Modelling the influence of hydrogen bond network on chemical shielding tensors description. GIAO-DFT study of WALP23 transmembrane α-helix as a test case

Léa Rougier, Alain Milon, Valérie Réat, Franck Jolibois

ARTICLE TYPE www.rsc.org/xxxxxx | XXXXXXXX

Modelling the influence of hydrogen bonds network on chemical shielding tensors description. GIAO - DFT study of WALP23 transmembrane α-helix as a test case.

Léa Rougier,[a,b,c] Alain Milon,[b,c] Valérie Réat[b,c] and Franck Jolibois*[a,c]

Received (in XXX, XXX) Xth XXXXXXXXX 200X, Accepted Xth XXXXXXXXX 200X
First published on the web Xth XXXXXXXXX 200X
DOI: 10.1039/b000000x

Density Functional Theory (B3LYP/6-31G(d,p)) calculations of ^{15}N amide and ^{13}C carbonyl NMR chemical shielding tensors have been performed on **WALP23** trans-membrane α-helix peptide and compared to solid state NMR experiment performed on [$^{13}C_1$-Ala$_{13}$, ^{15}N-Leu$_{14}$] specifically labelled peptide powder sample. Using either theoretical result obtained on the whole peptide or experimental data as reference, several simplest chemical models have been explored in order to reduce the computational cost while maintaining good theoretical accuracy. From this study, it appears that the hydrogen bond (N-H…O=C) network that exists in the α-helix has a major influence on the chemical shielding tensor and more specifically on the carbonyl ^{13}C σ_{22} eigenvalue. We show that a small truncated **WALP_7** model is not adequate for $^{13}C_1$ NMR description. The application of an external electric field in order to model the hydrogen bond network allows calculating chemical shielding tensors with accurate eigenvalues while the associated eigenvectors are largely modified. Finally, a 23 residues polyglycine peptide that includes the Alanine and Leucine residues for which NMR parameters must be calculated is proposed as the chemical model. This model is sufficient to mostly reproduce the calculation performed on **WALP23** with major gain in computational time. Moreover, the application of an external electric field allows reaching the experimental accuracy for the determination of the eigenvalues.

Introduction

Accurate determination of "static" chemical shielding anisotropic (CSA) tensors components remains a challenge in the field of solid state NMR (SSNMR) applied to biological systems whereas these quantities can be directly related to structure, dynamic and/or orientation [2-4]. Not only isotropic [5, 6] but also anisotropic chemical shifts [7, 8] are useful experimental data to probe secondary structure and short or long range electrostatic interactions. Since few years, ^{13}C and ^{15}N chemical shielding tensors of labelled biomolecules have received increasing interest for structure and dynamic analysis purpose using SSNMR.

If the backbone only is considered, carbonyl $^{13}C_1$ and amide ^{15}N CSA tensors are not only sensitive to secondary structure but can reveal insight into the nature of the hydrogen bonding pattern that can exits locally. More particularly, the $^{13}C_1$ σ_{22} tensor eigenvalue that is almost collinear to the C=O bond vector strongly depends on the distance of the hydrogen bonding partner [9, 10] thus giving valuable information concerning the backbone geometry [11]. The same trend is also observed for ^{15}N CSA tensor that are sensitive to local geometry but also to the distance to the C=O hydrogen bond partner [12-15]. ^{13}C and ^{15}N CSA tensors are not only helpful for structure analysis, but can also be employed for the study of the dynamic and/or the orientation of peptides and proteins. Concerning ^{15}N CSA tensor, the main SSNMR experiment used for this purpose is the PISEMA technique which uses ^{15}N chemical shift and ^{1}H-^{15}N dipolar coupling to define the orientation of a peptide structure relative to an oriented lipid bilayer [16]. We have recently demonstrated that ^{13}C CSA tensors can also be used to determine orientation angle and order parameter of a biomolecule immerged in a lipid bilayer [4]. In this study, dynamically averaged ^{13}C CSA tensors were measured and combined to static ^{13}C CSA tensors obtained from quantum chemistry calculations in order to define the orientation and dynamical parameters of ergosterol relative to a 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine bilayer. A similar approach has also been proposed combining $^{13}C_1$, ^{15}N chemical shift anisotropies and ^{2}H quadrupolar couplings dynamically averaged by large amplitude movements for the study of the dynamic and orientation of a transmembrane peptide[17, 18]. These studies have highlighted the need of accurate static CSA tensors in order to understand the dynamic of a membrane compound. Furthermore, determination of the high resolution structure of transmembrane peptides or proteins is also constrained by the knowledge of such static CSA tensors [16, 19]. While static CSA tensor eigenvalues can be determined experimentally on ^{13}C and/or ^{15}N specifically labelled small dry peptides,

the associated eigenvectors i.e. the tensor orientation relative to the molecular frame may be difficult to obtain using experimental approaches[12]. However, such quantities can be "easily" accessible through the use of theoretical approaches such as quantum chemical methods.

Since several decades, quantum chemical approaches have been employed for the calculations of NMR parameters with real success in the fields of chemistry, physical chemistry and biochemistry. While the whole chemical shielding tensor is generally computed, the large majority of NMR theoretical studies concerned isotropic chemical shift values. The chemical shielding tensor eigenvalues are less analyzed and this tendency is even more pronounced for ^{15}N [20-26] and for $^{13}C_1$ carbonyl [24, 27, 28] CSA tensors compared to aliphatic ^{13}C ones[29-37]. Moreover, the overall agreement for the ^{15}N and $^{13}C_1$ CSA tensor calculations with experiment is not as great as for aliphatic C_α tensors. If the targeted information is the chemical shielding tensor eigenvectors, the theoretical orientation of CSA tensors show large variations from molecule to molecule [22-25].

Consequently, one of the remaining issues in chemical shielding tensor calculation is not the development of new *ab initio* or DFT methods but mainly the correct modelling of molecular systems that are used in theoretical calculation compared to experimental real molecule under investigation. More precisely, the local and non-local environmental effects must be correctly modelled in order to obtain accurate magnitudes as well as orientations of CSA tensors[38]. For instance, it is known that ^{15}N CSA tensors are not only sensible on primary and secondary structures but also on long range electrostatic interactions up to 10 Å [25].

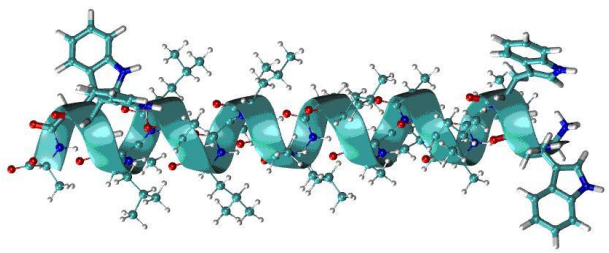


Fig. 1 Schematic representation of **WALP23** transmembrane peptide.

In this paper, we present a theoretical DFT investigation of the amide ^{15}N and carbonyl $^{13}C_1$ CSA tensors of a synthetic α-helical transmembrane peptide, namely **WALP23** (see Figure 1 for schematic representation)**,** which has been introduced and extensively studied in the group of A. Killian **(**see review [39]**)**. In a recent work, we have employed this peptide in order to develop a new strategy for the analysis of the orientation and dynamic of transmembrane peptide [18]. This study was based on the analysis of the ^{13}C and ^{15}N dynamically averaged CSA tensors measured by solid state NMR experiments. It required static $^{13}C_1$ and ^{15}N CSA tensors obtained by a quantum chemical approach. In the present work, from an experimental point of view, tensor amplitudes (or eigenvalues) have been measured on two different samples based on a $^{13}C_1$, ^{15}N specifically labelled peptide: dry WALP peptide and lyophilised WALP peptide inserted in detergent micelles. Using various chemical models, we have built an inexpensive and accurate strategy for the calculation of such NMR parameters. One of the most important results is that, on one hand, local hydrogen bond is necessary and sufficient for the accurate ^{15}N chemical shielding tensor calculation[40]. On the other hand, the whole hydrogen bond network has an important influence on the magnitude of the σ_{22} component of the $^{13}C_1$ tensor and must be explicitly taken into account during the calculation. The use of an external electric field is discussed and shown to improve the prediction of experimental data.

Experimental Section

Experimental procedures

$^{13}C_1$, ^{15}N isotopically labelled **WALP23** [$^{13}C_1$-Ala$_{13}$, ^{15}N-Leu$_{14}$] peptide (acetyl-GW$_2$(LA)$_8$LW$_2$A-NH2) was synthesized using Fmoc/tBu solid-phase synthesis[18]. The purity of the peptide was better than 95%. Two powder samples have been prepared for $^{13}C_1$ and ^{15}N static CSA eigenvalue evaluation. 6.5µmoles of **WALP23** peptide were dissolved in a mixture of tert-butanol/water (1:1) and lyophilized. The resulting powder was then packed in 3.2mm MAS rotor. For the second dry sample, 0.8µmoles of **WALP23** and 54.9 µmoles of dodecylphosphocholine (DPC) were mixed in 400µL of 2,2,2-trifluoroethanol (TFE). The organic solvent was evaporated under a nitrogen flow and further removed under vacuum overnight (ca. $1*10^{-2}$ mbar). The sample was hydrated with milli-Q water and lyophilized to yield a fluffy powder. The powder was packed in 3.2mm MAS rotor.

Solid-state magic angle spinning (MAS) experiments were carried out on a Bruker Avance narrow bore spectrometer operating at 700.13 MHz for ^1H. A Bruker 3.2mm MAS triple-tuned solenoid coil was used for $^{13}C_1$ and ^{15}N experiments. CP-MAS spectra were acquired using a ^1H excitation pulse length of 2 µs and a CP spin-lock field-strength of 50 kHz. The CP contact time was 2 ms and 1.25 ms for ^{13}C and ^{15}N spectra, respectively. The ^1H radio-frequency field-strength for heteronuclear two-pulse phase-modulation (TPPM) decoupling was 125 kHz during acquisition. The repetition delay was 2s. Spinning sideband ^{15}N and ^{13}C MAS spectra were recorded at 3 and 4 kHz, respectively. ^{13}C spectra were referenced to DSS and ^{15}N spectra to liquid ammonia, using the ^1H chemical shift of DSS in water as an external reference. The Static CSA eigenvalues δ_{11}, δ_{22} and δ_{33} (with $|\delta_{33}-\delta_{iso}| \geq |\delta_{11}-\delta_{iso}| \geq |\delta_{22}-\delta_{iso}|$ where δ_{iso}, the isotropic chemical shift) [41-43] were obtained by the analysis of spinning sidebands intensity. This analysis was performed using dmfit freeware (http://crmht-europe.cnrs-orleans.fr/dmfit/).

Eigenvalues determined for both samples are given in Table 1. In order to evaluate the influence of the chemical environment on both CSA values, dry powder of "pure" peptide and dodecylphosphoglycol (DPC) lyophilized micelles that contains WALP peptides have been employed (see previous paragraphs). The second sample allowed us to remove possible peptide – peptide interactions. DPC has been employed instead

of standard DMPC phospholipids in order to avoid difficulties arising from the lipids carbonyl ^{13}C resonances.

Quantum calculation

All calculations (geometry optimization and NMR properties determination) have been performed using the Gaussian03 suite of programs[44]. Because **WALP23** experimental structure has not been determined by standard methods, perfect canonical α-helix (φ = -58° and ψ = -47°) has been used as starting geometry for the optimization procedure. The molecular system being large (367 atoms), the nature of the stationary point found by geometry optimization procedure has not been verified by Hessien matrix calculation as usually. Three type of molecular systems have been investigated, the entire WALP peptide (noted **WALP23**), a model peptide reduced to 7 amino acids (noted **WALP_7**) and a model peptide of 23 residues where all amino acids were replace by Glycine except the specifically labeled ones (noted **polyGLY-AL**).

For the entire peptide (23 amino acids), geometry optimization have been performed using a 2 layers ONIOM approach [45, 46]. The high and low layers were defined according to experimental ^{13}C and ^{15}N labelling of **WALP23**. Because we are interested in calculating the NMR parameters of Alanine 13 carbonyl atom (noted Ala_{13} $^{13}C_1$) and of the Leucine 14 amide nitrogen atom (noted Leu_{14} ^{15}N), the principal structural feature that must be conserved is the intramolecular hydrogen bonds between one residue *i* and the residues *i+3* and *i-3* along the alpha helix. Consequently, the high layer of the ONIOM procedure was composed of residues 10 to 17, the cut being performed at the alpha carbon level. The low layer was constituted by the rest of the peptide. From a computational point of view, the high layer part was computed using the hybrid Density Functional Theory B3LYP method [47, 48] associated with a Pople type double ξ basis set augmented by polarization functions on all atoms (namely 6-31G(d,p)) [49] while the lower part was treated at the semi empirical AM1 level [50-52]. During the optimization process, Cartesian coordinates associated to the nuclei of the lower part were frozen in order to prevent suspicious geometrical deformation. Therefore, only the geometry associated to the high layer part (residues 10 to 17) was optimized. If the model peptide with 7 amino acids (namely **WALP_7**) is considered, geometry optimization (when performed) has been carried out at the B3LYP/6-31G(d,p) theoretical level, the Cartesian coordinates of the extremity residues being frozen for the same reason that previously exposed. For **polyGLY-AL** system, **WALP23** optimized structure has been used in order to build this new molecular system. All removed residue side chains have been replaced by hydrogen atoms. The additional hydrogen atoms have been placed on the C_β positions and the C_β-H distances have been adjusted to 1.07 Å. No further geometry optimization has been performed for this peptide.

NMR shielding tensors have been computed using the hybrid Density Functional Theory B3LYP method associated with a Pople type double ξ basis set augmented by polarization functions on all atoms (namely 6-31G(d,p)). NMR chemical shielding tensors have been computed using the Gauge Including Atomic Orbital method (GIAO) for the numerous advantages it offers [53-58]. According to the Haeberlen, Mehring and Spiess [41-43] convention, eigenvalues have been classified according to the following inequalities: $|\sigma_{iso} - \sigma_{33}| \geq |\sigma_{iso} - \sigma_{11}| \geq |\sigma_{iso} - \sigma_{22}|$ where σ_{iso} is the isotropic chemical shielding.

Results and Discussion

The ^{15}N and ^{13}C CSA tensors have been determined experimentally, by standard spinning side band analysis method. Two sample have been compared, pure peptide and peptide "diluted" in a DPC environment to assess possible intermolecular interaction effects. It appears that both samples give similar eigenvalues with a maximum difference of 1 and 0.5 ppm for ^{15}N and ^{13}C eigenvalues, respectively (Table 1). Consequently, it can be argued that no specific external interaction significantly influences CSA tensors eigenvalues of ^{15}N amide and ^{13}C carbonyl atoms located on the peptide backbone and that comparison with theoretical values based on an isolated peptide is relevant.

The first part of this work was devoted to the determination of the Ala_{13} $^{13}C_1$ and Leu_{14} ^{15}N NMR parameters of **WALP23**. First calculations have been performed on the entire peptide after geometry optimization (see materials and methods for details). Results are presented in Table 1 and exhibit a relatively good agreement when compared to experimental values. Comparison between experimental and theoretical isotropic chemical shift is dependent of the choice and the accuracy of the chemical shift reference and do not necessarily reflect the quality of tensor calculations. In order to avoid this problem, differences between chemical shielding eigenvalues and their associate isotropic chemical shielding have been calculated and will be the only data discussed thereafter. In the case of ^{13}C carbonyl CSA tensor, if absolute differences between theory and experiment are considered, the errors for σ_{ii}-σ_{iso} are relatively small (between 1 and 7 ppm). Moreover if the RMSD is considered, the overall agreement is rather good with a RMSD of 5.0 ppm for Ala_{13} $^{13}C_1$. From an experimental point of view, α-helix and β-sheet secondary structures exhibit significant differences if one consider ^{13}C carbonyl CSA tensors. In the case of poly-L-alanine peptide, a difference of 13 ppm has been experimentaly determined for σ_{22}-σ_{iso} depending on the secondary structure (14.7 and 1.9 ppm for α-helix and β-sheet respectively)[59, 60]. This difference is less pronounced for the two other tensor components (5 and 8 ppm for σ_{11}-σ_{iso} and σ_{33}-σ_{iso}, respectively). Compare to these experiments, our calculations reproduce correctly the general trend of ^{13}C carbonyl CSA tensor of α-helix secondary structure. In the case of ^{15}N amide, CSA tensor the difference between secondary structures is less pronounced and no tendency can be clearly extracted from experimental ^{15}N amide CSA tensor components[61]. In our case, the errors for σ_{ii}-σ_{iso} are the same than for ^{13}C atom (between 2 and 6 ppm) but with a smaller RMSD (4.5 ppm for Leu_{14} ^{15}N). It should be stated that our calculations were performed in vacuum and did not include any environment contribution such as others peptides or DPC molecules that may influence spectroscopic

data. One way to include these contributions would have been to use a QM/MM approach with an extended MM part including the surrounding molecules. This kind of approach, commonly employed in the field of biology to calculate properties of one fragment of a protein or to analyse the reactivity at a catalytic site, needs the knowledge of a three dimensional structure extracted from crystallographic data or classical molecular dynamics simulations. However, because no structure already exists concerning our specific samples, the calculation starting from scratch would have been too expensive for the present work. In conclusion, on can consider that theoretical eigenvalues are calculated with enough accuracy (4% and 3.5% errors relative to the ^{13}C and ^{15}N tensor experimental spans (tensor span = σ_{33}-σ_{11})) and can then safely be used for the study of the dynamic and orientation of a transmembrane peptides [17, 18].

Table 1 Experimental ssNMR and theoretical (see text for details) $^{13}C_1$ and ^{15}N chemical shielding tensors characteristic (in ppm) for **WALP23** specifically labelled at position Ala 13 ($^{13}C_1$) and Leu 14 (^{15}N). σ_{iso} is the isotropic chemical shielding. σ_{ii}-σ_{iso} (i=1-3) have been calculated in order to avoid isotropic reference problem for theoretical and experimental values. |*Max Error*| (in ppm) is the maximum absolute value of error found for the eigenvalue indicated in the round brackets. *RMSD* (in ppm) is the root mean square difference calculated for each atom using the three σ_{ii}-σ_{iso} values.

^{13}C	σ_{11} - σ_{iso}	σ_{22} - σ_{iso}	σ_{33} - σ_{iso}	|Max. Error|	RMSD
Experimental (dry peptide)	-67.7	-16.9	84.6		
Experimental (peptide/DPC)	-67.3	-16.8	84.1		
WALP23					
In vacuo	-66.4	-11.4	77.8	6.8 (σ_{33})	5.0
With field (20 10^{-4} a.u.)	-63.1	-16.3	79.4	5.2 (σ_{11})	3.8
WALP_7					
In vacuo	-72.8	-1.1	73.9	15.8 (σ_{22})	11.4
With field (140 10^{-4} a.u.)	-63.4	-14.9	78.3	6.2 (σ_{33})	4.3
polyGLY					
In vacuo	-71.8	-2.8	74.6	14.1 (σ_{22})	10.2
polyGLY-AL	-68.0	-9.4	77.4	7.5 (σ_{33})	5.9
polyGLY-AL +field (20 10^{-4} a.u.)	-65.2	-13.5	78.7	5.9 (σ_{33})	4.0
^{15}N	σ_{11} - σ_{iso}	σ_{22} - σ_{iso}	σ_{33} - σ_{iso}	|Max. Error|	RMSD
Experimental (dry peptide)	62.8	41.9	-104.7		
Experimental (peptide/DPC)	63.4	42.3	-105.7		
WALP23					
In vacuo	60.4	38.4	-98.8	6.9 (σ_{33})	4.5
With field (20 10^{-4} a.u.)	60.6	42.1	-102.7	3.0 (σ_{33})	2.0
WALP_7					
In vacuo	61.3	37.2	-98.5	7.2 (σ_{33})	4.9
With field (140 10^{-4} a.u.)	62.9	46.1	-108.9	4.3 (σ_{33})	3.2
polyGLY					
In vacuo	58.6	42.9	-101.5	4.8 (σ_{11})	3.4
polyGLY-AL	61.1	37.7	-98.8	6.9 (σ_{33})	4.6
polyGLY-AL +field (20 10^{-4} a.u.)	60.4	41.5	-101.8	3.9 (σ_{33})	2.6

In an effort to reduce computational time and to find a simpler molecular model allowing the reproduction of all experimental NMR parameters of **WALP23** we tried using a smaller peptide consisting of 7 amino acids (**WALP_7,** see materials and methods for details). NMR parameters calculations reveal a poorer agreement between theoretical and experimental values if ^{13}C is considered (See Table 1). More particularly, Ala_{13} $^{13}C_1$ σ_{22}-σ_{iso} and σ_{33}-σ_{iso} values exhibit an error almost twice larger than the maximum error found for **WALP23**. The same tendency is also observed for the RMSD value (11.4 ppm) compared to the one calculated for **WALP23** (5.0 ppm). On the other hand, the discrepancy between **WALP_7** and **WALP23** molecular model is considerably smaller in the case of ^{15}N atom with a RMSD value compared to experiment equal to 4.9 and 4.5 for **WALP_7** and **WALP23** respectively. It can be concluded that this simpler molecular model and more specifically the inclusion of the local hydrogen bonding pattern alone is not sufficient to reproduce the entire peptide and that some information has been lost by removing the 16 remaining amino acids.

In order to recover part or totality of this information, two theoretical physico-chemical perturbations have been applied on **WALP_7**. The first one has consisted in simulating the rest of the peptide (16 amino acids) using a background charge distribution. For the second perturbation, a homogenous electric field has been applied along the helix axis.

Background charge perturbation.

One way to correct the **WALP_7** molecular model was to simulate the part of the molecule that has been removed from the entire peptide by using a background charge distribution. In this context, atoms of this part of the peptide have been replaced by partial atomic charges located at the nuclei Cartesian coordinates. Three type of atomic charge models have been employed: Mulliken type atomic charges, natural atomic charges extracted from Natural Population Analysis (noted NBO [62]) and MK atomic charges which are charges fitted to reproduce the molecular electrostatic

potential at a number of points selected around the molecule according to the Merz-Singh-Kollman scheme [63]. These three type of charges were extracted from a DFT B3LYP/6-31G(d,p) calculation performed on **WALP23**. It is well known that Mulliken type atomic charges are strongly dependent on the basis set and these values must be treated with care. However, the two other types of atomic charge are less sensitive to the choice of the basis set. Thus, results associated to these two type of atomic charges can be considered more reliable. ^{13}C and ^{15}N chemical shift tensors parameters evolution compared to experiment calculated for **WALP_7**, **WALP23** and **WALP_7** surrounded by charge distribution are presented in Figure 2. Concerning ^{15}N nucleus, charge distribution has little effect on discussed NMR parameters, leaving these values closed to **WALP23** as already obtained for **WALP_7** alone. When Ala_{13} $^{13}C_1$ is considered, charge distribution displays a more significant influence on NMR parameters. While σ_{11}-σ_{iso} and σ_{33}-σ_{iso} are less perturbed by the presence of charge distribution, σ_{22}-σ_{iso} is clearly but not sufficiently improved, the difference between theory and experiment ranging from 15.8 ppm (**WALP_7** alone) to 10.4 ppm (**WALP_7** with Mulliken charges) compare to 5.5 ppm for **WALP23**. However, if partial charges are calculated for atoms of Ala_{13} $^{13}C_1$ and Leu_{14} ^{15}N using Mulliken approximation (Table 2), sum of atomic partial charges in N-H bond (that reflects the charge repartition inside the covalent bond) does not vary when applying charge distribution (difference compare with **WALP23** remains equal to 10%). On the other hand, while for **WALP_7** this sum is 50% larger than for **WALP23** in the case of C=O bond, the application of a charge distribution reduces this discrepancy to about 25%. Thus, it appears that the background charge distribution has a real influence on the polarization of the hydrogen bond and more specifically on its carbonyl part but is not sufficiently accurate to represent the extremities of the peptide in order to reproduce NMR parameters of the entire peptide.

Table 2 Mulliken partial charges of the atoms of Ala_{13} $^{13}C_1$ Leu_{14} ^{15}N as a function of chemical models. **WALP_7** means *in vacuo* **WALP_7**. NBO, MK, Mulliken mean **WALP_7** with background charge distribution. **WALP23** means *in vacuo* whole peptide.

Alanine 13	WALP_7	NBO	MK	Mulliken	WALP23
C_1	0.617	0.617	0.621	0.617	0.61
O	-0.532	-0.536	-0.545	-0.547	-0.55
C_1=O	0.085	0.081	0.076	0.070	0.057
Leucine 14	WALP_7	NBO	MK	Mulliken	WALP23
N	-0.556	-0.556	-0.554	-0.556	-0.55
H	0.307	0.308	0.305	0.312	0.33
N-H	-0.249	-0.248	-0.249	-0.244	-0.225

These observations concerning H-bond polarization associated to the analysis of the behaviour of chemical shielding tensor discussed above highlight the fact that good modelling of hydrogen bond type interactions and more specifically the consideration of the whole alpha helix hydrogen bond network has to be taken into account in order to determine accurate NMR chemical shielding tensors. For this purpose, we propose to apply a uniform electric field in order to reproduce the influence of the hydrogen bond network. From a physical point of view, this idea is similar to the one proposed by Torri H theoretical work on retinal molecule [64]. In this study, an electric field is applied along the principal axis of the molecule in order to reproduce the effects on its electronic absorption maximum of the dipolar residues that surround the retinal inside the opsin protein. In should be stressed that in our approach the role of the external electric field is not to reproduce the external environment but rather to model the effect on bonds polarization of the whole hydrogen bonds network that has been lost when shortening WALP23 into WALP_7. This approach is different from others studies that employ atomic charges or charge dipoles in order to reproduce specific effect arising from charges or polar groups in proteins or to generate dipolar perturbations [65-71].

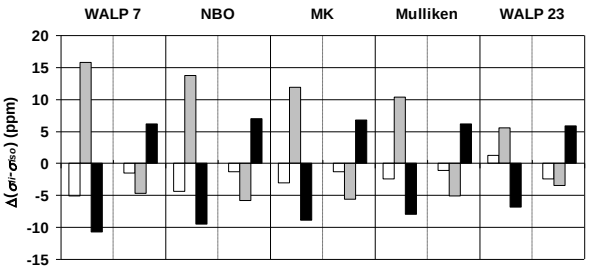


Fig. 2 Ala_{13} $^{13}C_1$ (left) and Leu_{14} ^{15}N (right) σ_{ii}-σ_{iso} (*i=1-3*) difference between theory and experiment as a function of punctual charge model used for charge distribution perturbation of **WALP_7** ($\Delta[\sigma_{ii}-\sigma_{iso}] = (\sigma_{ii}-\sigma_{iso})_{theo} - (\sigma_{ii}-\sigma_{iso})_{exp}$). In white *i=1*, grey, *i=2*, black *i=3*. **WALP23** values are shown for comparison purpose.

Uniform electric field perturbation

In order to essentially try to reproduce the influence of the alpha helix hydrogen bonds network, the electric field has been applied to **WALP_7** along the inertia axis of the helix fragment, from N terminal to C terminal atoms. Electric field intensities have been set from 0 to 300 10^{-4} atomic unit (noted a.u. with 1 a.u. = 514.10^9 $V.m^{-1}$). To estimate the influence of the electric field on the polarisation of the reduced hydrogen bonds network of **WALP_7**, the average value per amino acids of the dipole moment has been calculated and is given as a function of the electric field intensity in Figure 3. While the dipole moment value is definitively too low when **WALP_7** is considered (3.7 Debye / amino acids), the electric field significantly induces a regular increase of this value that can reach and goes beyond the one obtained for **WALP23** (5.2 Debye / amino acids) for electric field intensities larger than 100 10^{-4} a.u.. This first observation clearly demonstrates the effect of the electric field on the polarisation of the functionnal groups involved in the hydrogen bonds and thus as a possible way to model the hydrogen bond network existing in the entire peptide. However, if the electric field intensity become too large (superior to 200 10^{-4} for **WALP_7**), the dipole moment evolution as a function of electric field intensity presents a discontinuity. Before this electric field value, the dipole moment evolves linearly with a slope equals to 155.8

Debye/(electric field a.u.) for **WALP_7**. After this electric field intensity value, the slope becomes larger (282.1 Debye/(electric field a.u.)). This reflects the fact that with such intensity, charge separations become too large reflecting the unphysical nature of this calculation for such "high" intensity electric field.

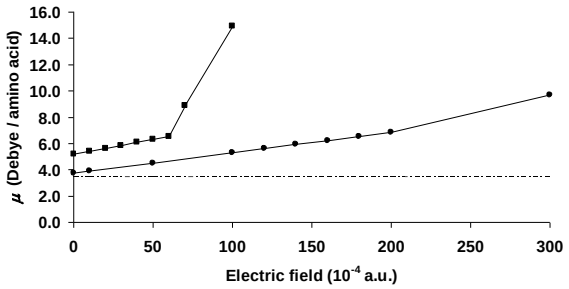


Fig.3 Dipolar moment (Debye) of **WALP_7** (circles) and **WALP23** (squares) calculated per amino acid as a function of the applied electric field intensity. The dashed line represents the usual theoretical dipole moment associated to the peptide group (3.45 D)[1]

The evolutions of ^{15}N and $^{13}C_1$ chemical shielding tensor parameters (σ_{ii}-σ_{iso}) compared to experiment are given in Figure 4 as a function of electric field intensity. If $^{13}C_1$ nucleus is considered, while, as previously stated, the main problem arises from σ_{22}-σ_{iso} and σ_{33}-σ_{iso} values for **WALP_7**, the application of the electric field clearly improves these values compared to experiment when electric field intensity is set between 120 10^{-4} and 180 10^{-4} a.u.. It must be noted that calculation using an electric field intensity set to 100 10^{-4} a.u. gives three eigenvalues almost identical to the ones obtained for **WALP23**. In the case of ^{15}N nucleus, the effect of the electric field on chemical shielding tensor parameters seems less sensitive, NMR calculations on **WALP_7** being already in relatively good agreement with experiment. While the absolute values of σ_{ii}-σ_{iso} are modified, their differences with experiment remain in an acceptable range of ±6 ppm. The best agreement over the 6 σ_{ii}-σ_{iso} values compared to experiment data has been found for an electric field of 140 10^{-4} a.u. (see values in Table 1) and is associated to the smallest total RMSD calculated using both atoms of 3.8 ppm. However, when compared to **WALP23**, a large deviation of all the ^{15}N σ_{ii}-σ_{iso} is observed when an electric field is applied. Consequently, while it is possible to reproduce **WALP23** ^{13}C CSA tensor with an appropriate electric field applied on **WALP_7**, such perturbation has an inverse results on ^{15}N CSA censor.

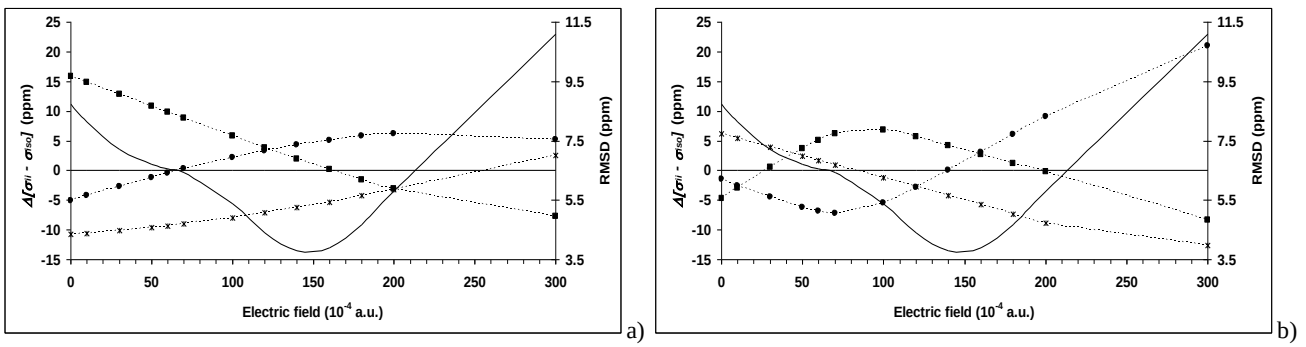


Fig.4 Ala_{13} $^{13}C_1$ (a) and Leu_{14} ^{15}N (b) σ_{ii}-σ_{iso} (*i=1-3*) difference between theory and experiment as a function of electric field applied to **WALP_7 (**$\Delta[\sigma_{ii}$-$\sigma_{iso}] = (\sigma_{ii}$-$\sigma_{iso})_{theo}$ - $(\sigma_{ii}$-$\sigma_{iso})_{exp}$). Circle *i=1*, square, *i=2*, star *i=3*. In solid line is represented the total RMSD between theory and experience calculated using the 6 σ_{ii}-σ_{iso} values as a function of electric field.

In order to highlight the effect of the electric field on the $^{13}C_1$ and ^{15}N chemical shielding tensors eigenvalues, we have analyzed the eigenvalues variation as a function of electric field intensity taking **WALP_7** eigenvalues without electric field as references. It can be stressed that the most influenced values are σ_{22} and σ_{33} for $^{13}C_1$ and ^{15}N, respectively. This can be explained by the fact that the tensor eigenvectors associated to these eigenvalues are more or less collinear to the direction of the applied electric field. In the case of ^{15}N, this eigenvector is about 20° deviated from the N-H bond direction while for $^{13}C_1$, the eigenvector direction is almost collinear to C=O bond (Figure 5). As a consequence, the applied electric field directly polarizes the electronic density along C=O and N-H bonds, the effect of the entire H-bond network being partly recreated. Thus, the electric field effect on $^{13}C_1$ σ_{22} eigenvalue is more important than on ^{15}N σ_{33}.

Considering that DFT methods and/or the atomic orbital basis set employed may have difficulties in correctly representing weak electrostatic interactions such as hydrogen bond and/or that geometry optimization has not been performed on the whole peptide, it can be argued that the application of an electric field on **WALP_7** has two consequences: modelling the hydrogen bond network of the entire peptide and correcting the theoretical method (method, basis set, optimization) for weak interactions. Furthermore, if the type of sample used for NMR experimental measurement is considered (powder sample or lyophilized DPC micelles), dipole – dipole intermolecular interactions have a small influence on the polarization of the hydrogen bond network and thus slightly modify NMR parameters compared to an isolated peptide in gas phase which is the computational condition (See experimental section). In order to verify these assumptions, NMR calculations have been performed on **WALP23** in the presence of an applied electric field. The same tendency,

improvement of NMR parameters with the application of electric field, is observed for **WALP23**, thus confirming the correction of our theoretical model by the use of an electric field (See Supp. Info). In this case, because the hydrogen bond network is explicitly present in the molecular structure, the intensity necessary for a good description of chemical shielding eigenvalues compared to experiment is less important, the best value being 20 10^{-4} a.u. compared to 140 10^{-4} a.u. for **WALP_7**. With these electric field intensities, differences of σ_{ii}-σ_{iso} compared to experiment are less than 6 ppm, the RMSD for the 6 eigenvalues being equal to 3.1 ppm and 3.8 ppm for **WALP23** and **WALP_7,** respectively. Dipolar moment evolution as a function of electric field has also been analysed in the case of WALP23 (see Figure 3). As previously observed with WALP_7, an increase of the dipolar moment per amino acid is observed with a slope higher than for WALP_7 (222.7 Debye/(electric.field a.u.)). Similarly, for a high intensity electric field (superior to 60.10^{-4} a.u.), one observes a discontinuity that indicates an unrealistic polarisation of all the bonds. With intensity set to 20.10^{-4} a.u. (the one for which best NMR values are obtained compared to experiment), the dipole moment per amino acid is close to the one determined with an intensity set to 140.10^{-4} a.u. for WALP_7 that gives the best NMR values compared to experiment.

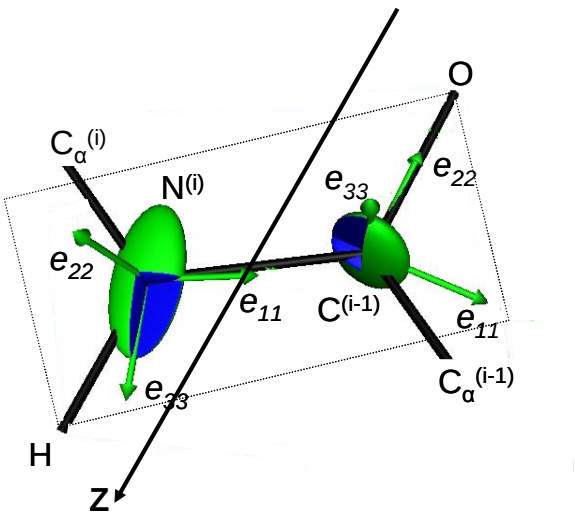


Fig.5 Orientation of eigenvectors of the 2 CSA tensors (Ala_{13} $^{13}C_1$ and Leu_{14} ^{15}N) and direction of the applied electric field.

Influence of peptide planes number.

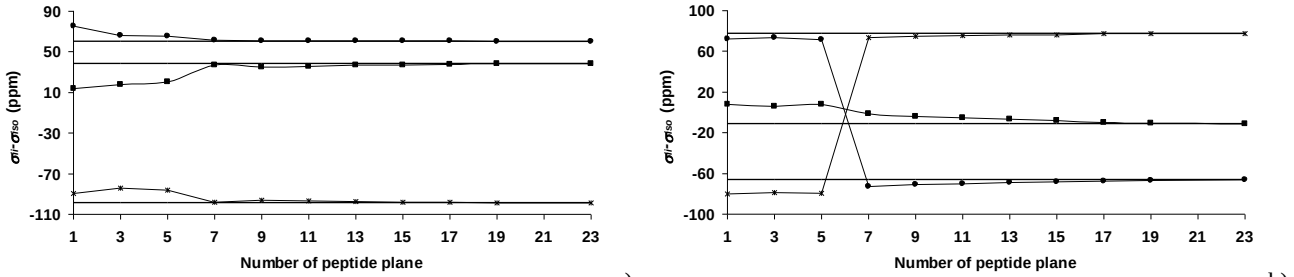
In order to obtain further insight on the influence of the hydrogen bond network extent, we have performed NMR calculations on WALP peptides of increasing size. Our chemical models were based on Ala13-Leu14 fragment (Me-CO-Ala13-Leu14-NH-Me). One peptide plane has been added at each side of this initial peptide, increasing its size by 2 peptide planes each time until we reach the whole **WALP23**. The **WALP23** structure has been used and no geometry optimization has been performed in order to evaluate only the hydrogen bond network effect on NMR parameters. If ^{15}N σ_{ii}-σ_{iso} NMR parameters are considered (Figure 6a), it appears that a minimum of 7 peptide planes is necessary for accurate results compared to **WALP23** data. This is consistent with our previous results obtained for **WALP_7**. Thus, in the case of ^{15}N nucleus, the local hydrogen bond is necessary and sufficient to reproduce the spectroscopic observables and no long range interaction affects the calculation of these parameters[40]. In the case of $^{13}C_1$ σ_{ii}-σ_{iso} NMR parameters, the situation is more complicated (6b). According to our ordering of chemical shielding eigenvalues relative to the isotropic value, one can see that the first hydrogen bond involving directly the carbonyl group is absolutely necessary for a correct description of the tensor. From one to five peptide planes, σ_{11} and σ_{33} are inverted thus leading to a chemical shielding anisotropy (σ_{33}-σ_{iso}) with an inverse sign (positive instead of negative). Moreover, while a number of peptide planes higher than seven, does not significantly improve σ_{11} and σ_{33} eigenvalues, σ_{22} eigenvalue needs a more extended hydrogen bond network (at least two helical turns) in order to achieve a good computational accuracy. Thus, the cooperativity, that proceeds from interactions among residues and that has been highlighted previously [72], has its most important effect on $^{13}C_1$ σ_{22} eigenvalue that is associated to e_{22} eigenvectors collinear to the hydrogen bond network.

Simplified WALP model: Polyglycine.

Previous calculations have demonstrated that one of the most important structural feature that must be conserved in the chemical model is the hydrogen bonds network. Indeed, while the alpha helix structure is conserved in the WALP_7 model, chemical shift tensor are not well reproduced. The hydrogen bond network not only influences the entire structure in a specific secondary structure but also provides bonds polarisation that affects chemical shift tensors principal components. Consequently, a simpler model of **WALP23** has been proposed by removing all the amino acids side chains in a first approximation. As expected, chemical shielding eigenvalues are correctly calculated for ^{15}N compared to experiment (see Table 1) while a poorer agreement is obtained for $^{13}C_1$ eigenvalues and more specifically for the σ_{22}-σ_{iso} value. In this case, an error of 14.1 ppm relative to experiment is obtained, compared to 15.8 ppm for **WALP_**7.

In order to correct this chemical model, Alanine 13 and Leucine 14 side chains were kept, the rest of the polypeptide amino acids being replaced by glycine residues (noted **polyGLY-AL**). In this case, chemical shielding eigenvalues are very close to the one obtained for *in vacuo* **WALP23** (max error equal to 7.5 ppm compared to 6.8 ppm, respectively). The Ala_{13} $^{13}C_1$ σ_{22}-σ_{iso} value has been considerably improved compared to calculation on

polyglycine (difference compared to experiment of 7.5 ppm instead of 13.4 ppm). Finally, on this last chemical model (polyglycine with Ala_{13} and Leu_{14}), an external electric field has been applied using an intensity of 20 10^{-4} a.u., this intensity being the best one obtained for **WALP23** (see previous paragraph). With this external perturbation, theoretical calculations are also improved compared to in *vacuo* calculations. ^{15}N σ_{ii}-σ_{iso} NMR parameters are in excellent agreement with experiment (RMSD (^{15}N) = 2.6 ppm).



X

Fig. 6 Leu_{14} ^{15}N (a) and Ala_{13} $^{13}C_1$ (b) evolution of σ_{ii}-σ_{iso} (*i=1-3*) as a function of number of peptide planes (helix size). Circle *i=1*, square, *i=2*, star *i=3*. Horizontal lines are theoretical **WALP23** references.

Paradoxically, it appears that compared to experiment, the simple polyglycine model gives better results for ^{15}N nuclei than **WALP23**. This highlights the fact that side chains may play an significant role for chemical shift tensor properties and that the theoretical **WALP23** model may not correctly reproduces these side chains contributions. It is suspected that side chain dynamics, i.e. rotation about C-C bonds, may play an role and influence the ^{15}N chemical shift tensor. This dynamic aspect is under investigation.

Table 3 Ala_{13} $^{13}C_1$ and Leu_{14} ^{15}N eigenvector orientation relative to the molecular frame determined without and with the optimal applied electric field (see text for details) for **WALP_7**, **WALP23** and **polyGLY-AL**.

	$\alpha(^{13}C_1)$	$\beta(^{13}C_1)$	$\gamma(1^{3}C_1)$	$\alpha(^{15}N)$	$\beta(^{15}N)$	$\gamma(^{15}N)$
WALP_7						
In vacuo	36.4	4.2	2.8	20.3	16.6	18.7
With field (140 10^{-4} a.u.)	47.7	18.3	11.9	53.0	52.5	18.1
WALP23						
In vacuo	31.8	2.4	2.0	20.7	17.7	19.5
With field (20 10^{-4} a.u.)	36.2	5.3	4.5	21.9	19.0	19.6
polyGLY-AL						
In vacuo	32.5	2.4	2.4	19.2	15.7	19.3
With field (20 10^{-4} a.u.)	35.3	4.5	4.1	17.5	13.3	19.2

Tensor orientation analysis.

As stated in the introduction, quantum chemical calculations are essential for the determination of the chemical shielding orientation with respect to molecular frame. Thus, in order to check our methodological approach, the description of chemical shift tensor eigenvectors in the molecular frame was determined. For practical reasons, the molecular frame has been chosen relative to the Ala_{13}-Leu_{14} peptide plane for all chemical models in use: *x-axis* is collinear with N-H, *z-axis* is normal to the peptide plane and *y-axis* is almost collinear to the N-C_α(*i-1*) vector. For ^{15}N, α_N is the angle between e_{11} and the N-C_α peptide bond, β_N the angle between e_{22} and the *z-axis* and γ_N, the angle between e_{33} and the N-H bond where e_{ii} is the eigenvectors associated to the σ_{ii} eigenvalue (see Figure 5 for e_{ii} representation). For $^{13}C_1$, α_C is the angle between e_{11} and the N-C_α peptide bond, β_C the angle between e_{22} and the C_1-O bond and γ_C, the angle between e_{33} and the *z-axis*. Because no experimental values for eigenvectors orientation exist for **WALP23** polypeptide, the *in vacuo* calculations has been taken as references for each molecular system. The angles obtained for *in vacuo* are in agreement with the standard values used to describe ^{15}N and $^{13}C_1$ chemical shift tensors eigenvectors [14, 73, 74]. Particularly, for ^{15}N, the angle between e_{33} and the N-H bond is equal to 19° and the e_{11} is tilted by about 19-20° from the peptide plane. For $^{13}C_1$, α_C angle ranges between 32° and 36°, e_{22} is almost collinear to the carbonyl C_1-O bond and e_{33} is perpendicular to the peptide plane. If *in vacuo* non perturbed chemical models are considered, chemical shift tensors orientations in the molecular frame are almost identical for **WALP23**, **WALP_7** and **polyGLY-AL** the maximum difference being equal to 5° for α_C compared to **WALP23** value (see Table 3).

However, when external electric field is applied, several angle values are significantly modified as a function of the field intensity. For clarity reason, only *in vacuo* results and calculations performed using the optimal electric field intensity (140 10^{-4} for **WALP_7** and 20 10^{-4} for **WALP23** and **polyGLY-AL)** are presented in Table 3. Others results with several electric field intensities are given in

supplementary materials, (Figure S1). It should be stressed that for the three molecular models, the ^{15}N e_{33} eigenvector orientation is almost not modified by the application of the electric field. As stated in previous paragraphs, an electric field with an intensity of 140 10^{-4} a.u. is necessary to obtain good eigenvalues for **WALP_7**. Nevertheless, using such intensity for the applied electric field modifies the eigenvectors orientation by up to 16° for $^{13}C_1$ and 35° for ^{15}N. Thus, while this electric field is adequate for correcting chemical shielding eigenvalues, it is not suitable for correcting chemical shielding eigenvectors. In **WALP23** or **polyGLY-AL** model, the intensity of the electric field (20 10^{-4} a.u.) applied for chemical shift tensor correction is less important. Consequently, the deviation of the eigenvector orientation from unperturbed chemical shielding tensors is rather small (max difference = 4°) and one can safely use such computational strategy in order to obtain accurate chemical shielding eigenvalues and eigenvectors with minimal computational effort.

Conclusions

In this paper, an extensive computational study of the NMR properties at the GIAO – DFT level has been proposed on **WALP23** α-helix transmembrane peptide. For this purpose, solid state NMR experimental data of a dry specifically labelled peptide ($^{13}C_1$ Alanine 13 and ^{15}N Leucine 14) have served as references in order to calibrate an accurate and "inexpensive" computational strategy. While calculation on the whole peptide reveals good agreement compared to experiment, this molecular model is too large to envisage performing calculations on this system routinely.

In order to reduce computational time, a smaller peptide **WALP_7** has been proposed that includes residues for which NMR parameters must be calculated and on helical step in order to include local hydrogen bond associated to the peptide bond chemical groups (amide and carbonyl groups). Our calculations clearly show that this molecular model is sufficient to reproduce ^{15}N amide chemical shift tensor while the $^{13}C_1$ carbonyl chemical shielding tensor is calculated with less accuracy. It is shown that this discrepancy is due to the bad reproduction of the hydrogen bonds network created by the whole peptide backbone that adopts an α-helix type secondary structure. Application of an external electric field is proposed in order to model the hydrogen bond network effects on local hydrogen bond. While this approach allows calculating chemical shielding tensors with accurate eigenvalues, the associated eigenvectors are largely modified. Moreover, at this stage of the study, no control on the choice of the electric field intensity is possible. It should be necessary to perform similar studies on peptides of various size in order to highlight possible correlation between the applied electric field intensity and the peptide size. Consequently, **WALP_7** peptide is not a good alternative for reproducing NMR characteristics of **WALP23**. In order to explicitly include the hydrogen bond network, a 23 residues polyglycine that includes the Alanine and Leucine residues for which NMR parameters must be calculated is proposed as the chemical model. It is shown that this molecular model associated (or not) to a low intensity external electric field allows calculating ^{15}N and $^{13}C_1$ CSA tensors (eigenvalues and eigenvectors) with a quantitative accuracy compared to experiment. These calculations are carried out with a computational cost divided by almost one order of magnitude compared to the calculations on the whole peptide.

In the future, the computational strategy presented in this paper will allow performing NMR calculations on larger and more complicated peptide in terms of sequence. However, further studies concerning the influence of amino acid side chains (more particularly, polar and charged residues) remain necessary in order to estimate to what extent these residues perturb chemical shielding tensors of ^{15}N amide and ^{13}C carbonyl atoms that are located on the peptide backbone. This strategy will also permit to carry out, within a reasonable time, calculations on a large number of molecular structures that can be extracted from molecular dynamic simulations, for example, in order to introduce fast nuclear motions in the theoretical description of chemical shielding tensors.

Acknowledgements

Authors wish to acknowledge financial supports from "Conseil Régional Midi-Pyrénées" and PRES "Université de Toulouse" for Léa Rougier Ph.D. funding and the CALcul en MIdi-Pyrénées (CALMIP) computing center for generous allocations of computer time.

Notes and references

1 D. Sengupta, R. N. Behera, J. C. Smith and G. M. Ullmann, *Structure*, 2005, **13**, 849-855.

2 B. J. Wylie and C. M. Rienstra, *J. Chem. Phys.*, 2008, **128**, 052207.

3 A. C. Dedios, J. G. Pearson and E. Oldfield, *Science*, 1993, **260**, 1491-1496.

4 O. Soubias, F. Jolibois, S. Massou, A. Milon and V. Reat, *Biophys. J.*, 2005, **89**, 1120-1131.

5 S. Spera and A. Bax, *J. Am. Chem. Soc.*, 1991, **113**, 5490-5492.

6 E. G. Paul and D. M. Grant, *J. Am. Chem. Soc.*, 1963, **85**, 1701.

7 N. Tjandra and A. Bax, *J. Am. Chem. Soc.*, 1997, **119**, 9576-9577.

8 D. A. Case, *Curr. Opin. Struct. Biol.*, 1998, **8**, 624-630.

9 T. Kameda, N. Takeda, S. Kuroki, H. Kurosu, S. Ando, I. Ando, A. Shoji and T. Ozaki, *J. Mol. Struct.*, 1996, **384**, 17-23.

10 Y. F. Wei, D. K. Lee and A. Ramamoorthy, *J. Am. Chem. Soc.*, 2001, **123**, 6118-6126.

11 B. J. Wylie, L. J. Sperling, H. L. Frericks, G. J. Shah, W. T. Franks and C. M. Rienstra, *J. Am. Chem. Soc.*, 2007, **129**, 5318.

12 K. W. Waddell, E. Y. Chekmenev and R. J. Wittebort, *J. Am. Chem. Soc.*, 2005, **127**, 9030-9035.

13 E. Y. Chekmenev, Q. W. Zhang, K. W. Waddell, M. S. Mashuta and R. J. Wittebort, *J. Am. Chem. Soc.*, 2004, **126**, 379-384.

14 C. J. Hartzell, M. Whitfield, T. G. Oas and G. P. Drobny, *J. Am. Chem. Soc.*, 1987, **109**, 5966-5969.

15 G. S. Harbison, L. W. Jelinski, R. E. Stark, D. A. Torchia, J. Herzfeld and R. G. Griffin, *J. Magn. Reson.*, 1984, **60**, 79-82.
16 T. A. Cross and S. J. Opella, *Curr. Opin. Struct. Biol.*, 1994, **4**, 574-581.
17 S. D. Cady, C. Goodman, C. D. Tatko, W. F. DeGrado and M. Hong, *J. Am. Chem. Soc.*, 2007, **129**, 5719-5729.
18 A. Holt, L. Rougier, V. Reat, F. Jolibois, O. Saurel, J. Czaplicki, J. A. Killian and A. Milon, *Biophys. J.*, 2009, **Accepted**.
19 W. Mai, W. Hu, C. Wang and T. A. Cross, *Protein Science*, 1993, **2**, 532-542.
20 H. B. Le and E. Oldfield, *J. Phys. Chem.*, 1996, **100**, 16423-16428.
21 H. B. Le and E. Oldfield, *J. Biomol. NMR*, 1994, **4**, 341-348.
22 A. Poon, J. Birn and A. Ramamoorthy, *J. Phys. Chem. B*, 2004, **108**, 16577-16585.
23 J. R. Brender, D. M. Taylor and A. Ramamoorthy, *J. Am. Chem. Soc.*, 2001, **123**, 914-922.
24 A. E. Walling, R. E. Pargas and A. C. deDios, *J. Phys. Chem. A*, 1997, **101**, 7299-7303.
25 C. Scheurer, N. R. Skrynnikov, S. F. Lienin, S. K. Straus, R. Bruschweiler and R. R. Ernst, *J. Am. Chem. Soc.*, 1999, **121**, 4242-4251.
26 M. Mirzaei and N. L. Hadipour, *Struct. Chem.*, 2008, **19**, 225-232.
27 A. C. Dedios and E. Oldfield, *J. Am. Chem. Soc.*, 1994, **116**, 5307-5314.
28 X. Chen and C.-G. Zhan, *Theochem-J. Mol. Struct.*, 2004, **682**, 73-82.
29 J. Bim, A. Poon, Y. Mao and A. Ramamoorthy, *J. Am. Chem. Soc.*, 2004, **126**, 8529-8534.
30 E. Czinki, A. G. Csaszar, G. Magyarfalvi, P. R. Schreiner and W. D. Allen, *J. Am. Chem. Soc.*, 2007, **129**, 1568-1577.
31 S. Wi, H. H. Sun, E. Oldfield and M. Hong, *J. Am. Chem. Soc.*, 2005, **127**, 6451-6458.
32 H. H. Sun, L. K. Sanders and E. Oldfield, *J. Am. Chem. Soc.*, 2002, **124**, 5486-5495.
33 E. Oldfield, *Annu. Rev. Phys. Chem.*, 2002, **53**, 349-378.
34 J. G. Pearson, H. B. Le, L. K. Sanders, N. Godbout, R. H. Havlin and E. Oldfield, *J. Am. Chem. Soc.*, 1997, **119**, 11941-11950.
35 R. H. Havlin, D. D. Laws, H. M. L. Bitter, L. K. Sanders, H. H. Sun, J. S. Grimley, D. E. Wemmer, A. Pines and E. Oldfield, *J. Am. Chem. Soc.*, 2001, **123**, 10362-10369.
36 R. H. Havlin, H. B. Le, D. D. Laws, A. C. deDios and E. Oldfield, *J. Am. Chem. Soc.*, 1997, **119**, 11951-11958.
37 J. Heller, D. D. Laws, M. Tomaselli, D. S. King, D. E. Wemmer, A. Pines, R. H. Havlin and E. Oldfield, *J. Am. Chem. Soc.*, 1997, **119**, 7827-7831.
38 O. Soubias, F. Jolibois, V. Reat and A. Milon, *Chem.-Eur. J.*, 2004, **10**, 6005-6014.
39 J. A. Killian and T. K. M. Nyholm, *Curr. Opin. Struct. Biol.*, 2006, **16**, 473-479.
40 L. Cai, D. Fushman and D. S. Kosov, *J Biomol NMR*, 2009, **45**, 245-253.
41 U. Haeberlen, ed. J.S.Waugh, academic press, New york, 1976, vol. Suppl. 1.
42 M. Mehring, *Principles of High Resolution NMR in solids 2nd. Ed.*, Springer Verlag, Berlin, 1983.
43 H. W. Spiess, ed. P. Diehl, E. Fluck and R. Kosfeld, Springer Verlag, Berlin, 1978, vol. 15.
44 M. J. Gaussian 03 Revision E.01.1 Frisch, et al., in *Gaussian, Inc, Wallingford CT, 2004.*
45 F. Maseras and K. Morokuma, *J Comput Chem*, 1995, **16**, 1170-1179.
46 T. Vreven and K. Morokuma, *J Comput Chem*, 2000, **21**, 1419-1432.
47 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
48 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.
49 A. E. Foresman, *Exploring chemistry with electronic structure methods*, Gaussian Inc., Pittsburgh, 1998.
50 M. J. S. Dewar, M. L. McKee and S. Rzepa, *J. Am. Chem. Soc.*, 1978, **100**, 3607.
51 M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, 1977, **99**, 2338.
52 E. Anders, R. Koch and P. Freunscht, *J Comput Chem*, 1993, **14**, 1301-1312.
53 K. Wolinski and A. J. Sadlej, *Mol. Phys.*, 1980, **41**, 1419.
54 K. Wolinski, J. F. Hinton and P. Pulay, *J. Am. Chem. Soc.*, 1990, **112**, 8251-8260.
55 F. J. London, *J. Phys. Radium*, 1937, **8**, 397.
56 R. McWeeny, *Phys. Rev.*, 1962, **126**, 1028.
57 R. Ditchfield, *Mol. Phys.*, 1974, **27**, 789.
58 J. L. Dodds, R. McWeeny and A. J. Sadlej, *Mol. Phys.*, 1977, **34**, 1779-1791.
59 N. Asakawa, S. Kuroki, H. Kurosu, I. Ando, A. Shoji and T. Ozaki, *J. Am. Chem. Soc.*, 1992, **114**, 3261-3265.
60 N. Asakawa, M. Takenoiri, D. Sato, M. Sakurai and Y. Inoue, *Magnetic Resonance in Chemistry*, 1999, **37**, 303-311.
61 E. Salnikov, P. Bertani, J. Raap and B. Bechinger, *J. Biomol. NMR*, 2009, **45**, 373-387.
62 A. E. Reed, R. B. Weinstock and F. Weinhold, *J. Chem. Phys.*, 1985, **83**, 735-746.
63 U. C. Singh and P. A. Kollman, *J Comput Chem*, 1984, **5**, 129-145.
64 H. Torii, *J. Am. Chem. Soc.*, 2002, **124**, 9272-9277.
65 M. A. S. Hass, M. R. Jensen and J. J. Led, *Proteins-Structure Function and Bioinformatics*, 2008, **72**, 333-343.
66 P. E. Hansen, J. Abildgaard and A. E. Hansen, *Chemical Physics Letters*, 1994, **224**, 275-282.
67 J. Augspurger, J. G. Pearson, E. Oldfield, C. E. Dykstra, K. D. Park and D. Schwartz, *J. Magn. Reson.*, 1992, **100**, 342-357.
68 J. Boyd, C. Domene, C. Redfield, M. B. Ferraro and P. Lazzeretti, *J. Am. Chem. Soc.*, 2003, **125**, 9556-9557.
69 A. D. Buckingham, *Can.J.Chem.*, 1960, **38**, 300-307.
70 M. P. Williamson and T. Asakura, *J. Magn. Reson. Ser. B*, 1993, **101**, 63-71.
71 W. T. Raynes and R. Ratcliffe, *Mol. Phys.*, 1979, **37**, 571-578.
72 Y. D. Wu and Y. L. Zhao, *J. Am. Chem. Soc.*, 2001, **123**, 5313-5319.
73 F. M. Marassi, *Concepts. Magn. Reson.*, 2002, **14**, 212-224.
74 C. H. Wu, A. Ramamoorthy, L. M. Gierasch and S. J. Opella, *J. Am. Chem. Soc.*, 1995, **117**, 6148-6149.