Charles University in Prague

Faculty of Mathematics and Physics

DOCTORAL THESIS

Ladislav Benda

Determination of structure and dynamics of biomolecules by theoretical calculations of NMR spectroscopic parameters

Institute of Organic Chemistry and Biochemistry, AS CR

Supervisor:	Dr. Vladimír Sychrovský
Study programme:	Physics
Specialization:	4F4 – Biophysics, chemical and macromolecular physics

Prague 2012

I declare that I carried out this doctoral thesis independently, and only with the cited sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

Prague, 29th June 2012 Ladislav Benda

Abstrakt

Název: *Určování strukturních a dynamických vlastností biomolekul pomocí teoretických výpočtů parametrů spekter NMR*

Autor: Ladislav Benda, `ladislav.benda@gmail.com`

Katedra/Ústav: Ústav organické chemie a biochemie AV ČR, v.v.i.

Vedoucí doktorské práce: Dr. Vladimír Sychrovský, Ústav organické chemie a biochemie AV ČR, v.v.i., `vladimir.sychrovsky@uochb.cas.cz`

Abstrakt: Předkládaná doktorská práce se zaměřuje na teoretické modelování parametrů spekter nukleární magnetické rezonance (NMR) v peptidech a nukleových kyselinách. Modelovány byly především závislosti NMR parametrů na molekulární struktuře a solvataci. Velký důraz byl kladen na porovnání vypočtených dat s experimentem. Studovanými modely byly především di-peptid L-alanyl-L-alanin (AA) a fosfátová skupina páteře nukleových kyselin. Na základě výpočtů se podařilo určit konformace všech tří nabitých forem AA v roztoku a vysvětlit experimentálně pozorované změny NMR parametrů při změnách pH. Byly kalibrovány závislosti NMR kros-korelovaných relaxačních rychlostí na geometrii molekuly AA. ^{31}P NMR parametry ve fosfátu nukleových kyselin byly systematicky modelovány v závislosti na konformaci a solvataci fosfátu. Navrhli jsme pravidla pro kvalitativní strukturní interpretaci dvojvazných jaderných spin–spinových konstant $^2J_{\mathrm{PC}}$. Podařilo se modelovat změny ^{31}P NMR parametrů způsobené koordinací iontu Mg^{2+} k fosfátu nukleových kyselin. Byly úspěšně simulovány nízkofrekvenční pásy v Ramanových spektrech vodných roztoků solí iontu Mg^{2+}.

Klíčová slova: kvantově-chemické výpočty, NMR parametry, peptidy, nukleové kyseliny, kovové ionty

Abstract

Subject: *Determination of structure and dynamics of biomolecules by theoretical calculations of NMR spectroscopic parameters*

Author: Ladislav Benda, ladislav.benda@gmail.com

Department/Institute: Institute of Organic Chemistry and Biochemistry, AS CR

Supervisor: Dr. Vladimír Sychrovský, Institute of Organic Chemistry and Biochemistry, AS CR, vladimir.sychrovsky@uochb.cas.cz

Abstract: This doctoral work was focused on theoretical modeling of nuclear magnetic resonance (NMR) parameters in peptides and nucleic acids. Dependences of NMR parameters on molecular structure and solvation were primarily modeled. Great emphasis was put on the comparison of the calculated data with the NMR experiment. The molecular models studied included the L-alanyl-L-alanine di-peptide (AA) and the phosphate group of nucleic acid backbone. Conformations of all three charged forms of AA in solution were determined and the respective pH-induced changes of experimental NMR chemical shifts and nuclear spin–spin coupling constants were explained. Dependences of NMR cross-correlated relaxation rates on the AA backbone geometry were calibrated. The ^{31}P NMR parameters in nucleic acid phosphate were systematically calculated in dependence on the backbone conformation and the phosphate solvation pattern. Qualitative rules for the structural interpretation of two-bond nuclear spin–spin couplings $^{2}J_{\mathrm{PC}}$ were proposed. Changes in the ^{31}P NMR parameters induced by the coordination of the Mg^{2+} cation to the nucleic acid phosphate were calculated. Specific low-frequency bands observed in the Raman spectra of aqueous Mg^{2+} salt solutions were successfully simulated.

Keywords: quantum-chemistry calculations, NMR parameters, peptides, nucleic acids, metal ions

Contents

Acknowledgements **3**

List of original publications **4**

1 Introduction **6**

1.1 Computational NMR studies of L-alanyl-L-alanine 7

1.2 Modeling the effects of molecular flexibility and solvation on NMR parameters in nucleic acid phosphate 9

1.3 Aims of the thesis 12

2 Computational methods **13**

2.1 Born–Oppenheimer approximation 13

2.2 Hartree–Fock approximation 14

2.2.1 LCAO method 16

2.3 Correlation energy 17

2.4 Density-functional theory 18

2.4.1 Hohenberg–Kohn theorem 19

2.4.2 Kohn–Sham formulation of DFT 20

2.4.3 Exchange-correlation functionals 21

2.4.4 LCAO method in DFT 25

2.4.5 Performance of DFT methods 26

2.5 Molecular properties 26

2.5.1 First-order properties 28

2.5.2 Time-independent perturbation theory 28

2.5.3 Coupled-perturbed scheme 29

2.6 General notes on the computational methodology 32

2.6.1 Quantum chemistry programs 33

3 Theory of NMR spectroscopic method **34**

3.1 NMR spectroscopy 34

3.1.1 Magnetic interactions in molecules 35

3.1.2 Effective spin Hamiltonian 37

3.1.3 NMR parameters 37

3.1.4 NMR relaxation 39

3.1.5 Structural interpretation of NMR parameters 41

3.2 Theory of NMR parameters 42

3.2.1 J-coupling Hamiltonian 43

3.2.2 Perturbation theory of J-coupling 44

3.2.3 Chemical shielding Hamiltonian 45
3.2.4 Perturbation theory of chemical shielding 46
3.3 DFT calculations of NMR parameters 46
3.3.1 DFT calculation of J-coupling 46
3.3.2 DFT calculations of chemical shielding 48

4 Results 51
4.1 Computational NMR studies of L-alanyl-L-alanine 51
4.1.1 Geometry dependencies of NMR chemical shifts and J-couplings . . 53
4.1.2 Geometry dependencies of cross-correlated relaxation rates 54
4.2 Modeling the effects of molecular flexibility and solvation on NMR parameters in nucleic acid phosphate 54
4.2.1 Benchmark NMR calculations 55
4.2.2 Geometry dependencies of ^{31}P NMR parameters 56
4.2.3 Response of ^{31}P NMR parameters to Mg^{2+} coordination 59
4.2.4 Magnesium ion imprints in Raman scattering of water 61

5 Summary 64

List of abbreviations 66

References 68

Published works 78

Acknowledgements

I would like to thank my supervisor Dr. Vladimír Sychrovský for his guidance and support throughout my Ph.D. studies. I'm sincerely grateful for all I learned from him. I also acknowledge valuable consultations with prof. Jaroslav Burda and doc. Jan Lang.

I'm grateful to my colleagues from the Department of Molecular Spectroscopy at the Institute of Organic Chemistry and Biochemistry for creating friendly and helpful working atmosphere. I specifically acknowledge kind cooperation with Jan Horníček, doc. Michal Straka, Dr. Jaroslav Šebestík, Petr Štěpánek, and Dr. Zuzana Sochorová Vokáčová. Very special thanks are addressed to the head of the department doc. Petr Bouř for his generous help, lots of encouragement, and countless hours of inspiring discussions.

Finally, I'm deeply thankful to my whole family for their unfailing support. This work is also their achievement.

List of original publications

The results presented in this doctoral thesis were published in the following works:

I L. Benda: “Calculation of the effective chemical shielding anisotropy in L-alanyl-L-alanine, conformational and charge dependence study.” *WDS’07 Proceedings of Contributed Papers, Part III*, **2007**, 84–89.

II V. Sychrovský, M. Buděšínský, L. Benda, V. Špirko, Z. Vokáčová, J. Šebestík, P. Bouř: “Dependence of the L-alanyl-L-alanine conformation on molecular charge determined from ab initio computations and NMR Spectra.” *J. Phys. Chem. B*, **2008**, *112*, 1796–1805, DOI: 10.1021/jp076557j.

III L. Benda, P. Bouř, N. Müller, V. Sychrovský: “Theoretical study of the effective chemical shielding anisotropy (CSA) in peptide backbone, rating the impact of CSAs on the cross-correlated relaxations in L-alanyl-L-alanine.” *J. Phys. Chem. B*, **2009**, *113*, 5273–5281, DOI: 10.1021/jp8105452.

IV J. Kapitán, M. Dračínský, J. Kaminský, L. Benda, P. Bouř: “Theoretical modeling of magnesium ion imprints in the Raman scattering of water.” *J. Phys. Chem. B*, **2010**, *114*, 3574–3582, DOI: 10.1021/jp9110508.

V L. Benda, B. Schneider, V. Sychrovský: “Calculating the response of the NMR shielding tensor $\sigma(^{31}\mathrm{P})$ and the $^{2}J(\mathrm{P,C})$ coupling constants in nucleic acid phosphate to coordination of Mg^{2+} cation.” *J. Phys. Chem. A*, **2011**, *115*, 2385–2395, DOI: 10.1021/jp1114114.

VI L. Benda, Z. Vokáčová, M. Straka, V. Sychrovský: “Correlating the ^{31}P-NMR chemical shielding tensor and the $^{2}J_{\mathrm{P,C}}$ spin–spin coupling constants with torsion angles ζ and α in backbone of nucleic acids.” *J. Phys. Chem. B*, **2012**, *116*, 3823–3833, DOI: 10.1021/jp2099043.

Other publications:

- V. Sychrovský, L. Benda, A. Prokop, V. Blechta, J. Schraml, V. Špirko: “Probing the flexibility of internal rotation in silylated phenols with the NMR scalar spin–spin coupling constants.” *J. Phys. Chem. A*, **2008**, *112*, 5167–5174, DOI: 10.1021/jp711809e.

- L. Benda, M. Straka, Y. Tanaka, V. Sychrovský: “On the role of mercury in the non-covalent stabilisation of consecutive U–Hg^{II}–U metal-mediated nucleic acid base pairs: Metallophilic attraction enters the world of nucleic acids.” *Phys. Chem. Chem. Phys.*, **2011**, *13*, 100–103, DOI: 10.1039/C0CP01534B.
  - presented on back cover of Phys. Chem. Chem. Phys. **2011**, *13*, issue 2

- M. Dračínský, L. Benda, P. Bouř: "Ab initio modeling of fused silica, crystal quartz, and water Raman spectra." *Chem. Phys. Lett.*, **2011**, *512*, 54–59, DOI: 10.1016/j.cplett.2011.06.077.
- T. Uchiyama, T. Miura, H. Takeuchi, T. Dairaku, T. Komuro, T. Kawamura, Y. Kondo, L. Benda, V. Sychrovský, P. Bouř, I. Okamoto, A. Ono, Y. Tanaka: "Raman spectroscopic detection of the T–Hg^{II}–T base pair and the ionic characteristics of mercury." *Nucleic Acids Res.*, **2012**, *40*, 5766–5774, DOI: 10.1093/nar/gks208.
- L. Benda, M. Straka, V. Sychrovský, Y. Tanaka, P. Bouř: "Detection of Mercury-TpT Dinucleotide Binding by the Raman Spectra. A Computational Study." *J. Phys. Chem. A*, *submitted.*

The author presented his work at 13 international conferences and workshops, giving 7 talks and 6 poster presentations. The published works were 31 times cited to date.

The author's contributions to the individual projects

I – All computational work was conducted by the author. The manuscript was written by the author.

II – All computations concerning local minima of the L-alanyl-L-alanine molecule were performed and analyzed by the author. The author actively participated in the manuscript preparation.

III – All computational work was conducted by the author. The author prepared all tabular and graphical matter including the Supporting Information and wrote approximately half of the manuscript.

IV – The author performed and analyzed Raman and NBO calculations for ab initio optimized clusters.

V – All computational work was conducted by the author. The author prepared all tabular and graphical matter including the Supporting Information and wrote approximately half of the manuscript.

VI – The author performed most of the calculations, analyzed all the data and prepared all tabular and graphical matter including the Supporting Information. The manuscript was completely written by the author.

1 Introduction

Determination of structure and dynamical behavior of biomolecules is a fundamental part of contemporary research. The knowledge of biomolecular structure is crucial for understanding their function and role in biological processes as well as for development of new drugs and nanodevices. Molecular spectroscopy techniques such as NMR, infrared and Raman spectroscopy, (vibrational) circular dichroism, Raman optical activity, and many others are well suited for detailed studies of molecular systems in liquid phase. Reliable interpretation of the experimental data is essential for the application of all spectroscopic methods. Improved interpretation schemes can substantially increase both the amount and relevance of information obtained from spectroscopic measurements.

During the past two decades, theoretical calculations have become an indispensable part of spectroscopic research, enabling unique correlation of the experimental data with molecular structure and dynamics at the atomic level. Many computational methods are available nowadays, ranging from very accurate and demanding quantum chemistry methods based on fundamental principles of quantum mechanics to rather rough semiempirical approximations. The *ab initio* wavefunction methods rigorously approximate the exact solution of Schrödinger equation and their accuracy can be systematically improved. However, due to their high computational cost they are mostly used only for calibration purposes. An alternative approach emerged with the onset of density functional theory (DFT). It was shown that satisfactory results can be obtained with the DFT at a fraction of the *ab initio* computational cost provided that the DFT was properly tested and calibrated for the particular purpose. The DFT methods turned out particularly useful for calculations of spectroscopic parameters. With the current computer power, DFT allows for affordable spectroscopic calculations even in large molecular systems, extensive conformational sampling, and averaging over statistical distribution of solvent molecules.

As explained in chapter 2, the quantum chemistry calculations of spectroscopic parameters are much more demanding than those of energy or gradient. Moreover, the effects of solvent and molecular motion are usually not negligible and their proper description further increases the time demands. The performance of individual techniques also depends on the calculated quantity and the particular system under study. Therefore, the accuracy of all approximations employed in the spectroscopic calculations should be balanced with their computational cost.

This doctoral thesis is aimed at theoretical modeling and interpretation of NMR spectroscopic properties of fundamental building blocks of peptide and nucleic acid structure. Special attention is dedicated to the effects of molecular flexibility and solvation. The presented material is for clarity divided into two sections according to the molecular system under study: (1) Computational NMR studies of L-alanyl-L-alanine (publications

I–III). (2) Modeling the effects of molecular flexibility and solvation on NMR parameters in nucleic acid phosphate (publications IV–VI). The rest of this chapter introduces the reader to these topics and states the aims of the thesis. Chapter 2 summarizes the applied computational techniques and chapter 3 gives the theoretical background for the NMR spectroscopy and NMR parameters. The important results of publications I–VI are presented in chapter 4 and the achieved goals are summarized in the final chapter 5. A complete account of the presented work is given in the explicitly enclosed publications.

1.1 Computational NMR studies of L-alanyl-L-alanine

Peptides and proteins are functional biomolecules consisting of amino acids linked with amide bonds. One of the most distinguished features of peptides is their ability to fold into active conformation in which they can exert their biological function. Peptide conformation is characterized by repetitive sequence of backbone torsion angles φ, ψ and ω (Figure 1.1). The torsions φ and ψ describe backbone conformation of each amino acid; the values found in peptides are usually visualized in the form of so-called Ramachandran plot. The torsion angle ω describing the conformation of the amide bond is mostly found around 180° (*trans* conformation) and rarely also around 0° (*cis* conformation).

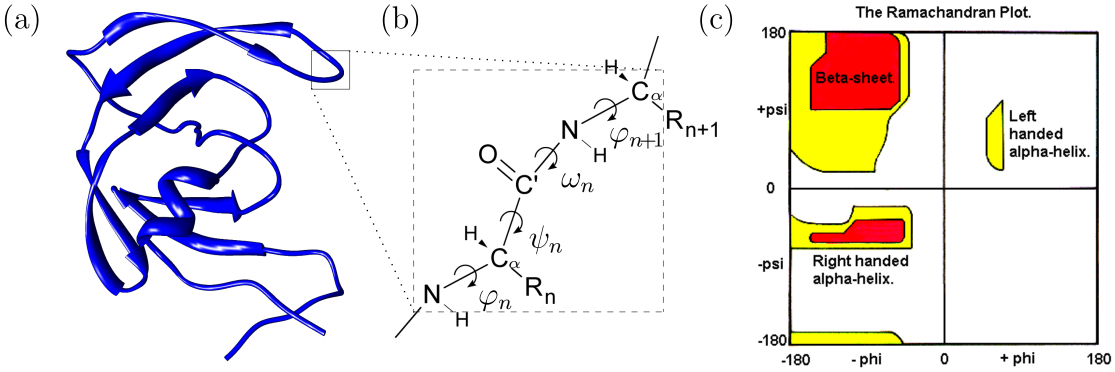


Figure 1.1 (a) A schematic example of peptide fold (visualized with `Chimera`). (b) A close-up of peptide backbone structure with definition of backbone torsion angles φ, ψ, ω, backbone atoms C_α, C', N, amino-acid side chains -R, and amino-acid residue numbering (starting from N-terminus). (c) Example of Ramachandran plot; adopted from `http://www.cryst.bbk.ac.uk/PPS95/course/3_geometry/rama.html`.

Determination of peptide structure with NMR spectroscopy usually relies on measurement and interpretation of the three-bond nuclear spin–spin coupling constants (^3J-couplings), the NMR relaxation enhancements due to the nuclear Overhauser effect (NOE), and the residual dipolar couplings (RDCs) in partially oriented media. The ^3J-couplings are interpreted as structural restraints on the peptide backbone torsion angles via known Karplus relations, the NOEs provide ^1H–^1H inter-atomic distances, and the RDCs contain information on spatial orientation of dipolar vectors.

Of the three backbone torsion angles, NMR determination of torsion ψ is typically the most problematic since there is no ${}^3J_{\mathrm{HH}}$ coupling that could be assigned to this angle with a Karplus relation. At the same time, the torsion ψ is very important as it distinguishes between two major peptide backbone patterns, the α-helix and the β-sheet (Figure 1.1). Other NMR parameters such as cross-correlated relaxation rates [1–3] could be used for determining the torsion angle ψ, provided that their relation to the peptide structure was calibrated.

Various models of peptide structure can be adopted for computational characterization of geometry dependencies of NMR parameters. We studied the L-alanyl-L-alanine dipeptide (AA; Figure 1.2) which correctly represents the $C_\alpha{-}C'{-}N{-}C_\alpha$ link of peptide backbone including its chemical surroundings. The AA molecule can be thus regarded as the smallest system suitable for realistic modeling of conformational dependencies of NMR parameters adjacent to the amide bond.

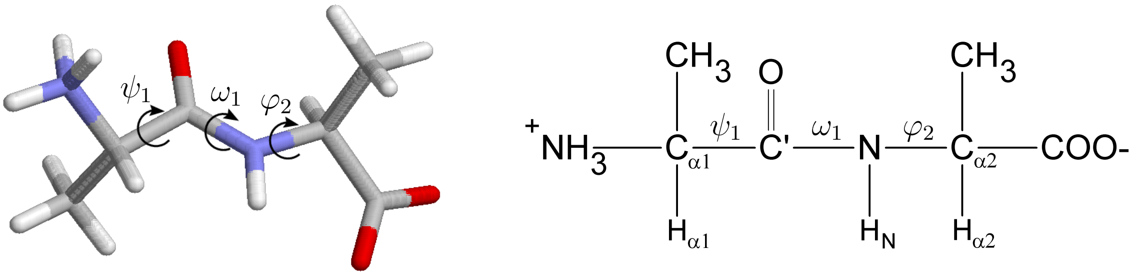


Figure 1.2 The L-alanyl-L-alanine zwitterion with the definition of three main torsion angles ψ_1, ω_1, φ_2 and notation of atoms $C_{\alpha1}$, $H_{\alpha1}$, C', H_N, $C_{\alpha2}$, $H_{\alpha2}$ numbered according to the amino-acid residue starting from the N-terminus.

The AA molecule in aqueous solution exists in three protonation forms, whose mutual ratios depend on pH. At neutral pH both amino (-NH_3^+) and carboxyl (-COO^-) groups are charged, forming thus a neutral AA zwitterion (further denoted AA^{zw}; Figure 1.2). In acidic environment at low pH, the carboxyl group accepts proton (-COOH) and the molecule is positively charged (AA^+). Finally, at high pH the amino group loses proton (-NH_2) and the AA is anionic (AA^-). The terminal amino and carboxyl groups strongly interact with solvent, what substantially influences conformational flexibility and spectroscopic properties of AA; the AA^{zw} form is actually unstable in gas phase. The solvation effects thus have to be included for reliable modeling of NMR parameters in AA.

The AA structure is mainly described with three torsion angles ψ_1, ω_1 and φ_2 (Figure 1.2). Other degrees of freedom such as rotation of methyl side chains and the amino and carboxyl terminal groups can be considered less important for modeling the NMR parameters in peptide backbone. Moreover, as stated above, the ω_1 torsion is typically restricted around 180°. There are thus two major structural descriptors of the AA molecule, the backbone torsions ψ_1 and φ_2. It is feasible to theoretically model the conformational dependence of spectroscopic parameters in the AA dipeptide by systematic variation of these two torsion angles.

Spectroscopic characterization of basic building blocks of peptide structure has been one of the central research topics in the Molecular Spectroscopy group at the Institute of Organic Chemistry and Biochemistry in Prague. Prior to publications I–III presented in this thesis, the AA^{zw} molecule was characterized with Raman optical activity (ROA) [4], its absolute conformation in solution was determined based on both NMR calculations and experiment [5], and the pH dependence of the electronic circular dichroism spectra was studied [6]. Directly related to this topic are also NMR conformational studies of other dipeptides [7, 8], simulations of molecular dynamics effects on ROA spectra of L-alanine [9] and L-proline [10], evaluation of anharmonic effects on optical [11] and NMR [12] spectra of these amino-acids, and modeling the solvent effects on NMR parameters in L-alanine [13, 14].

1.2 Modeling the effects of molecular flexibility and solvation on NMR parameters in nucleic acid phosphate

Nucleic acids (NAs) are nowadays recognized as rather versatile biomolecules with large variety of functions, far exceeding the traditional view of NAs as mere means of storage and translation of genetic information. Structure and dynamics of NAs have been extensively studied with many methods including NMR spectroscopy [15, 16]. Particular attention has been dedicated to ribonucleic acids (RNAs) [17, 18] that can form many unusual structures, some even exhibiting catalytic activity [19, 20].

NA molecule is a poly-anion constituting of a sequence of (2'-deoxy)ribonucleotides connected with phosphodiester bonds. The negative charge of NA phosphates is stabilized in aqueous solution by dissolved cations. Conformation of each nucleotide is described with seven torsion angles α, β, γ, δ, ϵ, ζ, χ, the firs six describing the conformation of the sugar-phosphate backbone and the last torsion determining the orientation of NA base (Figure 1.3). Only certain combinations of the backbone torsion angles are possible. DNA mostly occurs in the A-, B_I-, B_{II}-, or Z-form, whereas the conformational richness of RNA molecules is described with 46 RNA conformational classes [21].

Similarly as for peptides, the NMR studies of NA structure utilize mainly NOE restraints on inter-atomic distances, RDCs in partially aligned samples, and ^3J-couplings assigned to torsion angles and interpreted via Karplus relations [15, 16, 22]. The applicability of NMR for determining the NA structure is, however, more problematic than for peptides, mainly due to relatively low density of ^1H nuclei in NA molecules and also because the ^{31}P NMR spectra are typically very crowded. The size of NA structures that can be determined with NMR is therefore limited.

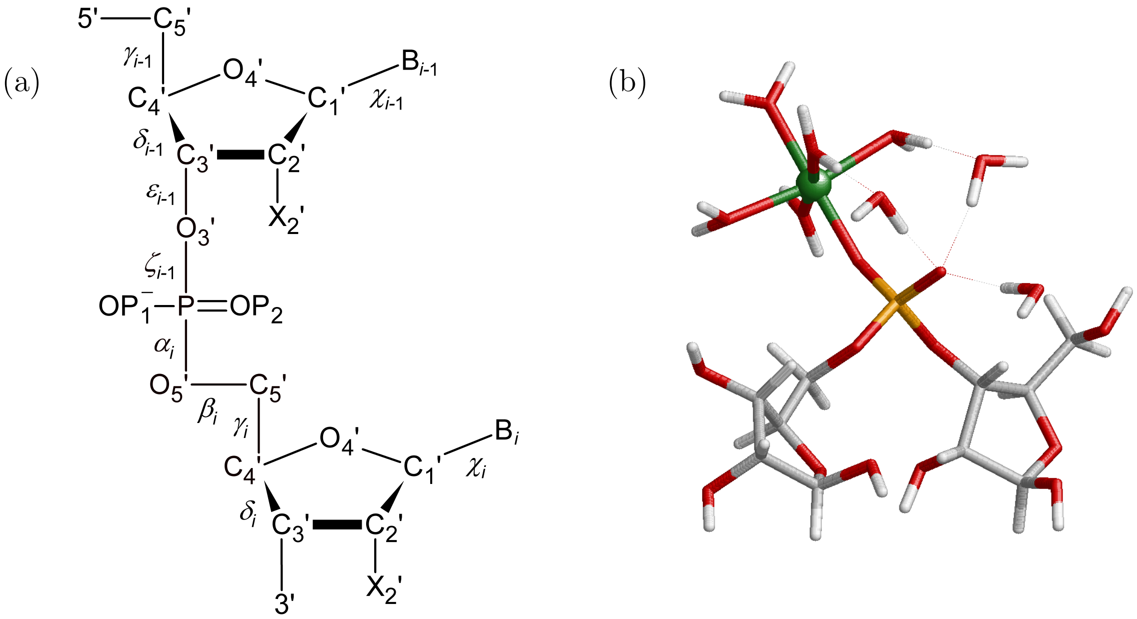


Figure 1.3 (a) Scheme of NA primary structure. Definition of NA backbone torsion angles α, β, γ, δ, ϵ and ζ, glycosidic torsion χ, notation of atoms of sugar-phosphate backbone ($X'_2 = H'_2$ in DNA or O'_2H in RNA), and nucleotide residue numbering starting from $5'$-terminus. B is a NA base. (b) Example of NA phosphate specific solvation: inner-sphere coordination of Mg^{2+} (green sphere) to phosphate oxygen OP1. First hydration shell of the phosphate oxygen OP2 and the Mg^{2+} ion is retained.

Phosphorus NMR is one of the key techniques for studying NA structure [23,24]. The $^3J_{\mathrm{PC}}$ and $^3J_{\mathrm{PH}}$ coupling constants interpreted via Karplus relations [25] are used for determining the NA backbone torsion angles ϵ and β. The P–O torsions ζ and α (Figure 1.3), however, cannot be obtained on the basis of three-bond coupling constants, since the $^3J_{\mathrm{OC}}$ couplings are not measurable. Instead, the ^{31}P chemical shifts δ_{P} are used as somewhat loose restraints on the (ζ, α) conformation of NA phosphate. Recently, the $\Gamma_{\mathrm{P,CH}}$ cross-correlated relaxation (CCR) rates were applied for unambiguous determination of torsion angles ζ and α [26,27]. The resulting structural restraints were demonstrated to substantially reduce the uncertainty of a RNA 14-mer structure determined with NMR [28]. A drawback of this approach might be the great complexity of NMR experiments needed for the $\Gamma_{\mathrm{P,CH}}$ measurement. Considering the gain in quality of NA structure determination, finding some further restraints on the (ζ, α) conformation of NA phosphate is still very attractive.

The relationship between δ_{P} and NA phosphate conformation has been for a long time known only semi-empirically [29]. A significant improvement became possible with development of accurate computational methods and reliable solvent models. Quite recently, the ^{31}P chemical shielding tensor has been calculated on SOS–DFPT level for several geometries of explicitly hydrated dimethyl phosphate corresponding to the most populated (ζ, α) conformers of NA backbone [30,31]. Přecechtělová et al. also successfully modeled the experimentally known δ_P difference of 1.6 ppm between the $\mathrm{B_{II}}$ and $\mathrm{B_I}$ forms of DNA [32]. However, a systematic first-principles calibration of the δ_P dependence on torsion angles ζ and α in the full $(0°, 360°)$ range has not been reported.

One of the few remaining options for monitoring the torsion angles ζ and α with NMR spectroscopy are the ${}^2J_{\mathrm{PC3'}}$ and ${}^2J_{\mathrm{PC5'}}$ coupling constants involving carbon atoms in the closest vicinity of NA phosphate (Figure 1.3). Previous theoretical calculations indicated that the *trans* orientation of either torsion ζ or α could be distinguished from *gauche* by a significantly decreased absolute value of ${}^2J_{\mathrm{PC3'}}$ or ${}^2J_{\mathrm{PC5'}}$ coupling (by $2 - 3$ Hz), respectively [22]. Similar variations of the two ${}^2J_{\mathrm{PC}}$ couplings were actually measured in NAs, but no structural interpretation was available [33–37]. Theoretical calculations of the ${}^2J_{\mathrm{PC}}$ coupling constants in dependence on torsion angles ζ and α thus may reveal their applicability in NA structure determination.

NMR parameters related to the phosphate group are generally sensitive to two factors: the NA backbone geometry and the phosphate group solvation. Solvation of NAs has been an important topic since the discovery of NA structure [38]. Negative charge of NA backbone is in water solution compensated by the surrounding cations, mostly $\mathrm{Na^+}$ and $\mathrm{K^+}$. Apart from the non-specific water mediated electrostatic solvation, metal cations can specifically coordinate to NA phosphate. Such specific interactions have turned out exceedingly important in biochemical processes like RNA folding [39, 40], ribozyme catalysis [19, 20, 41], or drug-DNA complexation [42]. Exceptional structural role of the $\mathrm{Mg^{2+}}$ ion has been recognized in all of these cases.

The most usual method for monitoring the NA metalation is X-ray crystallography. The identification of $\mathrm{Na^+}$ and $\mathrm{Mg^{2+}}$ is, however, slightly complicated by the fact that these cations are hard to distinguish from water molecules since they possess the same number of electrons. Metal coordination sites around phosphate are known from statistical analysis of X-ray structures of organic phosphate salts [43]. The statistical distributions of metal cations obtained by the method of Fourier averaging [44] are well localized around the phosphate, similarly to water molecules in the phosphate first hydration shell [45]. $\mathrm{Mg^{2+}}$ cation has two sharp distributions around the phosphate, one near each of the charged phosphate oxygens OP1 and OP2 (Figure 1.3). Only one of the two coordination sites can be populated at a time. This very distinct phosphate coordination pattern is clearly related to the exceptional properties of the $\mathrm{Mg^{2+}}$ ion in aqueous solution (see below).

Metalation of NA phosphate could be also detected with NMR spectroscopy, by monitoring NMR parameters adjacent to the phosphate group. From the available experimental evidence, only effects of direct (inner-sphere) coordination of metal ions to NA phosphate were observed in ${}^{31}\mathrm{P}$ NMR spectra. In particular, $\mathrm{Cd^{2+}}$ coordination to a thio-substituted metal binding site in hammerhead ribozyme lead to the decrease of δ_P of the phosphorothioate residue by $1.5 - 10$ ppm [46–48] and $\mathrm{Mg^{2+}}$ coordination to the minimal metal binding motif of hammerhead ribozyme caused decrease of δ_{P} by 0.3 ppm [49]. Theoretical investigation of Mg^{2+} binding to NA phosphate and associated changes in the ^{31}P NMR spectra may shed light on the molecular basis of metal-binding effects in NAs.

Divalent magnesium itself is very interesting because of its biological importance and unique properties [50,51]. Mg^{2+} in water solution strongly coordinates six water molecules, enforcing a regular octahedral structure of its first hydration shell. The corresponding water residence time of about 1 μs is by far the largest among all physiological cations [52]. Regularity of the Mg^{2+} hydration is preserved even in the limit case of $MgCl_2 \cdot 6\,H_2O$ (bischofite) crystal [53], where the water molecules just saturate the first hydration shell of the Mg^{2+} ions. Such high affinity of Mg^{2+} towards water and its exceptional solvent-structuring properties [54] lead to observable effects in optical spectra of magnesium salts in water solution. In particular, local symmetry of water solvent induced by Mg^{2+} and the strength of the Mg−O bond cause significant increase of certain Raman signals in low-wavenumber region [55, 56]. Theoretical studies of ionic liquids have been so far mainly concentrated on modeling the structure and dynamics of the ion hydration shell [55, 57–65], vibrational spectra of ion-water clusters [55, 66–68], behavior of the ions on water surface [69–71], and ionic solution acidity [72]. However, faithful simulation of Raman bands measured in ionic solutions and theoretical explanation of the origin of low-wavenumber spectral features are still challenging.

1.3 Aims of the thesis

- Thorough computational characterization of the L-alanyl-L-alanine dipeptide in its various protonation forms. Assessment of the relationship between the dipeptide conformation and NMR parameters. Theoretical analysis of cross-correlated relaxation mechanisms in L-alanyl-L-alanine.

- Exploring the relationship between nucleic acid backbone conformation and NMR parameters related to the phosphate group. Calculating the effect of Mg^{2+} coordination to nucleic acid phosphate on ^{31}P NMR spectra. Modeling of the magnesium ion properties in aqueous solution.

2 Computational methods

Virtually all chemistry is described at the level of relativistic quantum mechanics [73]. For the spectroscopic calculations carried out in this work, non-relativistic description is sufficient. Molecular systems including their interactions with electromagnetic fields are described in the non-relativistic limit with the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{\mathcal{H}}\Psi \ , \tag{2.1}$$

where $\hbar$ is the reduced Planck constant, $\hat{\mathcal{H}} = \hat{\mathcal{H}}(t)$ is the molecular Hamiltonian, and $\Psi = \Psi(t)$ is the wave function. In the absence of electromagnetic fields, the molecular Hamiltonian is time-independent and the corresponding Schrödinger equation is stationary with energy $\mathcal{E}$:

$$\hat{\mathcal{H}}\Psi = \mathcal{E}\Psi \ . \tag{2.2}$$

The molecular Hamiltonian $\hat{\mathcal{H}}$ consists of the following terms:

$$\hat{\mathcal{H}} = \hat{T} + \hat{T}_{\text{nuc}} + \hat{U} + \hat{U}_{\text{nuc}} + \hat{V} \ , \tag{2.3}$$

where $\hat{T}$ and $\hat{T}_{\text{nuc}}$ are the electronic and nuclear kinetic energy operators, $\hat{U}$ and $\hat{U}_{\text{nuc}}$ include all the electron-electron and nucleus-nucleus repulsion terms, and $\hat{V}$ describes the electrostatic attraction between electrons and nuclei. The individual terms in atomic units can be expressed as

$$\begin{aligned} &\hat{T} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 \ , \qquad \hat{U} = \sum_{i<j}\frac{1}{r_{ij}} \ , \qquad \hat{V} = -\sum_{i=1}^{N}\sum_{A=1}^{Q}\frac{Z_A}{r_{iA}} \ , \\ &\hat{T}_{\text{nuc}} = -\frac{1}{2}\sum_{A=1}^{Q}\frac{1}{2M_A}\nabla_A^2 \ , \qquad \hat{U}_{\text{nuc}} = \sum_{A<B}\frac{Z_A Z_B}{R_{AB}} \ , \end{aligned} \tag{2.4}$$

where N and Q are numbers of electrons and nuclei, ∇_i^2 and ∇_A^2 are Laplace operators acting on electrons and nuclei, M_A and Z_A are nuclear masses and charges, and r_{ij}, r_{iA} and R_{AB} are the electron-electron, electron-nucleus and nucleus-nucleus distances, respectively.

2.1 Born–Oppenheimer approximation

The Schrödinger equation (2.2) with Hamiltonian (2.3) can be solved exactly only in the simplest cases such as one-electron atoms. The solution for molecular systems can be approximated by assuming that the electronic and nuclear parts of the total wave function

Ψ are uncoupled and separable as

$$\Psi(\{\mathbf{x}_i\},\{\mathbf{R}_A\}) = \Phi(\{\mathbf{x}_i\};\{\mathbf{R}_A\}) \cdot \Theta(\{\mathbf{R}_A\}) \ , \tag{2.5}$$

where $\{\mathbf{x}_i\} = \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N = \{\mathbf{r}_i, s_i\}$ denote spatial and spin coordinates of electrons. The electronic wave function Φ implicitly depends also on the positions of nuclei $\{\mathbf{R}_A\}$. By further assuming $\nabla_A \Phi = \mathbf{0}$ and $\nabla_A^2 \Phi = 0$ we obtain the Born–Oppenheimer approximation [74], which allows to describe the electrons and nuclei separately. The respective Schrödinger equations for electrons and nuclei are:

$$\big(\hat{T} + \hat{U} + \hat{V}\big)\Phi = E\Phi \tag{2.6a}$$

$$\big(\hat{T}_{\text{nuc}} + \hat{U}_{\text{nuc}} + E\big)\Theta = \mathcal{E}\Theta \tag{2.6b}$$

The electronic energy $E = E(\{\mathbf{R}_A\})$ introduced in (2.6a) as an eigenvalue of the electronic Hamiltonian $\hat{H} = \hat{T} + \hat{U} + \hat{V}$ parametrically depends on the configuration of atomic nuclei and it acts as an external potential for the nuclei in (2.6b). The operator $\hat{V}$ is sometimes referred to as the external potential since it is the only term in (2.6a) involving interactions reaching out of the otherwise purely electronic system.

2.2 Hartree–Fock approximation

The electronic structure calculations in molecules require simplifications of the many-electron equation (2.6a). The N-electron wave function $\Phi(\{\mathbf{x}_i\})$ is in the Hartree–Fock (HF) method assumed in the form of antisymmetrized product of orthonormal one-electron spin orbitals, the Slater determinant

$$\Phi^{\text{SD}}(\{\mathbf{x}_i\}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \ , \tag{2.7}$$

where each spin orbital $\chi(\mathbf{x}_i)$ is defined as a product of molecular orbital (MO) $\psi(\mathbf{r}_i)$ and spin function $\sigma(s_i)$. The antisymmetry of the Slater determinant with respect to the exchange of electrons ensures that the Pauli exclusion principle is satisfied. Spin orbitals of the electronic ground state are found by applying the variational principle to minimize the energy functional

$$E[\Phi^{\text{SD}}] = \big\langle \Phi^{\text{SD}} \big| \hat{H} \big| \Phi^{\text{SD}} \big\rangle \tag{2.8}$$

upon the condition of orthonormality of spin orbitals $\langle \chi_i | \chi_j \rangle = \delta_{ij}$. This procedure leads to canonical HF equations [75]

$$\hat{f}\chi_i = \varepsilon_i \chi_i \ , \tag{2.9}$$

where $\hat{f}$ is the Fock operator, its eigenfunctions χ_i are the HF spin orbitals and the corresponding eigenvalues ε_i are the HF orbital energies. In the following we use the indexes a, b, c, d for occupied orbitals, p, q, r, s for unoccupied (virtual) orbitals and i, j, k, l for all orbitals. The Fock operator can be expressed as

$$\hat{f} = \hat{h} + \hat{g} \ , \tag{2.10a}$$

$$\hat{h} = -\frac{1}{2}\nabla^2 - \sum_{A=1}^{Q} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \ , \tag{2.10b}$$

$$\hat{g} = \sum_{b=1}^{N} \left(\hat{J}_b - \hat{K}_b \right) \ , \tag{2.10c}$$

where $\hat{h}$ is the 'core' operator containing all one-electron terms and $\hat{g}$ is the HF potential describing the interaction of one electron with all remaining $N-1$ electrons. $\hat{J}_b$ and $\hat{K}_b$ are the Coulomb and exchange operators, respectively, defined as

$$\hat{J}_b \chi_i(\mathbf{x}_1) = \left[\int \frac{1}{r_{12}} \chi_b^*(\mathbf{x}_2) \chi_b(\mathbf{x}_2) \, \mathrm{d}\mathbf{x}_2 \right] \chi_i(\mathbf{x}_1) \ , \tag{2.11a}$$

$$\hat{K}_b \chi_i(\mathbf{x}_1) = \left[\int \frac{1}{r_{12}} \chi_b^*(\mathbf{x}_2) \chi_i(\mathbf{x}_2) \, \mathrm{d}\mathbf{x}_2 \right] \chi_b(\mathbf{x}_1) \ . \tag{2.11b}$$

Since both $\hat{J}_b$ and $\hat{K}_b$ involve occupied spin orbitals, the Fock operator implicitly depends on its own eigenfunctions. The HF equations are thus non-linear and have to be solved iteratively.

The HF ground-state energy E_0^{HF} is obtained by evaluating the expectation value (2.8) with the Slater determinant formed by occupied HF spin orbitals χ_a. After introducing the MOs and integrating over the spin parts we obtain

$$E_0^{\mathrm{HF}} = \sum_{a=1}^{N} \langle a|\hat{h}|a\rangle + \frac{1}{2} \sum_{a=1}^{N} \sum_{b=1}^{N} \big((aa|bb) - \delta_{\sigma_a \sigma_b} (ab|ba) \big) \ , \tag{2.12}$$

where we used the usual notation for one-electron and two-electron integrals

$$\langle a|\hat{h}|a\rangle = \int \psi_a^*(\mathbf{r}) \hat{h} \psi_a(\mathbf{r}) \, \mathrm{d}\mathbf{r} \ , \tag{2.13a}$$

$$(ab|cd) = \iint \psi_a^*(\mathbf{r}_1) \psi_b(\mathbf{r}_1) \frac{1}{r_{12}} \psi_c^*(\mathbf{r}_2) \psi_d(\mathbf{r}_2) \, \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \ . \tag{2.13b}$$

The exchange integral $(ab|ba)$ represents the so-called 'static correlation' between same-spin electrons and its presence only for the same-spin electron pairs (note the Kronecker delta in (2.12)) is a pure consequence of Pauli exclusion principle invoked by the antisymmetry of the Slater determinant (2.7). The absence of such correlation between the opposite-spin electrons leads to the orbital instability of the HF method [76, 77] (see section 3.3.1). Note that for $b = a$ the Coulomb and HF exchange integrals $(aa|bb)$ and $(ab|ba)$ exactly cancel, which implies that the spurious self-interaction error is absent at the HF level.

This work involved only molecular systems with closed-shell electronic configuration, where the MOs for α and β electrons coincide. The closed-shell HF equations are in this case solved only for the $N/2$ lowest doubly-occupied MOs. In the following, if not stated otherwise, the theory will be formulated only for closed-shell systems.

2.2.1 LCAO method

In most cases of chemical interest, the HF equations cannot be solved analytically. In practice, the MOs are approximated as expansions into the basis of atomic orbitals (AOs) ϕ_μ centered at atomic positions $\mathbf{R}_\mu$, $\mathbf{r}_\mu = \mathbf{r} - \mathbf{R}_\mu$:

$$\psi_i(\mathbf{r}) \approx \sum_{\mu=1}^{M} C_{\mu i}\phi_\mu(\mathbf{r}_\mu) \ , \qquad M \geq N/2 \ , \tag{2.14}$$

where $C_{\mu i}$ is the matrix of AO expansion coefficients. In the rest of this section, the Greek letters $\mu, \nu, \lambda, \sigma$ are used as suffixes for the AO basis functions, whereas the Roman letters are used for the MOs. Each AOs is characterized with three quantum numbers n, l, m and formed by a product of angular part (spherical harmonic Y_{lm}) and radial part $R_{nl}(r) = P_{nl}(r)\exp(-\alpha_n r)$, where $P_{nl}(r)$ is a polynomial of order $n-1$. To facilitate integration, the exponential functions $\exp(-\alpha_n r)$ are in most quantum chemistry packages approximated by linear combinations of several Gaussian functions, forming thus the Gaussian-type orbitals (GTOs). Obviously, the accuracy of expansion (2.14) depends on the basis set size and quality. Nevertheless, if the AO basis is chosen suitably to the calculation type, results close to the HF complete basis set limit can be achieved with only a moderate effort.

Using the LCAO expansion, the HF equations (2.9) can be in the closed-shell case rewritten into the matrix form usually referred to as the Roothaan (or Roothaan–Hall) equations [78, 79]:

$$\sum_{\nu=1}^{M}(F_{\mu\nu} - \varepsilon_i S_{\mu\nu})C_{\nu i} = 0 \ , \tag{2.15}$$

where $F_{\mu\nu}$ are the elements of the Fock matrix

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda,\sigma}^{M} P_{\lambda\sigma}\big[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu)\big] \ . \tag{2.16}$$

$S_{\mu\nu}$, $h_{\mu\nu}$, and $P_{\mu\nu}$ are the elements of overlap matrix, one-electron 'core' matrix, and closed-shell density matrix defined as

$$S_{\mu\nu} = \langle\phi_\mu|\phi_\nu\rangle = \int \phi_\mu^*(\mathbf{r}_\mu)\phi_\nu(\mathbf{r}_\nu)\,\mathrm{d}\mathbf{r} \ , \tag{2.17a}$$

$$h_{\mu\nu} = \langle\phi_\mu|\hat{h}|\phi_\nu\rangle = \int \phi_\mu^*(\mathbf{r}_\mu)\hat{h}\phi_\nu(\mathbf{r}_\nu)\,\mathrm{d}\mathbf{r} \ , \tag{2.17b}$$

$$P_{\mu\nu} = 2\sum_{a=1}^{N/2} C_{a\mu}^* C_{\nu a} \tag{2.17c}$$

and the two-electron AO integrals in (2.16) are defined as

$$(\mu\nu|\lambda\sigma) = \iint \phi_\mu^*(\mathbf{r}_{1\mu})\phi_\nu(\mathbf{r}_{1\nu})\frac{1}{r_{12}}\phi_\lambda^*(\mathbf{r}_{2\lambda})\phi_\sigma(\mathbf{r}_{2\sigma})\,\mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2 \ . \tag{2.18}$$

Similarly as in the case of HF equations, the Roothaan equations have to be solved iteratively, using the self-consistent field (SCF) method. The HF ground-state energy (2.12) is then evaluated as

$$E_0^{\mathrm{HF}} = \sum_{\mu,\nu}^{M} P_{\mu\nu}h_{\mu\nu} + \frac{1}{2}\sum_{\mu,\nu,\lambda,\sigma}^{M} P_{\mu\nu}P_{\lambda\sigma}\big[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu)\big] \ . \tag{2.19}$$

2.3 Correlation energy

Accuracy of the HF method is inherently limited by the Ansatz (2.7). The factorization of the wave function to one-electron functions leads to unphysical separation of electronic degrees of freedom and, therefore, to a too large probability of finding the electrons near each other. The omitted part of the electron-electron interaction is usually referred to as the dynamical correlation. Furthermore, the HF approximation fails for the multi-reference systems, i.e. for those where the description using a single configuration of electrons is inappropriate. The missing interaction is usually referred to as the non-dynamical (or static) correlation. Both dynamical and non-dynamical correlation contribute to the correlation energy defined as a difference between the exact non-relativistic Born–Oppenheimer ground-state energy E_0 and the HF energy,

$$E_{\mathrm{c}} \equiv E_0 - E_0^{\mathrm{HF}} \ . \tag{2.20}$$

As follows from the variational principle, the correlation energy is always negative.

The electron correlation is generally very important for the calculation of molecular properties. Numerous approximate methods have been developed to include it in the quantum chemistry calculations. Rigorous approximations are provided by the correlated *ab initio* methods such as post-HF methods [75, 80–82] (Møller–Plesset perturbation theory, configuration interaction, coupled-clusters theory) or multi-reference approaches [82]. The correlated *ab initio* methods are regarded as very robust and reliable since they offer a well-defined accuracy and there is a clear hierarchy of *ab initio* methods enabling systematic improvement of the calculated results. However, the computational requirements grow rapidly with the increasing level of approximation, what limits the use of the high-level *ab initio* methods only to small systems. A computationally much more feasible alternative is provided by the density-functional theory (DFT), where the electron correlation is approximated by the exchange-correlation potential. Most of the results presented in this thesis were obtained using DFT.

2.4 Density-functional theory

DFT is, in principle, an exact theory of electronic ground state formulated in terms of electron density $\rho(\mathbf{r})$ [83–86]. The idea of describing many-electron systems with one-electron density instead of wave function originates in the works of Thomas, Fermi, and Dirac from late 1920s [87–89]. However, working theory applicable to real systems was developed only in 1960s by Hohenberg, Kohn, and Sham [90, 91]. Originally, DFT was designed for calculations in periodic systems and solids. When sufficiently accurate DFT functionals appeared in late 1980s, DFT became vastly successful also for molecules, mainly due to its relatively good accuracy and low computational cost.

The electron density $\rho(\mathbf{r})$ is a positive real function describing the ground-state electron distribution that is observable e.g. in x-ray diffraction experiments. It is conceptually much simpler than the many-electron wave function $\Phi_0(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$. The electron density is defined as

$$\rho(\mathbf{r}) = n \int \cdots \int |\Phi_0(\mathbf{x}, \mathbf{x}_2, \ldots, \mathbf{x}_N)|^2 \, \mathrm{d}\sigma \, \mathrm{d}\mathbf{x}_2 \ldots \mathrm{d}\mathbf{x}_N \ . \tag{2.21}$$

It integrates to the number of electrons,

$$\int \rho(\mathbf{r}) \, \mathrm{d}\mathbf{r} = N \ , \tag{2.22}$$

and it exponentially decays for $\mathbf{r} \to \infty$. There are other physically sound requirements on the electron density (or related quantities) such as density cusp at nuclear positions or boundary conditions for the pair density and the associated exchange-correlation hole

[85, 92], which are invoked as boundary conditions at different DFT approximations.

2.4.1 Hohenberg–Kohn theorem

Recall the electronic Hamiltonian (2.6a), where the external potential $\hat{V}$ is written as a sum of one-electron contributions,

$$\hat{V} = \sum_{a=1}^{N} v(\mathbf{r}_a) \ . \tag{2.23}$$

The Hohenberg–Kohn (HK) theorem [90] states that the electron density $\rho(\mathbf{r})$ uniquely determines the one-electron potential $v(\mathbf{r})$ and, by integration (2.22), also the number of electrons n. Therefore, $\rho(\mathbf{r})$ fully determines the electronic Hamiltonian $\hat{H}$ and implicitly also the exact ground-state wave function $\Phi_0[\rho]$ and energy $E_0[\rho]$. The HK theorem implies that once $\rho(\mathbf{r})$ is known, all other ground-state properties of a given system are readily available and can be exactly evaluated.

The ground-state energy functional $E_0[\rho]$ can be expressed as

$$\begin{aligned} E_0[\rho] &= \langle \Phi_0 | \hat{T} | \Phi_0 \rangle + \langle \Phi_0 | \hat{U} | \Phi_0 \rangle + \langle \Phi_0 | \hat{V} | \Phi_0 \rangle \\ &= T[\rho] + U[\rho] + \int v(\mathbf{r}) \rho(\mathbf{r}) \, \mathrm{d}\mathbf{r} \ , \end{aligned} \tag{2.24}$$

where we defined the kinetic and electron-electron energy functionals $T[\rho]$ and $U[\rho]$, respectively. The functionals $T[\rho]$ and $U[\rho]$ are independent of $v(\mathbf{r})$, wherefore they are universal functionals of $\rho(\mathbf{r})$, the same for all electronic structure problems. The explicit forms of $T[\rho]$ and $U[\rho]$ are unknown but their existence is *guaranteed* by the HK theorem.

The ground-state electron density is found by applying the variational principle to minimize the energy functional $E_0[\rho]$ (2.24) subject to the constraint (2.22). The existence of the variational principle for $E_0[\rho]$ is often stated as the second HK theorem.

The electron-electron energy functional $U[\rho]$ can be decomposed as

$$U[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}\mathbf{r} \, \mathrm{d}\mathbf{r}' + E_{\mathrm{ncl}}[\rho] \ , \tag{2.25}$$

where the first term is a classical Coulomb repulsion $J[\rho]$ appearing also in the HF theory (the second term in (2.12)) and $E_{\mathrm{ncl}}[\rho]$ is a very important non-classical energy contribution. Within the HK framework, the functionals $T[\rho]$ and $E_{\mathrm{ncl}}[\rho]$ are the unknown parts of $E_0[\rho]$, which need to be approximated. Unfortunately, direct approximations of these functionals turned out to be rather poor even for atoms, mainly due to unsatisfactory models of $T[\rho]$ [83].

2.4.2 Kohn–Sham formulation of DFT

Kohn and Sham [91] introduced the Thomas–Fermi model of non-interacting uniform electron gas as a reference system for calculation of $T[\rho]$. The non-interacting Hamiltonian

$$\hat{H}_{\mathrm{s}} = \hat{T} + \hat{V} = \sum_{a=1}^{N} \Big(-\tfrac{1}{2}\nabla_a^2 + v(\mathbf{r}_a) \Big) \tag{2.26}$$

contains only one-electron operators and, therefore, the non-degenerate ground-state wave function is exactly given by a single Slater determinant $\Phi_{\mathrm{s}}[\rho]$. The electron density can be expressed in terms of occupied spin orbitals as

$$\rho(\mathbf{r}) = \sum_{a=1}^{N} |\chi_a(\mathbf{r})|^2 \tag{2.27}$$

and the kinetic energy of the model non-interacting system is

$$T_{\mathrm{s}}[\rho] = \Big\langle \Phi_{\mathrm{s}}[\rho] \Big| \hat{T} \Big| \Phi_{\mathrm{s}}[\rho] \Big\rangle = \sum_{a=1}^{N} \langle \chi_a | -\tfrac{1}{2}\nabla_a^2 | \chi_a \rangle \ . \tag{2.28}$$

In the interacting system, the functional $T_{\mathrm{s}}[\rho]$ contains almost all of the true kinetic energy $T[\rho]$. Hence, the total electronic energy of the interacting system (2.24) can be in the Kohn–Sham (KS) formalism rewritten as

$$E_0[\rho] = T_{\mathrm{s}}[\rho] + J[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})\,\mathrm{d}\mathbf{r} + E_{\mathrm{xc}}[\rho] \ , \tag{2.29}$$

where the first three terms are dominant and can be calculated exactly. The residual functional $E_{\mathrm{xc}}[\rho]$ defined as

$$E_{\mathrm{xc}}[\rho] = T[\rho] - T_{\mathrm{s}}[\rho] + U[\rho] - J[\rho] \tag{2.30}$$

is the exchange-correlation energy that remains to be found. The exact form of $E_{\mathrm{xc}}[\rho]$ is unknown but it satisfies a number of boundary conditions which can be imposed also on approximate DFT functionals. Moreover, the existence of exact $E_{\mathrm{xc}}[\rho]$ is guaranteed by the HK theorem, what encourages constant development in this filed.

The central idea of the KS approach is based on transforming the Hamiltonian of a real (interacting) many-electron system into the form of model Hamiltonian (2.26), for which $T_{\mathrm{s}}[\rho]$ is the exact kinetic energy. This can be done by introducing the effective KS potential

$$v^{\mathrm{KS}}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\,\mathrm{d}\mathbf{r}' + v_{\mathrm{xc}}(\mathbf{r}) \ , \tag{2.31}$$

where $v_{\mathrm{xc}}(\mathbf{r})$ is the exchange-correlation potential defined as

$$v_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho} \ . \tag{2.32}$$

The electronic Hamiltonian then adopts the desired form

$$\hat{H}^{\mathrm{KS}} = \sum_{a=1}^{N} \Big(-\tfrac{1}{2} \nabla_a^2 + v^{\mathrm{KS}}(\mathbf{r}_a) \Big) \ , \tag{2.33}$$

which corresponds to a system of non-interacting electrons moving in the external potential $v^{\mathrm{KS}}(\mathbf{r})$. The Schrödinger equation with the Hamiltonian (2.33) has an exact solution, the Slater determinant Φ^{KS} formed out of occupied KS spin orbitals χ_a^{KS}. In other words, the insufficiency of the single-determinant approximation is in KS-DFT remedied by adjusting the Hamiltonian, so that Φ^{KS} is an exact eigenfunction of $\hat{H}^{\mathrm{KS}}$. Obviously, $\hat{H}^{\mathrm{KS}}$ and Φ^{KS} need not describe any real system, as they were constructed only to provide exact $E_0[\rho]$.

The KS spin orbitals χ_i^{KS} are found by applying the variational principle to minimize the ground-state energy functional $E_0[\rho]$ (2.29) under the orthonormality constraint $\langle \chi_i | \chi_j \rangle = \delta_{ij}$:

$$\hat{f}^{\mathrm{KS}} \chi_i^{\mathrm{KS}} = \varepsilon_i^{\mathrm{KS}} \chi_i^{\mathrm{KS}} \ , \tag{2.34}$$

where

$$\hat{f}^{\mathrm{KS}} = -\frac{1}{2} \nabla^2 + v^{\mathrm{KS}}(\mathbf{r}) \tag{2.35}$$

is the KS operator. The KS equations (2.34) formally resemble the HF equations (2.9) but, same as for the KS wave function Φ^{KS}, the spin orbitals χ_i^{KS} and the eigenvalues $\varepsilon_i^{\mathrm{KS}}$ have no strict physical significance [84]. Nevertheless, KS-DFT is in principle exact theory of the electronic ground state, where the only error (in a complete basis set limit) arises due to the approximation of $E_{\mathrm{xc}}[\rho]$. The KS operator $\hat{f}^{\mathrm{KS}}$ implicitly depends on the ground-state density $\rho(\mathbf{r})$ and thus, through (2.27), also on its own eigenfunctions χ_i^{KS}. The KS equations are thus non-linear and have to be solved by an iterative self-consistent procedure.

2.4.3 Exchange-correlation functionals

The single-determinant solution of KS equations suggests that the exchange-correlation energy functional $E_{\mathrm{xc}}[\rho]$ can be formally decomposed into the exchange and correlation contributions $E_{\mathrm{x}}[\rho]$ and $E_{\mathrm{c}}[\rho]$ analogous (but not equal) to those known from the HF theory:

$$E_{\mathrm{xc}}[\rho] = E_{\mathrm{x}}[\rho] + E_{\mathrm{c}}[\rho] \ . \tag{2.36}$$

The exchange energy is in the KS framework defined as

$$E_{\mathrm{x}}[\rho] = \left\langle \Phi^{\mathrm{KS}}[\rho] \middle| \hat{U} \middle| \Phi^{\mathrm{KS}}[\rho] \right\rangle - J[\rho] \tag{2.37}$$

and it contains major portion of $E_{\mathrm{xc}}[\rho]$. $E_{\mathrm{x}}[\rho]$ dominates by far the non-classical energy $E_{\mathrm{ncl}}[\rho]$ defined in (2.25) and it is usually very close to the exact HF exchange energy (third term in (2.12)).

The KS correlation energy $E_{\mathrm{c}}[\rho]$ is defined by (2.36). It can be rearranged as

$$E_{\mathrm{c}}[\rho] = T[\rho] - T_{\mathrm{s}}[\rho] + E_{\mathrm{ncl}}[\rho] - E_{\mathrm{x}}[\rho] \ , \tag{2.38}$$

where the kinetic energy difference $T[\rho] - T_{\mathrm{s}}[\rho]$ is always positive, while the potential energy contribution $E_{\mathrm{ncl}}[\rho] - E_{\mathrm{x}}[\rho]$ is about twice as large and always negative.

Local density approximation. The simplest form of $E_{\mathrm{xc}}[\rho]$ is the local density approximation (LDA) introduced already by Kohn and Sham [91]. It has a general form

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[\rho] = \int \epsilon_{\mathrm{xc}}[\rho]\, \rho(\mathbf{r})\, \mathrm{d}\mathbf{r} \ , \tag{2.39}$$

where $\epsilon_{\mathrm{xc}}[\rho]$ is the exchange-correlation energy per particle of uniform interacting electron gas with density ρ. The exchange part of $\epsilon_{\mathrm{xc}}[\rho]$ is in LDA given by the uniform-electron-gas exchange energy functional of Dirac [89],

$$\epsilon_{\mathrm{x}}^{\mathrm{LDA}}[\rho] = -C_{\mathrm{x}} \rho^{1/3} \ , \qquad C_{\mathrm{x}} = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} \ , \tag{2.40}$$

which is, apart from a prefactor, equivalent to the well-known X_α approximation of the HF exchange potential by Slater [93]. Approximating the correlation contribution $\epsilon_{\mathrm{c}}[\rho]$ required much more effort. The most widely used local approximations of $\epsilon_{\mathrm{c}}[\rho]$ were developed by Vosko, Wilk, and Nusair (functionals III and V of ref. [94], usually denoted VWN3 and VWN5, respectively).

Open-shell systems can be described with a spin-polarized extension of LDA, the local spin density approximation (LSDA). For example, the LSDA exchange energy is

$$E_{\mathrm{x}}^{\mathrm{LSDA}}[\rho_\alpha, \rho_\beta] = -C_{\mathrm{x}} \int \left(\rho_\alpha^{4/3}(\mathbf{r}) + \rho_\beta^{4/3}(\mathbf{r})\right) \mathrm{d}\mathbf{r} \ , \tag{2.41}$$

where ρ_α and ρ_β are the densities of α and β electrons, respectively, $\rho_\alpha + \rho_\beta = \rho$. LSDA is superior to LDA even for closed-shell systems because introducing two density functions instead of one increases the flexibility of any approximate exchange-correlation functional $E_{\text{xc}}[\rho_\alpha, \rho_\beta]$. All currently used DFT functionals are spin-polarized.

The local approximations are valid only for slowly varying densities [83]. It is thus somewhat surprising that LSDA performs well also in molecules, where it provides similar accuracy as the HF method for equilibrium geometries, vibrational frequencies, and charge moments [84]. The success of LSDA can be understood by realizing that the density of the uniform electron gas naturally fulfills the boundary conditions for the exchange-correlation hole [85,92], which were not a priori assumed in the LSDA model. Nevertheless, LSDA has a clear overbinding tendency and it fails completely for magnetic properties. Improved approximations of $E_{\text{xc}}[\rho]$ are therefore desired.

Generalized gradient approximation. The next logical step in improving the approximations of $E_{\text{xc}}[\rho]$ is including the dependence on the density gradient $\nabla\rho$. However, it turned out that a simple correction using a simple gradient expansion disturbs the naturally favorable properties of LSDA such as correct behavior of the exchange-correlation hole [85]. These properties are restored within a somewhat more elaborate generalized gradient approximation (GGA), which represents a major improvement over LSDA. The increase in computational cost due to the evaluation of $\nabla\rho_\alpha$ and $\nabla\rho_\beta$ is only modest. In most cases, the gradient-corrected exchange-correlation energy can be expressed as

$$\begin{aligned} E_{\text{xc}}^{\text{GGA}} &= \int f_{\text{xc}}(\rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\beta\beta}, \gamma_{\alpha\beta})\,\mathrm{d}\mathbf{r}\ , \\ \gamma_{\sigma\sigma'} &= \nabla\rho_\sigma \cdot \nabla\rho_{\sigma'}\ , \end{aligned} \tag{2.42}$$

where f_{xc} is a function only of densities ρ_α, ρ_β, and their gradient invariants $\gamma_{\sigma\sigma'}$, $\sigma, \sigma' = \{\alpha, \beta\}$.

The gradient correction is most important for the exchange part of $E_{\text{xc}}[\rho]$. The gradient-corrected exchange energy functionals typically have a general form introduced by Perdew in his P86 functional [95]:

$$\begin{aligned} E_{\text{x}}^{\text{GGA}} &= E_{\text{x}}^{\text{LSDA}} - \sum_{\sigma=\alpha,\beta} \int F(\xi_\sigma)\, \rho_\sigma^{4/3}(\mathbf{r})\,\mathrm{d}\mathbf{r} \\ \xi_\sigma &= \frac{|\nabla\rho_\sigma(\mathbf{r})|}{\rho_\sigma^{4/3}(\mathbf{r})} \end{aligned} \tag{2.43}$$

The function $F(\xi_\sigma)$ defines the particular GGA exchange functional and its argument ξ_σ is the dimensionless reduced density gradient, which can be interpreted as a local inhomogeneity parameter. Probably the most commonly used GGA exchange correction is the Becke's 1988 functional [96]

$$\Delta E_{\mathrm{x}}^{\mathrm{B88}} = -\sum_{\sigma=\alpha,\beta} \int \frac{b\,\xi_\sigma^2}{1 + 6b\,\xi_\sigma \sinh^{-1} \xi_\sigma} \rho_\sigma^{4/3}(\mathbf{r})\, \mathrm{d}\mathbf{r} \ , \tag{2.44}$$

where $b = 0.0042$ a.u. is the only parameter fitted such that the sum $E_{\mathrm{x}}^{\mathrm{LSDA}} + \Delta E_{\mathrm{x}}^{\mathrm{B88}}$ correctly reproduces the exact exchange energy of noble gas atoms. Apart from correct properties of the associated electron density, the exchange energy $E_{\mathrm{x}}^{\mathrm{B88}}$ also reproduces the asymptotic behavior of exact E_{x} for $r \to \infty$. Other notable GGA exchange functionals are the parameter-free functionals PW91 (Perdew and Wang, 1991) [97, 98] and PBE (Perdew, Burke, and Ernzerhof, 1996) [99].

The GGA corrections to the correlation energy E_{c} are much smaller and their forms are substantially more complicated than those for exchange. Among the most widely used GGA correlation functionals are P86 [100], the parameter-free functionals PW91 [97, 98] and PBE [99], and probably the most popular functional LYP (Lee, Yang, and Parr, 1988). Unlike other functionals mentioned so far, LYP is not based on the uniform electron gas but rather on the correlation energy of the helium atom [101]. Hence, LYP does not involve the LSDA correlation and it does not have an easily separable local component.

Reliable DFT approximations are usually developed by imposing physically rigorous requirements on density, exchange-correlation hole, and exchange-correlation functional itself. However, since DFT is inherently semi-empirical, the rigorous approximations are sometimes outperformed by more arbitrary functionals due to simple error cancellation. In other words, it is not the physics but the results which dictate the choice of the mathematical constructs [85]. The same pragmatic view so typical for DFT in general applies also for pairing of the exchange and correlation functionals. Thus, although they were not specifically designed to work together, the GGA combinations known as BP86, BPW91, or BLYP typically provide better results than the others. The principle of favorable error cancellation applies even more profoundly to hybrid DFT functionals.

Hybrid DFT functionals. The performance of GGA functionals may be improved by adding the exact HF exchange into E_{xc}. Although the reasons may not be immediately seen, Becke gave a straightforward explanation for the need of non-local exchange by invoking the concept of adiabatic connection between the non-interacting KS reference system (2.26) and the fully interacting system (2.33) [102]. Becke combined the admixtures of B88 exchange, PW91 correlation, and exact HF exchange into his famous three-parameter hybrid functional [103], which is known today as B3PW91. A slight modification was suggested soon afterwards, replacing the PW91 correlation with the LYP correlation [104] and arriving thus at the B3LYP functional in the well-known form

$$E_{\mathrm{xc}}^{\mathrm{B3LYP}} = (1-a)E_{\mathrm{x}}^{\mathrm{LSDA}} + aE_{\mathrm{x}}^{\mathrm{exact}} + b\,\Delta E_{\mathrm{x}}^{\mathrm{B88}} + (1-c)E_{\mathrm{c}}^{\mathrm{LSDA}} + cE_{\mathrm{c}}^{\mathrm{LYP}} \ , \tag{2.45}$$

where a, b, c are the fitted parameters of Becke [103],

$$a = 0.20 \ , \qquad b = 0.72 \ , \qquad c = 0.81 \ .$$

The portion of exact exchange is given by the parameter a; the values around $20 - 25\,\%$ seem to be the amount needed also for other general-purpose hybrid functionals. Note that Vosko, Wilk, and Nusair originally recommended to use their VWN5 variant for the LSDA correlation [94] but the implementation of B3LYP within the `Gaussian` program uses the VWN3 variant [105].

B3PW91 was originally parameterized to reproduce correct thermochemistry [103], but both B3PW91 and its twin B3LYP perform unexpectedly well also in many other areas. To date B3LYP has been applied probably in all major fields of computational chemistry, providing acceptable results in most of the cases. Such universality is unique and probably very fortunate but it keeps being massively exploited throughout the literature and it was extensively applied also in this work.

2.4.4 LCAO method in DFT

Same as in the HF method, the LCAO expansion (2.14) is usually introduced in KS-DFT, enabling the use of standard computational machinery developed for the *ab initio* methods. The matrix form of the KS equations is the same as that of the spin-dependent variant of the Roothaan equations (the Pople–Nesbet equations), where the Fock matrix is replaced with the KS matrix

$$F_{\mu\nu}^{\mathrm{KS}\alpha} = h_{\mu\nu} + \sum_{\lambda,\sigma}^{M} P_{\lambda\sigma}(\mu\nu|\lambda\sigma) + \int \phi_{\mu}^{*}(\mathbf{r}_{\mu}) v_{\mathrm{xc}}^{\alpha}(\mathbf{r}) \phi_{\nu}(\mathbf{r}_{\nu})\,\mathrm{d}\mathbf{r} \ , \tag{2.46}$$

where $P_{\lambda\sigma} = P_{\lambda\sigma}^{\alpha} + P_{\lambda\sigma}^{\beta}$; analogous expression applies for the spin β. The integrals involving $v_{\mathrm{xc}}^{\alpha}(\mathbf{r})$ and $v_{\mathrm{xc}}^{\beta}(\mathbf{r})$ are mostly not known analytically and have to be evaluated by numerical quadrature. An efficient scheme for the calculation of these integrals with the GGA functionals of the form (2.42) was presented by Pople et al. [106]. The KS ground-state energy is given by

$$E_0^{\mathrm{KS}} = \sum_{\mu,\nu}^{M} P_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu,\nu,\lambda,\sigma}^{M} P_{\mu\nu} P_{\lambda\sigma}(\mu\nu|\lambda\sigma) + E_{\mathrm{xc}}[\rho] \ . \tag{2.47}$$

2.4.5 Performance of DFT methods

Unlike the *ab initio* wave function methods, DFT does not have a strict hierarchy of approximations that would gradually improve towards the exact theory. There is no known systematic way within DFT that would allow to reach an arbitrarily high level of precision [84]. Nevertheless, following the development of exchange-correlation functionals as outlined above, the first few 'rungs' on the Jacob's ladder of DFT accuracy [92, 107] are nowadays well-established. Starting from the Hartree approximation (which can be regarded as the zeroth level of DFT) with no exchange or correlation whatsoever and passing the local approximations, the GGA functionals are the first that provide results of acceptable quality for a large scale of chemical problems and quantities at a very favorable computational cost. Hybrid DFT functionals mostly provide a distinct improvement over the GGAs. They have somewhat higher computational demands due to the calculation of HF exchange, but a moderate increase in computer time is usually not an issue for the calculation of molecular properties. Hybrid DFT functionals, and B3LYP in particular, are the workhorses of current computational spectroscopy.

As already mentioned, all DFT approximations are inherently semi-empirical and their accuracy is limited. Therefore, the DFT results should be always compared with the available experimental data. However, a simple comparison with the experiment provides only a very rough idea about the actual DFT error since there are several other possible sources of computational imperfections (see section 2.6). For many molecular properties, benchmark *ab initio* methods are nowadays available and the results for small model systems can be used to reliably monitor the DFT accuracy. Such an approach was employed also in this work.

2.5 Molecular properties

Molecular properties characterize the response of electronic system of a molecule to the external perturbations such as electric or magnetic fields, nuclear magnetic moments, electronic spin, nuclear displacements from equilibrium positions, and combinations of thereof. The electronic Hamiltonian $\hat{H}$, its exact eigenfunctions $|\Phi_i\rangle$ and energies E_i in the presence of two different static external perturbations can be expanded as

$$\hat{H} = \hat{H}^{(0)} + \kappa\hat{H}^{(10)} + \lambda\hat{H}^{(01)} + \kappa^2\hat{H}^{(20)} + \kappa\lambda\hat{H}^{(11)} + \lambda^2\hat{H}^{(02)} + \dots \tag{2.48a}$$

$$|\Phi_i\rangle = |i\rangle + \kappa|\Phi_i^{(10)}\rangle + \lambda|\Phi_i^{(01)}\rangle + \kappa^2|\Phi_i^{(20)}\rangle + \kappa\lambda|\Phi_i^{(11)}\rangle + \lambda^2|\Phi_i^{(02)}\rangle + \dots \tag{2.48b}$$

$$E_i = E_i^{(0)} + \kappa E_i^{(10)} + \lambda E_i^{(01)} + \kappa^2 E_i^{(20)} + \kappa\lambda E_i^{(11)} + \lambda^2 E_i^{(02)} + \dots \tag{2.48c}$$

where $\hat{H}^{(kl)}$, $|\Phi_i^{(kl)}\rangle$, and $E_i^{(kl)}$ are the $(k+l)^{\mathrm{th}}$-order perturbations, wave functions, and molecular properties of the i-th electronic electronic state. The zeroth-order quantities

Table 2.1 Ground-state molecular properties as energy derivatives.

Energy derivative[a]	Molecular property	Related observables
$\frac{\mathrm{d}E_0}{\mathrm{d}\mathbf{R}_A}$	Energy gradient	equilibrium geometry
$\frac{\mathrm{d}^2E_0}{\mathrm{d}\mathbf{R}_A\mathrm{d}\mathbf{R}_B}$	molecular force field	vibrational frequencies
$\frac{\mathrm{d}^2E_0}{\mathrm{d}\mathbf{m}_K\mathrm{d}\mathbf{B}_0}$	NMR chemical shielding tensor	chemical shift, chemical shielding anisotropy
$\frac{\mathrm{d}^2E_0}{\mathrm{d}\mathbf{m}_K\mathrm{d}\mathbf{m}_L}$	nucleus-independent J-coupling tensor	scalar spin–spin coupling (in Hz)
$\frac{\mathrm{d}^3E_0}{\mathrm{d}\mathbf{R}_A\mathrm{d}^2\boldsymbol{\varepsilon}}$	electric polarizability derivatives	Raman intensities

[a] with respect to nuclear coordinates $\mathbf{R}_A$, nuclear magnetic moments $\mathbf{m}_K$, static magnetic field $\mathbf{B}_0$, and electric field $\boldsymbol{\varepsilon}$

$\hat{H}^{(0)}$, $|i\rangle \equiv |\Phi_i^{(0)}\rangle$, and $E_i^{(0)}$ correspond to the unperturbed Schrödinger equation (2.6a). The $(k+l)^{\text{th}}$-order properties $E_i^{(kl)}$ can be expressed as energy derivatives with respect to the perturbations κ and λ [108],

$$E_i^{(kl)} = \left.\frac{\mathrm{d}^{(k+l)}E_i}{\mathrm{d}\kappa^k\mathrm{d}\lambda^l}\right|_{\kappa=\lambda=0} . \tag{2.49}$$

This work involved only ground-state properties $E_0^{(kl)}$ $(i = 0)$ of closed-shell molecular systems, as summarized in Table 2.1.

In practice, the energy derivatives can be evaluated either analytically (provided that higher-order wave functions can be calculated) or numerically, by finite-differentiation techniques. The analytical self-consistent scheme is usually preferable [108] since it provides reliable results with a well-defined accuracy. The numerical differentiation typically requires several lower-order energy evaluations and the control of its accuracy may not be guaranteed. Moreover, the use of the numerical approach for magnetic perturbations is complicated by the fact that the perturbed wave functions are no longer real-valued and need not be pure singlet even for closed-shell systems (see section 3.3). On the other hand, the numerical approach can be quite useful for obtaining geometry derivatives. The numerical calculation of geometry derivatives does not require extensive coding, it is accessible to massive parallelization, and it can be conveniently combined with the analytical scheme to obtain higher-order properties such as geometry derivatives of NMR parameters, anharmonic frequencies or Raman intensities.

2.5.1 First-order properties

Probably the most important first-order molecular property is the energy gradient, which is a key quantity for finding equilibrium molecular structures by minimization of energy and molecular forces (the procedure known as the geometry optimization). The geometry optimization was employed throughout this work as a basic computational tool and the method of calculation of molecular forces will be thus briefly introduced.

The energy gradient was the first molecular property calculated at a rigorous *ab initio* level [109] and, as such, it serves as a good example of practical issues arising in the calculations of molecular properties. A general first-order ground-state property $E_0^{(1)}$ corresponding to a perturbation $\hat{H}^{(1)}$ is given by the Hellmann–Feynman theorem [110]

$$E_0^{(1)} = \langle 0|\hat{H}^{(1)}|0\rangle \ . \tag{2.50}$$

This theorem, however, holds only for an exact ground-state wave function $|0\rangle$ (and for certain special cases such as the HF complete basis set limit). Evaluation of the molecular gradient according to (2.50) with an approximate wave function gives rise to a fictitious wave-function force (Pulay force) [109]. This computational artifact can be avoided by calculating the HF energy gradient as a total derivative of the energy $E_0^{\text{HF}} + U_{\text{nuc}}$ with respect to nuclear coordinates $\mathbf{R}_A$ [75,109]:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\mathbf{R}_A}(E_0^{\text{HF}} + U_{\text{nuc}}) = \sum_{\mu,\nu}^{M} P_{\mu\nu}\frac{\partial h_{\mu\nu}}{\partial \mathbf{R}_A} + \frac{1}{2}\sum_{\mu,\nu,\lambda,\sigma}^{M} P_{\mu\nu}P_{\lambda\sigma}\frac{\partial\big[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu)\big]}{\partial \mathbf{R}_A}\\ - \sum_{\mu,\nu}^{M}\frac{\partial S_{\mu\nu}}{\partial \mathbf{R}_A}\sum_{a=1}^{N}\varepsilon_a C_{\nu a}C_{a\mu}^{*} + \frac{\partial U_{\text{nuc}}}{\partial \mathbf{R}_A} \ .\end{aligned} \tag{2.51}$$

An analogous expression for efficient calculation of the KS energy gradient was given by Pople et al. [106].

2.5.2 Time-independent perturbation theory

The expressions for higher-order ground-state properties $E_0^{(kl)}$ can be derived using the Rayleigh–Schödinger pertubation theory [75,111]. The Hamiltonian $\hat{H}$, ground-state wave function $|\Phi_0\rangle$, and energy E_0 (2.48) in the presence of perturbations satisfy the time-independent Schrödinger equation

$$\hat{H}|\Phi_0\rangle = E_0|\Phi_0\rangle \ . \tag{2.52}$$

By collecting the terms of individual perturbation orders and applying the wave-function normalization condition, the higher-order molecular properties are obtained. In particular, a general second-order ground-state property can be expressed as

$$\begin{aligned} E_0^{(11)} &= \langle 0|\hat{H}^{(11)}|0\rangle + 2\,\langle 0|\hat{H}^{(10)}|\Phi_0^{(01)}\rangle \\ &= \langle 0|\hat{H}^{(11)}|0\rangle + 2\,\langle 0|\hat{H}^{(01)}|\Phi_0^{(10)}\rangle \ , \end{aligned} \tag{2.53}$$

where the first term can be readily evaluated using the unperturbed ground-state wave function $|0\rangle$ as a simple expectation value of the second-order Hamiltonian $\hat{H}^{(11)}$. The second term of (2.53) can be expressed in two equivalent ways using the first-order wave functions $|\Phi_0^{(10)}\rangle$ or $|\Phi_0^{(01)}\rangle$. Note that the second-order wave functions are not needed to calculate the second-order property. In general, Wigner's $(2n+1)$-rule applies, stating that the nth-order wave functions are sufficient to evaluate the $(2n+1)^{\text{th}}$-order properties [108, 112].

Assuming non-degenerate ground state $|0\rangle$, the first-order wave functions $|\Phi_0^{(10)}\rangle$ and $|\Phi_0^{(01)}\rangle$ can be expanded into a basis of virtual states $|m\rangle$, $m \neq 0$, and the second-order property $E_0^{(11)}$ can be rewritten into the sum-over-states (SOS) form

$$E_0^{(11)} = \langle 0|\hat{H}^{(11)}|0\rangle + 2\sum_{m\neq 0} \frac{\langle 0|\hat{H}^{(10)}|m\rangle\langle m|\hat{H}^{(01)}|0\rangle}{E_0^{(0)} - E_m^{(0)}} \ . \tag{2.54}$$

In the theory of NMR parameters, the first term in (2.54) is usually called diamagnetic and the SOS expansion is referred to as the paramagnetic term. Unfortunately, the exact virtual states $|m\rangle$ are almost never available and the SOS calculations employing approximate states $|m\rangle$ and energies $E_m^{(0)}$ can be quite inaccurate. The SOS expansions are mainly used to gain insight into the theoretical characteristics of molecular properties. The actual calculations are most usually carried out in the framework of perturbed HF or KS equations.

2.5.3 Coupled-perturbed scheme

For accurate and rigorous calculation of molecular properties, the perturbed wave functions have to be found in a self-consistent manner. In the coupled-perturbed (CP) scheme, instead of focusing on the total wave function, the perturbational expansion is applied directly to the HF or KS equations. Here we concentrate on the coupled-perturbed Hartree–Fock (CPHF) method [113] for ground-state molecular properties.

The perturbation $\hat{H}^{(kl)}$ can be decomposed into contributions acting on individual occupied spin orbitals χ_a:

$$\hat{H}^{(kl)} = \sum_{a=1}^{N} \hat{h}^{(kl)}(a) \ . \tag{2.55}$$

Assuming the perturbed ground-state wave function $|\Phi_0\rangle$ (2.48b) in a form of Slater determinant, the second-order ground-state property $E_0^{(11)}$ (2.53) can be expressed as

$$E_0^{(11)} = \sum_{a=1}^{N} \langle \chi_a^{(0)} | \hat{h}^{(11)} | \chi_a^{(0)} \rangle + 2 \sum_{a=1}^{N} \langle \chi_a^{(0)} | \hat{h}^{(10)} | \chi_a^{(01)} \rangle \ . \tag{2.56}$$

Same as in the perturbation theory, the calculation of expectation values of the second-order operator $\hat{h}^{(11)}$ with the unperturbed spin orbitals $\chi_a^{(0)}$ is straightforward. $\chi_a^{(01)}$ are the first-order spin orbitals that remain to be found.

The Fock operator $\hat{f}$, the occupied spin orbitals χ_a, and the orbital energies ε_a can be expanded in analogy to (2.48). Inserting these expansions into the HF equations (2.9) and collecting the '(01)' terms we get the perturbed HF equations [113–115]

$$\big(\hat{f}^{(01)} - \varepsilon_a^{(01)}\big)\chi_a^{(0)} + \big(\hat{f}^{(0)} - \varepsilon_a^{(0)}\big)\chi_a^{(01)} = 0 \tag{2.57}$$

for the unknown first-order spin orbitals $\chi_a^{(01)}$. The CPHF scheme thus requires previous solution of the unperturbed HF equations. The first-order Fock operator $\hat{f}^{(01)}$ can be decomposed into the one- and two-electron parts,

$$\hat{f}^{(01)} = \hat{h}^{(01)} + \hat{g}^{(01)} \ , \tag{2.58}$$

where the first-order two-electron operator $\hat{g}^{(01)}$ consists of the first-order Coulomb and exchange operators $\hat{J}_b^{(01)}$ and $\hat{K}_b^{(01)}$:

$$\hat{g}^{(01)} = \sum_{b=1}^{N} \Big(\hat{J}_b^{(01)} - \hat{K}_b^{(01)} \Big) \ , \tag{2.59a}$$

$$\hat{J}_b^{(01)} \chi_a^{(0)}(1) = \left[\int \frac{1}{r_{12}} \Big(\chi_b^{(01)*}(2)\chi_b^{(0)}(2) + \chi_b^{(0)*}(2)\chi_b^{(01)}(2) \Big) \, \mathrm{d}2 \right] \chi_a(1) \ , \tag{2.59b}$$

$$\begin{aligned} \hat{K}_b^{(01)} \chi_a^{(0)}(1) = & \left[\int \frac{1}{r_{12}} \chi_b^{(01)*}(2)\chi_a^{(0)}(2) \, \mathrm{d}2 \right] \chi_b^{(0)}(1) \\ & + \left[\int \frac{1}{r_{12}} \chi_b^{(0)*}(2)\chi_a^{(0)}(2) \, \mathrm{d}2 \right] \chi_b^{(01)}(1) \ , \end{aligned} \tag{2.59c}$$

where, for brevity, the indexes 1 and 2 were used instead of the electron coordinates $\mathbf{x}_1$ and $\mathbf{x}_2$. Through $\hat{g}^{(01)}$, the first-order Fock operator $\hat{f}^{(01)}$ depends on the first-order occupied spin orbitals $\chi_b^{(01)}$. At the same time, each perturbed spin orbital $\chi_a^{(01)}$ is generated by $\hat{f}^{(01)}$ as a solution of the perturbed HF equations (2.57). $\hat{f}^{(01)}$ and $\chi_a^{(01)}$ are thus coupled and the equations (2.57) have to be solved iteratively.

The first-order spin orbitals $\chi_a^{(01)}$ can be expanded into the basis of unperturbed virtual spin orbitals $\chi_q^{(0)}$,

$$\chi_a^{(01)} = \sum_{q>N} d_{qa}^{(01)} \chi_q^{(0)} \ , \tag{2.60}$$

where $d_{qa}^{(01)}$ are the expansion coefficients. By substituting (2.60) into (2.57), multiplying with $\langle\chi_q^{(0)}|$ and integrating over the electron coordinates we obtain the expansion coefficients

$$d_{qa}^{(01)} = \frac{\langle\chi_q^{(0)}|\hat{f}^{(01)}|\chi_a^{(0)}\rangle}{\varepsilon_a^{(0)} - \varepsilon_q^{(0)}} \ . \tag{2.61}$$

This is the basic equation used in the coupled-perturbed iterative scheme. In the initial step, the perturbation $\hat{h}^{(01)}$ is taken as a first guess for the perturbed Fock operator $\hat{f}^{(01)}$. The expansion coefficients $d_{qa}^{(01)}$ are determined from (2.61) and inserted into (2.60). The perturbed spin orbitals $|\chi_a^{(01)}\rangle$ are then used to construct new $\hat{f}^{(01)}$ according to (2.58) and (2.59). The CPHF procedure continues until the convergence of both $\hat{f}^{(01)}$ and $d_{qa}^{(01)}$. Finally, using the converged expansion coefficients $d_{qa}^{(01)}$, the first-order spin orbitals $\chi_a^{(01)}$ are found according to (2.60) and the corresponding one-electron contributions to the second-order ground-state property $E_0^{(11)}$ (2.56) are calculated as

$$\langle\chi_a^{(0)}|\hat{h}^{(10)}|\chi_a^{(01)}\rangle = \sum_{q>N} \langle\chi_a^{(0)}|\hat{h}^{(10)}|\chi_q^{(0)}\rangle d_{qa}^{(01)} \ . \tag{2.62}$$

Note that the CP equations need to be solved only for one of the perturbations '(01)' and '(10)'. Since the CP iterative procedure is the most time-consuming step in molecular property calculations, a suitable choice of the perturbation entering the CP scheme can substantially reduce the computational costs.

The general CP theory can be transformed into working self-consistent matrix equations by introducing the LCAO expansion (2.14). The actual implementations of the CP scheme are specific for each molecular property and usually exploit the particular form of the perturbation operators $\hat{h}^{(01)}$ and $\hat{h}^{(10)}$.

CPKS method. The KS variant of the coupled-perturbed scheme [116] can be obtained from the CPHF by substituting the exchange part of the first-order Fock operator $\hat{f}^{(01)}$ with the first-order exchange-correlation potential $v_{\mathrm{xc}}^{(01)}$:

$$\hat{f}^{\mathrm{KS}\,(01)} = \hat{h}^{(01)} + \sum_{b=1}^{N} \hat{J}_b^{(01)} + v_{\mathrm{xc}}^{(01)} \ . \tag{2.63}$$

The particular form of $v_{\mathrm{xc}}^{(01)}$ is specific for each DFT functional type.

Magnetic perturbations represent a specific case for the CPKS scheme since they involve imaginary perturbation operators [117, 118]. The pure imaginary perturbations lead to pure imaginary first-order KS spin orbitals $\chi_a^{\mathrm{KS}\,(01)}$, which yield zero contribution to the total density. Therefore, both Coulomb and exchange-correlation parts of $\hat{f}^{\mathrm{KS}\,(01)}$ vanish and the calculations of the corresponding perturbation terms with non-hybrid DFT functionals reduce to an uncoupled (non-iterative) case. Note, however, that the imaginary first-order spin orbitals contribute to the first-order current density. Therefore, if a functional dependent on current density would be used, the imaginary perturbation operators would enter the CPKS procedure [119].

2.6 General notes on the computational methodology

Equilibrium molecular geometries were usually optimized using the DFT method (mostly the B3LYP functional) with a moderately sized basis set (such as 6-31+G(d)) and the implicit polarizable continuum model (PCM) [120–122] or COSMO-PCM (CPCM) [123] hydration. The most typical NMR computational method also employed the B3LYP functional, as it is known to perform well for the NMR parameters [118, 124], the IGLO-III basis set [125], and the implicit PCM hydration. This approach, further denoted DFT/PCM, proved very useful for obtaining qualitative NMR results for large numbers of molecular conformers.

The DFT/PCM results were compared to the experimental data whenever possible and the differences were explained in terms of computational corrections due to various approximations underlying the DFT/PCM method. A perturbational view of these corrections was adopted, assuming their simple additivity. Within this ‘piecewise approximation’, the individual corrections were evaluated separately, using more elaborate computational techniques. The main sources of computational imperfections were:

(i) electronic structure method

(ii) atomic basis set

(iii) explicit solvent effects

(iv) (ro)vibrational corrections

(v) model system size and quality

As already discussed in section 2.4.5, the (i) DFT error can be estimated by a comparison with benchmark electronic structure methods such as coupled clusters. Testing the quality of (ii) atomic basis set and (v) computational model by gradually increasing the size of either of them is also quite obvious. The corrections due the (iii) explicit solvent are usually quite demanding as they require averaging of the calculated results over many solute-solvent configurations generated e.g. with molecular dynamics. The (iv) (ro)vibrational corrections to the spectroscopic parameters were not calculated in this work but they represent a very important and not yet fully explored area [11,13,126,127].

Indeed, the piecewise approximation is not exactly accurate. For example, the basis set convergence properties are certainly differ for different electronic structure methods, wherefore the corrections due to the points (i) and (ii) are strictly speaking not separable. The explicit solvent effects and (ro)vibrational corrections (iii) and (iv) are also very closely related and in the case of the solute internal dynamics they are completely intertwined. Nevertheless, the piecewise approximation is often the only feasible way how to estimate the combined effect of different computational approximations and it helps to explain the differences between the calculated and experimental data.

2.6.1 Quantum chemistry programs

Most of the calculations in this doctoral work were done using the `Gaussian 03` and `Gaussian 09` programs. The dynamical simulations of work V were performed within the `Amber 10` program package and the benchmark coupled-clusters NMR calculations in work VI employed the `Cfour` program.

3 Theory of NMR spectroscopic method

3.1 NMR spectroscopy

Nuclear magnetic resonance spectroscopy is one of the most powerful experimental techniques nowadays. It is widely used in many research fields, ranging from solid-state material science through biomolecular research to many applications of non-invasive imaging. High-resolution Fourier-transform NMR represents a unique method for obtaining detailed information on structure and dynamics of biomolecules. In this section we will very briefly outline the theoretical basis of liquid-phase biomolecular NMR.

The phenomenon of nuclear magnetism applies to atomic nuclei with non-zero spin $\hat{\mathbf{I}}$. Placing of such nucleus into magnetic field $\mathbf{B}$ gives rise to a Zeeman interaction described with a Hamiltonian

$$\hat{H}^{\mathrm{Z}} = -\hat{\mathbf{m}} \cdot \mathbf{B} \ , \tag{3.1}$$

where we introduced the nuclear magnetic moment operator

$$\hat{\mathbf{m}} = \gamma \hbar \hat{\mathbf{I}} \ . \tag{3.2}$$

The gyromagnetic ratio γ in $\mathrm{rad} \cdot \mathrm{s}^{-1} \cdot \mathrm{T}^{-1}$ is a constant specific for each nuclear isotope and it characterizes magnitude of the nuclear magnetic moment.

Magnetic field acting on a nucleus in NMR experiment generally consists of a static magnetic field $\mathbf{B}_0$ and a time-dependent part $\mathbf{B}_1(t)$:

$$\mathbf{B}(t) = \mathbf{B}_0 + \mathbf{B}_1(t) \ . \tag{3.3}$$

Expectation value $\langle \hat{\mathbf{m}} \rangle$ of the nuclear magnetic moment in magnetic field $\mathbf{B}(t)$ undergoes a time evolution corresponding to the time-dependent Hamiltonian (3.1). By inserting $\hat{H}^{\mathrm{Z}}$ into Schrödinger equation (2.1) and considering properties of the spin operator $\hat{\mathbf{I}}$ we obtain equation of motion

$$\frac{\mathrm{d}\langle \hat{\mathbf{m}} \rangle}{\mathrm{d}t} = \langle \hat{\mathbf{m}} \rangle \times \gamma \mathbf{B}(t) \ . \tag{3.4}$$

Static magnetic field. A sample placed inside NMR spectrometer experiences a highly homogeneous and extremely stable magnetic field $\mathbf{B}_0$ (also called static or longitudinal). It is typically much stronger than the time-dependent magnetic field $\mathbf{B}_1(t)$. The interaction of a nuclear magnetic moment with $\mathbf{B}_0$ is described by time-independent Zeeman

Hamiltonian

$$\hat{H}^{\mathrm{Z}} = -\hat{\mathbf{m}} \cdot \mathbf{B}_0 \ . \tag{3.5}$$

Solutions of the corresponding stationary Schrödinger equation are Zeeman eigenstates equidistant in energy. The equation (3.4) then corresponds to the *Larmor precession* of the expectation value of nuclear magnetic moment $\langle \hat{\mathbf{m}} \rangle$ at Larmor angular frequency $\boldsymbol{\omega}_0^{\mathrm{L}} = -\gamma \mathbf{B}_0$.

Time-dependent magnetic field. Two kinds of the time-dependent magnetic field $\mathbf{B}_1(t)$ are important in NMR: A local, stochastically fluctuating field $\mathbf{B}_{\mathrm{loc}}(t)$ and an external radio-frequency field $\mathbf{B}_{\mathrm{rf}}(t)$ oscillating in transversal plane (perpendicular to $\mathbf{B}_0$). Both these fields can induce transitions between Zeeman energy levels of a nuclear spin through time-dependent perturbation $\hat{H}^{\mathrm{Z}}(t) = -\hat{\mathbf{m}} \cdot \mathbf{B}_1(t)$.

The local magnetic field $\mathbf{B}_{\mathrm{loc}}$ experienced by a nucleus is created by the surrounding nuclear spins and by the electron cloud in a molecule. A real sample consists of a huge amount of statistically identical molecular systems; each nuclear spin in a molecule thus corresponds to an ensemble of statistically identical nuclear spins in the sample. The stochastic time evolution of $\mathbf{B}_{\mathrm{loc}}$ arises due to the thermal molecular motion which causes random variation of the local magnetic interactions (see section 3.1.1). This inherent magnetic noise enables the ensemble of statistically identical nuclear spins reach thermodynamic equilibrium in which populations of Zeeman energy levels correspond to Boltzmann distribution. The difference in populations of individual Zeeman states is demonstrated in creation of macroscopic magnetization in the sample. The stochastic magnetic field $\mathbf{B}_{\mathrm{loc}}$ is thus directly related to the phenomenon of NMR relaxation (see section 3.1.4).

One of the fundamental concepts in NMR is the application of transversal magnetic field $\mathbf{B}_{\mathrm{rf}}$ oscillating at Larmor frequency of a particular nucleus. It can be shown that the whole ensemble of statistically identical nuclear spins is then rotated around the direction of $\mathbf{B}_{\mathrm{rf}}$ with angular frequency $|\frac{1}{2}\gamma B_{\mathrm{rf}}|$ (the factor of $\frac{1}{2}$ arises due to linear polarization of the $\mathbf{B}_{\mathrm{rf}}$). Macroscopic magnetization in a sample thus can be manipulated by radio-frequency pulses applied in a transversal plane.

3.1.1 Magnetic interactions in molecules

Local magnetic field $\mathbf{B}_{\mathrm{loc}}$ interacting with nuclear magnetic moment $\hat{\mathbf{m}}_K$ gives rise to energy contributions in the total Zeeman Hamiltonian (3.1). In this work we were concerned only with spin-$\frac{1}{2}$ nuclei in diamagnetic (i.e. closed-shell) molecular systems. The main magnetic interactions in such systems are the chemical shielding and the direct and indirect spin–spin coupling.

Chemical shielding. According to the semi-classical view developed by Ramsey [128], if a molecule is placed in the external magnetic field $\mathbf{B}_0$, it induces currents in the electron cloud of the molecule, which in turn create a small secondary magnetic flux that adds to $\mathbf{B}_0$. The interaction of a nuclear magnetic moment $\hat{\mathbf{m}}_K$ with the induced magnetic flux is known as the chemical shielding. It can be described with a Hamiltonian

$$\hat{H}_K^{\mathrm{CS}} = \hat{\mathbf{m}}_K \cdot \overleftrightarrow{\boldsymbol{\sigma}}_K \cdot \mathbf{B}_0 \ , \tag{3.6}$$

where $\overleftrightarrow{\boldsymbol{\sigma}}_K$ is the dimensionless chemical shielding tensor, usually reported in ppm (parts per million, 10^{-6}). The orientation of $\overleftrightarrow{\boldsymbol{\sigma}}_K$ is fixed in the molecule and it is strongly dependent on the electronic environment of a nucleus K. In the most general case, all nine components σ_{uv} $(u, v = \{x, y, z\})$ are independent. However, in NMR spectra of isotropic liquids only the isotropic part $\sigma_K^{\mathrm{iso}} = \frac{1}{3}\mathrm{Tr}\,\overleftrightarrow{\boldsymbol{\sigma}}_K$ is retained. The anisotropic part of the chemical shielding tensor contributes to the NMR relaxation (see section 3.1.4).

Direct spin–spin coupling. Nuclear magnetic moments are subjected to the direct (through-space) dipole–dipole interaction which can be expressed in the explicit form

$$\begin{aligned} \hat{H}_{KL}^{\mathrm{DD}} &= -\frac{\mu_0}{4\pi}\left(\frac{3\hat{\mathbf{m}}_K \cdot \mathbf{R}_{KL}\mathbf{R}_{KL} \cdot \hat{\mathbf{m}}_L}{R_{KL}^5} - \frac{\hat{\mathbf{m}}_K \cdot \hat{\mathbf{m}}_L}{R_{KL}^3}\right) \\ &= \hat{\mathbf{m}}_K \cdot \overleftrightarrow{\mathbf{D}}_{KL} \cdot \hat{\mathbf{m}}_L \ , \end{aligned} \tag{3.7}$$

where $\mathbf{R}_{KL} = \mathbf{R}_K - \mathbf{R}_L$ is the position of nucleus K relative to nucleus L and $\overleftrightarrow{\mathbf{D}}_{KL}$ is the direct spin–spin coupling tensor. In isotropic liquids, the direct dipole–dipole interaction gets completely averaged out due to the molecular motion and it thus does not contribute to liquid-phase NMR spectra. It does, however, significantly contribute to NMR relaxation (see section 3.1.4).

Indirect spin–spin coupling. Finally, we must also consider an electron-mediated nuclear spin–spin interaction, usually called J-coupling. It is manifested by magnetic-field-independent splitting of NMR spectral lines that is specific to the electronic environment between the coupled nuclei K and L. The most general form of this interaction is again tensorial,

$$\hat{H}_{KL}^{\mathrm{J}} = h\hat{\mathbf{I}}_K \cdot \overleftrightarrow{\mathbf{J}}_{KL} \cdot \hat{\mathbf{I}}_L \ . \tag{3.8}$$

However, only the isotropic (scalar) J-coupling is retained in the NMR spectra of isotropic liquids,

$$\begin{aligned} \hat{H}_{KL}^{\mathrm{J,iso}} &= hJ_{KL}\hat{\mathbf{I}}_K \cdot \hat{\mathbf{I}}_L && \text{(homonuclear case)} \\ \hat{H}_{KL}^{\mathrm{J,iso}} &= hJ_{KL}\hat{I}_{K,z}\hat{I}_{L,z} && \text{(heteronuclear case)} \end{aligned} \tag{3.9}$$

where $J_{KL} = \frac{1}{3}\mathrm{Tr}\,\overleftrightarrow{\mathbf{J}}_{KL}$ is the scalar spin–spin coupling constant (J-coupling) denoting the magnitude of the spectral splitting in Hz.

3.1.2 Effective spin Hamiltonian

The individual magnetic interactions $\hat{H}_K^{\mathrm{Z}}$, $\hat{H}_K^{\mathrm{CS}}$, $\hat{H}_{KL}^{\mathrm{DD}}$ and $\hat{H}_{KL}^{\mathrm{J}}$ can be collected into the effective spin Hamiltonian

$$\hat{H}^{\mathrm{eff}} = -\sum_K \hat{\mathbf{m}}_K \cdot \left(\overleftrightarrow{\mathbf{1}} - \overleftrightarrow{\boldsymbol{\sigma}}_K\right) \cdot \mathbf{B}_0 + \frac{1}{2}\sum_{K \neq L} \hat{\mathbf{m}}_K \cdot \left(\overleftrightarrow{\mathbf{D}}_{KL} + \overleftrightarrow{\mathbf{K}}_{KL}\right) \cdot \hat{\mathbf{m}}_L \tag{3.10}$$

which empirically describes a general NMR spectrum of a diamagnetic molecular system with spin-$\frac{1}{2}$ nuclei. The summations are running through all nuclear spins and spin pairs in the system, respectively. In (3.10) we introduced a nucleus-independent indirect spin–spin coupling tensor $\overleftrightarrow{\mathbf{K}}_{KL}$ related to the J-coupling tensor according to

$$\overleftrightarrow{\mathbf{J}}_{KL} = h\frac{\gamma_K}{2\pi}\frac{\gamma_L}{2\pi}\overleftrightarrow{\mathbf{K}}_{KL} \ . \tag{3.11}$$

Assuming that the Born–Oppenheimer approximation holds also in the presence of magnetic field, we can distinguish between the through-space magnetic interactions ($\hat{H}_K^{\mathrm{Z}}$ and $\hat{H}_{KL}^{\mathrm{DD}}$) that contribute only to the nuclear energy and the electron-mediated magnetic interactions ($\hat{H}_K^{\mathrm{CS}}$ and $\hat{H}_{KL}^{\mathrm{J}}$) that perturb the electronic energy and can be thus evaluated using quantum chemistry methods. In the following, the nuclear magnetic moment operator $\hat{\mathbf{m}}$ can be replaced with a vector $\mathbf{m}$.

3.1.3 NMR parameters

In this work, when referring to NMR parameters, we mean specifically the tensors $\overleftrightarrow{\boldsymbol{\sigma}}_K$ and $\overleftrightarrow{\mathbf{J}}_{KL}$ and related quantities which are manifested either in NMR spectra or in NMR relaxations (see below). The NMR tensors are commonly presented as 3×3 Cartesian matrices while they are still referred to as tensors. In the following we will stick to this convention, omitting the formal difference between tensor and its Cartesian matrix representation.

NMR parameters as energy derivatives. The chemical shielding tensor $\overset{\leftrightarrow}{\boldsymbol{\sigma}}_K$ is a second-order molecular property (as defined in section 2.5) corresponding to the ground-state electronic energy E_0 perturbed by static magnetic field $\mathbf{B}_0$ and nuclear magnetic moment $\mathbf{m}_K$. The chemical shielding tensor components $\sigma_{K,uv}$, $u, v = \{x, y, z\}$, can be evaluated as

$$\sigma_{K,uv} = \left. \frac{\mathrm{d}^2 E_0}{\mathrm{d}m_{K,u} \mathrm{d}B_{0,v}} \right|_{\mathbf{m}_K=\mathbf{B}_0=0} . \tag{3.12}$$

Similarly, the J-coupling is also a second-order property. The Cartesian components of the nucleus-independent spin–spin coupling tensor $\overset{\leftrightarrow}{\mathbf{K}}_{KL}$ are given by

$$K_{KL,uv} = \left. \frac{\mathrm{d}^2 E_0}{\mathrm{d}m_{K,u} \mathrm{d}m_{L,v}} \right|_{\mathbf{m}_K=\mathbf{m}_L=0} . \tag{3.13}$$

Chemical shielding tensor invariants. The chemical shielding tensor $\overset{\leftrightarrow}{\boldsymbol{\sigma}}$ is in the Cartesian basis represented by a 3×3 matrix with nine independent components. It can be decomposed into irreducible tensors of rank 0, 1 and 2,

$$\overset{\leftrightarrow}{\boldsymbol{\sigma}} = \overset{\leftrightarrow}{\boldsymbol{\sigma}}^{(0)} + \overset{\leftrightarrow}{\boldsymbol{\sigma}}^{(1)} + \overset{\leftrightarrow}{\boldsymbol{\sigma}}^{(2)} . \tag{3.14}$$

The rank-0 irreducible tensor is the isotropic chemical shielding, $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{(0)} = \frac{1}{3} \mathrm{Tr}\, \overset{\leftrightarrow}{\boldsymbol{\sigma}} = \sigma^{\mathrm{iso}}$. The antisymmetric rank-1 tensor $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{(1)}$ with three independent components is manifested only in higher-order NMR effects and it will not be discussed here. The remaining five components of $\overset{\leftrightarrow}{\boldsymbol{\sigma}}$ correspond to the rank-2 irreducible tensor $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{(2)}$. The rank-0 and rank-2 tensors together constitute a symmetric tensor $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{\mathrm{symm}}$, which can be diagonalized, finding thus principal values σ_{ii} and the corresponding eigenvectors $\mathbf{e}_{ii}$, $i = \{1, 2, 3\}$. As the diagonalization does not alter the trace of the tensor, the isotropic chemical shielding is an average of the three principal values,

$$\sigma^{\mathrm{iso}} = \frac{1}{3} (\sigma_{11} + \sigma_{22} + \sigma_{33}) . \tag{3.15}$$

In the chemical shielding calculations, a convention $\sigma_{11} < \sigma_{22} < \sigma_{33}$ is usually adopted. An important characteristic of the $\overset{\leftrightarrow}{\boldsymbol{\sigma}}$ tensor is the chemical shielding anisotropy (CSA, $\Delta\sigma$). It can be expressed in several ways; we use the definition

$$\Delta\sigma = \sigma_{33} - \frac{\sigma_{11} + \sigma_{22}}{2} . \tag{3.16}$$

Chemical shift. Absolute chemical shielding σ is almost never reported in the literature since the chemical shielding scale originates at the Larmor frequency of a bare atomic nucleus. Instead, an expectedly inert molecule of so-called NMR standard (NMR reference) is added to the NMR sample, and the relative chemical shift δ defined as

$$\begin{aligned} \delta &= \frac{\omega - \omega_{\mathrm{std}}}{\omega_{\mathrm{std}}} = \frac{\sigma_{\mathrm{std}} - \sigma}{1 - \sigma_{\mathrm{std}}} \\ &\approx \sigma_{\mathrm{std}} - \sigma \end{aligned} \tag{3.17}$$

is reported. ω and ω_{std} are the nuclear resonance frequencies in the sample and in the NMR standard, respectively. The definition (3.17) also provides an easy way for obtaining the chemical shift computationally—by subtracting the chemical shielding σ calculated in the molecule of NMR sample from that of the NMR standard σ_{std}. Further, if both σ and σ_{std} are calculated with the same method, some computational errors may get canceled in the subtraction.

The σ_{std} might be difficult to calculate, as is the case of a widely used ^{31}P NMR standard, the 85% aqueous solution of phosphoric acid, H_3PO_4 (aq.). Problematic modeling of $\sigma_{H_3PO_4}$ can be circumvented by introducing secondary standard PH_3, as suggested by van Wüllen [129]. Experimental chemical shift of gas-phase PH_3 relative to H_3PO_4 (aq.) is known to be −266.1 ppm [130]. ^{31}P chemical shift δ_{P} thus can be calculated using the secondary standard approach as

$$\delta_{\mathrm{P}} = -266.1 + \sigma_{\mathrm{PH_3}} - \sigma_{\mathrm{P},} \tag{3.18}$$

where $\sigma_{\mathrm{PH_3}}$ is the ^{31}P chemical shielding calculated for the phosphine molecule in vacuum. Note that the accuracy of δ_{P} calculated this way is still limited by the precision of the computational method used.

3.1.4 NMR relaxation

The phenomenon of NMR relaxation has been utilized as a valuable source of information on molecular systems since the early days of NMR. The theory describing the evolution of nuclear magnetization in static magnetic field was first formulated by Bloch [131]. The origin of both longitudinal and transversal NMR relaxation has been ascribed to stochastic fluctuations of local magnetic fields acting on nuclear spins.

In molecular systems with spin-$\frac{1}{2}$ nuclei, the main sources of NMR relaxation are the direct dipole–dipole interaction (DD) and the chemical shielding anisotropy (CSA). Theory of NMR relaxation due to the dipole–dipole interaction was first given for a two-spin system by Solomon [132]. This simple case already included the theoretical basis for the explanation of the nuclear Overhauser effect.

A general description of relaxation processes in molecular systems was provided by the Bloch–Wangsness–Redfield theory [133, 134], usually called the Redfield theory. Essentially, it is a second-order time-dependent perturbation theory that approximately solves the Liouville–von Neumann equation for the density matrix $\hat{\rho}(t)$ of a nuclear spin ensemble in the presence of a stochastic magnetic perturbation. The magnetic interactions are expressed in terms of irreducible spherical tensors which are transformed from their principal axis frame to the laboratory frame using reduced Wigner coefficients. The time evolution of the expectation value $\langle\hat{M}\rangle$ of nuclear spin operator $\hat{M}$ is within the so-called secular approximation given by [135]

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{M}\rangle = -\frac{1}{2}\sum_{\lambda\lambda'} c_{\lambda}c_{\lambda'}\sum_{r} J^{\lambda\lambda'}(\omega_r)\cdot \mathrm{Tr}\left\{\left[[\hat{M},\hat{Q}_r^{\lambda}],\hat{Q}_r^{\lambda'\,\dagger}\right]\hat{\rho}(t)\right\} \ , \tag{3.19}$$

where the first sum runs over all pairs of magnetic interactions λ and λ' (NMR relaxation mechanisms) in the nuclear spin ensemble, the constants c_{λ} are $c_{\lambda'}$ are the magnetic interaction strengths, $J^{\lambda\lambda'}(\omega_r)$ is the spectral density function involving the stochastic parts of magnetics perturbations correlated at characteristic frequencies ω_r of the nuclear spin ensemble, and $\hat{Q}_r^{\lambda}$ are spin operators specific for the magnetic interaction λ.

The spectral density function $J^{\lambda\lambda'}(\omega)$ is a Fourier transform of the correlation function between the magnetic interactions λ and λ' represented by irreducible spherical tensors. The explicit form of the $J^{\lambda\lambda'}(\omega)$ can be very complex, depending on the employed model of molecular motion.

By particular choice of the nuclear spin operator $\hat{M}$ in equation (3.19) we find equations of motion for the spin populations (longitudinal NMR relaxation), single-quantum spin coherences (transversal NMR relaxation) and multiple-quantum coherences. The terms of (3.19) with $\lambda = \lambda'$ correspond to the so-called auto-correlated relaxation while the remaining terms with $\lambda \neq \lambda'$ describe the interference of two different relaxation mechanisms, the cross-correlated relaxation (CCR).

Cross-correlated relaxation. Three types of CCR can occur by interference of different DD or CSA relaxation mechanisms: the DD–DD, the CSA–DD and the CSA–CSA cross-correlated relaxation [136]. The DD–DD and CSA–DD cross-correlations are manifested in both longitudinal and transversal NMR relaxation; the CSA–CSA mechanism contributes only to the relaxation of multiple-quantum coherences. The DD–DD is usually the strongest CCR mechanism and has been successfully applied in the determination of biomolecular structure [1]. We focused on the CSA–DD mechanism that turned out important in structural studies of peptides [2, 3] and nucleic acids [26].

The cross-correlated relaxation rate $\Gamma^{\text{CSA-DD}}_{I,JK}$ involving CSA of a nucleus I and the DD interaction between nuclei J and K is given by [2, 135]

$$\Gamma^{\text{CSA-DD}}_{I,JK} = \frac{1}{2}\left(\frac{\mu_0}{4\pi}\right)\hbar\frac{\gamma_J\gamma_K}{r^3_{JK}}\gamma_I B_0\sum_r a_r J^{\text{CSA-DD}}(\omega_r) \ , \tag{3.20}$$

where $\gamma_I, \gamma_J, \gamma_K$ are the gyromagnetic ratios, r_{JK} is the length of the dipolar vector, B_0 is the magnetic field strength, a_r are the coefficients resulting from evaluating the double commutator in the equation (3.19) and $J^{\text{CSA-DD}}(\omega_r)$ is the CSA–DD spectral density function evaluated at resonance frequencies ω_r. Assuming that the mutual orientation of the CSA and DD principal axis frames is rigid, the CSA–DD spectral density function can be expressed in the form

$$J^{\text{CSA-DD}}(\omega) = \left[\sum_{i=1}^{3} \sigma_{I,ii} \cdot \mathcal{D}_{00}^{(2)}(\vartheta_{JK,ii})\right] J(\omega) = \Delta\sigma_{I,JK}^{\text{eff}}\ J(\omega)\ , \tag{3.21}$$

where $\sigma_{I,ii}$ are the principal values of the chemical shielding tensor of nucleus I, $\mathcal{D}_{00}^{(2)}(\vartheta_{JK,ii}) = (3\cos^2\vartheta_{JK,ii} - 1)/2$ is the reduced Wigner coefficient corresponding to the transition between CSA and DD principal axis frames, $\vartheta_{JK,ii}$ is the angle between the dipolar vector $\mathbf{r}_{JK}$ and the i-th principal axis of the chemical shielding tensor, and $J(\omega)$ is the ordinary spectral density function; in the special case of isotropic molecular tumbling characterized by the rotation correlation time τ_c, $J(\omega) = \frac{2}{5}\frac{\tau_c}{1+(\omega\tau_c)^2}$. The effective CSA $\Delta\sigma_{I,JK}^{\text{eff}}$ defined in (3.21) and the dipolar distance factor r_{JK}^{-3} in (3.20) constitute the molecular-geometry-dependent parts of CSA–DD cross-correlated relaxation rates that can be conveniently evaluated using quantum chemistry methods.

3.1.5 Structural interpretation of NMR parameters

The ability of NMR spectroscopy to sensitively probe the local environment of atomic nuclei is one of the main reasons for the vast success of NMR. The relationships between NMR parameters and molecular structure are essential tools for the application of NMR in biomolecular research. Empirical rules can be obtained experimentally for molecules with known geometry. However, a much deeper and more general insight into the structural dependence of NMR parameters is obtained from theoretical calculations. A prominent example of the theoretical approach was finding a general form of the dependence of three-bond vicinal J-coupling on the associated torsion angle by Karplus [137]. The famous Karplus relation can be expressed in the form

$${}^{3}J_{KL}(\alpha) = C + A_1\cos(\alpha - \phi) + A_2\cos^2(\alpha - \phi)\ , \tag{3.22}$$

where the constant C, the amplitudes A_1, A_2 and the phase ϕ are parameters specific for each pair of coupled nuclei K, L, and α is the torsion angle assigned to the coupled atoms. Recent advances in quantum chemistry computational methods enable finding more complex relations between NMR parameters and molecular structure [22, 32, 138]. Theoretical modeling of NMR parameters in dependence on molecular geometry was one of the central topics of this doctoral work.

3.2 Theory of NMR parameters

The theory of NMR parameters can be formulated following an approach outlined by Fukui [114, 115], starting from the Dirac equation [139, 140] for an electron in magnetic field,

$$\left[c\begin{pmatrix}0 & \boldsymbol{\sigma}\\ \boldsymbol{\sigma} & 0\end{pmatrix}\cdot\hat{\boldsymbol{\pi}}+\begin{pmatrix}\mathbb{1} & 0\\ 0 & -\mathbb{1}\end{pmatrix}m_{\mathrm{e}}c^2\right]\begin{pmatrix}\psi_{\mathrm{L}}\\ \psi_{\mathrm{S}}\end{pmatrix}=E_{\mathrm{rel}}\begin{pmatrix}\psi_{\mathrm{L}}\\ \psi_{\mathrm{S}}\end{pmatrix}\ , \tag{3.23}$$

where ψ_{L} and ψ_{S} are the large and small components of the total four-component wave function, E_{rel} is the relativistic electron energy, c is the speed of light in vacuum, m_{e} is the electron mass, $\mathbb{1}$ is a 2×2 unit matrix, $\boldsymbol{\sigma}$ is the vector of Pauli matrices $\sigma_x, \sigma_y, \sigma_z$, and $\hat{\boldsymbol{\pi}} = \hat{\mathbf{p}}+e\mathbf{A}$ is the mechanical momentum of electron constituting of canonical momentum $\hat{\mathbf{p}} = -i\hbar\nabla$ and a contribution from vector potential $\mathbf{A}$. The particular form of $\mathbf{A}$ is specific for each NMR parameter and can be inserted later.

Since there is no potential $\hat{V}$ in the Dirac Hamiltonian of equation (3.23), the components ψ_{L} and ψ_{S} are not coupled and the Dirac equation can be rewritten for the large component only,

$$\left[\frac{c^2(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{\pi}})^2}{E_{\mathrm{rel}}+m_{\mathrm{e}}c^2}+m_{\mathrm{e}}c^2\right]\psi_{\mathrm{L}} = E_{\mathrm{rel}}\psi_{\mathrm{L}}\ . \tag{3.24}$$

The one-electron Hamiltonian corresponding to E_{rel} is thus (discarding the negative-energy solution)

$$\begin{aligned}\hat{H}_{\mathrm{rel}} &= \sqrt{m_{\mathrm{e}}^2c^4+c^2(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{\pi}})^2}\\ &\approx m_{\mathrm{e}}c^2+\frac{(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{\pi}})^2}{2m_{\mathrm{e}}}+O(c^{-2}) = E_0+\hat{H}_1+O(c^{-2})\ ,\end{aligned} \tag{3.25}$$

where E_0 is the rest mass energy, the Hamiltonian $\hat{H}_1$ includes all non-relativistic terms, and the relativistic corrections are of the order of c^{-2}.

By choosing the Coulomb gauge $\nabla\cdot\mathbf{A} = 0$, the Hamiltonian $\hat{H}_1$ can be shown equivalent to the non-relativistic Pauli Hamiltonian $\hat{H}_{\mathrm{Pauli}}$ for an electron in magnetic field:

$$\hat{H}_{\mathrm{Pauli}} = -\frac{\hbar^2}{2m_{\mathrm{e}}}\Delta-\mu_{\mathrm{B}}g_{\mathrm{s}}\hat{\mathbf{s}}\cdot\nabla\times\mathbf{A}+\frac{e}{m_{\mathrm{e}}}\mathbf{A}\cdot\hat{\mathbf{p}}+\frac{e^2}{2m_{\mathrm{e}}}A^2\ , \tag{3.26}$$

where $\mu_{\mathrm{B}} = e\hbar/(2m_{\mathrm{e}})$ is Bohr magneton, $g_{\mathrm{s}} = 2.002319...$ is the electron g-factor, and $\hat{\mathbf{s}} = -\boldsymbol{\sigma}/g_{\mathrm{s}}$ is the electron spin operator. By introducing the semi-classical operator $\hat{\mathbf{s}}$ we definitively moved from the two-component non-relativistic limit of the Dirac Hamiltonian to the one-component case accessible to standard quantum chemistry calculations.

3.2.1 *J*-coupling Hamiltonian

Rigorous theoretical description of the J-coupling interaction was first given by Ramsey [141]. The J-coupling Hamiltonian can be found by inserting the vector potential of two magnetic dipoles (nuclear magnetic moments) $\mathbf{m}_K$ and $\mathbf{m}_L$

$$\mathbf{A} = \mathbf{A}_K + \mathbf{A}_L = \frac{\mu_0}{4\pi}\left(\frac{\mathbf{m}_K \times \mathbf{r}_K}{r_K^3} + \frac{\mathbf{m}_L \times \mathbf{r}_L}{r_L^3}\right) \tag{3.27}$$

into the one-electron Pauli Hamiltonian (3.26). μ_0 is the vacuum permeability and the vectors $\mathbf{r}_K = \mathbf{r} - \mathbf{R}_K$ and $\mathbf{r}_L = \mathbf{r} - \mathbf{R}_L$ describe the electron position relative to the two magnetic moments. The magnetic field of each magnetic dipole is given by

$$\nabla \times \mathbf{A}_K = \frac{\mu_0}{4\pi}\,\frac{3\mathbf{m}_K \cdot \mathbf{r}_K\mathbf{r}_K - \mathbf{m}_K r_K^2}{r_K^5} + \frac{2}{3}\mu_0\mathbf{m}_K\delta^3(\mathbf{r}_K)\ , \tag{3.28}$$

where the second term involving three-dimensional Dirac δ-function arises due to the differentiation in a distribution sense [142] and it is directly related to the Fermi-contact term of the J-coupling Hamiltonian. Griffiths gave a clear and intuitive explanation for the need of δ-function in the expression for magnetic field of point magnetic dipole in order to preserve the consistency of the theory for $r_K \to 0$ [143].

Using (3.27) and (3.28), the Pauli Hamiltonian (3.26) can be rewritten as

$$\hat{H}^{\mathrm{J}}_{\mathrm{Pauli}} = -\frac{\hbar^2}{2m_{\mathrm{e}}}\Delta + \frac{e^2}{2m_{\mathrm{e}}}(A_K^2 + A_L^2) + \hat{H}^{\mathrm{J}}_{KL}\ , \tag{3.29a}$$

$$\hat{H}^{\mathrm{J}}_{KL} = \hat{H}^{\mathrm{DSO}}_{KL} + \hat{H}^{\mathrm{PSO}}_{K} + \hat{H}^{\mathrm{PSO}}_{L} + \hat{H}^{\mathrm{FC}}_{K} + \hat{H}^{\mathrm{FC}}_{L} + \hat{H}^{\mathrm{SD}}_{K} + \hat{H}^{\mathrm{SD}}_{L}\ , \tag{3.29b}$$

where neither the electron kinetic energy in the first term of (3.29a) nor the terms quadratic in the vector potentials of magnetic dipoles contribute to the non-relativistic J-coupling. The full non-relativistic one-electron Hamiltonian $\hat{H}^{\mathrm{J}}_{KL}$ (3.29b) of J-coupling between nuclei K and L constitutes of the diamagnetic spin-orbit (DSO), paramagnetic spin-orbit (PSO), Fermi-contact (FC), and spin-dipolar (SD) operators:

$$\hat{H}^{\mathrm{DSO}}_{KL} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{e^2}{m_{\mathrm{e}}}\,\frac{\mathbf{m}_K \cdot \mathbf{m}_L\mathbf{r}_K \cdot \mathbf{r}_L - \mathbf{m}_K \cdot \mathbf{r}_L\mathbf{r}_K \cdot \mathbf{m}_L}{r_K^3 r_L^3} \tag{3.30a}$$

$$\hat{H}^{\mathrm{PSO}}_{X} = 2\left(\frac{\mu_0}{4\pi}\right)\frac{\mu_{\mathrm{B}}}{\hbar}\,\frac{\mathbf{m}_X \cdot \hat{\mathbf{l}}_X}{r_X^3} \tag{3.30b}$$

$$\hat{H}^{\mathrm{FC}}_{X} = -\frac{8\pi}{3}\left(\frac{\mu_0}{4\pi}\right)\mu_{\mathrm{B}}g_{\mathrm{s}}\mathbf{m}_X \cdot \hat{\mathbf{s}}\delta^3(\mathbf{r}_X) \tag{3.30c}$$

$$\hat{H}^{\mathrm{SD}}_{X} = -\left(\frac{\mu_0}{4\pi}\right)\mu_{\mathrm{B}}g_{\mathrm{s}}\frac{3\mathbf{m}_X \cdot \mathbf{r}_X\mathbf{r}_X - \mathbf{m}_X r_X^2}{r_X^5} \cdot \hat{\mathbf{s}}\ , \tag{3.30d}$$

where $X = \{K, L\}$ and $\hat{\mathbf{l}}_X = -i\hbar \mathbf{r}_X \times \nabla$ is the orbital angular momentum operator.

The DSO operator describes the interaction between the two magnetic moments mediated by the electron orbital motion. The PSO term corresponds to the interaction of nuclear magnetic moment with the electron orbital momentum. Both FC and SD operators describe the interaction between the nuclear magnetic moments and the electron spin, each corresponding to one part of the magnetic dipole field (3.28). Namely, the FC term arises due to the non-zero electron density at the position of nucleus. The decomposition of the electron spin-dependent part of $\hat{H}^{\mathrm{J}}_{KL}$ into the FC and SD terms is possible only in the non-relativistic case [142, 144] and it is thus somewhat artificial but convenient since the FC term of J-coupling is often dominant and relatively easy to evaluate.

3.2.2 Perturbation theory of J-coupling

According to (3.13), the nucleus-independent spin–spin coupling tensor $\overset{\leftrightarrow}{\mathbf{K}}_{KL}$ is a second-order molecular property corresponding to the electronic system perturbed by nuclear magnetic moments $\mathbf{m}_K$ and $\mathbf{m}_L$. Assuming a closed-shell molecule with non-degenerate ground state, the tensor components $K_{KL,uv}$, $u, v = \{x, y, z\}$, can be expressed using the sum-over-states expansion (2.54) as

$$K_{KL,uv} = \langle 0|\hat{h}^{\mathrm{DSO}}_{KL,uv}|0\rangle + 2\sum_{s>0} \frac{\langle 0|\hat{h}^{\mathrm{PSO}}_{K,u}|s\rangle\langle s|\hat{h}^{\mathrm{PSO}}_{L,v}|0\rangle}{E_0 - E_s} + \\ + 2\sum_{t} \frac{\langle 0|\hat{h}^{\mathrm{FC}}_{K,u} + \hat{h}^{\mathrm{SD}}_{K,u}|t\rangle\langle t|\hat{h}^{\mathrm{FC}}_{L,v} + \hat{h}^{\mathrm{SD}}_{L,v}|0\rangle}{E_0 - E_t} \,, \tag{3.31}$$

where $|0\rangle$ is the ground state, $|s\rangle$ and $|t\rangle$ are the singlet and triplet virtual states, E_0, E_s, and E_t are the corresponding electronic energies, and each of the perturbation operators $\hat{h}^{\mathrm{DSO}}_{KL,uv}$, $\hat{\mathbf{h}}^{\mathrm{PSO}}_X$, $\hat{\mathbf{h}}^{\mathrm{FC}}_X$, $\hat{\mathbf{h}}^{\mathrm{SD}}_X$, $X = K, L$, in the many-electron case is a sum of one-electron contributions obtained by differentiating the respective terms of the one-electron J-coupling Hamiltonian $\hat{H}^{\mathrm{J}}_{KL}$ (3.30):

$$\hat{h}^{\mathrm{DSO}}_{KL,uv} = \left(\frac{\mu_0}{4\pi}\right)^2 \frac{e^2}{m_\mathrm{e}} \sum_{a=1}^{N} \frac{\mathbf{r}_{aK} \cdot \mathbf{r}_{aL}\delta_{uv} - r_{aK,v} r_{aL,u}}{r^3_{aK} r^3_{aL}} \tag{3.32a}$$

$$\hat{\mathbf{h}}^{\mathrm{PSO}}_X = 2\left(\frac{\mu_0}{4\pi}\right) \frac{\mu_\mathrm{B}}{\hbar} \sum_{a=1}^{N} \frac{\hat{\mathbf{l}}_{aX}}{r^3_{aX}} \tag{3.32b}$$

$$\hat{\mathbf{h}}^{\mathrm{FC}}_X = -\frac{8\pi}{3}\left(\frac{\mu_0}{4\pi}\right) \mu_\mathrm{B} g_\mathrm{s} \hat{\mathbf{s}} \sum_{a=1}^{N} \delta^3(\mathbf{r}_{aX}) \tag{3.32c}$$

$$\hat{\mathbf{h}}^{\mathrm{SD}}_X = -\left(\frac{\mu_0}{4\pi}\right) \mu_\mathrm{B} g_\mathrm{s} \sum_{a=1}^{N} \frac{3\hat{\mathbf{s}} \cdot \mathbf{r}_{aX}\mathbf{r}_{aX} - \hat{\mathbf{s}} r^2_{aX}}{r^5_{aX}} \,. \tag{3.32d}$$

The PSO operator does not act on the spin part of wave-function and, therefore, it contributes only to the matrix elements between singlet ground state and singlet virtual states. The FC and SD operators on the other hand contain the electron spin operator and the corresponding integrals thus involve both singlet and triplet virtual states.

Therefore, in total five terms contribute to the tensor $\overset{\leftrightarrow}{\mathbf{K}}_{KL}$. Namely, the DSO term represents the diamagnetic part of the J-coupling and the PSO–PSO, FC–FC, SD–SD and FC–SD terms constitute the paramagnetic part. Moreover, the mixed FC–SD contribution is fully anisotropic [115]. The isotropic J-coupling in the non-relativistic limit thus consists of the four well-known terms J_{KL}^{DSO}, J_{KL}^{PSO}, J_{KL}^{FC} and J_{KL}^{SD}.

3.2.3 Chemical shielding Hamiltonian

The theory of chemical shielding in molecules was developed by Ramsey [128, 145]. The chemical shielding Hamiltonian can be obtained by inserting the vector potential of a magnetic dipole $\mathbf{m}_K$ and a static magnetic field $\mathbf{B}_0$

$$\mathbf{A} = \mathbf{A}_K + \mathbf{A}_0 = \frac{\mu_0}{4\pi}\frac{\mathbf{m}_K \times \mathbf{r}_K}{r_K^3} + \frac{1}{2}\mathbf{B}_0 \times \mathbf{r}_{\mathrm{O}} \tag{3.33}$$

into the one-electron Pauli Hamiltonian (3.26). Here $\mathbf{r}_{\mathrm{O}} = \mathbf{r} - \mathbf{R}_{\mathrm{O}}$ is the electron position relative to the origin of vector potential $\mathbf{R}_{\mathrm{O}}$ (gauge origin), which can be in theory placed anywhere in space since $\nabla \times \mathbf{A}_0 = \mathbf{B}_0$ for any choice of $\mathbf{R}_{\mathrm{O}}$. Similar to the J-coupling theory, the magnetic dipole field $\mathbf{B}_K$ corresponding to the vector potential $\mathbf{A}_K$ is given by (3.28). The Pauli Hamiltonian for an electron in the presence of magnetic perturbations $\mathbf{m}_K$ and $\mathbf{B}_0$ thus adopts the form

$$\hat{H}_{\mathrm{Pauli}}^{\mathrm{CS}} = -\frac{\hbar^2}{2m_{\mathrm{e}}}\Delta + \frac{e^2}{2m_{\mathrm{e}}}(A_K^2 + A_{\mathrm{O}}^2) - \frac{\mu_{\mathrm{B}}}{\hbar}g_{\mathrm{s}}\hat{\mathbf{s}} \cdot \mathbf{B}_0 + \hat{H}_K^{\mathrm{FC}} + \hat{H}_K^{\mathrm{SD}} + \hat{H}_K^{\mathrm{CS}} , \tag{3.34a}$$

$$\hat{H}_K^{\mathrm{CS}} = \hat{H}_K^{\mathrm{dia}} + \hat{H}^{\mathrm{orb}} + \hat{H}_K^{\mathrm{PSO}} , \tag{3.34b}$$

where neither the kinetic energy in the first term nor the terms quadratic in vector potential contribute to the chemical shielding. Moreover, the terms involving the electron spin operator (triplet operators FC and SD) do not play part in the non-relativistic chemical shielding, as can be seen from the perturbational expansion (3.36); these terms are involved in relativistic corrections to the chemical shielding. The non-relativistic chemical shielding Hamiltonian thus consists of the diamagnetic ("dia"), orbital Zeeman ("orb") and PSO (3.30b) terms:

$$\hat{H}_K^{\mathrm{dia}} = \left(\frac{\mu_0}{4\pi}\right)\frac{e^2}{2m_{\mathrm{e}}}\frac{\mathbf{B}_0 \cdot \mathbf{m}_K \mathbf{r}_{\mathrm{O}} \cdot \mathbf{r}_K - \mathbf{B}_0 \cdot \mathbf{r}_K \mathbf{r}_{\mathrm{O}} \cdot \mathbf{m}_K}{r_K^3} \tag{3.35a}$$

$$\hat{H}^{\mathrm{orb}} = \frac{\mu_{\mathrm{B}}}{\hbar}\mathbf{B}_0 \cdot \hat{\mathbf{l}}_{\mathrm{O}} , \tag{3.35b}$$

where $\hat{\mathbf{l}}_\mathrm{O} = -i\hbar\mathbf{r}_\mathrm{O} \times \nabla$ is the orbital angular momentum relative to the gauge origin.

3.2.4 Perturbation theory of chemical shielding

The chemical shielding tensor is a second-order property defined by (3.12). Its Cartesian components $\sigma_{K,uv}$, $u, v = \{x, y, z\}$, for a closed-shell molecule with non-degenerate ground state can be expressed using the perturbational expansion (2.54) as

$$\sigma_{K,uv} = \langle 0|\hat{h}^{\mathrm{dia}}_{K,uv}|0\rangle + 2\sum_{s>0} \frac{\langle 0|\hat{h}^{\mathrm{orb}}_{u}|s\rangle\langle s|\hat{h}^{\mathrm{PSO}}_{K,v}|0\rangle}{E_0 - E_s} \ , \tag{3.36}$$

where the summation runs only over singlet virtual states $|s\rangle$. Therefore, the chemical shielding tensor can be symbolically expressed as a sum of the diamagnetic and paramagnetic terms, $\overleftrightarrow{\boldsymbol{\sigma}} = \overleftrightarrow{\boldsymbol{\sigma}}^{\mathrm{dia}} + \overleftrightarrow{\boldsymbol{\sigma}}^{\mathrm{para}}$. The perturbation operators $\hat{h}^{\mathrm{dia}}_{K,uv}$ and $\hat{\mathbf{h}}^{\mathrm{orb}}$ for many-electron systems can be obtained similar to the J-coupling operators (3.32) by differentiating the terms of the chemical shielding Hamiltonian $\hat{H}^{\mathrm{CS}}_{K}$ (3.34b) and summing over all electrons:

$$\hat{h}^{\mathrm{dia}}_{K,uv} = \left(\frac{\mu_0}{4\pi}\right)\frac{e^2}{2m_\mathrm{e}}\sum_{a=1}^{N}\frac{\mathbf{r}_{a\mathrm{O}}\cdot\mathbf{r}_{aK}\delta_{uv} - r_{a\mathrm{O},v}r_{aK,u}}{r^3_{aK}} \tag{3.37a}$$

$$\hat{\mathbf{h}}^{\mathrm{orb}} = \frac{\mu_\mathrm{B}}{\hbar}\sum_{a=1}^{N}\hat{\mathbf{l}}_{a\mathrm{O}} \ , \tag{3.37b}$$

The form of the perturbation operator $\hat{\mathbf{h}}^{\mathrm{PSO}}_{K}$ was given above (3.32b).

3.3 DFT calculations of NMR parameters

From the theoretical point of view, there is a fundamental obstacle to using DFT in magnetic response calculations: the electronic energy and other molecular properties in magnetic field depend not only on the electron density but also on the current density. Fortunately, the dependence of NMR parameters on the current density is typically small and it can be neglected [146]. Therefore, the usual current-density independent DFT functionals can be used in NMR calculations.

3.3.1 DFT calculation of *J*-coupling

This work involved DFT calculations of all four Ramsey's terms of the isotropic part of $\overleftrightarrow{\mathbf{J}}_{KL}$ tensor, the scalar spin–spin coupling (J-coupling). DFT is very efficient for calculating the J-coupling. It includes the electron correlation, which is essential for correct description of the electron-mediated mechanism of J-coupling. Moreover, the DFT results are not much affected by the triplet instability problem of HF and related *ab initio* methods (see below).

First DFT J-coupling calculations were done by combining the finite-perturbation method for the FC term with the uncoupled DFT calculation of the DSO and PSO terms, and neglecting the difficult SD term [119, 147–149]. Rigorous CPKS implementations of the FC and SD terms were developed slightly later [118, 124]. Standard GGA functionals were shown to provide quite accurate J-couplings. Despite the triplet instability problem, the addition of a small portion of the HF exchange ($\sim 20\,\%$) on average improves the pure GGA results [118, 124].

Practical J-coupling calculations exploit the fact that if the CP equations are solved for the perturbations placed on one ('perturbing') nucleus, the J-couplings with all other ('responding') nuclei are available only at the cost of integral evaluation. A smart selection of the perturbing nuclei in the J-coupling calculations can thus substantially reduce the computational costs.

***J*-coupling terms.** DFT calculations of the individual J-coupling terms have certain specifics. The DSO term does not involve the perturbed orbitals and it thus does not enter the CP scheme. It only requires the evaluation of one-electron integrals of the type

$$\langle\phi_\mu|\frac{r_{K,u}}{r_K^3}\frac{r_{L,v}}{r_L^3}|\phi_\nu\rangle \ , \tag{3.38}$$

where ϕ_μ, ϕ_ν are basis functions and $u, v = \{x, y, z\}$. These integrals are not known analytically; they can be evaluated either by numerical quadrature or by applying the resolution of identity [118], converting thus the problem to the calculation of integrals of the PSO type (3.39). The DSO term is mostly small and of the four J-coupling terms it depends the least on the DFT functional choice.

The pure imaginary PSO perturbation operator $\hat{\mathbf{h}}_K^{\mathrm{PSO}}$ (3.32b) yields pure imaginary first-order spin orbitals. As mentioned in section 2.5.3, the imaginary perturbations do not enter the CPKS scheme and the calculation of the PSO term with non-hybrid DFT functionals thus reduces to the uncoupled DFT level. When HF exchange is added, all three Cartesian components of the PSO perturbation have to be evaluated by the CP procedure. The PSO calculations involve integrals of the type

$$\langle\phi_\mu|\frac{r_{K,u}}{r_K^3}|\phi_\nu\rangle \ , \tag{3.39}$$

which are zero for s-orbitals at the nucleus K. The PSO term can be large when low-lying virtual orbitals with a distinct non-s-character (p-, d-orbitals) are present [118].

The FC term is isotropic and, hence, only the zz-component $J_{KL,zz}^{\mathrm{FC}} = J_{KL}^{\mathrm{FC}}$ has to be calculated via the CP procedure. Moreover, the Dirac δ-function in the FC perturbation operator $\hat{\mathbf{h}}_K^{\mathrm{FC}}$ (3.32c) reduces the FC integrals to simple products of the orbital values at the position of the perturbing nucleus K:

$$\langle \chi^{(0)}_{a\sigma} | \hat{h}^{\mathrm{FC}}_{K,zz} | \chi^{(0)}_{q\sigma'} \rangle \sim \sigma \delta_{\sigma\sigma'} \psi^*_a(\mathbf{R}_K) \psi_q(\mathbf{R}_K) \ , \tag{3.40}$$

where $\sigma, \sigma' = \{\alpha, \beta\}$. Thus, the largest contributions to the FC term arise from the core s- and bonding sigma-electrons of the perturbing nucleus, which requires high precision on the DFT integration grid and basis set flexibility, mainly in the core region of the s-subspace. Nevertheless, the calculation of the FC term is very efficient and since it often dominates the total J-coupling, it can be used for qualitative estimates of J-coupling trends when other terms are too expensive.

The SD term is the most demanding as it requires the CP calculation for six perturbations. It is significant in similar cases as the PSO term, i.e. when low-lying non-s- virtual orbitals are present at the coupled nuclei [118].

In summary, the J-coupling calculations require the iterative solution of CP equations for 10 perturbations (7 for non-hybrid DFT) per each 'perturbing' nucleus. The importance of the individual J-coupling terms can be roughly estimated from the bonding situation in the molecule. Large basis sets including tight s-functions with large exponents (due to the FC term) and polarization functions (due to the PSO and SD terms) are needed for satisfactory results.

Triplet instability. The calculation of the FC and SD terms is complicated by the involvement of triplet virtual states and the consequent spin polarization of the perturbed wave function. In the HF approximation, the exact exchange energy contribution lowers the triplet-state energies too much relative to the singlet ground state. The corresponding triplet excitation energies can be near zero or even negative in certain cases (such as multiply-bonded systems) [77, 150] and the HF J-coupling calculation then provides meaningless results.

The description of singlet and triplet states in DFT is unbalanced as well, but the consequences are not so drastic as for HF. Unfortunately, the triplet instability is exhibited also by the correlated *ab initio* methods which involve HF orbitals relaxed in response to the magnetic perturbation [150]. Reliable *ab initio* alternatives, which avoid the instability problem, are provided by the multi-reference methods [151] or by the correlated single-reference methods that allow calculations without the orbital relaxation, such as the orbital-unrelaxed CCSD method [152, 153].

3.3.2 DFT calculations of chemical shielding

The chemical shielding calculations in this work were done for the whole $\overleftrightarrow{\boldsymbol{\sigma}}$ tensor including its principal component magnitudes σ_{ii} and orientations $\mathbf{e}_i$, $i = 1, 2, 3$. The CP-GIAO implementation of chemical shielding at the hybrid DFT level was used [117].

Performance of the HF and DFT methods for chemical shielding is different than for the J-coupling. While the HF method is inadequate for the J-coupling, the chemical shielding calculated with HF is already quite accurate. Actually, the popular B3LYP functional is on average only as accurate as the HF method [117]. The optimal amount of HF exchange in hybrid DFT calculations of NMR chemical shielding is usually close to $\sim 50\,\%$.

Gauge origin problem. The chemical shielding calculations are complicated by the fact that the NMR shielding perturbation operators (3.37a) and (3.37b) and the corresponding diamagnetic and paramagnetic terms of the chemical shielding tensor $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{\mathrm{dia}}$ and $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{\mathrm{para}}$ are dependent on the position of the origin of the Cartesian coordinate system (in this context usually called 'gauge origin'). In exact theory, the gauge-dependent parts of $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{\mathrm{dia}}$ and $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{\mathrm{para}}$ mutually cancel [114, 154] and the total chemical shielding is gauge-independent. In actual NMR calculations, however, the cancellation is only approximate and the calculated chemical shielding depends on the position of molecule relative to the gauge origin. The common-gauge chemical shielding calculations thus require extremely large basis sets.

The gauge origin dependence of molecular properties can be suppressed by several methods. Probably the most popular approach, which was used also in this work, is GIAO (gauge-including atomic orbitals) [155,156], where the atomic orbitals are assumed in the form of London orbitals [157] explicitly dependent on the external magnetic field:

$$\phi_{\mu}^{\mathrm{GIAO}}(\mathbf{r}_{\mu}) = \exp\Big[-i\frac{e}{2\hbar}(\mathbf{B}_0 \times \mathbf{R}_{\mu}) \cdot \mathbf{r}\Big]\phi_{\mu}(\mathbf{r}_{\mu}) \ . \tag{3.41}$$

The expressions for the GIAO-NMR calculations at the CPHF and CPKS levels are found by replacing the casual field-independent AOs with the London orbitals. Thus, apart from the usual expressions that are present in the common-gauge theory, additional terms appear in GIAO in both diamagnetic and paramagnetic shielding components, which compensate the gauge dependence [155, 156].

Chemical shielding terms. The one-electron contributions to the diamagnetic shielding tensor $\overset{\leftrightarrow}{\boldsymbol{\sigma}}^{\mathrm{dia}}$ are calculated as integrals of the diamagnetic perturbation operator (3.37a) with unperturbed spin orbitals $\chi_a^{(0)}$.

The calculations of paramagnetic chemical shielding with non-hybrid DFT methods differ from the HF calculations. The perturbation operators $\hat{\mathbf{h}}^{\mathrm{orb}}$ and $\hat{\mathbf{h}}_K^{\mathrm{PSO}}$ are pure imaginary and the non-hybrid DFT calculations of chemical shielding thus reduce to the uncoupled DFT level (see section 2.5.3) [158]. When the HF exchange is involved, the HF contribution to the paramagnetic chemical shielding term has to be calculated by the iterative CPHF procedure [117]. The operator $\hat{\mathbf{h}}^{\mathrm{orb}}$ is a natural choice for the perturbation entering the CPHF scheme since there are only three CPHF equations that have to be solved and the chemical shielding tensors in the whole molecule are then available at the cost of integral evaluation.

4 Results

This chapter gives an overview of the most important results of this doctoral thesis. The published works are appended in a separate section.

Computational NMR studies of L-alanyl-L-alanine comprise of publications I–III. The contribution I presented at a students' conference contained initial results of theoretical modeling of the effective CSAs in peptide backbone. The set of calculated effective CSAs was later greatly extended and published as a full paper III. The publication II resulted from a systematic computational and experimental work focused on the pH dependence of NMR parameters in AA.

Theoretical modeling of NMR parameters in nucleic acid phosphate was a subject of publications V–VI. The study V presents a thorough analysis of the effect of direct Mg^{2+} coordination to NA phosphate. Dependencies of ^{31}P NMR parameters on (ζ, α) conformation of NA phosphate were explored in the work VI. Properties of the Mg^{2+} ion itself were investigated in the publication IV, where the theoretical calculations were used for the interpretation of experimental Raman spectra of aqueous salt solutions.

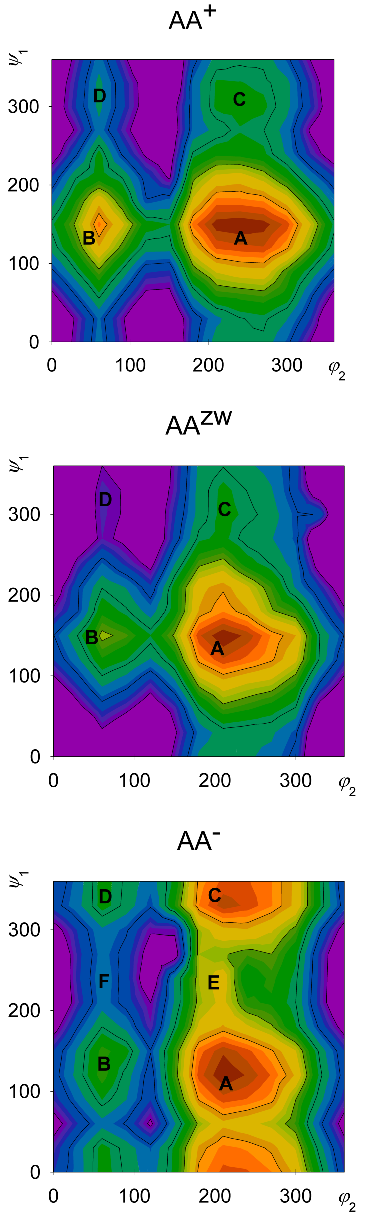
4.1 Computational NMR studies of L-alanyl-L-alanine

The cation, zwitterion and anion of AA were characterized in the work II by combining NMR calculations and experiment. Firstly, the potential energy surfaces (PES) of AA^{+}, AA^{zw} and AA^{-} in dependence on torsion angles ψ_1 and φ_2 were calculated with the BPW91/6-311++G(d,p) method including PCM of water solvent. In the case of AA^{-}, rotation and intramolecular hydrogen bonding of the -NH_2 group was taken into account by calculating three PESs for the three minima of -NH_2 rotation.

Four energy minima were found for AA^{zw} (denoted **A**–**D**), which was in agreement with the previous results obtained with sparser (ψ_1, φ_2) grid [5]. By comparing their relative energies we confirmed that only the lowest-energy AA^{zw} conformer $\mathbf{A} = (147°, 207°)$ is significantly populated at room temperature (Table 4.1). In the AA^{+} form, four energy minima analogous to those of AA^{zw} were found. The conformer $\mathbf{A} = (149°, 239°)$ is again global minimum, but the conformer $\mathbf{B} = (150°, 60°)$ is also thermally reachable. These theoretical findings indicated that the AA molecule exists in a well-defined conformation at room temperature and at pH ranging from neutral to low. The most complex case was the AA^{-} form, for which in total seventeen local minima were obtained. For all -NH_2 orientations, the conformers $\mathbf{A} \sim (125°, 210°)$ and $\mathbf{C} \sim (350°, 210°)$ were significantly populated and a new energy minimum $\mathbf{E} \sim (250°, 205°)$ appeared between **A** and **C**. We therefore concluded that the AA^{-} molecule is rather flexible and at higher temperatures it might even undergo hindered rotation along the ψ_1 dimension.

Table 4.1 Calculated (BPW91/6-311++G(d,p)/PCM=water) geometry parameters, relative energies and Boltzmann populations (300 K) of individual conformers of the three AA forms (left).[a] Graphical representation of the calculated potential energy surfaces (right).[b]

Conformer	ψ_1	φ_2	ΔE	p [c]
AA^+				
A	149	239	0.0	91.4
B	150	60	1.4	8.6
C	305	233	5.1	0.0
D	305	53	6.3	0.0
AA^{zw}				
A	147	207	0.0	99.8
B	150	66	3.6	0.2
C	309	210	5.1	0.0
D	312	64	9.2	0.0
AA^-				
A	127	208	0.0	45.4
B	135	63	4.3	0.0
C	339	210	1.4	4.0
D	343	62	6.1	0.0
E	255	206	2.7	0.5
F	244	61	6.9	0.0
A′	128	209	0.8	11.7
B′	136	63	5.2	0.0
C′	8	213	0.8	12.3
D′	14	62	5.0	0.0
E′	283	204	3.9	0.1
F′	–	–	–	–
A″	119	210	1.9	1.7
B″	128	62	6.3	0.0
C″	341	208	0.4	24.1
D″	344	63	5.0	0.0
E″	215	204	3.3	0.2
F″	233	61	7.5	0.0



[a] Torsion angles ψ_1 and φ_2 in degree, relative energies ΔE in kcal · mol^{-1} and relative conformer populations p in %. Unprimed, primed and doubly-primed AA^- conformers correspond to the three -NH_2 rotamers.

[b] For the AA^- form, only the PES for the unprimed rotamer with is depicted.

[c] Boltzmann populations calculated from the relative energies ΔE.

4.1.1 Geometry dependencies of NMR chemical shifts and *J*-couplings

In the experimental part of publication II we performed a complete assignment of ^{1}H, ^{13}C, and ^{15}N chemical shifts and *J*-couplings measured for AA and L-alanine in aqueous solution at pH 2, 7, and 12. By varying the pH in smaller steps we found that the AA dissociation constants are $pK_1 = 3.12$ and $pK_2 = 8.30$, confirming thus that at pH 2, 7, and 12 only the cationic, zwitterionic and anionic form of AA was present, respectively.

Computational part of work II was initiated by testing the performance of B3LYP functional, IGLO-III basis set, and PCM of water in NMR calculations. Theoretical NMR chemical shifts and *J*-couplings were obtained for the L-alanine cation, zwitterion, and anion and compared with NMR experiment for L-alanine at pH 2, 7, and 12 (work II, Table 2). The B3LYP/IGLO-III/PCM approach performed well for *J*-couplings and somewhat worse for chemical shifts. The largest discrepancy obtained for the ^{15}N chemical shift was probably caused by neglecting the multireference character of the ^{15}N NMR standard CH_3NO_2 [159]. Nevertheless, the pH-induced changes in NMR parameters were calculated in qualitative agreement with the experiment. The remaining errors were attributed to incomplete description of the solute–solvent interactions by PCM and also to the absence of vibrational corrections. The significance of both effects was confirmed in the follow-up studies [12, 13], where the explicit solvent averaging and the inclusion of vibrational effects were shown to substantially improve the agreement between NMR calculations and experiment.

Having tested our computational approach for the simple case of L-alanine, we applied it also to the AA molecule. A complete set of NMR chemical shifts and *J*-couplings was calculated for local energy minima of the three AA forms and also for the full set of AA geometries on the (ψ_1, φ_2) grid. We found that the NMR parameters (especially *J*-couplings) calculated for the populated minima were on average in better agreement with the NMR experiment than those for higher-energy minima (work II, Table 4 and Table 5). This confirmed our theoretical assignment of the most populated AA conformers in solution.

NMR parameters in the AA dipeptide are influenced by both molecular conformation and charge of terminal groups. The pH-induced changes of chemical shifts and *J*-couplings in AA were reproduced with similar accuracy as for the L-alanine. Naturally, the changes were the largest for NMR parameters in the vicinity of the AA (de)protonation site (work II, Figure 6 and Figure 7).

The $^{3}J_{\mathrm{NH}_{\alpha1}}$ and $^{3}J_{\mathrm{H_N H}_{\alpha2}}$ couplings assigned to peptide backbone torsion angles ψ_1 and φ_2, respectively, were compared with semi-empirical Karplus curves known from the literature (work II, Figure 9). A very good agreement was found for the $^{3}J_{\mathrm{H_N H}_{\alpha2}}$ coupling, but we noted a significant difference for $^{3}J_{\mathrm{NH}_{\alpha1}}$. The discrepancy was caused by comparison with an outdated $^{3}J_{\mathrm{NH}_{\alpha1}}$ parameterization by Demarco et al. [160]; the improved parameterization by Wang et al. [161] matches our calculations. As expected, the ${}^3J_{\mathrm{NH}_{\alpha1}}$ coupling in the AA$^-$ form differed from that in the AA^{zw} and AA$^+$ forms due to deprotonation of the -NH$_3^+$ group occurring close to the $\mathrm{H}_{\alpha1}$ atom.

Apart from vicinal ${}^3J$-couplings assigned to AA backbone torsion angles ψ_1 and φ_2, the variation of AA conformation also affected the geminal and one-bond spin–spin couplings. In particular, the ${}^2J_{\mathrm{C'H}_{\alpha1}}$ and ${}^1J_{\mathrm{C}_{\alpha1}\mathrm{H}_{\alpha1}}$ couplings were significantly modulated only in the ψ_1 dimension and the ${}^1J_{\mathrm{C}_{\alpha2}\mathrm{H}_{\alpha2}}$ coupling only in the φ_2 dimension. Although these couplings generally cannot be assigned to a single torsion in larger peptides, they can be used for determining the conformation of dipeptides in solution.

4.1.2 Geometry dependencies of cross-correlated relaxation rates

In the study I we analyzed the dependencies of N–NH$_\mathrm{N}$, N–C$_{\alpha1}$H$_{\alpha1}$ and N–C$_{\alpha2}$H$_{\alpha2}$ effective CSAs (as introduced in the section 3.1.4) on conformation and charge of the AA molecule. We proposed pH-independent Karplus-like parameterizations for the N–C$_{\alpha1}$H$_{\alpha1}$ and N–C$_{\alpha2}$H$_{\alpha2}$ effective CSAs.

The successive work III greatly extended the set of modeled effective CSAs, encompassing virtually all sensible combinations of the ^{15}N or ^{13}C$'$ chemical shielding tensors and adjacent dipolar vectors in the AA molecule (see work III, Figure 1). We even considered dipolar interactions between not directly bonded atoms, for which the distance factor $1/r^3_{JK}$ (see equation 3.20) significantly alters the magnitude of CCR rates. We therefore focused on conformational dependence of the product $1/r^3_{JK} \cdot \Delta\sigma^{\mathrm{eff}}_{I,JK}$. Effective CSAs involving the ^{1}H$_\mathrm{N}$ chemical shielding tensor were not analyzed since they are substantially influenced by hydrogen bonding, for which our computational model didn't account.

The applicability of each CSA–DD cross-correlation mechanism was assessed based on several criteria, the most preferential being smooth, large-scale, and one-dimensional dependence of the product $1/r^3_{JK} \cdot \Delta\sigma^{\mathrm{eff}}_{I,JK}$ on AA conformation. We confirmed that the CCR rates $\Gamma_{X,\mathrm{C}_\alpha\mathrm{H}_\alpha}$ ($X = \mathrm{N}, \mathrm{C}'$; $\alpha = \alpha1, \alpha2$) that have been probed experimentally [2, 3, 162–164] are the most convenient for structural interpretation, exhibiting large and smooth dependence on either ψ_1 or φ_2 (work III, Figure 6). We identified several other promising CCR mechanisms, namely the $\Gamma_{X,\mathrm{NH}_{\alpha2}}$ and $\Gamma_{X,\mathrm{C'H}_{\alpha1}}$ ($X = \mathrm{N}, \mathrm{C}'$) CCR rates, which could also serve for peptide structure determination, but their applicability has yet to be experimentally tested.

4.2 Modeling the effects of molecular flexibility and solvation on NMR parameters in nucleic acid phosphate

Three models of NA phosphate were employed in works V and VI (Figure 4.1): dimethyl phosphate (DMP), ethyl methyl phosphate (EMP), and abasic RNA dinucleotide monophosphate (nPn). The latter two models contain the inherent asymmetry of NA backbone and were thus used for modeling the dependencies NMR parameters on molecular conformation and solvation. The symmetry of the DMP molecule was exploited in benchmark coupled-clusters NMR calculations.

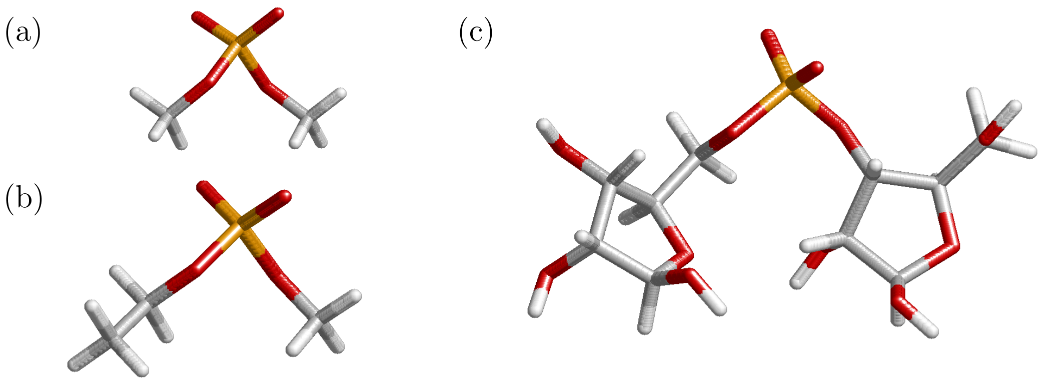


Figure 4.1 NA phosphate models: (a) dimethyl phosphate, (b) ethyl methyl phosphate, (c) abasic RNA dinucleotide monophosphate.

4.2.1 Benchmark NMR calculations

Testing of computational methods in works V and VI brought several methodological insights into the calculation of ^{31}P NMR parameters in NA phosphate. The performance of B3LYP functional was compared with the coupled-clusters method, we tested the convergence of electronic basis set and the influence of implicit and/or explicit solvent models on ^{31}P NMR parameters.

Basis set convergence of B3LYP calculations of isotropic ^{31}P chemical shielding σ_{P} and Fermi-contact term of $^{2}J_{\mathrm{PC}}$ coupling was explored in work V (Figure 2). Very good cost-effective results were obtained for both σ_{P} and $^{2}J_{\mathrm{PC}}^{\mathrm{FC}}$ with the 6-311++G(d,p), IGLO-III and cc-pCVTZ basis sets. Surprisingly, the performance of Jensen's (aug)-pcJ-n basis sets ($n = 0, 1, 2$) specifically designed for J-coupling calculations was only mediocre, what was probably caused by different bonding situations in NA phosphate and in model systems, for which these basis sets were optimized [165].

The isotropic ^{31}P chemical shift δ_{P} obtained with the secondary standard approach (3.18) was found quite sensitive to local phosphate geometry, electronic structure method and solvation model used in the NMR calculations. The δ_{P} decreased by $\sim$6 ppm when geometries of DMP and secondary standard PH_3 were optimized with the CCSD method and by additional $\sim$3.5 ppm when both molecular structures and NMR shielding tensors were obtained at the CCSD level (work VI, Table 2). Nevertheless, both B3LYP and CCSD methods provided essentially the same conformational differences of δ_{P}, indicating thus that the B3LYP functional is suitable for exploring the geometry dependence of δ_{P}.

The effect of water solvent on δ_P varied with the solvation model used (work V, Table 2), indicating that dynamical averaging of explicit solvent is necessary to obtain converged results. Importantly, we found that the effect of water solvent on δ_P is not uniform, but it varies with the phosphate conformation (work VI, Table 2). Reliable solvent description is therefore needed for accurate modeling of geometry dependencies of δ_P. Dynamical averaging of explicit solvent is probably the most reliable approach, but it becomes rather costly for a large number of conformers. In such a case, implicit solvation is a more feasible alternative enabling consistent treatment of hydration effects in all conformers of NA phosphate.

Principal components of the ^{31}P chemical shielding tensor $\sigma_{\mathrm{P},ii}$, $i = 1, 2, 3$ were much more sensitive to the NMR computational method than δ_P. The $\sigma_{\mathrm{P},ii}$ values calculated with the CCSD method differed by up to 14 ppm from those obtained with the B3LYP functional. The inclusion of PCM hydration changed the $\sigma_{\mathrm{P},ii}$ values even more, by up to 50 ppm. The B3LYP/PCM results differed from previous calculations with explicit phosphate hydration [30], implying that PCM accounts only partially for the solvent-induced changes in $\sigma_{\mathrm{P},ii}$. Interestingly, the orientation of the ^{31}P chemical shielding tensor in the phosphate-frame coordinate system was practically independent of the computational method.

The FC term of the $^2J_\mathrm{PC}$ coupling is dominant [22]. Unlike for δ_P, the difference between $^2J^\mathrm{FC}_\mathrm{PC}$ calculated with B3LYP and CCSD methods depends on phosphate conformation (work VI, Table 2). This fact limited the accuracy of structural assignment of the $^2J_\mathrm{PC}$ couplings calculated with the B3LYP functional. Similarly as for δ_P, the $^2J_\mathrm{PC}$ values were very sensitive to the solvation model used (work V, Table 2). Also in analogy to δ_P, the effect of implicit PCM hydration on $^2J^\mathrm{FC}_\mathrm{PC}$ varied with phosphate conformation (work VI, Table 2). The importance of reliable modeling of water solvent for the ^{31}P NMR calculations in NA phosphate was thus again confirmed.

4.2.2 Geometry dependencies of ^{31}P NMR parameters

The ^{31}P chemical shielding tensor, the $\Gamma_\mathrm{P,CH}$ CCR rates, and the $^nJ_\mathrm{PC}$ spin–spin coupling constants, $n = 2, 3$, were calculated in study VI in dependence on NA backbone torsion angles ζ and α (P–O torsions). Absolute accuracy of the calculated values was not crucial in this case, since we were mainly interested in the geometry trends. Thus, the B3LYP/IGLO-III method and implicit PCM hydration was used.

Values of δ_P calculated with the secondary standard approach were somewhat larger than the experimental δ_P scale, which could be explained by the imperfections of the B3LYP/PCM computational approach. Nevertheless, the experimentally known difference of 1.6 ppm between the $\mathrm{B_{II}}$ and $\mathrm{B_I}$ conformations of NA backbone [24] was reproduced quite well: the δ_P difference calculated between the $(t, g-)$ and $(g-, g-)$ phosphate conformers was 2.4 ppm in EMP and 1.8 – 2.6 ppm in nPn. The B3LYP/PCM computational approach thus proved suitable for modeling the geometry dependence of δ_P in NAs (work VI, Table 3). Our calculations provided a systematic first-principles calibration of the dependence of δ_P on torsion angles ζ and α in a full $(0°, 360°)$ range.

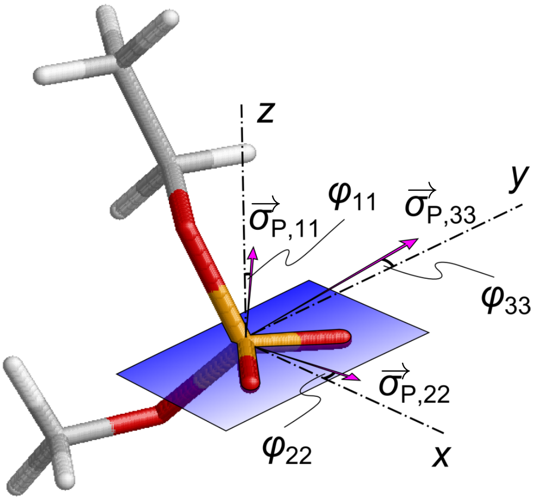

z
y
x
$\vec{\sigma}_{\mathrm{P},11}$
$\vec{\sigma}_{\mathrm{P},33}$
$\vec{\sigma}_{\mathrm{P},22}$
φ_{11}
φ_{33}
φ_{22}

Figure 4.2 Schematic representation of the ^{31}P chemical shielding tensor in EMP molecule: the principal components $\vec{\boldsymbol{\sigma}}_{\mathrm{P},ii} = \sigma_{\mathrm{P},ii}\,\mathbf{e}_i$, $i = 1, 2, 3$, where $\mathbf{e}_i$ are unit vectors, and their deviations φ_{ii} from the axes of the phosphate-frame coordinate system defined by Herzfeld et al. [166]. The correspondence between the principal components $\sigma_{\mathrm{P},11} < \sigma_{\mathrm{P},22} < \sigma_{\mathrm{P},33}$ and phosphate-frame axes is as follows: $11 \leftrightarrow z$, $22 \leftrightarrow x$, $33 \leftrightarrow y$.

Variation of principal components $\sigma_{\mathrm{P},ii}$ on the (ζ, α) grid (by up to 46 ppm) was much larger than the range of isotropic values δ_P ($\sim$13 ppm), what indicated mutual compensation of the components (work VI, Figure 3). Further analysis revealed that the geometry dependencies of both δ_P and chemical shielding anisotropy $\Delta\sigma_{\mathrm{P},}$ are dominated by the principal component $\sigma_{\mathrm{P},33}$, owing to mutual compensation of $\sigma_{\mathrm{P},11}$ and $\sigma_{\mathrm{P},22}$. Orientation of the ^{31}P chemical shielding tensor in the molecular frame (Figure 4.2) was found almost independent of torsion angles ζ and α. Deviations φ_{ii} of ^{31}P principal components from the phosphate-frame axes were quite small, rarely exceeding 6°. The deviations measured [166] in the BDEP crystal containing Ba^{2+} ions directly coordinated to the phosphate group [167] were somewhat larger ($\varphi_{11} \approx 7°$, $\varphi_{22} \approx 13°$, $\varphi_{33} \approx 9°$). Increased φ_{ii} values were also calculated in the study V for the phosphate group directly coordinated by Mg^{2+}. The correspondence between the ^{31}P chemical shielding tensor in BDEP and in NAs is quite important since the BDEP ^{31}P-tensor is used for the interpretation of ^{31}P NMR relaxation measurements in NAs. We proposed in work VI that the ^{31}P chemical shielding tensor measured in BDEP is influenced by the coordination of Ba^{2+} ions and that φ_{ii} values smaller than 6° are natural for all (ζ, α) conformations.

The quite large variation of $\sigma_{\mathrm{P},ii}$ on the (ζ, α) grid could affect the $\Gamma_{\mathrm{P,CH}}$ CCR rates. To test the validity of the rigid tensor approximation we calculated the dependence of

$\Gamma_{\mathrm{P,C5'H5'1}}$ and $\Gamma_{\mathrm{P,C5'H5'2}}$ CCR rates on torsion angles ζ and α using 'relaxed' and 'rigid' approach (for details see the Methods section of work VI). These CCR rates depend dominantly on torsion angles α and β [26, 27]. We found that they are only slightly modulated by torsion angle ζ (work VI, Figure 4). The 'rigid' approach thus provided $\Gamma_{\mathrm{P,CH}}$ values close to those obtained using the 'relaxed' approach. Validity of the rigid tensor approximation for the interpretation of $\Gamma_{\mathrm{P,CH}}$ CCR rates was thus confirmed.

Torsion angle dependencies of the $^2J_{\mathrm{PC}}$ couplings revealed that both ζ and α modulate the $^2J_{\mathrm{PC}}$ value, but only one of the P–O torsions is dominant for each coupling (work VI, Figure 5). Thus, the $^2J_{\mathrm{PC3'}}$ and $^2J_{\mathrm{PC5'}}$ couplings can be loosely assigned to torsion angles ζ and α, respectively. We also calculated dependence of the $^2J_{\mathrm{PC5'}}$ coupling on torsion β, using the EMP molecule with torsions ζ and α fixed to 290° (work VI, Supporting Information). The $^2J_{\mathrm{PC5'}}$ coupling was nearly constant in the whole $120 - 240°$ region, which is predominantly populated in NAs. A geometry change to one of the rare NA conformers with β near 80° [21] would be accompanied by increase the of $^2J_{\mathrm{PC5'}}$ absolute value by $\sim$1 Hz. Similar dependence of $^2J_{\mathrm{PC3'}}$ on torsion angle ϵ may be expected.

The absolute $^2J_{\mathrm{PC}}$ values calculated with the B3LYP/PCM method appeared too large when compared with the available experimental data. In particular, the B3LYP/IGLO-III/PCM results for the $(g-, g-)$ conformer were near -8 Hz, while the experimental range for the this conformer is $4.5 - 5.5$ Hz [35,37] (work VI, Table 4; the J-coupling sign is usually not determined experimentally). We were able to explain the discrepancy between $^2J_{\mathrm{PC}}$ calculations and experiment. Assuming simple additivity of computational errors as discussed in section 2.6, we separately evaluated the error of the B3LYP functional in $^2J_{\mathrm{PC}}$ calculations, the error due to inaccuracy of the B3LYP geometry, the basis-set error, and the influence of explicit water solvent dynamics. All of these contributions were similarly important, each amounting in absolute value ca. $0.5 - 1$ Hz. Their added effect shifted the calculated $^2J_{\mathrm{PC}}$ values to within $\sim$1 Hz from the experimental range (work VI, Table 4). Thus, the most important sources of computational errors were identified, allowing to use the calculated geometry dependencies for quantitative predictions.

Bearing in mind the limited accuracy of the B3LYP/IGLO-III/PCM computations as well as the inherently multi-dimensional character of the $^2J_{\mathrm{PC}}$ geometry dependencies, we proposed qualitative rules for the structural interpretation of $^2J_{\mathrm{PC}}$ couplings:

- If the $^2J_{\mathrm{PC}}$ coupling is in absolute value smaller than $\sim$3 Hz, the orientation of the assigned torsion angle (ζ or α) is different from typical *gauche* (near 70 or 290°).
- If the $^2J_{\mathrm{PC}}$ coupling is in absolute value larger than $\sim$4.5 Hz, the orientation of the assigned torsion angle (ζ or α) deviates from the *trans* region ($120 - 240°$).

If confirmed experimentally, these rules could serve as new restraints on the NA phosphate conformation, similar to those already existing for δ_{P} [24].

Modulation of the $^3J_{\mathrm{PC4'}}$ coupling by torsions ζ and α is also significant, up to $\sim$2 Hz. This could cause uncertainty as large as $\sim$25° in the determination of torsion angle β via standard $^3J_{\mathrm{PC4'}}$ Karplus equations [34, 168, 169]. This result represents the missing link between the $^3J_{\mathrm{PC}}$ spin–spin coupling constants and the P$-$O torsions of NA backbone.

4.2.3 Response of ^{31}P NMR parameters to Mg^{2+} coordination

Coordination sites of Mg^{2+} and water at NA phosphate were determined in the first part of the study V for a set of 53 RNA crystal structures containing over 7700 nucleotides by the method of Fourier averaging. Two sharp Mg^{2+} distributions analogous to those obtained previously for organic phosphate salts [43] were found, each localized at 1.9 Å from one of the charged phosphate oxygens OP1 and OP2. Five water coordination sites similar to those found in DNA [45] were confirmed; the sixth site completing the phosphate first hydration shell was somewhat less populated. Based on the bioinformatic analysis, the effects of direct (inner-shell) Mg^{2+} coordination to NA phosphate were evaluated with two approaches: (i) static NMR calculations for DFT-optimized nPn structures with explicit hydration or Mg^{2+} coordination patterns (see Figure 1.3 for an example) and (ii) averaging of NMR parameters in nPn solvated by clusters of explicit solvent generated with molecular dynamics (MD).

NMR calculations were carried out for phosphate geometries with static B-DNA and A-RNA solvation patterns. The trends in NMR parameters upon Mg^{2+} coordination obtained for the two kinds of phosphate solvation differed (work V, Table 3). In the case of B-DNA solvation, δ_{P} varied nonspecifically, $\Delta\sigma_{\mathrm{P},}$ decreased by $35 - 40$ ppm, and the $^2J_{\mathrm{PC3'}}$ and $^2J_{\mathrm{PC5'}}$ coupling magnitudes increased site-specifically upon Mg^{2+} coordination. For the solvation patterns found in A-RNA, δ_{P} decreased by $\sim$4 ppm, $\Delta\sigma_{\mathrm{P},}$ decreased by $20 - 40$ ppm, and the $^2J_{\mathrm{PC3'}}$ and $^2J_{\mathrm{PC5'}}$ couplings varied only slightly and without the site-specificity. Therefore, the NMR parameters obtained with the static approach depended more significantly on the particular phosphate solvation pattern than on whether the Mg^{2+} ion was coordinated or not. In other words, limited number of NA backbone solvation patterns, although selected reliably, cannot provide accurate trends for NMR parameters in NA phosphate and averaging over solvent reorientation is necessary.

The next step in theoretical calculation of the effects of Mg^{2+} coordination was thus modeling of solvent dynamics around NA phosphate. Three MD simulations were carried out, one with hydrated NA phosphate (`nomg`) and two for the cases of Mg^{2+} coordination at phosphate oxygen OP1 (`mg1`) or OP2 (`mg2`). The Mg^{2+} ions remained directly coordinated to the phosphate group throughout the simulations `mg1` and `mg2`. To eliminate the influence of geometry dependence of NMR parameters, the backbone conformation was restrained in the A-form by applying a rather strong 500 kcal $\cdot$ mol^{-1} $\cdot$ Å^{-2} energy penalty for displacement of backbone atoms from their initial DFT-optimized positions.

From each MD run, snapshot geometries including solvent molecules inside the sphere of radius 5.5 Å around the phosphorus atom were taken every 50 ps. It was verified that the dominant effect on NMR parameters comes from the first hydration shell surrounding the phosphate and Mg^{2+}, while the effects of higher solvent layers can be modeled with PCM (work V, Figure 3).

Although the NA backbone was restrained in the MD simulations, the bond distances and bond angles relaxed during MD equilibration and the MD snapshots geometries therefore differed from the DFT-optimized nPn structure. To reduce the influence of MD force-field on NA phosphate geometry while preserving the solvation shell structure, we re-optimized the positions of NA backbone atoms in the MD snapshots with the solvent kept fixed. NMR parameters were then calculated both for original MD snapshots ('unrelaxed' calculations) and for partially optimized geometries ('relaxed' calculations). The resulting data were analyzed in terms of mean values and standard errors of the mean (SEM). The 'unrelaxed' calculations provided different trends upon Mg^{2+} coordination and larger SEM values than the 'relaxed' ones (work V, Table 4). The latter approach can be considered more appropriate for modeling the NA phosphate solvation and was used to gather the following results.

Upon Mg^{2+} coordination, the mean values of ^{31}P chemical shift decreased by $5-6$ ppm, which agrees with the results obtained for static A-RNA solvation. Taking into the account the SEM values, δ_P decreased by $2-9.5$ ppm, which is in excellent agreement with the $1.5-10$ ppm decrease measured in thio-substituted hammerhead ribozyme upon Cd^{2+} coordination [46–48]. The measured decrease of δ_P due to Mg^{2+} coordination to phosphate in minimal metal binding motif of hammerhead ribozyme was much smaller, only 0.3 ppm [49], which could be explained by weaker binding and faster exchange of Mg^{2+} at the metal binding site. Detailed look on the principal components $\sigma_{\mathrm{P},ii}$ revealed that the calculated decrease of δ_P is solely due to the increase of the $\sigma_{\mathrm{P},33}$ component (work V, Supporting Information, Figure S1), whose importance for geometry dependence of δ_P was already noted above.

Apart from δ_P, that is up to our knowledge the only NMR parameter used so far for the detection of NA phosphate metalation, statistically significant effects of Mg^{2+} coordination were obtained also for $\Delta\sigma_{\mathrm{P},}$ $(0-25$ ppm increase) and $^{2}J_{\mathrm{PC5'}}$ $(0.2-1.8$ Hz decrease of its magnitude). The results for other NMR shifts and J-couplings in NA phosphate were statistically inconclusive. In contrary to the static results, none of the dynamically averaged NMR parameters distinguished between the two Mg^{2+} coordination sites. Considering the two limit cases in the study V—the static and dynamic phosphate solvation—we conclude that the solvent dynamics has a substantial impact on ^{31}P NMR parameters in NA phosphate.

4.2.4 Magnesium ion imprints in Raman scattering of water

The study IV correlated Raman spectra measured in magnesium salt solutions with those calculated with several methods for Mg^{2+}–water clusters.

Experimental Raman spectra of water and aqueous solutions of sodium, magnesium, and calcium chloride were recorded at 20 °C and a rather high 5 mol/kg concentration in the $200 - 2000$ cm^{-1} region. All salts notably increased the water Raman signal in the $200-900$ cm^{-1} region corresponding to water librations, rotations, and hydrogen bonding (work IV, Figure 1). The frequency and intensity of water bending signal at 1640 cm^{-1} also slightly increased, while the bandwidth decreased. In the $MgCl_2$ spectrum, a rather sharp peak appeared at 355 cm^{-1} and it was present also in the Raman spectra of $MgBr_2$ solution. The peak was thus assigned to specific Mg^{2+} interaction with water. The absence of such distinct feature in the low-frequency region of Raman spectra of NaCl and $CaCl_2$ solutions confirmed that the water interaction with Ca^{2+} and Na^+ is much weaker. The difference between the $MgCl_2$ and $MgBr_2$ spectra was quite small, indicating that the role of Mg^{2+} is dominant and the anions cause only minor spectral changes. In accord with current perspective [53, 68], there were no indications of direct Mg−Cl contacts even at high $MgCl_2$ concentrations. Raman spectra were also measured in D_2O solution of $MgCl_2$, where the Mg−O peak shifted to 337 cm^{-1}, i.e. by the mass factor $(M_{\mathrm{H2O}}/M_{\mathrm{D2O}})^{1/2}$, confirming thus the spectral assignment.

The computational part of work IV employed two different theoretical approaches: (i) static DFT calculations of geometry parameters, orbital properties, and Raman spectra of $[M(H_2O)_6]$ clusters with T_h symmetry (Figure 4.3a), $M = Mg^{2+}, Ca^{2+}, Na^+$, (ii) Raman spectra generation from Car–Parrinello molecular dynamics (CPMD [170]) simulations of pure water and $MgCl_2$ aqueous solution. The dynamical Raman spectra modeling was done either by using a modified Fourier-transform method or by averaging the Raman spectra calculated for $[Mg(H_2O)_6]^{2+}$ clusters extracted from random CPMD snapshot geometries. The latter method turned out clearly superior and the modified Fourier-transform technique is thus not discussed further. Note that classical MD employing empirical parameterization of intermolecular forces cannot be a priori expected to correctly simulate the dynamical behavior of Mg^{2+} microhydration and more reliable first-principles methods such as CPMD are thus preferential. Nevertheless, it was shown recently that for example the interfacial properties of aqueous salt solutions can be successfully modeled with a specifically adjusted MD force-field [68, 70, 71].

The static cluster calculations helped to elucidate the differences between Mg^{2+} and other cations and served also for the calibration of computational methods (work IV, Table 2). The Mg−O distance 2.084 Å optimized with the B3LYP/6-311++G(d,p) method and the CPCM hydration was in good agreement with the average Mg−O distance 2.063 Å found in crystals [61]. In the CPMD simulations discussed below, the first two maxima of

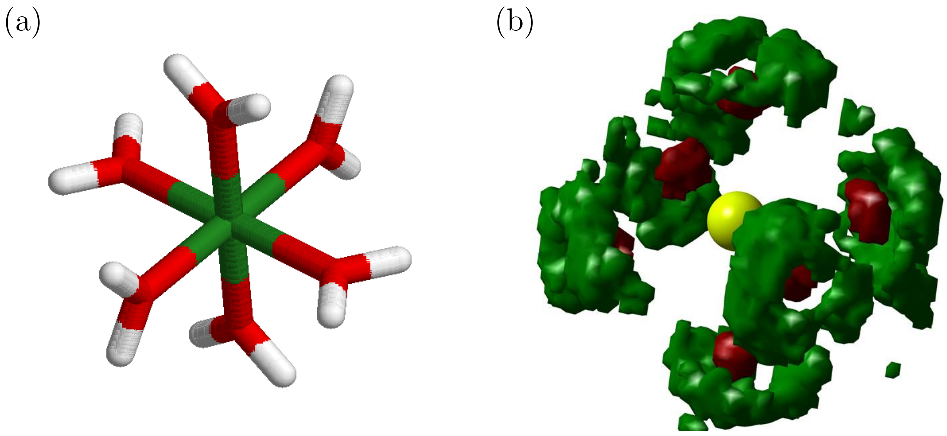


Figure 4.3 (a) The $[Mg(H_2O)_6]^{2+}$ cluster with T_h symmetry. (b) Calculated CPMD distribution of the water oxygen (red) and hydrogen (green) atoms in the first hydration shell of the Mg^{2+} ion (yellow sphere).

the Mg−O radial distribution function were at 2.1 Å and 4.2 Å (work IV, Figure 2), which agreed with previous CPMD studies [65] as well as with the average crystallographic values 2.063 Å and 4.312 Å [61]. The calculated Mg−O distance was by about 0.3 Å shorter than the Na−O and Ca−O distances, which corresponds with much tighter arrangement of water molecules in the Mg^{2+} first hydration shell. NBO analysis revealed that the strength of the Mg^{2+}–water interaction could be explained by a partially covalent character of the Mg−O bond (work IV, Figure 4 and Table 6). Natural charge of Mg^{2+} in the $[Mg(H_2O)_6]^{2+}$ complex amounts only 1.819, i.e. 91 % of its nominal charge. The remaining 9 % is compensated by electron density donated from oxygen lone pairs; this value is much smaller for Ca^{2+} and Na^+ (∼7 % both).

There are twenty-one unique normal modes in the T_h-symmetrized $[Mg(H_2O)_6]^{2+}$ cluster, of which lowest sixteen were analyzed in the work IV (omitting the O−H vibrations). The $[Mg(H_2O)_6]^{2+}$ vibrational frequencies calculated with the B3LYP and MP2 methods were quite close to the experimental values, indicating that a very good accuracy was achieved already at the DFT level (work IV, Table 3). The inclusion of higher hydration layers via PCM and CPCM provided somewhat inconsistent frequency changes; obviously, the calculations of low-frequency modes in a charged system pose a specific challenge for the numerical stability of continuum solvation models. Nevertheless, the frequency of the symmetrical Mg−O breathing mode, which clearly matches to the experimental Raman peak at 355 cm^{-1}, improved from 314 cm^{-1} in gas-phase to 337 (335) cm^{-1} with PCM (CPCM). Only eight of the sixteen lowest normal modes in $[Mg(H_2O)_6]^{2+}$ are Raman active. Their calculated Raman intensities were most sensitive to the quality of electronic basis set and to the effect of higher hydration shells modeled with PCM; the difference between B3LYP and MP2 methods was only minor (work IV, Table 4).

The static DFT calculations also shed light on the differences in low-frequency Raman spectra among different salt solutions (work IV, Table 5). Raman intensity of the symmetrical breathing mode calculated for the $[Mg(H_2O)_6]^{2+}$ cluster decreased by a factor of $\sim$4 in the case of $[Ca(H_2O)_6]^{2+}$ and it was almost zero for $[Na(H_2O)_6]^{+}$. At the same time, the calculated metal–O breathing frequency decreased in a series $Mg^{2+} > Ca^{2+} > Na^{+}$, consistently with the decreasing ion–water affinity [52]. The exceptionally high Mg–O stretching frequency is thus coherent with the strength of the Mg^{2+}–water interaction and the tight arrangement of the Mg^{2+} first hydration shell enables large polarizability changes along the Mg–O stretching, causing large Raman intensity of the symmetric breathing mode.

The combination of Raman spectra calculations with CPMD simulations proved to be a very good method for modeling the low-frequency Raman spectra of aqueous salt solutions. Similar to the methodology used in the work V, the CPMD snapshot geometries of the $[Mg(H_2O)_6]^{2+}$ clusters were partially relaxed by normal-mode optimization [171] of the lowest-frequency modes (< 200 cm^{-1}). Vibrational frequencies and Raman intensities were then calculated for the $[Mg(H_2O)_6]^{2+}$ clusters and the resulting Raman spectra were averaged. Moreover, by substituting the hydrogen mass with the deuterium mass and repeating the force field diagonalization, the Raman spectra of $[Mg(D_2O)_6]^{2+}$ clusters were also obtained. The averaged Raman spectra of the $[Mg(H_2O)_6]^{2+}$ and $[Mg(D_2O)_6]^{2+}$ clusters were in excellent agreement with the experimental Raman spectra of the H_2O and D_2O solutions of $MgCl_2$ (work IV, Figure 7). Not only the calculated Mg–O specific peak at 318 (307) cm^{-1} and the water bending band at 1645 (1200) cm^{-1} nicely corresponded with the experimental Raman bands, but even the broad band of water librations and rotations near 605 (495) cm^{-1} was successfully modeled around 560 (440) cm^{-1} for the H_2O (D_2O) solution. The CPMD simulations also included the chlorine anions, what allowed us to investigate their influence on the Raman spectra. We found that hydrated Cl^- clusters do not exhibit Raman signal that would interfere with the low-frequency features arising due to Mg^{2+}.

5 Summary

The aims of this doctoral thesis were successfully achieved. The most important results are summarized as follows:

- Potential energy surfaces of L-alanyl-L-alanine in dependence on the backbone torsion angles ψ_1 and φ_2 were thoroughly characterized by DFT calculations for all its charged forms AA^+, AA^{zw}, AA^-. The potential energy surfaces for AA^+ and AA^{zw} were both similar with a pronounced global minimum, implying that a single conformer is dominantly populated at the room temperature in solutions with low to neutral pH. The conformational behavior of AA^- significantly differed from the other two forms owing to the intramolecular interactions of the -NH_2 group. According to the calculations, several low-lying conformers of AA^- are populated at the room temperature and a hindered rotation about the ψ_1 angle occurs.

- NMR parameters calculated for the lowest-energy AA conformers were overall in better agreement with the NMR experiment than those calculated for the local minima with higher energy, what confirmed our theoretical assignment of the most populated AA conformers in solution. On the basis of DFT/PCM calculations we thus rationalized the experimental pH-induced changes of the NMR parameters in the AA di-peptide.

- We confirmed that the experimentally probed cross-correlated relaxation rates N–$C_{\alpha1}H_{\alpha1}$, C'–$C_{\alpha1}H_{\alpha1}$, N–$C_{\alpha2}H_{\alpha2}$, and C'–$C_{\alpha2}H_{\alpha2}$ are the most convenient CCR rates that can be used for the determination of peptide backbone geometry. The CCR mechanisms N–$NH_{\alpha2}$, C'–$NH_{\alpha2}$, N–$C'H_{\alpha1}$, and C'–$C'H_{\alpha1}$ were suggested as additional restraints for the NMR determination of peptide structure.

- Dependencies of ^{31}P NMR parameters on the nucleic acid phosphate geometry were systematically studied with DFT calculations. The geometry dependencies of the isotropic ^{31}P chemical shift and the principal components of the ^{31}P chemical shielding tensor were calibrated. Validity of the rigid-tensor approximation for the interpretation of $\Gamma_{P,CH}$ CCR rates was confirmed. New qualitative rules for the structural interpretation of the $^2J_{PC}$ spin–spin coupling constants were proposed.

- Response of the ^{31}P NMR parameters in nucleic acid phosphate to the coordination of Mg^{2+} was modeled using a combined molecular dynamics/DFT approach. The calculated decrease of the isotropic ^{31}P chemical shift upon Mg^{2+} coordination by $5-6$ ppm was in agreement with the available experimental data. A statistically significant increase of the ^{31}P chemical shielding anisotropy and a decrease of the $^2J_{PC5'}$ coupling magnitude upon the Mg^{2+} coordination were also calculated.

- Differences in the low-frequency Raman spectra measured for aqueous solutions of NaCl, $MgCl_2$, and $CaCl_2$ were explained by the Raman spectra calculations for the $[M(H_2O)_6]$ clusters, $M = Na^{+}, Mg^{2+}, Ca^{2+}$. The well-known tight coordination of water solvent to Mg^{2+} was rationalized in terms of bond distances, metal−water vibrational frequencies, and natural charges calculated for the $[Mg(H_2O)_6]^{2+}$ cluster. The theoretical approach combining the *ab initio* molecular dynamics with the DFT calculations for MD clusters was established as a reliable method for modeling the spectroscopic response of aqueous salt solutions.

List of abbreviations

AA	L-alanyl-L-alanine
AO	atomic orbital
B3LYP	Becke's 1993 three-parameter hybrid functional with Lee–Yang–Parr correlation functional
BLYP	Becke's 1988 functional with Lee–Yang–Parr correlation functional
BPW91	Becke's 1988 functional with Perdew–Wang correlation functional
CCR	cross-correlated relaxation
CCSD	coupled clusters singles and doubles
CP	coupled-perturbed
CPHF	coupled-perturbed Hartree–Fock
CPKS	coupled-perturbed Kohn–Sham
CPMD	Car–Parrinello molecular dynamics
CSA	chemical shielding anisotropy
DD	dipole–dipole
DFT	density functional theory
DMP	dimethyl phosphate
DNA	deoxyribonucleic acid
DSO	diamagnetic spin–orbit (operator)
ECP	effective core potential
EMP	ethyl methyl phosphate
FC	Fermi contact (operator)
GGA	generalized gradient approximation
GIAO	gauge including atomic orbitals
GTO	Gaussian type orbital
HOMO	highest occupied molecular orbital
IGLO	individual gauge for localized orbitals
IR	infrared
HF	Hartree–Fock
KS	Kohn–Sham
LCAO	linear combination of atomic orbitals
LSDA	local spin density approximation
LUMO	lowest unoccupied molecular orbital
MD	molecular dynamics
MO	molecular orbital
NA	nucleic acid
NBO	natural bond orbital

NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
nPn	abasic dinucleotide monophosphate
PCM	polarizable continuum model
PES	potential energy surface
PSO	paramagnetic spin–orbit (operator)
RDC	residual dipolar coupling
RNA	ribonucleic acid
SCF	self-consistent field
SD	spin–dipole (operator)
SEM	standard error of the mean
SOS	sum over states

References

[1] Reif, B., Hennig, M., and Griesinger, C. *Science* **276**, 1230–1233 (1997).

[2] Reif, B., Diener, A., Hennig, M., Maurer, M., and Griesinger, C. *J. Magn. Reson.* **143**, 45–68 (2000).

[3] Schwalbe, H., Carlomagno, T., Hennig, M., Junker, J., Reif, B., Richter, C., and Griesinger, C. *Method. Enzymol.* **338**, 35–81 (2001).

[4] Bouř, P., Kapitán, J., and Baumruk, V. *J. Phys. Chem. A* **105**, 6362–6368 (2001).

[5] Bouř, P., Buděšínský, M., Špirko, V., Kapitán, J., Šebestík, J., and Sychrovský, V. *J. Am. Chem. Soc.* **127**, 17079–17089 (2005).

[6] Šebek, J., Gyurcsik, B., Šebestík, J., Kejík, Z., Bednárová, L., and Bouř, P. *J. Phys. Chem. A* **111**, 2750–2760 (2007).

[7] Bouř, P., Sychrovský, V., Maloň, P., Hanzlíková, J., Baumruk, V., Pospíšek, J., and Buděšínský, M. *J. Phys. Chem. A* **106**, 7321–7327 (2002).

[8] Buděšínský, M., Daněček, P., Bednárová, L., Kapitán, J., Baumruk, V., and Bouř, P. *J. Phys. Chem. A* **112**, 8633–8640 (2008).

[9] Kapitán, J., Baumruk, V., Kopecký, V., and Bouř, P. *J. Phys. Chem. A* **110**, 4689–4696 (2006).

[10] Kapitán, J., Baumruk, V., Kopecký, V., Pohl, R., and Bouř, P. *J. Am. Chem. Soc.* **128**, 13451–13462 (2006).

[11] Daněček, P., Kapitán, J., Baumruk, V., Bednárová, L., Kopecký, V., and Bouř, P. *J. Chem. Phys.* **126**, 224513 (2007).

[12] Dračínský, M., Kaminský, J., and Bouř, P. *J. Chem. Phys.* **130**, 094106 (2009).

[13] Dračínský, M., Kaminský, J., and Bouř, P. *J. Phys. Chem. B* **113**, 14698–14707 (2009).

[14] Dračínský, M. and Bouř, P. *J. Chem. Theory Comput.* **6**, 288–299 (2010).

[15] Varani, G., Aboul-Ela, F., and Allain, F. H.-T. *Prog. Nucl. Magn. Res. Spectrosc.* **29**, 51–127 (1996).

[16] Wijmenga, S. S. and van Buuren, B. N. M. *Prog. Nucl. Magn. Res. Spectrosc.* **32**, 287–387 (1998).

[17] Fürtig, B., Richter, C., Wöhnert, J., and Schwalbe, H. *ChemBioChem* **4**, 936–962 (2003).

[18] Flinders, J. and Dieckmann, T. *Prog. Nucl. Magn. Res. Spectrosc.* **48**, 137–159 (2006).

[19] Zhou, D.-M. and Taira, K. *Chem. Rev.* **98**, 991–1026 (1998).

[20] Kuimelis, R. G. and McLaughlin, L. W. *Chem. Rev.* **98**, 1027–1044 (1998).

[21] Richardson, J. S., Schneider, B., Murray, L. W., Kapral, G. J., Immormino, R. M., Headd, J. J., Richardson, D. C., Ham, D., Hershkovits, E., Williams, L. D., Keating, K. S., Pyle, A. M., Micallef, D., Westbrook, J., and Berman, H. M. *RNA* **14**, 465–481 (2008).

[22] Sychrovský, V., Vokáčová, Z., Šponer, J., Špačková, N., and Schneider, B. *J. Phys. Chem. B* **110**, 22894–22902 (2006).

[23] Gorenstein, D. G., editor. *Phosphorus-31 NMR.* Academic press, Inc., Orlando, Florida, (1984).

[24] Gorenstein, D. G. *Chem. Rev.* **94**, 1315–1338 (1994).

[25] Lankhorst, P. P., Haasnoot, C. A. G., Erkelens, C., Westerink, H. P., van der Marel, G. A., van Boom, J. H., and Altona, C. *Nucleic Acids Res.* **13**, 927–942 (1985).

[26] Richter, C., Reif, B., Griesinger, C., and Schwalbe, H. *J. Am. Chem. Soc.* **122**, 12728–12731 (2000).

[27] Nozinovic, S., Richter, C., Fürtig, B., Duchardt-Ferner, E., Weigand, J. E., and Schwalbe, H. *J. Am. Chem. Soc.* **132**, 10318–10329 (2010).

[28] Nozinovic, S., Fürtig, B., Jonker, H. R. A., Richter, C., and Schwalbe, H. *Nucleic Acids Res.* **38**, 683–694 (2010).

[29] Gorenstein, D. G. and Luxon, B. A. *Biochemistry* **18**, 3796–3804 (1979).

[30] Přecechtělová, J., Munzarová, M. L., Novák, P., and Sklenář, V. *J. Phys. Chem. B* **111**, 2658–2667 (2007).

[31] Přecechtělová, J., Padrta, P., Munzarová, M. L., and Sklenář, V. *J. Phys. Chem. B* **112**, 3470–3478 (2008).

[32] Přecechtělová, J., Novák, P., Munzarová, M. L., Kaupp, M., and Sklenář, V. *J. Am. Chem. Soc.* **132**, 17139–17148 (2010).

[33] Blommers, M. J. J., Haasnoot, C. A. G., Walters, J., van der Marel, G. A., Vanboom, J. H., and Hilbers, C. W. *Biochemistry* **27**, 8361–8369 (1988).

[34] Mooren, M. M. W., Wijmenga, S. S., van der Marel, G. A., van Boom, J. H., and Hilbers, C. W. *Nucleic Acids Res.* **22**, 2658–2666 (1994).

[35] Zimmer, D. P., Marino, J. P., and Griesinger, C. *Magn. Reson. Chem.* **34**, S177–S186 (1996).

[36] Richter, C., Reif, B., Wörner, K., Quant, S., Marino, J. P., Engels, J. W., Griesinger, C., and Schwalbe, H. *J. Biomol. NMR* **12**, 223–230 (1998).

[37] O'Neil-Cabello, E., Wu, Z., Bryce, D. L., Nikonowicz, E. P., and Bax, A. *J. Biomol. NMR* **30**, 61–70 (2004).

[38] Watson, J. D. and Crick, F. H. C. *Nature* **171**, 737–738 (1953).

[39] Draper, D. E. *RNA* **10**, 335–343 (2004).

[40] Draper, D. E., Grilley, D., and Soto, A. M. *Annu. Rev. Bioph. Biom.* **34**, 221–243 (2005).

[41] Freisinger, E. and Sigel, R. K. O. *Coord. Chem. Rev.* **251**, 1834–1851 (2007).

[42] Sreedhara, A. and Cowan, J. A. *BioMetals* **15**, 211–223 (2002).

[43] Schneider, B. and Kabeláč, M. *J. Am. Chem. Soc.* **120**, 161–165 (1998).

[44] Schneider, B., Cohen, D. M., Schleifer, L., Srinivasan, A. R., Olson, W. K., and Berman, H. M. *Biophys. J.* **65**, 2291–2303 (1993).

[45] Schneider, B., Patel, K., and Berman, H. M. *Biophys. J.* **75**, 2422–2434 (1998).

[46] Suzumura, K.-i., Yoshinari, K., Tanaka, Y., Takagi, Y., Kasai, Y., Warashina, M., Kuwabara, T., Orita, M., and Taira, K. *J. Am. Chem. Soc.* **124**, 8230–8236 (2002).

[47] Suzumura, K.-i., Takagi, Y., Orita, M., and Taira, K. *J. Am. Chem. Soc.* **126**, 15504–15511 (2004).

[48] Osborne, E. M., Ward, W. L., Ruehle, M. Z., and DeRose, V. J. *Biochemistry* **48**, 10654–10664 (2009).

[49] Tanaka, Y., Morita, E. H., Hayashi, H., Kasai, Y., Tanaka, T., and Taira, K. *J. Am. Chem. Soc.* **122**, 11303–11310 (2000).

[50] Cowan, J. A. *BioMetals* **15**, 225–35 (2002).

[51] Maguire, M. E. and Cowan, J. A. *BioMetals* **15**, 203–210 (2002).

[52] Collins, K. D. *Biophys. Chem.* **119**, 271–281 (2006).

[53] Sugimoto, K., Dinnebier, R. E., and Hanson, J. C. *Acta Crystallogr. Sect. B* **63**, 235–242 (2007).

[54] Marcus, Y. *Chem. Rev.* **109**, 1346–1370 (2009).

[55] Pye, C. C. and Rudolph, W. W. *J. Phys. Chem. A* **102**, 9933–9943 (1998).

[56] Rudolph, W. W., Irmer, G., and Hefter, G. T. *Phys. Chem. Chem. Phys.* **5**, 5253–5261 (2003).

[57] Pavlov, M., Siegbahn, P. E. M., and Sandström, M. *J. Phys. Chem. A* **102**, 219–228 (1998).

[58] Ramaniah, L. M., Bernasconi, M., and Parrinello, M. *J. Chem. Phys.* **109**, 6839–6843 (1998).

[59] Ramaniah, L. M., Bernasconi, M., and Parrinello, M. *J. Chem. Phys.* **111**, 1587–1591 (1999).

[60] Tobias, D. J., Jungwirth, P., and Parrinello, M. *J. Chem. Phys.* **114**, 7036–7044 (2001).

[61] Markham, G. D., Glusker, J. P., and Bock, C. W. *J. Phys. Chem. B* **106**, 5118–5134 (2002).

[62] Adrian-Scotto, M., Mallet, G., and Vasilescu, D. *J. Mol. Struct. THEOCHEM* **728**, 231–242 (2005).

[63] Krekeler, C., Hess, B., and Delle Site, L. *J. Chem. Phys.* **125**, 054305 (2006).

[64] Ikeda, T., Boero, M., and Terakura, K. *J. Chem. Phys.* **126**, 034501 (2007).

[65] Ikeda, T., Boero, M., and Terakura, K. *J. Chem. Phys.* **127**, 074503 (2007).

[66] Mink, J., Németh, C., Hajba, L., Sandström, M., and Goggin, P. L. *J. Mol. Struct.* **661–662**, 141–151 (2003).

[67] Kamarchik, E. and Bowman, J. M. *J. Phys. Chem. A* **114**, 12945–12951 (2010).

[68] Callahan, K. M., Casillas-Ituarte, N. N., Roeselová, M., Allen, H. C., and Tobias, D. J. *J. Phys. Chem. A* **114**, 5141–5148 (2010).

[69] Jungwirth, P. and Tobias, D. J. *Chem. Rev.* **106**, 1259–1281 (2006).

[70] Callahan, K. M., Casillas-Ituarte, N. N., Xu, M., Roeselová, M., Allen, H. C., and Tobias, D. J. *J. Phys. Chem. A* **114**, 8359–8368 (2010).

[71] Casillas-Ituarte, N. N., Callahan, K. M., Tang, C. Y., Chen, X., Roeselová, M., Tobias, D. J., and Allen, H. C. *Proc. Natl. Acad. Sci. USA* **107**, 6616–6621 (2010).

[72] Bernasconi, L., Baerends, E. J., and Sprik, M. *J. Phys. Chem. B* **110**, 11444–11453 (2006).

[73] Dirac, P. A. M. *Proc. R. Soc. Lond. A* **123**, 714–733 (1929).

[74] Born, M. and Oppenheimer, J. R. *Ann. Phys.* **389**, 457–484 (1927).

[75] Szabo, A. and Ostlund, N. S. *Modern Quantum Chemistry.* McGraw-Hill, (1989).

[76] Salotto, A. W. and Burnelle, L. *J. Chem. Phys.* **52**, 2936–2945 (1970).

[77] Guest, M. F., Saunders, V. R., and Overill, R. E. *Mol. Phys.* **35**, 427–443 (1978).

[78] Roothaan, C. C. J. *Rev. Mod. Phys.* **23**, 69–89 (1951).

[79] Hall, G. G. *Proc. R. Soc. Lond. A* **205**, 541–552 (1951).

[80] Skála, L. *Kvantová teorie molekul.* Karolinum, Praha, (1995).

[81] Jensen, F. *Introduction to Computational Chemistry.* John Wiley & Sons Ltd, second edition, (2007).

[82] Widmark, P.-O., editor. *European Summerschool in Quantum Chemistry.* Lund University, fifth edition, (2007).

[83] Parr, R. G. and Yang, W. *Density-functional theory of atoms and molecules.* Oxford University Press, Inc., (1989).

[84] Kohn, W., Becke, A. D., and Parr, R. G. *J. Phys. Chem.* **100**, 12974–12980 (1996).

[85] Koch, W. and Holthausen, M. C. *A Chemist's Guide to Density Functional Theory.* Wiley-VCH Verlag GmbH, second edition, (2001).

[86] Burke, K. and friends. *The ABC of DFT.* University of California, Irvine, (2007). Available online: `http://dft.uci.edu/sites/default/files/g1.pdf` [cit. 2012-04-13].

[87] Thomas, L. H. *Proc. Cambridge Phil. Soc.* **23**, 542–548 (1927).

[88] Fermi, E. *Z. Physik* **48**, 73–79 (1928).

[89] Dirac, P. A. M. *Proc. Cambridge Phil. Soc.* **26**, 376–385 (1930).

[90] Hohenberg, P. and Kohn, W. *Phys. Rev.* **136**, B864–B871 (1964).

[91] Kohn, W. and Sham, L. J. *Phys. Rev.* **140**, A1133–A1138 (1965).

[92] Perdew, J. P. and Ruzsinszky, A. *Int. J. Quantum Chem.* **110**, 2801–2807 (2010).

[93] Slater, J. C. *Phys. Rev.* **81**, 385–390 (1951).

[94] Vosko, S. J., Wilk, L., and Nusair, M. *Can. J. Phys.* **58**, 1200–1211 (1980).

[95] Perdew, J. P. and Wang, Y. *Phys. Rev. B* **33**, 8800–8802 (1986).

[96] Becke, A. D. *Phys. Rev. A* **38**, 3098–3100 (1988).

[97] Perdew, J. P. In *Electronic Structure of Solids,* Ziesche, P. and Eschrig, H., editors, volume 17, 11–20. Akademie Verlag, Berlin, (1991).

[98] Perdew, J. P., Chevary, J. A., Vosko, S. H., Jackson, K. A., Pederson, M. R., Singh, D. J., and Fiolhais, C. *Phys. Rev. B* **46**, 6671–6687 (1992).

[99] Perdew, J. P., Burke, K., and Ernzerhof, M. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

[100] Perdew, J. P. *Phys. Rev. B* **33**, 8822–8824 (1986).

[101] Lee, C., Yang, W., and Parr, R. G. *Phys. Rev. B* **37**, 785–789 (1988).

[102] Becke, A. D. *J. Chem. Phys.* **98**, 1372–1377 (1993).

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

[104] Stephens, P. J., Devlin, F. J., Chabalowski, C. F., and Frisch, M. J. *J. Phys. Chem.* **98**, 11623–11627 (1994).

[105] Gaussian 09 User's Reference. Available online: `http://www.gaussian.com/g_tech/g_ur/g09help.htm` [cit. 2012-04-13].

[106] Pople, J. A., Gill, P. M. W., and Johnson, B. G. *Chem. Phys. Lett.* **199**, 557–560 (1992).

[107] Perdew, J. P., Ruzsinszky, A., Constantin, L. A., Sun, J., and Csonka, G. I. *J. Chem. Theory Comput.* **5**, 902–908 (2009).

[108] Gauss, J. In *Modern Methods and Algorithms of Quantum Chemistry,* Grotendorst, J., editor, volume 3 of *NIC series*, 541–592. John von Neumann Institute for Computing, Jülich, second edition, (2000).

[109] Pulay, P. *Mol. Phys.* **17**, 197–204 (1969).

[110] Feynman, R. P. *Phys. Rev.* **56**, 340–343 (1939).

[111] Schrödinger, E. *Ann. Phys.* **385**, 437–490 (1926).

[112] Löwdin, P.-O. *J. Math. Phys.* **6**, 1341–1353 (1965).

[113] Stevens, R. M., Pitzer, R. M., and Lipscomb, W. N. *J. Chem. Phys.* **38**, 550–560 (1963).

[114] Fukui, H. *Prog. Nucl. Magn. Res. Spectrosc.* **1**, 317–342 (1997).

[115] Fukui, H. *Prog. Nucl. Magn. Res. Spectrosc.* **35**, 267–294 (1999).

[116] Johnson, B. G. and Frisch, M. J. *J. Chem. Phys.* **100**, 7429–7442 (1994).

[117] Cheeseman, J. R., Trucks, G. W., Keith, T. A., and Frisch, M. J. *J. Chem. Phys.* **104**, 5497–5509 (1996).

[118] Sychrovský, V., Gräfenstein, J., and Cremer, D. *J. Chem. Phys.* **113**, 3530–3547 (2000).

[119] Dickson, R. M. and Ziegler, T. *J. Phys. Chem.* **100**, 5286–5290 (1996).

[120] Cances, E., Mennucci, B., and Tomasi, J. *J. Chem. Phys.* **107**, 3032–3041 (1997).

[121] Mennucci, B., Cances, E., and Tomasi, J. *The J. Phys. Chem. B* **101**, 10506–10517 (1997).

[122] Scalmani, G. and Frisch, M. J. *J. Chem. Phys.* **132**, 114110 (2010).

[123] Barone, V. and Cossi, M. *J. Phys. Chem. A* **102**, 1995–2001 (1998).

[124] Helgaker, T., Watson, M. A., and Handy, N. C. *J. Chem. Phys.* **113**, 9402–9409 (2000).

[125] Kutzelnigg, W., Fleischer, U., and Schindler, M. *NMR, Basic Principles and Progress*, volume 23. Springer, Heidelberg (1990).

[126] Daněček, P. and Bouř, P. *J. Comput. Chem.* **28**, 1617–1624 (2007).

[127] Dračínský, M. and Bouř, P. *J. Comput. Chem.* **33**, 1080–1089 (2012).

[128] Ramsey, N. F. *Phys. Rev.* **78**, 699–703 (1950).

[129] van Wüllen, C. *Phys. Chem. Chem. Phys.* **2**, 2137–2144 (2000).

[130] Jameson, C. J., de Dios, A. C., and Jameson, A. K. *Chem. Phys. Lett.* **167**, 575–582 (1990).

[131] Bloch, F. *Phys. Rev.* **70**, 460–474 (1946).

[132] Solomon, I. *Phys. Rev.* **99**, 559–565 (1955).

[133] Wangsness, R. K. and Bloch, F. *Phys. Rev.* **89**, 728–739 (1953).

[134] Redfield, A. G. *Adv. Magn. Reson.* **1**, 1–32 (1965).

[135] Kowalewski, J. and Mäler, L. *Nuclear spin relaxation in liquids: Theory, experiments, and applications.* Taylor & Francis group, New York, London, (2006).

[136] Kumar, A., Grace, R. C. R., and Madhu, P. K. *Prog. Nucl. Magn. Res. Spectrosc.* **37**, 191–319 (2000).

[137] Karplus, M. *J. Am. Chem. Soc.* **85**, 2870–2871 (1963).

[138] Czernek, J. and Bruschweiler, R. *J. Am. Chem. Soc.* **123**, 11079–11080 (2001).

[139] Dirac, P. A. M. *Proc. R. Soc. Lond. A* **117**, 610–624 (1928).

[140] Dirac, P. A. M. *Proc. R. Soc. Lond. A* **118**, 351–361 (1928).

[141] Ramsey, N. F. *Phys. Rev.* **91**, 303–307 (1953).

[142] Kutzelnigg, W. *Theor. Chim. Acta* **73**, 173–200 (1988).

[143] Griffiths, D. J. *Am. J. Phys.* **50**, 698–703 (1982).

[144] Kutzelnigg, W. chapter Fundamentals of Nonrelativistic and Relativistic Theory of NMR and EPR Parameters, 43–82. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim (2004).

[145] Ramsey, N. F. *Phys. Rev.* **86**, 243–246 (1952).

[146] Lee, A. M., Handy, N. C., and Colwell, S. M. *J. Chem. Phys.* **103**, 10095–10109 (1995).

[147] Malkin, V. G., Malkina, O. L., and Salahub, D. R. *Chem. Phys. Lett.* **221**, 91–99 (1994).

[148] Malkina, O. L., Salahub, D. R., and Malkin, V. G. *J. Chem. Phys.* **105**, 8793–8800 (1996).

[149] Bouř, P. and Buděšínský, M. *J. Chem. Phys.* **110**, 2–9 (1999).

[150] Auer, A. A. and Gauss, J. *Chem. Phys.* **356**, 7–13 (2009).

[151] Vahtras, O., Aagren, H., Jørgensen, P., Jensen, H. J. A., Padkjaer, S. B., and Helgaker, T. *J. Chem. Phys.* **96**, 6120–6125 (1992).

[152] Salter, E. A., Sekino, H., and Bartlett, R. J. *J. Chem. Phys.* **87**, 502–509 (1987).

[153] Perera, S. A., Nooijen, M., and Bartlett, R. J. *J. Chem. Phys.* **104**, 3290–3305 (1996).

[154] de Dios, A. C. *Prog. Nucl. Magn. Res. Spectrosc.* **29**, 229–278 (1996).

[155] Ditchfield, R. *Mol. Phys.* **27**, 789–807 (1974).

[156] Wolinski, K., Hinton, J. F., and Pulay, P. *J. Am. Chem. Soc.* **112**, 8251–8260 (1990).

[157] London, F. *J. Phys. Radium* **8**, 397 (1937).

[158] Malkin, V. G., Malkina, O. L., and Salahub, D. R. *Chem. Phys. Lett.* **204**, 80–86 (1993).

[159] Marynick, D. S., Ray, A. K., Fry, J. L., and Kleier, D. A. *J. Mol. Struct. THEOCHEM* **108**, 45–48 (1984).

[160] Demarco, A., Llinás, M., and Wüthrich, K. *Biopolymers* **17**, 2727–2742 (1978).

[161] Wang, A. C. and Bax, A. *J. Am. Chem. Soc.* **117**, 1810–1813 (1995).

[162] Yang, D., Konrat, R., and Kay, L. E. *J. Am. Chem. Soc.* **119**, 11938–11940 (1997).

[163] Hong, M., Gross, J. D., Hu, W., and Griffin, R. G. *J. Magn. Reson.* **135**, 169–177 (1998).

[164] Chiarparin, E., Pelupessy, P., Ghose, R., and Bodenhausen, G. *J. Am. Chem. Soc.* **121**, 6876–6883 (1999).

[165] Jensen, F. *J. Chem. Theory Comput.* **2**, 1360–1369 (2006).

[166] Herzfeld, J., Griffin, R. G., and Haberkorn, R. A. *Biochemistry* **17**, 2711–2718 (1978).

[167] Kyogoku, Y. and Itaka, Y. *Acta Crystallogr.* **21**, 49–57 (1966).

[168] Lankhorst, P. P., Haasnoot, C. A. G., Erkelens, C., and Altona, C. *J. Biomol. Struct. Dyn.* **1**, 1387–1405 (1984).

[169] Plavec, J. and Chattopadhyaya, J. *Tetrahedron Lett.* **36**, 1949–1952 (1995).

[170] Car, R. and Parrinello, M. *Phys. Rev. Lett.* **55**, 2471–2474 (1985).

[171] Bouř, P. and Keiderling, T. A. *J. Chem. Phys.* **117**, 4126–4132 (2002).

Published works

WDS'07 Proceedings of Contributed Papers, Part III, 84–89, 2007. *ISBN 978-80-7378-025-8 © MATFYZPRESS*

Calculation of the Effective Chemical Shielding Anisotropy in L-Alanyl-L-Alanine, Conformational and Charge Dependence Study

L. Benda

Faculty of Mathematics and Physics, Charles University Prague, Ke Karlovu 3, 121 16 Praha 2, Czech Republic.

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic,
Flemingovo nám. 2, 166 10 Praha 6, Czech Republic.

Abstract. DFT quantum-chemical calculations were carried out for the cationic, zwitterionic and anionic form of the L-Alanyl-L-Alanine di-peptide that can be experimentally accessed at different pH. The ^{15}N NMR chemical shielding tensor of the amide nitrogen was used for theoretical modeling of the cross-correlated relaxations. In particular, effective chemical shielding anisotropies (effective CSA, $\Delta\sigma^{\mathrm{eff}}$) were calculated in dependence on two major structural descriptors of the peptide backbone, the torsion angles φ and ψ. Proposed calibration of Karplus-like curves can be used for structural interpretation of the cross-correlated relaxation rates in peptides. The variation of modeled effective CSA's upon pH was discussed.

Introduction

Nuclear Magnetic Resonance (NMR) spectroscopy is a non-invasive experimental technique that provides unique information on the local magnetic field perturbations at the site of atomic nucleus. This spectroscopy method is based on manipulating the nuclear magnetization with external magnetic fields and probing the time evolution of the resulting magnetization. Detailed analysis of NMR spectra can thus unveil essential local properties of the examined system.

In the past decades, the high-resolution NMR spectroscopy has become an indispensable tool in biomolecular research. Variety of different NMR applications in this field has grown to an unprecedented extent [*Levitt*, 2001]. However, the accurate determination of molecular structure with NMR spectroscopy still remains challenging. Currently, the fundamental techniques used for the structural refinement are measurement of the NOE (Nuclear Overhauser Effect) that reflects the through-space distance between protons, and determination of the scalar spin-spin coupling constants (*J*-couplings) that can be used for the assignment of torsion angles with Karplus curves [*Karplus*, 1963]. However, the structural information gained in this way is often ambiguous, insufficient or inaccurate. Additional geometry restraints are therefore needed in NMR structural studies.

The cross-correlated relaxation is a specific NMR phenomenon that might be exploited to obtain further geometry restrains provided that its structural behavior has been calibrated [*Reif et al.*, 1997; *Reif et al.*, 2000]. Although the existence of this relaxation mechanism has been known since the early days of NMR [*Kumar et al.*, 2000], a systematic utilization of the cross-correlations has been established only in recent years due to the rapid development of the multi-spin NMR sequences and NMR metodology in general [*Kumar et al.*, 2000]. Theoretical research in this field is thus highly desired and can improve the currently used protocols for structural interpretation of the measured data.

Theoretical modeling of the NMR spectroscopic parameters using modern quantum-chemistry methods has become an essential tool that complements NMR experiment. Analysis of calculated results provides a deep insight into the structural behavior of NMR parameters and their relation to the electronic structure of investigated systems.

There are several NMR parameters that can be nowadays accessed from quantum mechanical calculations, namely chemical shielding tensor, *J*-coupling and electric field gradient [*Helgaker et al.*, 1999]. To the best of our knowledge only a limited number of theoretical studies dealing with theoretical modeling of NMR parameters involved in the cross-correlations is currently available [*Bartoschek et al.*, 2003; *Sychrovský et al.*, 2005].

In the present work we studied the structural behavior of ^{15}N chemical shielding tensor of the amide nitrogen of peptide bond in L-Alanyl-L-Alanine (LALA) di-peptide. This simple model of peptide chain has recently been studied in our group by both computational and experimental spectroscopic techniques [*Bouř et al.*, 2005; *Šebek et al.*, 2007].

Theory

A comprehensive theoretical description of the cross-correlated relaxations is given by the Bloch-Wangsness-Redfield theory, usually called the Redfield theory [*Kowalewski et al.*, 2006; *Kumar et al.*, 2000].

The phenomenon of NMR relaxation arises from the existence of fluctuating local magnetic fields at the site of atomic nucleus. There are several relaxation mechanisms of a nuclear spin: direct dipole-dipole (DD) interaction, chemical shielding anisotropy (CSA), quadrupolar interaction (for the nuclear spin > ½) and other weaker relaxation mechanisms.

The Redfield theory leads to a matrix form of relaxation equations. On the diagonal of the Redfield matrix we find the auto-relaxation rates while the off-diagonal terms are the desired cross-correlated relaxation rates. To be more specific, the cross-terms represent the mutual correlation between any pair of the NMR relaxation mechanisms mentioned above. The only limitation in practice is the experimental accesibility of a particular cross-correlation mechanism.

In the present work we were interested in structural interpretation of the cross-correlation between CSA and direct DD interaction that is described by the following equations [*Kumar et al.*, 2000]

$$\Gamma_{\boldsymbol{i},\boldsymbol{jk}}^{\text{DD-CSA}} = \frac{1}{2}\left(\frac{\mu_0 \hbar}{4\pi}\right)\frac{\gamma_{\boldsymbol{j}}\gamma_{\boldsymbol{k}}}{\boldsymbol{r}_{\boldsymbol{jk}}^3}\gamma_{\boldsymbol{i}}\boldsymbol{B}_0 \cdot \Delta\sigma_{\boldsymbol{i},\boldsymbol{jk}}^{\text{eff}} \sum_{\boldsymbol{q}} \boldsymbol{J}\left(\omega_{\boldsymbol{q}}\right) \qquad (1)$$

$$\Delta\sigma_{\boldsymbol{i},\boldsymbol{jk}}^{\text{eff}} = \sum_{\boldsymbol{n}=1}^{3} \sigma_{\boldsymbol{i}}^{\boldsymbol{nn}} \left(\frac{3\cos^2\theta_{\boldsymbol{jk}}^{\boldsymbol{nn}} - 1}{2}\right) \qquad (2)$$

$$\boldsymbol{J}\left(\omega_{\boldsymbol{q}}\right) = \frac{2}{5}\frac{\tau_{\boldsymbol{c}}}{1+\left(\omega_{\boldsymbol{q}}\tau_{\boldsymbol{c}}\right)^2}, \qquad (3)$$

where $\Gamma_{\boldsymbol{i},\boldsymbol{jk}}^{\text{DD-CSA}}$ stands for the cross-correlated relaxation rate, $\Delta\sigma_{\boldsymbol{i},\boldsymbol{jk}}^{\text{eff}}$ is the effective CSA of nucleus *i* with respect to the dipolar vector between nuclei *j* and *k*, and $\boldsymbol{J}\left(\omega_{\boldsymbol{q}}\right)$ is the spectral density function of the local magnetic field fluctuations at the frequency ω_q. A particular formula for the spectral density function depends on the model used for the description of molecular rotation. The spectral density (3) corresponds to the ideal case, the isotropic tumbling of a rigid spherical molecule. The autocorrelation time τ_c can be determined either by fitting the NMR experimental data or theoretically from the molecular dynamics simulation. However, theoretical modeling of the spectral density function was not the aim of this study. Our particular attention was paid to the conformational dependence of the effective CSA (2) that essentially contributes to the cross-correlated relaxation rate (1).

The effective CSA (2) can be accessed from quantum-chemical calculations since the only data needed for its evaluation are the principal components of the chemical shielding tensor $\sigma_{\boldsymbol{i}}^{\boldsymbol{nn}}$ and the angles $\theta_{\boldsymbol{jk}}^{\boldsymbol{nn}}$ between the *nn*-component of the chemical shielding tensor and the dipolar vector *jk*. The structural dependence of the effective CSA modeled in this work is twofold since the chemical shielding tensor inherently depends on molecular conformation and the projection of tensor components on the dipolar vector (the angular part of eq. (2)) reflects their mutual orientation in the molecular frame.

Methods

The L-Alanyl-L-Alanine (LALA) molecule consists of two amino-acid residues that are numbered starting from the N-terminus of the peptide. Backbone conformation of the *n*-th amino-acid residuum is defined by the torsion angles φ_n, ψ_n and ω_n (Fig. 1). In the vast majority of peptides, the torsion angle ω is confined around 180° due to the π-electron character of peptide bond. So the major structural descriptors of peptide backbone are the torsion angles φ_n and ψ_n.

For anionic and cationic form of the LALA di-peptide both ψ_1 and φ_2 torsion angles were varied with a step of 30°, and for each of the resulting 12×12=144 geometries all the remaining coordinates were fully relaxed by energy minimization. The calculations for zwitterion were performed with larger step of 60°.

All the calculations were carried out using the density functional theory (DFT) methods with the PCM model of water solvent included. The geometries and NMR parameters were obtained at the BPW91/6-311++G(d,p) and the GIAO B3LYP/IGLO-II level of theory, respectively. The Gaussian 03 program package was used for all quantum-chemical calculations done in this work.

Results

Local conformation of peptides is primarily described by the structural parameters of the peptide backbone. This theoretical study aims to suggest a possibility of structural interpretation of the cross-correlated relaxations. Three different DD-CSA cross-correlation mechanisms between atoms of the LALA backbone were considered.

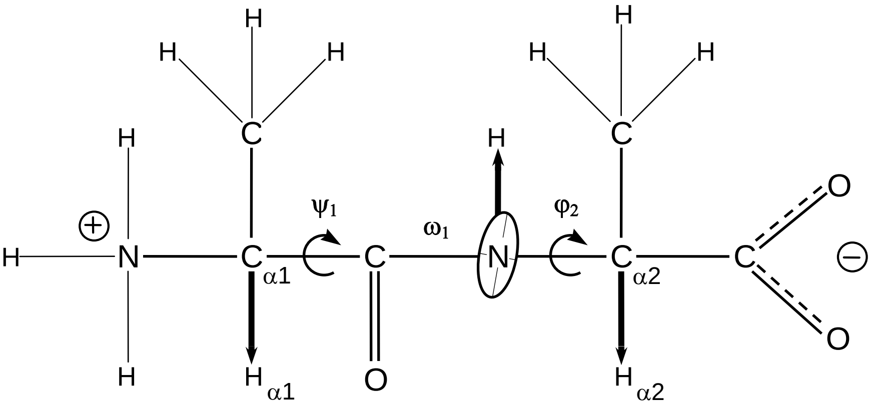


Figure 1. A sketch of the zwitterionic form of LALA di-peptide with the definition of backbone torsion angles and nuclear dipolar vectors ($C_{\alpha}H_{\alpha}$ and NH) involved in the cross-correlated relaxations.

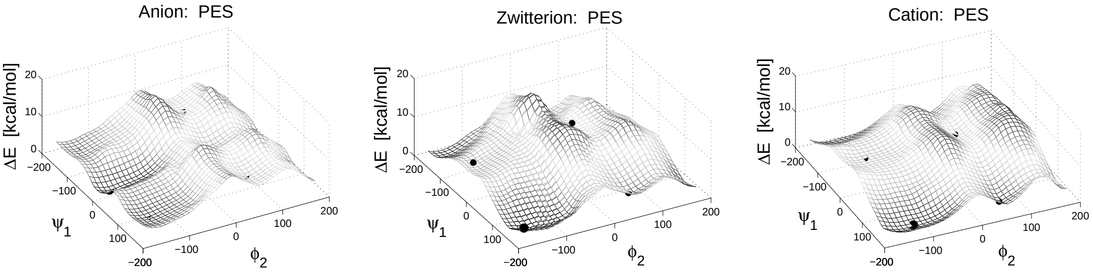


Figure 2. 2D potential energy surfaces (PES) calculated for the three forms of LALA di-peptide.

In particular, we calculated the N-NH, N-$C_{\alpha1}H_{\alpha1}$ and N-$C_{\alpha2}H_{\alpha2}$ effective CSA's that reflect mutual orientations between the ^{15}N chemical shielding tensor of amide nitrogen and the NH, $C_{\alpha1}H_{\alpha1}$, and $C_{\alpha2}H_{\alpha2}$ dipolar vectors, respectively (Fig. 1). Further we studied the dependence of these NMR parameters on the backbone torsion angles ψ_1 and φ_2 for the cationic, zwitterionic, and anionic form of LALA di-peptide.

The solution acidity, which is usually expressed in units of pH, essentially influences the peptide structure. At low pH the N-protonated cationic form of peptide prevails, at intermediate pH the double-charged zwitterionic form is the most populated one, and at high pH the C-deprotonated anionic form predominates. Therefore any theoretical investigation of the peptide structure has to consider the (de)protonation effects.

The importance of taking the molecular charge into account is clearly demonstrated in Fig. 2. The three plots exhibit appreciable differences not only in the relative energies and the positions of local minima but even the number of local minima differs.

The geometry changes obtained for different protonation states of the LALA are recapitulated in the Ramachandran plot in Fig. 3. The positions of global minima for the zwitterion and the cation coincide with the β-sheet region of the peptide secondary structure. However, for the anion the global minimum is shifted in the ψ_1 dimension (-21°) compared to the zwitterion (146°). This geometry change results from the intramolecular stabilization between the NH_2 terminal group and the amide proton of peptide bond in the anion. Only the second local minimum on the PES for the anion coincides with the global minima of zwitterion and cation (the β-sheet region).

The energy of local minima relative to the global minima calculated for all forms of the LALA, along with the particular angular coordinates are given in the first three columns of Table 1. The relative populations of the higher energy minima can be assessed from the Boltzman statistics at the room temperature of 300 K. For example, the relative population ratio of the conformers corresponding to the two lowest energy minima for the cation (energy difference 1.3 kcal/mol, Table 1) is approximately 1:9. This means that only global minima of the three LALA forms would be significantly populated, when not considering any external folding force. Thus only the NMR parameters for the conformers of global minima (Table 1) would be sufficent for the prediction of experimental observations. However, our aim was to model the general conformational behaviour of the effective CSA's in peptides with significant structural variations along the backbone that could appear in any region of Ramachandran plot. That's why we need to know the complete [ψ_1, φ_2] surface of each particular effective CSA.

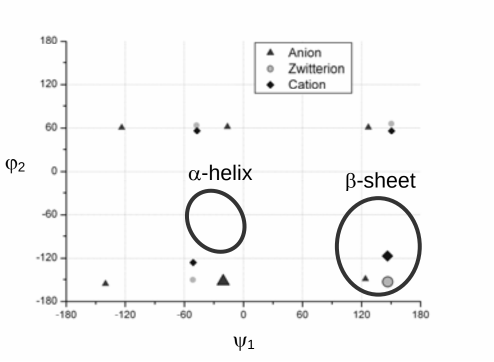


Figure 3. Ramachandran plot of the LALA with respect to the peptide backbone torsion angles ψ_1 and φ_2.

Table 1. Calculated geometry, energy and NMR parameters for the local minima of the three forms of LALA.

	Torsion angles		ΔE	$\Delta\sigma^{eff}$ [ppm]			^{15}N CS [ppm]	
LALA	ψ_1 [°]	φ_2 [°]	[kcal/mol]	N-NH	N-$C_{\alpha1}H_{\alpha1}$	N-$C_{\alpha2}H_{\alpha2}$	σ_N^{iso}	$\Delta\sigma_N$
Anion	-21	-153	0.0	-160	76	-70	99	98
	124	-149	1.6	-152	-84	-87	91	125
	-140	-156	2.9	-138	71	-58	98	92
	-16	62	4.5	-183	78	-24	103	119
	127	61	5.9	-171	-109	-2	98	116
	-124	61	7.1	-167	87	-21	103	95
Zwitterion	147	-153	0.0	-154	-97	-69	90	109
	150	66	3.6	-171	-98	-17	96	98
	-51	-150	5.1	-150	35	-81	95	94
	-48	64	9.2	-169	21	-23	102	104
Cation	146	-117	0.0	-157	-80	-124	101	106
	151	56	1.3	-161	-92	9	103	101
	-51	-126	5.1	-152	42	-119	106	93
	-47	56	6.2	-157	27	7	107	89

The surfaces of the N-NH effective CSA (first row in Fig. 4) obviously depend on both coordinates ψ_1 and φ_2. Although the structural interpretation of this cross-correlation mechanism is not straightforward, a detailed analysis of the surface(s) could provide a useful information for the confined regions of backbone conformation (like α-helix or β-sheet).

The amide nitrogen of peptide bond is involved in both NMR relaxation mechanisms, either in the ^{15}N CSA and in the NH direct dipolar interaction. Thus, the N-NH cross-correlated relaxation rate inherently depends only on variation of the chemical shielding tensor of the amide nitrogen since the origin of the NH dipolar vector coincides with the centre of the ^{15}N chemical shielding tensor. On the contrary, both N-$C_\alpha H_\alpha$ effective CSA's further depend on the mutual orientation of the ^{15}N chemical shielding tensor with respect to the $C_{\alpha1}H_{\alpha1}$ and $C_{\alpha2}H_{\alpha2}$ dipolar vectors, i.e. they are closely related to the torsion angles ψ_1 and φ_2, respectively (Fig. 1). These ideas help to understand the qualitatively different shapes of the calculated surfaces (Fig. 4).

The N-$C_{\alpha1}H_{\alpha1}$ effective CSA (the second row in Fig. 4) depends dominantly on the torsion angle ψ_1 whereas only a slight modulation was calculated in the φ_2-dimension. This is true for all forms of the LALA and it's rather consistent with the previous reasoning. The variation of the torsion angle ψ_1 modulates the angular part of the eq. (2) due to its structural overlap with the N-$C_{\alpha1}H_{\alpha1}$ cross-correlation mechanism (Fig. 1), while the angle φ_2 is spatially separated from this cross-correlation mechanism.

The prevalent one-dimensionality of the N-$C_{\alpha2}H_{\alpha2}$ effective CSA (the last row in Fig. 4) is again obvious. This time the torsion angle φ_2 is the determining variable, what is again supported by the geometric evidence from Fig. 1.

The N-$C_{\alpha1}H_{\alpha1}$ and N-$C_{\alpha2}H_{\alpha2}$ effective CSA's are worth further analysis, since their dependence on the backbone conformation is one-dimensional and relatively smooth and thus they might be parametrized rather simply. Firstly, the "weak" dependence of each effective CSA on one of the torsion angles may be neglected. Then an averaged one-dimensional projection of each surface was constructed (Figure 5). The maximal variation in the "weaker" dimension was 33 ppm for the N-$C_{\alpha1}H_{\alpha1}$ effective CSA and 41 ppm for the N-$C_{\alpha2}H_{\alpha2}$ effective CSA.

These smoothed true one-dimensional dependencies of the two effective CSA's were fitted with the Karplus-like curves [*Karplus*, 1963] (Fig. 5):

$$\boldsymbol{y} = \boldsymbol{C} + \boldsymbol{A}_1 \cos(\boldsymbol{x} - \boldsymbol{b}) + \boldsymbol{A}_2 \cos^2(\boldsymbol{x} - \boldsymbol{b}) \qquad (4)$$

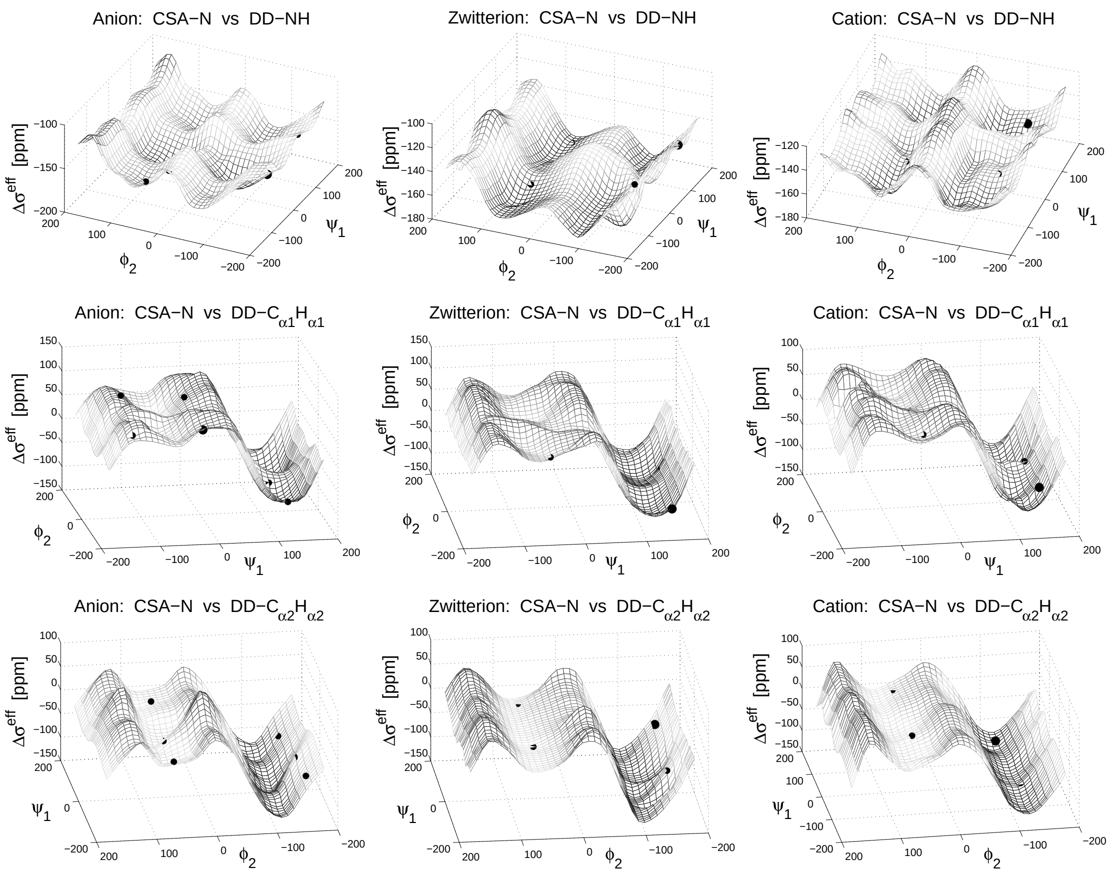


Figure 4. Surfaces of the effective CSA's. Columns represent the three different forms of the LALA and rows correspond to the three studied cross-correlation mechanisms, N-NH, N-$C_{\alpha1}H_{\alpha1}$ and N-$C_{\alpha2}H_{\alpha2}$.

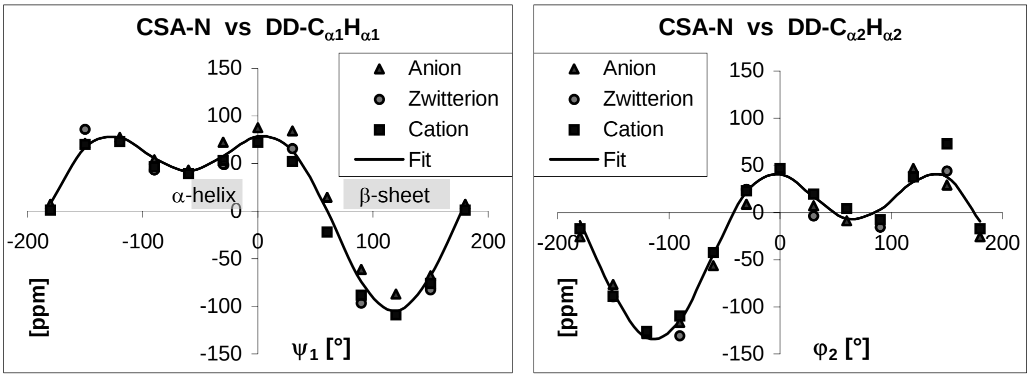


Figure 5. The fitted pH-independent Karplus-like curves for the N-$C_{\alpha1}H_{\alpha1}$ (left) and N-$C_{\alpha2}H_{\alpha2}$ (right) effective CSA averaged over all three forms of LALA di-peptide.

Before we started looking for the fitting parameters in eq. (4) we could make the second simplification of our results. Based on the great similarity of the surfaces for the three LALA forms (Fig. 4, Fig. 5) and taking into account the accuracy of the theoretical predictions as well as the possible experimental errors we could state that the calculated structural dependencies of the effective CSA's are pH-independent. Therefore we finally fitted all the calculated data for each effective CSA with only one pH-independent Karplus-like curve shown in Fig. 5 (see Appendix for the fitted parameters).

Conclusion

The calculated dependences of the N-NH, N-$C_{\alpha1}H_{\alpha1}$ and N-$C_{\alpha2}H_{\alpha2}$ effective CSA's on the torsion angles ψ_1 and φ_2 can be used for distinguishing between various conformations of peptide backbone. For example, the N-

$C_{\alpha1}H_{\alpha1}$ effective CSA is indicative for discrimination between the major peptide backbone conformations, since even the sign of this NMR parameter is specific with regard to the α-helix and the β-sheet (Fig. 5).

The N-$C_{\alpha1}H_{\alpha1}$ and N-$C_{\alpha2}H_{\alpha2}$ effective CSA's exhibit one-dimensional behaviour, while the N-NH effective CSA is strongly modulated by both backbone torsion angles ψ_1 and φ_2.

The dependence of the calculated effective CSA's on pH is very modest. For example, the splitting of the N-$C_{\alpha1}H_{\alpha1}$ effective CSA between α-helix and β-sheet is very similar for all three forms of the LALA molecule. We thus can conclude that major structural characteristics of the modeled effective CSA's remain almost identical in a wide range of pH. The effective CSA's that exhibit significant dependence on one dimension only were fitted as the pH-independent Karplus-like curves.

The calculated effective CSA's in LALA di-peptide can be generalized and used for the structural interpretation of the cross-correlated relaxation rates in peptides.

Acknowledgments. The author thanks the Molecular Spectroscopy Department at the IOCB for providing the computational resources. Special thanks to Dr. V. Sychrovský for critical reading of the manuscript and valuable discussions.

Appendix

Table 2. Parameters of the Karplus-like curves (4) fitted with the Least squares method.

N-$C_{\alpha1}H_{\alpha1}$		N-$C_{\alpha2}H_{\alpha2}$	
Parameter	Value	Parameter	Value
C [ppm]	65.3	C [ppm]	31.2
A_1 [ppm]	-73.7	A_1 [ppm]	63.9
A_2 [ppm]	-96.8	A_2 [ppm]	-102.6
b [°]	-118.6	b [°]	-67.1

References

Bartoschek, S.; Buurman, G.; Geierstanger, B. H.; Lapham, J.; Griesinger, C.; Measurement and ab Initio Calculation of CSA/Dipole-Dipole Cross-Correlated Relaxation Provide Insight into the Mechanism of a H_2-Forming Dehydrogenase. *J. Am. Chem. Soc.*, 125, 13308-13309, 2003.

Bouř, P.; Buděšínský, M.; Špirko, V.; Kapitán, J.; Šebestík, J.; Sychrovský, V.; A Complete Set of NMR Chemical Shifts and Spin-Spin Coupling Constants for L-Alanyl-L-Alanine Zwitterion and Analysis of Its Conformational Behavior *J. Am. Chem. Soc.*, 127, 17079-17089, 2005.

Helgaker, T.; Jaszunski, M.; Ruud, K.; Ab Initio Methods for the Calculation of NMR Shielding and Indirect Spin-Spin Coupling Constants. *Chem. Rev.*, 99, 293-352, 1999.

Karplus, M.; Vicinal Proton Coupling in Nuclear Magnetic Resonance. *J. Am. Chem. Soc.* 85, 2870-2871, 1963.

Kowalewski, J.; L. Mäler, L. in Nuclear spin relaxation in liquids: Theory, Experiments, and Applications. Taylor and Francis, New York – London, 2006.

Kumar, A.; Grace, R. C. R.; Madhu, P. K.; Cross-correlations in NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* 37, 191–319, 2000.

Levitt, M.H. in Spin Dynamics: Basics of Nuclear Magnetic Resonance. John Wiley & sons, Chichester, 2001.

Ravindranathan, S.; Kim, C. H.; Bodenhausen, G.; Cross correlations between ^{13}C-^{1}H dipolar interaction
s and ^{15}N chemical shift anisotropy in nucleic acids. *J. Biomol. NMR*, 27, 365-375, 2003.

Reif, B.; Hennig, C.; Direct measurement of angles between bond vectors in high resolution NMR. *Science*, 276, 1230-1233, 1997.

Reif, B.; Diener, A.; Hennig, M.; Maurer, M; Griesinger, C.; Cross-Correlated Relaxation for the Measurement of Angles between Tensorial Interactions. *J. Magn. Reson.* 143, 45–68, 2000.

Sychrovský, V.; Müller, N.; Schneider, B.; Smrečki, V.; Špirko, V.; Šponer, J; Trantírek, L.; Sugar Pucker Modulates the Cross-Correlated Relaxation Rates across the Glycosidic Bond in DNA. *J. Am. Chem. Soc.* 127, 14663-14667, 2005.

Šebek, J.; Gyurcsik, B.; Šebestík, J.; Kejík, Z.; Bednárová, L.; Bouř, P.; Interpretation of Synchrotron Radiation Circular Dichroism Spectra of Anionic, Cationic and Zwitterionic Dialanine Forms. *J. Phys. Chem. A*, 111, 2750-2760, 2007.

Dependence of the L-Alanyl-L-Alanine Conformation on Molecular Charge Determined from Ab Initio Computations and NMR Spectra

Vladimír Sychrovský,* Miloš Buděšínský,* Ladislav Benda, Vladimír Špirko,* Zuzana Vokáčová, Jaroslav Šebestík, and Petr Bouř*

Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Flemingovo nám. 2, 16610, Prague 6, Czech Republic

Received: August 15, 2007; In Final Form: October 21, 2007

The L-alanyl-L-alanine (AA) molecule behaves differently in acidic, neutral, and basic environments. Because of its molecular flexibility and strong interaction with the aqueous environment, its behavior has to be deduced from the NMR spectra indirectly, using statistical methods and comparison with ab initio predictions of geometric and spectral parameters. In this study, chemical shifts and indirect spin−spin coupling constants of the AA cation, anion, and zwitterion were measured and compared to values obtained by density functional computations for various conformers of the dipeptide. The accuracy and sensitivity of the quantum methods to the molecular charge was also tested on the (mono)-alanine molecule. Probable AA conformers could be identified at two-dimensional potential energy surfaces and verified by the comparison of the computed parameters with measured NMR data. The results indicate that, whereas the main-chain peptide conformations of the cationic (AA^+) and zwitterionic (AA^{ZW}) forms are similar, the anion (AA^-) adopts also another, approximately equally populated conformer in the aqueous solution. Additionally, the NH_2 group can rotate in the two main chain conformations of the anionic form AA^-. According to a vibrational quantum analysis of the two-dimensional energy surfaces, higher-energy conformers might exist for all three charged AA forms but cannot be detected directly by NMR spectroscopy because of their small populations and short lifetimes. In accord with previous studies, the NMR parameters, particularly the indirect nuclear spin−spin coupling constants, often provided an excellent probe of a local conformation. Generalization to peptides and proteins, however, has to take into account the environment, molecular charge, and flexibility of the peptide chain.

Introduction

NMR spectroscopy has a long history in the conformational analyses of peptide structures.[1,2] Empirical correlations of chemical shifts and nuclear spin−spin coupling constants (*J*-coupling) with the geometry were originally used to discriminate helical and sheetlike peptides and subsequently were extended to protein studies.[3−5] The possibility to calculate the NMR parameters for larger molecules with reasonable precision has lately provided an additional basis for the interpretation of the experiment, which consequently facilitated the verification of various conformational models. Particularly, the analytical approaches to chemical shifts[6] and the coupling constants[7−9] within the density functional theory (DFT) speeded up the computations and facilitated the conformational studies of interesting peptide systems.[10,11] The computations improved, for example, the empirical Karplus relations between the spectra and the structure.[12,13] However, peptide flexibility, solvent effects, and local vibrational motions have to be taken into account in accurate modeling.[14]

The dependence of the L-alanyl-L-alanine (AA) conformation on molecular charge studied in this work thus provides additional information about the behavior of the peptide chain in various pH conditions as well as about the accuracy and validity of current simulation techniques. NMR properties of charged and zwitterionic peptides themselves are notoriously difficult to calculate[15] mainly because of the electronic charge concentration requiring a large basis set and because of the strong interaction with the environment, in most cases with water.[16] The three AA forms also provide an experimentally well-accessible example of a simple molecular mechanical system controllable by pH. Therefore, we find it interesting to analyze in detail the two-dimensional potential energy surface and account for possible vibrational quantum effects.[17]

In the second part of this work, we use the statistical comparison of the experimental and computed chemical shifts and spin−spin coupling constants developed previously for the AA zwitterion.[11] The ability of the computation to discriminate between various charged forms is tested on the alanine molecule (A) labeled with stable ^{15}N and ^{13}C isotopes, where the conformational problem is simpler than in AA. The influence of the charge on the molecular potential energy surface, expressed as a function of the main-chain peptide torsion angles (φ, ψ), is computed with the inclusion of a continuum solvent correction. It appears that the pH (charge) change stabilizes a new anion (AA^-) conformer and that the conformational equilibrium can be proven from the NMR data. The quantum vibrational analysis predicts also other well-defined conformers, which, however, are neither significantly populated under normal conditions nor can be observed directly by NMR due to their short lifetimes.

Experimental Section

Isotopically labeled L-alanine (^{13}C,98%; ^{15}N,98%) was purchased from Stable Isotopes, Inc., whereas the nonlabeled AA

---

* To whom correspondence should be addressed: E-mail: vladimir.sychrovsky@uochb.cas.cz (V.S.), budes@uochb.cas.cz (M.B.), vladimir.spirko@marge.uochb.cas.cz (V.S.), bour@uochb.cas.cz (P.B.).

10.1021/jp076557j CCC: $40.75 © 2008 American Chemical Society
Published on Web 01/03/2008

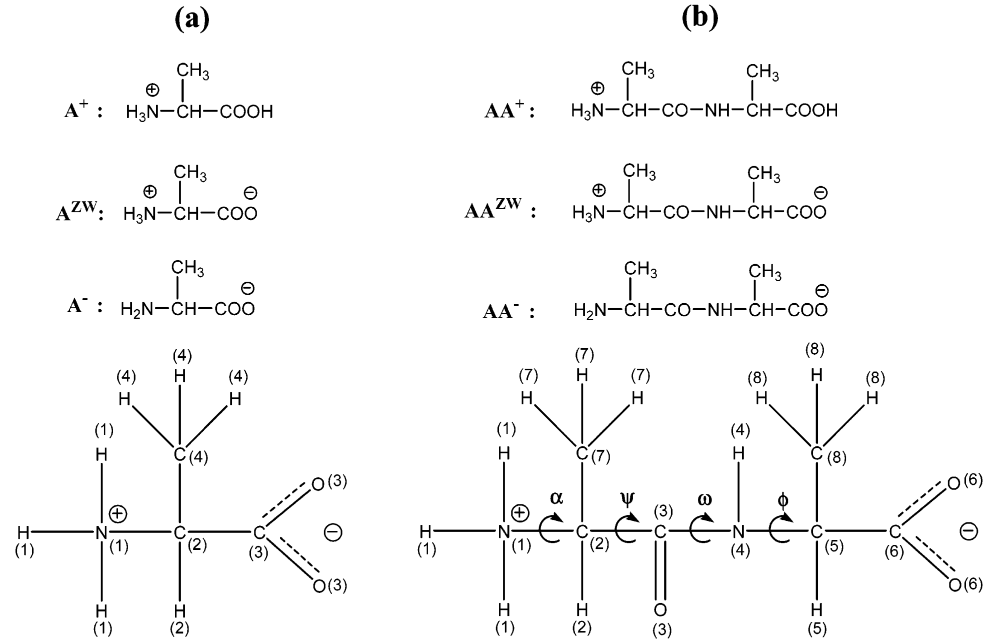


Figure 1. The ionic forms and symbols used for L-alanine (a) and AA (b). The numbering of the atoms for the definition of NMR parameters is shown in zwitterionic forms of both molecules. In addition to standard peptide torsion angles (ϕ, ψ, ω), we have introduced the angle α as the average angle of the two amine hydrogen −(C3−C2−N1−H1) torsion angles (for the NH_2 residue in AA^-).

was purchased from Sigma. The synthesis of labeled AA is described elsewhere.[11] NMR spectra of labeled alanine (A^+, A^-, A^{ZW}), natural AA (AA^+, AA^-, AA^{ZW}), and labeled AA^{ZW} were measured with Fourier transform (FT) NMR spectrometers Varian UNITY-500 and Bruker AVANCE-500 (^1H at 500 MHz, ^{13}C at 125.7 MHz, ^{15}N at 50.7 MHz, ^{17}O at 67.8 MHz) in D_2O and/or in the mixture H_2O/D_2O (9:1). The solution pH was varied by additions of 2 M HCl (pH ≈ 2) and NaOH (pH ≈ 12) solutions. For these pH values, the AA peptide exists exclusively in the AA^+ and AA^- forms, respectively, which was confirmed by measuring of the NMR titration curves (provided in the Supporting Information (SI), together with relevant p*K* constants for A and AA). The zwitterionic forms were obtained by dissolving the compounds in distilled water without any buffer. All spectra were measured at room temperature. Chemical shifts were referenced either to internal (DSS for ^1H and ^{13}C and H_2O for ^{17}O, but the oxygen is not discussed in this work) or external (nitromethane in the capillary for ^{15}N) standards. The structural assignment of the hydrogen and carbon chemical shifts was achieved using homonuclear and heteronuclear two-dimensional techniques with pulse field gradients (2D-^1H,^1H-PFG-COSY, ^1H,^{13}C-PFG-HSQC, and 2D-^1H,^{13}C-PFG-HMBC) in D_2O. The solvent mixture H_2O/D_2O (9:1) was used to observe the signals of NH and NH_3^+ protons. Only the couplings of amide NH could be observed in this mixture because of the fast exchange rate of amine NH_3^+ protons with water. The *J*(H,H) values were determined from the 1D-^1H NMR spectrum and the *J*(C,H) couplings from the non-decoupled 1D-^{13}C NMR spectrum. A series of selective ^1H-decoupled ^{13}C NMR spectra was used to assign individual *J*(C,H) couplings. The labeled ^{15}N and ^{15}N,^{13}C AA samples were used mainly in order to obtain *J*(N,H), *J*(N,C), and *J*(C,C) coupling constants using the 1D-^1H and ^{13}C NMR spectra, the 1D-^{13}C-INADEQUATE (Incredible Natural Abundance Double Quantum Transfer Experiment), and the 2D-^1H,^{15}N-PFG-HMBC spectra.

Calculations. The GAUSSIAN software[18] was used for the quantum chemical computations. The torsion angles φ and ψ (Figure 1) were varied with 30° steps, and for each of the resultant $12 \times 12 = 144$ geometries, all the remaining coordinates were fully relaxed by energy minimization. The angle ω was initially set to 180° so as to maintain the *trans*-peptide bond, because the experimental data do not suggest a presence of the *cis*-conformer. In the scans, the relaxed ω angle deviated from 180° by less than ~5°. The BPW91[19] functional and the standard Pople-type 6-311++G** basis were used with the PCM solvent model for water.[20] For the anion AA^-, three scans were performed (3×144 points), taking into account three initial orientations ($\alpha = -120$, 0, and 120°) of the NH_2 group. For this purpose, we defined the angle α as an average of the two amine hydrogen −(C3−C2−N1−H) angles. The geometries of the local minima located on the resultant surfaces were fully optimized without any constraints. The scan of the zwitterion mimics previous computations done with smaller grid steps.[11] For control computations, other potential energy scans were done in a vacuum for AA^+ and AA^- (AA^{ZW} is not stable without solvent), and local minima geometries were reoptimized at the MP2[21]/6-311++G** level of approximation.

For the grid points and local minima geometries, the NMR shielding tensors and the indirect NMR spin−spin coupling constants were calculated. The default gauge-invariant atomic orbital (GIAO)[22] method was used for the shieldings to prevent their origin dependence. For the *J*-couplings, all four important terms[7,23] were included in the analytical coupled-perturbed DFT computation. For the shieldings and *J*-couplings, we used the B3LYP functional[24] with the IGLOII or IGLOIII bases,[25] which are believed to be well-suited for computations of the NMR properties.[8] In all cases, the same PCM solvent model was applied. With the same method, the NMR parameters were calculated for the anion, cation, and zwitterion of (mono)-alanine in equilibrium geometries. Chemical shifts were related to standard molecules (DSS for ^1H and ^{13}C, nitromethane for ^{15}N, water for ^{17}O, same as in ref 11).

The quantum and dynamical effects of the torsional motions involving the angles (φ, ψ) were studied using an approximate Hamiltonian described in full in ref 11, where the potential was obtained by fitting of the computed two-dimensional surfaces, and the kinetic part was expressed in the curvilinear angular

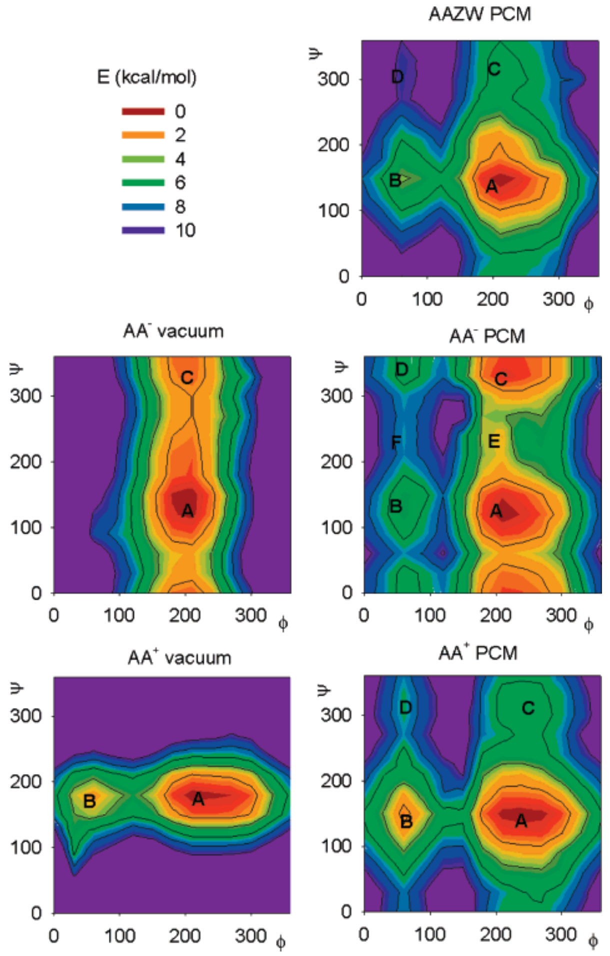


Figure 2. Contour plots of the computed (BPW91/6-311++G**) potential energy surfaces of the three AA forms. By default, the PCM solvent model was applied (right-hand side); for the AA^+ and AA^- forms stable in a vacuum, the dependencies without the solvent are shown on the left. Approximate positions of local minima are marked by the capital letters. To avoid splitting of the minimum well, the angles are plotted within the positive (0, 360°) interval instead of the usual (−180, 180°) range.

coordinates.[26] The eigenvalue vibrational problem $H\Psi = E\Psi$ was solved variationally in basis set functions expressed as products of the eigenfunctions of the corresponding uncoupled one-dimensional Schrödinger equations. The one-dimensional functions were determined numerically using the Numerov−Cooley integration procedure[27] and used for averaging the NMR parameters. The averaging, however, did not bring significant corrections with respect to the overall accuracy and is not discussed further. The lifetimes of selected localized excited vibrational states were estimated from simplified one-dimensional modeling.[11,28]

Results and Discussion

AA Conformers. The calculated adiabatic energy dependencies on the φ and ψ torsion angles are presented in Figure 2. Apparently, the energy profiles of the three hydrated forms exhibit several similarities. The global minimum **A** is associated with comparable angle values for all cases (see Table 1 for details). The similarity of the $\mathrm{AA}^{\mathrm{ZW}}$ and AA^+ surfaces is the most obvious. However, the protonation of $\mathrm{AA}^{\mathrm{ZW}}$ makes the resultant cation (AA^+) more flexible with respect to the φ rotation, and the potential well is elongated along this coordinate. The MP2 method provides an equilibrium value of the φ angle (−151°, see Table 1) even significantly shifted from the DFT angle of −121°. For AA^+, the local energy minimum **B** deepens when compared to the other forms, but its relative energy of 1.4 kcal/mol (cf. Table 1) and narrow potential well probably still prevent a significant population of this conformer in the sample. Obviously, the least probable is an occurrence of the other two, **C** and **D** conformers of $\mathrm{AA}^{\mathrm{ZW}}$ and AA^+.

The energy surface of the anion (AA^-, the middle of Figure 2) is different. Whereas the geometry of the lowest-energy **A** conformer is very close to that of the zwitterion, the minimum well **C** significantly broadens, and its energy is comparable with the global minimum **A**. Additionally, new, very shallow minima (**E**, **F**) appear for the anion; these are, along with the **B** and **D** extremes, not populated due to their high relative energies. For **C**, however, the computed relative energy (0.4 kcal/mol, Table 1) is probably comparable with the computational error and suggests a significant presence of this conformer in the sample at room temperature.

The more complicated behavior of the anion AA^- stems predominantly from the directional and ambivalent binding properties of the NH_2 group. The nitrogen electron lone pair can make an internal hydrogen bond to the amide hydrogen, or the NH_2 protons can be bonded to the carbonyl oxygen. This is documented in Figure 3, where for a fixed value of $\varphi = -150°$ the energy dependence on the ψ angle is plotted. Although detailed three- or more dimensional energy scans are currently not feasible, the coupling of the α, ψ, and φ torsional motions

TABLE 1: Computed (BPW91/PCM/6-311++G) Geometries and Relative Conformer Energies of the Three Charged AA Forms**[a]

	AA^+			$\mathrm{AA}^{\mathrm{ZW}}$			AA^-			
conformer	ψ	φ	E	ψ	φ	E	α	ψ	φ	E
A	149	−121	0.0	147	−153	0.0	−2	127	−152	0.0
A′							118	128	−151	0.8
A″							−125	119	−150	1.9
A[b]	178	−159					−11.3	138	−160	
A[c]	150	−151		146	−158		−13.1	133	−157	
A[d]	169	−160					−2.5	139	−160	
B[e]	150	60	1.4	150	66	3.6	−3	135	63	4.3
C	−55	−127	5.1	−51	−150	5.1	−136	−19	−152	0.4
C′							145	8	−147	0.8
C″							9	−21	−150	1.4
D	−55	53	6.3	−48	64	9.2	−138	−16	63	5.0
E							5	−105	−154	2.7
F							7	−116	61	6.9

[a] Torsion angles (φ, ψ, α) are given in degrees and the relative energies (*E*) in kcal/mol. [b] BPW91, gas phase. [c] MP2, PCM. [d] MP2, gas phase. [e] Calculated lifetimes of this conformer are approximately 1, 6−10, and 0.7 ms for AA^+, $\mathrm{AA}^{\mathrm{ZW}}$, and AA^-, respectively.

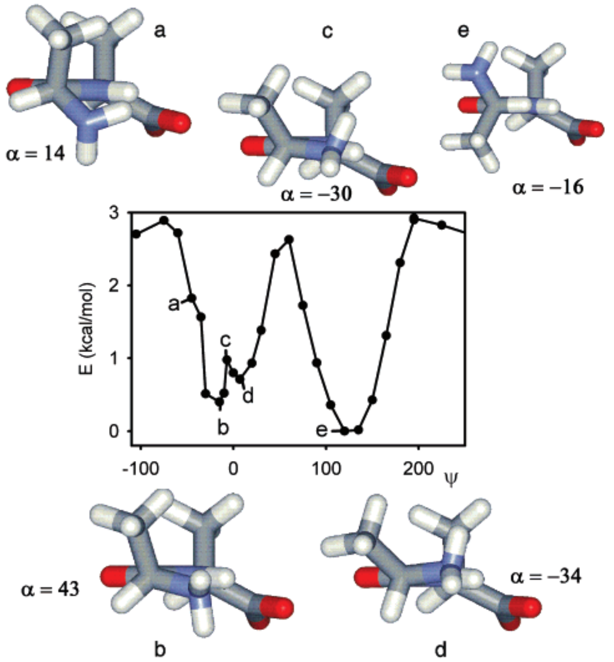


Figure 3. Detailed one-dimensional profile of the AA^- calculated potential energy surface for $\varphi = -150°$. The NH_2 binding pattern changes along the lowest-energy path: (a) the NH_2 hydrogens may be attracted by the COO^- group and the amide bond π-system, or (b) the NH_2 nitrogen electron lone pair creates a hydrogen bond to the amide hydrogen which is (c, d) conserved during the rotation over the syn ($\psi \approx 0°$) conformation. At the other minimum (e), the NH_2 hydrogens are attracted to the π-system as well as to the amide oxygen. The NH_2 rotation is indicated (α, in degrees).

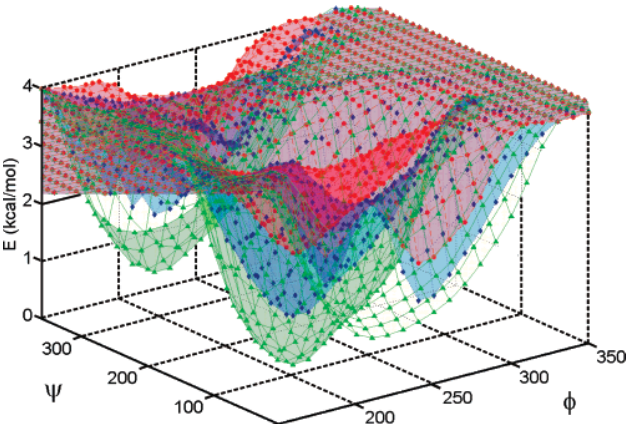


Figure 4. Part of the anionic (AA^-) potential energy surfaces (green, blue, and red) calculated for three NH_2 rotamers.

in AA^- nicely reveals the complexity and caveats that must be taken into account in the modeling of peptide conformational landscape. In longer molecules and proteins the folding of the backbone can presumably be influenced by side-chain interactions similarly as the (ψ, φ) energy map by the NH_2 group rotation. The one- and two-dimensional surfaces (Figures 2 and 3) can be also considered as projections of multidimensional surfaces. Another way illustrating this situation is represented in Figure 4, where three (φ, ψ) anionic surfaces are plotted as obtained from scans with three different NH_2 rotamers. The lowest-energy surface in Figure 2 was obviously obtained by taking the minimum energies for each (φ, ψ) pair in Figure 4.

However minor the purely electrostatic influence of the charged ends on the conformation might be in comparison with the role of the NH_2 group, a more detailed look at the optimal geometries in Table 1 reveals effects that can be attributed solely to the charge/pH changes. The protonation of the COO^- group ($AA^{ZW} \rightarrow AA^+$) results in a change of the φ torsion around the adjacent C−N bond by about 30° with the ψ torsion remaining nearly unaffected. Similarly, for the anionic form the ψ torsion decreases by about 20° under the $NH_3^+ \rightarrow NH_2$ deprotonation of the zwitterion. Whereas the effects of the pH and the NH_2 intrinsic hydrogen bonding can hardly be separated, the charge change undoubtedly further tweaks the dipeptide conformational properties.

Strictly speaking, two rotamers of the COOH group (and two other rotamers associated with the OH rotation are theoretically possible) should be considered in the potential energy surface of the cation AA^+ in a similar way as the NH_2 rotation was treated in the anion AA^-. However, the energy changes involved in the COOH rotation are minor (less than 1 kcal/mol) and so is the effect of these geometry variations on the NMR parameters. Therefore, the AA^+ (φ,ψ) surface was considered only for the lowest-energy COOH conformers.

Not all local minima on the potential energy surfaces in Figure 2 can support a stable quantum state. For both AA^+ and AA^-, only the two lowest-energy conformers are stable. This can be deduced from the two-dimensional Hamiltonian and the resultant localized wave functions shown for the cation and anion in Figure 5. A similar result was obtained for the AA zwitterion.[11] The transition barriers are clearly high enough to support at least the two distinct conformers with the wavefunctions plotted by the green lines in Figure 5. Moreover, we can deduce from the lifetimes of the second-lowest energy states listed in the footnotes of Table 1 (a few milliseconds) that it is in principle possible to prepare these conformers, although presently they cannot be detected by the inherently slow NMR experiment.

By comparison of the computations performed in a vacuum and with the solvent correction (the left- and right-hand parts of Figure 2, respectively), one can well estimate the effect of the environment on the peptide conformational properties. While it is true that the basic conformational characteristics of the dipeptide given by the covalent bonds do not change upon the solvation, the final energy profiles, the exact minima geometries, and, particularly, the steepness of the equilibrium potential wells do change. The MP2 method provides virtually the same conformers as the BPW91 functional (Table 1) except for the φ angle of the AA^+ form; in this case, however, the global minimum well is very broad and the equilibrium φ difference does not necessarily implicate a different behavior. Other backbone torsion angles obtained for the minima with the BPW91 and MP2 methods differ by less than 6°. The values of the torsion angles for the **A**-type AA conformers are close to those observed in β-sheet structures of longer peptides and proteins (antiparallel β-sheet: $\varphi,\psi = -139°,135°$; parallel β-sheets: $\varphi,\psi = -119°,113°$),[29] whereas the other types do not have canonical protein counterparts.

Similarly as for the zwitterion,[11] molecular dynamics (MD) simulations provided analogous conformer distribution to those deduced from the ab initio relative energies also for the positively and negatively charged AA forms (see Figure 1s in SI for the Amber force field). Particularly, the MD computations are consistent with the one (for AA^{ZW} and AA^+) and two (for AA^-) conformer presence in the sample under room temperature. Obviously, detailed MD results strongly depend on the force field parametrization, and presently we consider them inferior to those based on the DFT energy maps.

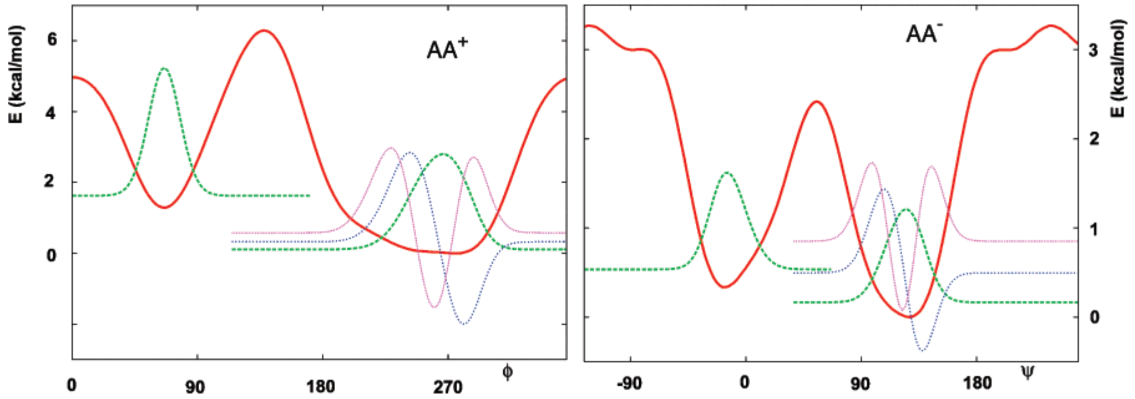


Figure 5. One-dimensional sections of the cation (AA^+) and anion (AA^-) smoothed potential energy surfaces (the red line) and vibrational wave functions of the lowest-energy states (the dashed/dotted lines, with the asymptotes corresponding to their energies).

TABLE 2: Experimental (in D_2O) and Computed[a] Changes of Alanine NMR Parameters under the pH Variations

	$\Delta(A^+ - A^{ZW})$		A^{ZW}		$\Delta(A^- - A^{ZW})$	
	calcd	exptl	calcd	exptl	calcd	exptl
			Chemical Shifts (ppm)			
N	−4.00	−2.20	−386.1	−339.6	−13.1	−6.6
C2	−0.17	−1.77	62.79	53.35	1.32	0.93
C3	−0.26	−3.11	188.38	178.66	11.74	8.94
C4	−0.27	−0.83	20.70	18.97	8.14	4.25
H2	0.81	0.37	4.17	3.78	−0.64	−0.48
H4	0.18	0.08	1.57	1.48	−0.44	−0.26
			Spin−Spin Coupling Constants (Hz)			
^{1}***J***:						
N1,C2	−1.6	−0.85	−3.2	−5.7	1.9	1.4
C2,C4	−4.1	−0.8	36.8	34.9	−1.9	0.3
C2,C3	12.2	5.6	48.5	54.0	3.1	−1.3
C2,H2	4.4	1.5	143.6	145.1	−10.4	−6.7
C4,H4	2.9	1.3	127.0	129.7	−3.4	−2.1
^{2}***J***:						
N1,H2	−0.2	0	−0.9	0	−2.4	−2.2
N1,C4	−0.1	0	−0.7	0	−3.1	0
N1,C3	0.1	0	−0.5	0	0	0
C2,H4	0.2	−0.2	−3.1	−4.4	0.3	0.1
C4,H2	−0.8	−0.35	−3.1	−4.55	1.0	−0.15
C3,C4	−0.3	−0.1	−0.7	−1.2	0.4	1.2
C3,H2	−0.8	−1.0	−4.6	−5.0	0.5	0.7
^{3}***J***:						
N1,H4	0.3	0.05	−3.3	−3.05	0.6	0.1
C3,H4	0.5	0.35	4.1	4.2	−0.2	0.1
H2,H4	0.1	0	7.3	7.3	−0.6	−0.2

[a] BPW91/6-311++G**/PCM geometries, B3LYP/IGLOIII NMR parameters. For A^+ and A^-, differences with respect to A^{ZW} are given; for A^-; the values calculated for three NH_2 rotamers were averaged.

Alanine NMR pH Dependence. The anion (A^-), cation (A^+), and zwitterion (A^{ZW}) of the ^{15}N,^{13}C-isotopically labeled alanine allowed us to investigate the dependence of the chemical shifts and spin−spin coupling constants for a system where no significant change of molecular shape can be induced by the change of the molecular charge. A complete set of experimental and calculated NMR parameters is given in Table 2. The absolute chemical shifts are reproduced with a notable error (e.g., computed 53.35 ppm, experimental 62.79 ppm for C2); nevertheless all signs of the differences between the charged and neutral forms are reproduced correctly. As expected, the hydrogen shifts could be calculated with a higher precision than those for the heavy atoms.

The alanine spin−spin coupling constants change under the deprotonation and protonation like the shifts, typically within 1−10%. The computation well reproduces the magnitudes of the individual coupling constants (e.g., for A^{ZW}, the experimental ^1J(C2,H2) = 145.1 Hz was calculated as 143.6 Hz, i.e., with 1% error). As in the case of the shifts, the computations reproduce most trends induced by the pH change, albeit with a limited precision. As an extreme case, the ^1J(C2,C3) coupling changes by −1.3 Hz while a +3.1 Hz change is predicted by the theory. With the exception of the error of the DFT method, we attribute the deviations to MD or solvent−solute interactions, incompletely covered by the present model. Other functionals (BPW91) and bases (IGLOII, 6-311++G**) provided similar results. Overall, we can see that the computations correctly reproduce the main changes in NMR shifts and coupling patterns induced by the change of molecular charge.

Accuracy of the Calculated NMR Parameters in AA. For the AA^+ cation, the accuracy of various approximate levels used for the computations of the NMR shifts and coupling constants are demonstrated in Table 3. The anion AA^- behaves similarly; the data can be found in the SI (Table 1s). As for alanine, the error of the chemical shifts computed for dialanine significantly exceeds the estimated experimental inaccuracy (estimated as 0.01, 0.02, and 4 ppm for the NMR shifts of hydrogen, carbon, and nitrogen, respectively). As observed earlier,[11,30] the hydrogen shifts are reproduced relatively accurately, whereas the DFT method becomes inaccurate for carbons, with the inaccuracy being even higher for the nitrogen atom. The parameters computed at different approximation levels do not vary dramatically. The best overall agreement of the calculated data with the experiment was achieved with the BPW91/6-311++G**/PCM equilibrium geometry and at the B3LYP/IGLOII/PCM level for the NMR parameters. The IGLOIII basis, although bigger, provides less accurate shifts than IGLOII.

Interestingly enough, the gas-phase computations appear to be reasonably accurate. For the geometry obtained with the MP2 method in the gas phase (vacuum), for example, the calculated carbonyl carbon shifts are even closer to experiment than when the PCM solvent correction is applied. For DFT, however, the PCM results are more precise. The differences in geometries (Table 1) and NMR shifts obtained with the MP2 and DFT and with PCM and gas phase might indicate that the current potential energy surfaces (Figure 2) are not quite accurate and that a better solvent model accounting for the directional hydrogen bonds[31] would be more appropriate; this is, however, impossible to achieve with the computer means available. On the other hand, some properties of the dipeptide are reproduced very well, e.g., the observed 84 ppm shift difference between N1 and N4 was calculated within the 92−95 ppm interval. The chemical environment (amine−amide) is thus perhaps more reproducible than solvent environment (vacuum−water), whose influence is weaker.

TABLE 3: Chemical Shifts and Spin−Spin Coupling Constants Calculated at Different Levels of Theory for Conformer A of the AA$^+$ Cation[a]

geometry (6-311++G**):	BPW91 PCM	BPW91 PCM	BPW91 (gas)	MP2 PCM	MP2 (gas)	
NMR (B3LYP):	IGLOIII/PCM	IGLOII/PCM	IGLOIII (gas)	IGLOIII/PCM	IGLOIII (gas)	exptl
Chemical Shifts (ppm)						
H2	4.63	4.45	4.07	4.40	3.87	4.10
H5	5.53	5.27	4.75	5.01	4.58	4.42
H7	1.58	1.58	1.77	1.41	1.56	1.55
H8	1.48	1.45	1.69	1.31	1.52	1.45
C2	63.2	61.4	63.3	59.9	60.0	51.8
C3	182.6	177.4	178.7	181.5	176.0	173.5
C5	56.4	54.6	61.8	55.7	58.9	51.6
C6	192.2	186.6	187.9	189.8	185.7	179.1
C7	22.0	21.5	21.0	21.8	22.2	19.4
C8	24.3	23.8	21.8	23.2	21.4	18.9
N1	−390	−379	−390	−392	−392	−342
N4	−295	−288	−297	−300	−299	−258
$\Delta\sigma$	11.09	8.27	10.67	10.89	10.11	
Coupling Constants (Hz)						
^1J:						
C7,H7	129.1	129.8	130.2	128.6	129.9	130.7
C8,H8	128.7	129.3	130.7	128.1	129.8	130.4
C5,H5	140.1	141.8	144.6	142.1	144.1	146.8
C2,C3	52.1	54.5	45.4	51.9	46.1	52.4
C5,C6	60.6	63.7	58.7	60.1	58.6	58.9
C2,C7	32.7	34.6	33.7	32.0	31.8	33.7
C5,C8	33.7	35.6	32.6	32.2	32.1	34.4
C2,N1	−4.6	−5.0	−6.0	−0.2	−0.6	−9.2
C3,N4	−18.4	−19.4	−23.0	−18.4	−22.5	−17.0
C5,N4	−11.8	−12.6	−10.8	−12.3	−11.3	−11.5
^2J:						
C2,H7	−2.6	−3.2	−2.6	−2.8	−2.6	−4.4
C5,H8	−2.9	−3.5	−2.9	−3.0	−3.0	−4.5
C7,H2	−2.6	−3.0	−2.8	−2.5	−2.4	−3.9
C8,H5	−4.7	−5.0	−3.8	−4.1	−3.7	−4.2
C6,H5	−6.2	−6.9	−5.7	−6.3	−5.7	−4.2
C6,C8	0.2	0.1	−1.2	−0.6	−1.3	1.1
C7,N1	−0.5	−0.5	−0.4	−0.5	−0.5	1.3
C2,N4	−11.1	−11.6	−8.5	−11.0	−9.2	−6.9
C6,N4	−1.0	−1.0	−1.1	−1.2	−1.0	−1.1
^3J:						
H2,H7	7.0	6.7	7.2	6.9	7.2	7.1
H5,H8	7.3	7.1	7.0	7.2	7.1	7.0
C3,H7	4.3	4.4	4.6	4.2	4.3	4.4
C6,H8	4.4	4.5	4.7	4.4	4.6	4.5
N1,H7	−3.8	−3.8	−3.7	−3.8	−4.0	−3.1
N4,H8	−2.8	−2.8	−2.9	−2.5	−2.9	−3.1
C2,C5	2.4	2.5	2.2	2.5	2.2	2.1
ΔJ	1.4	1.5	1.4	1.6	1.7	

[a] $\Delta\sigma$ and ΔJ are average absolute deviations. Isotropic shielding values of (31.83, 183.45, and −180.16 ppm, for the H, C, and N atoms, respectively) and (31.53, 181.10, and −194.36 ppm) were used for the IGLOII and IGLOIII computations.

The scalar couplings calculated using different approaches (the lower part of Table 3) also vary rather moderately. Like the shifts, the average absolute deviations ranging in a narrow interval of 1.4−1.7 Hz do not favor any particular method. No preferential approach was indicated even by a decomposition of the statistics into the absolute average deviations for the ^1J, ^2J, and ^3J couplings (not shown). The final precision seems to be an internal property of the B3LYP functional and perhaps the DFT methodology;[32] it was also discussed in previous works.[11,33] An occasional generalization, however, can be made. Particularly, the application of the PCM method improves some one-bond couplings (^1J(C2,C3) and ^1J(C3,N4)) for both the DFT and MP2 geometries. On the contrary, some couplings deviate more from the experiment upon the application of PCM (^2J(C2,N4) and ^1J(C5,H5)). As expected, the charged and polar groups are the most sensitive to the PCM−vacuum environment change.

Behavior similar to that of the calculated cation chemical shifts was also observed for the anion (Table 1s in the SI), although here the situation is complicated by the influence of the NH_2 group rotation. Indeed, a relatively large dispersion of the calculated shifts appeared for the nuclei in the vicinity of the amine group. For the three NH_2 rotamers of the conformer **A**, the calculated shifts are, for example, 24.7, 24.7, and 29.5 ppm for carbon C7, 63.3, 63.9, and 63.7 ppm for carbon C2, and 193.8, 190.7, and 190.6 ppm for carbon C3. The dispersion is even bigger for the conformer **C**: 27.1, 25.9, and 25.7 ppm for carbon C7, 61.9, 62.9, and 65.5 ppm for carbon C2, and 190.4, 191.5, and 193.7 ppm for carbon C3.

AA Chemical Shifts. Similarly as for the alanine, the computed and experimental chemical shift changes for the charged AA forms (related to AA^{ZW}, Figure 6) confirm that the theory can reproduce the experiment on average but with a limited accuracy. Extreme changes are usually better reproduced than the small ones. Especially the hydrogen shift changes are smaller than 0.5 ppm, with the exception of H1 (experimental values are $\sigma = 6.16$ ppm for AA^{ZW} and 8.09 ppm for AA$^+$)

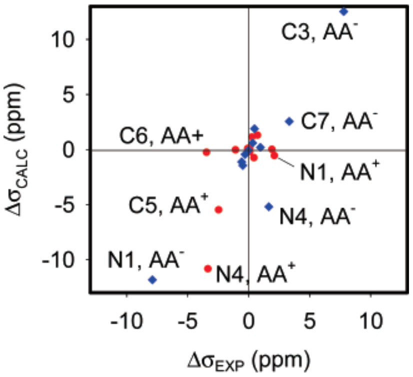


Figure 6. Comparison of calculated and experimental pH chemical shift changes (for the lowest-energy AA^+ and AA^- conformers, with respect to the zwitterionic form AA^{ZW}). For the biggest changes, the corresponding atoms are indicated in the plot.

and not easily reproducible. The amide NH group is also problematic: the computation overestimates the observed nitrogen (N4) shift both in the anion and cation. On the other hand, the amide carbon (C3) change in AA^- is easily reproduced; it is clearly caused by the loss of charge at the amine group. Large charge-induced shift variations (exceeding 1 ppm) have also been measured for the C5, C6, and N1 AA^+ atoms, all indicated in Figure 6. The calculated carbon shifts (conformer **A** in cation, zwitterion, and anion) of C3 (182.6, 181.3, 193.8 ppm), C5 (56.4, 61.9, 60.5 ppm), C6 (192.2, 192.5, 194.4 ppm), and C7 (19.4, 19.3, 22.6) thus nicely correspond to the trend observed under the pH change experimentally: C3 (173.5, 172.8, 180.6 ppm), C5 (51.6, 54.1, 53.6 ppm), C6 (179.1, 182.5, 183 ppm), and C7 (22.0, 22.2, 24.7 ppm). This reflects a general tendency of the ab initio/DFT computation to reproduce relative values of NMR parameters with a higher accuracy than for absolute ones.[34]

The accuracy of the calculated shifts (Table 2s in Supporting Information), however, is not sufficient to discriminate between the individual AA conformers **A**−**F**. For example, the absolute overall deviations from the experiment range within 11.1−11.7, 11.8−16.5, and 11.3−11.9 ppm (for the cation, zwitterion, and anion, respectively) only. All three AA forms thus behave similarly. The variations of the calculated NMR shifts should then be attributed to several factors, such as the reaction of the PCM continuum to the charge redistribution in the charged peptide forms, and do not directly reflect detailed conformational changes.

AA Spin−Spin Coupling. As documented in Figure 7, the coupling constant pH variation can be reproduced with the computation similarly as the shifts. Also here, smaller changes are less reliably calculated than the bigger ones, and the theoretical values concerning the vicinity of the charged residues are less accurate. The change of ^1J(C5,H5) in AA^+, for example, was predicted at the opposite direction. The magnitudes of the one-bond (^1J) constants change most, but the biggest relative changes can be found between the vicinal and geminal couplings (^2J,^3J).

Unlike for the shifts, calculated coupling constants provide useful information on AA conformation. This can be seen for the anion AA^- in Table 4, where the measurable coupling constants calculated for 10 conformers are listed and compared to the experiment. We can see that the results are consistent with the estimated relative conformer energies: The coupling constants calculated for the energetically inconvenient conformers (**B**, **D**−**F**) significantly deviate from the experimental values, which is also indicated by the average absolute deviations from the experiment listed at the bottom of the Table. The conformer **C** alone exhibits the lowest average deviations in the constants, whereas the **A** conformer is more preferred energetically. However, this can be explained by the conformational equilibrium. Indeed, the Boltzmann averaging, taking into account all the conformers **A**−**F** including the NH_2 group rotamers, provides a reasonably low average deviation of the couplings. Obviously, as discussed above and observed in previous works,[11,33] a future improvement of the computational model is desirable because more accurate theoretical constants can lead to a better discrimination between the peptide conformers.

For example, the NMR spectra would be sensitive to the NH_2 group rotation as the constants computed for the three rotamers differ significantly (for conformer **A**, the calculated ^1J(C2,H2) are 135.9, 134.0, and 136.8 Hz, ^1J(C2,C3) 48.3, 53.4, and 53.4 Hz, ^2J(N1,H2) −3.3, 0.7, and −3.9 Hz, etc.). The NH_2 rotation, however, influences the couplings only locally with remote atomic groups not being affected: The constants ^1J(N4,H4)

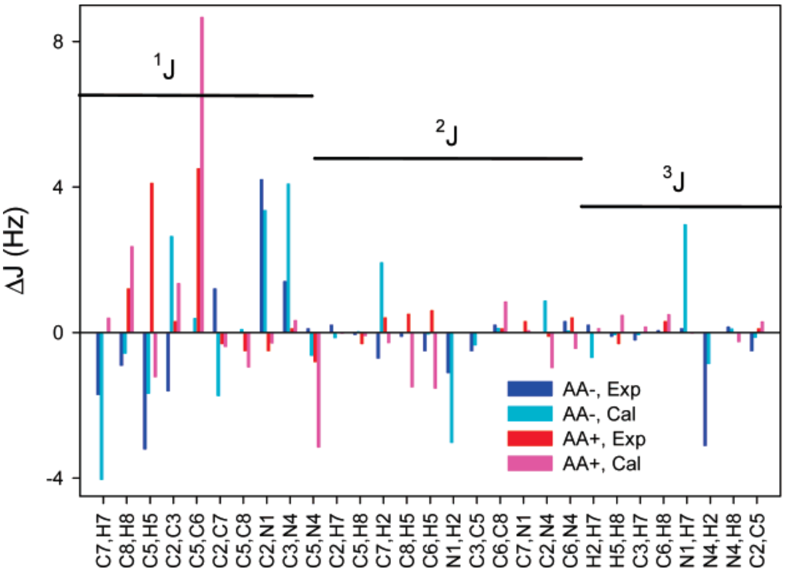


Figure 7. Computed and experimental pH-induced changes of spin−spin coupling constants (for the lowest-energy conformer of AA^+ and AA^-, with respect to that of the zwitterionic form AA^{ZW}).

TABLE 4: Spin−Spin Coupling Constants (Hz) Calculated for 10 AA$^-$ Conformers (A−F)[a] and Comparison with Experimental Values

	A	A′	A″	B	C	C′	C″	D	E	F		
η[b]	0.452	0.119	0.018	0.000	0.240	0.125	0.041	0.000	0.005	0.000	avg[c]	exptl[d]
J^1:												
C7,H7	125.0	124.7	125.0	125.0	125.5	124.9	125.5	124.9	123.9	123.7	125.0	129.0
C8,H8	125.8	125.8	125.7	125.0	125.7	125.6	125.8	124.8	125.9	125.1	125.7	128.3
C2,H2	135.9	136.8	134.0	135.4	137.7	130.6	137.8	130.5	136.2	135.1	135.7	142.2
C5,H5	139.6	139.6	139.5	127.6	139.2	138.8	139.6	127.6	140.1	127.6	139.4	139.5
C2,C3	48.3	53.4	53.4	47.8	52.5	52.6	47.1	52.5	48.8	48.2	50.5	50.5
C5,C6	52.3	52.3	52.4	52.3	52.5	52.4	52.1	52.5	51.3	52.4	52.3	54.4
C2,C7	36.5	31.3	35.9	36.0	34.0	35.3	35.2	35.6	36.6	36.6	35.0	35.2
C5,C8	34.6	34.7	34.6	39.2	34.6	34.9	35.0	39.0	35.2	39.2	34.7	34.9
C2,N1	−1.2	−1.0	−2.6	−1.2	−2.6	−2.0	−1.2	−2.2	−0.7	−0.8	−1.6	−4.5
C3,N4	−13.6	−14.6	−14.5	−13.8	−16.0	−15.9	−15.4	−16.7	−13.7	−13.9	−14.7	−15.7
C5,N4	−9.2	−9.3	−9.3	−9.0	−9.9	−9.9	−9.6	−9.8	−9.3	−9.1	−9.5	−10.6
J^2:												
C2,H7	−2.8	−2.8	−2.8	−2.8	−2.9	−3.1	−2.9	−3.1	−2.9	−3.0	−2.9	−4.2
C5,H8	−2.8	−2.8	−2.8	−3.3	−2.8	−2.8	−2.8	−3.3	−2.9	−3.3	−2.8	−4.25
C7,H2	−3.3	−0.4	−0.8	−3.5	−3.1	−3.5	−4.8	−3.8	−2.9	−3.1	−3.0	−5.0
C8,H5	−3.2	−3.2	−3.2	−3.8	−3.2	−3.2	−3.2	−3.5	−3.0	−3.8	−3.2	−4.8
C3,H2	−1.1	−3.9	−0.8	−1.1	−3.2	−6.1	−3.6	−5.9	−4.5	−4.2	−2.7	−5.0
C6,H5	−4.7	−4.7	−4.7	−5.9	−4.6	−4.6	−4.7	−5.9	−4.6	−5.9	−4.7	−5.3
N1,H2	−3.3	−3.9	0.7	−3.3	1.2	−1.2	−4.4	−0.7	−2.4	−2.4	−2.0	−1.1
C3,C5	−0.9	−0.6	−0.5	−0.8	−0.3	−0.3	−0.5	−0.1	−1.0	−0.9	−0.6	−0.5
C6,C8	−0.6	−0.6	−0.5	0.7	−0.6	−0.5	−0.5	0.7	−0.8	0.6	−0.6	1.2
C2,N4	−9.1	−9.3	−9.6	−9.3	−6.6	−6.4	−7.5	−6.5	−7.8	−7.6	−8.1	−6.8
C6,N4	−0.5	−0.5	−0.5	0.2	−0.7	−0.6	−0.6	0.2	−0.6	0.2	−0.6	−1.2
J^3:												
H2,H7	6.7	6.2	6.2	6.8	6.8	6.8	7.1	6.9	7.0	6.9	6.7	7.3
H5,H8	6.8	6.8	6.8	7.0	6.8	6.9	6.8	7.0	6.8	7.0	6.8	7.2
C3,H7	4.1	4.1	4.5	4.0	4.3	3.9	4.0	4.0	4.0	4.0	4.1	4.2
C6,H8	3.9	3.9	3.9	3.5	3.9	3.9	3.9	3.5	4.0	3.5	3.9	4.25
N1,H7	−3.7	−0.8	−3.4	−3.7	−2.9	−1.0	−3.7	−1.1	−3.8	−3.7	−2.8	−3.0
N4,H2	−0.8	−1.2	−1.2	−0.7	−0.7	−0.6	−1.1	−0.4	−1.2	−1.1	−0.8	−2.1
N4,H8	−2.5	−2.5	−2.5	−2.5	−2.5	−2.5	−2.5	−2.4	−2.5	−2.5	−2.5	−2.95
C2,C5	1.7	2.0	2.0	2.1	1.3	1.3	1.4	1.7	1.3	1.6	1.6	1.5
ΔJ[e]	1.6	1.7	1.7	2.1	1.2	1.4	1.2	1.9	1.5	2.0	1.3	0.0

[a] See Table 1 for the definition of the conformers. [b] Conformer ratios were estimated from the Boltzmann factor at 300 K. [c] The Boltzmann-weighted average. [d] pH = 12. [e] Average absolute deviations.

(−91.5, −91.4, and −91.6 Hz for the three rotamers), ^2J(C5,H4) (3.2, 3.2, and 3.1 Hz), or ^3J(H4,H5) (7.1, 7.4, and 7.2 Hz) are rather insensitive to the rotation.

Also the computed coupling constants of the AA$^+$ form listed in Table 5 agree best with the experiment for the energy-preferred conformer **A**. Average deviations of the **C** conformer are low as well, but this form can be excluded on the basis of the energy estimation. For the cation, especially the vicinal couplings (^3J) seem to be computed with significant errors. Fortunately, for AA$^+$, the number of conformers that can contribute to the averaged observed J's is smaller than for AA$^-$. In fact, the conformer **A** seems to be clearly dominant (~90%), with Boltzmann population of the conformer **B** estimated from the equilibrium geometry being about 10%. Given the narrow potential well (Figure 2) in comparison with **A**, the population of conformer **B**, obtained through a complete integration over the two-dimensional potential energy surface, would be even smaller. Additionally, the average error ($\Delta J = 1.9$ Hz, see Table 5) of the couplings for the **B** form is much higher than that for **A** ($\Delta J = 1.4$ Hz).

Finally, we can focus our attention on the conformational sensitivity of individual spin−spin coupling constants represented by the average absolute deviations plotted in Figure 8. In spite of the errors of the computed couplings, some constants clearly exhibit larger variations under the conformational change, and thus the NMR technique along with the quantum computation may be able to discriminate between peptide conformers. The ^1J(C5,H5) coupling is the most sensitive one, particularly to the change of the close φ angle, and can thus monitor the ratios of the (**A**, **C**) and (**B**, **D**) conformer classes (cf. Tables 4 and 5). The ψ torsion has little impact on this coupling. Similarly, the analogous ^1J(C2,H2) coupling constant is more sensitive to the ψ angle, which makes it possible for these two couplings alone in principle to determine the AA secondary structure. To be able to generalize the results, however, one has to realize also the dependence on the molecular charge (the experimental ^1J(C5,H5) constant is 146.8, 142.7, and 139.5 Hz for the cation, zwitterion, and anion, respectively), and in larger peptides similar variation can be expected for various amino acid side chains. The ^1J(C5,C6) coupling, for example, is predominantly driven by the molecular charge (the measured values for the cation, zwitterion, and anion are 58.9, 54.4, and 54.4 Hz, which are actually very well reproduced by the calculation as 60.6, 51.9, and 52.3 Hz; see Tables 4 and 5 and ref 11), and the conformational variance does not exceed 0.7 Hz.

The three-bond couplings, however, are more important for the peptide structural determination than the one- and two-bond interactions.[1,2,12,13] Therefore, their dependence on other factors than the torsion angle, such as charges of close molecular groups, is of paramount importance for peptide chemistry. As an example, two constants, ^3J(H4,H5) and ^3J(N4,H2), were selected and their dependence on the main-chain torsion angle plotted in Figure 9. Calculated curves for the AA cation, anion, and zwitterions are compared to the empirical Karplus-type curves derived in the literature on the basis of theoretical and

TABLE 5: Spin−Spin Coupling Constants (Hz) Calculated for Four AA$^+$ Conformers and a Comparison with the Experiment[a]

	A	B	C	D	
η	0.91	0.09	0.00	0.00	exptl
^1J:					
C7,H7	129.1	129.2	128.9	128.9	130.7
C8,H8	128.7	127.7	128.7	127.5	130.4
C5,H5	140.1	132.5	140.8	133.2	146.8
C2,C3	52.1	51.1	54.6	53.4	52.4
C5,C6	60.6	59.2	60.8	59.5	58.9
C2,C7	32.7	32.7	32.1	32.2	33.7
C5,C8	33.7	38.9	33.5	39.1	34.4
C2,N1	−4.6	−4.6	−0.7	−0.8	−9.2
C3,N4	−18.4	−17.6	−16.4	−15.6	−17.0
C5,N4	−11.8	−9.4	−11.8	−9.1	−11.5
^2J:					
C2,H7	−2.6	−2.6	−2.5	−2.5	−4.4
C5,H8	−2.9	−3.3	−2.9	−3.3	−4.5
C7,H2	−2.6	−2.7	−3.1	−3.1	−3.9
C8,H5	−4.7	−4.2	−4.7	−3.9	−4.2
C6,H5	−6.2	−7.2	−6.3	−6.8	−4.2
C6,C8	0.2	1.1	0.1	1.1	1.1
C7,N1	−0.5	−0.5	−0.4	−0.4	1.3
C2,N4	−11.1	−10.9	−8.7	−8.5	−6.9
C6,N4	−1.0	0.3	−1.2	0.5	−1.1
^3J:					
H2,H7	7.0	7.0	7.2	7.2	7.1
H5,H8	7.3	7.0	7.3	6.9	7.0
C3,H7	4.3	4.2	4.5	4.4	4.4
C6,H8	4.4	4.5	4.4	4.5	4.5
N1,H7	−3.8	−3.8	−4.1	−4.1	−3.1
N4,H8	−2.8	−2.6	−2.7	−2.7	−3.1
C2,C5	2.4	2.6	1.3	1.5	2.1
ΔJ	1.4	1.9	1.5	2.0	

[a] The symbols have the same reference as in Table 4.

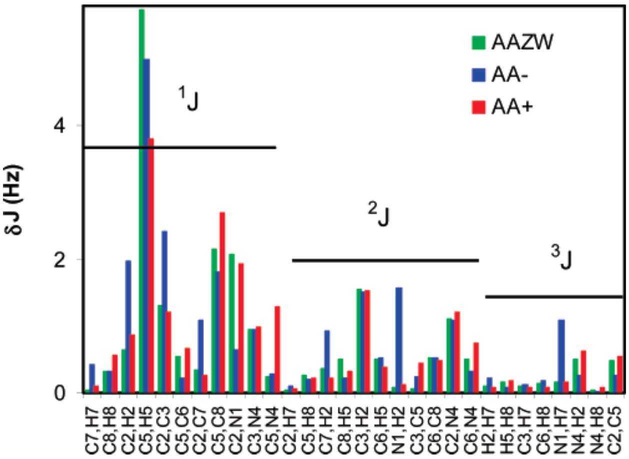


Figure 8. Computed sensitivities of selected spin−spin coupling constants to the conformational change (average absolute deviations from the average over individual conformers **A**−**F** are plotted for each form).

experimental data. Clearly, for ^3J(H4,H5), the angular dependence does not rely significantly on the molecular charge and nicely corresponds to the two approximations proposed previously. Within $\varphi \approx -180...-30°$, the cationic AA$^+$ ^3J(H4,H5) curve somewhat deviates from the anion and zwitterion, which can be explained by the vicinity of the COO$^-$ group, protonated in the cation. Some dispersion occurs also around $\varphi \approx 150°$; this is, however, rather minor with respect to the principal conformational dependence.

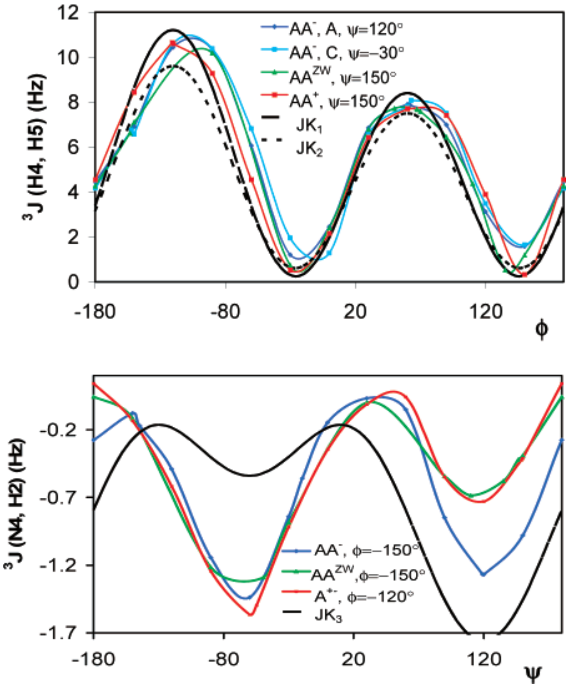


Figure 9. Calculated dependence of the ^3J(H4,H5) (top) and ^3J(N4,H2) (bottom) vicinal spin−spin coupling constants on the encompassed torsion angle for the three charged AA forms and a comparison to the JK$_1$,[35] JK$_2$,[36] and JK$_3$[37] semiempirical Karplus curves.

Apparently, the ^3J(N4,H2) coupling (lower part of Figure 9) is more sensitive to the molecular charge and deviates more from the previously proposed curve. As expected, the anion AA$^-$ curve deviates from the cation and zwitterion because of the deprotonation of the NH$_3^+$ group close to the rotating bond. However, a closer inspection reveals that the absolute coupling dispersion is like for the previous case, since the ^3J(N4,H2) constant varies in a much narrower range than ^3J(H4,H5).

Conclusions

On the basis of the DFT computations (BPW91/6-311++G**) of the two-dimensional potential energy surfaces, we were able to estimate the conformational behavior of the AA dipeptide under the pH changes. Whereas the neutral zwitterionic form AA^{ZW} and the cation (AA$^+$) adopt similar conformations of the main peptide chain (φ, ψ angles), the anion (AA$^-$) exists in two forms differing by the ψ-angle values. The forms are approximately equally populated in aqueous solutions at room temperature. The anion main chain folding is more complex than for the other forms because of the influence of the NH$_2$ group, which can serve both as a hydrogen donor and acceptor in an intramolecular hydrogen bonding. The results of the analysis of the potential energy surfaces are in agreement with both experimental and calculated NMR chemical shifts and spin−spin coupling constants. The NMR parameters could be calculated with a limited accuracy, but the pH dependence of the chemical shifts for the dipeptide as well as for the alanine monomer could be explained on the basis of the theory. Furthermore, the comparison of the experimental and calculated coupling constants is consistent with the energetic analysis.

Acknowledgment. The work was supported by the Czech Science Foundation (Grant Nos. 203/06/0420 and 202/07/0732), the Grant Agency of the Academy of Sciences (A4005507020, A400550701), and the Ministry of Education, Youth and Sports (Grant No. LC512).

Supporting Information Available: Cartesian coordinates, absolute energies of equilibrium structures, and additional computational and experimental details. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Kessler, H.; Seip, S. In *Two-Dimensional NMR Spectroscopy, Application for Chemists and Biochemists*; Croasmun, W. R., Carlson, R. M. K., Eds.; VCH Publ.: New York, 1994; p 619.

(2) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.

(3) Englander, S. W.; Mayne, L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 243.

(4) Evans, J. N. S. *Biomolecular NMR Spectroscopy*; Oxford University Press: Oxford, 1995.

(5) William, M. P.; Waltho, J. P. *Chem. Soc. Rev.* **1992**, *21*, 227.

(6) Wolinski, K.; Hilton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251.

(7) Sychrovský, V.; Gräfenstein, J.; Cremer, D. *J. Chem. Phys.* **2000**, *113*, 3530.

(8) Helgaker, T.; Watson, M.; Handy, N. C. *J. Chem. Phys.* **2000**, *113*, 9402.

(9) Barone, V.; Peralta, J. E.; Contreras, R. H.; Snyder, J. P. *J. Phys. Chem. A* **2002**, *106*, 5607.

(10) Bouř, P.; Sychrovský, V.; Maloň, P.; Hanzlíková, J.; Baumruk, V.; Pospíšek, J.; Buděšínský, M. *J. Phys. Chem. A* **2002**, *106*, 7321.

(11) Bouř, P.; Buděšínský, M.; Špirko, V.; Kapitán, J.; Šebestík, J.; Sychrovský, V. *J. Am. Chem. Soc.* **2005**, *127*, 17079.

(12) Karplus, M.; Weaver, D. L. *Nature* **1976**, *260*, 404.

(13) Perere, S. A.; Bartlett, R. J. *Magn. Reson. Chem.* **2001**, *39*, S183.

(14) Laws, D. D.; deDios, A. C.; Oldfield, E. *J. Biomol. NMR* **1993**, *3*, 607.

(15) Birn, J.; Poon, A.; Mao, Y.; Ramamoorthy, A. *J. Am. Chem. Soc.* **2004**, *126*, 8529.

(16) Sebastiani, D.; Rothlisberger, U. *J. Phys. Chem. B* **2004**, *108*, 2807.

(17) Čejchan, A.; Špirko, V. *J. Mol. Spectrosc.* **2003**, *217*, 142.

(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(19) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(20) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 404.

(21) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(22) Ruud, K.; Helgaker, T.; Bak, K. L.; Jorgensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1993**, *99*, 3847.

(23) Bouř, P.; Buděšínský, M. *J. Chem. Phys.* **1999**, *110*, 2836.

(24) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(25) Kutzelnigg, W.; Fleischer, U.; Schindler, M. *NMR - Basic Principles and Progress*; Springer: Heidelberg, 1990; Vol. 23.

(26) Podolsky, B. *Phys. Rev.* **1928**, *32*, 812.

(27) Ixaru, L. *Numerical Methods for Differential Equations and Applications*; Reidel: Dordrecht, 1984.

(28) Špirko, V.; Piecuch, P.; Bludský, O. *J. Chem. Phys.* **2000**, *112*, 189.

(29) Creighton, T. E. *Proteins*; New York, 1984.

(30) Fadda, E.; Casida, M. E.; Salahub, D. R. *J. Phys. Chem. A* **2003**, *107*, 9924.

(31) Bouř, P. *J. Chem. Phys.* **2004**, *121*, 7545.

(32) Kümmel, S.; Kronik, L.; Perdew, J. P. *Phys. Rev. Lett.* **2004**, *93*, 213002.

(33) Bouř, P.; Raich, I.; Kaminský, J.; Hrabal, R.; Čejka, J.; Sychrovský, V. *J. Phys. Chem. A* **2004**, *108*, 6365.

(34) *The Encyclopedia of Computational Chemistry*; Schleyer, P. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley & Sons: Chichester, 1998.

(35) Case, D. A.; Scheurer, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **2000**, *122*, 10390.

(36) Schmidt, J. M.; Blümel, M.; Löhr, F.; Rüterjans, H. *J. Biomol. NMR* **1999**, *14*, 1.

(37) Demarco, A.; Llinas, M.; Wührich, K. M. *Biopolymers* **1978**, *17*, 2727.

J. Phys. Chem. B **2009,** *113,* 5273–5281

Theoretical Study of the Effective Chemical Shielding Anisotropy (CSA) in Peptide Backbone, Rating the Impact of CSAs on the Cross-Correlated Relaxations in L-Alanyl-L-alanine

Ladislav Benda,*,† Petr Bouř,† Norbert Müller,*,‡ and Vladimír Sychrovský*,†

Institute of Organic Chemistry and Biochemistry v.v.i., Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Praha 6, Czech Republic, and Institute of Organic Chemistry, Johannes Kepler University, Altenbergerstrasse 69, 4040 Linz, Austria

Received: December 1, 2008; Revised Manuscript Received: February 3, 2009

The dependence of the effective chemical shielding anisotropy (effective CSA, $\Delta\sigma^{\mathrm{eff}}$) on the φ and ψ peptide backbone torsion angles was calculated in the L-alanyl-L-alanine (LALA) peptide using the DFT method. The effects of backbone conformation, molecular charge including the cation, zwitterion, and anion forms of the LALA peptide, and the scaling taking into account the length of the dipolar vector were calculated for the effective CSAs in order to assess their structural behaviors and to predict their magnitudes which can be probed for the β-sheet and α-helix backbone conformations via measurement of the cross-correlated relaxation rates (CCR rates). Twenty different CSA−DD cross-correlation mechanisms involving the amide nitrogen and carbonyl carbon chemical shielding tensors and the $C_\alpha H_\alpha$ (α-carbon group), NH_N (amide group), $C_\alpha H_N$, NH_α, $C'H_\alpha$, and $C'H_N$ ($\alpha = \alpha1, \alpha2$) dipolar vectors were investigated. The $X-C_\alpha H_\alpha$ ($X = N, C'$; $\alpha = \alpha1, \alpha2$) cross-correlations, which had already been studied experimentally, exhibited overall best performance of the calculated effective CSAs in the LALA molecule; they spanned the largest range of values upon variation of the ψ and φ torsions and depended dominantly on only one of the two backbone torsion angles. The $X-NH_N$ ($X = N, C'$) cross-correlations, which had been also probed experimentally, depended on both backbone torsions, which makes their structural assignment more difficult. The $N-NH_{\alpha2}$ and $N-C'H_{\alpha1}$ cross-correlations were found to be promising for the determination of various backbone conformations due to the large calculated range of the scaled effective CSA values and due to their predominant dependence on the ψ and φ torsions, respectively. The 20 calculated dependencies of effective CSAs on the two backbone torsion angles can facilitate the structural interpretation of CCR rates.

Introduction

During the past decades, high-resolution nuclear magnetic resonance (NMR) spectroscopy has become an indispensable tool in structural studies of molecules.[1,2] Applications of NMR spectroscopy for obtaining valuable structural information for biologically important molecules such as proteins and nucleic acids have grown to an unprecedented extent.[3−7]

Despite the marvelous progress in experimental technique, the NMR-resolved molecular structures may suffer from natural restrictions. Typically, the $^1H-^1H$ proton−proton distances probed with NOE (nuclear Overhauser effect) are not sufficient for an unambiguous determination of molecular structure. The NMR chemical shifts and scalar coupling constants can be assigned only to geometric parameters that are confined in the vicinity of the probed nuclei such as bond length or bond/torsion angle,[3] contrary to the long-range structural restraints by residual dipolar couplings[8] or cross-correlated relaxation rates (CCR rates).[9] High resolution of the NMR-determined structures thus still remains challenging and requires methods combining various NMR parameters assigned to specific structural descriptors. New reliable geometry restraints are therefore needed in NMR studies of global structural features, such as the protein tertiary structure.

The CCR is a specific NMR phenomenon known for a long time that was relatively recently exploited in dynamic studies of biomolecules[10−17] and for obtaining geometry restraints in proteins[9,18−25] and in nucleic acids.[26−33] The cross-correlation between chemical shielding anisotropy and dipole−dipole interaction (CSA−DD) was also used for the determination of the NMR chemical shielding tensor (σ-tensor) principal components and their relative orientation with respect to the molecular frame even in the liquid state.[34−37] The applicability of the CCRs exceeds the limited range of biomolecular studies.[38,39]

Measurement of the CSA−DD cross-correlated relaxation rates probes both the dynamics (specific fluctuations of molecular motions) and the σ-tensor (see Theory). While reliable prediction of the CCR dynamic component might be a challenging task,[40,41] the structural dependence of the σ-tensor and its orientation with respect to the dipolar vector can be obtained with many calculation methods. The theoretical modeling of NMR spectroscopy parameters using modern quantum-chemistry methods has become a practical tool that complements NMR experiments.[42] The theoretical analysis of NMR parameters can provide fast and deep insight into their dependence on molecular structure including the impact of specific solvent effects and molecular motion.[42] To the best of our knowledge, only a limited number of theoretical studies dealing with the ab initio modeling of NMR parameters appertained to the CCR phenomenon is currently available.[43−45] The goal of this theoretical study was

---

* Corresponding author.

† Academy of Sciences of the Czech Republic.

‡ Johannes Kepler University.

10.1021/jp8105452 CCC: $40.75 © 2009 American Chemical Society
Published on Web 03/20/2009

to shed light on the protocols used for the structural interpretation of relevant CSA−DD cross-correlations which can be used in probing the structure of peptide backbone. Some of the CSA−DD cross-correlation mechanisms investigated in this work were also probed in recent NMR studies.[18−20,24,37,46] However, the majority of them were theoretically modeled in this work probably for the first time.

Our study was focused on the calculation of the effective CSAs for different CSA−DD cross-correlation mechanisms involving the σ-tensors of amide nitrogen (^{15}N) and carbonyl carbon (^{13}C′) in the L-alanyl-L-alanine (LALA) peptide. The calculated dependence of NMR properties on backbone conformation was further extended by modeling the effects of protonation and deprotonation of the LALA terminal groups occurring at different pH values. The three forms of the LALA molecule (cation, zwitterion, and anion) were characterized in our recent joint computational and experimental NMR work.[47,48] Benchmark conformational analysis of short alanine peptides carried out with the molecular dynamics was successfully correlated with complex NMR data.[49] In the future we plan to use the structural information obtained previously for the LALA molecule with both NMR and optical spectroscopies[47,48,50] for calibration of measured CCR rates in peptides.

Theory

A comprehensive theoretical description of the CCR phenomenon is given by the Bloch−Wangsness−Redfield theory of relaxations, usually called the Redfield theory.[51] The matrix form of relaxation equations includes the autorelaxation rates on its diagonal, whereas the off-diagonal terms describe mutual correlations between different relaxation mechanisms (i.e., the CCR rates). The CCRs typically contribute by a very small amount to the overall relaxation of a nuclear spin, and the experimental accessibility of a particular CCR rate can be therefore limited.

On the assumption of isotropic overall molecular motion of a molecule that can be considered rigid on the time scale of the overall molecular motion, the CSA−DD cross-correlated relaxation rate $\Gamma_{i,jk}^{\mathrm{CSA-DD}}$ is described by the following equations:[51,52]

$$\Gamma_{i,jk}^{\mathrm{CSA-DD}} = \frac{1}{2}\left(\frac{\mu_0\hbar}{4\pi}\right)\frac{\gamma_j\gamma_k}{r_{jk}^3}\gamma_i \boldsymbol{B}_0(\Delta\sigma_{i,jk}^{\mathrm{eff}})\sum_q \boldsymbol{J}(\omega_q) \tag{1}$$

$$\Delta\sigma_{i,jk}^{\mathrm{eff}} = \sum_{n=1}^{3}\sigma_i^{nn}\left(\frac{3\cos^2\theta_{jk}^{nn} - 1}{2}\right) \tag{2}$$

$$\boldsymbol{J}(\omega_q) = \frac{2}{5}\frac{\tau_c}{1 + (\omega_q\tau_c)^2} \tag{3}$$

where $\boldsymbol{B}_0$ is the magnetic field strength, $\Delta\sigma_{i,jk}^{\mathrm{eff}}$ is the effective CSA that correlates the principal components of the σ-tensor of nucleus i with the nuclear dipolar vector interconnecting nuclei j and k, r_{jk} is the length of the dipolar vector, and θ_{jk}^{nn} is the angle between the nn-principal component of the σ-tensor and the dipolar vector. $\boldsymbol{J}(\omega_q)$ is the spectral density function describing the local magnetic field fluctuations at frequency ω_q. The other symbols in eq 1 have the meaning of usual physical constants: γ_i is the gyromagnetic ratio of nucleus i, μ_0 is the vacuum permeability, and $\hbar$ is the reduced Planck constant.

The particular form of the $\boldsymbol{J}(\omega_q)$ function depends on the physical model of molecular motion. The spectral density function in eq 3 describes the ideal case of isotropic tumbling of a rigid spherical molecule. The rate of this tumbling is characterized by the autocorrelation time τ_c which can be either determined experimentally by fitting the NMR data (T_1, T_2, and heteronuclear NOE) or modeled theoretically with the methods of molecular dynamics.[44]

The $\Delta\sigma^{\mathrm{eff}}$ essentially modulates the CCR rates (eq 1) via its angular dependence, projecting the principal components of the σ-tensor onto the dipolar vector of a particular CSA−DD mechanism.[43−45] The spectral density function $\boldsymbol{J}(\omega_q)$ and the magnetic field strength $\boldsymbol{B}_0$ can in that regard be considered as constants. The CSA−DD cross-correlation therefore probes the site-specific mutual orientation of the σ-tensor and the dipolar vector. This theoretical study models both effects that contribute to modulation of the effective CSAs: the dependence of the σ-tensor on molecular geometry that is implicitly considered in eq 2, and the projections of the σ-tensor principal components onto the dipolar vector. Although the latter effect is usually dominant, the former effect can be also significant, and models going beyond the so-called rigid tensor approximation are therefore more reliable.[45]

The experimental detection of the CSA−DD cross-correlations (generally involving three different nuclei A, B, C) requires the development of two-spin coherences in the first step followed by measurement of differential relaxation between the double- and zero-quantum coherences.[46] The CCR rates usually cannot be accessed directly since only the $\Gamma(\mathrm{A-BC}) + \Gamma(\mathrm{B-AC})$, $\Gamma(\mathrm{B-AC}) + \Gamma(\mathrm{C-AB})$, and $\Gamma(\mathrm{C-AB}) + \Gamma(\mathrm{A-BC})$ terms are observable. Therefore, in the most general case three independent NMR experiments have to be performed from which the three linked CCR rates can be extracted. However, if one of the CCR rates is much smaller than the other (e.g., due to a much smaller CSA value), it can be neglected and only one NMR experiment is then sufficient for the actual CCR rate determination. Further, in some cases the CCR rates can be obtained even from one-dimensional (1D) NMR spectra (for example, when only two different nuclei are involved as in the case of the N−NH$_{\mathrm{N}}$ cross-correlation).[37]

We used the usual definition of the chemical shielding anisotropy (CSA, $\Delta\sigma$):

$$\Delta\sigma = \sigma^{33} - \left(\frac{\sigma^{11} + \sigma^{22}}{2}\right) \tag{4}$$

where σ^{nn} are the σ-tensor principal values defined such that $\sigma^{33} > \sigma^{22} > \sigma^{11}$.

Methods

The LALA molecule consists of two L-alanine residues. Whenever unequivocal clarity was required, we used the indexes "1" and "2" (Figure 1) to distinguish between the N-terminal and the C-terminal residues of the LALA peptide, respectively. However, in the text we often omit the indexes, as in the case of the ψ and φ backbone torsions.

The backbone conformation of the LALA is determined by the ψ_1, ω_1, and φ_2 torsion angles (Figure 1). The values of the ω torsion angle are close to 180° in the majority of peptides due to the partial π-character of the peptide bond. The only geometry parameters describing the LALA backbone are thus the ψ and φ torsion angles.

For the anionic (NH$_2$ and COO$^-$ termini) and the cationic (NH$_3^+$ and COOH termini) forms of the LALA peptide both the ψ and the φ torsion angles were varied with steps of 30°,

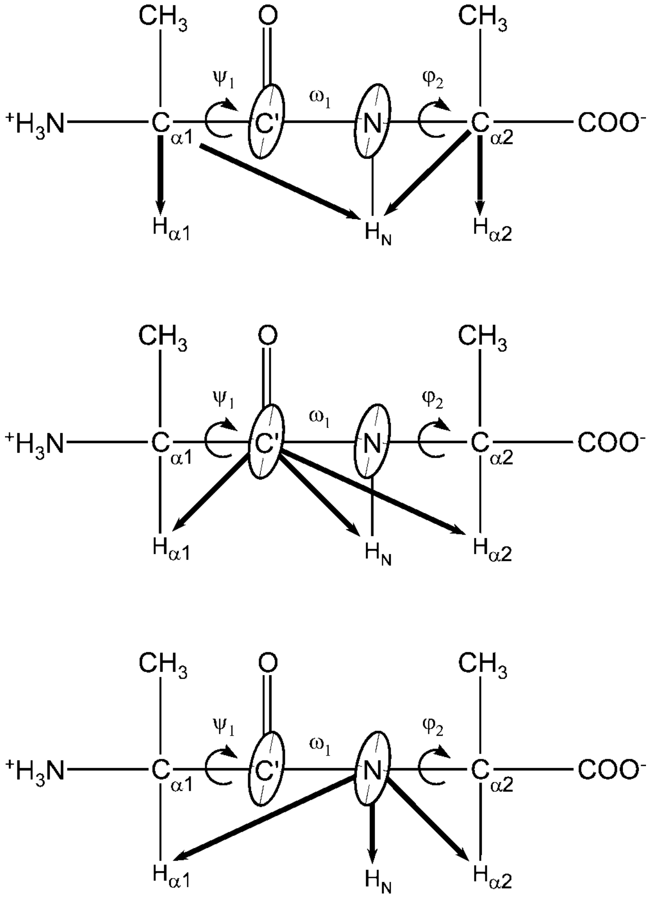


Figure 1. Schematic representation of the LALA molecule in its zwitterionic form with the indicated CSA−DD cross-correlated relaxation mechanisms between the anisotropy of σ-tensor (ellipses) and nuclear dipole−dipole vector (thick arrows).

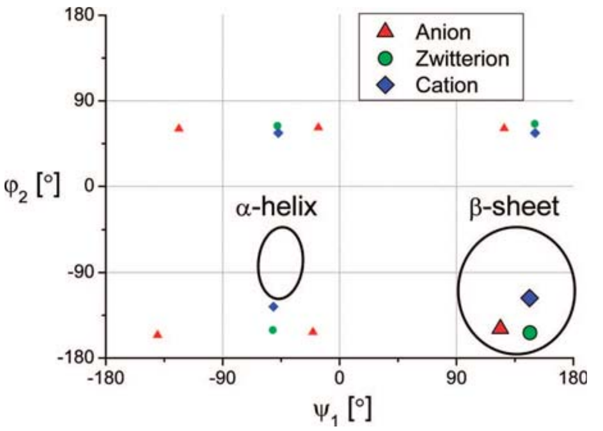


Figure 2. Localization of global (larger symbols) and local energy minima calculated for the three forms of the LALA molecule with respect to α-helix and β-sheet regions.[48]

and for each of the resulting 12 × 12 = 144 geometries all the remaining coordinates were fully relaxed by energy minimization. The calculations for the zwitterion (NH_3^+ and COO^- terminus) were performed with a larger step of 60°.[47,48]

The calculations were carried out using the density functional theory (DFT) method. The geometries and the NMR chemical shielding tensors were obtained at the BPW91[53,54]/6-311++G(d,p)[55] and the GIAO B3LYP[56,57]/IGLO-II[58] level of theory, respectively. The PCM dielectric model of water solvent was used consistently in both geometry optimization and NMR calculations. The Gaussian 03 program package[59] was used in all quantum-chemistry calculations. The numerical evaluation of eq 2 was carried out with the Matlab program package.

Results and Discussion

The φ and ψ backbone torsion angles are essential geometry parameters for describing peptide conformation. Once their values are determined in each residue along the peptide backbone, the global structural characteristics of proteins, in particular their fold, can be explicitly studied. From this point of view, the assignment of CCR rates to backbone torsion angles represents a rather versatile restraint, since in principle a plethora of CSA−DD cross-correlated relaxation rates involving different nuclei can be measured in oligopeptides and proteins.

The experimental accessibility of all CCR rates cannot be automatically taken for granted only on the basis of the calculated effective CSAs. For example, the r_{jk}^{-3} factor (eq 1) scales effectively down the CCR rates with longer dipolar vectors, i.e., those with the dipolar interaction between atoms that are not directly bonded. Also, the involvement of low-gamma nuclei reduces the magnitude of cross-correlation effects. Current experimental detection of the CCR rates in peptides was therefore carried out for the dipolar vectors corresponding to covalently bonded atoms (see the r_{DD} column and references in Table 2) and for the high-gamma nuclei preferentially. It should be also mentioned that from the experimental point of view the magnetization transfers necessary for detection of the CCR rates with the atoms appearing in the CSA−DD mechanism separated by more bonds can be rather inconvenient and inefficient. Further, the spectral density function (eq 3) reflects the specificity of molecular motions with respect to a given CCR mechanism and its actual form together with the approximations usually made can significantly alter the interpretation of measured CCR rates.

We therefore found it useful to model the effective CSAs for a certain class of the CSA−DD cross-correlations which actually complemented the CCR rates that were measured in recent experimental studies (as referenced in Table 2). The cross-correlations studied in this work involved the amide nitrogen and carbonyl carbon CSAs and the $C_{\alpha1}H_{\alpha1,}$ $C_{\alpha2}H_{\alpha2}$, $C_{\alpha1}H_N$, $C_{\alpha2}H_N$, NH_N, $NH_{\alpha1}$, $NH_{\alpha2}$, $C'H_N$, $C'H_{\alpha1}$, and $C'H_{\alpha2}$ dipolar vectors (Figure 1). The dipolar vectors systematically included the H_α and H_N hydrogen atoms on one side and the backbone C_α carbon, carbonyl carbon, and amide nitrogen on the other side. These dipolar vectors possess characteristic orientations with respect to the peptide backbone (Figure 1), and their lengths range from 1.0 to 3.0 Å. The calculated dependencies of the $\Delta\sigma^{eff}$ on the ψ and φ torsion angles for all three forms of the LALA peptide (cation, zwitterion, and anion) can be found in the Supporting Information.

Many aspects concerning the effective CSAs, and consequently also the CCR rates, could be theoretically modeled and analyzed. For practical applications to structural studies, a smooth dependence of the $\Delta\sigma^{eff}$ on the assigned geometry parameter(s) would be preferential. Even higher preference could be given to the cross-correlations with notable 1D character of the corresponding $\Delta\sigma^{eff}$, i.e. to those depending dominantly on only one geometry parameter. Strong differentiation between typical backbone conformation archetypes, such as α-helix and β-sheet, would be another obvious advantage.

We analyzed the calculated $\Delta\sigma^{eff}$ values with such complex criteria in order to determine whether (i) the α-helix and the β-sheet conformers of the LALA peptide could be distinguished due to sizable differences in the $\Delta\sigma^{eff}$ values (Tables 1 and 2) and (ii) the overall character of the $\Delta\sigma^{eff}$ (ψ,φ) surfaces was

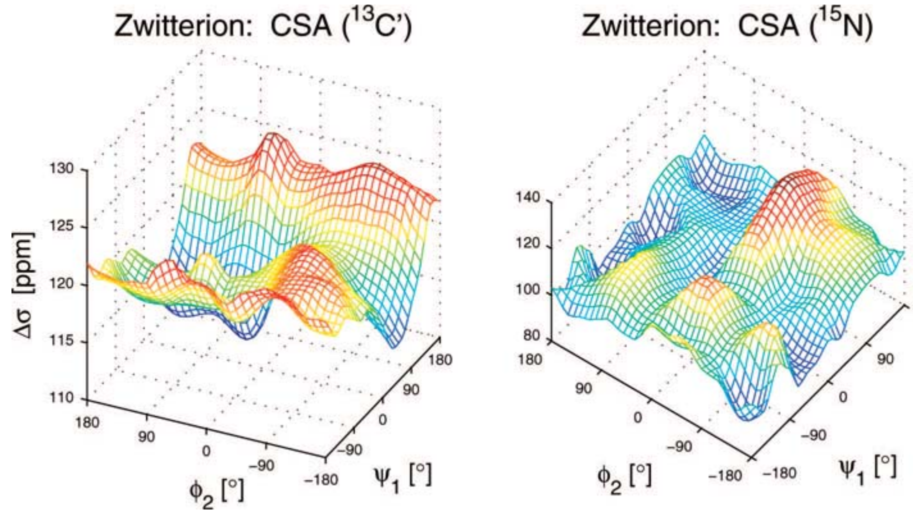


Figure 3. Dependence of the chemical shielding anisotropies $\Delta\sigma_N$ and $\Delta\sigma_{C'}$ on the ψ_1 and φ_2 backbone torsion angles calculated for the LALA zwitterion.

TABLE 1: Geometry[48] and NMR[a] Parameters Calculated for the β-Sheet and α-Helix Backbone Conformation of the LALA Peptide

	β-sheet[b]			α-helix[c]		
	anion	zwitterion	cation	anion	zwitterion	cation
ψ [deg]	119	147	149	−30	−30	−30
φ [deg]	210	207	239	−90	−90	−90
σ_N [ppm]	91.0	90.1	100.9	98.5	94.6	106.0
$\Delta\sigma_N$ [ppm]	125.0	108.8	105.7	117.6	101.1	109.1
$\sigma_{C'}$ [ppm]	−9.6	−0.2	−1.5	−12.4	−0.5	−0.9
$\Delta\sigma_{C'}$ [ppm]	137.4	123.5	124.4	144.8	122.5	123.8
			$\Delta\sigma^{eff}$ [ppm]			
CSA−DD						
$C'-C_{\alpha1}H_{\alpha1}$	−26.7	24.0	30.6	−32.3	5.9	1.7
$C'-C_{\alpha2}H_{\alpha2}$	32.7	44.7	2.6	−3.7	12.1	21.0
$C'-C_{\alpha1}H_N$	−107.9	−84.7	−94.6	−96.9	−88.6	−97.6
$C'-C_{\alpha2}H_N$	−59.6	−79.2	−70.8	−69.8	−69.8	−60.2
$C'-NH_N$	−18.4	−8.2	−3.3	−31.5	0.0	2.7
$C'-NH_{\alpha1}$	−93.1	−110.9	−108.3	−103.2	−102.4	−109.7
$C'-NH_{\alpha2}$	−36.1	−53.4	−58.7	−56.8	−62.3	−48.0
$C'-C'H_N$	−48.6	−13.5	−17.7	−51.1	−16.5	−23.1
$C'-C'H_{\alpha1}$	−37.9	−57.3	−47.2	−99.3	−105.4	−109.6
$C'-C'H_{\alpha2}$	−91.7	−101.4	−118.5	−96.6	−118.8	−112.6
$N-C_{\alpha1}H_{\alpha1}$	−84.0	−96.5	−79.6	83.4	62.7	74.4
$N-C_{\alpha2}H_{\alpha2}$	−87.3	−69.1	−124.4	−121.7	−137.3	−112.2
$N-C_{\alpha1}H_N$	122.9	101.7	104.1	92.9	86.9	96.6
$N-C_{\alpha2}H_N$	−118.7	−122.3	−129.8	−147.2	−136.0	−135.2
$N-NH_N$	−151.7	−154.5	−157.1	−181.7	−168.7	−154.0
$N-NH_{\alpha1}$	−26.7	−8.7	−13.7	96.0	94.7	103.5
$N-NH_{\alpha2}$	−103.4	−91.2	−146.1	−154.2	−147.0	−137.0
$N-C'H_N$	−3.4	−18.2	−15.8	−32.6	−36.4	−17.2
$N-C'H_{\alpha1}$	−158.5	−136.7	−138.7	60.1	72.0	77.5
$N-C'H_{\alpha2}$	34.5	33.0	−8.1	−9.8	−3.6	1.6

[a] The isotropic NMR shielding σ,[48] the chemical shielding anisotropy Δσ, and the effective chemical shielding anisotropy $\Delta\sigma^{eff}$ for the CSA−DD cross-correlations. [b] The geometries of global energy minima. [c] Grid-point geometries close to the α-helix region (Figure 2).

generally suitable for the structural studies in peptides (Table 2, Figures 4−6).

The composition of the solvent environment at different pH values clearly affects the backbone structure of the LALA peptide.[48] The protonation and deprotonation occurring at the LALA termini thus represent large perturbations for NMR properties.[48] The backbone geometry parameters calculated previously for the global energy minima of the three LALA forms unveiled that the zwitterion and cation forms are well localized within the β-sheet region, while the anion potential energy surface is rather flat along the ψ dimension, allowing thus transitions between the α-helix and β-sheet conformers (Figure 2).[48] For the CSA−DD cross-correlated relaxation rates, we implied that those CCR mechanisms which possessed a rather stable structural dependence irrespective of the actual protonation state could be also more resistant with respect to similar structural perturbations occurring in terminal peptide residues or in polar side chains.

The optimal values of the backbone torsions corresponding to the global energy minima of the three LALA forms (the β-sheet conformation) were site-specifically affected by the protonation state (Table 1, Figure 2). The optimal values of

TABLE 2: Summary for the Calculated Effective CSAs in the LALA Peptide

CSA−DD	$\Delta(\Delta\sigma^{\mathrm{eff}})^a$	$\Delta(r_{\mathrm{DD}}^{-3}\Delta\sigma^{\mathrm{eff}})^b$	$\Delta_{\alpha\beta}(\Delta\sigma^{\mathrm{eff}})^c$	$\Delta_{\alpha\beta}(r_{\mathrm{DD}}^{-3}\Delta\sigma^{\mathrm{eff}})^d$	$r_{\mathrm{DD}}{}^e$	$\Delta r_{\mathrm{DD}}{}^f$	surface[g]	pH dependence[h]	M^i	N^j	available NMR expts
$C'-C_{\alpha1}H_{\alpha1}$	200.1	146.1	18.1	13.5	1.1	1.4	ψ	+ +	3	1	refs 20 and 24[l]
$C'-C_{\alpha2}H_{\alpha2}$	151.9	112.8	32.6	24.4	1.1	1.0	φ	+	3	1	
$C'-C_{\alpha1}H_N$	19.7	2.0	3.9	−0.4	2.6	14.6	2D	C	3	1	
$C'-C_{\alpha2}H_N$	38.7	4.8	−9.4	−0.9	2.1	6.7	2D	C	3	1	
$C'-NH_N$	32.2	29.4	−8.2	−7.5	1.0	2.3	2D	C	3	3	refs 37 and 46[k]
$C'-NH_{\alpha1}$	47.9	4.5	−8.5	−3.4	2.9	31.1	ψ	+ +	3	1	
$C'-NH_{\alpha2}$	107.6	12.5	9.0	1.2	2.1	3.9	φ	+ +	3	1	
$C'-C'H_N$	34.6	3.8	3.0	0.3	2.1	6.7	2D	+ +	2	1	refs 37 and 46
$C'-C'H_{\alpha1}$	100.7	10.9	48.1	5.5	2.1	5.4	ψ	+ +	2	1	
$C'-C'H_{\alpha2}$	60.7	5.5	17.4	1.7	3.0	28.7	2D	+	2	1	
$N-C_{\alpha1}H_{\alpha1}$	215.6	160.4	−159.2	−118.9	1.1	1.4	ψ	0	3	1	ref 19
$N-C_{\alpha2}H_{\alpha2}$	222.5	166.6	68.2	51.1	1.1	1.0	φ	0	3	1	ref 18
$N-C_{\alpha1}H_N$	94.5	5.9	14.8	1.5	2.6	14.6	2D	+	3	1	
$N-C_{\alpha2}H_N$	62.3	5.9	13.8	1.6	2.1	6.7	2D (φ)	+	3	1	
$N-NH_N$	72.1	69.2	14.3	15.1	1.0	2.3	2D	0	2	1	refs 37 and 46
$N-NH_{\alpha1}$	148.1	5.8	−103.4	−3.1	2.9	31.1	ψ	0	2	1	
$N-NH_{\alpha2}$	274.9	31.6	55.7	6.6	2.1	3.9	φ	0	2	1	
$N-C'H_N$	68.8	8.7	18.2	2.0	2.1	6.7	2D (φ)	0	3	3	refs 37 and 46[k]
$N-C'H_{\alpha1}$	256.1	26.2	−208.7	−20.2	2.1	5.4	ψ	0	3	1	
$N-C'H_{\alpha2}$	141.8	5.0	36.6	2.0	3.0	28.7	φ	+	3	1	

[a] The maximal calculated difference between the maximum and minimum values on the $\Delta\sigma^{\mathrm{eff}}$ surface among the three forms of the LALA peptide; $\Delta(\Delta\sigma^{\mathrm{eff}})$ [ppm]. [b] The maximal calculated difference between the maximum and minimum values on the $r_{\mathrm{DD}}^{-3}\Delta\sigma^{\mathrm{eff}}$ surface among the three forms of the LALA peptide; $\Delta(r_{\mathrm{DD}}^{-3}\Delta\sigma^{\mathrm{eff}})$ [Å^{-3} ppm]. [c] The difference between the $\Delta\sigma^{\mathrm{eff}}$ values calculated for the β-sheet and the α-helix conformation in the LALA zwitterion (Table 1); $\Delta_{\alpha\beta}(\Delta\sigma^{\mathrm{eff}})$ [ppm]. [d] The difference between the $r_{\mathrm{DD}}^{-3}\Delta\sigma^{\mathrm{eff}}$ values calculated for the β-sheet and the α-helix conformation in the LALA zwitterion (Table 1); $\Delta_{\alpha\beta}(r_{\mathrm{DD}}^{-3}\Delta\sigma^{\mathrm{eff}})$ [Å^{-3} ppm]. [e] The average length of the dipolar vector [Å]. [f] The maximal deviation of the dipolar vector length on the (ψ,φ) grid relative to the average length; Δr_{DD} [%]. [g] The dominant structural dependence of the (ψ,φ) surfaces calculated for the $\Delta\sigma^{\mathrm{eff}}$ (Supporting Information) (the surfaces significantly modulated in both dimensions are denoted "2D"). [h] Qualitative classification of the $\Delta\sigma^{\mathrm{eff}}$ surfaces variation between the three charged LALA forms ("+ +" stands for "large", "+" for "small", "C" for "complex", and "0" for "negligible"). [i] The number of nuclei involved in the CSA−DD cross-correlation. [j] The number of independent NMR experiments needed for the experimental detection of CCR rate. [k] Only the linked $\Gamma(C',NH_N) + \Gamma(N,C'H_N)$, $\Gamma(N,C'H_N) + \Gamma(H_N,NC')$, or $\Gamma(H_N,NC') + \Gamma(C',NH_N)$ terms can be measured. [l] The inseparable term $\Gamma(C_\alpha,C'H_\alpha)$ was considered to be negligible due to the small C_α-CSA.

the ψ torsion angle obtained for the zwitterion and cation were similar since the NH_3^+ terminus remained unchanged in these two forms, while for the anion (the NH_2 terminus) ψ increased by ∼20° (Table 1). Similar behavior was found for the φ angle; the value calculated in the cation (COOH) was larger by ∼30° compared to both the anion and the zwitterion (COO^-). These backbone conformational changes were reflected by the values of the amide nitrogen (σ_N) and the carbonyl carbon ($\sigma_{C'}$) isotropic shielding constants that had been analyzed in detail previously[48] (see also Table 1). Similar correlations could be observed for the chemical shielding anisotropies $\Delta\sigma_N$ and $\Delta\sigma_{C'}$ (eq 4) with the exception of $\Delta\sigma_N$ in the zwitterion, which was closer to the value calculated for the cation (Table 1). The calculated dependencies of $\Delta\sigma_N$ and $\Delta\sigma_{C'}$ on the ψ and φ backbone torsion angles were significantly modulated in both dimensions (Figure 3).

The $\Delta\sigma^{\mathrm{eff}}$ values calculated for the β-sheet and α-helix conformers (Table 1) certainly provide an alternative approach for the identification of the respective conformers in LALA dipeptide and (due to the local character of NMR properties) probably also generally in oligopeptides and proteins. Large differences of the calculated $\Delta\sigma^{\mathrm{eff}}$ values between β-sheet and α-helix could be considered as a qualitative indicator for their possible distinguishing with the CCR rates (Tables 1 and 2). In order to estimate the $\Delta\sigma^{\mathrm{eff}}$ values in the α-helix conformation, we used one of the grid points in a vicinity of that region (Table 1, Figure 2), while the geometries of global energy minima for all LALA forms were used as models for the β-sheet conformation. The calculated absolute $\Delta\sigma^{\mathrm{eff}}$ differences between the β-sheet and α-helix conformations ($|\Delta_{\alpha\beta}\Delta\sigma^{\mathrm{eff}}|$) in the LALA zwitterion spanned a relatively wide range: 3−48 ppm and 14−209 ppm for the cross-correlations involving carbonyl carbon and amide nitrogen, respectively (Table 2). For the cross-correlations that were recently probed experimentally (see the references in Table 2), the $|\Delta_{\alpha\beta}\Delta\sigma^{\mathrm{eff}}|$ absolute differences in the LALA zwitterion were 18, 8, and 3 ppm ($C'-C_{\alpha1}H_{\alpha1}$, $C'-NH_N$, $C'-C'H_N$) and 160, 68, 14, and 18 ppm ($N-C_{\alpha1}H_{\alpha1}$, $N-C_{\alpha2}H_{\alpha2}$, $N-NH_N$, $N-C'H_N$).

The effective CSAs calculated for one backbone conformation in the three LALA forms reflected the structural changes due to different pH values in a similar way as calculated for the isotropic shieldings[48] and CSAs (Table 1). The $\Delta\sigma^{\mathrm{eff}}$ values calculated in β-sheet geometries for the $C'-C_{\alpha1}H_{\alpha1}$ (zwitterion and cation) and $C'-C_{\alpha2}H_{\alpha2}$ (anion, zwitterion) cross-correlations ranged over relatively narrow intervals (24−31 ppm and 33−45 ppm, respectively) compared to the $C'-C_{\alpha1}H_{\alpha1}$ effective CSA in the anion (−27 ppm) and the $C'-C_{\alpha2}H_{\alpha2}$ effective CSA in the cation (3 ppm) (Table 1).

The calculated $\Delta\sigma^{\mathrm{eff}}$ (ψ,φ) surfaces are rather universal since they show overall trends which should be more or less valid in peptides unless some larger effect perturbed the local nature of this NMR parameter (for example, the proximity of an aromatic ring or an additional charge). The dependencies of the $\Delta\sigma^{\mathrm{eff}}$ (ψ,φ) surfaces on pH can also help to estimate the stability of a particular cross-correlation mechanism with regard to the fluctuating total charge of a peptide molecule (Table 2). Global characteristics of the (ψ,φ) surfaces (Table 2), structural variation of the calculated $\Delta\sigma^{\mathrm{eff}}$ values (Supporting Information), and the r_{jk}^{-3} scaling (Table 2, Figure 6B) can provide a reliable estimate for the conformational dependence of the CCR rates.

The magnitudes of CSA−DD cross-correlation effects are of course very much dependent on the angular projections of the σ-tensor principal components on the particular dipolar vector (eq 2). In this sense, the CSA and the effective CSA can be regarded

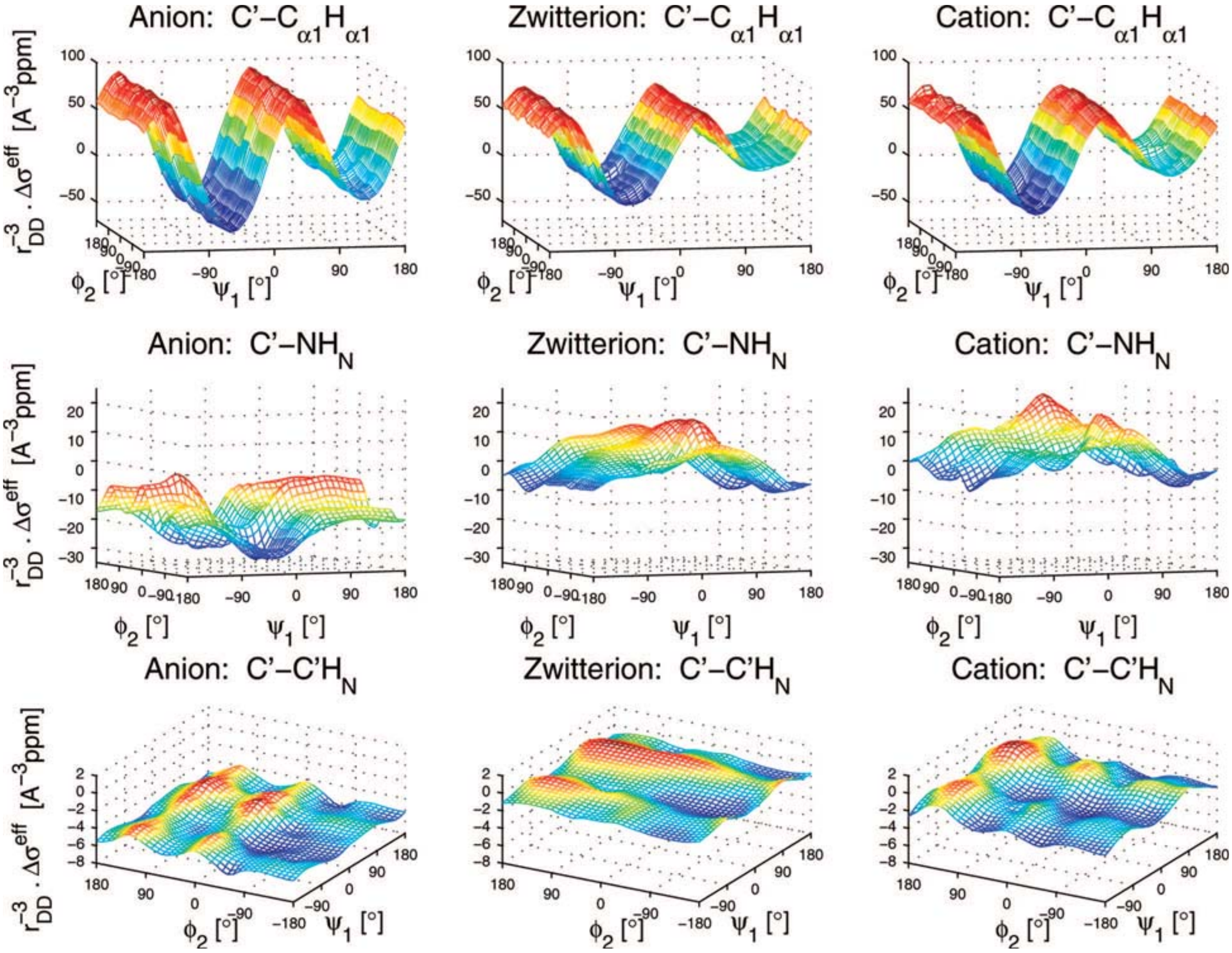


Figure 4. Dependence of selected $\Delta\sigma^{\mathrm{eff}}$ on the ψ_1 and φ_2 backbone torsion angles calculated for the three forms of the LALA molecule and scaled with the length of the dipolar vector. The cross-correlation mechanisms involved the carbonyl carbon σ-tensor and the $C_{\alpha1}H_{\alpha1}$, NH_N, and $C'H_N$ dipolar vectors.

as two different linear combinations of the σ-tensor principal values, one with constant coefficients (eq 4) and the other with molecular structure dependent coefficients (eq 2). The relative importance of the angular projections in comparison to the variation of the shielding tensor itself is thus immediately seen. The range of values calculated for the $\Delta\sigma_N$ and $\Delta\sigma_{C'}$ CSAs on the (ψ,φ) grid (Figure 3) was relatively small compared to the ranges calculated for the effective anisotropies $\Delta\sigma^{\mathrm{eff}}$ (eq 2, Figure 6A, Supporting Information), which were modulated specifically by the particular choice of the DD vector. Moreover, the calculated (ψ,φ) surfaces of the $\Delta\sigma_N$ and $\Delta\sigma_{C'}$ anisotropies possessing a significant "2D character" (i.e., the dependence on both torsion angles, see Figure 3) could have been smoothed or even may have lost their 2D character for a suitable choice of the DD vector. Such effective CSAs are rather attractive with regard to their direct structural interpretation (see the surface classification in Table 2 and the Supporting Information). Vice versa, the choice of the σ-tensor may also have resulted in a qualitatively different performance of the respective effective CSAs for one dipolar vector (see, for example, the $C'-C'H_{\alpha2}$ and $N-C'H_{\alpha2}$ effective CSAs in Table 2 and in the Supporting Information).

The experimentally rather easily accessible $N-NH_N$ cross-correlation is a particularly interesting parameter due to the amide nitrogen being the CSA center and the origin of the DD vector. The orientation of the DD vector with respect to the molecular frame remained quite similar for all backbone conformations. The $N-NH_N$ effective CSA therefore depended mainly on reorientation and magnitude alteration of the nitrogen σ-tensor principal components upon variation of the backbone torsions. The calculated (ψ,φ) surfaces of the $N-NH_N$ effective CSA (Supporting Information) were thus significantly modulated in both ψ and φ dimensions (Figure 5), similarly to the chemical shielding anisotropy $\Delta\sigma_N$ of the amide nitrogen (Figure 3). It should be however noted that in the case of hydrogen bonding both the direction and length of the NH_N vector may vary considerably.

The $\Delta\Delta\sigma^{\mathrm{eff}}$ values, i.e., the differences between maximum and minimum of the calculated $\Delta\sigma^{\mathrm{eff}}$ on the whole (ψ,φ) surface, were used to estimate the extent of the effective CSA modulation by the variation of the two torsion angles. For the $X-C_\alpha H_\alpha$ ($X = N, C'$; $\alpha = \alpha1, \alpha2$) cross-correlations, the $\Delta\Delta\sigma^{\mathrm{eff}}$ values were large (Table 2, Figure 6A) and the $\Delta\sigma^{\mathrm{eff}}$ (ψ,φ) surfaces depended predominantly on the φ ($X-C_{\alpha2}H_{\alpha2}$) or ψ ($X-C_{\alpha1}H_{\alpha1}$) torsion angles (Table 2, Supporting Information). These cross-correlations thus appear to be well suited for the peptide structural studies since they exhibit smooth and specific dependences on the backbone conformation, which is at the same time notably modulated (Table 2, Figures 4–6, and Supporting Information).

A somewhat lower applicability for the peptide structural studies could be expected from the rest of the cross-correlations involving the carbonyl carbon CSA. The $|\Delta_{\alpha\beta}\Delta\sigma^{\mathrm{eff}}|$ absolute differences were mostly small (exceeding 20 ppm in only two cases; see Tables 1 and 2), and the corresponding surfaces had usually pronounced 2D character often combined with significant pH dependence (Table 2) or shallow modulation (Figure 6A).

The effective CSAs involving the amide nitrogen σ-tensor performed significantly better. The calculated $\Delta\Delta\sigma^{\mathrm{eff}}$ values

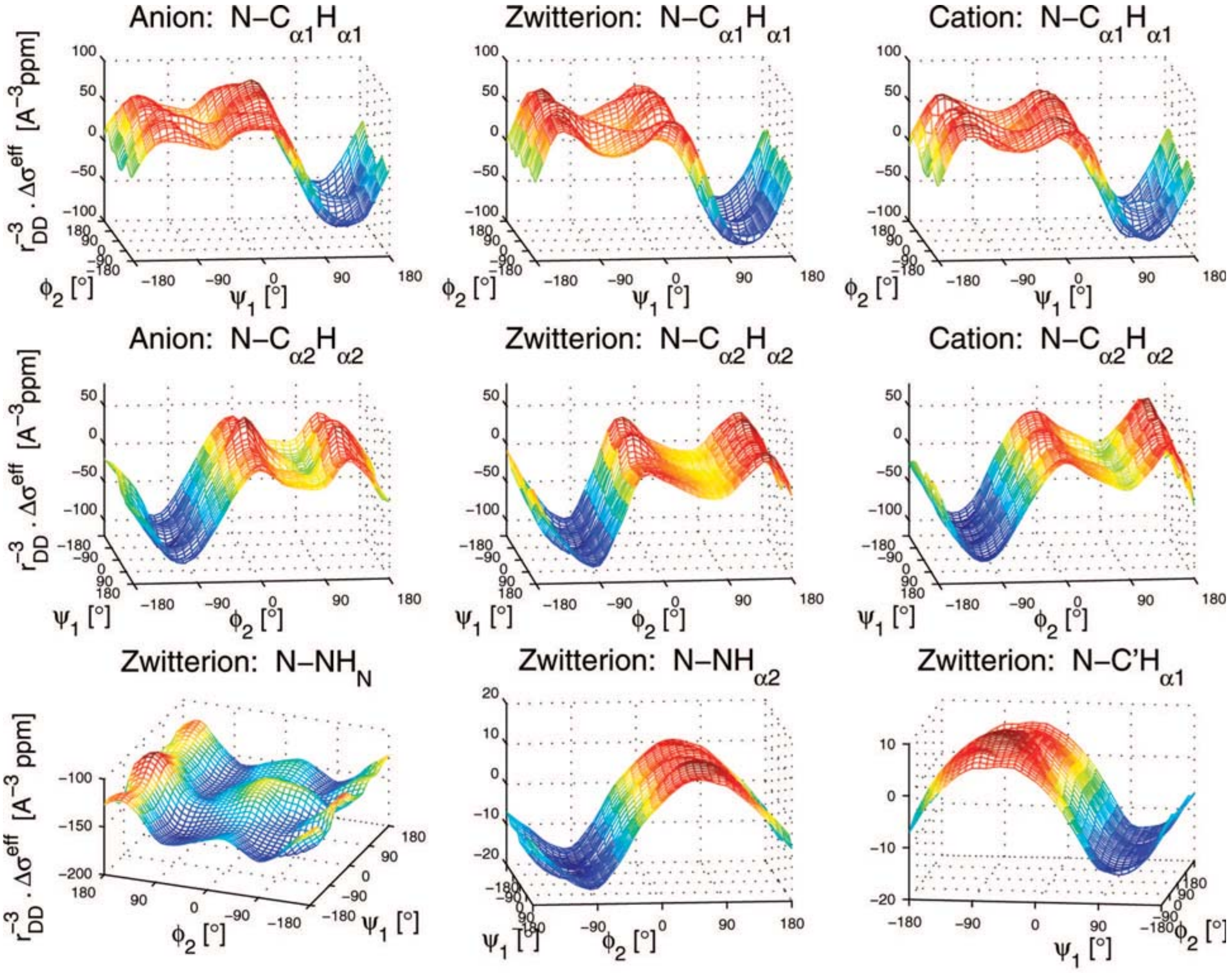


Figure 5. Dependence of selected $\Delta\sigma^{\mathrm{eff}}$ on the ψ_1 and φ_2 backbone torsion angles scaled with the length of the dipolar vector. The cross-correlation mechanisms involved the amide nitrogen σ-tensor and the $C_{\alpha1}H_{\alpha1}$, $C_{\alpha2}H_{\alpha2}$, NN_H, $NH_{\alpha2}$, and $C'H_{\alpha1}$ dipolar vectors.

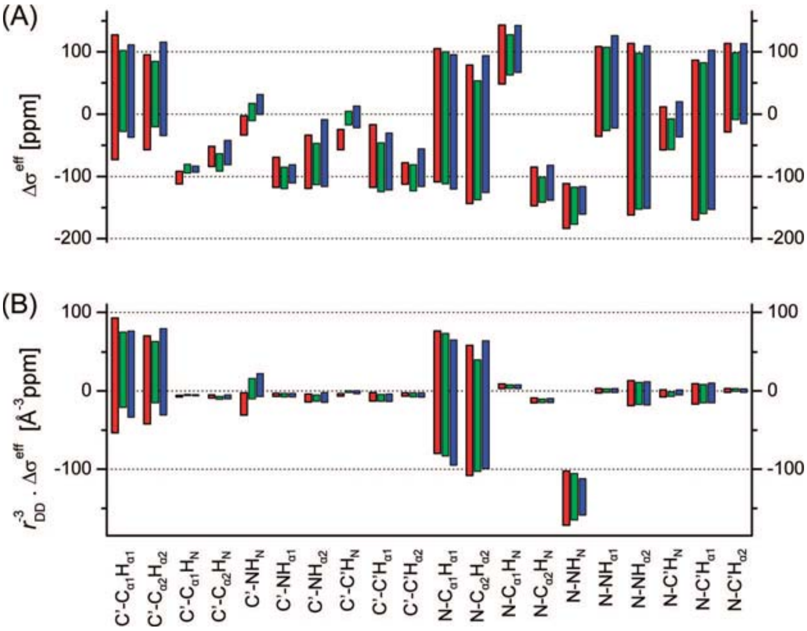


Figure 6. (A) Ranges of the effective CSA values corresponding to their maxima/minima on the (ψ,φ) surfaces calculated in the anion (red), zwitterion (green), and cation (blue) forms of the LALA molecule. (B) Calculated ranges of $\Delta\sigma^{\mathrm{eff}}$ scaled with the r_{DD}^{-3} factor, where r_{DD} is the length of the dipolar vector.

were overall larger than those involving the carbonyl carbon CSA (Figure 6). Also the (ψ,φ) surfaces of the nitrogen effective CSA were more often dominated by only a single torsion, and their pH dependence was smaller (Table 2, Supporting Information). In addition to the four cross-correlations involving the amide nitrogen CSA, which were probed in recent experimental studies (references in Table 2), we found similar $\Delta\Delta\sigma^{\mathrm{eff}}$ magnitudes for the $\mathrm{N{-}NH_{\alpha}}$ and $\mathrm{N{-}C'H_{\alpha}}$ ($\alpha = \alpha1, \alpha2$) effective CSAs. These cross-correlations might be also good candidates for the structural studies in peptides, since they depend mostly on a single torsion angle and their pH variation is negligible.

A comprehensive analysis of the calculated effective CSAs that can be used for a reliable estimate of the CCR rates must take into account their scaling with the length of the dipolar vector (r_{jk}^{-3} factor in eq 1, r_{DD}^{-3} column in Table 2). The scaling substantially reduced the $\Delta\sigma^{\mathrm{eff}}$ values for the cross-correlations with longer dipolar vectors (Table 2, Figure 6B). The length of the DD vector thus can be a rather limiting factor for the experimental accessibility of the CCR rates. On the other hand, the grid point dependent lengths of the investigated DD vectors varied only up to 2.3, 6.7, and 31.1% (relative to the average value) for the DD vectors shorter than 1.1, 2.1, and 3.0 Å, respectively (Table 2). Therefore, the calculated conformational dependence of the studied effective CSAs should not be significantly altered by the r_{DD}^{-3} scaling. Several selected (ψ,φ) surfaces of the scaled effective CSAs (Figures 4 and 5) can be compared to the original $\Delta\sigma^{\mathrm{eff}}$ surfaces (Supporting Information).

The differences of the scaled $\Delta\sigma^{\mathrm{eff}}$ values calculated between the β-sheet and α-helix conformers (Table 2) as well as the scaled $\Delta\sigma^{\mathrm{eff}}$ ranges (Table 2, Figure 6B) were evaluated consistently with the original effective CSAs. Significantly large values of the scaled $\Delta\sigma^{\mathrm{eff}}$ were obtained for the experimentally probed $\mathrm{C'{-}C_{\alpha1}H_{\alpha1}}$, $\mathrm{C'{-}NH_{N}}$, $\mathrm{N{-}C_{\alpha1}H_{\alpha1}}$, $\mathrm{N{-}C_{\alpha2}H_{\alpha2}}$, and $\mathrm{N{-}NH_{N}}$ cross-correlations (Table 2 and references therein, Figure 6B). The $\mathrm{C'{-}C_{\alpha2}H_{\alpha2}}$, $\mathrm{N{-}NH_{\alpha2}}$, and $\mathrm{N{-}C'H_{\alpha1}}$ cross-correlations, for which an experimental reference was not available, and which were theoretically modeled only in this work, appeared after the scaling with similarly good prerequisites for application in conformational studies of peptide backbone as those CCRs already experimentally tested.

In particular, the $\mathrm{C'{-}C_{\alpha2}H_{\alpha2}}$ cross-correlation had a range of $\Delta\sigma^{\mathrm{eff}}$ values equivalent to its complement $\mathrm{C'{-}C_{\alpha1}H_{\alpha1}}$ involving the $\mathrm{C_{\alpha1}H_{\alpha1}}$ dipolar vector (Table 2, Figure 6B). While the $\Delta\sigma^{\mathrm{eff}}$ calculated for the $\mathrm{C'{-}C_{\alpha1}H_{\alpha1}}$ cross-correlation depended dominantly on the ψ torsion, the choice of the $\mathrm{NH_{N}}$ dipolar vector for the same CSA dramatically changed the shape and other characteristics of the $\Delta\sigma^{\mathrm{eff}}$ surface (Figure 4, Table 2). Of the other cross-correlations involving the carbonyl carbon, the $\mathrm{C'{-}NH_{N}}$ was the only scaled $\Delta\sigma^{\mathrm{eff}}$ that still possessed a relatively large range of values. However, the scaled $\Delta\sigma^{\mathrm{eff}}$ surfaces were significantly modulated in both the ψ and the φ dimensions for all forms of the LALA molecule (Figure 4), and the charge of the terminal groups also had a notable impact (Figures 4 and 6B).

The cross-correlations involving the amide nitrogen CSA were shown to be overall more suitable for peptide structural studies than those involving the carbonyl carbon CSA. For a given dipolar interaction, the range of the scaled $\Delta\sigma^{\mathrm{eff}}$ values was almost always larger for the N–DD type of cross-correlation than for the C′–DD type (Table 2, Figure 6B). Such behavior could have been expected from a well-known fact that the amide ^{15}N CSA is larger than the carbonyl ^{13}C CSA.[60] Interestingly, the scaled $\Delta\sigma^{\mathrm{eff}}$ (ψ,φ) surfaces for the experimentally probed $\mathrm{N{-}C_{\alpha1}H_{\alpha1}}$ and $\mathrm{N{-}C_{\alpha2}H_{\alpha2}}$ cross-correlations[18,19] were rather smooth and depended almost exclusively on one backbone torsion only. On the contrary, the scaled $\Delta\sigma^{\mathrm{eff}}$ surfaces for the experimentally most convenient $\mathrm{N{-}NH_{N}}$ cross-correlation[10,37,46] were significantly modulated in both dimensions, which renders its structural interpretation difficult (Table 2, Figure 5), although not unsuitable in the form of restraints.

The promising $\mathrm{N{-}NH_{\alpha2}}$ and $\mathrm{N{-}C'H_{\alpha1}}$ cross-correlations, which had not been reported previously, involved the amide nitrogen CSA and relatively long DD vectors ($r_{\mathrm{DD}} \sim 2.1$ Å, Table 2). Although the calculated ranges of the scaled $\Delta\sigma^{\mathrm{eff}}$ values were somewhat smaller compared to the experimentally probed cross-correlations discussed above (Table 2, Figure 6B), they exhibited a rather smooth dependence on only one of the two backbone torsions (Table 2, Figure 5) and a negligible pH variation. Moreover, the $\mathrm{N{-}NH_{\alpha2}}$ cross-correlation mechanism involves only two nuclei, which is experimentally favorable. These two cross-correlations probably represent the best performing examples of the CSA−DD mechanism involving the dipolar interaction between not directly bonded atoms.

The following cross-correlations, which were not yet measured, can be probably also considered useful for the peptide backbone structural studies, namely because they depended dominantly on a single torsion only: $\mathrm{C'{-}NH_{\alpha1}}$, $\mathrm{C'{-}NH_{\alpha2}}$, $\mathrm{C'{-}C'H_{\alpha1}}$, $\mathrm{N{-}NH_{\alpha1}}$, and $\mathrm{N{-}C'H_{\alpha2}}$ (Table 2, Figure 6B). The other cross-correlation mechanisms, among which the $\mathrm{C'{-}C'H_{N}}$ and $\mathrm{N{-}C'H_{N}}$ experimentally probed CCRs can be counted, possessed either a small range of the scaled $\Delta\sigma^{\mathrm{eff}}$ values on the (ψ,φ) grid, or the surfaces were modulated in both ψ and φ dimensions. Their usability in peptide structural studies is therefore limited.

Conclusion

The calculated dependencies of the effective chemical shielding anisotropies on the main backbone torsion angles ψ and φ in the cationic, zwitterionic, and anionic forms of the L-alanyl-L-alanine peptide were analyzed in order to assess the applicability of the corresponding cross-correlated relaxation rates in structural studies of oligopeptides and proteins. The analysis of the calculated effective CSAs was focused on the overall character of their dependence on backbone conformation and molecular charge, on their usability for distinguishing between the β-sheet and α-helix backbone conformations, and on the effect of scaling by the length of the dipolar vector.

The 20 cross-correlated relaxation mechanisms studied in this work involved either the amide nitrogen CSA or the carbonyl carbon CSA and dipolar vectors connecting α-carbon, carbonyl carbon, or amide nitrogen to amide hydrogen or α-hydrogen.

The effective CSAs generally depend on both backbone torsion angles, but for the $\mathrm{X{-}C_{\alpha}H_{\alpha}}$, $\mathrm{X{-}NH_{\alpha}}$, $\mathrm{X{-}C'H_{\alpha1}}$, and $\mathrm{N{-}C'H_{\alpha2}}$ ($\mathrm{X} = \mathrm{C'}, \mathrm{N}$; $\alpha = \alpha1, \alpha2$) effective CSAs, a dominant dependence on a single torsion angle was observed. The amide nitrogen effective CSAs showed an overall clearer dependence on only one of the two backbone torsions with a smaller variation by the molecular charge compared to the carbonyl carbon effective CSAs. The cross-correlations involving the amide nitrogen also exhibited larger ranges of the effective CSA values, which probably implies their better usability in peptide structural studies.

By taking into account the length of the dipolar vector, a more conclusive estimate of the cross-correlated relaxation rates and their relative magnitudes was obtained. The cross-correlations involving the dipolar vectors with a length corresponding to a single covalent bond ($\mathrm{X{-}C_{\alpha}H_{\alpha}}$, $\mathrm{X{-}NH_{N}}$; $r_{\mathrm{DD}} \sim 1.1$ Å) expectedly provided the largest calculated ranges of the scaled effective CSAs. The complex analysis carried out for the effective CSAs showed that the $\mathrm{X{-}C_{\alpha}H_{\alpha}}$ cross-correlations probed in recent experimental studies[18−20,24] are best suited for the determination of both ψ and φ peptide backbone torsion angles. The dipolar scaling damped down most of the other investigated effective CSAs that would otherwise be also attractive due to their large effective CSA ranges and dominant dependence on a single torsion angle (namely the $C'{-}NH_{\alpha1}$, $C'{-}NH_{\alpha2}$, $C'{-}C'H_{\alpha1}$, $N{-}NH_{\alpha1}$, and $N{-}C'H_{\alpha2}$ effective CSAs). Finally, the $N{-}NH_{\alpha2}$ and $N{-}C'H_{\alpha1}$ cross-correlations involving the dipolar vectors of intermediate length ($r_{DD} \sim 2.1$ Å) were classified as the most promising candidates for peptide backbone structural studies, which should be probed experimentally in the future. To assess the weaker cross-correlation effects predicted in this paper reliably, highly sensitive experiments and high accuracy of the measurements are a stringent requirement. The symmetrical reconversion schemes introduced by Bodenhausen's group[37,61] provide valuable approaches for this.

Acknowledgment. This work was supported by the Academy of Sciences of the Czech Republic, Grants IAA400550701, A4005507020, and AV0Z40550506. V.S. was supported by a Human Frontier Science Program (HFSP) Young Investigator's Grant. We also acknowledge support by the ÖAD Kontakt (AT−CZ) program Project No. CZ 08/2007-KONTAKT 2007/5.

Supporting Information Available: The calculated surfaces of all 20 effective CSAs in the three forms of the LALA molecule. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Downing, K. A. *Protein NMR Techniques*, 2nd ed.; Humana Press: Totowa, NJ, 2004; Vol. 278.

(2) James, T. L.; Dotsch, V.; Schmitz, U. *Nuclear Magnetic Resonance of Biological Macromolecules Part A*, 1st ed.; Elsevier: New York, 2001; Vol. 338.

(3) Wijmenga, S. S.; van Buuren, B. N. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287.

(4) Mollova, E. T.; Hansen, M. R.; Pardi, A. *J. Am. Chem. Soc.* **2000**, *122*, 11561.

(5) Grzesiek, S.; Cordier, F.; Jaravine, V.; Barfield, M. *Prog. NMR Spectrosc.* **2004**, *45*, 275.

(6) Flinders, J.; Dieckmann, T. *Prog. NMR Spectrosc.* **2006**, *48*, 137.

(7) Tanaka, Y.; Ono, A. *Dalton Trans.* **2008**, 4965.

(8) Lipsitz, R. S.; Tjandra, N. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 387.

(9) Reif, B.; Hennig, M.; Griesinger, C. *Science* **1997**, *276*, 1230.

(10) Kroenke, C. D.; Loria, J. P.; Lee, L. K.; Rance, M.; Palmer, A. G. *J. Am. Chem. Soc.* **1998**, *120*, 7905.

(11) Pang, Y. X.; Wang, L. C.; Pellecchia, M.; Kurochkin, A. V.; Zuiderweg, E. R. P. *J. Biomol. NMR* **1999**, *14*, 297.

(12) Carlomagno, T.; Maurer, M.; Hennig, M.; Griesinger, C. *J. Am. Chem. Soc.* **2000**, *122*, 5105.

(13) Banci, L.; Bertini, I.; Felli, I. C.; Hajieva, P.; Viezzoli, M. S. *J. Biomol. NMR* **2001**, *20*, 1.

(14) Chiarparin, E.; Rudisser, S.; Bodenhausen, G. *ChemPhysChem* **2001**, *2*, 41.

(15) Ravindranathan, S.; Kim, C. H.; Bodenhausen, G. *J. Biomol. NMR* **2003**, *27*, 365.

(16) Vugmeyster, L.; Perazzolo, C.; Wist, J.; Frueh, D.; Bodenhausen, G. *J. Biomol. NMR* **2004**, *28*, 173.

(17) Wang, T. Z.; Frederick, K. K.; Igumenova, T. I.; Wand, A. J.; Zuiderweg, E. R. P. *J. Am. Chem. Soc.* **2005**, *127*, 828.

(18) Hong, M.; Gross, J. D.; Hu, W.; Griffin, R. G. *J. Magn. Reson.* **1998**, *135*, 169.

(19) Reif, B.; Diener, A.; Hennig, M.; Maurer, M.; Griesinger, C. *J. Magn. Reson.* **2000**, *143*, 45.

(20) Yang, D. W.; Konrat, R.; Kay, L. E. *J. Am. Chem. Soc.* **1997**, *119*, 11938.

(21) Yang, D. W.; Kay, L. E. *J. Am. Chem. Soc.* **1998**, *120*, 9880.

(22) Pelupessy, P.; Chiarparin, E.; Ghose, R.; Bodenhausen, G. *J. Biomol. NMR* **1999**, *13*, 375.

(23) Pelupessy, P.; Chiarparin, E.; Ghose, R.; Bodenhausen, G. *J. Biomol. NMR* **1999**, *14*, 277.

(24) Chiarparin, E.; Pelupessy, P.; Ghose, R.; Bodenhausen, G. *J. Am. Chem. Soc.* **1999**, *121*, 6876.

(25) Chiarparin, E.; Pelupessy, P.; Ghose, R.; Bodenhausen, G. *J. Am. Chem. Soc.* **2000**, *122*, 1758.

(26) Carlomagno, T.; Felli, I. C.; Czech, M.; Fischer, R.; Sprinzl, M.; Griesinger, C. *J. Am. Chem. Soc.* **1999**, *121*, 1945.

(27) Felli, I. C.; Richter, C.; Griesinger, C.; Schwalbe, H. *J. Am. Chem. Soc.* **1999**, *121*, 1956.

(28) Richter, C.; Griesinger, C.; Felli, I.; Cole, P. T.; Varani, G.; Schwalbe, H. *J. Biomol. NMR* **1999**, *15*, 241.

(29) Richter, C.; Reif, B.; Griesinger, C.; Schwalbe, H. *J. Am. Chem. Soc.* **2000**, *122*, 12728.

(30) Riek, R. *J. Magn. Reson.* **2001**, *149*, 149.

(31) Carlomagno, T.; Blommers, M. J. J.; Meiler, J.; Cuenoud, B.; Griesinger, C. *J. Am. Chem. Soc.* **2001**, *123*, 7364.

(32) Duchardt, E.; Richter, C.; Ohlenschlager, O.; Gorlach, M.; Wohnert, J.; Schwalbe, H. *J. Am. Chem. Soc.* **2004**, *126*, 1962.

(33) Rinnenthal, J.; Richter, C.; Ferner, J.; Duchardt, E.; Schwalbe, H. *J. Biomol. NMR* **2007**, *39*, 17.

(34) Pang, Y. X.; Zuiderweg, E. R. P. *J. Am. Chem. Soc.* **2000**, *122*, 4841.

(35) Cornilescu, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 10143.

(36) Cisnetti, F.; Loth, K.; Pelupessy, P.; Bodenhausen, G. *ChemPhysChem* **2004**, *5*, 807.

(37) Loth, K.; Pelupessy, P.; Bodenhausen, G. *J. Am. Chem. Soc.* **2005**, *127*, 6062.

(38) Konrat, R.; Sterk, H. *Chem. Phys. Lett.* **1993**, *203*, 75.

(39) Ghalebani, L.; Bernatowicz, P.; Aski, S. N.; Kowalewski, J. *Concepts Magn. Reson., Part A* **2007**, *30A*, 100.

(40) Tugarinov, V.; Shapiro, Y. E.; Liang, Z. C.; Freed, J. H.; Meirovitch, E. *J. Mol. Biol.* **2002**, *315*, 155.

(41) Meirovitch, E.; Shapiro, Y. E.; Polimeno, A.; Freed, J. H. *J. Phys. Chem. A* **2006**, *110*, 8366.

(42) Kaupp, M.; Bühl, M.; Malkin, V. *Calculation of NMR and EPR parameters*; Wiley-VCH Verlag: Weinheim, 2004.

(43) Bartoschek, S.; Buurman, G.; Geierstanger, B. H.; Lapham, J.; Griesinger, C. *J. Am. Chem. Soc.* **2003**, *125*, 13308.

(44) Markwick, P. R. L.; Sattler, M. *J. Am. Chem. Soc.* **2004**, *126*, 11424.

(45) Sychrovský, V.; Müller, N.; Schneider, B.; Smrečki, V.; Špirko, V.; Šponer, J.; Trantírek, L. *J. Am. Chem. Soc.* **2005**, *127*, 14663.

(46) Brutscher, B.; Skrynnikov, N. R.; Bremi, T.; Bruschweiler, R.; Ernst, R. R. *J. Magn. Reson.* **1998**, *130*, 346.

(47) Bouř, P.; Buděšínský, M.; Špirko, V.; Kapitán, J.; Šebestík, J.; Sychrovský, V. *J. Am. Chem. Soc.* **2005**, *127*, 17079.

(48) Sychrovský, V.; Buděšínský, M.; Benda, L.; Špirko, V.; Vokáčová, Z.; Šebestík, J.; Bouř, P. *J. Phys. Chem. B* **2008**, *112*, 1796.

(49) Graf, J.; Nguyen, P. H.; Stock, G.; Schwalbe, H. *J. Am. Chem. Soc.* **2007**, *129*, 1179.

(50) Šebek, J.; Gyurcsik, B.; Šebestík, J.; Kejík, Z.; Bednárová, L.; Bouř, P. *J. Phys. Chem. A* **2007**, *111*, 2750.

(51) Kowalewski, J.; Mäler, L. *Nuclear Spin Relaxation in Liquids: Theory, Experiments, and Applications*, Taylor & Francis Group: New York, London, 2006.

(52) Kumar, A.; Grace, R. C. R.; Madhu, P. K. *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *37*, 191.

(53) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(54) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

(55) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(56) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(57) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(58) Kutzelnigg, W.; Fleischer, U.; Schindler, M. *NMR, Basic Principles and Progress*; Springer: Heidelberg, 1990; Vol. 23.

(59) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(60) Hartzell, C. J.; Whitfield, M.; Oas, T. G.; Drobny, G. P. *J. Am. Chem. Soc.* **1987**, *109*, 5966.

(61) Pelupessy, P.; Espallargas, G. M.; Bodenhausen, G. *J. Magn. Reson.* **2003**, *161*, 258.

JP8105452

3574 *J. Phys. Chem. B* **2010,** *114,* 3574–3582

Theoretical Modeling of Magnesium Ion Imprints in the Raman Scattering of Water

Josef Kapitán,*,† Martin Dračínský,*,‡ Jakub Kaminský,‡ Ladislav Benda,‡ and Petr Bouř*,‡

Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, United Kingdom, and Institute of Organic Chemistry and Biochemistry, Academy of Sciences, 166 10 Prague, Czech Republic

Received: November 20, 2009; Revised Manuscript Received: February 4, 2010

Hydration envelopes of metallic ions significantly influence their chemical properties and biological functioning. Previous computational studies, nuclear magnetic resonance (NMR), and vibrational spectra indicated a strong affinity of the Mg^{2+} cation to water. We find it interesting that, although monatomic ions do not vibrate themselves, they cause notable changes in the water Raman signal. Therefore, in this study, we used a combination of Raman spectroscopy and computer modeling to analyze the magnesium hydration shell and origin of the signal. In the measured spectra of several salts (LiCl, NaCl, KCl, $MgCl_2$, $CaCl_2$, $MgBr_2$, and MgI_2 water solutions), only the spectroscopic imprint of the hydrated Mg^{2+} cation could clearly be identified as an exceptionally distinct peak at $\sim$355 cm^{-1}. The assignment of this band to the Mg−O stretching motion could be confirmed on the basis of several models involving quantum chemical computations on metal/water clusters. Minor Raman spectral features could also be explained. Ab initio and Fourier transform (FT) techniques coupled with the Car−Parrinello molecular dynamics were adapted to provide the spectra from dynamical trajectories. The results suggest that even in concentrated solutions magnesium preferentially forms a $[Mg(H_2O)_6]^{2+}$ complex of a nearly octahedral symmetry; nevertheless, the Raman signal is primarily associated with the relatively strong metal−H_2O bond. Partially covalent character of the Mg−O bond was confirmed by a natural bond orbital analysis. Computations on hydrated chlorine anion did not provide a specific signal. The FT techniques gave good spectral profiles in the high-frequency region, whereas the lowest-wavenumber vibrations were better reproduced by the cluster models. Both dynamical and cluster computational models provided a useful link between spectral shapes and specific ion−water interactions.

Introduction

Simple monatomic ions in aqueous environment are essential for many important biological processes in living organisms, such as osmotic regulation, neural impulses, muscle contraction, and protein conformational changes.[1−3] The metal ions often react quite specifically, exhibiting strong affinity to particular binding sites.[4] Anions, such as halogens, appear more important for water bulk properties.[5] It is clear that ionic physical properties are modulated by the hydration shell and interactions with the water molecules, but determination of the exact mechanisms is often hindered by the lack of experimental data. Some biological systems that could enlighten the role of the ions are difficult to crystallize. Nevertheless, many techniques, such as X-ray crystallography or NMR spectroscopy, did provide valuable insights into the origin of specific ionic interactions.[1,6] Vibrational spectra allow one to estimate the metal−water bond strength.[7] To contribute to this process, we explore the ability of the Raman scattering to detect specific ion−water interactions, particularly those of the magnesium ion providing the most distinct signal in the spectrum.

Vibrational properties probed by the Raman techniques and infrared absorption (IR) proved previously as sensitive indicators of the water structure. For example, the IR and Raman signal of the OH stretching seems to indicate clustering in liquid water as dependent on the temperature.[8−10] Librational motions (within 700−900 cm^{-1}) are sensitive to the ice structure and isotopomer composition.[11] Vibrational investigations of more complex ionic solutions are rarer.[7,12]

The information about structure encoded into the spectral band frequencies and intensities can be to a great extent decoded by computer simulations. Unlike for isolated molecules, however, bulk water and aqueous ionic solutions provide a limited amount of spectral features. A large part of the structural information is reflected in an inhomogeneous broadening of spectral shapes and requires special interpretation procedures.[13−15] Molecular dynamics (MD) with empirical[9,16−18] and ab initio[19,20] force fields provide the most flexible means for modeling of the structure and its development in time. Most often, the computationally efficient Car−Parrinello formulation of the ab initio molecular dynamics (CPMD) is preferred.[21] The CPMD and other ab initio MD techniques particularly enhanced structural studies of ionic solutions that, being complicated electronic systems, cannot be described by empirical force fields.[2,22,23]

IR and Raman spectral profiles can be simulated by various Fourier transform techniques from the time dependence of the dipole and polarizability functions, respectively.[18,20,24,25] The Fourier simulations can partially account for anharmonic effects and inhomogeneous broadening of spectral lines; in principle, however, they are constrained to the harmonic approximation. Alternatively, the instantaneous mode approximation and cluster approaches reproduce most of the spectral features.[26,27] The static models do not include the motion explicitly but typically allow one to calculate spectroscopic properties at a higher approximation level, and provide more flexibility, potentially including also the anharmonic effects. In the present study, both ap-

---

* To whom correspondence may be addressed. E-mail: kapitan@chem.gla.ac.uk (J.K.); dracinsky@uochb.cas.cz (M.D.); bour@uochb.cas.cz (P.B.).

† University of Glasgow.

‡ Academy of Sciences.

10.1021/jp9110508 © 2010 American Chemical Society
Published on Web 02/22/2010

proaches are used, as they reveal different aspects of the origin of the Raman signal.

The metal ions obviously have no vibrational spectra themselves. The specific spectroscopic signal comes from the perturbed liquid water structure in the vicinity of the ion. Potentially, the weak water−water, metal−water, and other noncovalent interactions are involved, which makes the Raman technique very attractive, as it enables measurement to very low vibrational frequencies comprising fundamental vibrational modes that correspond to these weak interactions.[7,28,29]

In this work, we have measured Raman spectra of several salts dissolved in water in order to map their influence on the water light scattering and, in final effect, on the water and hydration shell structure. As only the magnesium ion among the light metals exhibited an exceptionally specific band at ∼355 cm^{-1}, we concentrated MD and ab initio modeling to this ion in aqueous solutions. Classical and Car−Parrinello MD were used to estimate the dynamical behavior of the hydration shell. Cluster and Fourier transform techniques were tested for generation of the Raman spectra. The computations confirm that the signal stems from the tight magnesium−oxygen binding and relatively rigid structure of the magnesium first hydration shell. The experimental results and CPMD computations indicate that the signal is not exclusively associated with an ideal octahedral structure but rather with individual Mg−O bond stretchings. A natural bond orbital analysis suggests that the Mg−O bond adopts a partially covalent character.

Methods

Experimental Section. Samples of the salts (LiCl, NaCl, KCl, $MgCl_2$, $CaCl_2$, $MgBr_2$, and MgI_2) were supplied by Sigma-Aldrich and used without further purification. Raman spectra were measured in a backscattering mode using the previously described ChiralRAMAN instrument (BioTools, Inc.).[30] The laser wavelength was 532 nm, laser power at the sample ∼0.3 W, spectral resolution 10 cm^{-1}, and acquisition times 10−30 min. All spectral intensities were divided by accumulation time so that they are normalized to the same intensity scale. The Raman spectrum of an empty quartz cell was subtracted from all spectra. Note that the normalization is very approximate, as the collection angle of Raman scattering depends, among other factors, on the concentration and hence on the refractive index of the solution. For the HOH bending signals, Lorentzian bands were fitted to the experimental spectra with a correction to linear baseline using the SpectraCalc (Galactic Industries) software to get central peak positions and full widths at half-height.

Ab Initio Cluster Calculations. Geometries of metal (Na^+, Ca^{2+}, Mg^{2+}) and water clusters were optimized by energy minimization at the HF, MP2, MP4, and B3LYP[31] approximation levels and standard 6-311++G** and aug-cc-pVTZ basis sets, using the Gaussian software.[32] At the same level as for the optimization, harmonic frequencies and backscattered Raman transition intensities were calculated as[33]

$$I_{180} = 7S_1 + S_0 \tag{1}$$

where $S_1 = \sum_{i=1...3,j=1...3} \alpha_{ij}\alpha_{ij}$ and $S_0 = \sum_{i=1...3,j=1...3} \alpha_{ii}\alpha_{jj}$ are the isotropic invariants of the transition polarizability. The spectra were generated using a temperature correction factor[33] and Lorentzian bands of full width at half-maximum $\Delta = 5$ cm^{-1} as

$$S(\omega) = I_{180}\left[1 - \exp\left(-\frac{\omega_i}{kT}\right)\right]^{-1}\frac{1}{\omega_i}\left[4\left(\frac{\omega - \omega_i}{\Delta}\right)^2 + 1\right]^{-1} \tag{2}$$

where k is the Boltzmann constant, T temperature, ω_i transition frequency, and ω the light frequency. The Gaussian program was also used for the natural bond orbital (NBO) analysis[34] of the wave function.

CPMD Simulations. Periodic boxes containing 33 water molecules and salt ions were created by the HyperChem program.[35] The sizes of the boxes approximately corresponded to the experimental densities of salt solutions at 300 K. The systems included pure water, hydrated Mg^{2+}, Ca^{2+}, Na^+, and Cl^- ions, and a box containing three Mg and six Cl atoms corresponding to the experimentally investigated 5 mol/kg $MgCl_2$ solution. Within HyperChem and the Amber99 force field,[36,37] classical molecular dynamics (MD) was run for 1 ns with 1 fs integration time steps and a temperature of 400 K to equilibrate the systems. The TIP3P[38] model as a part of Amber99 was used for water. Because this approach did not provide symmetric Mg^{2+} hydration, alternatively, the starting geometry with ideal octahedral $[Mg(H_2O)_6]^{2+}$ geometry was created manually. Both approaches, however, gave very similar spectra and energies. The box geometry was optimized (Amber99) and input into the Car−Parrinello CPMD software package.[39] The same periodic boundary conditions, 4 au = 0.09676 fs time step, and energy cutoff of 85 Ry were used for all CPMD calculations performed with the BLYP[31] functional and Troulier−Martins[40] norm-conserving pseudopotentials. The initial configuration was relaxed by six short CPMD runs comprising 200 steps. After each run, the system was quenched to the Born−Oppenheimer surface by reoptimizing the wave function. Subsequently, longer 2.5−10 ps production runs were performed. A temperature of 300 K was maintained with the Nosé−Hoover algorithm,[41] which also kept the system in the canonical (NVT) ensemble. During the long simulations, the trajectory was saved at every step for the FT spectra generation and geometry analyses.

Spectral Simulations from Dynamical Trajectories. The spectral response was derived for an ensemble of N molecules with dipoles μ_i. A local field $\mathbf{E}_i$ sensed by molecule i is the external electric field $\mathbf{E}_0$ plus the field from other dipoles[42,43]

$$\mathbf{E}_i = \mathbf{E}_0 + \frac{1}{4\pi\varepsilon_0}\sum_{j\neq i}\frac{3\boldsymbol{\mu}_j \cdot \mathbf{R}_{ij}\mathbf{R}_{ij} - \boldsymbol{\mu}_j R_{ij}^{\,2}}{R_{ij}^{\,5}} \tag{3}$$

where $\mathbf{R}_{ij} = \mathbf{R}_i - \mathbf{R}_j$ is the distance vector. The total dipole moment is

$$\begin{aligned}\boldsymbol{\mu}_{\text{total}} &= \sum_i (\boldsymbol{\mu}_{i0} + \boldsymbol{\alpha}_i \cdot \mathbf{E}_i) \qquad (4)\\ &\approx \sum_i (\boldsymbol{\mu}_{i0} + \boldsymbol{\alpha}_i \cdot \mathbf{E}_0) + \\ &\frac{1}{4\pi\varepsilon_0}\sum_i\sum_{j\neq i}\boldsymbol{\alpha}_i \cdot \left(\frac{3\boldsymbol{\mu}_{0j} \cdot \mathbf{R}_{ij}\mathbf{R}_{ij} - \boldsymbol{\mu}_{0j}R_{ij}^{\,2}}{R_{ij}^{\,5}} + \right.\\ &\left.\frac{3\boldsymbol{\alpha}_j \cdot \mathbf{E}_0 \cdot \mathbf{R}_{ij}\mathbf{R}_{ij} - \boldsymbol{\alpha}_j \cdot \mathbf{E}_0 R_{ij}^{\,2}}{R_{ij}^{\,5}}\right)\end{aligned}$$

and the total polarizability

$$\boldsymbol{\alpha} = \frac{\partial \boldsymbol{\mu}_{\mathrm{total}}}{\partial \mathbf{E}_0} = \sum_i \boldsymbol{\alpha}_i + \frac{1}{4\pi\varepsilon_0}\sum_i \sum_{i\neq j} \frac{3\boldsymbol{\alpha}_j \cdot \mathbf{R}_{ij}\boldsymbol{\alpha}_i \cdot \mathbf{R}_{ij} - \boldsymbol{\alpha}_i \cdot \boldsymbol{\alpha}_j R_{ij}^{\,2}}{R_{ij}^{\,5}} \tag{5}$$

A derivative of (5) with respect to the ε-coordinate of an atom λ present in molecule m is

$$\frac{\partial \boldsymbol{\alpha}}{\partial r_\varepsilon^\lambda} = \boldsymbol{\alpha}_m^{(\lambda\varepsilon)} + \frac{1}{2\pi\varepsilon_0}\sum_{j\neq m} \boldsymbol{\alpha}_m^{(\lambda\varepsilon)} \cdot \frac{3\boldsymbol{\alpha}_j \cdot \mathbf{R}_{mj}\mathbf{R}_{mj} - \boldsymbol{\alpha}_j R_{mj}^{\,2}}{R_{mj}^{\,5}} + \frac{3}{2M\pi\varepsilon_0}\sum_{j\neq m}\left[\frac{\boldsymbol{\alpha}_{j\varepsilon}\boldsymbol{\alpha}_m \cdot \mathbf{R}_{mj} + \boldsymbol{\alpha}_j \cdot \mathbf{R}_{mj}\boldsymbol{\alpha}_{m\varepsilon} + \boldsymbol{\alpha}_m \cdot \boldsymbol{\alpha}_j R_{mj\varepsilon}}{R_{mj}^{\,5}} - 5\frac{\boldsymbol{\alpha}_j \cdot \mathbf{R}_{mj}\boldsymbol{\alpha}_m \cdot \mathbf{R}_{mj} R_{mj\varepsilon}}{R_{mj}^{\,7}}\right] \tag{6}$$

In the derivation of the previous formula, we used the definition of molecular position as a geometrical center of M atoms with position vectors $\mathbf{r}^i$, $\mathbf{R}_m = M^{-1}\sum_{i=1}^{M}\mathbf{r}^i$, so that $M(\partial\mathbf{R}_m/\partial\mathbf{r}^\lambda) = 1$ is the identity matrix. We can thus formally carry out the computations with polarizability derivatives obtained for one molecule ($\boldsymbol{\alpha}_m^{(\lambda\varepsilon)}$) corrected by the interactions with the rest of the system. A similar approach using the "dressed" tensors was proposed for mixed molecular and colloidal systems.[44] The approximate expansions of the induced dipole in colloidal systems could be avoided using a matrix formulation;[45] for molecules of relatively weak polarizability, however, the perturbational approach presented above is sufficient. In eq 6, we originally adopted the atomic polarizability model[43] where all atoms were treated as polarizable spheres of isotropic polarizabilities α_i, $i = 1...NAT$. While this approach was useful for program tests and relating our method to previous works, we show more advanced results only, with the polarizability derivatives calculated ab initio by Gaussian.

From the coordinate derivatives of the polarizability (eq 6), the time derivatives were calculated by the chain rule, $\partial\alpha_{\alpha\beta}/\partial t = \sum_{k\varepsilon}(\partial\alpha_{\alpha\beta}/\partial r_\varepsilon^k)(\partial r_\varepsilon^k/\partial t)$, where $\partial r_\varepsilon^k/\partial t = v_\varepsilon^k$ are the MD (CPMD) velocities. The Fourier transforms of the polarizability derivatives

$$F_{\alpha\beta}(\dot{\alpha}_{\alpha\beta})(\omega) = \int_0^t \alpha_{\alpha\beta}(\tau)\mathrm{e}^{-i\omega\tau}\,\mathrm{d}\tau = \int_0^t \frac{i}{\omega}\dot{\alpha}_{\alpha\beta}(\tau)\mathrm{e}^{-i\omega\tau}\,\mathrm{d}\tau \tag{7}$$

provided the Raman intensity of the backscattered (180°) signal as

$$I_{180} = \frac{\omega K(7S_1{}' + S_0{}')}{t[1 - \exp(-\hbar\omega/(kT))]} \tag{8}$$

where $S_0 = \sum_{\alpha\beta}[(\mathrm{Im}\,F_{\alpha\alpha})(\mathrm{Im}\,F_{\beta\beta}) + (\mathrm{Re}\,F_{\alpha\alpha})(\mathrm{Re}\,F_{\beta\beta})]$, $S_1 = \sum_{\alpha\beta}[(\mathrm{Im}\,F_{\alpha\beta})(\mathrm{Im}\,F_{\alpha\beta}) + (\mathrm{Re}\,F_{\alpha\beta})(\mathrm{Re}\,F_{\alpha\beta})]$, k is the Boltzmann constant, T is the temperature, K is an arbitrary intensity multiplier, and $\hbar$ is the Planck constant. For all of the simulated spectra, the y-scale is arbitrary, since the absolute intensities are not measured.

Spectral Simulations from CPMD Clusters. In a static model, 20 $[\mathrm{Mg(H_2O)_6}]^{2+}$ and $[\mathrm{Cl(H_2O)}_N]^-$ ($N = 4-6$) complex geometries were selected from randomly chosen CPMD snapshots and subjected to restricted normal mode optimization[46,47] performed at the B3LYP/6-311++G**/PCM(H_2O) level. The program Qgrad[48] interfaced to Gaussian was used for the optimization; normal modes within i300(imaginary)...200 cm^{-1} were kept fixed. Thus, the water molecules could approximately be maintained at instantaneous CPMD positions, while higher-frequency modes could be relaxed. Note that the relaxation is required to get the true harmonic vibrational frequencies. For the optimized geometries, Raman spectra were calculated by Gaussian at the B3LYP/6-311++G**/CPCM(H_2O) approximation level and averaged. A Lorentzian bandwidth of 15 cm^{-1} was used for the spectra plotting.

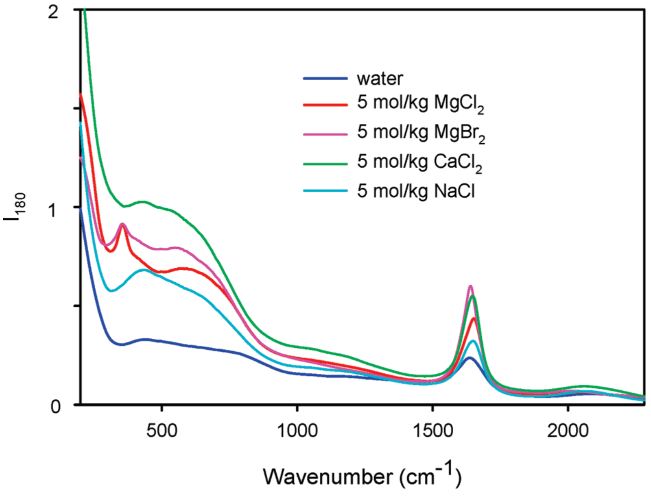


Figure 1. Experimental Raman spectra of water and $MgCl_2$, $MgBr_2$, $CaCl_2$, and NaCl solutions at 20 °C.

Results and Discussion

Experimental Spectra. In Figure 1, we display typical examples of the spectra for $MgCl_2$, $MgBr_2$, $CaCl_2$, and NaCl solutions. The signal from a quartz cell was subtracted, and spectra were normalized to the same accumulation time; nevertheless, the intensity scale should be considered rather informative. All salt solutions notably modify the Raman signal. However, such a significant increase in intensity of the HOH bending mode (~1640 cm^{-1}) in the salt solutions was recorded that it cannot be explained solely by variations in the collection angle of Raman scattering caused by changes of the solution refractive indices. This result is surprising, since there are fewer water molecules per volume in the salt solutions than in pure water. Moreover, in the set of $MgCl_2$ solutions in D_2O, an increase of DOD bending and decrease of OD stretching mode intensity with increasing concentration were recorded (see later, Figure 7), which also suggests that the changes in aqueous Raman scattering caused by the salts reflect the solution structure.

The changes in relative intensities (spectral shapes) are more specific and occur mostly within the 200−1000 cm^{-1} range. This region comprises the librational and rotational motion of water molecules, probably inseparable form the hydrogen bond stretching.[16,18,27,49] Most probably, the observed line shapes reflect the structure of the metal cation hydration spheres, while the anions give less specific signals. For example, the divalent Mg and Ca metals perturb the spectra more than the monovalent sodium. Potassium and lithium chlorides behaved similarly to the sodium chloride and are not shown. The specific "355 cm^{-1}" magnesium band is present in Raman spectra of both in $MgCl_2$ and $MgBr_2$ salt solutions. There are no indications that the

TABLE 1: Experimental Raman Peak Frequencies (cm^{-1}) of Various Aqueous Salt Solutions

	$\nu(M\cdots OH_2)$	$\delta(HOH)^a$	$\delta(HOH)^b$	$\nu(H\cdots O) + \delta(HOH)$
Water Solutions				
pure water 20 °C		1636	1638(115)	2090
$MgBr_2$ 5 mol/kg 20 °C	355	1642	1639(70)	2011
MgI_2 5 mol/kg 20 °C	325	1625	1628(62)	
$CaCl_2$ 5 mol/kg 20 °C		1647	1644(72)	2060
NaCl 5 mol/kg 20 °C		1649	1645(89)	2063
LiCl 5 mol/kg 20 °C		1648	1645(95)	2063
$MgCl_2$ 5 mol/kg 20 °C	355	1651	1647(83)	2052
$MgCl_2$ 2.5 mol/kg 20 °C	357	1648	1646(101)	2068
$MgCl_2$ 1.3 mol/kg 20 °C	361	1644	1644(111)	2076
$MgCl_2$ 5 mol/kg 0 °C	355	1651	1647(80)	2048
$MgCl_2$ 5 mol/kg 48 °C	352	1651	1646(83)	2018
D_2O Solutions				
pure D_2O 20 °C		1205	1204(72)	1554
$MgCl_2$ in D_2O 5 mol/kg 20 °C	337	1216	1210(57)	1516
$MgCl_2$ in D_2O 2.5 mol/kg 20 °C	337	1210	1208(70)	1522
$MgCl_2$ in D_2O 1.3 mol/kg 20 °C	332	1206	1206(73)	1552

[a] Peak maximum. [b] Peak center and width (in parentheses) from a Lorentzian band fit.

halogen anions would form a stable complex with Mg^{2+}. Nevertheless, from the comparison of the bromide and chloride spectra, it is clear that also the anions cause minor spectral changes. However, their interpretation goes beyond the scope of this study. Supposedly, the role of the metal cation is dominant and the interaction of simple halogenide ions with water is weaker.[1,5,12] For example, the computations indicate that the shortest Mg−O distance (∼1.95 Å) is much smaller than a Cl−H hydrogen bond (∼2.2 Å).

Individual peak positions are summarized in Table 1 for a wider variety of salts and experimental conditions. For example, the Mg−O stretching band moves to 352 cm^{-1} at 48 °C; in MgI_2 solution, its line shape (not shown) is significantly distorted and it is centered even lower, at 325 cm^{-1}. In fact, a new band appears also at 225 cm^{-1} for MgI_2. A similar shoulder at 202 cm^{-1} is apparent also for $MgBr_2$.

The HOH bending vibration (at 1636 cm^{-1} in pure water) moves up for most salts (except for MgI_2, where it goes down by ∼10 cm^{-1}), up to 1651 cm^{-1} for concentrated $MgCl_2$ solutions. This band also becomes significantly narrower for the salt solutions. The fitted Lorenzian bandwidth (cf. Table 1) changes from 115 cm^{-1} in pure water to 62 cm^{-1} for MgI_2. The combination band (HOH bending and hydrogen bond stretching) centered at 2090 cm^{-1} in pure water moves down, up to 2011 cm^{-1}, and disappears completely for MgI_2. However, it should be noted that the MgI_2 solution is a brown absorbing liquid and its baseline may be significantly distorted. Especially the combination signal (>2000 cm^{-1}) may be hidden in the background radiation. The temperature of the sample around the laser beam may be elevated, unlike for the other salts.

The $MgCl_2$ spectral trends in D_2O mostly follow those for H_2O. All vibrational frequencies are shifted down due to the larger mass of deuterium. Also, the bandwidth of the DOD bending band in pure D_2O (72 cm^{-1}) is significantly smaller than that for the HOH (115 cm^{-1}) in H_2O, even more than it would correspond to the central frequency ratio (1204/1638).

Hydrated Metal-Ion Geometries. Although the dynamical coordination number of the Mg^{2+} is known to be lower than six even in ideal conditions ($N \sim 5.85$),[1] the ideal octahedral geometry with $N = 6$ well represents the predominant structure of the first hydration shell of the magnesium ion.[6] In Table 2, the metal−water oxygen and OH bond distances calculated for the $[Mg(H_2O)_6]^{2+}$ complex at different levels are compared to the $Mg^{2+}\cdot H_2O$ system, and to the $[Ca(H_2O)_6]^{2+}$ and $[Na(H_2O)_6]^{+}$ complexes. More ab initio results are compiled in Table 1s of the Supporting Information. The calculated vacuum Mg−O distances are similar to those reported previously;[6] however, embedding the system into the polarizable dielectrics (CPCM) shortens the distance in $[Mg(H_2O)_6]^{2+}$ by ∼0.016 Å and improves thus the agreement of the calculations with experimental X-ray studies, where an average Mg−O distance of 2.063−2.068 Å is encountered.[6] On the other hand, the distances in the $Mg^{2+}\cdot H_2O$ bimolecular system are too short, even with the CPCM model, which reflects known deficiencies of the continuum approach in representations of strong polar interactions.[50,51] The Amber MD geometry parameters are reasonable; nevertheless, the quantum CPMD clearly provides more reliable data.

The calcium and sodium ions also form computationally stable octahedral complexes, but the distances between the metal and the water oxygen are significantly larger than those for Mg. For Ca^{2+}, this can be partially attributed to its larger covalent radius than Mg^{2+}. The calculated metal−water oxygen distances agree with the previously reported theoretical results[52] as well as with the known ionic radii.[1] Interestingly, the calculated Na−O and Ca−O distances are quite similar and, depending on the approximation used, one can be smaller or longer than the other. The OH distance does not seem to be significantly perturbed by the water coordination to the metals. The HF and higher-correlated MP2/MP4/CCSD(T) and B3LYP method give very similar results; the larger aug-cc-pVTZ basis brings only minor shortening of the Mg−O distance. On the basis of the comparison with the presumably most reliable MP4 and CCSD(T) computations, we can conclude that the computationally faster B3LYP method is adequate for a reliable description of the metal−water cluster geometries. The results thus suggest that the Mg−O link behaves to some extent like a stable single covalent bond. An analogous strong Mg−O interaction was also observed for Mg^{2+} methanol solutions.[22]

The geometry parameters obtained from the static CPMD computations (Table 2) were quite similar to those observed for the ab initio optimized clusters. Next, we look at the average geometries during the dynamics. The simulations with the HyperChem starting geometries and those starting with the arbitrarily T_h symmetrized $Mg(H_2O)_6$ initial structure provided very similar results. In accord with previous studies,[2,4] the radial distribution function of water oxygens around the magnesium ion (Figure 2) exhibits an exceptionally sharp and narrow peak at ∼2.1 Å. A similar peak can also be observed for concentrated $MgCl_2$ solution (dashed line in Figure 2), where the number of water molecules coordinated around Mg^{2+} was limited to N ∼3. Nevertheless, the CPMD approach reveals more differences between the calcium and sodium ions than the static computations summarized in Table 2; the Ca^{2+} binds the water molecules much more efficiently and perturbs the water ordering to a larger distance than Na^{+}, although less than Mg^{2+} exhibiting a secondary peak at ∼4.2 Å. Test MD computations reproduced this peak also with the Amber99 force field, and excluded an artifact influence of the limited box size (Figure 1s in the Supporting Information). The CPMD distribution gives more realistic Mg···O distances than MD; both dynamical models predicted a much broader chlorine···hydrogen distance distribution than for the Mg (cf. Figure 2s in the Supporting Information).

TABLE 2: Calculated Geometry Parameters (Å) for Some Model Metal−H_2O Clusters

	$[MgH_2O]^{2+}$		$[Mg(H_2O)]_6{}^{2+}$		$[Ca(H_2O)_6]^{2+}$		$[Na(H_2O)_6]^{+}$	
level	$d_{Mg-O}{}^{a}$	d_{H-O}	d_{Mg-O}	d_{H-O}	d_{Ca-O}	d_{H-O}	d_{Na-O}	d_{H-O}
Amber99	1.878	0.996	2.000	0.981	2.643	0.975	2.421	0.964
BLYP/CPMD			2.120	0.979	2.427	0.979		
CCSD(T)/6-311++G**	1.956	0.975	2.100	0.965	2.399	0.965	2.355	0.961
B3LYP/6-311++G**	1.942[a]	0.980[a]	2.111[a]	0.948[a]	2.414[b]	0.974[b]	2.386	0.963
B3LYP/aug-cc-pVTZ	1.916	0.979	2.100	0.966			2.393	0.963
$B3LYP_{CPCM}$/6-311++G**	2.039	0.978	2.084	0.965	2.368	0.977	2.397	0.972
Exp.			2.063[a]		2.397[c]			

[a] Reference 6. [b] Reference 58. [c] Reference 59.

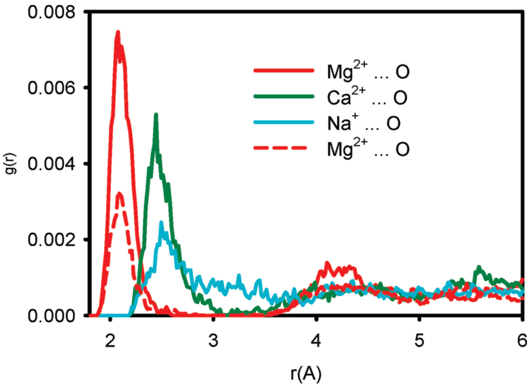


Figure 2. Calculated (CPMD) radial distribution functions for the Mg^{2+}, Ca^{2+}, and Na^{+} hydrated metal ions. For the $Mg^{2+}\cdots O$ distance, the calculated RDF in a 5 mol/kg $MgCl_2$ solution is also plotted as the dashed line.

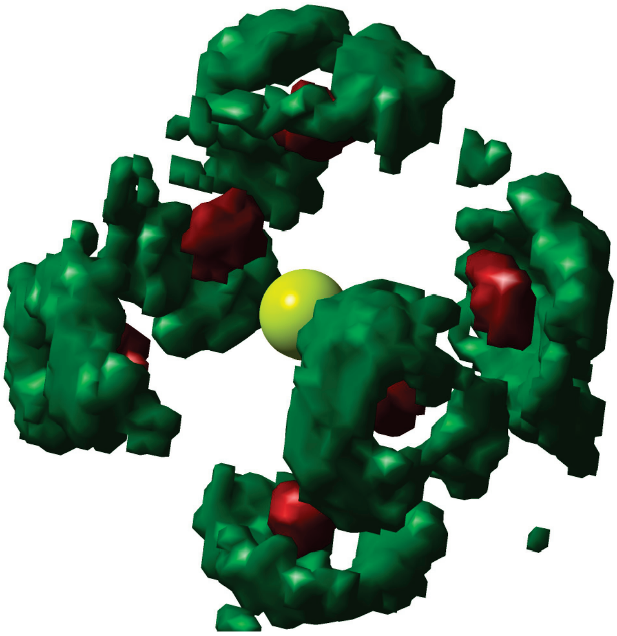


Figure 3. Calculated (CPMD) distributions of the aqueous oxygen (red) and hydrogen (green) atoms around the central Mg^{2+} ion.

From the calculated water 3D probability distribution in the octahedral magnesium complex in Figure 3, we see that the oxygen atoms indeed oscillate only moderately around the equilibrium positions. On the other hand, the hydrogen positions are dispersed more as the water molecules bound in the complex are perturbed by hydrogen bonds to molecules from the second hydration sphere.

Bond Strengths and Vibrational Frequencies. The relatively strong Mg^{2+}−O binding is reflected in the vibrational properties of the $[Mg(H_2O)_6]^{2+}$ complex. Harmonic frequencies calculated for optimized structures at seven approximation levels are summarized in Table 3. The mode assignment is based on potential energy distributions of internal vibrational coordinates as well as on a visual inspection of the normal mode displacement. The six Mg^{2+}−O bonds generate three normal modes of A_g, T_u (triple degenerate), and E_g (double degenerate) symmetries. Note that, for the T_h point group only, the gerade symmetry modes (A_g, E_g, T_g) are Raman active (cf. Table 4).[53] Ungerade modes are active in IR and can be assigned to previous observations (Table 3, ref 7). Although due to dynamical symmetry distortions the T_h symmetry rules may not hold in solution, from the calculated intensities (Table 4), the totally symmetric "breathing" mode clearly dominates in the Raman spectrum within the Mg−O stretching region. Unlike Mg^{2+}, the Ca^{2+} and Na^{2+} octahedral aqueous clusters exhibit metal−O stretching frequencies (Table 5) too low to be detectable.

As discussed before, although lacking the d-orbitals, magnesium binds to water as strongly as much heavier elements.[7] Our NBO analysis indicates that the relative strength of the Mg−O bond is also reflected in its partially covalent character. A "non-bonding" natural lowest unoccupied molecular orbital (LUMO)[34] centered on the metal ion (Figure 4) was identified that could clearly be associated with the magnesium binding to the water molecules in the complex. Its partial occupations and the metal natural charges are summarized in Table 6, also for other metals and clusters with chlorine. The nonzero occupancy of the natural LUMO indicates that it can be associated with the metal binding to the water molecules for all of the Na^{+}, Ca^{2+}, and Mg^{2+} metal ions. The highest value of the occupancy for Mg thus correlates with the strongest metal−water bond. Moreover, the calculated natural charge of the metal ion is lower than an ideal ionic charge (+1 for Na, +2 for Mg and Ca) by approximately the amount of the natural LUMO occupancy. Perhaps surprisingly, for the $[Mg(H_2O)_6]^{2+}$ cluster, there are negligible differences in the occupancy and the charge for the vacuum and solvent environments (Table 6).

The water lone-pair orbital aligned along the metal−O bond slightly differs for the three ions (Figure 4) as well. It is most extended in the $[Mg(H_2O)_6]^{2+}$ complex, where it is even partially localized on the metal. This simplified picture suggests that, although the metal−O bond strengths are different, the nature of the metal−water interaction is similar for all of the ions. Interestingly, the shape (not shown) and occupation (last two lines in Table 6) of the "unoccupied" central orbital remain similar in complexes where some water atoms are replaced by Cl^{-}.

All quantum methods in Table 3 provide reasonable estimates of the complex vibrations. For HF, the HOH bending frequencies (1792−1799 cm^{-1}) are somewhat higher, and the Mg−O breathing frequency (320 cm^{-1}) is lower than the experimental ones (1653 and 355 cm^{-1}, respectively). The larger aug-cc-pVTZ basis set brings only a minor improvement over the 6-311++G** basis set. Embedding the $[metal\cdot(H_2O)_6]$ clusters in implicit

TABLE 3: Fundamental Vibrational Frequencies of the $[Mg(H_2O)_6]^{2+}$ Complex[a]

	HF	MP2	B3LYP					exp.	(gas phase)
				b	*c*	*d*	*b*, *c*		
16 A_g	**1799**	**1678**	**1670**	**1664**	**1575**	**1644**	**1574**	**1653**	δ(HOH)
15 E_g	**1794**	**1674**	**1665**	**1660**	**1569**	**1635**	**1571**		δ(HOH)
14 T_u	1792	1673	1664	1659	1571	1634	1570		δ(HOH)
13 T_u	607	579	566	416	503	540	486	627[e]	H_2O wagging
12 T_g	**589**	**549**	**543**	**516**	**481**	**528**	**459**	**591**	wag + in plane rot.
11 T_u	487	417	439	415	407	419	405	421[e]	in plane H_2O rot.
10 A_u	433	406	427	407	277	372	256		axial H_2O rot.
9 A_g	**320**	**319**	**314**	**318**	**337**	**335**	**360**	**355**	$\nu_{breathing}(Mg-O)$
8 T_g	**453**	**380**	**410**	**392**	**203**	**309**	**185**		in plane H_2O rot.
7 T_u	388	386	378	368	163	300	163	178[e]	ν(Mg−O)
6 T_g	**299**	**237**	**286**	**275**	**132**	**285**	**134**		axial H_2O rot.
5 E_g	**253**	**266**	**249**	**246**	**263**	**263**	**273**	**309**[e]	$\nu(Mg-O)$
4 E_u	230	186	213	205	153	250	134		axial H_2O rot.
3 T_u	166	155	155	160	99	165	102	133[e]	δ(OMgO)
2 T_g	**136**	**92**	**122**	**129**	**59**	**135**	**46**		δ(OMgO)
1 T_u	100	97	96	99	42	88			δ(OMgO)

[a] By default, the 6-311++G** basis set was used. Raman-active modes are indicated in bold. [b] aug-cc-pVTZ basis set. [c] PCM solvent model (Gaussian). [d] COSMO solvent model (Turbomole). [e] IR spectra, ref 7, tentative assignment only.

TABLE 4: $[Mg(H_2O)_6]^{2+}$ Complex, Raman Intensities of Fundamental Vibrations[a]

	MP2		B3LYP	
			b	*c*
16 A_g	0.02	0.00	0.12	0.25
15 E_g	4.42	3.88	3.11	5.55
14 T_u	0	0	0	0
13 T_u	0	0	0	0
12 T_g	0.06	0.00	0.00	1.00
11 T_u	0	0	0	0
10 A_u	0	0	0	0
9 A_g	*0.89*	*1.13*	*2.45*	*1.58*
8 T_g	1.89	1.88	1.27	0.85
7 T_u	*0*	*0*	*0*	*0*
6 T_g	0.05	0.34	0.31	0.44
5 E_g	*0.18*	*0.20*	*0.23*	*0.17*
4 E_u	0	0	0	0
3 T_u	0	0	0	0
2 T_g	0.02	0.02	0.01	0.63
1 T_u	0	0	0	0

[a] By default, the 6-311++G** basis set was used; Mg−O stretching vibrations are indicated in italics. [b] aug-cc-pVTZ basis set. [c] PCM solvent model (Gaussian).

TABLE 5: [Metal·$(H_2O)_6$] Clusters, Calculated (B3LYP/6-311++G) Frequencies, and Intensities of Fundamental Vibrations**

	ω (cm^{-1})			*I* (au)		
M:	Mg^{2+}	Ca^{2+}	Na^+	Mg^{2+}	Ca^{2+}	Na^+
16 A_g	**1670**	**1673**	**1632**	0.00	0.09	0.22
15 E_g	**1665**	**1670**	**1630**	3.88	4.13	4.35
14 T_u	1664	1669	1629			
13 T_u	566	506	351			
12 T_g	**543**	**503**	**354**	0.00	0.08	0.04
11 T_u	439	444	131			
10 A_u	427	325	326			
9 A_g	*314*	*262*	*172*	*1.13*	*0.25*	*0.01*
8 T_g	**410**	**415**	**211**	1.88	1.22	1.67
7 T_u	*378*	*311*	*210*			
6 T_g	**286**	**220**	**205**	0.34	0.76	0.09
5 E_g	*249*	*237*	*138*	*0.20*	*0.02*	*0.02*
4 E_u	213	157	163			
3 T_u	155	88	54			
2 T_g	**122**	**77**	**i56**	0.02	0.32	
1 T_u	96	57	33			

(PCM, COSMO) solvent models improves the Mg−O stretching frequencies; however, the Gaussian implementation of PCM (and also that of COSMO, not shown) provides HOH bending frequencies that are too low. From Table 4, we see that the intensities are even more sensitive to the computational model than the frequencies. The PCM environment, for example, causes a modest (∼40%) increase against vacuum; the intensity calculated with the larger aug-cc-pVTZ basis is more than doubled against that obtained with the smaller 6-311++G** set.

As the ideally symmetric clusters may be disturbed in real solutions, we also investigated spectra obtained with a varying number of water molecules around the Mg^{2+} ion (Figure 3s in the Supporting Information). The position of the strongest Mg−O stretching band was found to be relatively stable for the higher-coordinated ($N > 2$) clusters, drifting only within 330−375 cm^{-1}. The minor fluctuations of the frequency are consistent with the minor differences observed experimentally for different anions that presumably influence the number of water molecules present around Mg^{2+}. For instance, $MgBr_2$ gives a broader peak than $MgCl_2$, there is a shoulder around 410 cm^{-1} and a broadband at 600 cm^{-1} for chloride, and a maximum 550 cm^{-1} for bromine. The computations indicate that even partially hydrated magnesium ions in concentrated solutions may be visible in the spectra due to the Mg−O vibration. The narrowness of the experimental peak and previous ab initio and CPMD computations[1,2] nevertheless suggest a strong tendency toward the hexahydrated complexes. The higher-frequency wagging modes (∼450−530 cm^{-1}) also significantly contribute to the Raman scattering. However, their resulting signal is much broader (cf. Figure 1), as they are mixed with the water hydrogen bond stretching and water librational motions.

Many features observed in the experimental Raman spectra could also be obtained from the CPMD trajectories by the modified Fourier transform method. For example, the simulated pure water spectrum in Figure 5 exhibits the main features observed experimentally, in particular the HOH bending band at 1608 cm^{-1} (observed at ∼1650 cm^{-1}) and the broad signal around 460 cm^{-1}. The simulated HOH bandwidth (∼130 cm^{-1}) is somewhat larger than the experimental one (115 cm^{-1}, Table 1). The anharmonic combination signal above 2000 cm^{-1} cannot be modeled using this approach, as the simulations based on classical trajectories are restricted to harmonic potentials.[25] For

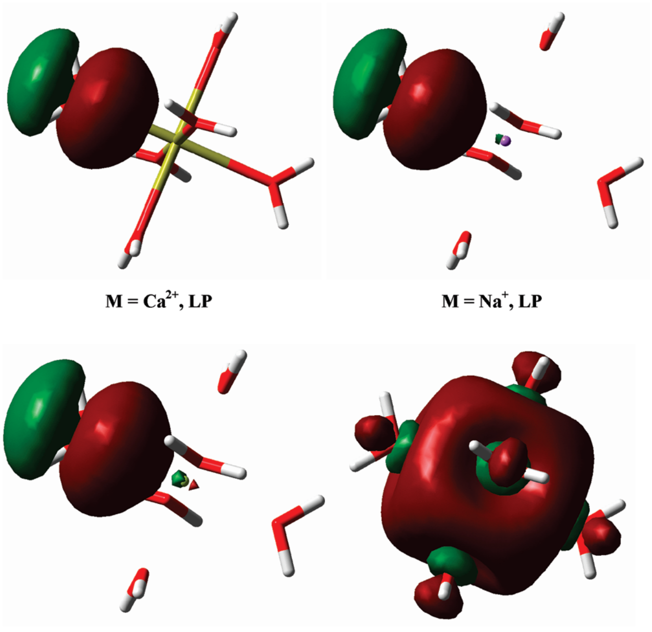

M = Ca^{2+}, LP
M = Na^{+}, LP
M = Mg^{2+}, LP
M = Mg^{2+}, LUMO

Figure 4. Comparison of the [M(H_2O)$_6$] NBO lone pair (LP) and NBO-LUMO (Mg only) orbitals (B3LYP/6-311++G**/PCM(H_2O) calculation).

TABLE 6: Occupancy of the Metal Nonbonded LUMO and the Natural Charge (*q*) in Model Complexes as Obtained by the NBO Analysis[34] of the B3LYP/6-311++G/PCM Wave Function**

complex	occ.	*q*
$[Na(H_2O)_6]^+$	0.073	0.927
$[Ca(H_2O)_6]^{2+}$	0.102	1.855
$[Mg(H_2O)_6]^{2+}$(vac)	0.172	1.821
$[Mg(H_2O)_6]^{2+}$	0.175	1.819
$[Mg(H_2O)_5Cl]^+$	0.201	1.783
$[Mg(H_2O)_4Cl_2]$	0.222	1.752

the $MgCl_2$ solution (red curve in Figure 5), the signal of the Mg−O stretching is clearly apparent; however, its Raman intensity is significantly underestimated and the width is too broad. Also, the narrowing of the HOH bending band observed for the salt solutions (Table 1) is not reproduced. Thus, we see that, although the CPMD-FT model has the potential to comprise both geometrical and dynamical factors of the metal−water interactions, its inherent approximations (Newtonian trajectories and harmonic potentials) may cause inaccuracies in the simulated spectra.

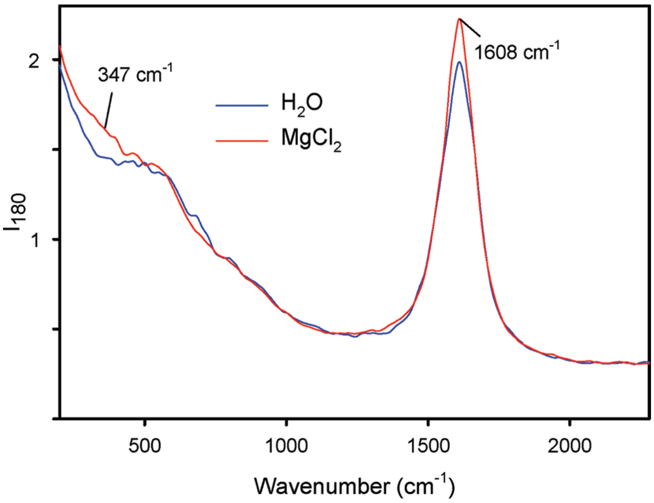


Figure 5. Raman spectra of pure water and 5 mol/kg $MgCl_2$ solution obtained from the CPMD trajectories (∼4 ps) by the Fourier transform method.

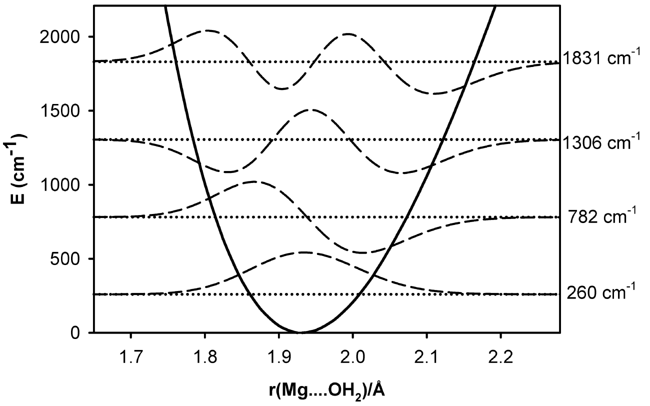


Figure 6. Calculated (MP2/cc-pVTZ) dependence of the relative energy of the $Mg^{2+}\cdot H_2O$ complex on the Mg−O distance, the adiabatic vibrational energy levels, and the corresponding wave functions.

To investigate a larger part of the potential surface governing the Mg−H_2O interaction at least qualitatively, we calculated energies of a complex consisting of Mg^{2+} cation and one water molecule, as dependent on the metal−oxygen distance (Figure 6). The C_{2v} symmetry was conserved, and all other coordinates were relaxed. Indeed, the symmetry of the potential is apparently distorted already for modest deviations from the minimum. Thus, the partial anharmonic character may account for some of the inconsistencies in the FT simulation above. When exact adiabatic one-dimensional wave functions are calculated numerically as dependent on the $r(Mg\cdots OH_2)$ coordinate,[54−57] the anharmonicity transfers to the asymmetry of the wave functions (Figure 6). Nevertheless, the vibrational energy levels remain approximately equidistant, similarly as in the harmonic case, within 2−5 cm^{-1}.

Finally, the Raman spectra of the octahedral hydrated $[Mg(H_2O)_6]^{2+}$ complex were generated by the force field diagonalization directly from 20 randomly selected CPMD clusters (subjected to the constrained normal-mode optimization, see Methods) and averaged. Such an approach corresponds to the instantaneous mode approximation,[26,27] as can be justified by the predominantly harmonic character of the vibrational motions (cf. Figure 6). The backscattered Raman spectra for the H_2O and D_2O complexes thus obtained are compared to the experimental spectra in Figure 7. The cluster approach provided similar spectral shapes as the FT method; however, the Mg−O signal (at ∼318 cm^{-1} for H_2O) is more distinct and more evidently corresponds to the experimental one (355 cm^{-1}). Strictly speaking, the calculated water bands (shoulder at 559 cm^{-1} and HOH bending peak at 1645 cm^{-1}) should not be directly compared to experiment which includes also free water molecules. Nevertheless, we see that also these frequencies and bandwidths reasonably correspond to the observed ones. The experimental frequency of the shoulder (605 cm^{-1}) is higher than the theoretical one, but this is not surprising, as the limited clusters lack the water hydrogen bond network. More importantly, the deuteration effects are well reproduced by the calculation, in particular the red shift of the Mg−O stretching signal (experimentally at 355 cm^{-1} in H_2O and 338 in D_2O, i.e., a difference of 17 cm^{-1}), calculated as 318−307 = 11 cm^{-1}. Thus, the assignment of the band to the Mg−O bond stretching is ultimately confirmed.

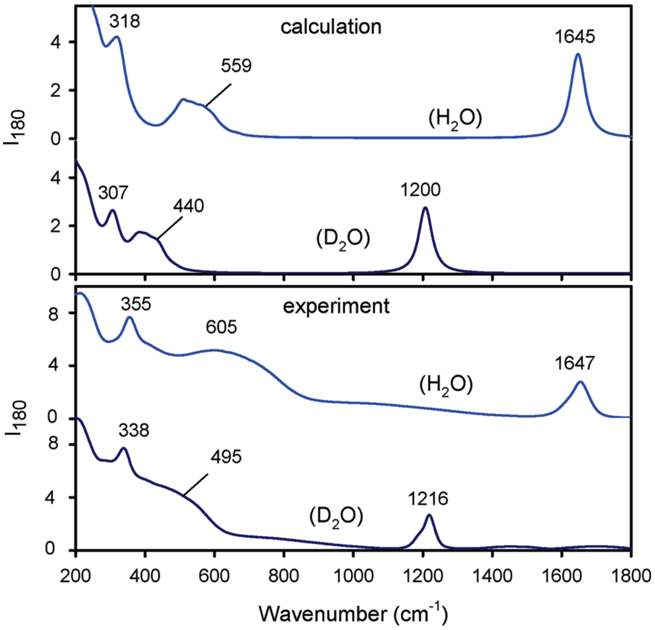


Figure 7. Upper panel: Calculated (B3LYP/6-311++G**/PCM) Raman spectra of the $[Mg(H_2O)_6]^{2+}$ and $[Mg(D_2O)_6]^{2+}$ complexes, average from 20 instantaneous CPMD geometries. Bottom panel: Corresponding experimental spectra of the 5 mol/kg $MgCl_2$ H_2O and D_2O solutions measured at 20 °C.

The cluster models are thus superior to the CPMD-FT simulations for the lowest-frequency vibrations. The FT method, nevertheless, can well reproduce the inhomogeneous HOH bandwidth, or the HOH → DOD bending frequency shift (1620 → 1210 cm^{-1}, experimentally 1647 → 1216 cm^{-1}), and the shift of the lowest-frequency signal (<1000 cm^{-1}, see Figure 4s in the Supporting Information). To obtain insight also into the Raman scattering of the hydrated chlorine anions, we simulated the spectra for 20 clusters of hydrated Cl$^-$ obtained by CPMD, containing four to six water molecules from the first hydration sphere. The normal mode optimization was not used in this case, not to affect the CPMD geometry distribution; otherwise, the same approximation (B3LYP/6-311++G**/PCM) was used as for Mg^{2+}. In comparison with the magnesium, chlorine anion causes a high-frequency shift of the HOH bending band (cf. Figure 5s in the Supporting Information) and an overall Raman intensity increase but no specific signal in the lowest-frequency region. Finally, spectra of hexacoordinated Mg^{2+} clusters with a varying number of water molecules and chlorine ions (not shown) and spectra of the large CPMD cluster corresponding to the 5 mol/kg $MgCl_2$ solution (Figure 6s in the Supporting Information) were simulated, which are, however, consistent with the previous simulations. They suggest that the spectral signal around 355 cm^{-1} originates in the Mg−O binding, and survives in distorted or incomplete hydration shells, in particular when some water molecules are replaced by chlorine.

Conclusions

Among the Raman spectra of concentrated salt solutions, we identified a strong signal for Mg^{2+} that could be, in accord with previous computational studies and experimental data, attributed to the strong magnesium−water binding. On the basis of the quantum-chemical computations, we could also determine that the specific 355 cm^{-1} peak is not exclusively associated with a complete octahedral hydration of Mg^{2+} but stems mostly from the vibrations of the Mg−O linkage. Partially covalent character of the bond was confirmed by the natural bond orbital analysis. Both the static cluster and dynamical models reproduced the most significant experimental spectral features. The isotopic shifts observed for the D_2O solutions were reproduced by the computations, and thus, the normal mode assignment could be verified. Hydrated chlorine anion clusters did not exhibit a Raman signal interfering with that of the metal. The data confirmed the large potential of the Raman spectroscopy for the studies of solution structure and dynamics. In particular, although the monatomic ions do not vibrate themselves, they moderate the solvent signal in a specific way that faithfully reflects their hydration patterns.

Acknowledgment. The present study was undertaken owing to a support from the Grant Agency of the Czech Republic (202/07/0732, P208/10/P356), Grant Agency of the Academy of Sciences (A400550702, M200550902, IAA400550701), research grant from EPSRC, the Luna (FZU) and University of Trømso computer facilities.

Supporting Information Available: CPMD-FT Raman spectra simulation of Mg^{2+} in H_2O and D_2O, hydrated Cl$^-$ spectral simulations, and CPMD and MD radial distribution function comparisons. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Collins, K. D. *Biophys. Chem.* **2006**, *119*, 271.
(2) Ikeda, T.; Boero, M.; Terakura, K. *J. Chem. Phys.* **2007**, *127*, 074503.
(3) Maguire, M. E.; Cowan, J. A. *BioMetals* **2002**, *15*, 203.
(4) Spångberg, D.; Hermansson, K. *J. Chem. Phys.* **2004**, *120*, 4829.
(5) Jungwirth, P.; Tobias, D. J. *J. Phys. Chem. A* **2002**, *106*, 379.
(6) Markham, G. D.; Glusker, J. P.; Bock, C. W. *J. Phys. Chem. B* **2002**, *106*, 5118.
(7) Mink, J.; Németh, C.; Hajba, L.; Sandström, M.; Goggin, P. L. *J. Mol. Struct.* **2003**, *661−662*, 141.
(8) Lenz, A.; Ojamäe, L. *J. Phys. Chem. A* **2006**, *110*, 13388.
(9) Schmidt, J. R.; Roberts, S. T.; Loparo, J. J.; Tokmakoff, A.; Fayer, M. D.; Skinner, J. L. *Chem. Phys.* **2007**, *341*, 143.
(10) Starzak, M.; Mathlouthi, M. *Food Chem.* **2003**, *82*, 3.
(11) Severson, M. W.; Devlin, J. P.; Buch, V. *J. Chem. Phys.* **2003**, *119*, 4449.
(12) Marcus, Y. *Chem. Rev.* **2009**, *109*, 1346.
(13) Kapitán, J.; Baumruk, V.; Kopecký, V., Jr.; Bouř, P. *J. Am. Chem. Soc.* **2006**, *128*, 2438.
(14) Kapitán, J.; Baumruk, V.; Kopecký, V., Jr.; Pohl, R.; Bouř, P. *J. Am. Chem. Soc.* **2006**, *128*, 13451.
(15) Kaminský, J.; Kapitán, J.; Baumruk, V.; Bednárová, L.; Bouř, P. *J. Phys. Chem. A* **2009**, *113*, 3594.
(16) Svishchev, I. M.; Kusalik, P. G. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1405.
(17) Halgren, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 7828.
(18) Ricci, M. A.; Ruocco, G.; Sampoli, M. *Mol. Phys.* **1989**, *67*, 19.
(19) Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I. F. W.; Mundy, C. J. *J. Phys. Chem.* **2006**, *110*, 3685.
(20) Dyer, P. J.; Cummings, P. T. *J. Chem. Phys.* **2006**, *125*, 144519.
(21) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, 55.
(22) Faralli, C.; Pagliai, M.; Cardini, G.; Schettino, V. *J. Chem. Theory Comput.* **2008**, *4*, 156.
(23) Amira, S.; Spångberg, D.; Hermansson, K. *J. Chem. Phys.* **2006**, *124*, 104501.
(24) Pagliai, M.; Cavazzoni, C.; Cardini, G.; Erbacci, G.; Parrinello, M.; Schettino, V. *J. Chem. Phys.* **2008**, *128*, 224514.
(25) Horníček, J.; Kaprálová, P.; Bouř, P. *J. Chem. Phys.* **2007**, *127*, 084502.
(26) Ahlborn, H.; Space, B.; Moore, P. B. *J. Chem. Phys.* **2000**, *112*, 8083.
(27) Bouř, P. *Chem. Phys. Lett.* **2002**, *365*, 82.
(28) Tu, A. T. *Raman Spectroscopy in Biology*; Wiley: New York, 1982.
(29) *Advances in Infrared and Raman Spectroscopy*; Clark, R. J. H., Hester, R. E., Eds.; Heyden: London, 1985; Vol. 12.
(30) Barron, L. D.; Zhu, F.; Hecht, L.; Tranter, G. E.; Isaacs, N. W. *J. Mol. Struct.* **2007**, *834−836*, 7.
(31) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(33) Polavarapu, P. L. *Vib. Spectra Struct.* **1984**, *13*, 103.

(34) Weinhold, F. Natural Bond Orbital Methods. In *Encyclopedia of Computational Chemistry*; Schleyer, P. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F. I., Schreiner, P. R., Eds.; Wiley: Chichester, U.K., 1998; Vol. 3, p 1792.

(35) *HyperChem 8.0.3*; Hypercube, Inc.: Gainesville, FL, 2007.

(36) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

(37) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.

(38) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926.

(39) CPMD version 3.11, 2006 IBM Corp., MPI für Festkörperforschung Stuttgart (http://www.cpmd.org).

(40) Troulier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993.

(41) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.

(42) Barron, L. D. *Molecular Light Scattering and Optical Activity*; Cambridge University Press: Cambridge, U.K., 2004.

(43) Polavarapu, P. L. *Vibrational spectra: principles and applications with emphasis on optical activity*; Elsevier: Amsterdam, The Netherlands, 1998; Vol. 85.

(44) Janesko, B. G.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 124704.

(45) Bouř, P. *J. Chem. Phys.* **2007**, *127*, 136101.

(46) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2002**, *117*, 4126.

(47) Bouř, P. *Collect. Czech. Chem. Commun.* **2005**, *70*, 1315.

(48) Bouř, P. *Qgrad*; Academy of Sciences, Prague: Prague, Czech Republic, 2002−2006.

(49) Ricci, M. A.; Signorelli, G.; Mazzacurati, V. *J. Phys.: Condens. Matter* **1990**, *2*, SA183.

(50) Bouř, P. *J. Chem. Phys.* **2004**, *121*, 7545.

(51) Caricato, M.; Ingrosso, F.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2005**, *122*, 154501.

(52) Pavlov, M.; Siegbahn, P. E. M.; Sandström, J. *J. Phys. Chem. A* **1998**, *102*, 219.

(53) Charney, E. *The Molecular Basis of Optical Activity*; Wiley-Interscience: New York, 1979.

(54) Čejchan, A.; Špirko, V. *J. Mol. Spectrosc.* **2003**, *217*, 142.

(55) Bouř, P.; Buděšínský, M.; Špirko, V.; Kapitán, J.; Šebestík, J.; Sychrovský, V. *J. Am. Chem. Soc.* **2005**, *127*, 17079.

(56) Sychrovský, V.; Buděšínský, M.; Benda, L.; Špirko, V.; Vokáčová, Z.; Šebestík, J.; Bouř, P. *J. Phys. Chem. B* **2008**, *112*, 1796.

(57) Podolsky, B. *Phys. Rev.* **1928**, *32*, 812.

(58) Bock, C. W.; Katz, A. K.; Markham, G. D.; Glusker, J. P. *J. Am. Chem. Soc.* **1999**, *121*, 7360.

(59) Carugo, O.; Djinovic, K.; Rizzi, M. *J. Chem. Soc., Dalton Trans.* **1993**, 2127.

JP9110508

THE JOURNAL OF
PHYSICAL CHEMISTRY A

ARTICLE

pubs.acs.org/JPCA

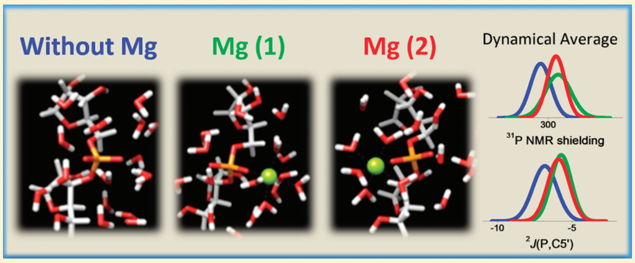
Calculating the Response of NMR Shielding Tensor $\sigma(^{31}P)$ and $^2J(^{31}P,^{13}C)$ Coupling Constants in Nucleic Acid Phosphate to Coordination of the Mg^{2+} Cation

Ladislav Benda,[†] Bohdan Schneider,[*,‡] and Vladimír Sychrovský[*,†]

[†]Institute of Organic Chemistry and Biochemistry, v.v.i., Academy of Sciences of the Czech Republic, Flemingovo Square 2, 166 10 Prague 6, Czech Republic

[‡]Institute of Biotechnology, v.v.i., Academy of Sciences of the Czech Republic, Vídeňská 1083, 142 20 Prague 4, Czech Republic

S *Supporting Information*



ABSTRACT: Dependence of NMR ^{31}P shielding tensor and ^2J(P,C) coupling constants on solvation of nucleic acid phosphate by Mg^{2+} and water was studied using methods of bioinformatic structural analyses of crystallographic data and DFT B3LYP calculations of NMR parameters. The effect of solvent dynamics on NMR parameters was calculated using molecular dynamic. The NMR calculations for representative solvation patterns determined in crystals of B-DNA and A-RNA molecules pointed out the crucial importance of local Mg^{2+} coordination geometry, including hydration by explicit water molecules and necessity of dynamical averaging over the solvent reorientation. The dynamically averaged ^{31}P chemical shift decreased by 2−9.5 ppm upon Mg^{2+} coordination, the chemical shielding anisotropy increased by 0−20 ppm, and the ^2J(P,C5′) coupling magnitude decreased by 0.2−1.8 Hz upon Mg^{2+} coordination. The calculated decrease of the ^{31}P chemical shift is in excellent agreement with the 1.5−10 ppm decrease of the phosphorothioate ^{31}P chemical shift upon Cd^{2+} coordination probed experimentally in hammerhead ribozyme (Suzumura; et al. J. Am. Chem. Soc. 2002, 124, 8230−8236; Osborne; et al., Biochemistry 2009, 48, 10654−10664). None of the dynamically averaged NMR parameters unequivocally distinguishes the site-specific Mg^{2+} coordination to one of the two nonesterified phosphate oxygen atoms of the phosphate determined by bioinformatic analyses. By comparing the limit cases of static and dynamically averaged solvation, we propose that mobility of the solvent has a dramatic impact on NMR parameters of nucleic acid phosphate and must be taken into account for their accurate modeling.

INTRODUCTION

Nucleic acids (NAs) perform a multitude of biological functions: DNA stores genetic information, mRNA mediates between DNA and the final protein, rRNA forms scaffold and catalytic center for protein synthesis, tRNA fetches amino acids to a growing polypeptide chain, various interference RNA molecules fine-tune the translation process by controlling splicing of mRNA, signal recognition particles steer transport of proteins across membranes. Many aspects of these functions are directly related to the exact three-dimensional structure. Both kinetic and thermodynamic stability of the folded NA molecules essentially depend on counterions balancing out the negative charge of the phosphodiester backbone. Most RNA folding pathways depend on the presence of a particular metal cation and its concentration.[1]

Metal solvation can be classified as either nonspecific or specific. A diffuse "cloud" of positively charged ions binds nonspecifically to a NA molecule to compensate its negative charge. This interaction is critical for stability of both DNA and RNA and can be described by phenomenological physical models.[2−5] A useful comparison of methods and approaches that integrates empirical facts and theoretical concepts has been reviewed.[6]

Metals are also known to bind to "specific" sites of DNA and especially RNA molecules; "specific" in this context means sufficiently long residence times or occupancies for metals to be identified by NMR and/or by crystallography. Specifically bound cations interact directly with NA atoms forming so-called inner shell complexes, in contrast to nonspecifically bound outer shell complexes. Specific binding requires at least partial dehydration of the interacting metal cation and NA atoms, typically highly hydrated phosphates. Energetically costly desolvation of both the phosphate and the ion causes a large local electrostatic attraction between the metal and NA to be stabilizing overall RNA or DNA structure only in such relatively rare cases when the metal binds to sites of highly negative potential with advantageous stereochemistry.[7] Therefore, nonspecific solvation by fully

Received: December 1, 2010
Revised: February 3, 2011
Published: March 02, 2011

ACS Publications © 2011 American Chemical Society

dx.doi.org/10.1021/jp1114114 | *J. Phys. Chem. A* 2011, 115, 2385–2395

hydrated metal cations is a thermodynamically generally preferred mode of binding.

Invaluable data about solvation of biomolecules are obtained from thermodynamic measurements. Various thermodynamic and volumetric properties of monovalent metals have been studied in the process of formation and melting of double helical RNA and DNA.[8]

Structural aspects of the first solvation shell have been studied in great detail in crystal structures of DNA double helical forms for bases[9] as well as for phosphates.[10,11] These studies clearly demonstrated that the first solvation shell is localized into well-defined sites that are specific for DNA form and sequence. For a long time structural data on RNA were scarce but hydration of the double helical A-RNA has been analyzed.[12] Recently, hydration of various base pairs in RNA can be visualized online using a web service SwS.[13] An interesting aspect of NA solvation, anion binding, has been studied by analysis of crystal structures and by MD simulations[14] that have been critically reviewed.[15]

Specific, or "site" binding (albeit less frequent) is important for its role in folding, stability, and also catalytic activity of NAs. Either type can bind to phosphates or to bases. In the case of RNA, bases bind metals, mostly Mg^{2+}, in the deep and narrow major groove of the double helical regions that has highly negative potential. One of such major groove motifs is for instance a pocket of deep negative electrostatic energy formed by the G-U tandems.[16] Also a dinucleotide motif, originally called adenosine platform[17] (that is not limited to adenines), binds dehydrated K^+.

Effects of metal ions on ^{31}P NMR shielding tensor of NA phosphate can be detected by NMR spectroscopy.[18−21] Different coordination sites in NAs have different affinity for metal ions and bind them with different residence times; an Mg^{2+} ion coordinated to a solvent exposed NA phosphate group cannot be expected to have as long residence time as the same ion trapped, for example, between two phosphates in the deep pocket of folded RNA. The structural flexibility of NAs also affects values of NMR parameters that respond to interactions with a metal ion. Therefore, the NMR detection of a metal ion coordinated to structurally rigid NA is likely to provide more structural details than that for highly flexible NA. For example, NMR shifts of the inner ring atoms in nucleobase measured for increasing concentrations of Cd^{2+} ion in hammerhead ribozyme[22] can be interpreted safely as the effect of direct coordination of Cd^{2+} to N7 nitrogen of guanine.[23] In contrast, structural assignments of NMR parameters at direct neighborhood of the phosphorus atom of the phosphodiester link pose a special challenge because phosphate groups represent structurally the most variable segment of NAs.[24−29] In addition, NMR parameters of NA backbone are difficult to interpret in structural terms because different backbone conformations can result in similar values of NMR parameters.[30−33] What causes further ambiguity of structural interpretation of metal coordination to the flexible phosphate group is a small range of the ^{31}P NMR shift in NAs ∼ 1−8 ppm[30,34−36] and the fact that the chemical shift changes due to backbone variation and metal coordination are likely to fall to the same interval.

For most common physical methods such as X-ray crystallography, NMR spectroscopy, UV/vis, IR, or Raman spectroscopies localization of alkali and alkali earth metals in NAs is a complicated task. NMR measurements probing inner-shell metal ion coordination to the NA phosphate are exceptional. One of the prominent examples is metal binding in the hammerhead ribozyme.[37] The so-called "Cd^{2+} rescue"[38] experiments unveiled the decrease of ^{31}P NMR shift by 1.5−10 ppm due to coordination of Cd^{2+} metal ion to the phosphorothioate.[39−41] The decrease of ^{31}P shift by 0.3 ppm upon addition of Mg^{2+} was measured in the native phosphate group of minimal metal-binding motif of hammerhead ribozyme.[42] We think that improvement of metal-ion detection techniques is desirable. That is why we initiated this theoretical work focused on calculation of ^{31}P NMR shielding tensor and indirect ^2J(P,C) coupling constants in NA backbone and their dependence on solvation by Mg^{2+} ion.

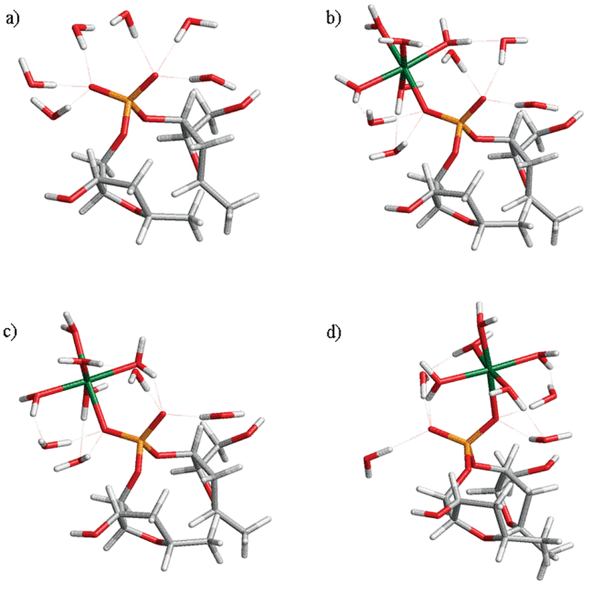


Figure 1. (a) Reference hydration model of NA phosphate B-6w with six water molecules as observed in B-DNA in crystals.[10] (b) B-Mg1-10w model of NA phosphate solvated by Mg-5w coordinated to the OP1 oxygen containing ten water molecules. (c) B-Mg1-9w model of NA phosphate solvated by Mg-5w coordinated to the OP1 oxygen containing nine water molecules. (d) B-Mg2-9w model of NA phosphate solvated by Mg-5w coordinated to the OP2 oxygen containing nine water molecules.

■ METHODS

This work consisted of three parts. First, we developed representative molecular models of NA phosphate solvation using bioinformatic analysis of NA crystals. Obtained models were then used for DFT calculations of NMR parameters. Finally, we calculated the effect of solvent reorientation by averaging the NMR parameters using molecular dynamics (MD).

Molecular Models. The phosphate group of NA backbone was modeled as a dinucleotide monophosphate with both bases replaced by methyl groups (Figure 1a). The "abasic dinucleotide" (sPs) was built in two conformations (Table 1) reflecting the "canonical" BI- and AI-forms of NA architecture described by seven backbone torsion angles (torsion values from Table 3 of Svozil et al.[27] were used). We note that the BI and AI patterns differ only slightly and particularly the ζ and α torsions adopt the same g-,g-conformation (Table 1). Two prochiral (nonesterified) phosphate oxygen atoms were distinguished by indexes OP1 and OP2 in accordance with the IUPAC

Table 1. Geometry Constraints Imposed on the sPs Molecule

torsion angle[a]	B-DNA	A-RNA
$\delta 0$	129	81
$\varepsilon 0$	184	212
$\zeta 0$	265	288
$\alpha 1$	303	295
$\beta 1$	176	174
$\gamma 1$	48	53
$\delta 1$	129	81

[a] In degrees. The torsion angles are sorted according to their occurrence in the sPs molecule from the 5′-terminus. The residue number is included in the torsion angle label.

nomenclature (in the B-form the C5′−O5′−P−OP1 torsion angle is ~55° and the C5′−O5′−P−OP2 one is ~185°).

Three hydration models were built: one for the B-form and two for the A-form. The former conformer was used as a reference model of hydration because the geometry of the first hydration shell in B-DNA is well described.[10] This reference model, denoted hereafter B-6w (Figure 1a), consisted of the sPs molecule in the BI conformation and six water molecules with their oxygens in positions determined previously by the analysis of B-DNA crystal structures.[10] The resulting complex was optimized (see below) with backbone torsions fixed to values presented in Table 1 and with spatial arrangement of the six waters constrained by fixing pseudotorsion angles O*w*−OP1−P−OP2 and O*w*−OP2−P−OP1 to the values from the bioinformatic model;[10] O*w* is oxygen of the *w*th water molecule ($w = 1, 2, ..., 6$) coordinating OP1 or OP2 (Figure 1a). All remaining coordinates in the water-coordinated sPs were freely optimized. All geometries in this study were obtained similarly, by neglecting small geometry fluctuations of solvent molecules in late cycles of energy minimization, because the default criteria of energy minimum convergence could not be fulfilled exactly in all hydrated and solvated sPs complexes. The negligible energy and geometry fluctuations of a particular model observed toward the end of the optimization ensure that the relaxed geometries provide representative structures for NMR calculations.

The two remaining hydration models were based on the A-form and hydrated according to the bioinformatic analysis of water positions in crystals of RNA. The protocol for obtaining averaged water and Mg^{2+} positions was the same as in our previous works, e.g., ref 9. The actual statistics for the RNA ensemble were as follows: 53 RNA crystal structures containing over 7700 nucleotides were analyzed; about two-thirds of them were classified as A-RNA. Out of 16 000 water molecules in these structures 13 000 were found within 3.2 Å from any nucleotide atom. Corresponding numbers for Mg^{2+} cations were orders of magnitude lower; of all 511, a total of 365 were located in the neighborhood of an A-RNA nucleotide. Individual nucleotides adopting the A-form were then overlapped including the crystallographic waters and magnesium ions that were found near them. In the last step, water and Mg^{2+} positions were averaged by the method of Fourier averaging.[9,43]

Resulting hydration sites around the A-RNA phosphate had spatial a arrangement similar to that in B-DNA,[10] but there were quantitative differences in their relative positions and importance. While hydration in B-DNA was formed by two "cones of hydration" with each cone consisting of three sites, one hydration site in A-RNA had a very low density and the presence of water at this site is questionable. Therefore, we built two models of A-RNA hydration, one with five and one with six water molecules around the sPs phosphate. A model further denoted A-5w contained five water molecules in their statistically determined positions. To form a A-6w model, the sixth water was added to A-5w at the position of crystallographically rather weak phosphate hydration site.

Optimization of geometries of RNA models A-5w and A-6w was done in a slightly different fashion compared to that of B-DNA model B-6w. We allowed full relaxation of water molecules coordinated to the phosphate and kept only the backbone torsions fixed to A-RNA values (Table 1).

Phosphate solvation models including the inner-sphere coordinated Mg^{2+} ion were built for both B-DNA and A-RNA forms. Geometry arrangement of Mg^{2+} ions present in RNA crystals is similar to that observed around organic phosphates:[11] there are two populated Mg^{2+} sites each lying symmetrically at 1.9 Å from OP1 and OP2, respectively, near the OP1−P−OP2 plane. The distance between the two sites is about 5 Å but it is unlikely that both the sites can be occupied by Mg^{2+} ion at the same time. Structural similarity between Mg^{2+} sites in a general diesterified organic phosphate with charge −1[11] and in A-RNA nucleotides suggests that mutual spatial arrangement between the phosphate and Mg^{2+} ion is a robust local property that depends marginally on the broader neighborhood of the complex.

Initial geometries of B-DNA models with Mg^{2+} ion were assembled by replacing the water molecule in the B-6w hydration shell nearest to the position observed for the Mg^{2+} ion with the preoptimized complex of penta-hydrated Mg^{2+} (Mg-5w), as illustrated in Figure 1b. In the next step, another water molecule coordinated to the other prochiral phosphate oxygen was removed and replaced by one water molecule from Mg-5w by rotating Mg-5w around the OP1/2-Mg axis; the optimized complexes are shown in Figure 1c,d. Initial geometries of Mg^{2+} complexes in A-RNA conformation were constructed in a slightly different fashion. Water molecules coordinated to one of the two prochiral phosphate oxygens were removed from the A-5w geometry and were replaced with the Mg-5w complex so that the magnesium atom would coincide with the maximum of Mg^{2+} statistical crystallographic distribution.

The sPs phosphate was therefore solvated by one Mg^{2+} ion and five (models B-Mg1-10w and B-Mg2-10w), four (models B-Mg1-9w and B-Mg2-9w), or three (A-Mg1-8w and A-Mg2-8w) water molecules in the first solvation shell. In all these cases the first solvation shell of the ion was saturated completely. During geometry optimizations of sPs complexes with Mg^{2+} ion only the backbone torsions were constrained to achieve systematic comparison of the calculated NMR parameters; other geometry parameters were freely relaxed.

DFT Computational Details. DFT calculations were performed using the B3LYP functional[44] as implemented in Gaussian 03, revision C02 program.[45] A standard 6-31+G(d) basis set[46−49] was used in all geometry optimizations. GIAO calculations of ^{31}P NMR shielding tensor[50] and indirect $^{2}J(^{31}P,^{13}C)$ scalar coupling constants[51−53] were carried out with the IGLO-III[54] or cc-pCVTZ[55−58] atomic basis set for H, C, O, and P atoms and the Ahlrichs VTZ[59] or cc-pCVTZ basis set for Mg atom. In the following we skip the isotope indexing for $^{2}J(P,C)$ couplings. All geometry optimizations and NMR calculations (if not specified otherwise) were done with the implicit polarizable continuum model (PCM) of water solvent.[60] The chemical

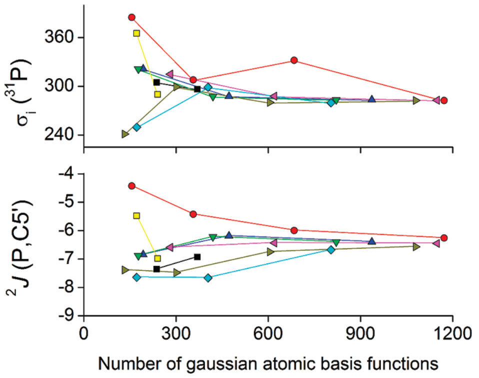


Figure 2. Basis set convergence analysis for NMR isotropic shielding σ_i (in ppm) and the Fermi-contact contribution of ^2J(P,C5′) coupling (in Hz) calculated using the B3LYP functional in ethyl methyl phosphate with cc-pVnZ (red circles; n = D, T, Q, 5), cc-pCVnZ (blue triangles pointing up; n = D, T, Q), cc-pxVnZ (light green triangles pointing down; n = D, T, Q), aug-cc-pxVnZ (pink triangles pointing left; n = D, T, Q), IGLO-n (black squares; n = II, III), pcJ-n (tan triangles pointing right; n = 0, 1, 2, 3), aug-pcJ-n (cyan diamonds; n = 0, 1, 2), and Pople's 6-31+G(d) and 6-311++G(d,p) basis sets (yellow squares).

shielding anisotropy (CSA, $\Delta\sigma$) was defined in the usual way:

$$\Delta\sigma = \sigma_{33} - \left(\frac{\sigma_{11} + \sigma_{22}}{2}\right)$$

where σ_{nn} are the σ-tensor principal values defined such that $\sigma_{11} < \sigma_{22} < \sigma_{33}$.

To test the quality of different atomic basis sets, we ran NMR calculations for the ethyl methyl phosphate molecule (ePm) optimized on the same level as all complexes in this study. The convergence test included the use of IGLO-II/III,[54] Dunning's cc-pVnZ[55,56] (n = D, T, Q, 5), (aug-)cc-pCVnZ[55−58,61] (n = D, T, Q), Jensen's (aug-)pcJ-n[62−64] (n = 0−3), and Pople's 6-31+G(d)[46−49] and 6-311++G(d,p)[49,65,66] basis sets. All the basis sets were downloaded from the EMSL Basis Set Exchange web portal.[67,68] A possibility to save computer time in NMR calculations by applying locally dense basis sets[69] was also tested.

The effect of atomic basis set on NMR parameters of the phosphate group was summarized in Figure 2. The cc-pVnZ basis without the tight-core functions converged very slowly with the number of atomic basis functions; inclusion of these functions in the cc-pCVnZ basis dramatically improved the convergence. The cc-pCVTZ calculations got very close to the basis set limit with the computations being still affordable. Keeping the tight core basis functions in the cc-pCVnZ basis only for the phosphorus and all carbon atoms in the ePm molecule (the cc-pxVnZ calculations) actually resulted in the same trends as with the complete cc-pCVnZ basis but with slightly lower computational demands. Addition of diffuse functions on all atoms (the aug-cc-pxVnZ calculations) improved the cc-pxVnZ results even further at somewhat higher computational cost. The popular NMR basis sets IGLO-II/III performed rather well and IGLO-III represents a reasonable choice for bigger systems. Surprisingly, Jensen's basis sets pcJ-n and aug-pcJ-n designed specifically for J-coupling NMR calculations converged slower than some other general-purpose basis sets. The reason for that might be that our optimization of the pcJ-n and aug-pcJ-n basis sets was based on calculations of one-bond scalar couplings in selected molecules[63] with a bonding situation different from that in NA phosphate. Very good results for all monitored NMR parameters were obtained with rather small 6-311++G(d,p) basis set of triple-ζ quality. The cost-effectivness of this basis encouraged us in using it further in dynamical averaging of NMR parameters.

Dynamical Solvent Averaging. NMR parameters of polar or charged groups are known to be significantly influenced by molecular environment,[70,71] and dynamical averaging over the solvent reorientation is, therefore, necessary for their reliable calculation. Consequently, static picture was complemented by calculation of dynamical motion of the solvent with MD method. GpU dinucleotide monophosphate with A-RNA backbone conformation (Table 1) was arbitrarily selected as a model system.

Three phosphate solvation patterns were considered: direct coordination of pentahydrated Mg^{2+} ion (Mg-5w) to the OP1 oxygen further denoted "mg1", coordination of Mg-5w to the OP2 oxygen denoted "mg2", and the hydrated phosphate denoted "nomg". Each starting structure was embedded in an orthogonal box of water molecules (for details, see Supporting Information). The MD simulations were performed using standard Amber 99 force field[72] with parmbsc0[73] refinement for NAs as implemented in the Amber 10 program package.[74] The MD protocol consisted of initial solvent minimization, full system minimization, solvent heating to 300 K, NPT equilibration (300 K, 1 atm), and 2 ns NPT production run (300 K, 1 atm). Further details are given in the Supporting Information.

The NMR parameters depend both on the solvent orientation/coordination and on the adjacent torsion angles of the phosphate.[30−32,75] To calculate separately, the exclusive effect of the solvent positions of backbone atoms were restrained to A-RNA geometry using a 500 kcal mol^{-1} Å^{-2} Cartesian force constant (Supporting Information).

Snapshots were taken every 50 ps from the three MD trajectories (mg1, mg2, nomg) resulting in 3 × 41 MD geometries in total. In the subsequent NMR calculations only the nearest layers of the solvent surrounding the phosphate group were preserved, i.e., only those water molecules with oxygen atoms appearing within the defined sphere of radius r around the phosphorus atom. Furthermore, the G and U bases of GpU molecule were replaced by hydrogen atoms at the C1′ carbon. Validity of this approach was tested previously for all backbone J-couplings in an ApU dinucleotide;[75] the effect of removing NA bases on ^{n}J(P,X) coupling constants calculated with inclusion of PCM hydration was smaller than 0.3 Hz.

Convergence of NMR parameters with the thickness of the solvent shell around the phosphate was carefully tested for r = 0, 4.5, 5.5, 6.5, 7.5, and 8.5 Å (Figure 3). The resulting set of MD geometries was directly, i.e., without any geometry relaxation, subjected to NMR calculation including implicit PCM hydration on top of explicit solvent layer (the level of theory is described below). The contributions to NMR parameters due to extension of explicit hydration by including individual solvent layes in Figure 3 are de facto equal to errors that one would make by substituting explicit hydration with implicit PCM solvent.

Water molecules inside the sphere with r = 4.5 Å covered the first hydration shell of the phosphate. Because its effect is apparently large, this level of explicit hydration must be included in the NMR calculations. Higher layers of explicit solvent with r > 4.5 Å had only a minor effect on the NMR parameters. We anticipate that after averaging over MD snapshots (Figure 3) the overall effect of explicit treatment of higher solvent layers can be sufficiently modeled with PCM.

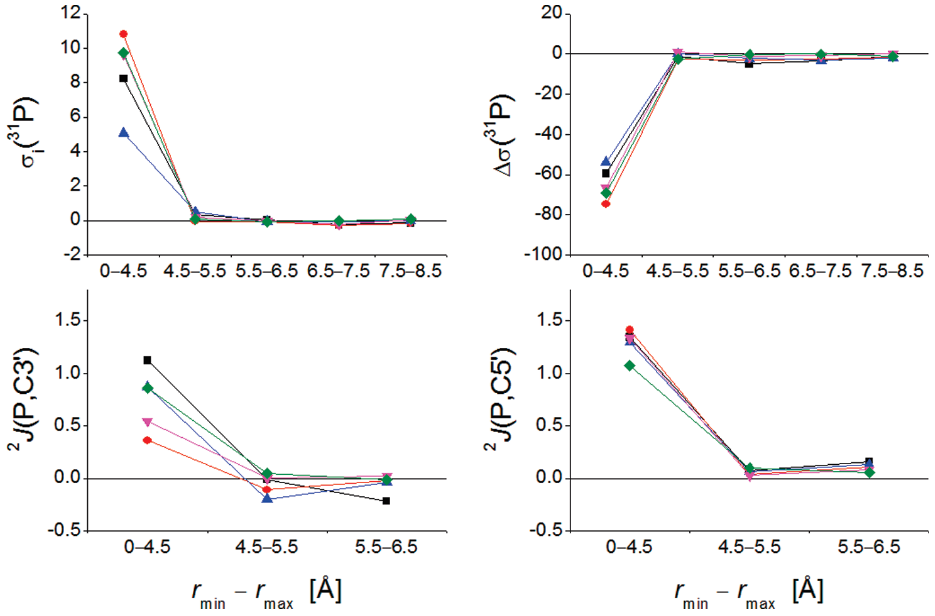


Figure 3. Contributions of individual solvent layers to isotropic ^{31}P NMR shielding (σ_i; in ppm), ^{31}P shielding anisotropy ($\Delta\sigma$; in ppm) and the two ^2J(P,C) spin–spin coupling constants (in Hz) evaluated for five selected MD snapshot geometries (depicted with different colors), calculated as a difference between the values for r_{max} and r_{min}.

Table 2. NMR Parameters[a] Calculated with B3LYP/IGLO-III Method in the sPs Molecule with Different Models of Hydration

complex[b]	$\delta_i(^{31}P)$	$\Delta\sigma(^{31}P)$	$^2J(P,C3')$	$^2J(P,C5')$
B	−1.3	243	−9.2	−8.6
B-PCM	2.7	189	−8.2	−8.2
B-6w	−1.3	158	−8.1	−7.8
B-6w-PCM	0.0	143	−8.1	−8.2
B-OP1-3w	−1.7	207	−7.3	−8.8
B-OP1-3w-PCM	1.3	170	−7.8	−8.4
B-OP2-3w	−1.4	215	−9.6	−7.2
B-OP2-3w-PCM	1.6	172	−8.5	−7.9
A-5w-PCM	−0.5	154	−7.3	−7.1
A-6w-PCM	1.5	144	−6.1	−7.0

[a] The isotropic chemical shift δ_i in ppm (referenced to the B-6w-PCM isotropic shielding σ_i^{ref} = 294.3 ppm), the chemical shielding anisotropy $\Delta\sigma$ in ppm, and the indirect spin–spin coupling $^2J(P,X)$, X = C3′, C5′ in Hz. [b] Characterization of the models: B: sPs molecule in B-DNA conformation in the gas phase; B-PCM: B immersed in PCM model of water solvent; B-6w: explicit hydration of B-DNA from crystal data; B-OP1-3w: explicit hydration of the OP1 oxygen only; B-OP2-3w: explicit hydration of the OP2 oxygen only; A-5/6w-PCM: explicit hydration of A-RNA with five and six water molecules respectively immersed in PCM.

The results encouraged us to use a rather economic solvent model in dynamical averaging of NMR parameters. The solvent sphere with r = 5.5 Å covered the first hydration shell of both the phosphate and coordinated Mg^{2+} ion. This solvation model including PCM hydration was used in the B3LYP NMR calculations for the 3 × 41 MD snapshot geometries. The 6-311++G** basis set was used for the atoms of solute, 6-31+G* for the Mg^{2+} ion, and 6-31G for the water molecules. All DFT calculations in MD part of this work were carried out using Gaussian 09, revision A02.[76]

Optimal MD geometry of NA backbone could differ from that obtained with DFT methods due to inherent difference between classical force-field potential and quantum-mechanical energy surface. Even a small difference in geometry could then result in a substantial variation of calculated NMR parameters. We, therefore, calculated NMR parameters for both the unrelaxed MD snapshot geometries and the MD geometries with partially DFT optimized NA backbone (Supporting Information).

■ RESULTS AND DISCUSSION

Effect of Hydration on NMR Parameters, the Static Model. The calculations included various combinations of explicit and implicit PCM hydration of the sPs molecule in B-DNA and A-RNA conformations. Models starting with "B-" in Table 2 show pure electronic effect of different theoretical hydration models on NMR parameters since geometry of the sPs molecule in all models was identical with that optimized with explicit hydration (Figure 1a) and PCM. The explicit hydration in A-5w and A-6w models was optimized by unconstrained relaxation (Methods).

The isotropic chemical shielding $\sigma_i(^{31}P)$ of hydrated complexes was referenced to the value σ_i^{ref} = 294.3 ppm calculated for the B-6w model including PCM. The isotropic chemical shifts $\delta_i(^{31}P)$ reported in Table 2 were thus calculated as $\delta_i = \sigma_i^{ref} - \sigma_i$.

The $\delta_i(^{31}P)$ calculated with both explicit and implicit PCM hydration ranged from −0.5 to +1.6 ppm; values calculated with explicit hydration only ranged from −1.7 to −1.3 ppm (Table 2). Surprisingly, the gas-phase value of −1.3 ppm was in the explicit-hydration-only interval. The PCM-only value +2.7 ppm was more than 1 ppm off from any other hydration model (Table 2), which points to the necessity of explicit treatment of phosphate–water interactions that are omitted in the PCM-only model.

In contrast to $\delta_i(^{31}P)$, the chemical shielding anisotropy $\Delta\sigma(^{31}P)$ depended systematically and more significantly on the level of hydration. The calculated $\Delta\sigma(^{31}P)$ ranged from

143 to 154 ppm for full explicit hydration with PCM, from 170 to 172 ppm for site-specific explicit hydration with PCM, and from 207 to 215 ppm for site-specific hydration without PCM. The explicit-hydration-only value was 158 ppm, the PCM-only value rose to 189 ppm, and the gas-phase value was 243 ppm. The large span of the calculated $\Delta\sigma(^{31}P)$ values contrasts with a modest range calculated for chemical shifts $\delta_i(^{31}P)$. This could be attributed to the changes in ^{31}P shielding tensor induced by the solvent environment, which were mutually compensated in the case of $\delta_i(^{31}P)$, whereas the same changes led to a large variation of $\Delta\sigma(^{31}P)$. A similar effect occurring upon variations of both phosphate conformation and hydration was observed earlier by Přecechtělová et al.[30] We, therefore, propose that the notable $\Delta\sigma(^{31}P)$ sensitivity is a general property of the phosphate ^{31}P chemical shielding tensor exploitable in experimental ^{31}P NMR.

Calculated scalar spin−spin coupling constants $^2J(P,C3')$ and $^2J(P,C5')$ ranged from −9.6 to −6.1 Hz and from −8.8 to −7.0 Hz, respectively. The gradual extension of the phosphate hydration was accompanied by decrease of both *J*-coupling magnitudes. We noted a selective increase of the difference between $^2J(P,C3')$ and $^2J(P,C5')$ couplings when only one of the two prochiral oxygens (OP1 or OP2) was explicitly hydrated (models B-OP1/2-3w in Table 2). The lack of hydration when one of these oxygens was sterically blocked by the folded part of the NA molecule or by another molecule, such as an interacting protein, could, thus, be in principle detectable with NMR.

Variations of $\delta_i(^{31}P)$ (2.0 ppm), $\Delta\sigma(^{31}P)$ (11 ppm), and $^2J(P,C)$ (2.0 Hz) calculated with the three rather equivalent explicit hydration models (B-6w, A-5w, A-6w, all including PCM, Table 2) are not negligible. We anticipate that only a relevant dynamic model of the phosphate hydration could provide more realistic values of the NMR parameters. Nevertheless, calculations utilizing static models of hydration derived from crystals provided useful trends for NMR parameters in representative NA phosphate geometries.

Explicit hydration of the phosphate group cannot be readily modeled by PCM, as seen for the $\delta_i(^{31}P)$ calculations in B-PCM, B-6w, and B-6w-PCM models (Table 2). Additionally, the effect of PCM hydration surrounding the explicit phosphate first hydration shell was not negligible. The two $^2J(P,C)$ spin−spin couplings were much less sensitive to the particular level of hydration than $\delta_i(^{31}P)$.

NA Phosphate Solvated by Mg^{2+} Ion. Coordination of Mg-5w complex to sPs phosphate induced geometry changes that were specific for the A- and B-forms (Supporting Information). In both forms, the P−O3′ and P−O5′ bonds shrank and the P−OP*n* bond adjacent to the Mg-5w complex lengthened upon the ion coordination.

Coordination of Mg-5w in the B-form was slightly hindered by two water molecules coordinated at the same OP*n* that were absent in the A-form. This small solvation pattern difference resulted in a significantly larger Mg-OP*n* distance and much smaller Mg-OP*n*-P valence angle in the B-form. The observed phosphate geometry variations were reflected in the Mulliken atomic charges (Supporting Information) and mainly in the response of NMR parameters to the Mg-5w coordination that was qualitatively different in the two forms.

In this section, we aimed to obtain NMR parameters close to the basis set limit. On the basis of our atomic basis set convergence check, we used the cc-pCVTZ basis set as an accurate yet affordable alternative.

Table 3. NMR Parameters[a] Calculated with the B3LYP/cc-pCVTZ Method Including PCM in the sPs Complexes with Mg-5w

complex[b]	$\delta_i(^{31}P)$	$\Delta\sigma(^{31}P)$	$^2J(P,C3')$	$^2J(P,C5')$	θ_{11}	θ_{22}	θ_{33}
B-6w	0.0	148	−7.3	−7.4	3.8	3.3	1.8
B-Mg1-9w	1.5	109	−7.0	−9.1	5.7	8.8	8.0
B-Mg2-9w	1.7	107	−9.1	−7.9	3.4	5.8	6.4
B-Mg1-10w	−1.9	113	−7.4	−9.1	6.9	11.4	10.1
B-Mg2-10w	−0.9	110	−10.1	−7.8	3.7	8.0	7.9
A-5w	0.3	159	−6.6	−6.4	2.5	4.9	4.3
A-6w	1.9	148	−5.4	−6.3	2.6	3.3	2.1
A-Mg1-8w	−4.0	126	−5.4	−6.6	1.3	1.4	0.5
A-Mg2-8w	−5.6	121	−6.1	−5.7	3.5	3.5	1.6

[a] The isotropic chemical shift δ_i in ppm (referenced to B-6w isotropic shielding σ_i^{ref} = 285.8 ppm), the chemical shielding anisotropy $\Delta\sigma$ in ppm, the indirect spin−spin coupling $^2J(P,X)$, X = C3′, C5′ in Hz, and the θ_{nn} angles between the σ_{nn} principal shielding tensor components ($\sigma_{11} < \sigma_{22} < \sigma_{33}$) and the respective *z*-, *x*-, *y*-coordinates in the phosphate-frame coordinate system[77] in degrees. [b] The complexes are specified in the Methods.

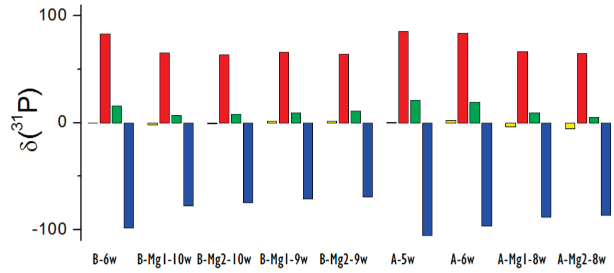


Figure 4. Isotropic values δ_i (yellow) and principal components δ_{11} (red) > δ_{22} (green) > δ_{33} (blue) of the ^{31}P chemical shift tensor calculated with the B3LYP/cc-pCVTZ method including PCM in hydrated and solvated sPs complexes referenced to isotropic chemical shielding σ_i^{ref} = 285.8 ppm in B-6w complex ($\delta_{nn} = \sigma_i^{ref} - \sigma_{nn}$).

The chemical shifts $\delta_i(^{31}P)$ in Table 3 were referenced to the B-6w isotropic shielding σ_i^{ref} = 285.8 ppm in analogy with the previous section. The three hydrated models provided values dispersed similarly (0−1.9 ppm) to the results with IGLO-III basis set in Table 2. The $\delta_i(^{31}P)$ calculated for the ion coordination to OP1 (−1.9, +1.5, −4.0 ppm) and to OP2 (−0.9, +1.7, −5.6 ppm) in the three subclasses of complexes solvated by Mg-5w (-10w, -9w, -8w) indicated that the ion coordination effect on $\delta_i(^{31}P)$ also reflected the phosphate solvation pattern and backbone conformation, as discussed above.

The principal components δ_{nn} (*nn* = 11, 22, 33) of the ^{31}P chemical shift tensor, $\delta_{nn} = \sigma_i^{ref} - \sigma_{nn}$, responded sensitively to ion proximity and local arrangement of solvent molecules (Figure 4, Supporting Information). Upon Mg-5w coordination, all the δ_{nn} changed toward the isotropic value $\delta_i(^{31}P)$ (Figure 4), which was consistent with our previous analysis implying mutual compensation of the molecular environment induced changes in ^{31}P shielding tensor principal components. $\Delta\sigma(^{31}P)$ necessarily decreased with increased isotropy of the chemical shielding tensor upon Mg-5w coordination (Table 3). Orientations of all the δ_{nn} components with respect to the phosphate-frame coordinate system[77] changed upon Mg-5w coordination by 8° at the most (Table 3).

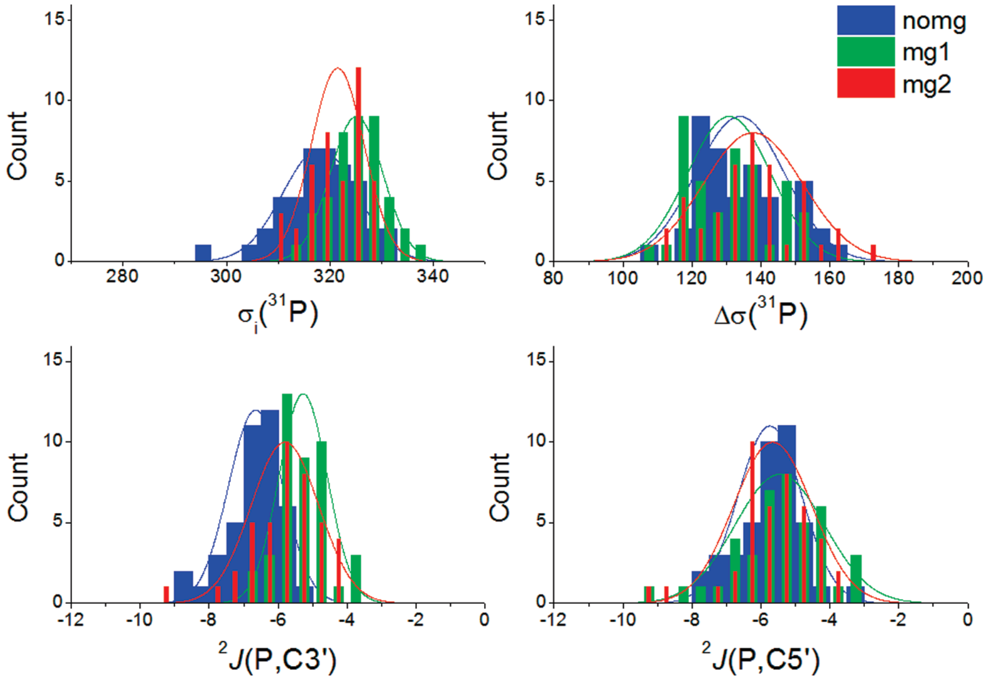


Figure 5. Statistical representation of σ_i and $\Delta\sigma$ (in ppm) and ${}^2J$(P,C) (in Hz) calculated for unrelaxed snapshot geometries taken along nomg, mg1, and mg2 MD runs (Methods). Parameters of the Gaussian functions, added for better clarity, were mean and standard deviation of the respective statistical distribution.

${}^2J$(P,C3′) and ${}^2J$(P,C5′) spin−spin couplings calculated for B-form models depended site-specifically on the ion coordination. The ${}^2J$(P,C5′) coupling magnitude increased by 1.7 Hz in the B-Mg1-10w/-9w complexes relative to the B-6w value, while the simultaneous change of ${}^2J$(P,C3′) coupling was smaller than 0.4 Hz (Table 3). Similarly, ${}^2J$(P,C3′) coupling increased by 2.8 and 1.8 Hz for the ion coordination to OP2 (the B-Mg2-10w/-9w complexes, respectively), while, at the same time, the calculated variation of ${}^2J$(P,C5′) coupling was smaller than 0.6 Hz. The static NMR calculations in the B-form, therefore, suggested that site-specific coordination of the ion in phosphates could in principle be probed by the two ${}^2J$(P,C) couplings. However, the ${}^2J$(P,C) spin−spin couplings calculated in the A-Mg*n*-8w (n = 1, 2) complexes failed to display the site-specificity observed for the B-form.

Variations in the phosphate geometry and, especially, the solvation pattern altered significantly the interpretation of the static NMR calculations. Thus, dynamical modeling of the phosphate solvation was the next logical step.

Effect of Solvent Dynamics on NMR Parameters. We repeatedly observed in NMR calculations utilizing the solvation patterns from crystals that the NMR parameters depended on the local solvent geometry. In the dynamical regime, solvent fluctuations around solute induce variations of NMR parameters. Magnitude of the NMR response to the solvent reorientation increases locally at highly polar or charged groups of solute and with the increase of solvent polarity. The fluctuations of solvent molecules are typically several orders of magnitude faster than the time-scale distinguishable in NMR experiment. Therefore, an experimentally determined NMR parameter is typically a time average over all solvent configurations and fast-enough solute conformation changes. In this sense, static calculations of NMR parameters could have provided misleading trends despite the fact that they were determined for representative molecular structures.

The effect of solvent reorientation on NMR parameters was calculated for the sPs molecule in the A-form and the mg1, mg2, and nomg solvation model (Methods). We checked the geometry of the phosphate in the three MD runs (Supporting Information) and verified that the sPs conformation remained conserved during the simulations. As a result, only the solvent dynamics was reflected in subsequent DFT calculations of NMR parameters. We emphasize that in mg1 and mg2 simulations the ion remained coordinated at the same position throughout the whole MD run (as documented by Mg-OP*n* distance and Mg-OP*n*-P valence angles, n = 1, 2, Supporting Information). The agreement of the position and its stability with the crystallographically determined Mg^{2+} distribution around the phosphate verified the validity of the MD simulations.

The NMR parameters were calculated for unrelaxed (Figure 5) and partially optimized (Figure 6) MD snapshot geometries. The ensembles of NMR data were statistically characterized in Table 4. The arithmetic mean (μ) of the N-point data set (N = 41) provided a reasonable estimate of experimentally probed long-time average. The mean values were determined with the accuracy given by a standard deviation of the mean $s_M = [(\Sigma(x_k - \mu)^2/(N(N - 1)))]^{1/2}$. The long-time average should lie within the $\mu \pm 3s_M$ interval with 99.7% confidence provided that the data distribution is normal. The phosphate solvated by Mg^{2+} could, thus, be distinguished from the hydrated phosphate with one of the NMR parameters when the two corresponding confidence intervals ($\mu \pm 3s_M$) do not overlap. The OP1 and OP2 coordination sites of Mg^{2+} ion could in principle be distinguished by the same statistical criterion.

The NMR data for unrelaxed MD snapshot geometries (Table 4) indicated that the phosphate solvation by Mg^{2+} is

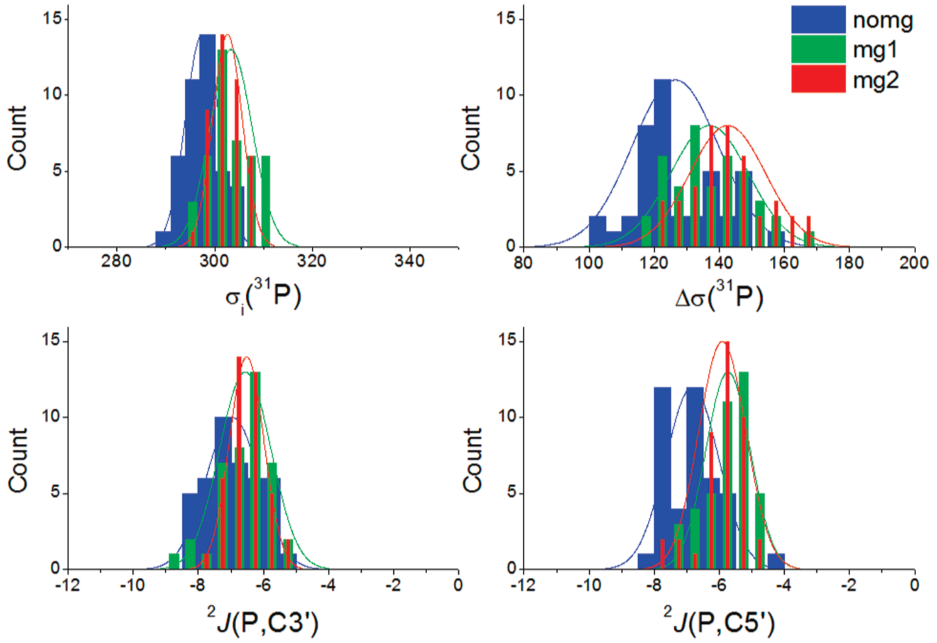


Figure 6. Statistical representation of σ_i and $\Delta\sigma$ (in ppm) and ^2J(P,C) (in Hz) calculated for partially optimized snapshot geometries taken along nomg, mg1, and mg2 MD runs (Methods). Parameters of the Gaussian functions, added for better clarity, were mean and standard deviation of the respective statistical distribution.

Table 4. Statistical Parameters of Calculated NMR Data Distributions

NMR parameter[a]/ MD run	unrelaxed MD geometries		DFT optimized MD geometries	
	μ^b	$3s_M{}^c$	μ^b	$3s_M{}^c$
		$\sigma_i(^{31}P)$		
nomg	318.0	3.4	297.5	1.6
mg1	325.1	2.4	303.3	2.0
mg2	321.5	2.4	302.6	1.4
		$\Delta\sigma(^{31}P)$		
nomg	133.9	6.1	126.5	6.2
mg1	130.8	5.7	137.0	5.6
mg2	137.7	6.7	142.5	5.5
		^2J(P,C3′)		
nomg	−6.66	0.37	−6.95	0.38
mg1	−5.28	0.33	−6.56	0.37
mg2	−5.81	0.46	−6.52	0.25
		^2J(P,C5′)		
nomg	−5.74	0.43	−6.83	0.38
mg1	−5.45	0.61	−5.73	0.32
mg2	−5.66	0.53	−5.91	0.34

[a] The isotropic shielding $\sigma_i(^{31}P)$ in ppm, the chemical shielding anisotropy $\Delta\sigma(^{31}P)$ in ppm, and the indirect spin−spin coupling ^2J(P,X), X = C3′, C5′ in Hz. [b] Arithmetic mean. [c] Three times the standard deviation of the mean $s_M = [(\Sigma(x_k - \mu)^2/(N(N-1)))]^{1/2}$.

observable through statistically significant decrease of ^2J(P,C3′) spin−spin coupling magnitude. Similarly, the coordination of Mg^{2+} to OP1 was manifested by increase of $\sigma_i(^{31}P)$, i.e., decrease of the ^{31}P chemical shift. The results for the chemical shielding anisotropy $\Delta\sigma(^{31}P)$ and ^2J(P,C5′) spin−spin coupling were statistically inconclusive.

Partial DFT optimization of the MD snapshot geometries brought the phosphate geometries in all three models much closer to fully optimized structures in the A-form (Supporting Information). The largest change of NMR parameters due to the phosphate geometry relaxation was a decrease of $\sigma_i(^{31}P)$ by about 20 ppm (Table 4). The mean values of $\sigma_i(^{31}P)$ increased upon the ion coordination by 5−6 ppm (Table 4), which was consistent with static calculations for the A-form structures (Table 3). Importantly, the s_M values decreased considerably after geometry relaxation, which enabled us to conclude that a phosphate coordinated by Mg^{2+} to either OP1 or OP2 should be distinguishable experimentally from the hydrated-only phosphate by a significant decrease of the ^{31}P chemical shift.

The $\Delta\sigma(^{31}P)$ mean values with relaxed phosphate geometry indicated an increase of chemical shielding anisotropy upon Mg^{2+} coordination, which was actually the opposite trend to that calculated for static models (Table 3). However, the statistical distributions of all three solvation patterns (Figure 6) were sufficiently wide to encompass the respective static values thus explaining the apparent discerpancy.

Analysis of the calculated $\sigma(^{31}P)$ tensor unveiled an increase of the largest σ_{33} principal component by about 13−16 ppm upon Mg^{2+} coordination (Supporting Information), which readily explained the Mg^{2+}-induced increase of $\sigma_i(^{31}P)$. Interestingly, response of the two remaining components and the three tensor orientations θ_{nn} to Mg^{2+} coordination was negligible.

After the phosphate geometry relaxation, ^2J(P,C5′) spin−spin coupling dereased upon Mg^{2+} coordination while ^2J(P,C3′) remained almost the same. Both couplings failed to respond site-specifically to the ion coordination.

■ CONCLUSIONS

This study focused on theoretical calculation of ^{31}P shielding tensor and ^{2}J(P,C) coupling constants in dependence on solvation of nucleic acid phosphate by Mg^{2+} ion. The following conclusions can be drawn on the basis of (i) NMR calculations for solvation patterns determined with bioinformatic analysis of B-DNA and A-RNA molecules in crystals and (ii) NMR calculations including dynamically averaged reorientation of solvent with molecular dynamics.

The bioinformatic analysis unveiled two coordination sites of the phosphate with Mg^{2+} lying symmetrically at the distance of 1.9 Å from the two nonesterified phosphate oxygen atoms. However, only one of the two sites can be exclusively populated, which gives rise to site-specific coordination of the Mg^{2+} ion.

NMR calculations for the distinguished solvent patterns found in B-DNA and A-RNA molecules, here called "the static calculations", provided important insight into the NMR parameters dependence on solvent variation because backbone conformations were practically the same, while local coordination of solvent molecules differed. In the static calculations for B-DNA solvation ^{31}P isotropic shifts varied unspecifically, ^{31}P anisotropies decreased by 35−40 ppm and ^{2}J(P,C3′) and ^{2}J(P,C5′) coupling magnitudes increased site-specifically upon Mg^{2+} coordination. For the ion coordination found in A-RNA the calculated ^{31}P isotropic shifts decreased by more than 4 ppm, ^{31}P anisotropies decreased by 20−40 ppm, and ^{2}J(P,C3′) and ^{2}J(P,C5′) couplings varied only slightly and nonspecifically. To conclude, the NMR parameters obtained for the static model depended more on particular coordination of solvent than on whether the phosphate was only hydrated or solvated by Mg^{2+} ion. In other words, the limited number of nucleic acid backbone patterns, although selected reliably, cannot provide accurate trends for the NMR parameters and averaging over solvent reorientation is indispensable.

An important aspect of dynamical averaging concerned dependence of calculated NMR parameters on the thickness of explicit water-solvent layer. It was verified here that the dominant effect on NMR parameters comes from the first hydration shell surrounding the phosphate and Mg^{2+}, while the effect of higher solvent layers can be modeled with PCM.

The dynamically averaged ^{31}P chemical shift decreased by 2−9.5 ppm upon Mg^{2+} coordination, which is in excellent agreement with the 1.5−10 ppm decrease of phosphorothioate ^{31}P chemical shift upon Cd^{2+} coordination probed experimentally in hammerhead ribozyme.[39−41] The decrease of the ^{31}P chemical shift by 0.3 ppm measured[42] in the native phosphate group of minimal metal-binding motif upon addition of Mg^{2+} is also in qualitative agreement with our results although its magnitude is smaller. This could probably be attributed to the fast exchange rate of Mg^{2+} in minimal metal-binding motif. We propose that the inner-shell Mg^{2+} coordination to the nucleic acid phosphate is detectable via decrease of the ^{31}P chemical shift and the magnitude of the decrease is modulated by the metal-ion exchange rate and local mobility of solvent.

To the best of our kowledge the isotropic ^{31}P chemical shift is the only NMR parameter used so far for detection of phosphate metalation. The dynamically averaged chemical shielding anisotropy increased by 0−20 ppm and the ^{2}J(P,C5′) coupling magnitude decreased by 0.2−1.8 Hz upon Mg^{2+} coordination. We propose that the ^{2}J(P,C5′) experiment represents an alternative approach for the detection of metal ion coordination in site-specifically ^{13}C-labeled nucleoside. We emphasize that this experiment would allow NMR detection of the metal in native nucleic acids, i.e., without introducing the thio-phosphate which moreover prefers binding with other metals like Cd^{2+}.

None of the four dynamically averaged NMR parameters distinguished unequivocally the site-specific Mg^{2+} coordination. Considering the two limit cases of this study, the static and the dynamic pictures of solvation, we can conclude that mobility of solvent has a dramatic impact on NMR parameters of nucleic acid phosphate and must be taken into account for their accurate modeling.

■ ASSOCIATED CONTENT

Ⓢ **Supporting Information.** Ab initio optimized geometries of NA complexes with solvent, additional results from static DFT calculations, MD computational details, details on partial optimizations of MD snapshot geometries, additional solvent averaging results, and geometry parameters in MD trajectories. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*V.S.: e-mail, vladimir.sychrovsky@uochb.cas.cz; tel, +420 220 183 234; fax, +420 220 183 578. B.S.: e-mail, bohdan@img.cas.cz; tel, +420 241 063 624.

■ ACKNOWLEDGMENT

This work was supported by the Academy of Sciences of the Czech Republic, Grant No. IAA400550701, the Grant Agency of the Czech Republic P205/10/0228, and the Academy of Sciences of the Czech Republic, Institutional grants AV0Z40550506 and AV0Z50520701. V.S. was supported by the Young Investigator's Grant of the Human Frontier Science Program (HFSP). We acknowledge Frank Jensen for discussion concerning the pcJ-*n* basis sets.

■ REFERENCES

(1) Pyle, A. M. *J. Biol. Inorg. Chem.* **2002**, *7*, 679.
(2) Manning, G. *Q. Rev. Biophys.* **1978**, *11*, 179.
(3) Record, M. T.; Anderson, C. F.; Lohman, T. M. *Q. Rev. Biophys.* **1978**, *11*, 103.
(4) Chin, K.; Sharp, K. A.; Honig, B.; Pyle, A. M. *Nat. Struct. Biol.* **1999**, *6*, 1055.
(5) Misra, V. K.; Draper, D. E. *J. Mol. Biol.* **2000**, *299*, 813.
(6) Makarov, V.; Pettitt, B. M.; Feig, M. *Acc. Chem. Res.* **2002**, *35*, 376.
(7) Misra, V. K.; Draper, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12456.
(8) Tikhomirova, A.; Chalikian, T. V. *J. Mol. Biol.* **2004**, *341*, 551.
(9) Schneider, B.; Berman, H. M. *Biophys. J.* **1995**, *69*, 2661.
(10) Schneider, B.; Patel, K.; Berman, H. M. *Biophys. J.* **1998**, *75*, 2422.
(11) Schneider, B.; Kabeláč, M. *J. Am. Chem. Soc.* **1998**, *120*, 161.
(12) Egli, M.; Portmann, S.; Usman, N. *Biochemistry* **1996**, *35*, 8489.
(13) Auffinger, P.; Hashem, Y. *Bioinformatics* **2007**, *23*, 1035.
(14) Auffinger, P.; Bielecki, L.; Westhof, E. *Structure* **2004**, *12*, 379.
(15) Auffinger, P.; Hashem, Y. *Curr. Opin. Struct. Biol.* **2007**, *17*, 325.
(16) Cate, J. H.; Doudna, J. A. *Structure* **1996**, *4*, 1221.

(17) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Szewczak, A. A.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1696.

(18) Tran-Dinh, S.; Neumann, J. M. *Nucleic Acids Res.* **1977**, *4*, 397.

(19) Smithers, G. W.; O'Sullivan, W. J. *J. Biol. Chem.* **1982**, *257*, 6164.

(20) Johannsen, S.; Korth, M. M. T.; Schnabl, J.; Sigel, R. K. O. *Chimia* **2009**, *63*, 146.

(21) Rinnenthal, J.; Richter, C.; Nozinovic, S.; Furtig, B.; Lopez, J. J.; Glaubitz, C.; Schwalbe, H. *J. Biomol. NMR* **2009**, *45*, 143.

(22) Tanaka, Y.; Kojima, C.; Morita, E. H.; Kasai, Y.; Yamasaki, K.; Ono, A.; Kainosho, M.; Taira, K. *J. Am. Chem. Soc.* **2002**, *124*, 4595.

(23) Sychrovský, V.; Šponer, J.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 663.

(24) Murray, L. J.; Arendall, W. B., 3rd; Richardson, D. C.; Richardson, J. S. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13904.

(25) Schneider, B.; Morávek, Z.; Berman, H. M. *Nucleic Acids Res.* **2004**, *32*, 1666.

(26) Richardson, J. S.; Schneider, B.; Murray, L. W.; Kapral, G. J.; Immormino, R. M.; Headd, J. J.; Richardson, D. C.; Ham, D.; Hershkovits, E.; Williams, L. D.; Keating, K. S.; Pyle, A. M.; Micallef, D.; Westbrook, J.; Berman, H. M. *RNA-Publication RNA Soc.* **2008**, *14*, 465.

(27) Svozil, D.; Kalina, J.; Omelka, M.; Schneider, B. *Nucleic Acids Res.* **2008**, *36*, 3690.

(28) Florián, J.; Štrajbl, M.; Warshel, A. *J. Am. Chem. Soc.* **1998**, *120*, 7959.

(29) Leulliot, N.; Ghomi, M.; Scalmani, G.; Berthier, G. *J. Phys. Chem. A* **1999**, *103*, 8716.

(30) Přecechtělová, J.; Munzarová, M. L.; Novák, P.; Sklenář, V. *J. Phys. Chem. B* **2007**, *111*, 2658.

(31) Přecechtělová, J.; Padrta, P.; Munzarová, M. L.; Sklenář, V. *J. Phys. Chem. B* **2008**, *112*, 3470.

(32) Vokáčová, Z.; Buděšínský, M.; Rosenberg, I.; Schneider, B.; Šponer, J.; Sychrovský, V. *J. Phys. Chem. B* **2009**, *113*, 1182.

(33) Přecechtělová, J.; Novák, P.; Munzarová, M. L.; Kaupp, M.; Sklenář, V. *J. Am. Chem. Soc.* **2010**, *132*, 17139.

(34) Ribasprado, F.; Giessnerprettre, C.; Pullman, B.; Daudey, J. P. *J. Am. Chem. Soc.* **1979**, *101*, 1737.

(35) Gorenstein, D. G. *Phosphorus-31 NMR*; Academic Press, Inc.: Orlando, FL, 1984.

(36) Gorenstein, D. G. *Methods Enzymol.* **1992**, *211*, 254.

(37) Takagi, Y.; Warashina, M.; Stec, W. J.; Yoshinari, K.; Taira, K. *Nucleic Acids Res.* **2001**, *29*, 1815.

(38) Peracchi, A.; Beigelman, L.; Scott, E. C.; Uhlenbeck, O. C.; Herschlag, D. *J. Biol. Chem.* **1997**, *272*, 26822.

(39) Suzumura, K.; Yoshinari, K.; Tanaka, Y.; Takagi, Y.; Kasai, Y.; Warashina, M.; Kuwabara, T.; Orita, M.; Taira, K. *J. Am. Chem. Soc.* **2002**, *124*, 8230.

(40) Suzumura, K.; Takagi, Y.; Orita, M.; Taira, K. *J. Am. Chem. Soc.* **2004**, *126*, 15504.

(41) Osborne, E. M.; Ward, W. L.; Ruehle, M. Z.; DeRose, V. J. *Biochemistry* **2009**, *48*, 10654.

(42) Tanaka, Y.; Morita, E. H.; Hayashi, H.; Kasai, Y.; Tanaka, T.; Taira, K. *J. Am. Chem. Soc.* **2000**, *122*, 11303.

(43) Schneider, B.; Cohen, D. M.; Schleifer, L.; Srinivasan, A. R.; Olson, W. K.; Berman, H. M. *Biophys. J.* **1993**, *65*, 2291.

(44) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(46) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

(47) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(48) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; Defrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982** *77*, 3654.

(49) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. *J. Comput. Chem.* **1983**, *4*, 294.

(50) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251.

(51) Sychrovský, V.; Grafenstein, J.; Cremer, D. J. *Chem. Phys.* **2000**, *113*, 3530.

(52) Helgaker, T.; Watson, M.; Handy, N. C. J. *Chem. Phys.* **2000**, *113*, 9402.

(53) Peralta, J. E.; Scuseria, G. E.; Cheeseman, J. R.; Frisch, M. J. *Chem. Phys. Lett.* **2003**, *375*, 452.

(54) Kutzelnigg, W.; Fleischer, U.; Schindler, M. *NMR, Basic Principles and Progress*; Springer: Heidelberg, 1990; Vol. 23.

(55) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(56) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358.

(57) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1995**, *103*, 4572.

(58) Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **2002**, *117*, 10548.

(59) Schafer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.

(60) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43.

(61) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(62) Jensen, F. *J. Chem. Phys.* **2002**, *117*, 9234.

(63) Jensen, F. *J. Chem. Theory Comput.* **2006**, *2*, 1360.

(64) Jensen, F. *J. Phys. Chem. A* **2007**, *111*, 11198.

(65) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(66) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

(67) Feller, D. *J. Comput. Chem.* **1996**, *17*, 1571.

(68) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L. S.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. *J. Chem. Inf. Model.* **2007**, *47*, 1045.

(69) Chesnut, D. B.; Moore, K. D. *J. Comput. Chem.* **1989**, *10*, 648.

(70) Dračínský, M.; Kaminský, J.; Bouř, P. *J. Phys. Chem. B* **2009**, *113*, 14698.

(71) Dračínský, M.; Bouř, P. *J. Chem. Theory Comput.* **2010**, *6*, 288.

(72) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.

(73) Pérez, S. *Methods Enzymol.* **1991**, *202*, 510.

(74) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, 2008.

(75) Sychrovský, V.; Vokáčová, Z.; Šponer, J.; Špacková, N.; Schneider, B. *J. Phys. Chem. B* **2006**, *110*, 22894.

(76) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.;

Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(77) Herzfeld, J.; Griffin, R. G.; Haberkorn, R. A. *Biochemistry* **1978**, *17*, 2711.

THE JOURNAL OF
PHYSICAL CHEMISTRY B

Article

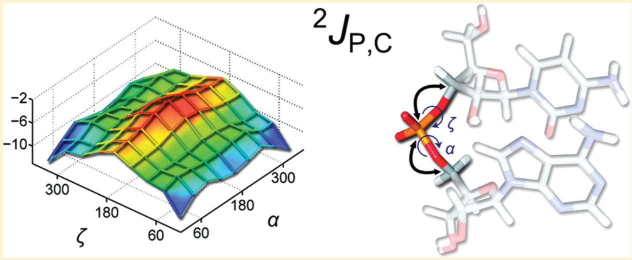
pubs.acs.org/JPCB

Correlating the ^{31}P NMR Chemical Shielding Tensor and the $^{2}J_{P,C}$ Spin−Spin Coupling Constants with Torsion Angles ζ and α in the Backbone of Nucleic Acids

Ladislav Benda,* Zuzana Sochorová Vokáčová, Michal Straka, and Vladimír Sychrovský*

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v.v.i., Flemingovo náměstí 2, 166 10 Prague 6, Czech Republic

Ⓢ *Supporting Information*



ABSTRACT: Determination of nucleic acid (NA) structure with NMR spectroscopy is limited by the lack of restraints on conformation of NA phosphate. In this work, the ^{31}P chemical shielding tensor, the $\Gamma_{P,C5'H5'1}$ and $\Gamma_{P,C5'H5'2}$ cross-correlated relaxation rates, and the $^{2}J_{P,C3'}$, $^{2}J_{P,C5'}$, and $^{3}J_{P,C4'}$ coupling constants were calculated in dependence on NA backbone torsion angles ζ and α. While the orientation of the ^{31}P chemical shielding tensor was almost independent of the NA phosphate conformation, the principal tensor components varied by up to ∼40 ppm. This variation and the dependence of the phosphate geometry on torsion angles ζ and α had only a minor influence on the calculated $\Gamma_{P,C5'H5'1}$ and $\Gamma_{P,C5'H5'2}$ cross-correlated relaxation rates, and therefore, the so-called rigid tensor approximation was here validated. For the first time, the $^{2}J_{P,C}$ spin−spin coupling constants were correlated with the conformation of NA phosphate. Although each of the two *J*-couplings was significantly modulated by both torsions ζ and α, the $^{2}J_{P,C3'}$ coupling could be structurally assigned to torsion ζ and the $^{2}J_{P,C5'}$ coupling to torsion α. We propose qualitative rules for their structural interpretation as loose restraints on torsion angles ζ and α. The $^{3}J_{P,C4'}$ coupling assigned to torsion angle β was found dependent also on torsions ζ and α, implying that the uncertainty in determination of β with standard Karplus curves could be as large as ∼25°. The calculations provided a unified picture of NMR parameters applicable for the determination of NA phosphate conformation.

■ INTRODUCTION

The biological function of nucleic acids (NAs) is closely related to their three-dimensional architecture, which is determined by the geometry of the NA backbone. The NA backbone consists of chemically identical segments appearing repetitively along the NA strand (Figure 1a). Although full rotational freedom of NA backbone torsion angles ε, ζ, α, β, γ, and δ would correspond to a rather extreme number of hypothetical conformers, the backbone geometries were shown to cluster only into a relatively small number of distinguished structural patterns.[1] These so-called NA conformational classes determined both for DNA and RNA molecules in crystals[2] are nowadays used for classification of NA backbone geometries. While the DNA molecules predominantly occur in A-, B_I-, B_{II}-, and Z-form, the conformational richness of RNA molecules is described with 46 conformational classes.[1]

Nuclear magnetic resonance (NMR) spectroscopy has become an indispensable technique in structural studies of NAs.[3−6] It provides a unique alternative to X-ray crystallography, enabling to probe conformation and dynamics of NAs in their native environment. The NMR structural studies of the NA backbone utilize mainly NMR relaxation enhancements due to the nuclear Overhauser effect (NOE), residual dipolar couplings (RDCs) in partially oriented media, three-bond spin−spin coupling constants (^{3}J-couplings), and newly also

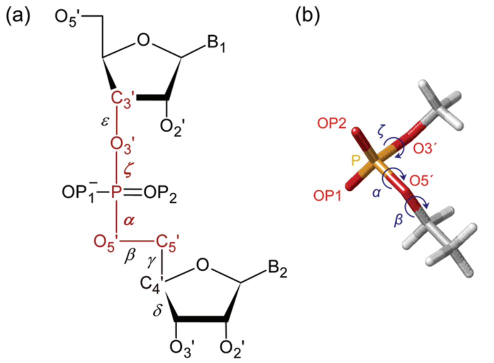


Figure 1. (a) Definition of NA backbone torsion angles. The symbols B_1 and B_2 stand for NA bases, which were in the nPn model replaced by methyl groups. (b) Ethyl methyl phosphate (EMP) with atoms and torsion angles defined in analogy to NA backbone. The nPn and EMP models of NA phosphate were used in this study for calculating the dependence of NMR parameters on torsion angles ζ and α.

Received: October 14, 2011
Revised: February 10, 2012
Published: March 1, 2012

ACS Publications © 2012 American Chemical Society

dx.doi.org/10.1021/jp2099043 | *J. Phys. Chem. B* 2012, 116, 3823−3833

cross-correlated relaxation (CCR) rates. The NOEs provide interatomic distances, the RDCs contain information on spatial orientation of dipolar vectors, and the ^3J-couplings[3,4,7] and the $\Gamma_{P,CH}$ CCR rates[8,9] are interpreted as structural restraints on the NA backbone torsion angles.

The NMR measurements in the NA backbone are conditioned by the abundance of spin-$^1/_2$ nuclei ^{1}H, ^{13}C, and ^{31}P. The C3′−O3′−P−O5′−C5′ segment including the phosphate group is in this regard particularly difficult due to the multiple occurrences of backbone oxygen atoms (Figure 1). The torsion angles ζ and α thus cannot be determined on the basis of ^3J-couplings, and alternative NMR restraints on NA phosphate conformation are needed.[3] Virtually, the only available NMR parameters directly responding to the conformation of torsion angles ζ and α are the isotropic ^{31}P chemical shift δ_P, the $\Gamma_{P,CH}$ CCR rates, which are closely related to the ^{31}P chemical shielding anisotropy (CSA), and the $^2J_{P,C}$ spin−spin coupling constants. Their complex structural dependencies, however, require accurate calibration, which was the main aim of this work. It should be noted that an empirical correlation between torsions ε and ζ in NA structures is known and a combined measurement of δ_P and the $^2J_{P,H3'}$ coupling is used to restrict the value of torsion ζ.[10]

There are further obstacles concerning the ^{31}P NMR studies of the NA backbone. The highly uniform chemical structure of the NA backbone is manifested in a relatively narrow range of δ_P.[11] This fact limits the assignment and interpretation of δ_P in large NA molecules exceeding ca. 40 nucleotides. Specific interactions of metal ions with the NA phosphate may also affect the structural interpretation of ^{31}P NMR parameters.[12] The NMR determination of torsion angles ζ and α is therefore still challenging, and restraints on the (ζ, α) conformation even of a qualitative nature would help to improve the precision of NA structures.

The torsions ζ and α typically adopt one of the orientations commonly called *gauche* (*g*+, 60° or *g*−, 300°) and *trans* (*t*, 180°). Historically, δ_P was used as a very loose restraint on the NA backbone structure, allowing one to exclude the *trans* conformation of either torsion ζ or α if the δ_P value was within the −5 to −4 ppm range.[3] Structural interpretation of the δ_P exceeding this interval was considered problematic. Early *ab initio* calculations for the dimethyl phosphate (DMP) suggested that δ_P decreases when the (ζ, α) conformation of NA phosphate changes from (*g*−, *g*−) to (*t*, *g*−).[11,13] The decrease of δ_P by 1.6 ppm was actually measured and assigned to DNA conformational transition $B_I \rightarrow B_{II}$ corresponding to the (*t*, *g*−, *g*−) → (*g*−, *t*, *g*−) change of (ε, ζ, α) torsion angles.[10] The same δ_P difference was recently modeled by combining molecular dynamics simulations of the Dickerson−Drew dodecamer with DFT NMR calculations.[14] Přecechtělová et al. also calculated the structural dependence of the ^{31}P chemical shift in the DMP model with an explicitly hydrated phosphate group, where the (ζ, α) torsions were confined around the (*g*−, *g*−) and (*t*, *g*−) conformations.[15] Despite the effort, general rules for structural interpretation of δ_P in NA phosphate are still missing. The reason is that, besides the dependence on torsion angles ζ and α, δ_P is sensitive also to the O−P−O bond angles,[10,11,16] torsion angles ε and β,[17,18] and phosphate group solvation.[12,19]

The anisotropic components of the ^{31}P chemical shielding tensor can be exploited through their contribution to NMR relaxations. It was shown that the torsion angles ζ and α can be determined on the basis of CCR rates $\Gamma_{P,CH}$ in combination with the ^3J-couplings assigned to torsions ε and β.[8] This approach was modified in a recent work by Schwalbe's group,[9] where the torsion angles ζ and α were unambiguously determined by combined measurement of $\Gamma_{P,CH}$ and NOE. Structural interpretation of the $\Gamma_{P,CH}$ CCR rates usually relies on the so-called "rigid tensor approximation", which assumes that the principal components of the ^{31}P chemical shielding tensor including their orientation in the molecular frame remain the same for different phosphate geometries. The approximation usually employs the ^{31}P chemical shielding tensor measured in barium diethyl phosphate (BDEP) in the solid state.[20] Validity of the rigid tensor approximation in the interpretation of $\Gamma_{P,CH}$ CCR rates can be tested, e.g., by theoretical calculations of their dependencies on NA phosphate conformation.

The only remaining NMR parameters that could be used for the determination of torsion angles ζ and α are the $^2J_{P,C3'}$ and $^2J_{P,C5'}$ spin−spin coupling constants involving the carbon atoms in the closest vicinity of NA phosphate (Figure 1). While a significant modulation of $^2J_{P,C3'}$ and $^2J_{P,C5'}$ by both torsion angles ζ and α can be expected, their dependence on NA phosphate geometry has not been systematically studied yet. Our previous calculations indicated that the *gauche* and *trans* orientations of torsions ζ and α could be distinguished with ∼2−3 Hz differences in magnitudes of the $^2J_{P,C3'}$ and $^2J_{P,C5'}$ couplings, respectively.[7] Similar variation of the two $^2J_{P,C}$ couplings was actually measured in NAs, but no structural interpretation was available.[21−25]

In this theoretical work, we therefore particularly addressed the dependencies of $^2J_{P,C}$ coupling constants on torsion angles ζ and α in order to assess their applicability for determining the NA phosphate conformation. We further studied the conformational behavior of the ^{31}P chemical shielding tensor and addressed the validity of the rigid tensor approximation commonly accepted in the interpretation of CCR rates. By this, we wish to fill the existing gap in utilization of the ^{31}P NMR parameters in structural studies of the NA backbone.

■ THEORY

Chemical Shielding and Chemical Shift. In this work, we consistently distinguish ^{31}P chemical shielding σ_P defined as a *decrease* of magnetic field at the site of the ^{31}P nucleus due to the surrounding electronic environment, $B = B_0(1 - \sigma_P)$, from the ^{31}P chemical shift δ_P, which represents the relative NMR scale defined with respect to the chemical shielding of a NMR standard σ_{std}

$$\delta_P = \frac{\sigma_{std} - \sigma_P}{1 - \sigma_{std}} \approx \sigma_{std} - \sigma_P \tag{1}$$

Accordingly, the theoretical chemical shift is obtained as a difference between chemical shielding calculated in NMR standard and in the molecule of interest. In this work, σ_P and σ_{std} were always calculated at the same level.

The usual ^{31}P NMR standard is 85% aqueous solution of phosphoric acid (H_3PO_4 (aq)). Calculation of the ^{31}P chemical shielding in H_3PO_4 (aq) is a challenging task, which can be circumvented, e.g., by introducing secondary standard PH_3, as was suggested by van Wüllen.[26] The isotropic ^{31}P chemical shift in the NA phosphate is then evaluated as

$$\delta_P = -266.1 + \sigma_{PH3} - \sigma_P \tag{2}$$

where −266.1 ppm is the experimental chemical shift of the gas-phase PH_3 relative to the 85% H_3PO_4 (aq)[27] and σ_{PH_3} and σ_P are the isotropic ^{31}P chemical shieldings calculated in the secondary standard PH_3 and in NA phosphate, respectively. If not specified otherwise, the value σ_{PH_3} = 566.0 ppm calculated with the B3LYP/IGLO-III method was used. We note that, while σ_{PH_3} was calculated in the gas phase to comply with the experimental conditions,[27] the calculation of σ_P required inclusion of water solvent effects.

Chemical Shielding Anisotropy. The ^{31}P CSA was defined as

$$\Delta\sigma_P = \sigma_{P,33} - \frac{\sigma_{P,11} + \sigma_{P,22}}{2} \tag{3}$$

where $\sigma_{P,11} < \sigma_{P,22} < \sigma_{P,33}$ are the principal components of the ^{31}P chemical shielding tensor.

CSA−DD Cross-Correlated Relaxation. The CSA−DD CCR is a NMR phenomenon resulting from the interference of the CSA and dipole−dipole (DD) relaxation mechanisms.[28,29] The CSA−DD CCR rates depend on mutual orientation of the CSA and DD interaction tensors and thus also on local molecular geometry. The so-called "remote" CSA−DD CCRs, which correlate CSA and DD interactions centered on different nuclei, were previously applied for determination of torsion angles in peptides[30,31] and NAs.[8,9,32,33] Theoretical studies provided nonempirical insights into the geometry dependence of the CSA−DD CCR rates, including the effect of geometry dependence of the chemical shielding tensor.[34,35]

In this work, we focused on the remote CSA−DD cross-correlation between the ^{31}P chemical shielding tensor and the adjacent C5′−H5′ dipolar vectors. According to the Redfield theory of relaxations, assuming invariable relative orientation of the CSA and DD principal axis frames, the remote transversal CSA−DD CCR rate $\Gamma_{P,CH}$ is given by[28,29]

$$\Gamma_{P,CH} = \frac{1}{2}\frac{\mu_0}{4\pi}\hbar\gamma_P B_0 \frac{\gamma_C\gamma_H}{r_{CH}^3}\Delta\sigma_{P,CH}^{eff}\cdot\frac{4}{3}J(0) \tag{4}$$

where μ_0 is the vacuum permeability, $\hbar$ is the reduced Planck constant, γ_P, γ_C, and γ_H are the gyromagnetic ratios of the ^{31}P, ^{13}C, and ^1H nuclei, B_0 is the magnetic field strength, r_{CH} is the length of the C−H dipolar vector $\mathbf{r}_{CH}$, $J(0)$ is the spectral density function at zero frequency, and

$$\Delta\sigma_{P,CH}^{eff} = \sum_{i=1}^{3}\sigma_{P,ii}\frac{3\cos^2\vartheta_{CH,ii} - 1}{2} \tag{5}$$

is the effective CSA, which can be conveniently evaluated by quantum-chemistry methods. Here, $\sigma_{P,ii}$ is the ith principal component of the ^{31}P chemical shielding tensor and $\vartheta_{CH,ii}$ is the angle between $\sigma_{P,ii}$ and $\mathbf{r}_{CH}$. We further assume overall isotropic molecular tumbling, for which $J(0) = 2/5\tau_c$, where τ_c is the rotation correlation time. The value τ_c = 2.27 ns determined experimentally for a RNA 14-mer[9] and the magnetic field B_0 = 14.09 T corresponding to the ^1H resonance frequency 600 MHz were used in our calculations.

It should be noted that the isotropic part of the ^{31}P chemical shielding tensor does not contribute to the effective CSA value, because

$$\sum_{i=1}^{3}\frac{3\cos^2\vartheta_{CH,ii} - 1}{2} \equiv 0 \tag{6}$$

for any orientation of the chemical shielding tensor. Therefore, only the relative values of principal components $\sigma_{P,ii}^r = \sigma_{P,ii} - \sigma_P$ are essential in the effective CSA, and they can be used instead of $\sigma_{P,ii}$ for calculating the $\Delta\sigma_{P,CH}^{eff}$ according to eq 5.

■ METHODS

Computational Details. All DFT calculations were performed using the B3LYP functional[36,37] as implemented in the Gaussian 09 package.[38] Water solvent in the DFT calculations was simulated with the polarizable continuum model (PCM).[39] The coupled-cluster singles and doubles[40] (CCSD) calculations of geometry and NMR parameters of DMP and PH_3 molecules were performed with the CFOUR program.[41] The DFT and CCSD geometry optimizations employed the 6-31+G(d) and 6-311++G(d,p) basis sets as obtained from the EMSL basis set exchange web portal, respectively.[42]

The NMR parameters were calculated at the coupled-perturbed B3LYP and CCSD level. The GIAO method[43−45] was applied in all ^{31}P chemical shielding calculations. The principal components $\sigma_{P,ii}$ and the corresponding eigenvectors were obtained by diagonalizing the symmetric part of the calculated ^{31}P chemical shielding tensor. The Fermi contact (FC), spin−dipolar (SD), paramagnetic spin−orbit (PSO), and diamagnetic spin−orbit (DSO) terms of the total J-coupling were evaluated at the B3LYP level.[46,47] The FC term of the $^2J_{P,C}$ coupling was calculated also with the CCSD method. The triplet instability problem[48] was avoided by applying the orbital-unrelaxed approach in the coupled-cluster calculations (keyword DIFF_TYPE = UNRELAXED). All NMR calculations utilized the IGLO-III basis set; this choice was inspired by our previous work, where the IGLO-III basis provided both ^{31}P chemical shielding and $^2J_{P,C}$ coupling close to the B3LYP basis-set limit.[12]

Conformational dependencies of the $\Gamma_{P,CH}$ CCR rates in the EMP molecule were evaluated with two different approaches: (a) both the ^{31}P chemical shielding tensor variation and the molecular geometry relaxation were included in the $\Gamma_{P,CH}$ calculations ("relaxed approach"); (b) the rigid tensor approximation was adopted, using the ^{31}P chemical shielding tensor calculated in the EMP global minimum (g−, g−) and transferring it on rigidly rotated EMP conformers ("rigid approach").

Molecular Models. We used three different molecular models of NA phosphate. The DMP molecule was employed to compare the performance of the B3LYP and CCSD computational methods. Dependence of NMR parameters on the NA phosphate (ζ, α) conformation was calculated for the EMP molecule (Figure 1b) and the abasic RNA dinucleotide (nPn, Figure 1a), where the NA bases were replaced by methyl groups.

The EMP molecule is the smallest NA phosphate model respecting the asymmetry of the NA backbone. Its conformation is described with torsion angles ζ, α, and β (Figure 1). The global energy minimum with conformation (ζ, α, β) = (g−, g−, t) was optimized at the B3LYP level. The conformational space of EMP was explored by varying the torsion angles ζ and α with 20° steps, resulting in 18 × 18 = 324 geometries, while the torsion β was fixed to 180°. The constraint on torsion β reduced the complexity of the problem, and it is consistent with the most usual value of torsion β in NAs.[1,2] Furthermore, we adopted the (ζ, α, β) values corresponding to the 46 RNA conformational classes found

Table 1. Selected Geometry Parameters[a] of NA Phosphate Models

model	geometry method	ζ	α	β	O3′−P−O5′	OP1−P−OP2	P−OP[b]	P−O3′	P−O5′	O3′−C3′	O5′−C5′
DMP	B3LYP	286	286		99.5	125.5	1.504	1.683	1.683	1.418	1.418
DMP	B3LYP/PCM	293	293		102.3	121.7	1.512	1.655	1.655	1.435	1.435
DMP	CCSD	288	288		99.2	125.7	1.488	1.663	1.663	1.412	1.412
EMP	B3LYP/PCM	293	292	180	102.4	121.6	1.512	1.655	1.653	1.434	1.442
nPn1[c]	Amber99	56	164	152	102.0	120.7	1.471	1.599	1.611	1.416	1.420
nPn2[c]	"	286	162	163	101.7	121.1	1.473	1.594	1.603	1.414	1.420
nPn3[c]	"	205	296	153	103.5	121.3	1.474	1.594	1.599	1.417	1.412
nPn6[c]	"	163	293	167	99.9	121.4	1.473	1.584	1.596	1.411	1.419
nPn9[c]	"	83	74	172	101.1	120.5	1.473	1.590	1.602	1.414	1.416
nPn12[c]	"	165	289	166	102.0	121.4	1.472	1.590	1.600	1.414	1.416
nPn15[c]	"	210	69	195	102.1	121.3	1.473	1.592	1.596	1.413	1.414
nPn18[c]	"	260	63	179	102.2	120.2	1.474	1.606	1.608	1.417	1.415
nPn20[c]	"	285	294	179	103.0	120.9	1.473	1.599	1.603	1.416	1.415
DMP[d]	MP2	70	70		99	126	1.51	1.68	1.68	1.43	1.43
BDEP[e]	Exp.	72	68	184[f]	103.5	121.7	1.513	1.620	1.593	1.474	1.442

[a] Torsion and bond angles in deg, bond lengths in Å. [b] Average P−OP*n* bond length, $n = 1, 2$, calculated as P−OP = (P−OP1 + P−OP2)/2. [c] Abasic RNA dinucleotide.[7] [d] Obtained previously by Florián et al. with the 6-31G(d,p) basis set.[49] [e] X-ray geometry of the barium diethyl phosphate (CSD ID BADETP).[50] Differences in O−P−O−C torsions (ζ and α) and P−O and C−O bond lengths in the structurally symmetric BDEP molecule were caused by crystal packing. [f] Average value of the P−O−C−C torsion in the BDEP crystal.

Table 2. Isotropic ^{31}P Chemical Shifts[a] in ppm and $^2J_{P,C}$ Coupling Constants in Hz Calculated with Different Methods and the IGLO-III Basis Set for the ($g-$, $g-$) and (t, t) Conformers of DMP

geometry method	NMR method	δ_P[a] ($g-$, $g-$)/(t, t)	$^2J^{FC}_{P,C}$ ($g-$, $g-$)/(t, t)	$\Delta_{tt-gg}(\delta_P)$	$\Delta_{tt-gg}(^2J^{FC}_{P,C})$
B3LYP	B3LYP	5.2/13.9[b]	−7.6/−2.1	8.6	5.5
B3LYP/PCM	B3LYP/PCM	8.9/15.3[b]	−7.5/−2.9	6.5	4.6
CCSD	B3LYP	−0.7/8.2[c]	−6.8/−1.8	8.9	5.1
CCSD	CCSD	−4.2/4.3[d]	−6.4/−2.2	8.6	4.2

[a] Relative to 85% H_3PO_4 (aq), using the PH_3 secondary standard as described in the Theory section. The calculated values of the ^{31}P shielding in PH_3 were: [b] σ_{PH_3} = 566.0 ppm, [c] σ_{PH_3} = 568.0 ppm, [d] σ_{PH_3} = 608.4 ppm.

by Richardson et al.[1] For each (ζ, α, β) geometry of EMP, all remaining coordinates were fully relaxed by B3LYP energy minimization. By optimizing the EMP geometries in this way, we obtained adiabatically relaxed phosphate group geometries, namely, the P−O bond distances and O−P−O valence angles, which most significantly influence the ^{31}P NMR parameters in NA phosphate.[11]

Nine different conformations of nPn (Table 1) were adopted from our previous work.[7] These structures—further denoted nPn1, nPn2, nPn3, nPn6, nPn9, nPn12, nPn15, nPn18, and nPn20—represent RNA conformational classes 1, 2, 3, 6, 9, 12, 15, 18, and 20 (canonical A-form) as defined by Schneider et al.[2] The geometries were obtained previously with the Amber99 force field and were used in this work as starting structures for rigid variation of either ζ or α torsion angle, thus sampling the important regions of the (ζ, α) conformational space.

The effects of water solvent on both molecular geometries and NMR parameters were simulated solely with the PCM. Explicit hydration was not applied to avoid steric clashes when sampling the complete (ζ, α) conformational space. The PCM is known to provide only a rough approximation of water solvent effects in NA phosphate.[12] However, absolute accuracy of NMR calculations was not crucial in our case, since we were mainly interested in differences of NMR parameters between different phosphate conformations ("conformational differences"). In this sense, PCM performed rather well and allowed for consistent treatment of water solvent effects in all NA phosphate models used.

For brevity, we used a notation $\Delta_{\zeta'\alpha'-\zeta\alpha}(X)$ standing for the difference of NMR parameter X between (ζ', α') and (ζ, α) conformers of NA phosphate. For example, the difference of δ_P between the B_{II} and B_I conformers of DNA ("B_{II} − B_I" difference) is denoted $\Delta_{tg-gg}(\delta_P)$.

RESULTS AND DISCUSSION

Geometry Parameters of NA Phosphate Models. The geometries of phosphate group in DMP molecule optimized with the B3LYP and CCSD methods were rather similar, differing only in the P−O bond lengths (Table 1). The gas-phase optimized structures agreed with the DMP geometry obtained previously by Florián et al. at the MP2 level.[49] The inclusion of water solvent modeled with PCM had a sizable effect on the DMP geometry; the influence of solvent on the NMR (and other) properties of NA phosphate thus originates also from geometry changes.

Molecular geometries optimized with the B3LYP/PCM and Amber99 methods differed (Table 1), which had a sizable effect on the calculated NMR parameters (see below). The geometry of the EMP global minimum ($g-$, $g-$) could be directly compared with the nPn20 structure corresponding to the A-form of NAs. Torsion angles α and β obtained in EMP at the B3LYP/PCM level were in very good agreement with the nPn20 force-field values and also with the averaged crystallographic data for the A-form α = 295°, β = 174°.[1,2] The slight discrepancy in torsion ζ could be attributed to the fact that the C3′ methyl group in EMP is a poor substitute for a ribose ring. This limitation of the EMP model was also reflected in the calculated ${}^2J_{\mathrm{P,C3'}}$ couplings (see below). Bond angles optimized with B3LYP/PCM and Amber99 methods also nicely agreed, but the bond lengths differed, especially for P−O*Pn*, n = 1, 2 (by ∼0.04 Å) and P−Om', m = 3, 5 (by ∼0.05 Å). This fact has serious consequences for force-field modeling of NA phosphate properties, some of which have been recently addressed.[12]

Benchmark NMR Calculations. We tested the performance of the B3LYP and B3LYP/PCM methods in calculations of the ^{31}P chemical shift δ_P and the FC term of the ${}^2J_{\mathrm{P,C}}$ spin−spin coupling against the benchmark CCSD method. The global minimum ($g-$, $g-$) and transition state (t, t) of the DMP molecule were used for this purpose.

The δ_P was quite sensitive to various conditions such as the molecular geometry, the PCM solvation, and the electronic structure method used in the NMR calculations (Table 2). Namely, the calculated δ_P values decreased by 5.7−6.0 ppm when the molecular geometries were optimized with the CCSD method instead of B3LYP and by an additional 3.5−3.9 ppm when both structure and NMR shielding were obtained at the CCSD level. The inclusion of PCM solvation, that is currently available only for DFT, led to an increase of δ_P by 1.4−3.7 ppm.

The $\Delta_{tt-gg}(\delta_P)$ conformational differences of 8.6−8.9 ppm were to a large extent independent of the electronic structure method, indicating that the B3LYP method is well-suited for theoretical modeling of conformational variation of δ_P in NA phosphate. The decrease of $\Delta_{tt-gg}(\delta_P)$ by ∼2 ppm induced by the PCM hydration implies that the effect of water solvent on δ_P is not uniform but varies with the DMP conformation. Hence, both the DMP polarization and accessibility of the phosphate group by water solvent must be considered for accurate δ_P calculations.

The FC term of the ${}^2J_{\mathrm{P,C}}$ coupling is dominant.[7] The sum of the remaining SD, PSO, and DSO terms ranged from −0.4 to −0.6 Hz for different computational methods and DMP conformers. The J-coupling difference $\Delta_{tt-gg}({}^2J^{\mathrm{FC}}_{\mathrm{P,C}})$ was overestimated by the B3LYP method relative to CCSD by 1.3 Hz, mainly due to the different performance of B3LYP and CCSD for the ($g-$, $g-$) conformer (Table 2). The accuracy of the ${}^2J_{\mathrm{P,C}}$ coupling calculated at the B3LYP level therefore varied with the NA phosphate conformation. This fact was taken into account when estimating the accuracy of the calculated structural dependencies of ${}^2J_{\mathrm{P,C}}$ couplings. Other DFT functionals performed similarly as the B3LYP; the ${}^2J^{\mathrm{FC}}_{\mathrm{P,C}}$ values for the ($g-$, $g-$) conformer of DMP calculated with BLYP, BPW91, and B3PW91 functionals were 7.3, 7.2, and 7.4 Hz.

We also compared the performance of the B3LYP, B3LYP/PCM, and CCSD methods in calculations of the full ^{31}P chemical shielding tensor (Supporting Information, Table S1). The principal components $\sigma_{\mathrm{P},ii}$, ii = 11, 22, 33, calculated with the CCSD method differed by 6−14 ppm from those obtained with B3LYP. The inclusion of PCM hydration in the B3LYP calculations had a much larger impact; the $\sigma_{\mathrm{P},ii}$ values changed by 22−50 ppm. The ^{31}P-CSA calculated at the CCSD level was by 21 ppm smaller than that obtained using B3LYP. The inclusion of PCM hydration decreased the ^{31}P-CSA by more then 60 ppm. The principal components $\sigma_{\mathrm{P},ii}$ were thus much more sensitive to the NMR computational method and water solvent description than the isotropic chemical shift δ_P itself. Interestingly, the orientation of the ^{31}P chemical shielding tensor in the phosphate−frame coordinate system[20] represented by angles φ_{ii}, ii = 11, 22, 33 (Figure 2), was practically independent of the computational method.

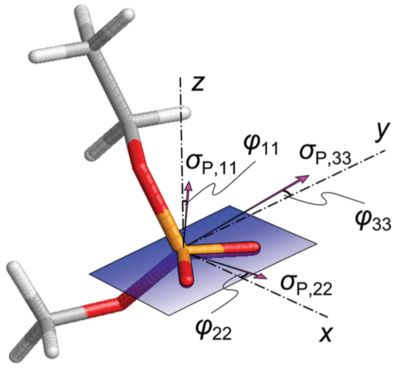


Figure 2. Schematic representation of the ^{31}P chemical shielding tensor in the EMP molecule: principal components $\sigma_{\mathrm{P},ii}$ and their orientations φ_{ii}, ii = 11, 22, 33, in the phosphate−frame coordinate system.[20]

Our results are consistent with previous calculations by Přecechtělová et al., who found that the ^{31}P-CSA values calculated in DMP with two explicit hydration patterns were by up to 130 ppm smaller than the ^{31}P-CSA in the gas phase.[15] The PCM hydration thus accounted for approximately half of the effect of static hydration on ^{31}P-CSA and also on principal components $\sigma_{\mathrm{P},ii}$ (Supporting Information, Table S1). We note that neither implicit PCM nor static hydration can be regarded as a correct description of phosphate solvation. According to the current state of knowledge, dynamical averaging of explicit solvent is probably the best theoretical method for modeling the solvation effects on NMR parameters in NA phosphate.[12,14]

The benchmark NMR calculations showed that the B3LYP method performs rather well for both ^{31}P chemical shift and ${}^2J_{\mathrm{P,C}}$ coupling. We found that hydration has a larger impact on the calculated NMR parameters than the particular choice of electronic structure method. Solvation effects should be therefore always included in NMR calculations in NA phosphate.

Geometry Dependence of ^{31}P Chemical Shift. The δ_P calculated in the EMP molecule ranged from 0.7 ppm for the rather unrealistic (0°, 0°) phosphate conformer to 13.8 ppm for the (t, t) conformer (Figure 3). The δ_P values of ∼7.5−8.0 ppm calculated for the ($g-$, $g-$) region dominantly populated in NAs were somewhat larger than the values of −2.0−0.5 ppm typically measured in NAs.[10,11] This difference can be attributed to intrinsic errors of the DFT chemical shielding calculations, to insufficient accounting for water solvent by PCM, and to improperly described ^{31}P NMR reference via the secondary standard approach. However, these systematic computational errors tend to mutually cancel when calculating the chemical shift differences between different conformers (see Table 2). Hence, the geometry trends calculated for δ_P can be trusted more than the absolute δ_P values.

The δ_P values calculated for the nPn structures were by ∼25 ppm smaller than those obtained for the corresponding (ζ, α) conformers of EMP (Supporting Information, Figure S1). This difference originates from the differences in local phosphate geometry optimized with the Amber99 and B3LYP methods (Table 1). Interestingly, the dependencies of δ_P on the (ζ, α) conformation calculated for both EMP and nPn possessed a very similar shape, thus indicating that a qualitatively correct geometry dependence of δ_P can be obtained even with the smaller EMP model.

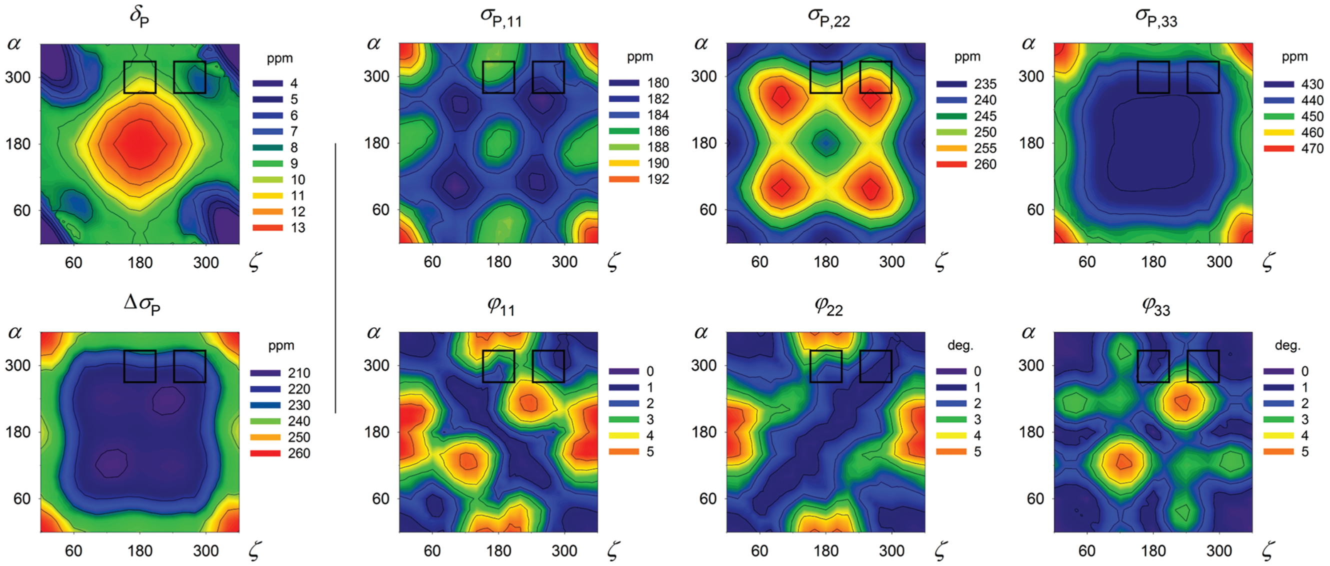


Figure 3. Dependencies of isotropic ^{31}P chemical shift δ_P, chemical shielding anisotropy $\Delta\sigma_P$, principal components $\sigma_{P,ii}$ of the ^{31}P chemical shielding tensor and their orientations in the molecular frame φ_{ii}, ii = 11, 22, 33, on the (ζ, α) conformation of the EMP molecule calculated with the B3LYP/PCM method. Black squares indicate the (ζ, α) regions that were explored previously by Přecechtělová et al.[15]

Table 3. Comparison of the $B_{II} - B_I$ Differences of δ_P in ppm, Principal Components $\sigma_{P,ii}$ in ppm, ii = 11, 22, 33, and $^2J_{P,C3'}$ Coupling in Hz, Obtained with Different Computational Approaches

			Δ_{tg-gg}				
model	geometry method	NMR method	δ_P	$\sigma_{P,11}$	$\sigma_{P,22}$	$\sigma_{P,33}$	$^2J_{P,C3'}$
EMP[a]	B3LYP	B3LYP	3.8	6.9	4.3	−11.2	3.3
EMP−PCM[a]	B3LYP/PCM	B3LYP/PCM	2.3	4.0	−0.4	−3.6	2.2
nPn[b]	Amber99	B3LYP/PCM	1.8, 1.9, 2.6				2.2, 2.2, 2.3
DMP[a,c]	B3LYP	SOS-DFPT-IGLO-PWP86	6.9	8.8	2.5	−11.3	
DMP−IC[a,c]	"	"	5.9	8.5	−2.2	−6.4	
DMP−SC[a,c]	"	"	3.8	7.9	1.2	−8.9	
B-DNA[d]	Amber99	"	2.1, 1.6				
B-DNA[e]	Exp.	Exp.	1.6				

[a] Calculated as differences between the (180°, 300°) and (300°, 300°) conformers. [b] Results obtained as differences between the respective conformers of nPn3, nPn12, and nPn20. [c] Reference 15; IC and SC denote two different explicit hydration patterns. [d] Reference 14; results for two phosphate residues of a Dickerson−Drew DNA dodecamer obtained by averaging of the respective δ_P values over MD trajectory. [e] Reference 10.

Conformational differences of δ_P were in good agreement with available experimental and theoretical data. The B3LYP/PCM calculations satisfactorily reproduced the ^{31}P chemical shift difference between the B_I and B_{II} conformations of the NA backbone (Table 3). Relative chemical shifts $\Delta_{tg-gg}(\delta_P)$ calculated in EMP (2.3 ppm) and nPn (1.8, 1.9, and 2.6 ppm) nicely coincided with the experimental $B_{II} - B_I$ difference[10] (1.6 ppm). Previous theoretical calculations of $\Delta_{tg-gg}(\delta_P)$ in DMP with two different explicit hydration patterns provided somewhat dispersed values (5.9 and 3.8 ppm, Table 3),[15] which anticipated the need of dynamic averaging of explicit hydration surrounding phosphate to obtain converged results (2.1, 1.6 ppm).[14] The need of dynamic averaging of explicit solvent was also concluded in calculations of solvation effects on ^{31}P NMR parameters in NA phosphate.[12] However, in the case of a large number of conformations like here, the dynamic approach becomes rather costly and implicit solvation may be the only feasible option. The PCM actually performed rather well, because it accounted for a large part of the conformation-dependent water solvent effect on the ^{31}P chemical shift: the $\Delta_{tg-gg}(\delta_P)$ calculated in the gas phase was 3.8 ppm, with PCM it was 2.3 ppm, and the experimental value is 1.6 ppm. The comparison of calculated $\Delta_{tg-gg}(\delta_P)$ values with the experimentally known $B_{II} - B_I$ difference thus indicates that conformational differences of δ_P calculated with the B3LYP/PCM method are qualitatively correct.

To the best of our knowledge, the theoretical δ_P data available in the literature were calculated either for a limited set of (ζ, α) phosphate conformations[15,18] or for full variation of torsion angles ζ and α with a crude semiempirical approach.[51] Our DFT calculations provide consistent theoretical insight into the complete dependence of δ_P on the (ζ, α) conformation of NA phosphate.

Geometry Dependences of ^{31}P Chemical Shielding Tensor Components. Variation of principal components $\sigma_{P,ii}$, ii = 11, 22, 33, by up to 15, 31, and 46 ppm, respectively (Figure 3), calculated with the B3LYP/PCM method was much larger than the range of isotropic values δ_P (~13 ppm). This could be explained by mutual compensation of the components. Namely, the $\sigma_{P,11}$ and $\sigma_{P,22}$ surfaces were nearly complementary, implying that the geometry dependencies of both δ_P and $\Delta\sigma_P$ are to a large extent dominated by the $\sigma_{P,33}$ component (Figure 3).

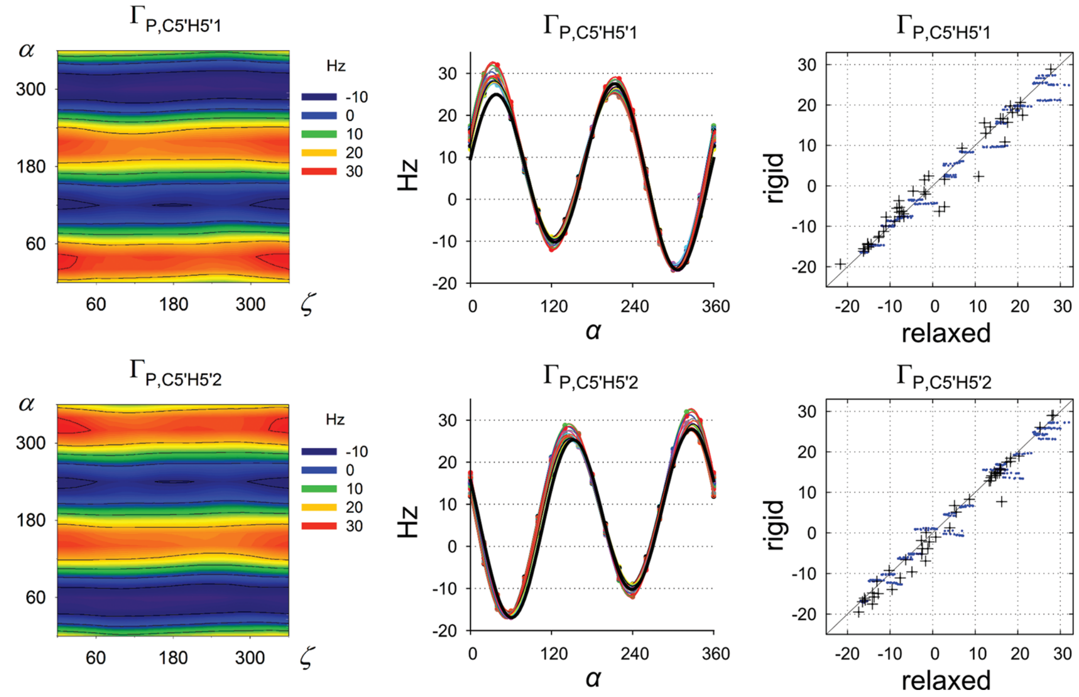


Figure 4. (left) Dependencies of the $\Gamma_{\mathrm{P,C5'H5'1}}$ and $\Gamma_{\mathrm{P,C5'H5'2}}$ cross-correlated relaxation rates on the (ζ, α) conformation of EMP calculated using the "relaxed" approach. (middle) Comparison of the "rigid" (thick black line) and "relaxed" approach (colored lines; each line corresponds to a specific value of torsion ζ ranging from 0 to 340°), here displayed in dependence on torsion α only. (right) Correlation between the "rigid" and "relaxed" approach in the calculations of $\Gamma_{\mathrm{P,C5'H5'1}}$ and $\Gamma_{\mathrm{P,C5'H5'2}}$ (in Hz) for EMP grid-point geometries with β = 180° (blue dots) and for EMP geometries sampling the RNA structural classes (black crosses).

The geometry dependences of principal components of the ^{31}P *chemical shift* tensor $\delta_{\mathrm{P},ii}$ calculated with the secondary standard approach (eq 2) were in qualitative agreement with previously calculated data for explicitly hydrated DMP in (t, $g-$) and ($g-$, $g-$) conformations[15] (Supporting Information, Figure S3). In the two conformational regions, the principal component $\delta_{\mathrm{P},22}$ consistently increased and $\delta_{\mathrm{P},33}$ decreased with the increase of torsion α. The $\delta_{\mathrm{P},ii}$ values, ii = 22, 33, calculated in this work varied in the (t, $g-$) and ($g-$, $g-$) regions by ca. 15−20 ppm, which is significantly less than calculated previously in explicitly hydrated DMP[15] (by ca. 25−30 ppm). Variation of the $\delta_{\mathrm{P},11}$ component was only a few ppm both in this and previous work. The isotropic δ_{P} values consistently decreased with increasing torsion α. The decrease, which was larger in the (t, $g-$) region, was more pronounced in the previous work by Přecechtělová et al. than in this work, which was probably caused by the difference between explicit and implicit phosphate hydration. Interestingly, our calculations provided quite similar $\mathrm{B_{II}} - \mathrm{B_I}$ differences of $\sigma_{\mathrm{P},ii}$ as obtained previously for the explicitly hydrated DMP−IC model[15] (Table 3).

The orientation of the ^{31}P chemical shielding tensor in the phosphate−frame coordinate system was nearly independent of the EMP conformation because the calculated angles φ_{ii} were smaller than 6° (Figure 3). This agrees with the previous theoretical results[15] but conflicts with the experimental solid-state NMR data for a crystal of BDEP (φ_{11} = 7°, φ_{22} = 13°, φ_{33} = 9°).[20] We have to keep in mind that the first coordination shell of the phosphate group in the BDEP crystal structure contains only Ba^{2+} ions.[50] In our previous work, we showed that direct coordination of divalent ion to the phosphate group has a significant impact on the ^{31}P chemical shielding tensor.[12] Therefore, we propose that the ^{31}P shielding tensor measured in the BDEP crystal was influenced by direct coordination of Ba^{2+} ions and that the φ_{ii} values smaller than 6° are natural for all (ζ, α) conformations of NA phosphate.

$\Gamma_{\mathrm{P,C5'H5'}}$ Cross-Correlated Relaxation Rates. We showed above that the principal components $\sigma_{\mathrm{P},ii}$ vary significantly with rotation of torsion angles ζ and α, by up to 46 ppm (Figure 3). The question is to what extent this behavior affects the $\Gamma_{\mathrm{P,CH}}$ CCR rates and how large could be the error when applying the rigid tensor approximation. We could answer this question by calculating the $\Gamma_{\mathrm{P,C5'H5'1}}$ and $\Gamma_{\mathrm{P,C5'H5'2}}$ CCR rates in EMP with the "relaxed" and "rigid" approach described in the Methods section.

The $\Gamma_{\mathrm{P,C5'H5'1}}$ and $\Gamma_{\mathrm{P,C5'H5'2}}$ CCR rates depend dominantly on torsion angles α and β.[8,9] The EMP calculations with β = 180° show only the dependence on torsions ζ and α (Figure 4, left). To see the effect of β variation, we calculated the CCR rates also for the EMP geometries sampling the RNA structural classes,[1] where the torsion β varies from 83 to 248° (Supporting Information, Figure S2).

The CCR rates obtained with the "rigid" approach are inherently independent of torsion angle ζ. The application of the "relaxed" approach only slightly perturbed the one-dimensional character of the two dependencies (Figure 4, middle). Both $\Gamma_{\mathrm{P,C5'H5'1}}$ and $\Gamma_{\mathrm{P,C5'H5'2}}$ ranged from ca. −17 to 32 Hz. The difference between CCR rates calculated using the "rigid" and "relaxed" approach was on average over all EMP geometries smaller than 2 Hz. By detailed analysis of the calculated data (not shown), we found that the differences were caused by both conformational variation of the ^{31}P chemical shielding tensor and geometry relaxation of the C5′−H5′ bond vectors affecting the angles $\vartheta_{\mathrm{CH},ii}$ in eq 5. The r_{CH}^{-3} factor varied by 1.6% at most; its influence on the CCR rates was thus

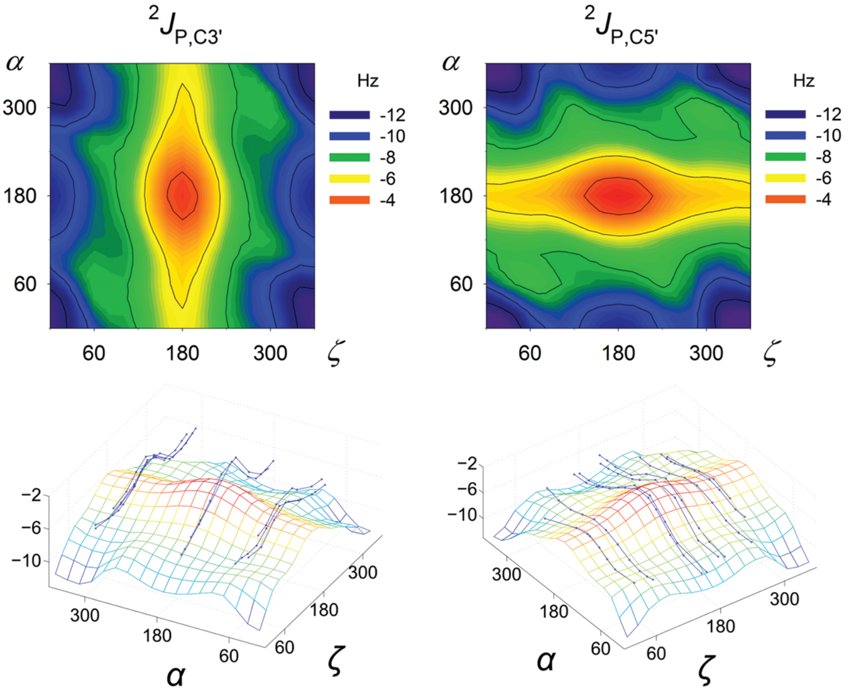


Figure 5. (top) Dependencies of the $^2J_{P,C3'}$ and $^2J_{P,C5'}$ coupling constants in Hz on the (ζ, α) conformation of EMP. (bottom) Comparison of $^2J_{P,C}$ couplings calculated in EMP (colored surfaces; the grid represents the calculated data points) and nPn (blue lines; dots mark the calculated data points) molecules.

negligible. The rigid tensor approximation can be therefore considered valid for the $\Gamma_{P,CH}$ CCR rates in NA phosphate.

$^2J_{P,C}$ Coupling Constants. For the first time, the dependencies of $^2J_{P,C3'}$ and $^2J_{P,C}$ coupling constants on the (ζ, α) conformation of NA phosphate are reported (Figure 5). We obtained two coherent sets of theoretical data for EMP and nPn molecules. While the EMP calculations sample the complete (ζ, α) dependence of the two *J*-couplings, the nPn data were obtained for selected NA backbone classes by stepwise variation of one of the two torsion angles. The $^2J_{P,C5'}$ coupling depends dominantly on torsion angle α and the $^2J_{P,C3'}$ coupling on torsion ζ, but the dependence on the second P−O torsion is also not negligible. The two-dimensional dependencies of $^2J_{P,C3'}$ and $^2J_{P,C5'}$ are similar owing to the symmetry of NA phosphate. The $^2J_{P,C}$ couplings are negative for all phosphate conformations. For better clarity, and also because the sign of the $^2J_{P,C}$ coupling is usually not determined experimentally, we further discuss only the absolute $^2J_{P,C}$ values.

The two $^2J_{P,C}$ couplings calculated in the EMP molecule range approximately from 3 to 13 Hz (Figure 5). Some conformers of NA phosphate can be clearly distinguished with the $^2J_{P,C}$ couplings. The lowest values near 3 Hz were obtained for the (*t*, *t*) conformer, while the values corresponding to the most populated NA conformer (*g*−, *g*−) were around 8 Hz. Unfortunately, the *g*+ and *g*− conformation of both ζ and α torsion angles are hardly distinguishable, since they both correspond to similar values of $^2J_{P,C}$. The $B_I \rightarrow B_{II}$ transition calculated as a change of the (ζ, α) conformation from (*g*−, *g*−) to (*t*, *g*−) was accompanied by ∼2−3 Hz decrease of the $^2J_{P,C3'}$ coupling (Table 3, Figure 5), which is in agreement with our previous work for NA conformational classes.[7] Similarly, the transition of torsion α from *gauche* to *trans* corresponds to a significant decrease of $^2J_{P,C5'}$. Also, the effect of neighboring torsions ε and β may not be negligible. For example, the $^2J_{P,C5'}$ coupling notably increases when β exceeds the 120−240° interval (Supporting Information, Figure S5).

Experimentally, the $^2J_{P,C}$ couplings typically occur between 1 and 6.5 Hz.[21−25] Structural interpretation of the measured values in terms of torsions ζ and α was so far unavailable. Here, we correlated the experimental $^2J_{P,C}$ couplings with torsion angles ζ and α for NA molecules with known structures: cyclic d<pApA> dinucleotide,[21] 10 base-pair (bp) long DNA duplex oligonucleotide[23,52] (PDB ID 1NEV), and 8 bp long stem of hairpin-35 of 23S rRNA[25] (PDB ID 2GBH). The NMR data for the most populated NA conformer (*g*−, *g*−) were used to validate our theoretical approach. The experimental $^2J_{P,C}$ couplings for the (g−, g−) conformer ranged from 4.5 to 5.5 Hz (Supporting Information, Table S2 and Figure S4), while the values calculated in EMP were near 8 Hz (Table 4). The deviation from experiment can be explained by imperfections of the B3LYP method in both geometry and NMR calculations, the effect of incomplete basis set, and the error of PCM in describing the phosphate hydration (Table 4). The dynamically averaged effect of solvent on $^2J_{P,C}$ values would be even larger when considering NA phosphate solvation with metal cation, as was reported in our previous work.[12] Somewhat larger deviations of the calculated $^2J_{P,C}$ values from experiment were obtained for other phosphate conformations; similar sources of computational errors can be expected, but their values are not known. Nevertheless, we can anticipate that the geometry dependencies of the $^2J_{P,C}$ couplings calculated with the B3LYP/PCM method are qualitatively correct, albeit the $^2J_{P,C}$ magnitudes are overestimated.

The nPn calculations provide an important link between the simple EMP model and actual NA backbone structure. The $^2J_{P,C}$ couplings calculated in the EMP and nPn molecules for the same (ζ, α) conformers differ (Figure 5), mainly due to different local geometries of NA phosphate (Table 1) and also

Table 4. Comparison of the Calculated $^2J_{P,C}$ Coupling Values and Corrections Δ*J* (in Hz) with Available Experimental Data for the Most Populated Phosphate Conformer (*g*−, *g*−)

J-coupling contributions	$^2J_{P,C3'}$	$^2J_{P,C5'}$
J: B3LYP/PCM[a]	−8.1	−7.7
Δ*J*: geometry[b]	+0.7	
Δ*J*: electron correlation[c]	+0.5	
Δ*J*: basis set[d]	+0.5	
Δ*J*: explicit solvent[e]	+0.9	−0.5
J: total	−5.4	−6.6
|*J*|: experiment[f]	4.5−5.5	

[a] Calculated for the ($\zeta = 280°$, $\alpha = 300°$, $\beta = 180°$) conformer of EMP, using the IGLO-III basis set. [b] The effect of molecular geometry obtained as a difference between the $^2J_{P,C}$ values calculated for DMP optimized with the CCSD and B3LYP methods (Table 2). [c] The effect of the NMR calculation method obtained as a difference between the $^2J^{FC}_{P,C}$ values calculated in DMP with the CCSD and B3LYP methods (Table 2). [d] The effect of basis set obtained as a difference between $^2J_{P,C}$ values calculated in EMP with the B3LYP/aug-cc-pCVQZ (near basis-set limit) and B3LYP/IGLO-III methods.[12] [e] The effect of dynamically averaged explicit solvent obtained as a difference between the $^2J_{P,C}$ values calculated for A-RNA dinucleotide with dynamically averaged explicit water solvent (−6.9 and −6.8 Hz with the uncertainty of ~0.4 Hz; see our previous work[12]) and the respective values calculated here for the same model with the same computational method employing PCM of water solvent (−7.9 and −6.3 Hz). [f] The experimental $^2J_{P,C}$ couplings for the (*g*−, *g*−) conformer taken from refs 23 and 25 (see also the Supporting Information, Table S2 and Figure S4).

because of the improper termination of the EMP molecule at C3′ carbon. On the other hand, the dependencies obtained with the two NA phosphate models correspond qualitatively to each other, showing the same conformational trends (Figure 5). The $^2J_{P,C3'}$ and $^2J_{P,C5'}$ couplings calculated for the (*g*−, *g*−) conformer of nPn ranged from 4.9 to 6.3 Hz and from 6.7 to 7.5 Hz, respectively. The *J*-couplings calculated with presumably lower quality force-field geometries were thus closer to the experimental range 4.5−5.5 Hz than those obtained for the DFT-optimized EMP structures. This was probably caused by error cancellation. Nevertheless, the correspondence between geometry trends calculated in EMP and nPn indicates that the dependencies shown in Figure 5 should be generally valid for the NA backbone.

Structural interpretation of the $^2J_{P,C3'}$ and $^2J_{P,C5'}$ couplings can be only qualitative, namely, because the two-dimensional character of their dependences implies that similar $^2J_{P,C}$ values may correspond to different phosphate conformers. Furthermore, the computational imperfections mentioned above and the uncertainty of experimental determination of $^2J_{P,C}$ couplings (which may be larger than 1 Hz[23,24]) prohibited so far their accurate calibration. Nevertheless, we found that the $^2J_{P,C3'}$ and $^2J_{P,C5'}$ couplings can be loosely assigned to NA backbone torsion angles ζ and α, respectively. On the basis of our calculations and comparison with available experimental data, we propose two qualitative rules: (a) The $^2J_{P,C}$ coupling smaller than ~3 Hz should indicate the orientation of the assigned torsion angle (ζ or α) that is different from typical *gauche* (near 70 or 290°). (b) The $^2J_{P,C}$ coupling larger than ~4.5 Hz should indicate that the orientation of the assigned torsion angle deviates from the 120−240° region. The $^2J_{P,C}$ values between 3 and 4.5 Hz are difficult to interpret in terms of torsions ζ and α because the dominant *J*-coupling dependence on the assigned torsion is largely perturbed by variation of the second torsion. These rules may serve as qualitative restraints on phosphate conformation in NMR structural studies of NAs. Their further validation based on more extended comparison with accurate experimental data is clearly advisable.

$^3J_{P,C4'}$ Coupling. The $^3J_{P,C4'}$ coupling is assigned to torsion angle β with a standard Karplus equation.[22,53,54] The structural interpretation of $^3J_{P,C4'}$ relies on its independence of other torsion angles. To validate this assumption, we calculated the dependence of the $^3J_{P,C4'}$ coupling on torsion angles ζ and α in the EMP molecule. The calculated variation of $^3J_{P,C4'}$ was quite large, ca. 2 Hz (Supporting Information, Figure S6). Such variation in the $^3J_{P,C4'}$ value interpreted with standard Karplus equations can cause error in determination of the assigned torsion β as large as ~25°.

The calculations thus revealed that variation of the NA phosphate conformation may affect the spin−spin coupling pathways between phosphorus and relatively distant atoms. This could have been foreseen, since the complicated electronic structure of negatively charged NA phosphate sensitively responds to the local geometry deformations and a perturbation at one side of the spin−spin coupling pathway is in this way promoted to the total *J*-coupling value.

■ CONCLUSIONS

In this theoretical work, the isotropic ^{31}P chemical shift, the principal components $\sigma_{P,ii}$ of the ^{31}P chemical shielding tensor, and their orientations in the molecular frame, the $\Gamma_{P,C5'H5'1}$ and $\Gamma_{P,C5'H5'2}$ cross-correlated relaxation rates, as well as the $^2J_{P,C3'}$, $^2J_{P,C5'}$, and $^3J_{P,C4'}$ coupling constants were calculated in dependence on nucleic acid backbone torsion angles ζ and α.

Although the NMR calculations did not provide absolute accuracy of the isotropic ^{31}P chemical shift, the relative values for different phosphate conformations were in good agreement with previous experimental and theoretical data. Geometry dependencies of both isotropic ^{31}P chemical shift and chemical shielding anisotropy were dominated by the $\sigma_{P,33}$ principal component, owing to mutual compensation of the $\sigma_{P,11}$ and $\sigma_{P,22}$ components. Deviations of the ^{31}P principal components from phosphate frame axes were smaller than 6°. The orientation of the ^{31}P chemical shielding tensor in the molecular frame is thus almost independent of torsion angles ζ and α.

We found that the $\Gamma_{P,C5'H5'1}$ and $\Gamma_{P,C5'H5'2}$ cross-correlated relaxation rates are nearly independent of torsion angle ζ, being only slightly modulated by the conformational variation of the ^{31}P chemical shielding tensor and local phosphate geometry. The approximation assuming a rigid ^{31}P chemical shielding tensor can be thus considered valid for the interpretation of $\Gamma_{P,CH}$ cross-correlated relaxation rates in nucleic acids.

DFT calculations of $^2J_{P,C}$ coupling constants in nucleic acid backbone revealed their strong dependence on both torsion angles ζ and α with dominant modulation by only one of the torsions. The $^2J_{P,C3'}$ and $^2J_{P,C5'}$ couplings thus could be structurally assigned to torsions ζ and α, respectively, and interpreted as loose restraints on the phosphate conformation. The absolute $^2J_{P,C}$ values calculated with the B3LYP/PCM method were overall overestimated relative to the experiment. Computational errors for the most populated phosphate conformer (*g*−, *g*−) were ascribed to imperfect performance of the DFT method, which was tested against the benchmark CCSD method, incomplete atomic basis set, and incomplete description of phosphate hydration by PCM. Our results provide the missing link between the $^{2}J_{\mathrm{P,C3'}}$ and $^{2}J_{\mathrm{P,C5'}}$ spin−spin coupling constants and the NA backbone torsion angles ζ and α.

ASSOCIATED CONTENT

Supporting Information

Comparison of the dependencies of ^{31}P chemical shift on the (ζ, α) conformation of EMP and nPn. Correlation between backbone torsion angles ζ, α, and β in RNA conformational classes. Comparison of calculated principal components of the ^{31}P chemical shielding tensor and their orientations in the phosphate−frame coordinate system with previous theoretical results and NMR experiment. Correlation of available experimental $^{2}J_{\mathrm{P,C}}$ couplings with torsion angles ζ and α. Dependence of the $^{2}J_{\mathrm{P,C5'}}$ coupling on torsion angle β. Dependence of the $^{3}J_{\mathrm{P,C4'}}$ coupling on the (ζ, α) conformation of EMP. Optimized geometries of the DMP and EMP models. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ladislav.benda@uochb.cas.cz (L.B.); vladimir.sychrovsky@uochb.cas.cz (V.S.). Phone: +420 220 183 234. Fax: +420 220 183 578.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Czech Science Foundation, Grant No. 205/10/0228. Z.S.V. and M.S. acknowledge support by the Czech Science Foundation, Grant Nos. P208/10/P398 and 203/09/2037, respectively. V.S. was supported by a Human Frontier Science Program (HFSP) Young Investigator's Grant.

REFERENCES

(1) Richardson, J. S.; Schneider, B.; Murray, L. W.; Kapral, G. J.; Immormino, R. M.; Headd, J. J.; Richardson, D. C.; Ham, D.; Hershkovits, E.; Williams, L. D.; Keating, K. S.; Pyle, A. M.; Micallef, D.; Westbrook, J.; Berman, H. M. *RNA* **2008**, *14*, 465−481.

(2) Schneider, B.; Morávek, Z.; Berman, H. M. *Nucleic Acids Res.* **2004**, *32*, 1666−1677.

(3) Varani, G.; Aboul-ela, F.; Allain, F. H.-T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996**, *29*, 51−127.

(4) Wijmenga, S. S.; van Buuren, B. N. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287−387.

(5) Fürtig, B.; Richter, C.; Wohnert, J.; Schwalbe, H. *ChemBioChem* **2003**, *4*, 936−962.

(6) Flinders, J.; Dieckmann, T. *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *48*, 137−159.

(7) Sychrovský, V.; Vokáčová, Z.; Šponer, J.; Špacková, N.; Schneider, B. *J. Phys. Chem. B* **2006**, *110*, 22894−22902.

(8) Richter, C.; Reif, B.; Griesinger, C.; Schwalbe, H. *J. Am. Chem. Soc.* **2000**, *122*, 12728−12731.

(9) Nozinovic, S.; Richter, C.; Rinnenthal, J.; Fürtig, B.; Duchardt-Ferner, E.; Weigand, J. E.; Schwalbe, H. *J. Am. Chem. Soc.* **2010**, *132*, 10318−10329.

(10) Gorenstein, D. G. *Chem. Rev.* **1994**, *94*, 1315−1338.

(11) Gorenstein, D. G. *Phosphorus-31 NMR*; Academic Press: Orlando, FL, 1984.

(12) Benda, L.; Schneider, B.; Sychrovský, V. *J. Phys. Chem. A* **2011**, *115*, 2385−2395.

(13) Ribas Prado, F.; Giessner-Prettre, C.; Pullman, B.; Daudey, J. P. *J. Am. Chem. Soc.* **1979**, *101*, 1737−1742.

(14) Přecechtělová, J.; Novák, P.; Munzarová, M. L.; Kaupp, M.; Sklenář, V. *J. Am. Chem. Soc.* **2010**, *132*, 17139−17148.

(15) Přecechtělová, J.; Munzarová, M. L.; Novák, P.; Sklenář, V. *J. Phys. Chem. B* **2007**, *111*, 2658−2667.

(16) Gorenstein, D. G. *J. Am. Chem. Soc.* **1975**, *97*, 898−900.

(17) Giessner-Prettre, C.; Pullman, B.; Parado, F. R.; Cheng, D. M.; Iuorno, V.; Ts'O, P. O. P. *Biopolymers* **1984**, *23*, 377−388.

(18) Přecechtělová, J.; Padrta, P.; Munzarová, M. L.; Sklenář, V. *J. Phys. Chem. B* **2008**, *112*, 3470−3478.

(19) Lerner, D. B.; Becktel, W. J.; Everett, R.; Goodman, M.; Kearns, D. R. *Biopolymers* **1984**, *23*, 2157−2172.

(20) Herzfeld, J.; Griffin, R. G.; Haberkorn, R. A. *Biochemistry* **1978**, *17*, 2711−2718.

(21) Blommers, M. J. J.; Haasnoot, C. A. G.; Walters, J.; Vandermarel, G. A.; Vanboom, J. H.; Hilbers, C. W. *Biochemistry* **1988**, *27*, 8361−8369.

(22) Mooren, M. M. W.; Wijmenga, S. S.; Vandermarel, G. A.; Vanboom, J. H.; Hilbers, C. W. *Nucleic Acids Res.* **1994**, *22*, 2658−2666.

(23) Zimmer, D. P.; Marino, J. P.; Griesinger, C. *Magn. Reson. Chem.* **1996**, *34*, S177−S186.

(24) Richter, C.; Reif, B.; Worner, K.; Quant, S.; Marino, J. P.; Engels, J. W.; Griesinger, C.; Schwalbe, H. *J. Biomol. NMR* **1998**, *12*, 223−230.

(25) O'Neil-Cabello, E.; Wu, Z. R.; Bryce, D. L.; Nikonowicz, E. P.; Bax, A. *J. Biomol. NMR* **2004**, *30*, 61−70.

(26) van Wüllen, C. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2137−2144.

(27) Jameson, C. J.; de Dios, A. C.; Jameson, A. K. *Chem. Phys. Lett.* **1990**, *167*, 575−582.

(28) Kumar, A.; Grace, R. C. R.; Madhu, P. K. *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *37*, 191−319.

(29) Kowalewski, J.; Mäler, L. *Nuclear Spin Relaxation in Liquids: Theory, Experiments, and Applications*; Taylor & Francis Group: New York, London, 2006.

(30) Reif, B.; Hennig, M.; Griesinger, C. *Science* **1997**, *276*, 1230−1233.

(31) Reif, B.; Diener, A.; Hennig, M.; Maurer, M.; Griesinger, C. *J. Magn. Reson.* **2000**, *143*, 45−68.

(32) Duchardt, E.; Richter, C.; Ohlenschlager, O.; Gorlach, M.; Wohnert, J.; Schwalbe, H. *J. Am. Chem. Soc.* **2004**, *126*, 1962−1970.

(33) Rinnenthal, J.; Richter, C.; Ferner, J.; Duchardt, E.; Schwalbe, H. *J. Biomol. NMR* **2007**, *39*, 17−29.

(34) Sychrovský, V.; Müller, N.; Schneider, B.; Smrečki, V.; Špirko, V.; Šponer, J.; Trantírek, L. *J. Am. Chem. Soc.* **2005**, *127*, 14663−14667.

(35) Benda, L.; Bouř, P.; Müller, N.; Sychrovský, V. *J. Phys. Chem. B* **2009**, *113*, 5273−5281.

(36) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(37) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(39) Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2010**, *132*, 114110.

(40) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910−1918.

(41) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. *CFOUR*, a quantum chemical program package, 2010, http://www.cfour.de.

(42) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L. S.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. *J. Chem. Inf. Model.* **2007**, *47*, 1045−1052.

(43) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.

(44) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1995**, *102*, 251−253.

(45) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1995**, *103*, 3561−3577.

(46) Sychrovský, V.; Grafenstein, J.; Cremer, D. *J. Chem. Phys.* **2000**, *113*, 3530−3547.

(47) Helgaker, T.; Watson, M.; Handy, N. C. *J. Chem. Phys.* **2000**, *113*, 9402−9409.

(48) Auer, A. A.; Gauss, J. *Chem. Phys.* **2009**, *356*, 7−13.

(49) Florián, J.; Štrajbl, M.; Warshel, A. *J. Am. Chem. Soc.* **1998**, *120*, 7959−7966.

(50) Kyogoku, Y.; Iitaka, Y. *Acta Crystallogr.* **1966**, *21*, 49.

(51) Gorenstein, D. G.; Luxon, B. A. *Biochemistry* **1979**, *18*, 3796−3804.

(52) Barbic, A.; Zimmer, D. P.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 2369−2373.

(53) Lankhorst, P. P.; Haasnoot, C. A. G.; Erkelens, C.; Altona, C. *J. Biomol. Struct. Dyn.* **1984**, *1*, 1387−1405.

(54) Plavec, J.; Chattopadhyaya, J. *Tetrahedron Lett.* **1995**, *36*, 1949−1952.