	2.3
C	3 ± 1	1.03 ± 0.01	0.9985	5.3	1.6
C′	1 ± 1	1.07 ± 0.01	0.9990	6.4	1.2
E	3.2 ± 0.7	1.041 ± 0.007	0.9994	6.1	1.1
E′	1.5 ± 0.5	1.078 ± 0.007	0.9993	7.2	1.0

[a] See footnotes to Table 1. [b] B3LYP/cc-pVTZ always used for NMR calculations. The prime indicates that NMR properties have been calculated by using the PCM solvation model (see text).

to be 5.5–5.6 Hz,[41] while the calculated value is only 0.45 Hz. The second case is ^{3}J(C-10,H-12), determined as 5.8 Hz (calcd 1.29 Hz). The correlation coefficient between calculated and experimental data by Keeler when method I of ref. [41] is used is slightly better and these new values correlate significantly better with our calculations with respect to previously available ones.[42–45]

In Figures 6 and 7 we show the simulated ^{1}H and ^{13}C NMR spectra (9.4 T) of strychnine obtained by using the calculated and experimental parameters. The aromatic ^{1}H region is well-reproduced even if the H-4 proton is too much deshielded; coupling patterns are also preserved. The apparent broad triplet of the H-22 vinylic proton is seen to arise from an unresolved complex coupling of H-22 with H-14, H-20a, H-23a, and H-23b. The crowded aliphatic region, spanning across 3 ppm, is predictably the most difficult as discussed above. However, coupling patterns seem to be preserved (as far as one can judge graphically). In particular, calculated proton spin–spin coupling constants are very well correlated with the experimental values.

The strychnine case shows that a good level of accuracy can be attained by DFT calculations, even though some shortcomings remain that concern essentially only chemical shifts. Indeed, in some cases where a disagreement is noted the peaks are so close that, probably, factors such as concentration and temperature may alter the experimental data.

We now turn our attention to some recently discovered molecules, for which the structures have been determined by means of NMR spectroscopy and for which some ambiguities

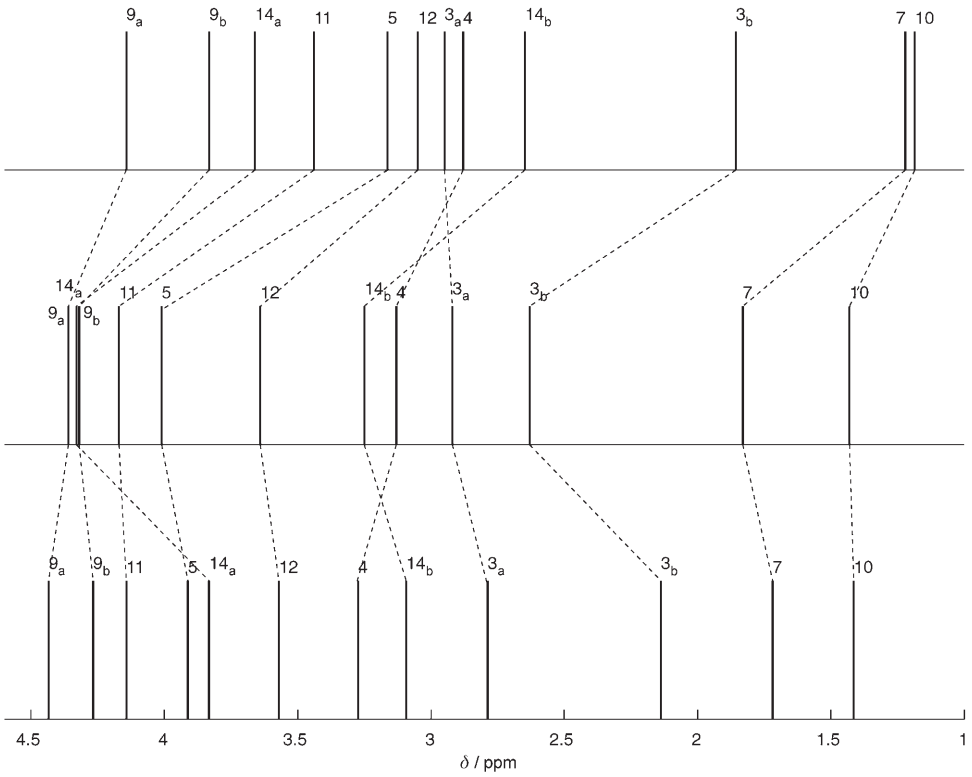


Figure 8. ^{1}H NMR spectra of corianlactone. Experimental (middle); method A (top); B′ (bottom).

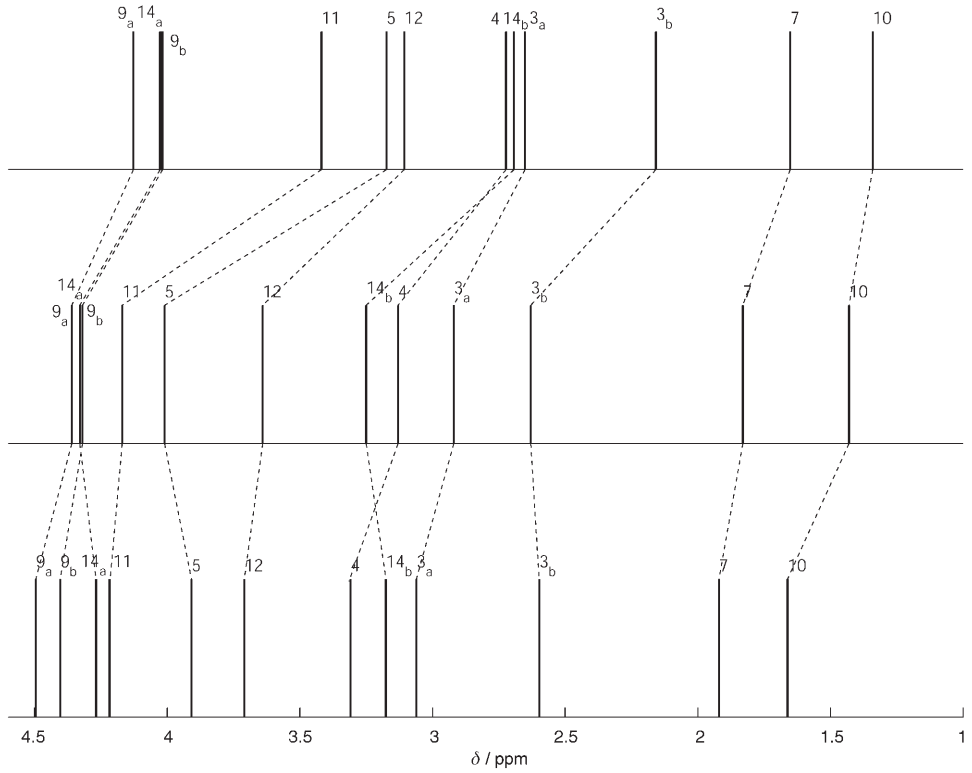


Figure 9. ^{1}H NMR spectra of corianlactone. Experimental (middle); method C (top); E′ (bottom).

have emerged in the literature. Since the modeling of coupling constants is much more satisfactory than that of chemical shifts, we will focus on the chemical shifts only, except in specific cases. Also, the pertinent spectra were often obtained in more polar solvents than chloroform, so that solvent effects will also be tested.

Corianlactone: Corianlactone is a new sesquiterpene recently isolated and characterized by Sun and co-workers from *Coriaria nepalensis*.[46] Its structure has been determined by NMR spectroscopy in [D_5]pyridine and by X-ray diffraction; it is rigid, sterically congested, and features 12 ^{1}H and 15 ^{13}C signals. In our calculations we have used the following models: A) X-ray geometry;[47] B) X-ray geometry for the skeleton but with the coordinates

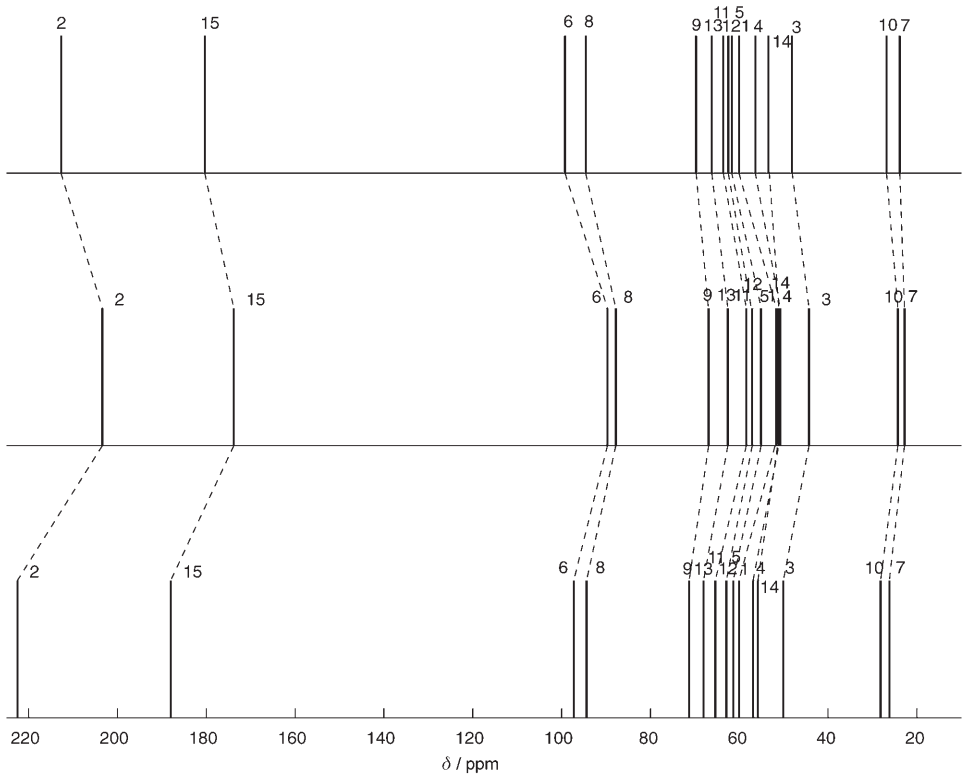


Figure 10. ^{13}C NMR spectra of corianlactone, full range. Experimental (middle); method C (top); E′ (bottom).

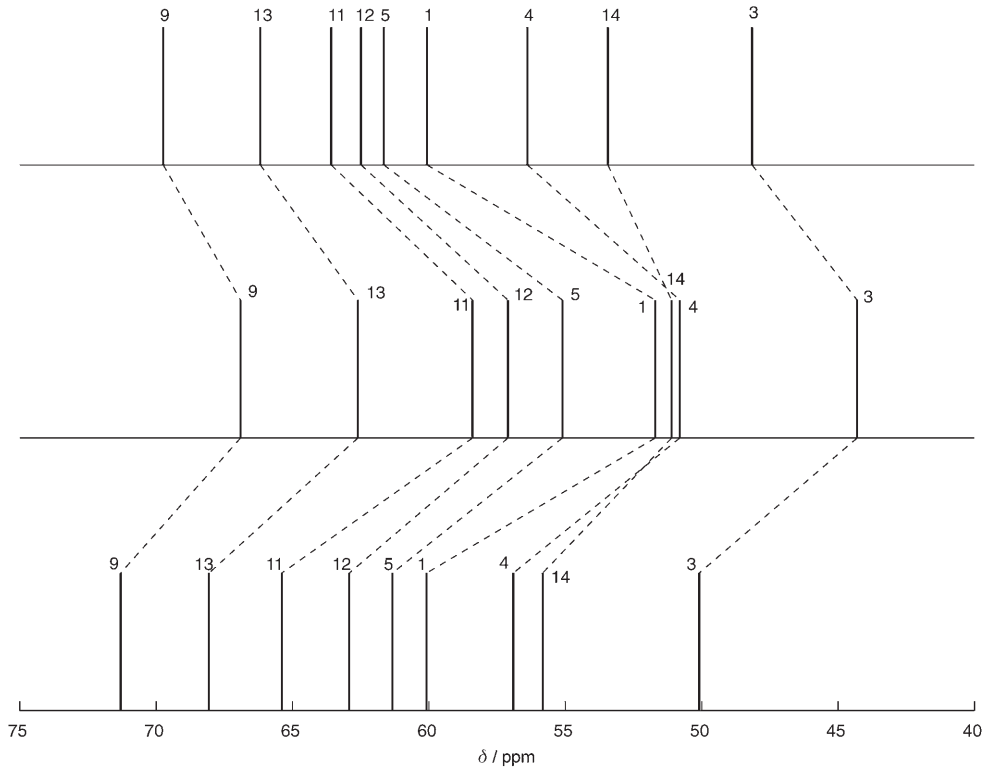


Figure 11. ^{13}C NMR spectra of corianlactone between 40 and 80 ppm. Experimental (middle); method C (top); E′ (bottom).

of hydrogen atoms optimized with B3LYP/6-31G(d,p); C) fully optimized geometry at B3LYP/6-31G(d,p) in vacuo; D) fully optimized geometry (B3LYP/6-31G(d,p)) with solvent reaction field of pyridine ($\varepsilon = 12.5$); E) fully optimized geometry (MP2/cc-pVDZ) in vacuo. NMR properties have always been calculated at the B3LYP/cc-pVTZ level, both in vacuo and with the reaction field of pyridine (indicated by a prime). In Table 2 we report the correlation parameters for chemical shifts, which span over δ 2.9 and 44 ppm for the ^{1}H and ^{13}C NMR data.

The results from the X-ray structure A) are very poor especially for the proton chemical shifts; little improvement is observed when the hydrogen atom positions are minimized (B). A substantial improvement occurs when the solvent effect on the hydrogen-relaxed X-ray structure is introduced (B′). Still, H-14a and H-3b deviate by about 0.5 ppm; also, H-14b and H-4 are reversed (Figure 8).

Since the experimental data have been recorded in pyridine, we have also optimized the structure in pyridine at the PCM-B3LYP/6-31G(d,p) level (Table 2(D)), but the minimized structures are almost indistinguishable, that is, if there is a solvent effect on the structure it is not modeled by the reaction field method. Therefore, by using the structure optimized in vacuo (C), we have calculated the NMR properties both in vacuo and in pyridine. There is no significant improvement when the NMR properties are calculated in vacuo; however, the improvement is substantial when the NMR properties are calculated by using the solvent reaction field (C′). In this case, the only two protons which appear to be less well-correlated are H-14a and H-14b which are both calculated too much shielded.

Since the C-14 methylene group is close to a carbonyl group, we expect the geometry (and electronic structure) to be affected by its π electrons. Subtle effects related to electron correlation are not always correctly modeled by DFT; to check for this, we optimized the geometry at the ab initio MP2/cc-pVDZ level (Table 2(E)). The distance between the C-15 carbonyl oxygen atom and the C-14 atom changes from 2.94 (B3LYP) to 2.83 Å (MP2) (exptl 2.76 Å). As can be seen in Figure 9, the calculation of NMR properties including the solvent reaction field on the MP2 structure (E′) improves the agreement also for the H-14a and H-14b protons facing the carbonyl group. Unfortunately, even at this level the H-14b and H-4 proton chemical shifts are still reversed.

The ^{13}C chemical shifts are also in better agreement with experiment when calculated on the MP2 structure; little effect of the reaction field is observed. In Figure 10 we show the comparison between calculated and experimental ^{13}C NMR spectra across the full chemical shift range, while an enlarged region is shown in Figure 11. Even at level E′ the C-4 and C-14 chemical shifts ($\Delta\delta = 0.3$ ppm) are reversed (again, this inconsistency should be easily addressed experimentally). The solvent effect is much less important, as expected. The choice of geometry is the only important factor and, as expected, the MP2 geometry gives better results than the B3LYP one. This example highlights the need for accurate methods for sterically crowded molecules where electron correlation effects may be important to obtain the correct structure. In addition, solvent effects have to be included to obtain a good agreement.

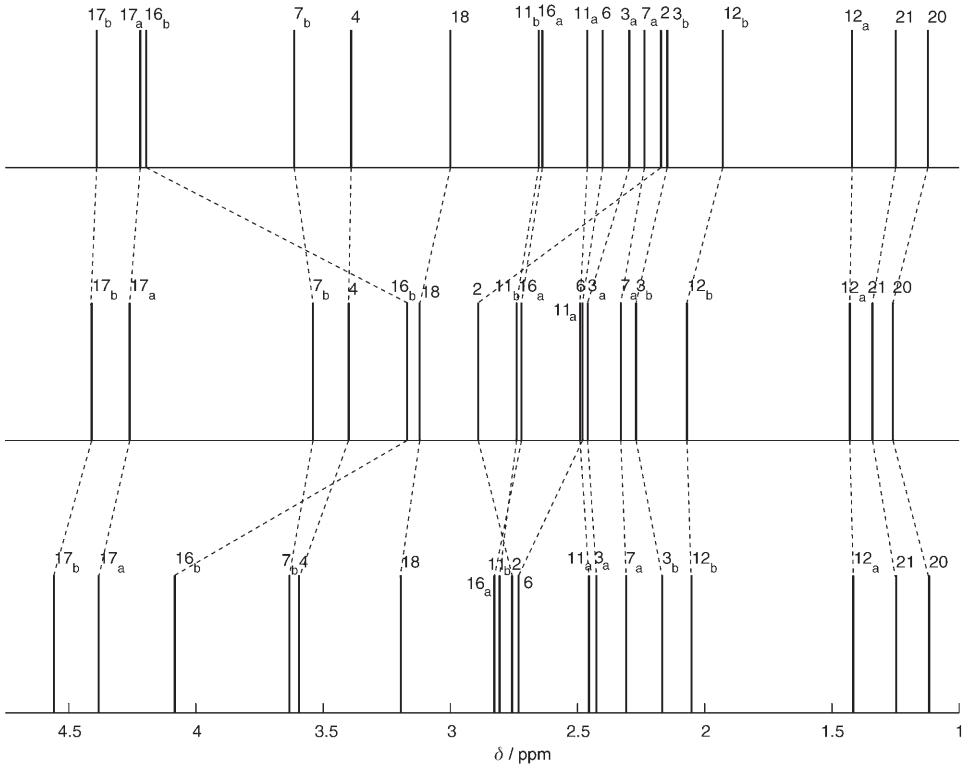


Figure 12. ^{1}H NMR spectra of daphnipaxinin. Experimental (middle); method A (top); A′ (bottom).

Table 3. Correlations and fitting parameters of NMR chemical shifts of daphnipaxinin.[a]

Structure	a	b	r^2	MAE	CMAE
$\delta(^{1}H)$					
A	-0.2 ± 0.2	1.06 ± 0.09	0.9028	0.17	0.13
A′	-0.2 ± 0.2	1.11 ± 0.06	0.9543	0.14	0.11
$\delta(^{13}C)$					
A	3 ± 1	1.009 ± 0.008	0.9986	3.9	1.6
A′	2.1 ± 0.7	1.030 ± 0.006	0.9993	4.7	1.2

[a] See footnotes to Table 1.

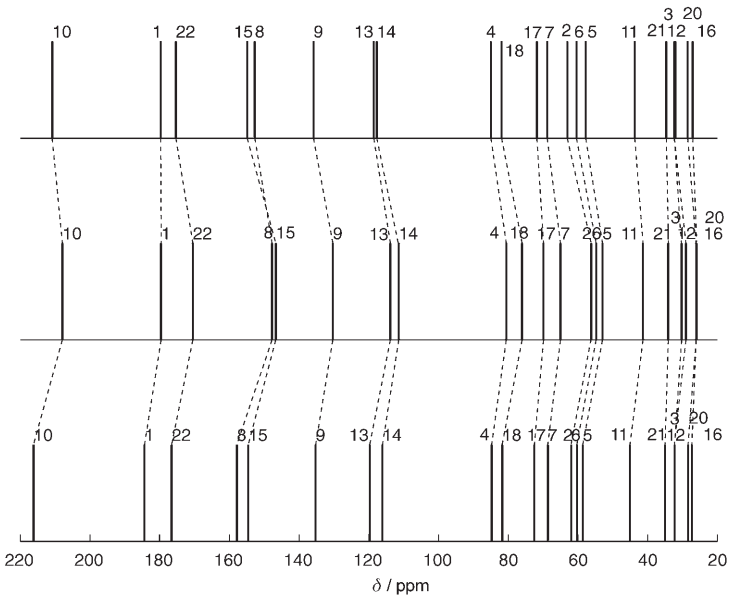


Figure 13. ^{13}C NMR spectra of daphnipaxinin. Experimental (middle); method A (top); A′ (bottom).

Daphnipaxinin: Daphnipaxinin was isolated by Yue and co-workers and is the first alkaloid of the genus *Daphniphyllum* with an unprecedented hexacyclic fused skeleton.[48] The structure was determined mainly by using NMR data (in [D_4]MeOH) since it was not possible to obtain high-quality crystals for X-ray analysis. The proposed structure[48] was optimized at the B3LYP/6-31G(d,p) level. NMR properties (18 ^1H and 21 ^{13}C signals) were calculated in vacuo (Table 3(A)) and with the solvent reaction field of methanol ($\varepsilon = 32.63$) (A′). Correlation parameters (Table 3) do not indicate a fully satisfactory correlation for protons, while the situation is remarkably better for carbon atoms.

Concerning protons, the disagreement between gas-phase calculated and experimental results is large only for H-2 (calcd δ 4.20, exptl 3.17 ppm) and H-16b (calcd δ 2.17, exptl 2.89 ppm). The remaining shifts are in very good agreement with experiment (excluding H-2 and H-16b protons, r^2 becomes 0.9968). These two protons are close to the hydrogen-bonding sites of this molecule (the amino and keto groups, respectively). In fact, when using the PCM method the agreement is improved only for the H-2 proton (Figure 12). The situation is again much better for ^{13}C, with a very good correlation between experimental and calculated resonances (Figure 13). These observations suggests that, when hydrogen bonds or other specific interactions with the solvent occur, a full dynamic treatment of the solvent molecules should be considered,[49,50] but such an approach is not yet feasible for molecules of this size.

Boletunone B: Finally we present the case of boletunone B, a fungal metabolite recently isolated from the mushroom *Boletus calopus* by Yoo and co-workers.[51] This is a representative example of situations that may arise in this context. The structure firstly proposed on the basis of NMR evidence was soon revised by Steglich and co-workers,[52] still based mainly on NMR data in [D_6]DMSO. Thus, we have calculated the ^1H and ^{13}C spectra (12 and 15 signals, respectively) both for the original proposal (Table 4(B1)) and the revised structure (B2). Drawing on the previous results, the geometries have been optimized at the B3LYP/6-31G(d,p) level in the gas phase, and the NMR properties calculated with B3LYP/cc-pVTZ, in the gas phase and with the solvent reaction field of DMSO ($\varepsilon = 46.7$), denoted with a prime as before (Table 4).

Table 4. Correlations and fitting parameters of the ^1H and ^{13}C chemical shifts of the two proposed structures of boletunone B.[a]

Structure	*a*	*b*	r^2	MAE	CMAE
$\delta(^1H)$					
B1[b]	0.2 ± 0.2	0.99 ± 0.08	0.9439	0.31	0.29
B1′	0.1 ± 0.2	1.07 ± 0.06	0.9675	0.36	0.22
B2[c]	0.1 ± 0.1	1.04 ± 0.04	0.9856	0.24	0.16
B2′	0.1 ± 0.1	1.11 ± 0.04	0.9844	0.39	0.15
$\delta(^{13}C)$					
B1	6 ± 2	0.98 ± 0.02	0.9944	5.3	3.9
B1′	5 ± 2	1.01 ± 0.02	0.9952	6.0	3.7
B2	4 ± 2	1.03 ± 0.01	0.9986	6.2	1.8
B2′	3 ± 2	1.05 ± 0.01	0.9984	7.2	1.9

[a] B3LYP/cc-pVTZ//B3LYP/6-31G(d,p); the prime indicates that NMR properties have been calculated by using the PCM solvation model. See also the footnotes to Table 1. [b] Original structure.[51] [c] Revised structure.[52]

The fit parameters (Table 4) indicate that a better correlation is obtained by using the revised structure (B2). However, some ^1H chemical shifts are still not in perfect agreement, for example, the geminal protons H-11a,b. In particular, H-11b is more deshielded (Figure 14). The ^{13}C spectrum calculated by using the revised structure is also in better agreement with the experimental spectrum than the one calculated by using the originally proposed structure (see Figure 15). An inversion of resonances still occurs (C-4 and C-11) but the two signals differ by only 2.5 ppm.

One of the arguments given in support of the revised structure over the original proposal[52] was that methyl protons H-13 are unusually shielded (0.94 ppm) compared with the chemical shift of 1.57 ppm exhibited by the same methyl group in the structurally related boletunone A (not discussed here). According to Steglich and co-workers this difference can be accounted for by the revised structure B2 (Table 4), where the methyl group lies above a conjugated system. However, the calculated chemical shift of the H-13 proton is, for both B1 and B2, in very close agreement with the experimental data: we obtain δ 1.02 ppm for the revised structure (B2) and δ 1.08 ppm for the original proposal (B1) (see Supporting Information). Thus, this difference appears too small to favor one structure or the other. Therefore, subtle details of the geometry cannot be ignored when unusual structures occur, as we have found in the case of corianlactone. Another argument against the original proposal, according to Steglich and co-workers, is that in structure B1

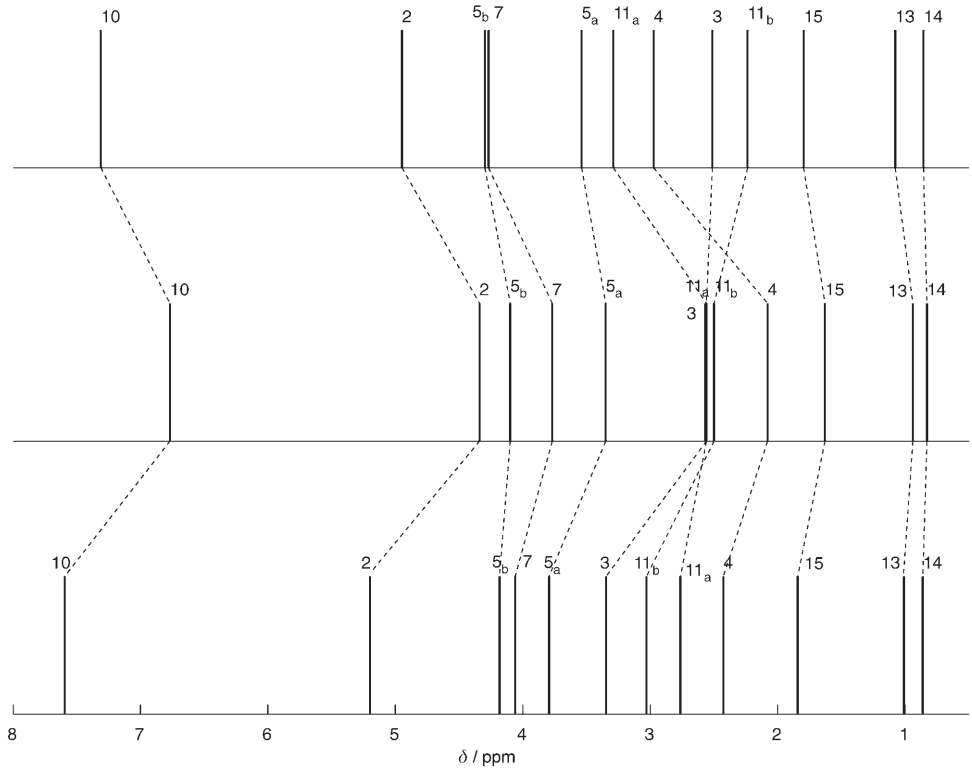


Figure 14. ^{1}H NMR spectra of boletunone B. Experimental (middle); calculated for the original structure B1′ (top); revised structure B2′ (bottom).

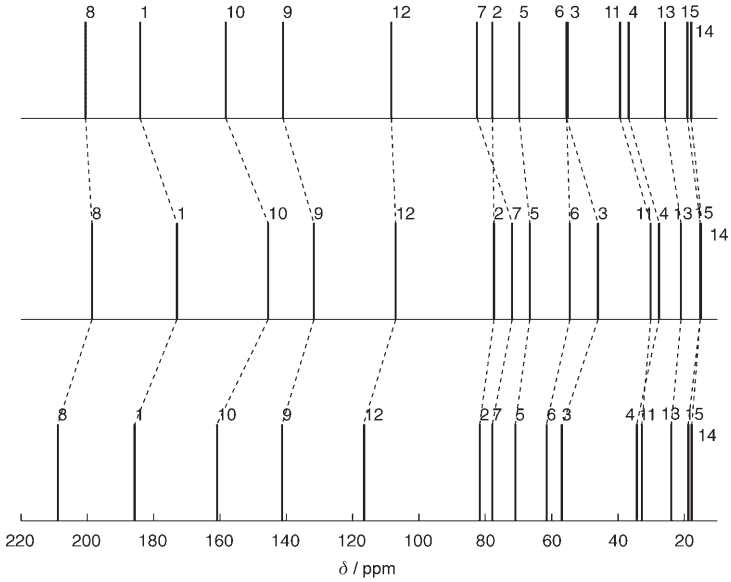


Figure 15. ^{13}C NMR spectra of boletunone B. Experimental (middle); calculated for the original structure B1′ (top); revised structure B2′ (bottom).

there should be a relatively strong vicinal ^{3}J(H-3,H-7) coupling. Indeed, our calculated value of 6.9 Hz, obtained by using the originally proposed structure, should have been easily detected in the spectrum.

We have also optimized the structures at the MP2/cc-pVDZ level, but the changes in geometry were very small and NMR properties were very similar. The disagreement for protons H-11 may be explained by the fact that DMSO is a strong hydrogen-bond acceptor and is likely to interact with the hydroxyl group on the C-7 atom; this specific interaction may affect H-11b, lying on the same side as the OH. Clearly, the above considerations on specific solvent effects apply here as well.

We have also calculated all coupling constants for structure B1; some of the experimentally observed couplings and calculated values are reported in Table 5. Together with the previously mentioned ^{3}J(H-3,H-7), we also note that the observable correlation ^{2}J(C-4,H-5a) was calculated to be larger in structure B2 than in B1.

Table 5. Selected coupling constants of boletunone B.

J_{expt} [Hz][a]		J_{calcd} [Hz][b]	
		B1	B2
not observed[c]	J(H-3,H-7)	6.4	0.0
H-2 (d, 3.6)	J(H-2,H-3)	1.9	3.8
H-5a (dd, 11.5, 6.8)	J(H-5a,H-5b)	−10.9	−10.6
	J(H-5a,H-4)	5.6	6.1
H-5b (dd, 11.5, 11.5)	J(H-5b,H-5a)	−10.9	−10.6
	J(H-5b,H-4)	10.6	11.4
observed[d]	^{3}J(C-3,H-5a)	6.8	5.5
observed[d]	^{2}J(C-4,H-5a)	−0.8	−1.7
observed[d]	^{3}J(C-12,H-5a)	8.6	6.9

[a] From refs. [51] and [52]. [b] B3LYP/cc-pVTZ//B3LYP/6-31G(d,p), Fermi contact term only. [c] Not observed in the correlation spectrum. [d] Observed in the correlation spectrum but value not determined.

Conclusions and Outlook

DFT calculations can attain a considerable degree of accuracy in the prediction of the ^{13}C and ^{1}H NMR spectra of complex organic molecules such as naturally occurring substances. In most cases the correct ordering of ^{1}H and, especially, of ^{13}C signals is predicted. When experimental spectra are recorded in polar solvents, the modeling of solvent by means of a reaction field generally leads to a substantial improvement. The importance of geometry optimization is also highlighted; an ab initio method such as MP2 is shown

to perform better than DFT in critical cases. While some shortcomings remain, they are probably related to the need for a wider sampling of conformations in flexible molecules, a better treatment of specific solvent effects, or possibly to minor factors like concentration and temperature effects. Nevertheless, such calculations can considerably aid in the assignment of crowded, poorly resolved, or unusual spectra, without recourse to external sources like database lookup.

Acknowledgements

This work was financially supported by the University of Padova (Progetto di ricerca di Ateneo CPDA045589).

[1] a) K. C. Nicolaou, S. A. Snyder, *Angew. Chem.* **2005**, *117*, 1036–1069; *Angew. Chem. Int. Ed.* **2005**, *44*, 1012–1044.
[2] S.-G. Lee, *Magn. Reson. Chem.* **2002**, *40*, 311–312.
[3] M. J. P. Ferreira, V. P. Emerenciano, G. A. R. Linia, P. Romoff, P. A. T. Macari, G. V. Rodrigues, *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *33*, 153–206.
[4] For a recent entry to this approach see: J. Junker, W. Maier, T. Lindel, M. Köck, *Org. Lett.* **1999**, *1*, 737–740.
[5] C. A. Hunter, M. J. Packer, C. Zonta, *Prog. Nucl. Magn. Reson. Spectrosc.* **2005**, *47*, 27–39.
[6] T. Helgaker, M. Jaszuñski, K. Ruud, *Chem. Rev.* **1999**, *99*, 293–352.
[7] J. Vaara, J. Jokisaari, R. E. Wasylishen, D. L. Bryce, *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *41*, 233–304.
[8] R. H. Contreras, V. Barone, J. C. Facelli, J. E. Peralta, *Annu. Rep. NMR Spectrosc.* **2003**, *51*, 167–260.
[9] *Calculation of NMR and EPR Parameters* (Eds.: M. Kaupp, M. Bühl, V. G. Malkin), Wiley-VCH, Weinheim, **2004**.
[10] Special Issue on "Theoretical Methods in Magnetic Resonance" (Ed.: M. Bühl), *Magn. Reson. Chem.* **2004**, *42*, S1–S206.
[11] D. A. Case, *Calculation of NMR and EPR Parameters* (Eds.: M. Kaupp, M. Bühl, V. G. Malkin), Wiley-VCH, Weinheim, **2004**.
[12] a) F. Cloran, Y. Zhu, J. Osborn, I. Carmichael, A. S. Serianni, *J. Am. Chem. Soc.* **2000**, *122*, 6435–6448; b) J. Czernek, J. Lang, V. Sklenář, *J. Phys. Chem. A* **2000**, *104*, 2788–2792; c) O. L. Malkina, M. Hricovíni, F. Bízik, V. G. Malkin, *J. Phys. Chem. A* **2001**, *105*, 9188–9195; d) R. Stenutz, I. Carmichael, G. Widmalm, A. S. Serianni, *J. Org. Chem.* **2002**, *67*, 949–958; e) E. A. Larsson, J. Ulicny, A. Laaksonen, G. Widmalm, *Org. Lett.* **2002**, *4*, 1831–1834; f) M. U. Roslund, K. D. Klika, R. L. Lehtila, P. Tähtinen, R. Sillanpää, R. Leino, *J. Org. Chem.* **2004**, *69*, 18–25.
[13] A. Bagno, *Chem. Eur. J.* **2001**, *7*, 1652–1661.
[14] A. Bagno, F. Rastrelli, G. Saielli, *J. Phys. Chem. A* **2003**, *107*, 9964–9973.
[15] P. Tähtinen, A. Bagno, K. D. Klika, K. Pihlaja, *J. Am. Chem. Soc.* **2003**, *125*, 4609–4618.
[16] G. Barone, L. Gomez-Paloma, D. Duca, A. Silvestri, R. Riccio, G. Bifulco, *Chem. Eur. J.* **2002**, *8*, 3233–3239.
[17] G. Barone, D. Duca, A. Silvestri, L. Gomez-Paloma, R. Riccio, G. Bifulco, *Chem. Eur. J.* **2002**, *8*, 3240–3245.
[18] P. Cimino, L. Gomez-Paloma, D. Duca, R. Riccio, G. Bifulco, *Magn. Reson. Chem.* **2004**, *42*, S26–S33.
[19] P. Cimino, G. Bifulco, A. Evidente, M. Abouzeid, R. Riccio, L. Gomez-Paloma, *Org. Lett.* **2002**, *4*, 2779–2782.
[20] G. Bifulco C. Bassarello, R. Riccio, L. Gomez-Paloma, *Org. Lett.* **2004**, *6*, 1025–1028.
[21] a) D. Colombo, P. Ferraboschi, F. Ronchetti, L. Toma, *Magn. Reson. Chem.* **2002**, *40*, 581–588; b) A. Aiello, E. Fattorusso, P. Luciano, A. Mangoni, M. Menna, *Eur. J. Org. Chem.* **2005**, 5024–5030.
[22] M. A. Watson, P. Salek, P. Macak, M. Jaszuñski, T. Helgaker, *Chem. Eur. J.* **2004**, *10*, 4627–4639.
[23] a) C. Ochsenfeld, J. Kussmann, F. Koziol, *Angew. Chem.* **2004**, *116*, 4585–4589; *Angew. Chem. Int. Ed.* **2004**, *43*, 4485–4489.
[24] K. K. Baldridge, J. S. Siegel, *J. Phys. Chem. A* **1999**, *103*, 4038–4042.
[25] a) P. Csaszar, P. Pulay, *J. Mol. Struct.* **1984**, *114*, 31–34; b) Ö. Farkas, PhD (CsC) Thesis, Eötvös Loránd University and Hungarian Academy of Sciences, Budapest (Hungary), **1995**; c) Ö. Farkas, H. B. Schlegel, *J. Chem. Phys.* **1999**, *111*, 10806–10814.
[26] a) M. T. Cancès, B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3032–3041; b) M. Cossi, V. Barone, B. Mennucci, J. Tomasi, *Chem. Phys. Lett.* **1998**, *286*, 253–260; c) B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *106*, 5151–5158; this method is selected with the keyword "pcm" in Gaussian 03.
[27] 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, and J. A. Pople, Gaussian, Inc., Wallingford CT, **2004**.
[28] a) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652; b) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
[29] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
[30] 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; see also: http://www.scm.com
[31] a) A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100; b) J. P. Perdew, *Phys. Rev. B* **1986**, *33*, 8822–8824.
[32] S. Bourg, J.-M. Nuzillard, *J. Magn. Reson.* **1998**, *133*, 173–176.
[33] J. C. Carter, G. W. Luther III, T. C. Long, *J. Magn. Reson.* **1974**, *15*, 122–131.
[34] W. J. Chazin, L. D. Colebrook, J. T. Edward, *Can. J. Chem.* **1983**, *61*, 1749–1755.
[35] G. E. Martin, C. E. Hadden, R. C. Crouch, V. V. Krishnamurty, *Magn. Reson. Chem.* **1999**, *37*, 517–528.
[36] C. Benzi, O. Crescenzi, M. Pavone, V. Barone, *Magn. Reson. Chem.* **2004**, *42*, S57–S67.
[37] X. Miao, R. Freeman, *J. Magn. Reson. Ser. A* **1995**, *116*, 273–276.
[38] D. A. Craig, G. E. Martin, *J. Nat. Prod.* **1986**, *49*, 456–465.
[39] E. Kupce, J.-M. Nuzillard, V. S. Dimitrov, R. Freeman, *J. Magn. Reson. Ser. A* **1994**, *107*, 246–250.
[40] J. C. Cobas, V. Constantino-Castillo, M. Martín-Pastor, F. del Río-Portilla, *Magn. Reson. Chem.* **2005**, *43*, 843–848.
[41] R. A. E. Edden, J. Keeler, *J. Magn. Reson.* **2004**, *166*, 53–68.
[42] V. Blechta, F. del Río-Portilla, R. Freeman, *Magn. Reson. Chem.* **1994**, *32*, 134–137.
[43] T. Facke, S. Berger, *J. Magn. Reson. Ser. A* **1996**, *119*, 260–263.
[44] T. Parella, J. Belloc, *J. Magn. Reson.* **2001**, *148*, 78–87.
[45] B. L. Marquez, W. H. Gerwick, R. T. Williamson, *Magn. Reson. Chem.* **2001**, *39*, S499–S530.
[46] Y.-H. Shen, S.-H. Li, R.-T. Li, Q.-B. Han, Q.-S. Zhao, L. Liang, H.-D. Sun, Y. Lu, P. Cao, Q.-T. Zheng, *Org. Lett.* **2004**, *6*, 1593–1595.
[47] F. H. Allen, *Acta Crystallogr. Sect. B* **2002**, *58*, 380–388; deposition number CCDC-235021.
[48] S.-P. Yang, J.-M. Yue, *Org. Lett.* **2004**, *6*, 1401–1404.
[49] a) M. Cossi, O. Crescenzi, *J. Chem. Phys.* **2003**, *118*, 8863–8872; b) O. Crescenzi, M. Pavone, F. De Angelis, V. Barone, *J. Phys. Chem. B* **2005**, *109*, 445–453.

[50] a) B. Mennucci, J. M. Martínez, J. Tomasi, *J. Phys. Chem. A* **2001**, *105*, 7287–7296; b) C. Cappelli, B. Mennucci, S. Monti, *J. Phys. Chem. A* **2005**, *109*, 1933–1943.
[51] W.-G. Kim, J.-W. Kim, I.-J. Ryoo, J.-P. Kim, Y.-H. Kim, I.-D. Yoo, *Org. Lett.* **2004**, *6*, 823–826.
[52] W. Steglich, V. Hellwig, *Org. Lett.* **2004**, *6*, 3175–3177.

Received: December 16, 2005
Revised: January 17, 2006
Published online: May 8, 2006

