RISM-BASED PRESSURE-DEPENDENT COMPUTATIONAL SPECTROSCOPY

Dissertation zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)

Die Dissertation wurde im Zeitraum 10.2017 bis
11.2022 angefertigt und der Fakultät Chemie und
Chemische Biologie der Technischen Universität
Dortmund vorgelegt

von Tim Pongratz

DORTMUND, 2022

Erstgutachter: Prof. Dr. Stefan M. Kast
Zweitgutachter: Prof. Dr. Roland Winter

Danksagung

An dieser Stelle möchte ich mich zunächst bei Prof. Dr. Stefan Kast bedanken. Vielen Dank, dass du mir die Möglichkeit gegeben hast, diese Arbeit und meine vorherigen Abschlussarbeiten in deiner Gruppe anzufertigen. Während der gesamten Zeit habe ich mich sehr wohl gefühlt und viel gelernt. Bei wissenschaftlichen Fragen und anderen Problemen konnte man jederzeit zu dir kommen und du hast mich unterstützt, dafür möchte ich mich bedanken.

Des Weiteren möchte ich mich bei Prof. Dr. Roland Winter für die Übernahme des Zweitgutachtens bedanken. Sie haben mich während meiner gesamten akademischen Laufbahn begleitet und ohne die experimentellen Arbeiten in Ihrer Arbeitsgruppe wäre die Dissertation in dieser Form nicht möglich gewesen.

Außerdem möchte ich mich bei allen Mitgliedern der Arbeitsgruppe für die angenehme Arbeitsatmosphäre und die ständige Hilfsbereitschaft bedanken. Bei Patrick Kibies und Lukas Eberlein möchte ich mich dafür bedanken, dass ihr damals die Betreuer meiner Abschlussarbeiten wart und wir danach noch viele lustige und schöne Momente zusammen erlebt haben, unter anderem in Regensburg und Tokio. Unter anderem möchte ich mich bei Nico, Lenny, Julia und Martin für die vielen Tischtennis-Matches bedanken, die wir zusammengespielt haben. Außerdem möchte ich noch Stefan Maste, Chris, Fabian. Lars und Yannic danken.

Mein Dank gilt den Mitgliedern der DFT-Forschungsgruppe FOR1979 und des Exzellenzclusters RESOLV, ohne deren Beiträge diese Arbeit nicht möglich gewesen wäre. Especially, I want to thank Van Anh Tran, Laura Galazzo and Bikramjit Sharma, for the productive, fruitful and pleasant cooperation. Danken möchte ich ebenfalls Saskia Körning, Christoph Hölzl, Markus Teucher, Markus Beck Erlach und Paul Schummel.

Außerdem möchte ich mich bei meiner Fußballmannschaft Teutonia Überruhr 3 bedanken, ohne die ich neben der Promotion wahrscheinlich nicht genügend Ausgleich gefunden hätte. Bedanken möchte ich mich bei Fabian, Alex, Julian, Philipp, Max, Benett und Steffen.

Meiner Familie möchte ich natürlich danken, die mich während der gesamten Zeit immer unterstüzt hat. Zu letzt möchte ich natürlich Rudi und Fini danken.

Publications

Parts of this thesis were already published in the following publications:

Pongratz, T., Kibies, P., Eberlein, L., Tielker, N., Hölzl, C., Imoto, S., Beck Erlach,M., Kurrmann, S., Schummel, P. H., Hofmann, M., Reiser, O., Winter, R., Kremer, W., Kalbitzer, H. R., Marx, D., Horinek, D. & Kast, S. M. Pressure-dependent electronic structure calculations using integral equation-based solvation models. *Biophys. Chem.* **257,** 106258; 10.1016/j.bpc.2019.106258 (2020).

Sharma, B., Tran, V. A., Pongratz, T., Galazzo, L., Zhurko, I., Bordignon, E., Kast, S. M., Neese, F. & Marx, D. A Joint Venture of Ab Initio Molecular Dynamics, Coupled Cluster Electronic Structure Methods, and Liquid-State Theory to Compute Accurate Isotropic Hyperfine Constants of Nitroxide Probes in Water. *J. Chem. Theory Comput.* **17,** 6366–6386; 10.1021/acs.jctc.1c00582 (2021).

Table of contents

Zusammenfassung

Spektroskopische Messungen sind ein unverzichtbares Werkzeug in der chemischen Analyse, auch unter extremen Bedingungen wie zum Beispiel hohen hydrostatischen Drücken können sie wertvolle Erkenntnisse liefern. Zur Bestätigung der gewonnenen Ergebnisse können theoretische Methoden eingesetzt werden, die Observablen in Lösung verlässlich wiedergeben können. Ein gängiges theoretisches Modell ist das *Reference Interaction Site Model* (RISM), welches in dieser Arbeit verwendet wurde.

Im Rahmen dieser Arbeit wurden die Spektroskopie-Arten Infrarotspektroskopie (IR), Elektronenspinresonanzspektroskopie (ESR/EPR) und die Kernspinresonanzspektroskopie (NMR) theoretisch untersucht und mit experimentellen Referenzdaten verglichen. Die Arbeit gliedert sich dabei in drei Abschnitte, die wie folgt aufgebaut sind.

Im ersten Teil wurde für die IR-Spektroskopie eine bereits entwickelte Methode zur numerischen Berechnung von IR-Frequenzen mit *embedded cluster*(EC)-RISM unter Gleichgewichtsbedingungen auf die nicht-Gleichgewichts-thermodynamik erweitert. Als Modellsysteme wurden die druckabhängigen IR-Frequenzverschiebungen von TMAO und dem Cyanid-Anion untersucht. Darüber hinaus wurde der Einfluss elektrostatischer Modelle und so genannter Lösemittelsuszeptibilitäten eingehend untersucht. Es wurde gezeigt, dass die Druckabhängigkeit am besten durch eine Mischung aus Nichtgleichgewichts- und Gleichgewichts-Solvatisierung beschrieben werden kann.

Des Weiteren wurde hier EC-RISM zum ersten Mal für die Berechnung von EPR-Observablen bei Umgebungsdruck verwendet. Als Modellsystem wurde das starre Molekül 2,2,3,4,5,5,-hexamethylidaolidin-1-oxyl (HMI) verwendet, das eine sehr starke Abhängigkeit der Hyperfeinkopplungskonstante vom pH-Wert zeigt. Erste Versuche mit der geometrisch optimierten Struktur zeigten, dass EC-RISM trotz großer Abweichung vom Experiment deutlich bessere Ergebnisse liefert als eine Standard-Kontinuumsberechnung. Eine wesentliche Verbesserung in Richtung der experimentellen Werte, konnte dadurch erzielt werden, dass die Berechnungen nicht mit einer einzelnen Geometrie durchgeführt wurden, sondern mit einer großen Anzahl von *Snapshots*, die aus einer *ab initio*-Moleküldynamiksimulation (AIMD) stammen. Weitere Fortschritte konnten durch die Verwendung der *domain-based pair natural orbital coupled-cluster theory* (DLPNO-CCSD)-Methode erzielt werden.

Generell kann im Zusammenhang mit der theoretischen Beschreibung von Hochdruckeffekten auf Proteinen die kritische Frage gestellt werden, ob die Verwendung von Kraftfeldern, die für Umgebungsbedingungen parametrisiert sind, für Hochdruckbedingungen geeignet ist. Um diese Frage zu beantworten, wurde die Druckabhängigkeit des Peptidrückgrats im dritten Teil untersucht und die kleinen Moleküle *N*-Methyl-Acetamid (NMA) und Ac-Gly/Ala-NHMe als Modellsysteme verwendet. Das experimentelle Gleichgewicht zwischen *cis*- und *trans*-NMA konnte sehr gut reproduziert werden. Auf der Grundlage dieser Ergebnisse

wurden die Populationen für die Hauptkonformere der Dipeptide bestimmt, die eine gute Übereinstimmung mit experimentellen Referenzdaten zeigen. Infolge der Druckerhöhung wurde nur eine geringfügige Änderung der Populationen beobachtet. Zur weiteren Überprüfung der Ergebnisse wurden druckabhängige NMR-Berechnungen durchgeführt. Der korrekte Drucktrend wurde für alle Amidkerne beobachtet; eine weitere Bestätigung, dass EC-RISM in der Lage ist, die Druckeffekte akkurat zu reproduzieren. Diese Ergebnisse konnten durch neuartige druckabhängige EC-RISM-Geometrieoptimierungen verbessert werden. Schließlich wurde zur Beantwortung der Ausgangsfrage die vollständige Ramachandran-Ebene mit verschiedenen EC-RISM- und Kraftfeldmethoden berechnet. Es wurde festgestellt, dass die Druckabhängigkeit nur gering ist und dass der durch die Kraftfelder verursachte Fehler deutlich größer ist als der Fehler, der entsteht, durch die Verwendung eines umgebungsdruckoptimierten Kraftfeldes bei Hochdruck.

Im Rahmen dieser Arbeit konnte gezeigt werden, dass EC-RISM eine geeignete Methode der Wahl zur Berechnung von spektroskopischen Observablen in Lösung ist. Speziell wenn nicht Umgebungsbedingungen traktiert werden sollen, spielt EC-RISM seine Stärke aus, da es relativ leicht erweiterbar ist für z.B. Hochdruckumgebungen.

Abstract

Spectroscopic measurements are an indispensable tool in chemical analysis; even under extreme conditions such as high hydrostatic pressures, they can provide valuable insights. Theoretical methods that can reliably reproduce observables in solution can be used to validate the obtained results. A common theoretical model is the Reference Interaction Site Model (RISM), which was used in this work.

Especially the spectroscopy types infrared spectroscopy (IR), electron spin resonance spectroscopy (ESR/EPR), and nuclear magnetic resonance spectroscopy (NMR) were theoretically investigated and compared with experimental reference data. The work is divided into three sections as follows.

In the first part, a previously developed method for calculating IR frequencies with the embedded cluster(EC)-RISM under equilibrium conditions was extended to non-equilibrium thermodynamics for IR spectroscopy. The pressure-dependent IR frequency shifts of TMAO and the cyanide anion were investigated as model systems. In addition, the influences of electrostatic models and so-called solvent susceptibilities were studied in detail. It was shown that the pressure dependence is best described by a mixture of non-equilibrium and equilibrium solvation.

Furthermore, EC-RISM was used here for the first time to calculate EPR observables at ambient conditions. The rigid molecule 2,2,3,4,5,-hexamethylidaolidine-1-oxyl (HMI) was used as a model system, which shows a very strong dependence of the hyperfine coupling constant on the pH. First, experiments with the geometrically optimized structure showed that EC-RISM gives significantly better results than a standard continuum calculation despite a large deviation from the experiment. A significant improvement in the direction of the experimental values was achieved by using a large number of snapshots from an ab initio molecular dynamics simulation (AIMD) instead of a single geometry. Further progress was made using the domain-based pair natural orbital coupled-cluster theory (DLPNO-CCSD) method.

In general, in the context of the theoretical description of high-pressure effects on proteins, the critical question can be raised whether using force fields parameterized for ambient conditions is appropriate for high-pressure conditions. To answer this question, the pressure dependence of the peptide backbone was investigated in the third part, and the small molecules *N*-methyl acetamide (NMA) and Ac-Gly/Ala-NHMe were used as model systems. The experimental equilibrium between *cis*- and *trans*-NMA could be reproduced very well. After this, populations were determined for the main conformers of the dipeptides, showing good agreement with experimental reference data. As a result of the pressure increase, only a marginal change in the populations was observed. Pressure-dependent NMR calculations were performed to verify the results further. The correct pressure trend was observed for all amide nuclei, confirming that EC-RISM can accurately reproduce the pressure effects. Novel pressure-dependent

EC-RISM geometry optimizations could improve these results. Finally, the full Ramachandran plane was calculated using different EC-RISM and force field methods to answer the initial question. It was found that the pressure dependence is only small and that the error caused by the force fields is significantly larger than the error caused by using an ambient pressure optimized force field at high pressure.

In this work, it was shown that EC-RISM is a suitable method of choice for the calculation of spectroscopic observables in solution. Especially when non-ambient conditions are to be examined, EC-RISM shows its strength since it is relatively easily extensible, e.g., high-pressure environments.

1 Introduction

1.1 Overview and outline

Despite its extreme conditions, both in terms of temperature and pressure, the deep sea represents one of the largest habitats on Earth.[1] This result is surprising since proteins, in particular, are very sensitive to changes in temperature and pressure. For example, it can be observed that pressure has a particularly strong influence on the intermolecular interactions, the hydrophobic interactions, and the conformational states of proteins.[2] The effect can be so strong that the denaturation sets in in the 1-4 kbar pressure range.[3] The pressure-dependent unfolding is caused by the fact that the partial molar volume of folded proteins is larger than that of denatured proteins. Under pressure, states with smaller partial molar volumes are thermodynamically favored.[4] The differences in the partial molar volume can be explained by the non-ideal packing of the folded protein and the resulting voids, which can be closed by water molecules solvating the peptide chain, leading to the protein's unfolding.[5,6] The pressure-dependent unfolding is generally slow due to the unfavorable activation volume under high pressure.[3] Another example of the effect of high pressure on proteins is reversible inhibition of the aggregation amyloid Aβ up to a pressure of 180 MPa.[7]

Important applications of high pressure can be found in industry, science, and engineering. Especially in the food industry, high pressure can be used as a preservation method and represents a gentler alternative in retaining texture, consistency, smell, and color compared to the commonly used high-temperature treatment.[8,9]

Since experiments under high-pressure conditions are elaborate, difficult, and costly, theoretical models to evaluate pressure dependence are a desirable alternative to get more insights into the underlying mechanisms. A common method would be to utilize pressure-dependent molecular dynamics (MD) simulations with explicit water molecules. In the past, MD simulations addressing this problem were utilized[10–16]; nevertheless, the total number of pressure-dependent unfolding simulations lags behind the number of temperature-dependent simulations.[17] However, all these simulations used force fields, which were all parameterized for ambient conditions, and consequently, the question arises if these force fields are reliable for extreme conditions like, e.g., high pressure. To address this problem, in 2016, Hölzl et al.[18] showed that for a specific force field representing the small osmolyte trimethylamine *N*-oxide (TMAO), a pressure-dependent reparameterization of the atomic point charges was necessary to represent the increasing dipole moment upon pressurization. In contrast, for the counteracting molecule of TMAO, urea, the force field for ambient pressure showed the best results in representing pressure-dependent density fluctuations, whereas scaling the charges led to worse results.[19]

Alternative approaches based on quantum chemical (QC) calculations use either implicit solvation models, e.g., continuum models[20], or integral equation theory-based methods. The so-called extreme pressure polarizable continuum model (PCM-XP) was developed to extend the widely used continuum models to the high-pressure regime.[21–23] The idea behind this approach is based on one of the fundamental thermodynamic equations, which states that the partial derivative of the free energy with respect to the pressure is equal to the partial molar volume. In the PCM-formalism, the cavity is adapted to hold this thermodynamic relation. So far, these models have been used to make predictions of reaction equilibria and vibrational spectroscopic data in non-aqueous environments.

An established model to represent aqueous solutions based on integral equation theory is the embedded cluster reference interaction site model (EC-RISM).[24] The idea behind the EC-RISM formalism is to calculate the solute's free energy in a solvent where the solute and the solvent mutually polarize each other. The solvent perturbation due to the solute is encoded in the solute-solvent interaction potential. Additionally, a so-called solvent susceptibly χ is used to represent the solvent response. Normally, χ is precomputed, and it contains the pressure information via the density and dielectric constant of the solvent at the corresponding pressure.[18] Two ways to calculate χ are normally used; on the one hand, the information can be extracted from a pressure-dependent MD simulation or, on the other hand, from a 1D RISM calculation. The interaction potential consists of the Lennard-Jones interactions and the Coulomb potential. For the former, known force field parameters are used for both solute and solvent; in the case of the electrostatic interactions, the solute partial charges are derived from QC calculations, and only the solvent charges are taken from a force field. In contrast to the continuum models, EC-RISM retains the granularity of the solvent and can represent the effect of H-bonds.

For example, the EC-RISM model showed excellent results in predicting pK_a values in the SAMPL6 challenge for ambient conditions.[25] The first benchmark to evaluate the performance at high pressure was the study of the pressure-dependent change of the dipole moment of TMAO compared with ab initio MD (AIMD) data.[18] Another study, with the collaboration of the author, was the prediction of the pressure-dependent chemical NMR shifts. [26] In the first place, the shielding constants of the NMR-standard DSS were calculated and used as a reference for pressure-dependent NMR calculation of the peptide fragment NMA. In a recent publication[27], a pressure-dependent correction of excess chemical potential by scaling the partial molar volume was introduced. By applying this correction to calculating the pressure-dependent ion product of water, remarkable agreement with experimental reference data could be obtained. Furthermore, new pressure-dependent NMR calculations were performed on DSS, TMAO, and purine nucleotides.[27–29]

The upcoming section provides a few fundamentals about common force fields, the application of spectroscopy in solution, and a short history and overview of nitroxides. Theoretical descriptions of common solvent models are presented, where the man focus lies on the continuum and statistical solvent models.

Subsequently, a detailed description of the theory of IR, EPR, and NMR spectroscopy is provided. In the methodology chapter, the underlying principles of some calculation methods are explained, and computational details are provided.

The first topic of the results section of this thesis addresses the pressure-dependent infrared frequency shifts of TMAO and the cyanide anion. Here, two different aspects are analyzed in detail; on the one hand, the influence of equilibrium and non-equilibrium solvation on the wavenumbers is investigated. On the other hand, the impact of different combinations of calculating the electrostatic interaction within EC-RISM and representing the solvent response with the solvent susceptibility is benchmarked.

The second part of the results section is dedicated to the first calculations of EPR parameters with EC-RISM at ambient conditions. In the first instance, the performance of EC-RISM and conductor-like polarizable continuum model (CPCM) calculations in reproducing the isotropic hyperfine couplings are compared on the minima geometries of the test system HMI. Snapshots from an AIMD simulation of HMI in water were provided for more sophisticated results. The trajectory was divided into two subsets, one where all water molecules were removed and another where only the first two solvation shells around the N-O moiety were retained. EC-RISM and CPCM DFT-based calculations were compared with a Quantum-mechanics/molecular mechanics (QM/MM) approach for the two subsets. DLPNO-CCSD calculations, in combination with EC-RISM, will be utilized for the first time to improve the quality of the calculations further. The other parts of this section will deal with the calculation of the full W-band spectrum in solution at different pH values. In the last part, a short outlook of pressure-dependent calculations will be given however experimental reference values are not available now.

The last chapter of the results part aims to answer how pressure affects the relative occupancy of the stable peptide backbone model systems. Therefore, on the one hand, geometries are optimized at ambient conditions using the PCM solvation model and then subjected to pressure-dependent EC-RISM calculations. On the other hand, the novel EC-RISM-based geometry optimizations are used to see the influence of pressure on the geometries, minima populations, and NMR parameters. Initially, the pressure dependence of the *cis*/*trans* equilibrium of *N*-methylacetamide (NMA) is investigated, and based on these results, similar analyses are performed for the dipeptides *N*-acetyl-L-glycine-N'-methyl amide (Ac-Gly-NHMe) and *N*-acetyl-L-alanine-N'-methyl amide (Ac-Ala-NHMe). Additionally, the Ramachandran surfaces of Ac-Gly-NHMe are calculated with a wide range of different levels of theories and force fields, and the influence of pressure is discussed.

In the end, an overall summary of the results is provided, and an outlook is given

1.2 Force field parameters and typical protein backbone models

Accurate force fields are the foundation for reliable MD simulations. Critical parameters within the simulations of the dynamics of proteins are the backbone torsion and the non-bonded parameters, which have an essential impact on the conformational ensemble of proteins. Famous examples of protein force fields are AMBER,[30–37] CHARMM[38–40], and GROMOS.[41] For example, the AMBER force field has the following general form:

$$\begin{aligned} u_{\text{total}} = \sum_{\text{bonds}} K_b \left(b - b_{\text{eq}}\right)^2 + \sum_{\text{angles}} K_\theta \left(\theta - \theta_{\text{eq}}\right)^2 \\ + \sum_{\text{dihedrals}} \frac{V_n}{2}\left[1 + \cos(n\phi - \gamma)\right] \\ + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right] \end{aligned} \tag{1}$$

where u_{total} is the total potential energy of the system, K_b is the bond force constant, b is the bond length and b_{eq} is the equilibrium value of the bonds. The terms K_θ, θ and θ_{eq} are the corresponding quantities for the bond angles. V_n is the force constant for the torsion, n is periodicity, ϕ is the torsion angle, and γ is the phase angle. In the last term, the non-bonded interactions are described, whereby A_{ij} and B_{ij} are the Lennard-Jones terms, respectively. q_i and q_j are the atomic charges, and ε is the dielectric constant of the medium and resembles the effect of the solvent if it is not explicitly represented. When the solvent is treated explicitly, the value is set to 1. In the case of protein force fields, the bond, angle, and structure parameters are normally parameterized to represent experimentally measured vibrational frequencies reliably.[33,34] The parameters for the van der Waals interactions are difficult to determine using pure quantum mechanical calculations. A combination of an electron correlation method and a very large basis set is needed. Due to the time-consuming nature of such calculations, the parameters are often also fitted to experimental data such as evaporation enthalpies or experimental densities.[30]

One of the main problems in the parameterization of force fields is the generation of partial charges. Since partial charges are not quantum mechanical observables, they cannot be extracted directly from the wave function. Instead, the molecular electrostatic potential (ESP) can be extracted from the wave function. The partial charges are then fitted to minimize the difference error between the wave function ESP and the ESP generated from the partial charges. Among the best-known algorithms for calculating the electrostatic potential are the Merz-Singh-Kollmann[42,43], the Chelp[44], and the ChelpG[45] algorithms. The algorithms differ significantly in the structure of the grid.

Correcting the torsion angle parameters is usually the last step in the force field parameterization. This correction is intended, among other things, to eliminate the weaknesses in treating the so-called 1-4 interactions, which occur due to the fixed partial charges.[32]

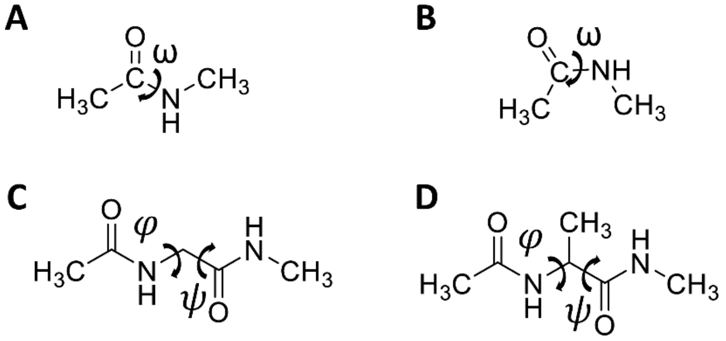

Figure 1 Typical model systems to represent the peptide backbone. A: *trans*-NMA, B: *cis*-NMA, C: Ac-Gly-NHMe, D: Ac-Ala-NHMe.

These three typical model systems used in the parameterization process of the protein backbone are NMA for the amide bond, respectively ω dihedral, (Ac-Gly-NHMe), and (Ac-Ala-NHMe) for the ψ and φ dihedrals. The ω dihedral of NMA occurs in two different conformations, *cis* and *trans* (see Figure 1), with a strong preference for the *trans* conformer of approximately 98 % occupation probability at room temperature.[46–55] These results were confirmed with many experiments, e.g., NMR spectroscopy or vibrational spectroscopy.[49,56,57] Nearly any kind of theoretical approach ranging from MD or Monte Carlo (MC) simulations to quantum-chemical (QC) calculations and AIMD simulations in both vacuum[58] and different solvents were applied to this problem. A long list of different models was utilized to represent the solvent, e.g., integral equation theory,[52] explicit water molecules, or continuum models.

The conformational landscapes of Ac-Gly-NHMe and Ac-Ala-NHMe are much more complex and were extensively studied theoretically and experimentally in the past.[59–83] For Ac-Ala-NHMe, four main structures regions are found in aqueous solution, which are the P_{II} (φ 75°, ψ 145°), C_5 (φ 180°, ψ 180°), $C_{7,\mathrm{eq}}$ and α_R (φ 60°, ψ -40°) conformations. These four regions are the main basins for every amino acid, except proline and glycine. For Ac-Gly-NHMe additional basins are observed, due to the missing side chain, leading to a mirror symmetry in the Ramachandran plane.[84] One of the most common experiments to determine the conformational preferences of the protein backbone is the measurement of the $J(H_N, H_\alpha)$ coupling constants,[63], which is directly correlated with the φ angle through the Karplus relation.[85] However, the NMR experiments have a problem determining the conformational preferences due to the long time scale of the measurements. Another disadvantage of the Karplus relation is that the parameterization is only valid for ambient pressure conditions. An efficient alternative is vibrational spectroscopy, where the amide III mode is a characteristic beacon to detect conformational preferences of the backbone and is directly correlated to the dihedral angle ψ.[86–88] In two recently published studies by Cormanich[66,67], the conformational preferences of Ac-Gly-NHMe and Ac-Ala-NHMe were investigated using the PCM continuum model. This solvation model prefers the α_R conformations as the main conformation in an aqueous solution. Grouleff and Jensen[68] performed an intensive study about the conformational preferences of Ac-Gly-NHMe and Ac-Ala-NHMe utilizing the OPLS force field combined with Tinker. The main conformer found for Ac-Gly-NHMe was the P_{II} conformer, and for Ac-Ala-NHMe, the $C_{7,eq}$ conformer. It can be seen that a wide range of different results was obtained. In this thesis, results obtained with the EC-RISM approach will be appended to the long list, which fits well with the theoretical and experimental results obtained in aqueous solutions.

1.3 The power of spectroscopic methods to resolve solvent-caused effects under HP conditions

Spectroscopic methods are among the indispensable instruments in the analysis of chemical substances. The idea behind spectroscopy is to induce a time-varying electric field through electrical or magnetic interactions, e.g., in the form of radiation, and thus to change the intra- or intermolecular interactions.[89] The energy change is mainly described in the context of the dipole approximation by the transition dipole moment, whereby a dipole moment must be present in the molecule under consideration, at least temporarily. The interaction with radiation leads to absorption, emission, and scattering. The frequency of the irradiated energy is largely responsible for which energy levels are excited.[89] For example, X-Ray radiation can stimulate internal nuclear electron transitions. *UV*/VIS spectroscopy is ideal for inducing valence electron transitions. To treat vibrational transitions, there are possibilities to use either infrared radiation (IR spectroscopy) or inducing Raman scattering with a laser pulse. With microwave radiation, it is possible to investigate rotational transitions. Radio frequencies can be used to identify magnetic resonance transitions, e.g., in NMR.[90] This work focuses on IR spectroscopy and magnetic resonance methods, which can also be extended to the high-pressure regime.

Infrared spectroscopy is a powerful tool to identify and investigate functional groups in different aggregation states. An infrared spectrometer samples the frequency range from 400 cm^{-1} to 4000 cm^{-1}. The most common way to record an IR spectrum today is using the so-called Fourier-transform infrared (FTIR) spectrometer.[91] The huge advantage of this specific spectrometer is the possibility of detecting all frequencies simultaneously instead of setting every frequency one after another. Besides chemical substance analysis, where especially the so-called fingerprint area in the IR spectrum is a major pillar[92], IR spectroscopy is used to identify secondary structural elements in proteins, especially the amide group of the protein backbone considered, where three characteristic amide bands occur. These are the amide-I band at 1700 cm^{-1}, the amide-II band at 1500 $cm^{-1,}$ and the amide-III band at 1400 cm^{-1}.[93]

FTIR spectroscopy can also be used to reveal the molecular responses of, e.g., piezolytes in water under extreme biophysical conditions.[94] Such extreme biophysical conditions are high hydrostatic pressures, temperature extremes, or high salt concentrations. Furthermore, FTIR spectroscopy can also be used to measure the conformational dynamics between different equilibrium states. This is of particular interest when studying protein unfolding processes. Especially for the latter, high pressure is a useful tool, as it allows the selective controlling of the unfolding processes.

NMR spectroscopy is undoubtedly one of the most important and powerful methods to analyze a compound's chemical structure. The idea behind NMR is to expose the molecules to a strong magnetic field, which leads to the orientation change of the nuclear spins, resulting in an energetic splitting (Zeeman splitting, more detailed background information follows in section 2.3).[95] If electromagnetic radiation is now irradiated onto the sample, a transition between the energy levels can be observed. An important property of NMR is that the energy difference between the levels depends strongly on the direct chemical environment. This effect is also known as shielding. Values for the shielding or, more specifically, for the chemical shift, since NMR measurements are always relative to a so-called standard, can be assigned to individual chemical groups. In addition, the different nuclear spins within a molecule are coupled, which allows a spatial assignment of signals via the so-called J-couplings.[95]

NMR spectroscopy was used to study the unfolding and changes in folding states of proteins due to pressurization.,[6,96,97] For example, the group of Kalbitzer showed that for the RAS interaction with the GTP analog GppNHp, a second state is detectable, and the population of this state is increased due to pressurization.[98] The pressure-dependent changes of the ^{1}H and ^{15}N chemical shifts of the amide bond are strong beacons to detect pressure-caused conformational changes. The ^{1}H amide shifts strongly correlate with the H-O bond length of the NH-OC intermolecular interaction. On the other hand, the ^{15}N shifts are very sensitive to changes in the backbone dihedral angles ψ and φ and even in changes of side-chain torsions χ.[99] These changes are responsible for the linear pressure-dependent changes in the chemical shifts (linear term in the quadratic term to describe the pressure-dependence of chemical shifts).[97]

While NMR focuses on the nuclear spin in molecules, the electron’s spin can also be perturbated by an applied magnetic field. Electron spin resonance spectroscopy (EPR) is associated with an unpaired electron, and the underlying principles are similar to NMR spectroscopy. In contrast to pulsed methods in NMR, most of the EPR spectra are recorded by applying a continuous wave (CW) experiment since it is not feasible to record a frequency spectrum most of the time. The frequency of the operating spectrometer is usually in the microwave range.[100] Spectroscopic parameters extracted from an EPR spectrum are the isotropic hyperfine coupling constants, the isotropic g-value, and the g- and A-tensors. [101] In the following section, more details regarding EPR-spectroscopy will be provided.

1.4 Nitroxide as EPR-spin labels

The so-called Frémy's salt, a long-living potassium nitrodisulfonate radical, was first synthesized in 1845[100] and can be described as the ancestor of modern nitroxides. In 1961 and 1964, the foundations for modern nitroxides were laid by synthesizing the molecules di-tert-butyl nitroxide[102] and 2,2,6,6-tetramethyl-1-piperidinyloxyl.[103] By introducing steric blocking groups to the NO moiety, stability of the radicals is achieved. An effective spin-Hamiltonian can be used to fit the EPR spectrum of nitroxide radicals and to evaluate the influences of the microenvironment.[100,101] A detailed description of the spin-Hamiltonian, especially for nitroxides, can be found in chapters 2.3 and 3.2.

Nitroxide can be relatively easily incorporated into biomolecules, like membranes or proteins. For example, fatty acids can be functionalized by attaching a nitroxide to the hydrocarbon chain. These systems can then be used as molecular probes to identify different membrane properties, such as hydrophobicity, oxygen solubility, diffusion rates, membrane order and fluidity, and bilayer penetration of water molecules.[104] Normally, one percent of the fatty acids in a membrane bilayer are modified. In this manner, avoiding spin-spin interactions is guaranteed. [104]

In the year 1965, for the first time, nitroxides were introduced as a spin-label in proteins.[105] The groundbreaking work to make nitroxide the spin probe of choice was performed by Hubbell in the late 1980s and the 1990s.[106–108] The stability and the small molecular weight made nitroxides attractive. Additionally, reactive groups can be easily incorporated into nitroxides, which bind to, e.g., cysteines. The most commonly used spin-label example is the 1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl (MTSL). With its highly reactive methanethiosulfonate group, binding to external cysteines can be achieved by nearly 100%. Since MTSL has five rotatable groups, it is very adaptable to most protein sites.[100] EPR spectroscopy can then be used to study the topography of the polypeptide chain[107,109] or electrostatic potentials at any surface site.[110] By applying these analyses, secondary structures and their orientation in the protein can be identified[108] Additionally, the distance between two spin-labels can be measured.[111] EPR can also be used to resolve conformational equilibria of spin-label proteins. For example, McCoy made pressure-dependent EPR measurements for a spin-labeled mutant of T4 lysozyme.[112] Here, it was demonstrated that with EPR, the equilibrium between spectrally observable states can be measured and that the origin of the states can be either attributed to a conformational or a rotameric exchange. Additionally, by applying EPR under pressure, the relative partial molar volume and isothermal compressibility between conformational substates can be determined.

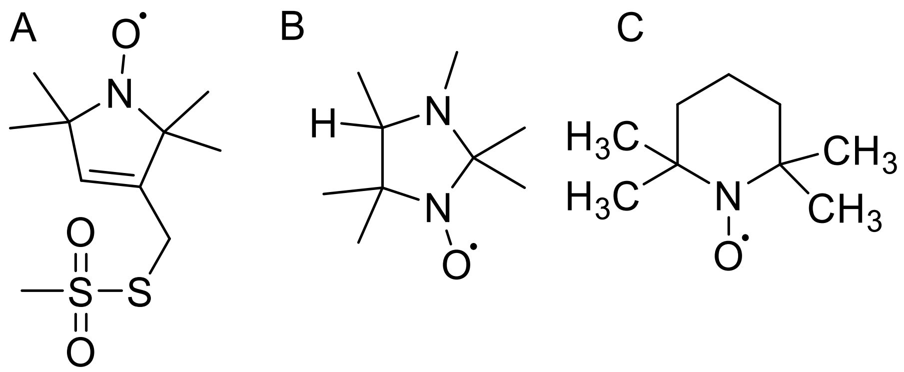

Figure 2 Structures of the nitroxides MTSL (A), HMI (B), and TEMPO (C).

As mentioned above, an important feature of nitroxides is that the spectroscopic parameters are very sensitive to the microenvironment. Especially the isotropic hyperfine coupling constant and the A_{zz} component increase in polar environments. On the other hand, the g_{xx} component of the g-tensor decreases in more polar environments.[113] The A_{zz} and g_{xx} components can be used to investigate the local polarity and H-bond network. One nitroxide especially suitable for the microenvironment is the 2,2,3,4,5,5-hexamethylidaolidin-1-oxyl (HMI). This spin label is used to analyze surface potentials, local polarity, and pK_a-values.[114–116] For example, a change of 2.5 MHz is observed for the transition from the low pH range to the high pH range (see chapter 4.2).

Some points need to be considered to model EPR observables accurately from a theoretical point of view. On the one hand, a suitable statistical ensemble must be generated, whereby the solvation must be considered. On the other hand, high-level QC calculations must be performed to get satisfying results compared to the experimental data. Fortunately, these two tasks can be separated, e.g., hybrid models like QM/MM can be used.[117] For the ensemble generation, MD simulations could be utilized; however, these methods have the drawback that specific force fields need to be parameterized to accurately model the intra- and intermolecular interactions.[118] A more general and elegant method but more time-consuming is the Car-Parrinello or AIMD simulation.[119–121] After generating a sufficiently large ensemble of solvated structures, QC calculations can be utilized to determine the final parameters. For example, if the target parameter is the isotropic hyperfine coupling constant, theoretical methods and basis sets need to be utilized to represent the spin densities in the vicinity of the target nuclei accurately. Thus, the basis set needs to be flexible in the core region and should provide an accurate description of core polarization.[122] The performance of Hartree-Fock methods to calculate the latter is poor.[123,124] Perturbation theory (e.g., second-order MP2 theory) cannot solve this problem[101], and coupled-cluster approaches like CCSD are too time-consuming. Most of the publications focused, therefore, on different DFT methods.[125–128] Recently, the so-called domain-based pair natural orbital coupled-cluster theory (DLPNO-CCSD(T))[129–131] was extended for open-shell systems and showed promising results in calculating EPR obervables.[127,132,133]

2 Theory

2.1 Solvation models

2.1.1 Explicit and implicit solvation models

A majority of the relevant biochemical processes occur in solution. Thus, models are needed to describe the solvation process within theoretical calculations to represent the experimental observations reliably. The most obvious decision would be to represent the solvent explicitly, i.e., include each solvent atom-coordinate in the calculation. However, this is problematic because the calculation time significantly increases, limiting oneself to a few solvent molecules, which can not resolve the whole picture. Due to this disadvantage, only these explicit solvent molecules directly involved in intermolecular interactions, e.g., hydrogen bonds, are reasonable in theoretical calculations. The so-called implicit solvent models represent a significantly more cost-sensitive method in terms of computing time. In these models, the solvent is not atomically resolved; instead, it is characterized by its macroscopic properties. With this representation, it is unnecessary to calculate the interactions of the individual solvent molecules with each other, leading to massive time-saving. However, this results in the loss of any atomic information, leading to possible systematic errors.

Among the best-known implicit solvent models are the so-called continuum models; examples are PCM[134] (polarizable continuum model), COSMO[135,136] (conductor-like screening model), and CPCM (conductor-like polarizable continuum model).[137] The idea behind continuum models is to represent the solvent as a dielectric continuum in which the solute particle is embedded within a cavity. The solvent is solely characterized by its dielectric constant. The basis of this formalism is the classical Poisson equation in the following form:[138]

$$-\vec{\nabla} \times \left[\varepsilon(\boldsymbol{r})\vec{\nabla}V(\boldsymbol{r})\right] = 4\pi\rho_M(\boldsymbol{r}), \tag{2}$$

where $\varepsilon(\boldsymbol{r})$ is the position-dependent $(\boldsymbol{r})$ permittivity and $\rho_M(\boldsymbol{r})$ is the solute (M) charge distribution. It is assumed that the solute charge distribution is embedded in a molecular cavity C, resulting in the following positional dependence for the permittivity:

$$\varepsilon(\boldsymbol{r}) = \begin{cases} 1 & \boldsymbol{r} \in C \\ \varepsilon & \boldsymbol{r} \notin C' \end{cases} \tag{3}$$

where ε is the dielectric constant of the pure solvent. $V(\boldsymbol{r})$ is the potential that describes the solute potential (V_M) plus the contribution of the apparent charge distribution $(\sigma(\boldsymbol{s})$, on the surface $\boldsymbol{s}$ of the cavity Γ. The corresponding equations are:

$$V(\boldsymbol{r}) = V_M(\boldsymbol{r}) + V_\sigma(\boldsymbol{r}), \tag{4}$$

and

$$V_\sigma(\boldsymbol{r}) = \int_\Gamma \frac{\sigma(\boldsymbol{s})}{|\boldsymbol{r}-\boldsymbol{s}|} \mathrm{d}^2 s. \tag{5}$$

A general problem of continuum models is to find a suitable description of the apparent surface charges (more details below). If an appropriate surface charge is determined, the electrostatic component of the solvation-free energy G^{el} is given by:

$$G^{\mathrm{el}} = \frac{1}{2}\int_\Gamma \sigma(\boldsymbol{s}) \left[\iint_V \frac{\rho_M(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{s}|} d^3 r\right] d^2 s. \tag{6}$$

The different PCM formalisms differ mainly in the design of the surface charges. The historical first formulation is the DPCM, whereby the first letter D stands for dielectric and is added to PCM to distinguish it from the alternative successors.[134] A well-known alternative is the COSMO[135,136] model mentioned above. In contrast to the original formulation, where the permittivity was described by the dielectric constant of the solvent, here, the continuum is treated like a conductor, and its ε value is set to infinity. A minor modification of the COSMO model leads to the CPCM.[137] A common alternative is the so-called integral equation formulation of PCM, the IEFPCM.[139] In this thesis, either IEFPCM or CPCM was used. A significant advantage of continuum models is that they can be combined with various QC methods.

In 2012 an extension of the PCM model for high-pressure environments was developed.[21,22,140] Pressure is introduced to the XP-PCM through an increased Pauli repulsion between the molecule and the medium. Generally, the density and permittivity increase upon pressurization. For modeling the effect of pressure on the solute, the boundary of the cavity is shrunk. The final pressure is obtained by differentiating the free energy with respect to the cavity volume in the following manner:

$$p = -\left(\frac{\partial G_{e-r}}{\partial V_c}\right). \tag{7}$$

Here, G_{e-r} is the free energy of the whole environment system, and it is defined as:

$$G_{e-r} = \langle \Psi | \hat{H}_0 + \hat{V}_e + \hat{V}_r | \Psi \rangle + \hat{V}_{nn}. \tag{8}$$

The reference state is defined as a hypothetical state composed of non-interacting electrons and nuclei of the solute and by the unperturbed medium at the chosen thermodynamic conditions. So far, XP-PCM has been used to make predictions of reaction equilibria and vibrational spectroscopic data in non-aqueous environments.

2.1.2 Statistical-mechanical solvation models

In addition to explicit and implicit solvent models, statistical-mechanical solvent models are popular for representing the solvent. In contrast to the implicit solvent models, the statistical solvent models retain information about the structure and granularity of the solvent, which allows, for example, to model the effect of hydrogen bonds. The integral equation theory is an important example in the cosmos of statistical solvent models. It aims to calculate the statistical distribution of a solvent around a solved particle. The pillar on which this theory rests is the so-called classical density functional theory (cDFT). Analogous to quantum mechanical DFT, where the Hohenberg-Kohn[141] theorem states that a unique electron density can only describe the minimum energy of a system, the free energy $F[\rho]$ of a liquid is minimized by exactly one unique particle density. In the context of this work, a few central aspects should be summarized, and a rough overview of the subject should be given. Both Evans[142] and Hansen[143] have written detailed reviews of the cDFT in the past.

One of the central quantities within the classical DFT is the grand potential $\Omega[\rho]$, which can be defined in the grand-canonical ensemble as

$$\Omega[\rho] = \int d\boldsymbol{r}\, \rho(\boldsymbol{r}) V_{\text{ext}}(\boldsymbol{r}) + F[\rho] - \mu \int d\boldsymbol{r}\, \rho(\boldsymbol{r}). \tag{9}$$

The grand potential is a functional of the particle density $\rho(\boldsymbol{r})$. $V_{\text{ext}}(\boldsymbol{r})$ is an arbitrary external potential and μ is the chemical potential.

Mermin extended the previously mentioned Hohenberg-Kohn theorem to an electron gas at finite temperature and showed that by applying the principle *reductio ad absurdum,* a specific electron density minimizes the grand potential.[144] Transferred to cDFT, it follows that the equilibrium particle density minimizes the grand potential. This fact can be shown by the functional derivation of the grand potential according to the particle density as

$$\left.\frac{\delta\Omega[\rho]}{\delta\rho(\boldsymbol{r})}\right|_{\rho_0} = 0. \tag{10}$$

From equations (9) and (10) follows

$$0 = V_{\text{ext}}(\boldsymbol{r}) + \frac{\delta F[\rho]}{\delta\rho} - \mu, \tag{11}$$

and with the definition of the intrinsic chemical potential

$$\mu_{\text{in}}[\rho;\boldsymbol{r}] \equiv \frac{\delta F[\rho]}{\delta\rho(\boldsymbol{r})}, \tag{12}$$

equation (11) becomes

$$\mu = V_{\text{ext}}(\boldsymbol{r}) + \mu_{\text{in}}[\rho_0;\boldsymbol{r}]. \tag{13}$$

An elementary part in describing a fluid system are the mutual interactions of the particles. The Free Energy functional depends on these interactions and can be divided into an ideal part $F^{\text{id}}[\rho]$ and a functional $\Phi[\rho]$, which represents the interaction part of $F[\rho]$:

$$F[\rho] = F^{\text{id}}[\rho] + \Phi[\rho]. \tag{14}$$

The ideal part can be derived from the partition function of the ideal gas via:

$$F^{\text{id}}[\rho] = \beta^{-1} \int d\boldsymbol{r}\, \rho(\boldsymbol{r}) \big(\ln\big(\lambda^3 \rho(\boldsymbol{r})\big) - 1\big), \tag{15}$$

with $\lambda = (h^2\beta/2m\pi)^{1/2}$ where h is the Planck constant, β is the inverse temperature, and m is the mass. The intrinsic chemical potential can be written as:

$$\beta\mu_{\text{in}}[\rho;\boldsymbol{r}] = \beta\frac{\delta F^{\text{id}}[\rho]}{\delta\rho(\boldsymbol{r})} + \beta\frac{\delta\Phi[\rho]}{\delta\rho(\boldsymbol{r})}. \tag{16}$$

Where the last term will be defined as:

$$\beta\frac{\delta\Phi[\rho]}{\delta\rho(\boldsymbol{r})} \equiv c^{(1)}[\rho,\boldsymbol{r}]. \tag{17}$$

$c^{(1)}[\rho,\boldsymbol{r}]$ is the so-called one-particle direct correlation function and determines the equilibrium density. The one-particle direct-correlation functional is the first in a series of higher-order correlation functionals that can be obtained by further differentiation:

$$c^{(2)}[\rho;\boldsymbol{r}_1,\boldsymbol{r}_2] = \frac{\delta c^{(1)}[\rho;\boldsymbol{r}_1]}{\delta\rho(\boldsymbol{r}_2)} = \frac{\beta\delta^2\Phi[\rho]}{\delta\rho(\boldsymbol{r}_2)\delta\rho(\boldsymbol{r}_1)} = c^{(2)}[\rho;\boldsymbol{r}_2,\boldsymbol{r}_1]. \tag{18}$$

And another differentiation leads to:

$$c^{(3)}[\rho;\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3] = \frac{\delta c^{(1)}[\rho;\boldsymbol{r}_1]}{\delta\rho(\boldsymbol{r}_2)\delta\rho(\boldsymbol{r}_3)} = c^{(3)}[\rho;\boldsymbol{r}_2,\boldsymbol{r}_1,\boldsymbol{r}_3], \tag{19}$$

etc. Now the following definition is introduced:

$$u(\boldsymbol{r}) = \mu - V_{\text{ext}}(\boldsymbol{r}), \tag{20}$$

and the grand potential is written as follows:

$$\Omega[\rho] = -\int d\boldsymbol{r}\,\rho(\boldsymbol{r})u(\boldsymbol{r}) + F[\rho], \tag{21}$$

and the following differentiation w.r.t. the intrinsic chemical potential yields:

$$\frac{\delta\Omega[\rho]}{\delta u(\boldsymbol{r})} = -\rho(\boldsymbol{r}) + \int d\boldsymbol{r}'\frac{\delta\rho(\boldsymbol{r}')}{\delta u(\boldsymbol{r})}\left(\frac{\delta F[\rho]}{\delta\rho(\boldsymbol{r}')} - u(\boldsymbol{r})\right). \tag{22}$$

If $\rho(\boldsymbol{r}) = \rho_0(\boldsymbol{r})$ and with equations (12) and, it follows:

$$\frac{\delta\Omega[\rho_0]}{\delta u(\boldsymbol{r})} = -\rho_0(\boldsymbol{r}). \tag{23}$$

In this manner, the grand potential is also a generating functional for the equilibrium density with respect to the variable $u(\boldsymbol{r})$. It becomes interesting when $\Omega[\rho_0]$ is differentiated once again w.r.t. $u(\boldsymbol{r})$:

$$G(\boldsymbol{r}_1,\boldsymbol{r}_2) \equiv \beta^{-1}\frac{\delta\rho_0(\boldsymbol{r}_1)}{\delta u(\boldsymbol{r}_2)} = \frac{\delta\Omega[\rho_0]}{\delta u(\boldsymbol{r}_1)\delta u(\boldsymbol{r}_2)}. \tag{24}$$

Here $G(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is the density fluctuation function or density-density correlation function. $G(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is closely related to the pairwise distribution function $\rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ via:

$$G(\boldsymbol{r}_1,\boldsymbol{r}_2) = \rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2) + \rho_0(\boldsymbol{r}_1)\delta(\boldsymbol{r}_1-\boldsymbol{r}_2) - \rho_0(\boldsymbol{r}_1)\rho_0(\boldsymbol{r}_2), \tag{25}$$

with the relation:

$$\rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \rho_0^2 g(r), \tag{26}$$

the pairwise distribution function is directly correlated to the well-known radial distribution function $g(r)$. Another important property is the inverse of the density-density-correlation function since it is linked to the functional derivative of the intrinsic chemical potential via:

$$G^{-1}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \frac{\beta\delta u(\boldsymbol{r}_1)}{\delta\rho_0(\boldsymbol{r}_2)}, \tag{27}$$

with the general definition of the functional inverse:

$$\int d\boldsymbol{r}_3\, G^{-1}(\boldsymbol{r}_1,\boldsymbol{r}_3)G(\boldsymbol{r}_3,\boldsymbol{r}_2) = \delta(\boldsymbol{r}_1-\boldsymbol{r}_2). \tag{28}$$

At this point, another relationship must be introduced in order to obtain the most fundamental equation of static solvent models, the Ornstein-Zernike equation. From equations (11), (16), (17), and (13), the following expression for the equilibrium density can be extracted:

$$\rho_0(\boldsymbol{r}) = z\exp\left[-\beta V_{\mathrm{ext}}(\boldsymbol{r}) + c[\rho_0;\boldsymbol{r}]\right]. \tag{29}$$

Here z is the so-called fugacity, which is directly proportional to the thermal wavelength λ:

$$z = \lambda^{-3}\exp[-\beta\mu]. \tag{30}$$

Equation (29) can now be re-written in the following form:

$$c^{(1)}[\rho;\mathbf{r}_1] = \ln\left(\lambda^3\rho_0(\mathbf{r}_1)\right) - \beta u(\mathbf{r}_1). \tag{31}$$

If equation (31) is differentiated with respect to $\rho_0(\boldsymbol{r}_2)$, it follows:

$$c^{(2)}[\rho_0;\boldsymbol{r}_1,\boldsymbol{r}_2] = c^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \frac{\delta(\boldsymbol{r}_1-\boldsymbol{r}_2)}{\rho_0(\boldsymbol{r}_1)} - \frac{\beta\delta u(\boldsymbol{r}_1)}{\delta\rho_0(\boldsymbol{r}_2)}. \tag{32}$$

After several substitutions and transformations, one can now extract the relationship that is called the Ornstein-Zernike (OZ)[145] equation for uniform fluids with the density ρ_0:

$$g(r) - 1 = c^2(r) + \rho_0\int d\boldsymbol{r}'^{(g(r')-1)}\, c^2(|\boldsymbol{r}-\boldsymbol{r}'|), \tag{33}$$

The discovery of the OZ equation was somewhat serendipitous.[145] The key feature of the OZ equation is the connection of the direct correlation function with the total correlation or pair distribution function. The OZ equation cannot be solved in the above form alone but only iteratively with the help of another relation, which links the direct correlation function with the pair-distributions function. To obtain this relation, the functional $\Phi[\rho]$ is expanded via a Taylor series w.r.t. the density of the homogenous reference system:

$$\begin{aligned}\Phi[\rho] = \Phi[\rho_0] + \int \left.\frac{\delta\Phi[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r})}\right|_{\rho=\rho_0} \left(\rho(\boldsymbol{r}) - \rho_0(\boldsymbol{r})\right)d\boldsymbol{r} \\ + \frac{1}{2}\iint \left.\frac{\delta\Phi^2[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r})^2}\right|_{\rho=\rho_0} \left(\rho(\boldsymbol{r}) - \rho_0(\boldsymbol{r})\right)\left(\rho(\boldsymbol{r}') - \rho_0(\boldsymbol{r}')\right)d\boldsymbol{r}d\boldsymbol{r}' \\ + O(\rho(\boldsymbol{r})^3).\end{aligned} \tag{34}$$

With equations (17) and (18), the following relations can be obtained:

$$\begin{aligned}\int \left.\frac{\delta\Phi[\rho(\boldsymbol{r})]}{\delta\rho(\boldsymbol{r})}\right|_{\rho=\rho_0} \left(\rho(\boldsymbol{r}) - \rho_0(\boldsymbol{r})\right)d\boldsymbol{r} = \int -\beta^{-1}c_0^{(1)}(\boldsymbol{r})\Delta\rho(\boldsymbol{r})d\boldsymbol{r} \\ = \int \mu^{\mathrm{ex}}\Delta\rho(\boldsymbol{r})d\boldsymbol{r},\end{aligned} \tag{35}$$

and

$$\iint \frac{\delta \Phi^2[\rho(\boldsymbol{r})]}{\delta \rho(\boldsymbol{r})^2}\bigg|_{\rho=\rho_0} (\rho(\boldsymbol{r})-\rho_0(\boldsymbol{r}))(\rho(\boldsymbol{r}')-\rho_0(\boldsymbol{r}'))d\boldsymbol{r}d\boldsymbol{r}' \\ = -\beta \iint c_0^{(2)}(\boldsymbol{r},\boldsymbol{r}')\Delta\rho(\boldsymbol{r})\Delta\rho(\boldsymbol{r}')d\boldsymbol{r}d\boldsymbol{r}'. \tag{36}$$

For the excess part of the free energy functional, it follows:

$$\Phi[\rho] = \Phi[\rho_0] - \int \mu^{\mathrm{ex}} \Delta\rho(\boldsymbol{r})d\boldsymbol{r} - \frac{\beta}{2}\iint c_0^{(2)}(\boldsymbol{r},\boldsymbol{r}')\Delta\rho(\boldsymbol{r})\Delta\rho(\boldsymbol{r}')d\boldsymbol{r}d\boldsymbol{r}'. \tag{37}$$

This relation is correct up to the second-order if equation (37) is now combined with equation (17) and inserted into the grand-potential definition in equation (21), the expression for the density that minimizes the Helmholtz free energy is given by:

$$\rho(\boldsymbol{r}) = \rho_0 \exp\left(-\beta V_{\mathrm{ext}}(\boldsymbol{r}) + \int \rho(\boldsymbol{r}')c_0^{(2)}(|\boldsymbol{r}-\boldsymbol{r}'|)\right). \tag{38}$$

In the following step, the so-called Percus idea can be used. The assumption is that the external system's source is an extracted particle from the uniform homogeneous fluid. The position of this particle is fixed in space, and the external potential is then represented as the sum of pair potentials between the reference particle and all other particles. Based on these assumptions, all quantities in equation (38) become radially symmetric, and the density can be represented as a pair distribution function in the following way:

$$g(\boldsymbol{r}) = \exp\left(-\beta u(\boldsymbol{r}) + \rho \int c(|\boldsymbol{r}-\boldsymbol{r}'|)h(\boldsymbol{r}')d\boldsymbol{r}'\right) \\ = \exp(-\beta u(\boldsymbol{r}) + h(\boldsymbol{r}) - c(\boldsymbol{r})), \tag{39}$$

where $u(\boldsymbol{r})$ is the pair potential. The latter equation is also called the closure relation.

2.1.3 Reference interaction site model (RISM)

The previous chapter derived the system of equations consisting of the OZ equation and the closure relation. In this form, the equations system only applies to simple spherical fluids. In the year 1972, Blum et al.[146] extended the OZ equation to the molecular Ornstein-Zernike equation (MOZ) in the following form:

$$h^{(2)}(\boldsymbol{r},\boldsymbol{\Omega},\boldsymbol{r}',\boldsymbol{\Omega}') \\ = c^{(2)}(\boldsymbol{r},\boldsymbol{\Omega},\boldsymbol{r}',\boldsymbol{\Omega}') \\ + \rho \iint c^{(2)}(\boldsymbol{r}'',\boldsymbol{\Omega}'',\boldsymbol{r}',\boldsymbol{\Omega}')\, h^{(2)}(\boldsymbol{r}'',\boldsymbol{\Omega}'',\boldsymbol{r}',\boldsymbol{\Omega}')d\boldsymbol{r}d\boldsymbol{\Omega}. \tag{40}$$

In the MOZ, not only the distances between two particles are included, but also the angle dependence is taken into account, which is represented by the well-known Euler angles $\boldsymbol{\Omega}$. Equation (40) is the exact analytical expression for the MOZ but cannot be solved due to the high dimensionality of this equation. To avoid this problem, Chandler and Andersen published a series of papers[147,148] in 1972 in which the dimensionality of the MOZ was significantly reduced. For this purpose, a so-called site-site MOZ was derived, in which the molecules are described via spherical rigid interaction sites. To model the structure of the molecule, an additional correlation function is introduced, the so-called intramolecular correlation function ω in the form:[149]

$$\omega_{\alpha\gamma}(r) = \frac{\delta\left(r - l_{\alpha\gamma}\right)}{4\pi l_{\alpha\gamma}^{2}}. \tag{41}$$

Eq. (41) contains the interaction sites α and γ, and the intramolecular distance l between these two sites, δ is the Dirac delta function. The direct correlation function can be described by a summation of the pairwise site-site interactions and can be represented as follows:

$$c_{\alpha\gamma}(r) = \sum_{a} \rho_{\alpha} \sum_{\alpha} \sum_{\gamma} c_{\alpha\gamma}\left(\left|r_{\alpha} - r_{\gamma}\right|\right). \tag{42}$$

For example, a water molecule could be taken as a reference here, where three sites represent the three atoms. However, a two-site model can also be developed due to symmetry considerations. In today's literature, the model developed by Chandler and Anderson is called the one-dimensional reference interaction site model (1D RISM). The associated system of equations can be described in the following matrix form:[150]

$$\boldsymbol{h} = \boldsymbol{\omega} * \boldsymbol{c} * \boldsymbol{\omega} + \boldsymbol{\rho}\boldsymbol{\omega} * \boldsymbol{c} * \boldsymbol{h}. \tag{43}$$

Equation (43) can be separated into three parts, a solvent-solvent (*vv*) part, a solvent-solute part (*uv*), and a solute-solute part (*uu*). However, these equations cannot be solved simultaneously but must be solved one after another. The basis is the *vv*-calculation in the form:

$$\boldsymbol{h}^{vv} = \boldsymbol{\omega}^{v} * \boldsymbol{c}^{vv} * \boldsymbol{\omega}^{v} + \boldsymbol{\omega}^{v} * \boldsymbol{c}^{vv} * \boldsymbol{\rho}^{v}\boldsymbol{h}^{vv}. \tag{44}$$

However, equation (44), like the previously described OZ equation, cannot be solved alone, so the closure relationship is required here again. For a site-site model, the closure takes the following form:

$$h_{\alpha\gamma}(r) + 1 = \exp\left(-\beta u_{\alpha\gamma}(r) + h_{\alpha\gamma}(r) - c_{\alpha\gamma}(r) + B_{\alpha\gamma}(r)\right), \tag{45}$$

where u is the pair potential and $B_{\alpha\gamma}$ is the so-called bridge function. Unfortunately, since the bridge function is unknown and ill-defined in 1D/3D RISM, it is usually set to zero, and the hypernetted chain closure (HNC) is obtained. Equations (44) and (45) can now be solved iteratively. Before the *uv* equation is solved, another practical quantity is introduced, the so-called solvent susceptibility

$$\boldsymbol{\chi} = \boldsymbol{\rho}^{v}\boldsymbol{\omega}^{v} + \boldsymbol{\rho}^{v}\boldsymbol{h}^{vv}\boldsymbol{\rho}^{v}, \tag{46}$$

which reflects the density response of the solvent. Since in equation (46), the density of the pure solvent is considered, $\boldsymbol{\chi}$ is the perfect way to integrate information regarding pressure or temperature. For the solvent-solute (*uv*) equation, it follows:

$$\boldsymbol{h}^{uv} = \boldsymbol{\omega}^{u} * \boldsymbol{c}^{uv} * (\boldsymbol{\rho}^{v})^{-1}\boldsymbol{\chi}. \tag{47}$$

A simple way to generate pressure-dependent $\boldsymbol{\chi}$ is to use experimental data, specifically the solvent's bulk density and dielectric constants (besides, one also needs the molecular geometry and accurate parameters for the pairwise potential, which will be discussed later). However, this method is limited to applying the HNC closure or other simple closures, which neglect the bridge function. To still use the contribution from the bridge function, it is possible to constrain the radial distribution function from the 1D RISM calculation to the values obtained from a radial distribution function calculated using a MD simulation. Roux and co-workers developed the following expression for the bridge-function:[151]

$$B_{\alpha\gamma}(r) = f(r)\left[\ln\left(g_{\alpha\gamma}^{MD}(r)\right) + \beta u_{\alpha\gamma}(r) - h_{\alpha\gamma}(r) + c_{\alpha\gamma}(r)\right]. \tag{48}$$

where $f(r)$ is a cubic switching function, which varies between 1 and 0. The application of the switching function is necessary because the MD simulations are performed in a simulation box of finite size. Beyond the range of the simulation box, the correlation functions are extrapolated using the known HNC closure.

A considerable advantage of the RISM methodology is the possibility of getting a closed form for the excess chemical potential. Within the HNC approximation, the corresponding expression for the excess chemical potential has the following form:[152,153]

$$\Delta\mu^{\text{ex}} = -\frac{\rho}{\beta}\sum_{\alpha,\gamma}\int d\boldsymbol{r}\left(\frac{1}{2}{h_{\alpha\gamma}}^{2} - c_{\alpha\gamma} + \frac{1}{2}h_{\alpha\gamma}c_{\alpha\gamma}\right). \tag{49}$$

A disadvantage of the RISM formalism in this form is that the RISM calculations are performed for rigid geometries at a time, and every small change in geometry results in a new iterative cycle. To elegantly circumvent this problem, Schmidt and Kast[154] developed the following formalism, based on previous works by Sato et al.[155] The premise was that a slight change in geometry has only a negligible influence on the direct correlation function $c_{\alpha\gamma}$. Consequently, it is possible to calculate the difference in excess chemical potentials between two geometries $\boldsymbol{\Gamma}$and $\boldsymbol{\Gamma}'$ as follows:

$$\begin{aligned}&\Delta\mu(\boldsymbol{\Gamma}') - \Delta\mu(\boldsymbol{\Gamma}) \approx \\ &-\frac{1}{2(2\pi)^{3}\beta}\sum_{\alpha,\alpha',\gamma,\gamma'}\int d\boldsymbol{k}\hat{c}_{\alpha,\gamma}\,\hat{c}_{\alpha',\gamma'}\chi_{\gamma\gamma'}\left[\widehat{\omega}_{\alpha\alpha'(\boldsymbol{\Gamma}')} - \widehat{\omega}_{\alpha\alpha'(\boldsymbol{\Gamma})}\right],\end{aligned} \tag{50}$$

where the carets denote Fourier transforms of the corresponding correlation functions.

During the 1990s, Beglov and Roux[156], and Kovalenko and Hirata[157], developed the three-dimensional extension of RISM formalism, the three-dimensional reference interaction site model (3D RISM). The resulting 3D RISM equation has the following form:

$$h_{\gamma}(\boldsymbol{r}) = \sum_{\gamma'}\left(\rho_{\gamma}\right)^{-1}\int c_{\gamma'}(\boldsymbol{r} - \boldsymbol{r}')\chi_{\gamma\gamma'}(|\boldsymbol{r}'|)\,d\boldsymbol{r}', \tag{51}$$

where the quantity $\chi_{\gamma\gamma'}(|\boldsymbol{r}'|)$ is a precomputed solvent susceptibility function, which can be calculated via eq.(46). The 3D RISM equation calculates the solvent's anisotropic total correlation function under the influence of an infinitely diluted solute. The corresponding 3D HNC closure is given by:

$$h_\gamma(\boldsymbol{r}) = \exp\left(-\beta u_\gamma(\boldsymbol{r}) + h_\gamma(\boldsymbol{r}) - c_\gamma(\boldsymbol{r})\right) - 1. \tag{52}$$

As a result of the frequent numerical instability of the HNC-closure,[158] Kast and Kloss developed a partial series expansion of the HNC closure, the so-called PSE-*n* closures,[159]

$$h_\gamma(\mathbf{r}) = \begin{cases} \dfrac{\sum_{i=0}^{k}\left(-\beta u_\gamma(\mathbf{r}) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})\right)^i}{i!} & \Leftrightarrow -\beta u_\gamma(\mathbf{r}) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r}) > 0 \\ \exp\left[-\beta u_\gamma(\mathbf{r}) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})\right] - 1 & \Leftrightarrow -\beta u_\gamma(\mathbf{r}) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r}) \leq 0 \end{cases}, \tag{53}$$

the PSE-1 closure is also known as the Kovalenko-Hirata-closure (KH).[160] In equation (53), *u* describes the interaction component between the solute and the solvate site γ. The corresponding potential function has the following form:

$$u_\gamma(\boldsymbol{r}) = \sum_\alpha \left(\frac{1}{4\pi\varepsilon_0}\frac{q_\alpha q_\gamma}{|\boldsymbol{r} - \boldsymbol{R}_\alpha|}\right) + 4\varepsilon_{\alpha\gamma}\left(\frac{\sigma_{\alpha\gamma}^{12}}{|\boldsymbol{r} - \boldsymbol{R}_\alpha|^{12}} - \frac{\sigma_{\alpha\gamma}^{6}}{|\boldsymbol{r} - \boldsymbol{R}_\alpha|^{6}}\right) \tag{54}$$

The first term describes the electrostatic Coulomb interactions, and the second describes the dispersive Lennard-Jones interactions between the solute sites α and the solvent site γ. In the latter, the mixed Lennard-Jones parameters σ and ε occur, where σ describes the contact distance between two atoms and ε is a measure of the depth of the applied potential (in most cases, the Lorentz-Berthelot[161] mixing rules are applied). Usually, these parameters are taken from known force fields like the GAFF[162,163] force field, as was done in this thesis. An alternative option is to use 3D RISM to perform a self-consistent optimization of the parameters to represent thermodynamic observables correctly. During the evaluation of the electrostatic interactions, the same truncation errors occur that can be observed during the MD simulation in a finite box. To avoid truncation errors, the idea of Ewald summation is used, where the electrostatic interactions are divided into two parts,

$$u_\gamma^{\text{elec}}(\boldsymbol{r}) = u_\gamma^{\text{elec,short}}(\boldsymbol{r}) + u_\gamma^{\text{elec,long}}(\boldsymbol{r}), \tag{55}$$

a short-range part

$$u_\gamma^{\text{elec,short}}(\boldsymbol{r}) = \sum_\alpha \frac{1}{4\pi\varepsilon_0}\frac{q_\alpha q_\gamma}{|\boldsymbol{r} - \boldsymbol{R}_\alpha|}\operatorname{erfc}(\kappa|\boldsymbol{r} - \boldsymbol{R}_\alpha|) \tag{56}$$

and a long-range part

$$u_\gamma^{\text{elec,long}}(\boldsymbol{r}) = \sum_\alpha \frac{1}{4\pi\varepsilon_0}\frac{q_\alpha q_\gamma}{|\boldsymbol{r} - \boldsymbol{R}_\alpha|}\operatorname{erf}(\kappa|\boldsymbol{r} - \boldsymbol{R}_\alpha|). \tag{57}$$

After a Fourier transformation, the latter term can be calculated elegantly in the reciprocal space, where it becomes short-range using a proper smearing factor κ. As in the one-dimensional case, the excess chemical potential can also be calculated with the help of 3D RISM. However, this depends on the selected closure; for the HNC-closure, the following equation follows:

$$\mu_{\mathrm{HNC}}^{\mathrm{ex}} = \frac{1}{\beta}\sum_{\gamma}\rho_{\gamma}\int\left(\frac{1}{2}h_{\gamma}^{2}(\boldsymbol{r}) - c_{\gamma}(\boldsymbol{r}) - \frac{1}{2}h_{\gamma}(\boldsymbol{r})c_{\gamma}(\boldsymbol{r})\right)d\boldsymbol{r}. \tag{58}$$

And for the PSE-closures, the following expression is used:

$$\mu_{\mathrm{PSE}}^{\mathrm{ex}} = \mu_{\mathrm{HNC}}^{\mathrm{ex}} - \frac{1}{\beta}\sum_{\gamma}\rho_{\gamma}\int\frac{\Theta\left(h_{\gamma}(\boldsymbol{r})\right)\left(h_{\gamma}(\boldsymbol{r}) - \beta u_{\gamma}(\boldsymbol{r}) - c_{\gamma}(\boldsymbol{r})\right)^{k+1}}{(k+1)!}d\boldsymbol{r}. \tag{59}$$

2.1.4 The fusion between quantum chemistry and RISM –EC-RISM, RISM-SCF, 3D RISM SCF

In 1D RISM and 3D RISM, the pair potential appears, determining the interactions between solute and solvent or solvent and solvent. While the Lennard-Jones parameters mainly originate from known force fields, for the electrostatic interactions, the partial charges or electrostatic potentials can be calculated by QC calculations or polarizable force fields, respectively.

Ten-No *et al.* made the first attempts to couple the RISM theory with a QC description of the solute in the early 1990s.[155,164,165] They developed the so-called RISM-SCF procedure, where the idea is that the solute partial charges obtained from an SCF calculation are plugged into a 1D RISM calculation to obtain a solvent charge distribution around the solute. This solvent charge distribution is then added to the Fock-Operator in the SCF calculation to represent the effect of solvent polarization.[164] A combination of 3D RISM and Kohn-Sham DFT was developed by Kovalenko and Hirata[166,167], which was later implemented into the Amsterdam density functional package (ADF)[168], whereby analytical first derivatives were developed.[169,170] Alternative formulations of the 3D RISM-SCF were designed by Minezwa[171] and Aono[172]. In contrast to the previous RISM-SCF versions, where the solute partial charges were used for solvent polarization, the exact QC-derived electrostatic potential was applied in these implementations. In 2020, Reimann and Kaupp added to the ADF implementation of 3D RISM-SCF the exact electrostatic potential to polarize the solvent and developed numerical second derivatives.[173]

In 2008 Heil, Kloss and Kast[24] developed a combination of the 3D RISM solvent model with QC calculations of the partial charges of the solute, the so-called embedded cluster reference interaction site model (EC-RISM). In 2011 an alternative variant was presented, in which the 3D RISM model was coupled with the AMOEBA polarizable force field.[174] Another implementation of EC-RISM, which is limited to point charge solute-solvent interactions only, was developed by Ganyecz and Kállay in 2022.[175]

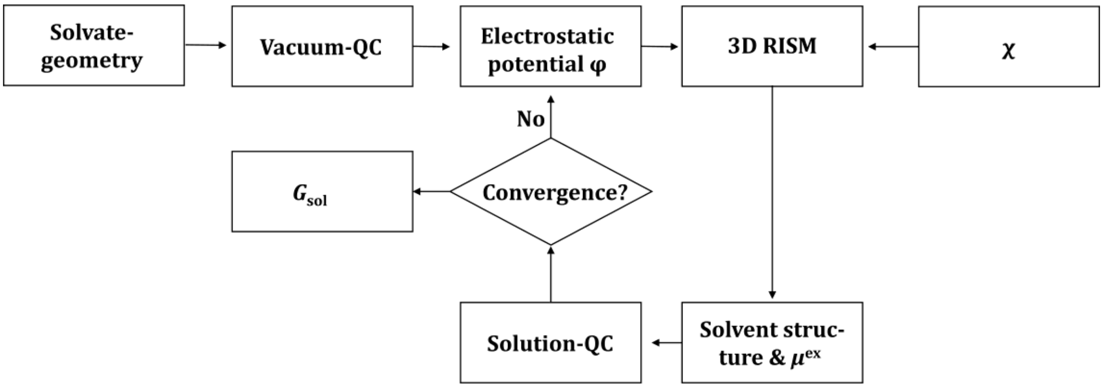

Figure 3 Schematic representation of one EC-RISM cycle. [24]

The idea behind EC-RISM is to calculate the mutual polarization between solute and solvent. On the one hand, a polarized wave function of the solute can be obtained, and on the other hand, the Gibbs free energy G_{sol} can be calculated via :

$$G_{\mathrm{sol}} = E_{\mathrm{sol}} + \mu^{\mathrm{ex}} \tag{60}$$

where $E_{\mathrm{sol}}$is the intramolecular energy of the solute, which can be taken from a QC-calculation and $\mu^{\mathrm{ex}}$is the excess chemical potential, which is calculated via eqs. (58) and (59). In this formulation, contributions from the so-called thermal corrections are neglected.

The corresponding scheme of an EC-RISM cycle is depicted in Figure 3. Before the EC-RISM cycle starts, a suitable solute geometry must be first calculated, usually using a PCM or vacuum geometry optimization. Furthermore, an appropriate force field for the Lennard-Jones parameters must be selected. The starting point is a vacuum QC calculation from which a QC electrostatic potential or the corresponding partial charges are obtained. The electrostatic potential is used in the subsequent 3D RISM calculation to polarize the solvent. At this point, the solvent susceptibility is needed from a pre-computed 1D RISM calculation. From this calculation, a polarized solvent environment is obtained, which is then embedded in the QC calculation as background point charges. To calculate these point charges, the charge density ρ_q is obtained by:

$$\rho_q(\boldsymbol{r}) = \sum_{\gamma} q_{\gamma}\, \rho_{\gamma} g_{\gamma}(\boldsymbol{r}) \tag{61}$$

where q_{γ} is the partial charge of the interaction solvent site γ, usually taken from a known force field. The charge density can then be discretized on a grid, resulting in the background charges,

$$q(\boldsymbol{r}_i) = \rho_q(\boldsymbol{r}_i)\Delta V \tag{62}$$

whereby $\Delta V$is the volume of one grid cell. The corresponding Hamilton operator under the influence of the background charges can be written in the following form:

$$\widehat{H}_{\mathrm{tot}} = \widehat{H}_{\mathrm{ne}} + \widehat{H}_{\mathrm{ee}} + \widehat{H}_{\mathrm{nn}} + \widehat{H}_{\mathrm{nq}} + \widehat{H}_{\mathrm{eq}} + \widehat{H}_{\mathrm{qq}}, \tag{63}$$

where $\widehat{H}_{\mathrm{tot}}$ is the total Hamiltonian, the terms on r.h.s. describe the interactions between the nuclei (n), the electrons (e), and the point charges q. Note that the target value is not the total energy of the solvent-solute system, but only the solute's intramolecular energy under the solvent's influence is required. To calculate these intramolecular energies, there are two variants; in the first one, the total wave function of the system is taken, and only the intramolecular operators are applied to it in the following form:

$$E_{\mathrm{sol}} = \langle \Psi_{\mathrm{tot}} | \widehat{H}_{\mathrm{ne}} + \widehat{H}_{\mathrm{ee}} + \widehat{H}_{\mathrm{nn}} | \Psi_{\mathrm{tot}} \rangle. \tag{64}$$

This is usually the most elegant way to calculate the intramolecular energy of the solute; however, some QC codes do not support these zero SCF-cycle procedures. The alternative variant is to subtract the remaining contribution from the total energy. Normally the term $\widehat{H}_{\mathrm{qq}}$ is suppressed during the calculations. The remaining two terms describe the electrostatic interaction of the solute with the background point charges. To calculate this contribution, the electrostatic potential from the QC calculation is multiplied with the final charge density and integrated over the whole box cell via:

$$E_q = \int \rho_q(\boldsymbol{r})\, \varphi(\boldsymbol{r}) d\boldsymbol{r}. \tag{65}$$

After every step of the cycle, the energy residue between the actual and last steps is calculated, and if the residual value falls below a specific limit, the cycle is considered converged. From the converged wave function, the energy of the solute can be obtained and a variety of spectroscopic observables, for example, chemical shielding constant or the isotropic hyperfine coupling constant, even under high hydrostatic pressures.

If the exact quantum mechanical electrostatic potential $\varphi(\boldsymbol{r})$ is used to calculate the electrostatic interactions for the corresponding interaction energy function $U^{\mathrm{elec},\varphi}$ at location **r** follows:

$$u^{\mathrm{elec},\varphi}(\boldsymbol{r}) = \sum_{\gamma} \varphi(\boldsymbol{r})\, q_{\gamma}. \tag{66}$$

The full electrostatic potential is then built using a renormalization ansatz (derived by Hoffgaard, Heil, and Kast)[174] via:

$$u_{\gamma}^{\mathrm{elec}}(\boldsymbol{r}) = u_{\gamma}^{\mathrm{q,elec,short}} + \Delta u_{\gamma}^{\varphi,\mathrm{elec}}(\boldsymbol{r}) + u_{\gamma}^{\mathrm{q,elec,long}}(\boldsymbol{r}), \tag{67}$$

whereas $\Delta u_{\gamma}^{\varphi,\mathrm{elec}}(\boldsymbol{r})$ is the difference between the exact electrostatic potential and the point charge representation at the corresponding grid point. $u_{\gamma}^{\mathrm{q,elec,short}}$ is the point charge-based short-range potential, which is calculated according to the Ewald sum scheme in real space. $u_{\gamma}^{\mathrm{q,elec,long}}$ is the long-range potential, which is calculated in reciprocal space. Divergence problems can occur if the difference $\Delta u_{\gamma}^{\varphi,\mathrm{elec}}(\boldsymbol{r})$ does not disappear at the edges of the box. In 2018, P. Kibies and S. Kast[176] developed a smooth switching approach in which the QC electrostatic potential describes the interactions in the vicinity of the solute, and the potential is switched to the point charge potential in the vicinity of the box edges. The resulting cubic switching function has the following form:

$$s\left(r, r_{\min}, r_{\max}\right) = \begin{cases} 1 & \forall r < r_{\min} \\ s_0 + s_1 r + s_2 r^2 + s_3 r^3 & \forall r_{\max} \leq r \leq r_{\min} \,, \\ 0 & \forall r < r_{\max} \end{cases} \tag{68}$$

with the parameters:

$$s_0 = (3r_{\min} r_{\max}^2 - r_{\max}^3)\left(r_{\min} - r_{\max}\right)^3, \tag{69}$$

$$s_1 = -6r_{\min} r_{\max}\left(r_{\min} - r_{\max}\right)^3 \quad , \tag{70}$$

$$s_2 = 3(r_{\min} + r_{\max})\left(r_{\min} - r_{\max}\right)^3, \tag{71}$$

and

$$s_3 = -2\left(r_{\max} - r_{min} \quad\right)^3. \tag{72}$$

The following term corrects the error in the excess chemical potential caused by manipulating the electrostatic potential:

$$\begin{aligned} \Delta\mu^{ex} &= \int_V g(\boldsymbol{r})\rho_\gamma \, \Delta u_\gamma^{\varphi,\mathrm{elec}} \\ &= \int_V g(\boldsymbol{r})\rho_\gamma \left(1 - s(r_{\max} - r_{\min})\left(U_\gamma^{\varphi,\mathrm{elec}} - U_\gamma^{q,\mathrm{elec}}\right)\right). \end{aligned} \tag{73}$$

EC-RISM was mainly used to calculate energy differences in the early years, e.g., tautomer equilibria or -ΔpK_a values. However, when absolute energies were tried to be calculated, there were significant errors in calculating the excess chemical potential. It was observed that this error is directly correlated to the solute-sized cavity in the solvent and, therefore, correlated to the partial molar volume. For correction of this error, different empirical corrections were developed. For example, Palmer established the so-called universal correction[177,178], whereby Sergiievskyi developed the pressure correction of solvation free energy, called PC/PC+ correction.[179] Based on this preliminary work, Tielker[180] and Tomazic[181] developed empirical corrections, which correct the excess chemical potential to represent experimental absolute hydration-free energies. The corresponding equation for the corrected excess chemical potential is:

$$\mu^{\mathrm{ex,corr}} = \mu^{ex} + c_{V_m} V_m + c_q q, \tag{74}$$

where $\mu^{\mathrm{ex,corr}}$is the corrected excess chemical potential, V_m is the partial molar volume (PMV), c_{V_m} is the empirical parameter, which corrects the PMV, q is the net charge of the molecule of interest and c_q is the empirical parameter that corrects the net charge. In 2018, Tielker developed an extension of the PMV correction for high pressures to calculate accurate excess chemical potentials even under the influence of high pressure.[27] The major problem was to find experimental data where hydration energies were measured under high pressure. As an alternative, pressure-dependent thermodynamic integrations (TI) were performed instead, and the reference value at high pressure was extrapolated from the relative pressure difference between the pressure-dependent Tis and the ambient condition value for the correct EC-RISM solvation free energy. The expression for the reference solvation free energy (using Ben-Naim reference states) $\Delta_{\mathrm{solv}} G_{\mathrm{ref}}^0$ is:

$$\Delta_{\mathrm{solv}}G_{\mathrm{ref}}^{0}(p) = \Delta_{\mathrm{solv}}G_{\mathrm{TI}}^{0}(p) + \Delta_{\mathrm{solv}}G_{\text{EC-RISM}}^{0}(1\mathrm{bar}) - \Delta_{\mathrm{solv}}G_{\mathrm{TI}}^{0}(1\mathrm{bar}), \tag{75}$$

whereby $\Delta_{\mathrm{solv}}G_{\mathrm{TI}}^{0}(p)$ is the pressure-dependent solvation free energy from the TI, $\Delta_{\mathrm{solv}}G_{\text{EC-RISM}}^{0}(1\mathrm{bar})$ is the PMV-corrected EC-RISM-based solvation free energy. The reference data was calculated for a small subset of the MNSOL (details can be found in Ref[27]). The expression for the pressure-dependent PMV-corrected excess chemical potential is:

$$\mu^{\mathrm{ex,corr,HP}} = \mu^{\mathrm{ex,corr}} + c_{\mathrm{HP}}(p - 1\mathrm{bar})V_{m}(p), \tag{76}$$

with c_{HP} , which is the additional pressure-dependent correction parameter. The correction parameter c_{HP} equals -1.0108 10^{-5} kcal mol^{-1} Å^{-3} bar^{-1}, for the level of theory MP2/6-311+G(d,p)/EC-RISM.

A limitation of the EC-RISM calculations is the point that the calculations are performed on optimized geometries based on continuum solvation. In the publication of Hölzl, pressure-dependent geometry optimizations for TMAO were performed.[18] A disadvantage of this procedure is at an external QC program was used to calculate the gradient of the solute electronic energy. Recently a direct implementation of EC-RISM geometry optimizations within the QC program *ORCA*[182] was realized by P. Kibies and M. A. Garcia-Rates. Here, the gradient of the free energy is defined as:

$$\frac{\partial G}{\partial \boldsymbol{R}_{\alpha}} = \frac{\partial E_{\mathrm{sol}}}{\partial \boldsymbol{R}_{\alpha}} + \frac{\partial \mu^{\mathrm{ex}}}{\partial \boldsymbol{R}_{\alpha}}. \tag{77}$$

The gradient consists of two parts; the first one is the derivative of the intramolecular energy w.r.t. the coordinates of the solute, and the second one is the gradient of the excess chemical potential. The former is directly calculated within *ORCA*, whereas the latter is calculated within 3D RISM in the following manner:[169]

$$\frac{\partial \mu^{\mathrm{ex}}}{\partial \boldsymbol{R}_{\alpha}} = \sum_{\gamma} \rho_{\gamma} \int d\,\boldsymbol{r} g_{\gamma}(\boldsymbol{r}) \left(\frac{\partial u_{\gamma}^{\mathrm{LJ}}}{\partial \boldsymbol{R}_{\alpha}} + \frac{q_{\gamma}\partial \phi(\boldsymbol{r}, \boldsymbol{R}_{\alpha})}{\partial \boldsymbol{R}_{\alpha}} \right). \tag{78}$$

2.2 IR Spectroscopy

2.2.1 Foundations

As mentioned in the introduction, vibrational spectroscopy deals with transitions between vibrational states. The vibrations can be described both in a classical and quantum mechanical way. An outstanding overview and a detailed derivation of the mathematical fundamentals of vibration spectroscopy can be found in the book written by Wilson, Decius, and Cross[183], and the following chapter will mainly be based on this book. Here, the author wants to present only a tiny part of the mathematics to better understand the methods used later in this work. The first hurdle is to describe the vibrations accurately and choose a suitable coordinate system. It was shown that the best coordinate system is based on three cartesian center of mass coordinates for the translation, the three Eulerian angles of a rotating system, and the cartesian coordinates of the molecule atoms with respect to the rotating system. Since there are only $3N$ possible degrees of freedom for a molecule, there are six too many coordinates in the aforementioned list.

Consequently, the six coordinates do not have to be independent of each other. To define the rotating coordinate system for nonlinear molecules, only six conditions have to be defined. The first three conditions place the origin of the rotating system in the center of mass so that the rotating system moves together with the molecule. The second three conditions bind the coordinate system and the molecule together so that they both rotate together.

With the choice of this coordinate system, it is now possible to accurately describe vibration transitions. The kinetic energy T of the system can be described by:

$$T = \frac{1}{2}\sum_{\alpha=1}^{N} m_\alpha \left[\left(\frac{d\Delta x_\alpha}{dt}\right)^2 + \left(\frac{d\Delta y_\alpha}{dt}\right)^2 + \left(\frac{d\Delta z_\alpha}{dt}\right)^2\right], \tag{79}$$

where m_α is the mass of the αth atom and dt the time derivative of the displacement coordinate:

$$\Delta x_\alpha = x_\alpha - \alpha_\alpha, \tag{80}$$

here x_α the coordinate of the αth atom in the moving system is and α_α is the equilibrium position. At this point, the definition of mass-weighted coordinates is introduced, the advantage of which will be shown later in this chapter. The mass-weighted displacement coordinates q_i are defined as:

$$q_1 = \sqrt{m_1}\Delta x_1 \quad q_2 = \sqrt{m_1}\Delta y_1 \quad q_3 = \sqrt{m_1}\Delta z_1 \quad q_4 = \sqrt{m_2}\Delta x_2 \dots. \tag{81}$$

Based on this definition, the kinetic energy can now be written in the following compact representation:

$$T = \frac{1}{2}\sum_{i=1}^{3N} \dot{q}_i^2, \tag{82}$$

with $\dot{q}$ as the time derivative of the mass-weighted displacement coordinates. The potential energy function is dependent on the displacements. If the displacements are relatively small, a Taylor series expansion in the displacement coordinate q can be used to describe the potential energy via:

$$V = \frac{1}{2}V_0 + \frac{1}{2}\sum_{i=1}^{3N}\left(\frac{\partial V}{\partial q_i}\right)_0 q_i + \sum_{i=1}^{3N}\left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_0 q_i q_j + \text{etc..} \tag{83}$$

Since the position of the energy zero-point can be chosen freely, V_0 can be eliminated under the assumption that $V_0 = 0$. All q's should be zero in the equilibrium, and the molecule's energy should be at a minimum. The latter results in the following relations:

$$\left(\frac{\partial V}{\partial q_i}\right)_0 = f_i = 0 \quad \text{and} \quad \left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_0 = f_{ij}, \tag{84}$$

where f_{ij} are constants, in which the mixed partial derivatives of the potential energy with respect to the displacement coordinates are encoded. Higher-order terms will be neglected (this is reasonable as long as the vibration amplitude is sufficiently small). Since the kinetic energy depends only on the velocities of the atoms and the potential energy of the molecule only on the atomic position, Newton's equation of motion can be written as follows in the context of Lagrange formalism:

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} + \frac{\partial V}{\partial q_j} = 0 \quad j = 1,2,\ldots,3N. \tag{85}$$

If equations (79) and (83) are substituted in equation (85), the following relation is obtained:

$$\ddot{q}_j + \sum_{i=1}^{3N} f_{ij} q_i = 0, j = 1,2,\ldots.,3N. \tag{86}$$

A possible solution to this differential equation system is:

$$q_i = A_i \, cos\left(\lambda^{1/2} t + \varepsilon\right), \tag{87}$$

where A_i, λ and ε are constants, which need to be appropriately chosen. Inserting the latter relation leads to the following:

$$\sum_{i=1}^{3N} \left(f_{ij} - \lambda \delta_{ij}\right) A_i = 0 \quad j = 1,2,\ldots.,3N, \tag{88}$$

whereby δ_{ij} is the well-known Kronecker delta symbol. Equation (88) consists of a composition of homogenous linear algebraic equations in the $3N$ amplitudes. The trivial solutions λ correspond to no vibrational modes with amplitudes of A=0,1,2. All non-trivial solutions λ can be obtained by solving the secular equation:

$$\begin{vmatrix} f_{11} - \lambda & f_{12} & \cdots & f_{1,3N} \\ f_{21} & f_{22} - \lambda & \cdots & f_{2,3N} \\ \cdots & \cdots & \cdots & \cdots \\ f_{3N,1} & f_{3N,2} & \cdots & f_{3N,3N} - \lambda \end{vmatrix} = 0. \tag{89}$$

From the elements of the determinant of this equation, the coefficients of the unknown amplitudes can be obtained. It is now possible to select a fixed value λ_k so that the determinant vanishes, and the coefficients of the amplitudes A become fixed. For example, the possible solutions are not unique solutions to the equations system; instead, they can only calculate ratios. A unique solution can be obtained via the following relationship:

$$l_{ik} = \frac{A'_{ik}}{\left[\sum_i (A'_{ik})^2\right]^{\frac{1}{2}}}, \tag{90}$$

where A'_{ik} is an arbitrary set of solutions and l_{ik} amplitudes that are normalized in the following fashion:

$$\sum_i l_{ik}^2 = 1. \tag{91}$$

To obtain the solution for the real physical problem, the following ansatz can be solved:

$$A_{ik} = K_k l_{ik}, \tag{92}$$

with K_k as constants, which are determined by the initial values of the coordinates and velocities. Equation (89) consists of $3N$x$3N$ rows and columns leading to $3N$ possible solution for unknown A_i. However, only $3N$-6 roots (in the case of non-linear molecules) are not zero and correspond to vibrational modes. The frequency of the vibrational mode oscillating around its equilibrium position is:

$$v = \frac{\sqrt{\lambda_k}}{2\pi}. \tag{93}$$

For a given λ_k, the movement of each coordinate shows the same frequency and phase but not necessarily the same amplitude. As a result, the phase and frequency are identical; all atoms show the same course along the displacement coordinate. This means that all atoms simultaneously experience their maximum displacement or the passing of their equilibrium. Vibrational modes that accomplish these criteria are called normal modes of vibration. Not all frequencies need to have different values; instead, identical values can be assigned to so-called degenerated modes.

A so-called normal coordinate can be assigned to each normal mode. The introduction of normal coordinates is useful to accurately describe, e.g., the quantum mechanical treatment of molecular vibrations. The normal coordinates Q_k can be described in dependency of the mass-weighted cartesian displacement coordinates q_i as follows:

$$Q_k = \sum_{k=1}^{3N} l''_{ki} q_i \quad k = 1,2,..,3N. \tag{94}$$

The coefficients l''_{ki} must be selected so that no cross-products occur in the potential energy. To achieve this, the kinetic and potential energies need to have the following form:

$$T = \frac{1}{2}\sum_{k=1}^{3N} \dot{Q}_k^2 \, V = \frac{1}{2}\sum_{k=1}^{3N} \lambda'_k Q_k^2. \tag{95}$$

The calculation of the potential energy depends only on the squares of Q_k. It is necessary to mention that the indices *i* and *j* refer to the *q* coordinates and *k* and *l* indicate the *Q* coordinates. The coordinates *q* can be described in terms of the normal coordinates via:

$$q_i = \sum_{k=1}^{3N} l'_{ik} Q_k \quad i = 1,2,\ldots,3N. \tag{96}$$

Equations (89) and (96) can be substituted in one and vice versa, and the following result is obtained:

$$\sum_{i=1}^{3N} l'_{kk} l'_{ll} = \delta_{kl} \quad \sum_{k=1}^{3N} l'_{ik} l'_{kj} = \delta_{ij}. \tag{97}$$

The Newtonian equation of motion can now be expressed in terms of the normal coordinates as follows:

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{Q}_k} + \frac{\partial T}{\partial Q_k} = \ddot{Q}_k + \lambda'_k Q_k = 0 \quad k = 1,2,\ldots.,3N, \tag{98}$$

and the corresponding solutions are:

$$Q_k = K'_k \cos\left(\sqrt{\lambda'_k} t + \varepsilon'_k\right) \quad k = 1,2,\ldots,3N. \tag{99}$$

with the constants K'_k and ε'_k

The potential function, which appears in the equation, can have different origins. Both classical force fields and QC methods are suitable. Depending on the choice of QC methods, it is possible to calculate the Hesse matrix analytically or numerically. For example, analytical expressions can be found for HF, MP2[184], and many DFT methods. Analytical calculations have great advantages: they are more accurate and less time-consuming. However, the differences between the individual QC programs must be considered. For example, the calculation of the reduced mass for the Gaussian program is different from other programs. J. W. Ochterski has written a detailed essay on this topic, describing the steps for calculating the vibration analysis within the Gaussian program. [185]

2.3 Magnetic-resonance spectroscopy

2.3.1 Principles

Magnetic resonance spectroscopy methods are indispensable in chemical analysis. In this chapter, the basic principles needed to understand and interpret everyday experiments will be shown, and then, in the following sections, the concrete calculation methods that can be utilized in QC calculations will be derived. First, the magnetic moments of an electron spin and the nuclear magnetic moment shall be defined. For the electron spin magnetic moment $\vec{m}_i$ of electron *i,* it follows:

$$\vec{m}_i = \gamma_s \vec{s}_i = -g_e \frac{\mu_B}{\hbar} \vec{s}_i = -g_e \frac{e}{2m_e} \vec{s}_i, \tag{100}$$

where γ_s is the gyromagnetic ratio, $\vec{s}$ is the spin-vector, g_e the Lande-factor of 2.0023, μ_B is the Bohr magneton, *e* is the elementary charge, m_e is the mass of the electron. The corresponding expression for the nuclear magnetic moment $\vec{\mu}_C$ of nucleus *C* results as follows:[95]

$$\vec{\mu}_C = \gamma_C \vec{I}_C = +g_C \frac{e}{2m_p} \vec{I}_C = +g_C \frac{\mu_N}{\hbar} \vec{I}_C, \tag{101}$$

with γ_C the gyromagnetic ratio of the nucleus *C*, g_C is the g-factor of the nucleus *C*, which shows a strong dependence on the nucleus of choice, m_p is the mass of the proton; the connection between the elementary charge and m_p is given by the nucleus magneton μ_N. The nuclear spin $\vec{I}_C$ is defined as:

$$|\vec{I}_C| = \sqrt{I(I+1)\hbar^2}, \tag{102}$$

where I is the nuclear spin quantum number. The nuclear magnetic moment becomes zero for nuclei with an even number of protons and neutrons; consequently, these nuclei cannot be detected via NMR experiments. For systems with an odd number of protons and/or neutrons, the quantum spin number becomes a non-zero integer or half number. In magnetic resonance experiments, the sample is transferred into an external magnetic field $\vec{B}$. This external magnetic field leads to an energy splitting which has the following form in the case of the electron spin:

$$E_{m_s} = -\vec{m}_i \cdot \vec{B} = -m\ {}_s\gamma_s\hbar B_0 = \pm\frac{1}{2}g_e\mu_B B_0, \tag{103}$$

where m_s is the magnetic spin quantum number. The energy difference between the two energy states is given via:

$$\varDelta E_{m_s} = \frac{1}{2}g_e\mu_B B_0 = h\nu. \tag{104}$$

The effect of the energy splitting due to the induction of the external magnetic field is the so-called Zeeman effect. The resonant frequency that induces the energy transfer between the two-state can be determined as follows:

$$\nu = \frac{\gamma_s B_0}{2\pi}. \tag{105}$$

Where ν is also called the Larmor-frequency, straightforwardly, analog expressions can also be determined for the nuclear spins in an external magnetic field. The corresponding expressions are:

$$E_{m_I} = -m_I\gamma_I\hbar B_0 = \mp\frac{1}{2}g_I\mu_N B_0, \tag{106}$$

$$\varDelta E_{m_I} = g_I\mu_N B_0, \tag{107}$$

$$\nu = \frac{\gamma_I B_0}{2\pi}. \tag{108}$$

In this example, the nucleus of interest has a nuclear spin quantum number I of ½. This is the case for some of the most frequently measured nuclei in everyday experiments, e.g., ^{1}H, ^{15}N, ^{13}C, and ^{19}F. An important quantity within NMR spectroscopy, less so in EPR spectroscopy, where the energy difference between the two energy levels is much greater than in NMR spectroscopy, is the relative occupation number of the individual energy levels. These occupation numbers can be determined with the help of Boltzmann statistics as follows:

$$\frac{N_{(m=-\frac{1}{2})}}{N_{(m=+\frac{1}{2})}} = \exp(-\beta\Delta E) \approx 1 - \frac{\gamma_I\hbar B_0}{k_B T}, \tag{109}$$

where β is the inverse temperature($k_B T^{-1}$), which is directly connected to the Boltzmann-constant k_B and the temperature T. For protons or nuclei, the energy difference is only very small, so the two levels are virtually identically occupied.

Equation (109) shows that both a stronger magnetic field and a lower temperature lead to a larger population difference. The decisive advantage, which causes the outstanding importance of magnetic resonance spectroscopy methods, is that the resonance condition occurring in equation (107) does not depend solely on the external magnetic field but depends on a so-called effective magnetic field. This effective magnetic field depends on the chemical environment around the observed nucleus or electron spin, leading to different signals in the spectrum. The external magnetic field is weakened by the so-called shielding constant, which results in the following relationship for the effective magnetic field:

$$B_{\text{eff}} = B_0 - \sigma B_0 = (1-\sigma)B_0, \tag{110}$$

and for the resonance condition, it follows:

$$\nu = \frac{\gamma_I}{2\pi} B_{\mathbf{0}}(1-\sigma). \tag{111}$$

Different solvents have different effects on the shielding constant. In general, it is not the shielding constant that is specified but the so-called chemical shift. Since the resonance frequency depends on the strength of the external magnetic field, different spectra result from different magnetic field strengths. The resonance frequencies are referenced relative to a so-called standard for better comparability between different experiments. The chemical shift δ is defined as:

$$\delta = \frac{\nu - \nu^0}{\nu^0} \cdot 10^6 = \frac{(1-\sigma)B_0 - (1-\sigma_0)B_0}{(1-\sigma_0)B_0} \cdot 10^6 \approx (\sigma_0 - \sigma) \cdot 10^6. \tag{112}$$

where ν^0 and σ_0 are the resonance frequency and the shielding constant of the reference substance, respectively. The most common reference substance for referencing ^{1}H and ^{13}C is TMS in unipolar solvents. If the measurement is made in a polar solvent, DSS is often used as a standard. Liquid ammonia is used to reference ^{15}N chemical shifts. A similar environmental effect can also be observed in EPR spectroscopy. In this case, the shielding is combined with the g_e-factor to a g-factor and leads to:

$$B_{\text{eff}} = (1-\sigma)B_0 = (g/g_e)B_0. \tag{113}$$

In addition to chemical shifts, a splitting of the signals (fine structure) can be observed in the NMR spectrum. These fine structures result from the coupling between the magnetic moments of different nuclei. In principle, there are two different coupling mechanisms. On the one hand, the so-called direct dipolar coupling, where the nuclei couple directly through space. This phenomenon becomes relevant in the context of solid-state NMR spectroscopy since the contributions average to zero in liquid media with low viscosity. However, the dipolar couplings are the reason behind the nuclear Overhauser effect (NOE).[95]

On the other hand, there is the so-called indirect spin-spin coupling, which is mediated through bonds. The indirect coupling is called *J*-coupling and results from a complex interaction between the electron spins and nuclear magnetic moments. In EPR spectra, splitting can occur due to the coupling of the free electron spins with the nuclear magnetic moments. As long as systems are considered in solution, the description of hyperfine splitting is usually achieved simply by specifying the isotropic hyperfine coupling constant. The next sections of this thesis deal with the concrete calculation of the quantities observed in the experiment using QC formulations. For this purpose, the concept of the effective spin Hamiltonian is first introduced, and then explicit expressions for magnetic observables are obtained.

2.3.2 The effective spin Hamiltonian

A spatial electronic wave function can describe a molecule with N atoms, with the electron spin state quantum numbers m_S and the nuclear spin quantum numbers m_I. The two states, which are created by a spin transition ϕ, can then be described as follows:[101]

$$\phi^a = |\Psi, m_S^a, m_1^a, \ldots, m_N^a\rangle \quad \Leftrightarrow \quad \phi^b = |\Psi, m_S^b, m_1^b, \ldots, m_N^b\rangle, \tag{114}$$

where a and b indicate the spin states. It is reliable to assume that the spatial electronic wave function does not change much during the spin transition.[101] Therefore, the energy difference due to this spin transition can be represented as follows:

$$\Delta E^{ab} = E(\phi^b) - E(\phi^a) = E\left(m_S^b, m_1^b, \ldots., m_N^b\right) - E(m_S^a, m_1^a, \ldots., m_N^a). \tag{115}$$

This spin transition energy difference is relatively small compared to the orders of magnitude resulting from nucleus-electron and electron-electron interactions. Although the surrounding electrons influence the magnetic interactions, it is possible to describe the energy transfer detected by NMR and EPR spectroscopy solely by a pure spin Hamiltonian, a so-called effective spin Hamiltonian. The electrons no longer appear in this effective spin Hamiltonian, except via their spin state information, and the nuclei are described only by their intrinsic spin and the resulting magnetic moments.[186] The energy for a spin state can then be described in the following way:

$$\begin{aligned}\langle\phi^a|\hat{H}_S|\phi^b\rangle &= \langle m_S^a, m_1^a, \ldots., m_N^a|\hat{H}_S|m_S^a, m_1^a, \ldots., m_N^a\rangle \\ &= E(m_S^a, m_1^a, \ldots., m_N^a).\end{aligned} \tag{116}$$

As mentioned before, magnetic resonance spectroscopic methods describe the oscillations between electronic and nuclear magnetic dipoles with an external magnetic field. In the effective spin Hamiltonian, therefore, the pairwise interactions of all magnetic moments with each other and with the external magnetic field $\vec{B}$ must be described. The contributions are coming from the electron spin $\vec{S}$ and the collection of nuclear spins, which are defined in the following manner:

$$\vec{I} = \sum_{C=1}^{N} \vec{I}_C, \tag{117}$$

and the electron orbital angular moment $\vec{L}$. The corresponding spin Hamiltonian has the following form:

$$\begin{aligned}\hat{H}_S = {} & \hat{H}_S(S,S) + \hat{H}_S(S,I) + \hat{H}_S(S,L) + \hat{H}_S(S,B) + \hat{H}_S(I,I) \\ & + \hat{H}_S(I,L) + \hat{H}_S(I,B) + \hat{H}_S(L,L) + \hat{H}_S(L,B).\end{aligned} \tag{118}$$

The influence of the electron angular moment is rather small compared to the impact originating from the external magnetic field and can be treated separately.[101] Thus, the Hamiltonian reduces to:

$$\hat{H}_S = \hat{H}_S(S,S) + \hat{H}_S(S,I) + \hat{H}_S(S,B) + \hat{H}_S(I,I) + \hat{H}_S(I,B). \tag{119}$$

The next question arises: What do the individual interactions look like in concrete terms to formulate the Hamiltonian? The classical expressions for the interactions between magnetic dipoles can be used as a starting point. Therefore, the interactions of the magnetic dipoles with each other and the external magnetic field are represented by simple scalar products. Thus, for the Hamiltonian, the following relation is obtained:

$$\hat{H}_S = C_{SS}\vec{S}\cdot\vec{S} + C_{SI}\vec{S}\cdot\vec{I} + C_{SB}\vec{S}\cdot\vec{B} + C_{II}\vec{I}\cdot\vec{I} + C_{IB}\vec{I}\cdot\vec{B}, \tag{120}$$

where the constants C are the parameters that include the different influences of the dipole-dipole interaction on the transition energy. Depending on whether an NMR or an EPR experiment is to be represented, there are different formulations for the effective spin Hamiltonians. For an NMR experiment, the common Hamiltonian operator looks like this:

$$\begin{aligned}\hat{H}_S(\mathrm{NMR}) = {} & -\sum_{C=1}^{N} \hbar\gamma_C \vec{B}\left(\overleftrightarrow{1} - \overleftrightarrow{\sigma}_C\right)\cdot\vec{I}_C \\ & + \frac{\hbar^2}{2}\sum_{C=1}^{N}\sum_{\substack{D=1\\ D\neq C}}^{N} \gamma_C\gamma_D\,\vec{I}_C\cdot\left(\overleftrightarrow{D}_{CD} + \overleftrightarrow{J}_{CD}\right)\cdot\vec{I}_D,\end{aligned} \tag{121}$$

here $\overleftrightarrow{1}$ is a unitary tensor, $\overleftrightarrow{\sigma}_C$ the nuclear shielding tensor, $\overleftrightarrow{D}_{CD}$ the classical dipolar and $\overleftrightarrow{J}_{CD}$ are the indirect nuclear spin-spin J-coupling tensors. Equation (121) is the paramagnetic generalization, which is able to address the relevant interactions in most NMR spectra. However, there are also cases where the effective spin Hamiltonian needs to be extended. If, for example, atoms in the molecule under consideration have high spin nuclear magnetic dipoles ($I_C > 0.5$), an additional term must be included, including the resonance effects resulting from the electric quadrupole moment. The corresponding expression now looks as follows:

$$\hat{H}_Q(I,I) = \sum_{\substack{C=1\\ |I_C|\geq 1}}^{N} \vec{I}_C\cdot\overleftrightarrow{Q}_{CC}\cdot\vec{I}_C, \tag{122}$$

with $\overleftrightarrow{Q}_{CC}$ as the quadrupole coupling tensors.

Analogous to the effective NMR spin Hamiltonian, an effective spin Hamiltonian can also be formulated to describe an EPR experiment. The most common version of this spin Hamiltonian has the following form:

$$\hat{H}_S(\mathrm{EPR}) = \mu_B \vec{S}\cdot\overleftrightarrow{g}\cdot\vec{B} + \sum_{C=1}^{N}\vec{S}\cdot\overleftrightarrow{A}_C\cdot\overleftrightarrow{I}_C \tag{123}$$

whereby the first term is the description of the Zeeman effect, with the g-tensor $\overleftrightarrow{g}$, that can be interpreted as an equivalent to the shielding tensor $\overleftrightarrow{\sigma}_C$. In the second term, the hyperfine splitting is described with $\overleftrightarrow{A}_C$ as hyperfine coupling tensor. Equation (123) neglects the interactions between the nuclear magnetic dipoles and between the nuclear magnetic dipoles and the external magnetic field. This is justifiable since the strength of a nuclear magnetic dipole is only about 1/2000 of the electron spin magnetic dipole. Therefore, in most experiments, the chemical shifts and J-couplings occur on different energy scales than the g-value and hyperfine coupling constants. However, if experiments like the ENDOR (electron-nuclear double resonance) are to be quantified quantum mechanically, it is necessary to consider the interactions of the nuclear magnetic dipole moments. The effective spin Hamiltonian for an ENDOR experiment is the combination of equation (121) and equation (123). The next section of this thesis will deal with the concrete calculations of the parameters $\overleftrightarrow{\sigma}_C$, $\overleftrightarrow{A}_C$, $\overleftrightarrow{J}_{CD}$ and $\overleftrightarrow{g}$, based on the evaluation of first-principle physics.

2.3.3 Spectroscopic parameters

The absolute goal is to find expressions to extract the above-introduced parameters from a previously calculated electronic wave function. As a starting point, it is useful to extend the expression for the electronic energy under the perturbation of the magnetic field $\vec{B}$ and the nuclear magnetic dipole moments $\vec{\mu}_C$. (at this point, an expression for the NMR parameters shall be derived first, but the EPR parameters can be determined according to the same principles). The series development takes place around zero fields and zero magnetic moments and leads to:[186,187]

$$E(\vec{B},\vec{\mu}) = E_0 + \frac{1}{2}\vec{B}\cdot\overleftrightarrow{E}^{(20)}\vec{B} + \sum_{C}\vec{B}\cdot\overleftrightarrow{E}_C^{(11)}\vec{\mu}_C + \frac{1}{2}\sum_{D\neq C}\vec{\mu}_C\cdot\overleftrightarrow{E}_{DC}^{(02)}\vec{\mu}_D. \tag{124}$$

Where second derivatives are defined as:

$$\overleftrightarrow{E}^{(20)} = \left.\frac{d^2E(\vec{B},\vec{\mu})}{d\vec{B}^2}\right|_{\vec{B}=0,\vec{\mu}=0}, \tag{125}$$

$$\overleftrightarrow{E}_C^{(11)} = \left.\frac{d^2E(\vec{B},\vec{\mu})}{d\vec{B}d\vec{\mu}}\right|_{\vec{B}=0,\vec{\mu}=0}, \tag{126}$$

$$\overleftrightarrow{E}_{DC}^{(02)} = \left.\frac{d^2E(\vec{B},\vec{\mu})}{d\vec{\mu}_C d\vec{\mu}_D}\right|_{\vec{B}=0,\vec{\mu}=0}. \tag{127}$$

For closed-shell systems, the first-order derivatives disappear and are therefore not included in the series development. Higher-order terms are also neglected, and equation (124) is exact for non-rotating rigid molecules. If a comparison between equation (121) and equation (124) is made, it is noticeable that the parameters from equation (121) can be assigned to the second derivatives as follows:[186]

$$\overleftrightarrow{\sigma}_C = \overleftrightarrow{E}_C^{(11)} + \overleftrightarrow{1}, \tag{128}$$

and

$$\overleftrightarrow{J}_{CD} = \overleftrightarrow{E}_{CD}^{(02)} - \overleftrightarrow{D}_{CD}. \tag{129}$$

The tensor $\overleftrightarrow{E}^{(20)}$ is not represented in the effective Hamiltonian; however, it also has a physical meaning as the molecular magnetizability or, in certain contexts, referred to as magnetic susceptibility χ. The tensors $\overleftrightarrow{\sigma}_C$, and $\overleftrightarrow{J}_{CD}$ are molecular properties, which can generally be obtained in quantum mechanics as coefficients of a series expansion, like in equation (124). Within the framework of Rayleigh-Schrödinger time-independent perturbation theory, the following expressions exist for the first and second-order molecular properties:[188,189]

$$\frac{dE(\boldsymbol{x})}{dx_i} = \langle \Psi_0 | \frac{d\hat{H}}{dx_i} | \Psi_0 \rangle, \tag{130}$$

and

$$\frac{d^2E(\boldsymbol{x})}{dx_i dx_j} = \langle \Psi_0 | \frac{d^2\hat{H}}{dx_i dx_j} | \Psi_0 \rangle - 2 \sum_{n \neq 0} \frac{\langle \Psi_0 | \frac{d\hat{H}}{dx_i} | \Psi_n \rangle \langle \Psi_n | \frac{d\hat{H}}{dx_j} | \Psi_0 \rangle}{E_n - E_0}, \tag{131}$$

with the derivatives taken at x=0, which reflects, in our case, the effect of zero-field and zero magnetic moments. Equation (130) is a first-order property and requires only knowledge of the unperturbed reference state $|\Psi_0\rangle$. The first part of the second derivative has an analog first-order property expression. The second part contains a sum-over state's contribution from each excited state $|\Psi_n\rangle$. In case of magnetics properties, the first term is known as the diamagnetic contribution, and the second part is referred to as the paramagnetic part. [186]

The next goal is to find expressions for the derivatives which can be derived directly from the electronic Hamiltonian. For this purpose, the non-relativistic electronic Hamiltonian is extended under the influence of the external magnetic field, the nuclei's magnetic moments, and the electrons' magnetic moments. The corresponding Hamiltonian has the following form:

$$\begin{aligned} \hat{H}(\vec{B},\vec{\mu}) = \frac{1}{2}\sum_i \pi_i^2 - \sum_i \vec{m}_i \cdot \vec{B}^{\mathrm{tot}}(\boldsymbol{r}_i) - \sum_{iC} \frac{Z_C}{r_{iC}} + \\ \frac{1}{2}\sum_{i \neq j} \frac{1}{r_{ij}} + \frac{1}{2}\sum_{C \neq D} \frac{Z_C Z_D}{R_{CD}} - \sum_C \vec{\mu}_C \cdot \vec{B}^{\mathrm{tot}}(\boldsymbol{r}_i) + \sum_{C>D} \vec{\mu}_C \cdot \overleftrightarrow{D}_{CD} \cdot \vec{\mu}_C. \end{aligned} \tag{132}$$

The first term describes the kinetic energy,

$$\vec{\pi}_i = -i\nabla_i + \vec{A}^{\mathrm{tot}}(\boldsymbol{r}_i), \tag{133}$$

containing the vector potential $\vec{A}^{\mathrm{tot}}(\boldsymbol{r}_i)$ of electron *i*, that is constructed in a way that its curl or rotation reproduces the magnetic induction $\vec{B}^{\mathrm{tot}}(\boldsymbol{r}_i)$, that results from the nuclei's magnetic moments and the external magnetic field. The corresponding expression for the total magnetic induction is:

$$\vec{B}^{\mathrm{tot}}(\boldsymbol{r}_i) = \nabla_i \times \vec{A}^{\mathrm{tot}}(\boldsymbol{r}_i). \tag{134}$$

It is reasonable to split the vector potential into two parts, the first part arises from the contribution of the external field, and the other part is the sum of all nuclei contributions. For the vector potential and the magnetic induction, the following expressions are obtained:

$$\vec{A}^{\mathrm{tot}}(\boldsymbol{r}_i) = \vec{A}_O(\boldsymbol{r}_i) + \sum_C \vec{A}_C(\boldsymbol{r}_i), \tag{135}$$

$$\vec{B}^{\mathrm{tot}}(\boldsymbol{r}_i) = \vec{B}(\boldsymbol{r}_i) + \sum_C \vec{B}_C(\boldsymbol{r}_i). \tag{136}$$

The relation between the two latter expressions is given in equation (134). The vector potential that corresponds to a magnetic-induction vector, which represents a homogeneous external magnetic field, can be expressed as follow:

$$\vec{A}_O(\boldsymbol{r}_i) = \frac{1}{2}\vec{B} \times \boldsymbol{r}_{iO}. \tag{137}$$

The subscript O is the origin of the vector potential. It is also referred to as the gauge origin since the vector potential vanishes at this origin. It is very important to mention that the vector potential depends on the choice of the gauge origin, but the physical magnetic field due to the induction of B does not. However, as will be discussed later, the choice of the gauge origin is not completely free when the approximate wave function is used. The next contribution to the vector potential arises from the nuclear magnetic moments. A possible potential can be written in the following form:

$$\vec{A}_K(\boldsymbol{r}_i) = \alpha^2 \frac{\vec{\mu}_C \times \boldsymbol{r}_{iC}}{{r_{iC}}^3}. \tag{138}$$

In contrast to the external vector potential, here, the position of the nucleus is the preferred gauge origin, and for these terms, no errors regarding the choice of the gauge origin occur. To formulate an expression for the induced magnetic field of the nuclei, the curl of equation (133) has to be taken, and the following result is obtained:

$$\vec{B}_K(\boldsymbol{r}_i) = -\alpha^2 \frac{{r_{iC}}^2\overleftrightarrow{1} - 3\boldsymbol{r}_{iC}\boldsymbol{r}_{iC}^T}{{r_{iC}}^5}\overleftrightarrow{\mu}_C + \frac{8\pi a^2}{3}\delta(\boldsymbol{r}_{iC})\overleftrightarrow{\mu}_C, \tag{139}$$

where $\delta(\boldsymbol{r}_{iC})$ is the Dirac delta function. (139) can be interpreted as the internal magnetic field of a dipole. Regarding the delta function, its use is only in the case of a non-relativistic consideration necessary.[190] Through the delta-function, the case where $\boldsymbol{r}_{iC} = 0$ i.e. the magnetic field directly at the nuclear position is represented.[191] The first term gives rise to the dipole-dipole interaction between the electrons and the nucleus. The second term is the foundation for the Fermi-contact interaction, which will become important later in this part as a major contributor to the hyperfine coupling tensor. Since the electronic Hamiltonian was formulated under the influence of different magnetic inductions, it is now possible to calculate the derivatives to obtain the molecular properties.

First, the first-order paramagnetic interaction shall be derived. The first derivative is with respect to the magnetic induction field $\vec{B}$. At zero magnetic moments and field. It follows:

$$\frac{d\hat{H}}{d\vec{B}} = \hat{H}_B^{\text{orb}} + \hat{H}_B^{\text{spn}}, \tag{140}$$

whereby the first term connects the orbital motion of the electron via the orbital angular momentum operator with the external field in the following way:

$$\hat{H}_B^{\text{orb}} = \frac{1}{2}\sum_i \vec{l}_{iO} = -i\frac{1}{2}\sum_i \boldsymbol{r}_{iO} \times \nabla_i. \tag{141}$$

The second term couples the magnetic induction field to the magnetic moments of the electrons via:

$$\hat{H}_B^{\text{spn}} = -\sum_i \vec{m}_i = \sum_i \vec{s}_i. \tag{142}$$

These two interactions are also called Zeeman interactions. The focus will now be on the first-order interactions, which are a function of nuclear magnetic moments. The derivative with respect to $\vec{\mu}_C$ leads to:

$$\frac{d\hat{H}}{d\vec{\mu}_C} = \hat{H}_C^{\text{pso}} + \hat{H}_C^{\text{sd}} + \hat{H}_C^{\text{fc}}, \tag{143}$$

with the paramagnetic spin-orbit operator

$$\hat{H}_C^{\text{pso}} = \alpha^2 \sum_i \frac{\vec{l}_{iC}}{{r_{iC}}^3}, \tag{144}$$

the spin-dipole operator

$$\hat{H}_C^{\text{sd}} = \alpha^2 \sum_i \frac{{r_{iC}}^2 \vec{m}_i - 3(\vec{m}_i \cdot \boldsymbol{r}_{iC})\boldsymbol{r}_{iC},}{{r_{iC}}^5} \tag{145}$$

and finally, the Fermi-contact contribution:

$$\hat{H}_C^{\text{fc}} = -\frac{8\pi\alpha^2}{3}\sum_i \delta(\boldsymbol{r}_{iC})\vec{m}_i. \tag{146}$$

The paramagnetic spin-orbit operator couples the nuclear magnetic moments to the electron's orbital motion. The last two operators couple the nuclear magnetic moments with the spin of the electrons and are the most important contributions to the hyperfine coupling, which will be discussed in more detail (vide infra). At this point, the behavior of the first-order term in a closed-shell system shall be analyzed. In the case of field-dependent terms, the following relations are obtained:

$$\frac{d\hat{H}}{d\vec{B}}|\Psi^0\rangle = \begin{array}{cl} \hat{H}_B^{\text{orb}}|\Psi^0\rangle & \leftarrow\text{imaginary singlet} \\ +\hat{H}_B^{\text{spn}}|\Psi^0\rangle & \leftarrow \quad \text{vanishes} \end{array}, \tag{147}$$

where the first term leads to an imaginary singlet wave function, and the second term vanishes because the sum of spin magnetic moments equals zero. The hyperfine interactions show the following picture:

$$\frac{d\hat{H}}{d\vec{\mu}_C}|\Psi^0\rangle = \begin{array}{ll} \hat{H}_C^{\mathrm{pso}}|\Psi^0\rangle & \leftarrow \text{imaginary singlet} \\ \hat{H}_C^{\mathrm{sd}}|\Psi^0\rangle & \leftarrow \quad \text{real triplet} \\ \hat{H}_C^{\mathrm{fc}}|\Psi^0\rangle & \leftarrow \quad \text{real triplet} \end{array}, \tag{148}$$

where the paramagnetic spin-orbit operator causes an imaginary singlet wave function, and real triplet wave functions are obtained for electron spin-dependent terms. All these contributions vanish if they are applied to a closed-shell system.[186] The Hamiltonians, including the orbital angle-momentum operator, do not contribute since imaginary Hermitian operators lead to zero expectation values on a closed-shell system.[191] This phenomenon is also called the quenching of the orbital angular momentum. Since the hyperfine operators generate only triplet wave function, they do not affect a closed-shell system, in contrast to their behavior when the system is open-shell (see below).

Before the final expressions for the magnetic molecular properties can be derived, the definition of the second derivatives needs to be introduced. The first second-order diamagnetic interaction is obtained by derivation of the molecular electron Hamiltonian with respect to the external magnetic induction and the nuclear magnetic moments:

$$\frac{d^2\hat{H}}{d\vec{B}d\vec{\mu}_C} = -\overleftrightarrow{1} + \hat{H}_{BC}^{\mathrm{dia}}, \tag{149}$$

with the diamagnetic operator H_{BC}^{dia} defined as:

$$\hat{H}_{BC}^{\mathrm{dia}} = \frac{\alpha^2}{2}\sum_i \frac{(\boldsymbol{r}_{iO}\cdot\boldsymbol{r}_{iC})\overleftrightarrow{1} - \boldsymbol{r}_{iC}\boldsymbol{r}_{iO}^T}{{r_{iC}}^3}. \tag{150}$$

The second-order derivatives with respect to the nuclear magnetic moments lead to the following expressions:

$$\frac{d^2\hat{H}}{d\vec{\mu}_D d\vec{\mu}_C} = \overleftrightarrow{D}_{CD} + \hat{H}_{CD}^{\mathrm{dso}}, \tag{151}$$

with H_{CD}^{dso} as the diamagnetic spin-orbit operator defined as:

$$\hat{H}_{BC}^{\mathrm{dso}} = \frac{\alpha^4}{2}\sum_i \frac{(\boldsymbol{r}_{iO}\cdot\boldsymbol{r}_{iD})\overleftrightarrow{1} - \boldsymbol{r}_{iC}\boldsymbol{r}_{iL}^T}{{r_{iC}}^3{r_{iL}}^3}. \tag{152}$$

Finally, all derivatives are defined, and the expressions for the magnetic shielding tensor $\overleftrightarrow{\sigma}_K$ and the nuclear-nuclear coupling tensor $\overleftrightarrow{J}_{CD}$ can be obtained. To do this, the expressions for the derivatives are substituted into equation (130), considering the definitions made in equations (127) and (128). For the shielding tensor, the following expression is obtained:

$$\overleftrightarrow{\sigma}_K = \langle\Psi^0|\hat{H}_{BC}^{\mathrm{dia}}|\Psi^0\rangle - 2\sum_{ns\neq 0}\frac{\langle\Psi^0|\hat{H}_{BC}^{\mathrm{orb}}|\Psi^{ns}\rangle\langle\Psi^0|\left(\hat{H}_{BC}^{\mathrm{pso}}\right)^T|\Psi^{ns}\rangle}{E_{ns}-E_0}, \tag{153}$$

and for the nuclear spin-spin coupling tensor:

$$\overleftrightarrow{J}_{CD} = \langle\Psi^0|\hat{H}^{\mathrm{dso}}|\Psi^0\rangle - 2\sum \frac{\langle\Psi^0|\hat{H}_{DC}^{\mathrm{pso}}|\Psi^{ns}\rangle\langle\Psi^{ns}|\left(\hat{H}_{DC}^{\mathrm{pso}}\right)^T|\Psi^0\rangle}{} \tag{154}$$

$$-2\sum_{n_T}\frac{\langle\Psi^0|\hat{H}_C^{\mathrm{sd}}+\hat{H}_C^{\mathrm{fc}}|\Psi^T\rangle\langle\Psi^T|\left(\hat{H}_D^{\mathrm{sd}}\right)^T|\Psi^0\rangle}{E_{ns}-E_0}.$$

These relations are also known as the Ramsey expressions. Ψ^{ns} is a single excited state and Ψ^{T} denotes a triplet excited state. The trace of the shielding sensor then leads to the experimentally observed shielding:

$$\sigma=\frac{1}{3}\mathrm{Tr}(\overleftrightarrow{\sigma}). \tag{155}$$

As long as solutions are considered, only isotropic contributions to the shielding tensor are relevant because the anisotropic contributions are canceled due to the fast movements and rotations. The g-tensor from equation (123) can also be obtained as a second-order molecular property. This is done by the second derivative of the energy w.r.t. the electron spin and the magnetic field in the following way:

$$\overleftrightarrow{g}=\left.\frac{d^2E\left(\vec{B},\vec{S}\right)}{d\vec{B}d\vec{S}}\right|_{\vec{B}=0,\vec{S}=0}, \tag{156}$$

in contrast to the molecular properties presented above, the hyperfine-coupling tensor in equation (123) can be obtained as a first-order molecular property with respect to the nuclear .magnetic moment. At this point, it can be referred to equations (143) and (145), and the following expression for the hyperfine-coupling tensor is obtained:

$$\overleftrightarrow{A}=\langle\Psi^0|\hat{H}^{\mathrm{fc}}+\hat{H}^{\mathrm{sd}}|\Psi^0\rangle. \tag{157}$$

For light nuclei, the spin-orbit coupling is neglectable. The Fermi-contact term is responsible for the isotropic hyperfine coupling, whereas the spin-dipolar term normally vanishes in bulk solution. For the isotropic hyperfine-coupling constant of the nucleus C it follows: [192]

$$A_C^{\mathrm{iso}}=\langle\Psi^0|-\frac{8\pi\alpha^2\mu_N g_C}{3}\sum_i\delta(\boldsymbol{r}_{iC})\vec{m}_i\,|\Psi^0\rangle. \tag{158}$$

The term $\mu_N g_C$ needs to be added since in the effective Spin-Hamiltonian in equation (123), only the nuclear-spin-operator occurs, but the differentiation was w.r.t the nuclear magnetic moment. Since the spins are aligned along the magnetic field direction (normally called the z-axis) in the experiment, only the s_z spin is considered. We obtain:

$$A_C^{\mathrm{iso}}=\frac{8\pi\alpha^2\mu_N g_C\mu_B g_e}{3}\langle\Psi^0|\sum_i\delta(\boldsymbol{r}_{iC})\hat{s}_{zi}\,|\Psi^0\rangle. \tag{159}$$

In equation (137), the vector potential representing the external magnetic field via eq. (137) was defined. In the formulation of equation (137), an arbitrary origin of the gauge was chosen. In the context of gauge theory, this is a legitimate process and should not influence the observables of the system, such as electronic energy, nuclear shielding constants, or g-tensors. However, it turns out that this statement did not necessarily apply to approximate wave functions.[193] To guarantee the independence from the gauge origin, the wave function or, more concretely, the basis functions need to be modified. To do this, first, a gauge transformed vector potential is defined in the following way:[191]

$$\vec{A}_{O'}(\boldsymbol{r}) = \vec{A}_O(\boldsymbol{r}) + \nabla f(\boldsymbol{r}), \tag{160}$$

where $f(\boldsymbol{r})$ is a scalar function. At this point, it is necessary to notice that for any scalar function, the curl of the gradient equals zero:

$$\nabla \times \nabla f(\boldsymbol{r}) = 0. \tag{161}$$

Due to this, the so-called gauge transformation in equation (160) does not affect the induced magnetic field when the relation in equation (134) is applied. The gauge transformation leads to a gauge-transformed wave function in the following manner:

$$\psi'^{(r)} = \exp\big(-f(\boldsymbol{r})\big)\,\psi(\boldsymbol{r}). \tag{162}$$

The scalar function acts here as a complex phase factor. A scalar function could have the following form:

$$f(\boldsymbol{r}) = \frac{1}{2}\vec{B} \times (\boldsymbol{O} - \boldsymbol{O}') \cdot \boldsymbol{r}. \tag{163}$$

With this type of gauge transformation, the invariance of the energy and other molecular properties is maintained. It can be stated that the following relationship is obtained for exact wavefunctions:

$$\langle \psi' | \hat{H}' | \psi' \rangle = \langle \psi | \hat{H} | \psi \rangle. \tag{164}$$

Now the question arises of how the picture looks like in the case of approximate wave functions. Here, the results depend on the choice of the gauge origin.[194] It can be seen that the gauge error depends on the distance between the wave function and the gauge origin. Many different approaches have been developed to solve this problem in the past. Among the first is the so-called Individual Gauge for Localized Orbitals (IGLO)[195] or LORGs[196]. Among the more recent methods are those in which an explicit dependence of the basic functions on the magnetic field is introduced, analogous to equation (162). These methods are known as London Atomic Orbitals (LAO)[197,198] or Gauge Invariant Atomic Orbitals (GIAO).[199] The magnetic field dependence is again realized by introducing a complex phase factor. It follows:

$$X_C(\boldsymbol{r} - \boldsymbol{R}_C) = \exp\big(-i\vec{A}_C \cdot \boldsymbol{r}\big)\,\chi_C(\boldsymbol{r} - \boldsymbol{R}_C), \tag{165}$$

where $\vec{A}_C$ is defined as:

$$\vec{A}_C = \frac{1}{2}\vec{B} \times (\boldsymbol{R}_C - \boldsymbol{R}_O). \tag{166}$$

X_C is the gauge-independent atomic orbital. However, the term gauge-independent is kind of misleading since an explicit dependence on the gauge origin is introduced into the atomic orbitals. For the GIAO method, it can be shown that integrals over the GIAOs are independent of the gauge origin. For the overlap and potential energy, this can be easily shown, and it follows for the overlap integral:

$$\langle X_C|X_D\rangle = \langle \chi_C|\exp\big(i(\vec{A}_C-\vec{A}_D)\cdot\boldsymbol{r}\big)\,|\chi_D\rangle. \tag{167}$$

And for the potential energy, the following expression is obtained:

$$\langle X_C|\hat{V}|X_D\rangle = \langle \chi_C|\exp\big(i(\vec{A}_C-\vec{A}_D)\cdot\boldsymbol{r}\big)\,\hat{V}|\chi_D\rangle. \tag{168}$$

The term in the exponential function is not dependent on the gauge origin since:

$$\vec{A}_C-\vec{A}_D = \frac{1}{2}\vec{B}\times(\boldsymbol{R}_C-\boldsymbol{R}_D). \tag{169}$$

In the case of kinetic energy, the solution is a bit more tricky. The expression for the kinetic energy is:[198–200]

$$\begin{aligned}\langle X_C|\hat{T}|X_D\rangle &= \langle X_C|\frac{1}{2}\pi^2|X_D\rangle \\ &= \langle \chi_C|\exp\big(i\vec{A}_C\cdot\boldsymbol{r}\big)\frac{1}{2}\big[-i\nabla+\vec{A}(\boldsymbol{r})\big]\exp\big(i\vec{A}_D\cdot\boldsymbol{r}\big)\,|\chi_D\rangle.\end{aligned} \tag{170}$$

To get a gauge-independent expression of the integral, at first, the derivatives of X_D need to be calculated via:

$$\big[-i\nabla+\vec{A}(\boldsymbol{r})\big]X_D = \big[-i\nabla+\vec{A}(\boldsymbol{r})\big]\exp\big(i\vec{A}_D\cdot\boldsymbol{r}\big)\,|\chi_D\rangle. \tag{171}$$

After multiple applies of the product rule of derivation and the use of the 2nd binominal formula, the following relation is obtained:[200]

$$\big[-i\nabla+\vec{A}(\boldsymbol{r})\big]X_D = \exp\big(-i\vec{A}_D\cdot\boldsymbol{r}\big)\big[-i\nabla+\big(\vec{A}(\boldsymbol{r})-\vec{A}_D\big)\big]\chi_D. \tag{172}$$

Reinserting into equation 170) leads to:

$$\langle X_C|\hat{T}|X_D\rangle = \langle \chi_C|\exp\big(i(\vec{A}_C-\vec{A}_D)\cdot\boldsymbol{r}\big)\frac{1}{2}\big[-i\nabla+\big(\vec{A}(\boldsymbol{r})-\vec{A}_D\big)\big]|\chi_D\rangle. \tag{173}$$

In this way, the integral of the kinetic energy is also gauge-independent. A pragmatic ansatz to handle the gauge-origin dependence of electronic g-tensor calculations was introduced by Glasbrenner[194], where the gauge-origin is chosen to be the spin density center. This approach is especially reliable for molecules with one spin center.

2.4 Overview of computational spectroscopy in solution

While in the previous chapters, a detailed derivation was given, how, on the one hand, the solvent can be described within computational methods and, on the other hand, how spectroscopic observables can be calculated; in this chapter, an overview of practical applications will be given. One of the first questions that must be asked in the QC-supported computation of spectroscopic observables is the choice of an appropriate level of theory. Here always, the question arises, how much costs in the form of computing power and time can be spent. Usually, for inexpensive but still efficient calculations, different DFT methods are used, whereas, for more accurate results, MP-n, double hybrid functional[201], or even coupled-cluster methods are utilized.[191] A further question addresses the problem of accurate solvent representation if either explicit or implicit models should be used. A third question deals with the number of conformers used for the calculation and corresponding creation of these.

In the case of vibrational frequencies, normally, all levels of theories cause specific intrinsic errors in the calculation results.[202] The reasons are that all calculations are approximate and, on the other hand, that the vibrational frequencies are calculated based on a harmonic potential, which is not fully correct. In reality, bond stretches are better described by the so-called Morse potentials, including the anharmonicity nature of the stretching. Normally, the calculated wavenumbers are too high; therefore, specific scaling factors were determined and can be accessed on the computational chemistry comparison and benchmark database weak to match the calculated results with experimental data. Here, a relatively strong dependence on the level of theory is observed, whereas the basis set influence is small. [202] Regarding the absolute accuracy besides scaling factors, CCSD[203] and CCSD(T)[204] methods show the best performances. However, these results are almost limited to small molecules. In recent years double-hybrid methods have become a great alternative to calculated vibrational frequencies with nearly coupled-cluster quality.[205]

Regarding calculating NMR chemical shielding constants, formalisms for Hartree-Fock[206], DFT[101,207–209], MP2[210–212], and coupled cluster methods[213–215] were developed. Especially for the correlated wave function methods, Gauss derived the corresponding procedures in the 1990s.[210–215] A good balance between computational costs and the accuracy of the results was shown for MP2, which will be employed throughout this thesis for NMR chemical shielding calculations. [216] A popular alternative is the DFT method, where the problem arises that, most commonly, DFT variants are not dependent on the external magnetic field, and thus the produced exchange and correlation energies are unphysically constant in the presence of the external field.[217] In the past explicit field dependence was introduced[207]; however, the more commonly used alternative is the so-called current-density functional (CDFT)[208,218] Although CDFT shows to have a significant impact on the chemical shielding,[209,219] the obtained shielding constants are not automatically greatly better than those calculated with standard DFT methods.[220] Therefore, for reasons of practicality, the "normal" DFT methods are mostly used, whereby the results' quality is strongly dependent on the functional.[216,221] Good results are obtained, e.g., with the OLYP and OPBE functionals for determining ^{13}C and ^{15}N nuclear shielding constants.[222] The meta-GGA M06L also showed very good results in the past. [223] Going from so-called pure functionals to hybrid functionals does not necessarily improve the results, especially for ^{15}N shielding constants.[221] This is probably due to the inheritance of the bad performance of the HF method in describing the ^{15}N shielding constants.[217] Recently, Stoychev *et al.* made predictions of nuclear shielding tensors by applying the double-hybrid DFT methods[224], representing the 5th rung of the so-called Jacob's ladder of DFT approaches.[225] Here, DSD-PBEP86 showed the smallest deviation from the benchmark dataset with 1.9%, which performs significantly better than MP2 (4.1%) and the meta-GGA M06L (5.4%).[217] For further improvement of the calculations with DFT methods, Jensen developed an optimized polarization consistent basis set, the pcS-n basis set.[226,227] This one and derivatives[228] are nowadays widely used in benchmarking the performance of computing nuclear shielding.[217]

Pople and Beveridge did one of the first attempts to describe EPR isotropic hyperfine coupling constants with semi-empirical methods using an approximated self-consistent molecular-orbital theory. This approach was termed intermediate neglect of differential overlap (INDO).[229,230] These results were quite promising; however, when in the late 1970s, the computational power became available to make calculations with the unrestricted HF (UHF) formalisms, the results were sobering.[231,232] In the ongoing years, it became clear that for the accurate representation of HFCCs, high-level *ab initio* methods (like multireference configuration interaction) with large basis sets are necessary.[233] Thus, the evaluation of HFCCs was limited to small molecules. Until the 1990s, the calculations of g-tensors were sparse and were based on *ab initio* methods with small basis sets and semi-empirical methods.[234] With the rapid development of DFT methods in the early 1990s, accurate coupling constants and g-tensors calculations became possible.[125–128,235–239] Regarding the calculation of HFCCs, coupled cluster approaches have been found to be always superior compared to DFT methods.[240] In recent years, the development of DLPNO-CCSD[132] (see chapter 4.2) and double-hybrid methods[127,241] led to more time-saving calculations of HFCCs, while still maintaining a good level of accuracy compared to CCSD methods.

Continuum models have a long history in calculating spectroscopic observables in solution.[138] In 1998, Cammi[242,243] developed the first formulation for calculating NMR spectroscopic parameters in solution using the PCM formalism. In parallel, this group developed the implementation of analytical second derivatives within HF and DFT calculations and calculated the solvatochromic IR shifts of some compounds.[244] Combining PCM solvation and the computation of EPR HFCC parameters was done by Barone and coworkers.[245–247] While the different PCM variants are implemented in many QC codes and analytic second derivatives are available, the continuum description of the solvent causes intrinsic errors due to ignoring directional interactions, higher-order multipoles, and dynamic solvent contributions.[248,249] An extended continuum method that includes the directional interactions is the direct COSMO model for real solvents (direct COSMO-RS[250–252]). It was applied in calculating solvent shifts on g-tensors and the results were in good agreement with calculations that used a supermolecule approach with explicit water molecules.[251] Pressure-dependent vibrational spectroscopy was calculated in non-aqueous solvents by applying the X-PCM approach (see 2.1.1).[21,22]

Further alternative approaches based on statistical solvent models are the RISM-SCF[155,167] or EC-RISM[24] methods, which combine the RISM solvation theory with a QC description of the solvent. (see 2.1.4) The field of solvatochromism based on the time-dependent density functional theory (TD-DFT) was one of the main applications of the RISM-SCF variants.[174,253,254] For example, Nishiyama studied the stokes shifts of coumarin 153 in 13 different organic solvents using 1D RISM SCF. The deviations stoke shifts magnitudes were 400 cm^{-1} less than the experimentally observed ones.[253] An alternative approach to describe the stokes shifts more reliably was developed by Yokogawa[255], in which explicitly spatial electron density distribution was included in the RISM-SCF calculations, leading to the RISM-SCF-SEDD approach. The experimentally observed data could be reproduced by applying it to stokes shifts of acetone, indole, and 5-cyanoindole.[255] Further usage of the RISM-SCF-SEDD method was done by Hirano[256] and Suda[257]. TD-DFT calculations combined with the 3D RISM-SCF formalism were realized by Kaminiski[258] and Yang[259]. Recently, Reimann and Kaupp did computations on the ground- and excited states properties of acetone in water and betaine 30 in various solvents. The obtained results agree very well with experimental reference data and with more costly QM/MM-based MD simulations.[173]

Kast and co-workers[26,27,29,248] (including the author of this thesis) did the first NMR calculations based on EC-RISM calculations. Frach[248] made the first steps in calculating the chemical shifts of NMA at ambient pressure. It was shown that EC-RISM leads to a considerable and systematic improvement over PCM calculations. A possible explanation for this improvement could be the retaining of the solvent granularity in comparison to PCM. In a further publication, the pressure-dependent shifts of NMA relative to the DSS NMR standard were calculated.[26] It was shown that the effect of pressure on the DSS's shielding constants is much smaller than on the shielding constants of the ^{1}H amide proton. Thus, it was concluded that DSS is a very good NMR standard for pressure-dependent NMR measurements.

Further pressure-dependent NMR studies were calculated for TMAO[27] and purine nucleotides.[28,29] A limitation of all these calculations is that the structures used are based on PCM-optimized geometries at ambient conditions, and the effect of pressure on the geometries is not included. To tackle this issue, pressure-dependent geometry optimizations were applied in this thesis (details see 2.1.4 and 4.3.3). 3D-RISM SCF with consistently optimized geometries was used to calculate ^{17}O chemical shifts of *N*-Methylformamide in different solvents.[173] A second spectroscopy method analyzed with EC-RISM is vibrational spectroscopy, where the pressure-dependent vibrational wavenumber shifts of TMAO were calculated and compared to the experimental reference data. A detailed discussion of this study can be found in chapters 3.1 and 4.1.

Another limitation of all these EC-RISM studies is that using only a limited number of optimized geometries and explicit water molecules were neglected. However, for some spectroscopic observables, it is essential to include a thermal averaged ensemble of structures in the calculations. For example, the nitrogen HFCC of nitroxides depends strongly on the out-of-plane movement of the nitroxide and the surrounding solvent molecules.[260–262] To generate a reliable statistical ensemble, one can apply full AIMD[263], QM/MM[120] methods, or force field based molecular mechanics[118]. AIMD simulations were performed, for example, to calculate the dynamical effect on the electronic *g*-tensors of benzosemiquinone in aqueous solution under ambient conditions.[119] In this thesis also, an AIMD-based conformational ensemble was used for calculating EPR spectroscopic observables, and it was shown that the quality of the results is improved compared to single optimized structures (see 4.2.)

If force field based methods should be used to describe nitroxide in solutions, reliable parameters must be developed. Most parameterizations were targeted to specific small molecules and lacked generality.[117,264–266] Especially the accurate description of the out-of-plane motion of the nitroxide group was neglected. Stendardo developed a general nitroxide force field based on the AMBER99 force field.[118] In this work, new atom types were included, and the relevant parameters were fitted on geometries, potential energy surfaces, and vibrational frequencies at the DFT level for several nitroxides. This adapted force field is capable of providing reliable geometries, energies, and EPR spectroscopic observables.[118]

A thermally averaged ensemble can also improve the quality of NMR parameter calculations, and AIMD simulations are very well suited for this purpose. Recently, a detailed review of the combination of AIMD and NMR parameter calculation was written by Mazurek.[267] One particular example of this combination was performed by Exner[268], where the small peptide NMA was investigated in an aqueous solution. Here, it was shown that with conformational sampling and by taking explicit water molecules into account, a very good agreement between calculated and experimental shifts is observed. Another QM/MM study by Flaig showed that by including up to about 300-1200 atoms into the QM region, convergence in calculating NMR shielding tensors is achieved.[269]

IR wavenumbers based on QC calculations are normally determined by applying the normal-mode analysis (NMA). Alternatively, for determining computational-based IR wavenumbers, one can also use an MD simulation or AIMD simulation, respectively.[270,271] The use of a proper sampling of fluctuations becomes indispensable when finite temperature effects, as the broadening of band or shifting occur due to the fluctuations.[270] AIMD simulations can be used to calculate the linear absorption cross-sections since it relates to the Fourier transform of the dipole autocorrelation function.[270–273] A further advantage of utilizing an AIMD simulation to determine IR spectra is that extreme conditions can also be considered. For example, AIMD was also used to calculate the pressure-dependent IR spectrum of TMAO at 1 bar and 10 kbar in an aqueous solution.[94]

3 Methodology

3.1 Calculation of IR frequencies with EC-RISM

In 2016, P. Kibies and S. M. Kast made the first attempts to perform IR calculations with EC-RISM.[94,176] The system under investigation was the small osmolyte TMAO, focusing on the pressure-dependent changes of the IR frequencies. A blue shift of the main observed frequencies with increasing pressure was observed in the experiment.[94] By performing a pressure-dependent AIMD simulation, this trend could also be confirmed by theoretical methods (see chapter 2.4). At this point, it should be elucidated if EC-RISM can also reproduce this trend. However, a big disadvantage of the EC-RISM formalism is the lack of analytical expressions for the second derivatives of the excess chemical potential with respect to the displacement coordinates. Therefore, a numerical alternative must be introduced at this point (For a detailed description, see the next two subchapters). The calculated frequencies with EC-RISM could correctly reflect the real pressure-dependent trend but led to a massive overestimation of the pressure dependence. There are two possible causes for this:

i. The EC-RISM calculations were made with a legacy protocol, using χ_{Sim}[18] and point charges to calculate the electrostatic interactions instead of the exact electrostatic potential
ii. It is plausible to assume that the process of vibration is much faster, compared to the relaxation of the surrounding water, due to the vibration

IR frequencies with different solvent susceptibilities and electrostatic variants were calculated to determine the origin of these possible problems. On the other hand, a new approach is presented in which no solvent relaxation due to vibration is considered. The test systems used here are, on the one hand, the previously investigated pressure-dependent modes of TMAO (by P. Kibies and published in ref[94]) and, on the other hand, the cyanide anion since there is only a single vibrational mode.

3.1.1 Equilibrium case

To calculate IR frequencies numerically with EC-RISM in an equilibrium scenario, the following workflow, developed by P. Kibies and S. M. Kast, is used:

1. QC Optimization of the molecule of interest with PCM solvation
2. Normal modes determination of the optimized structure
3. Displace the atoms along the normal vectors starting from the equilibrium structures and select two structures in every direction
4. EC-RISM calculations for all displaced structures at different pressures

Parabolic fit or numerical five-point stencil to obtain the second derivative of the energy w.r.t. displacement coordinates

In practice, the Cartesian force constants $\boldsymbol{F}$ are first extracted from a Gaussian 16 revB01 calculation. Since mass-weighted quantities are required for the calculation of force constants and normal modes, the individual atomic masses, which are also used in the Gaussian calculations, are extracted and collected in a diagonal mass matrix via:

$$\boldsymbol{M} = \begin{pmatrix} m_1 & 0 & \cdots & 0 \\ 0 & m_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & m_N \end{pmatrix} \tag{174}$$

Then the mass-weighted Hessian is constructed via:

$$\boldsymbol{F}_M = \boldsymbol{M}^{1/2}\boldsymbol{F}\boldsymbol{M}^{1/2} \tag{175}$$

Subsequently, the corresponding eigenvalues and eigenvectors are calculated. The eigenvector can be associated with the mass-weighted cartesian displacement coordinates for the individual normal modes. To obtain the displacements of every atom, the eigenvectors are multiplied with the inverse of the square root of the diagonal mass matrix. This procedure leads to cartesian unit displacement vectors, which are then scaled by the amplitudes $-A_{\text{max}}$, $-A$, A, A_{max}. These scaled vectors are then added to the original equilibrium PCM-optimized geometry. With the four displaced vectors and the equilibrium structure, it is now possible to calculate the second derivatives, either numerically with a five-point stencil:

$$k(G_{\text{sol}}) = \frac{-G(\boldsymbol{x} - 2d\boldsymbol{q}_i) + 16G(\boldsymbol{x} - d\boldsymbol{q}_i) - 30G(\boldsymbol{x}) + 16G(\boldsymbol{x} + d\boldsymbol{q}_i) - G(\boldsymbol{x} + 2d\boldsymbol{q}_i)}{12d^2} \tag{176}$$

where $k(G_{\text{sol}})$ is the force constant associated with normal mode i, d is the step size of the displacement, $\boldsymbol{q}_i$ is the displacement vector of normal mode *i,* which is added to the equilibrium geometry. The alternative would be to perform a quadratic fit with the five structures and obtain the force constant as the second derivative of the received parabolic equation. The frequency is finally obtained via:

$$\nu(k, \mu) = \frac{\sqrt{\frac{k}{\mu}}}{2\pi} \tag{177}$$

where μ is the reduced mass according to the Gaussian[274] conventions (see J.W. Ochterski[185])

3.1.2 Non-equilibrium case

The non-equilibrium approach aims to develop a formalism where the solvent structure cannot adapt to the displacements due to vibration. In a regular EC-RISM calculation, an iterative cycle is performed where the solvent and the solute can polarize each other, i.e., the solvent adapts to the perturbation of the solute. To avoid this effect, the following workflow was developed by S. M. Kast developed:

1. The first three steps of the equilibrium approach are reused.
2. Full equilibrium EC-RISM calculation with the PCM optimized equilibrium structure.
3. Extracting the solvent structure around the equilibrium structure and subsequently using this solvent structure as a perturbation for single-point calculations of the displaced structures.
4. To extract the intramolecular energy of the displaced structures under the influence of solvent; an evaluation of the purely intramolecular Hamiltonian on the solvated wave function is needed.
5. The excess chemical potential is calculated via 1D RISM:
    a. The partial charges of the equilibrium structure under the perturbation of the solvent are needed, together with LJ-Parameters, that were used in the 3D RISM calculation.
    b. A full *uv*-1D RISM calculation for the equilibrium structure is performed, and the excess chemical potential is calculated.
    c. For the displaced structures, only the relative change of the excess chemical potential is calculated.
6. Excess chemical potential and intramolecular energy are added, and the frequencies can be calculated with the procedures mentioned above.

Regarding step 2, e.g., for the QC program Gaussian the solvent structure can be stored in a checkpoint file and then efficiently read in for the single-point calculations of the displaced structures. The intramolecular energy can be calculated according to the following:

$$E_{\mathrm{sol}} = \langle \Psi_{\mathrm{sol}} | \hat{H}_{\mathrm{ne}} + \hat{H}_{\mathrm{ee}} + \hat{H}_{\mathrm{nn}} | \Psi_{\mathrm{sol}} \rangle. \tag{178}$$

This is identical to the calculation of intramolecular E_{sol} energy in equation (64). Unfortunately, identical 3D RISM cannot be used to calculate the excess chemical potential at this point since even the tiniest displacements cause singularities in the solute-solvent interactions, which appear in a product with the frozen $g(\mathrm{r})$ under the integral if an approximation of the closure is applied.[159] Instead, we must evaluate the excess chemical potential using a 1D RISM-*uv* calculation. The 1D RISM- *uv* calculation for the equilibrium structure is the basis again, and for the other structures, only the relative changes are calculated. It is assumed that the direct correlation function provides identical values for small displacements of the geometry. Therefore, the change of the excess chemical potential is determined by the geometry alone, which is encoded by the ω function.[154] The equation (50) can then be used to calculate this relative change in excess chemical potential. In the end, the intramolecular energies and the excess chemical potentials are added, and the known methods for calculating the frequencies can be used again. The method is referred to as M1 in the context of this chapter. Additionally, a second method is introduced, which is not a real non-equilibrium approach, but where instead, the calculations of the excess chemical potential are performed via 1D RISM, where the point charges are taken from the optimized structure EC-RISM calculations. This method will be termed M2.

3.1.3 Computational details

For TMAO, geometry optimizations were performed with B3LYP/6-311+G(d,p) and MP2/6-311+G(d,p) in PCM solvent. These optimized structures were used as basis structures. Additionally, frequency calculations were performed on these structures to obtain normal modes, reduced masses, and IR intensities. The normalized Cartesian displacement vectors along the normal modes were calculated using a *Mathematica* script originally developed by P. Kibies and S. Kast. Four displaced structures were generated with displacement steps of 0.02 Å relative to the basis structure. For the equilibrium approach, EC-RISM calculations with MP2/6-311+G(d,p)/PSE3 or B3LYP/6-311+G(d,p)/PSE-3 on the corresponding optimized structures, respectively. For the EC-RISM calculations, a cubic box of 240^3 grid points with a spacing of 0.1 Å between two grid points was utilized. For 3D RISM, the convergence criteria were set to 10^{-6} w.r.t to the change in direct correlation function in successive steps. The EC-RISM cycle converged when the energy difference between two successive steps was below 10^{-4} kcal/mol. Every iteration was performed with this level of theory for the B3LYP/6-311+G(d,p) calculations. In the case of MP2/6-311+G(d,p), the iteration used HF/6-311+G(d,p), and only in the final calculation MP2/6-311+G(d,p) were utilized. Two approaches regarding the electrostatics calculation were tested; in the first one, the atom-centered partial charges, which were determined using the ChelpG algorithm, were utilized to calculate the electrostatic interactions. The full correct QC electrostatic potential was used in the second one. The number of point charges representing the solvent in the QC calculations was reduced by combining distant point charges using a Voronoi algorithm.[275] For TMAO, the optimized TMAO force field version 3 (see Table 36) taken from Ref[18,276] were utilized. In the case of cyanide, the optimized parameters from Lee et al. [277] were used. The LJ parameters can be found in 6.9.

Additionally, for the calculation of solvent-mediated contribution to the free energy under non-equilibrium conditions, 1D RISM calculations were performed. Here, the atom-centered point charges from the converged equilibrium EC-RISM calculations on the PCM optimized structure. For both 1D RISM and 3D RISM, the temperature was set to 298.15 K, and the solvent susceptibilities were taken from Ref[18]. The partial charges for the 1D RISM calculations can be found in SI_4.1 (SI is accessible in the electronic appendix). The HNC closure was used, and a convergence criterium of 10^{-8} difference in the direct correlation function between two successive steps. A logarithmic grid of 512 points ranging from 0.00598 Å to a maximum distance of 164.02 Å was used.

3.2 EPR-parameters of the pH-dependent nitroxide probe HMI

In this thesis, the EPR parameters of nitroxide (3R,4S)-2,2,3,4,5,5-hexamethylimidazolidin-1-oxyl (HMI) in solution were analyzed by different theoretical methods. From the above-described dependency of the EPR parameters on the geometry of the nitroxide, it is necessary to model an accurate conformational ensemble previous to the parameter calculations. Therefore, AIMD simulations in water were performed by B. Sharma (more details below) to generate a sufficient ensemble of structures. This simulation generated snapshots and was subsequently used for the EPR parameter calculations. These calculations were performed with three different solvation models EC-RISM, CPCM, and a QM/MM approach (done by V. A. Tran[132]). This project was part of the RESOLV collaboration. One of the subjects of this project is to analyze local solvent effects, where HMI is a very good sample to study. The HFCC parameter of HMI changes significantly when the molecule is protonated (The protonated version of HMI will be called HHMI throughout this thesis). Thus, the EPR parameters of HHMI were also analyzed. For the analysis of the results, some aspects of EPR parameters will be shown in the following.

As mentioned in chapter 0, nitroxide spin probes are one tool of choice to characterize the specific properties of the microenvironment around an organic material in which the nitroxide probe is incorporated. On the one hand, the characteristic parameters are the g-values, which are related to the chemical shielding due to the surroundings around the spin, and on the other hand, the hyperfine interactions reflect the interactions between electron and nuclear spins. In chapter 2.3.2, the concept of the effective spin Hamiltonian was introduced, allowing a full description of a measured EPR or NMR spectrum. In the case of nitroxide spin probes, the spin Hamiltonian generally has the following form:

$$\hat{H} = \mu_N \vec{B}\overleftrightarrow{g}\vec{S} + \vec{S}\overleftrightarrow{A}\vec{I} + \vec{I}P\vec{I}. \tag{179}$$

The last term describes the nuclear quadrupole interactions, which can be neglected in molecules synthesized with ^{15}N isotopes. The parameters $\overleftrightarrow{g}$ and $\overleftrightarrow{A}$ can be obtained by fitting the experimental spectrum to the spin Hamiltonian.

In EPR spectroscopy, different frequencies are usually taken for the constant irradiated frequency field, so-called bands, depending on which spectroscopic parameter is the object of interest. For example, if there is only interest in the isotropic hyperfine coupling constant, a so-called X-band (8.0–12.0 GHz) is enough to resolve the hyperfine coupling in solution at ambient conditions. If, instead, a resolution of the g-anisotropy is to be determined in frozen environments, a much higher band must be chosen, e.g., the Q-band (35 GHz) or W-band (95 GHz).

EPR spectroscopy is particularly well suited to resolve the microenvironment since the spectroscopic parameters depend very strongly on the polarity of the surrounding environment. In the case of nitroxide spin probes, this behavior can be described very well using an approximate LCAO approach. E.g., the principal components of the g-tensor can be described as:[100]

$$g_{xx} \approx g_e + 2\lambda^O \rho_\pi^O c_{ny}^2 / (E_{n\pi*} - E_0), \tag{180}$$

$$g_{yy} \approx g_e + 2\lambda^O \rho_\pi^O c_{nx}^2 / (E_{n\pi*} - E_0) \tag{181}$$

and

$$g_{zz} \approx g_e \approx 2.0023. \tag{182}$$

Here g_{xx}, g_{yy}, and g_{zz} are the principal components of the g-tensor, g_e is the g-factor of the free electron, λ^O is the spin-orbit-coupling constant of the electrons on the oxygen, ρ_π^O is the spin density in the π^* orbital and c_{nx}^2, c_{ny}^2 are the coefficients of the $2p_y$ and $2p_x$ orbitals in the lone-pair orbital on the oxygen. From this simple model, it can be stated that the g_{zz} not deviates from the free-electron value, and the other two values increase compared to the free-electron value. This model can also predict the influence of hydrogen bonds on the g_x component. Due to hydrogen bonding the c_{ny}^2 decreases, which arises from the delocalization of the lone-pair electron into the H_2O orbitals and a decrease of ρ_π^O is observed.

To describe the influence of different structural and environmental effects on the isotropic hyperfine coupling constant (the terms HFCC or A^{iso} will be used to abbreviate this expression within this work) of the nitrogen from eq (159), it is known that it depends strongly on the spin density at the nucleus N, which can be described via:

$$\rho_N = \sum_{\mu,\nu} P_{\mu,\nu}^{\alpha-\beta} \langle \phi_\mu(\mathbf{r}) | \delta(\mathbf{r} - \mathbf{r}_N) | \phi_\nu(\mathbf{r}) \rangle, \tag{183}$$

where $P_{\mu,\nu}^{\alpha-\beta}$ is the difference between the density matrices for an electron with α and β spins, the so-called spin density. From theoretical papers published in the 1950 two major contributions to ρ_N were derived.[128,278–280] The first one is the so-called direct contribution from the nucleus's spin density. This effect is due to the nominally unpaired orbital (also called the singly occupied molecular orbital (SOMO)), which is delocalized. The second contribution is the spin polarization or indirect contribution, which originates in the interactions of the σ electrons with the unpaired spin electron.[264] In absolute terms, the latter contribution is smaller than the direct one. However, the second term becomes important when the nucleus lies in the proximity or exactly in a nodal plane of the SOMO, leading to a vanishing of the direct term. In the case of nitroxides, the SOMO is basically the antibonding π^* orbital. The larger the contribution of the nitrogen atomic orbital to the antibonding π^* orbital, the larger the spin density is on the nitrogen and vice versa. However, the SOMO is also influenced by the surrounding orbitals and especially on the doubly occupied π bonding orbital. To address this problem, it is useful to consider the nitroxide resonance structures, that are depicted in Figure 4.

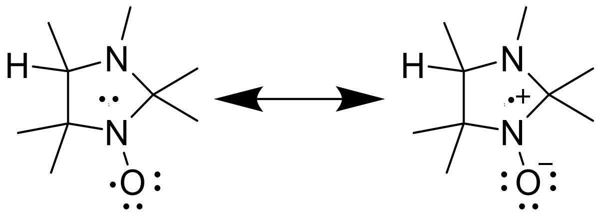

Figure 4: Resonance structures of the N-O nitroxide moiety in HMI.

In the left resonance structure, the unpaired electron is located on the O-atom, and the *N*-atom possesses a lone pair, which corresponds to the π bonding orbital. In this situation, a pyramidal structure is preferred, for example, in NH_3. In this case, the SOMO is constituted by the atom's singly occupied p orbital. In the right resonance structure, the N atom has a single electron in the p-like orbital: in this constellation, the planar geometry is favored since it minimizes the repulsion between the three σ bonds on the nitrogen. Due to the location of the free electron on the nitrogen, a higher HFCC is obtained for this structure. However, if small deviations from the planar geometry occur due to thermal fluctuations, the s orbital contributes to the SOMO, and the spin density even increases on the N atom, and therefore the HFCC increases.[118,261]

To describe the influence of the solvent on the HFCC, one can also refer to Figure 4, where the right zwitterionic structure will be stabilized in a polar environment. In many experiments, it was shown that the HFCC increases by increasing the dielectric constant.[261]

3.2.1 Ab-initio molecular dynamics simulation

B. Sharma performed and evaluated all AIMD simulations for HMI and HHMI.[132] The workflow is taken from ref[132]. The simulations were performed with one solute in a cubic periodic box of 128 water molecules, and the starting configurations were taken from a classical molecular dynamics simulation.

The AIMD simulations were calculated with the *CP2K* program. Kohn-Sham orbitals were represented by the atom-centered TZV2P basis set using Goedecker-Teter-Hutter pseudopotentials. Kohn-Sham orbitals and the corresponding total electron density were represented by a plane wave basis with a kinetic energy cutoff of 500 Ry. To speed up the computation of the Fock-exchange terms within the revPBE0-D3 hybrid functional, the auxiliary density matrix method (ADMM) with the cpFIT3 auxiliary basis was used.

The spin-polarized Kohn-Sham equations were solved in the simulations of the open-shell solute in solution. The NVT ensemble was simulated using the Nose-Hoover chain thermostat, resulting in every Cartesian coordinate getting its thermostat. A time step of 0.5 fs was used to solve the integrated equations of motion. The total length of the AIMD trajectory was 206 ps, where the first 6 ps were used to equilibrate. Two different configuration subsets were generated with an increment of 200 fs and 500 fs, respectively. The former subset was used for DFT calculations, whereas the latter was utilized for the DLPNO-CCSD calculations. Therefore, a total of 1000 configurations were used for the DFT calculations and 400 configurations for the DLPNO-CCSD calculations.

The same procedure was performed for the protonated version of HMI, called HHMI. For a more detailed description, see Ref [132]

3.2.2 Computational details

The two different configuration subsets (1000 and 400 snapshots) were used for either DFT or DLPNO-CCSD calculations to calculate the isotropic hyperfine couplings. In the case of DFT, the hybrid functional revPBE0-D3 was used. For all calculations, the basis set def2-TZVPP with decontracted s-functions was applied. A previous study showed that this basis set is well suited to reproduce hyperfine coupling constant, with a good balance between computational cost and accuracy.[132] The first subset containing the 1000 solvent-solute configurations was further divided into a set in which all water molecules were removed, and only the solvent-relaxed structure of HMI was retained. This subset will be called vertically desolvated (VD) in the remaining chapter. All water molecules that could not be assigned to the first two solvation shells around the N-O motif were removed in the second subset. V. A. Tran developed the corresponding script. Only the vertically desolvated structures were considered for the subset used for the DLPNO-CCSD calculations. This subset contains 400 structures that were extracted every 500 fs. For the protonated HHMI, the exact same process was performed.

The EC-RISM calculations were done on a cubic grid with 120 points in each dimension with a distance of 0.5 Å between two grid points. The SPC/E HNC water solvent susceptibility was used for 1 bar (see chapter 3.1). The GAFF version 1.4 was used as the source for the LJ parameters used in the 3D RISM calculations (see Table *37*). When the maximum residual norm of the direct correlation between two successive iterations is smaller than 10^{-6} convergence of the 3D RISM calculations was assumed. The convergence criterium for the whole EC-RISM cycle was 0.01 kcal mol^{-1} for the maximum free energy difference between two consecutive EC-RISM cycles. All EPR-parameter calculations were performed with *ORCA 4.2.1*.[281] For the DFT calculations in all iterations, the corresponding level of theory was utilized to extract the exact electrostatic potential from QC-calculated electron density. The atom-centered point charges were calculated using the CHelpG algorithm utilizing Breneman-Wiberg radii with a 0.3 Å grid spacing and a maximum distance of all atoms to any grid point of 2.8 Å. A constraint of the dipole moment to reproduce the QC-derived dipole moment was not enforced. The compression of distant point charges into a cluster of point charges was applied during the calculations.[275]

Geometry optimizations were made of 400 vertically desolvated structures generated from the first 80 ps of the AIMD using B3LYP/6-311+G(d,p) and the

CPCM solvation model for water. A clustering was then applied to the optimized structures, whereas the principle of the clustering algorithm is to calculate the RMSD between the optimized structures and assign structures that have an RMSD less than 0.1 Å to one cluster. The structure with the smallest energy was used as a reference for the corresponding cluster.

Additionally, geometry optimizations were performed with B3LYP, B3LYP-D3, revPBE0, and revPBE0-D3 using the def2-TZVPP with decontracted s-functions basis set with the CPCM solvation model for water and methanol using *ORCA Version 4.2.1*. [281,282] For these structures, HFCC single-point calculations were performed using EC-RISM and CPCM. For the methanol EC-RISM calculations, the solvent susceptibility for ambient pressure developed in ref[26], and the author's bachelor thesis was applied. Here, the solvent susceptibilities are based on 1D RSIM HNC calculations. The free energy of these optimized structures was calculated with the corresponding DFT level of theory and without an explicit PMV correction of the excess chemical potential. Additionally, MP2/6-311+G(d,p) calculations were performed for the B3LYP and revPBE0-D3 optimized structures. Here, for the water calculations, the PMV correction developed by N. Tielker was applied.[25] For the methanol calculation, a PMV correction based on the PC+ correction scheme was used, whereby the correction parameter is: c_{pc+}=-0.09696 kcal mol^{-1} $Å^{-3}$

Regarding the *g*-tensor calculations, the def2/j auxiliary basis set was combined with revPBE0-D3/def2-TZVPP with decontracted s-functions. To tackle the gauge origin problem, the GIAO approach was utilized. To simulate the W-band spectra, the *g*- and A-tensors were extracted and used as inputs for *Easyspin*. Within *Easyspin*[283], the *pepper* module was utilized for spectra simulation with a linewidth of 0.45 mT. For every snapshot, a separate spectrum was simulated, and afterward, the 1000 (or 400) spectra were added up and averaged. In this manner, a spectrum based on a sum of spectra is obtained. The W-band spectra were simulated at a frequency of 93.993 GHz

3.3 The pressure-dependence of the protein backbone model NMA, Ac-Gly-NHMe and Ac-Ala-NHMe

In the introduction, the problem was explained that pressure-dependent MD simulations were performed only with force fields parameterized for ambient pressure only. The three typical model systems, NMA, Ac-Gly-NHMe, and Ac-Ala-NHMe, were used to investigate how strong pressure affects the peptide backbone. Therefore reliable reference data is needed to evaluate how strong the effect of pressure is on the free energy hyperplane of the three typical peptide backbone models. Pressure-dependent EC RISM calculations are a well-suited tool to address this problem.

In experiments, it is observed that proteins change their structure due to pressurization. The main driver for this is that a folded protein has a larger volume than an unfolded one. In terms of thermodynamic relations, this effect can be described via:[284]

$$\Delta G(p) = \Delta G_0(p_0) + (\Delta V_0)(p - p_0) - \left(\frac{\Delta\kappa}{2}\right)(p - p_0)^2 + \cdots \tag{184}$$

where ΔG is the free energy change between the unfolded and folded state, $\mathit{\Delta G_0}$ is the corresponding free energy difference at ambient conditions, p is the pressure and ΔV_0 and $\Delta\kappa$ are the changes in volume and compressibility, respectively. ΔG_0 is normally small and positive; thus, the folded state is slightly preferred under ambient conditions. However, the second term in (184) is negative, whereby for a typical globular protein, the volume difference is in the range of -20 to 100 cm^{-3} mol^{-1}. This volume difference results in a -4 to -20 kJ mol^{-1} free energy difference at 2 kbar.[284] Due to this relatively small free energy difference, only small free energy changes on the peptides' level are expected. For EC-RISM calculations, two possible ways exist to calculate the pressure-dependent change in Gibbs energy. The first is to use the pressure-dependent partial molar volume correction, introduced in equation (76). A second approach is closely related to equation (184). The idea is to integrate the pressure-dependent change of the partial molar volume of the corresponding molecule over the pressure range, leading to the following relation:

$$G(p) = G^0 + \int_{p^0}^{p} V_m(p)\,\mathrm{d}p'. \tag{185}$$

Ideally, the calculated free energies with both methods should deliver similar results.

Another quantity to validate the quality of EC-RISM is the determination of pressure-dependent chemical shifts. In the experiments[6] a non-linear trend was observed that is mostly described via:

$$\delta(p) = B_1 \cdot (p - p_0)^2 + B_2 \cdot (p - p_0) + \delta_0. \tag{186}$$

In the latter equation, a is the linear coefficient, b is the nonlinear (second-order) coefficient, and c is the chemical shift at 1 bar. With EC-RISM, pressure-dependent chemical shifts were calculated in the past. Although the pressure trend could be reproduced correctly, some errors relative to the experiment remain. These errors could be attributed to small geometry changes, the influence of hydrogen bonds, the level of theory, errors in calculating the reference standard, or significant population changes. The first two issues shall be addressed by utilizing pressure-dependent EC-RISM geometry optimizations that were unavailable in the past. In equation 10 of ref[27], the pressure-dependent change of the ideal part of the chemical potential of the solute going from p_0 to p is calculated via:

$$\int_{p_0}^{p} \frac{\partial \mu_u^{\mathrm{id}}(p,T)}{\partial p}\,\mathrm{d}p = RT \int_{p_0}^{p} \kappa(p,T)\,\mathrm{d}p. \tag{187}$$

Whereby μ_u^{id} is the ideal part of the chemical potential, and κ is the solvent compressibility, which can be derived from experimental reference data. This equation is derived from the Kirkwood-Buff theory and follows from setting the solute-solvent Kirkwood-Buff integral to zero. Where on the one hand, this term was essential to reproduce the experimental shift in the ion product of water,[27,176] the term is not relevant for the description of the relative free energy changes between different conformers since this term is the same for all conformers.

3.3.1 Generation of conformers

In the beginning, reliable conformers for the three test systems, NMA, Ac-Gly-NHMe, and Ac-Ala-NHMe, are required. To sample the conformational space, PCM calculations at ambient conditions are used. Regarding NMA, two conformers, representing the *cis* and *trans* conformation, respectively, were optimized at the B3LYP/6-311+G(d,p)/PCM level of theory utilizing *Gaussian09 Rev E.01*.[274] For Ac-Gly-NHMe and Ac-Ala-NHMe an exhaustive dihedral scan around the backbone torsions φ and ψ were performed, leading to 400 structures for each molecule. The scans were performed at B3LYP/6-311+G(d,p)/PCM level and followed by a free relaxation optimization for every structure using *Gaussian16 Rev B.01*[285]. After the optimization, a clustering was applied, where all structures within a cartesian RMSD of 0.01 Å were considered a single minimum. This procedure leads to eight conformers for Ac-Gly-NHMe and seven conformers for Ac-Ala-NHMe. A hard restraint was implemented to guarantee identical dihedral values for the force field structures and the quantum mechanical optimized ones, with a dihedral barrier of 1000 kcal mol^{-1}. For the implementation of the restraint, the existing parameters for the backbone angles were deleted with the *Parmed* (*AmberTools 2021*[286]) program, and then a new parameter set was introduced for each of the two dihedral angles. For this parameter set, a torsion angle amplitude of 1000 kcal mol^{-1} was chosen. For the phase angle, a value was chosen that is 180° larger than the angle to be scanned. This high torsion angle amplitude ensured that the molecule retained the torsion angle during optimization. Within the scan, all other internal degrees of freedom were freely variable. The amber ff14SB[32], amber ff19SB[287], and the CHARMM36/mm[39] force field were applied for Ac-Gly-

NHMe with a Poisson Boltzman[288,289] implicit solvation model. The CHARMM36/mm force field implementation for Amber simulations was realized with the *chamber* option of *parmed*. The simulations were done using *sander*.[286] For Ac-Ala.-NHMe, ff14SB and ff19SB dihedral scans with analytical linearized Poisson-Boltzmann (ALPB) solvations were performed. The partial charges and Lennard Jones parameters of the atoms can be found in chapter 6.10

3.3.2 Pressure-dependent EC-RISM calculations on B3LYP/6-311+G(d,p)/PCM optimized structures

The EC-RISM calculations were strictly based on the already successfully applied approach to predicting different thermodynamic[25,180] and spectroscopic quantities under high-pressure conditions.[27] The EC-RISM calculations were performed at the MP2/6-311+G(d,p)/PSE-2 level with *Gaussian09 Rev E.01*.[274]. The Hartree-Fock method was used during the iterations, and after convergence of the EC-RISM iteration cycle, a final MP2 calculation was carried out. In all iterations, including the final one, the HF-electron density was used to compute the exact electrostatic potentials. Calculations of NMA were performed on a cubic grid, with 140 points per principal axis and a grid spacing of 0.3 Å. Ac-Gly-NHMe and Ac-Ala-NHMe utilized a cubic grid with 120 x 120 x 120 points with 0.3 Å between two grid points. The DRISM/HNC[18,290,291] approach was used for calculating solvent susceptibilities with densities, and dielectric constants were calculated with an EOS taken from Floriano[292]. ff14SB-based Lennard-Jones parameters (see chapter 6.9) were used for all molecules. Convergence criteria were set to 10^{-6} for 3D RISM calculations and 0.01 kcal mol^{-1} for the maximum energy difference between two consecutive EC-RISM cycles. Atom-centered point charges were determined with the ChelpG[45] algorithm using default radii to reproduce the quantum mechanically obtained dipole moment. The quantum-chemical calculations' expense was reduced by condensing far-away solvent point charges into a single point charge.[275] To estimate NMR parameters, MP2/6-311+G(d,p)/EC-RISM[26,248] calculations were applied to all molecules and conformations using the GIAO method. Additionally, DLPNO-CCSD/def2-TZVPP/PSE-3 calculations using *Orca 4.2.1*[281] were utilized. Thermal corrections were calculated with B3LYP/6-311+G(d,p)/PCM and were added to the free energies obtained from the EC-RISM calculations.

3.3.3 Pressure-dependent geometry optimizations based on EC-RISM

For the novel geometry optimizations based on the EC-RISM formalism, *ORCA 5.0.2*[182] was used, in which the EC-RISM formalism was integrated. The method B3LYP/g was utilized since this version of B3LYP uses the same parameters also included in Gaussian. As a basis set, the 6-311+G(d,p) was used to guarantee comparability with the previously made optimizations. A grid of 120^3 points with a distance of 0.3 Å between the grid points was used. In contrast to the old EC-RISM versions, the 3D RISM calculations are performed during the SCF calculations, and no complete SCF calculation is performed on one solvent polarization. Convergence was achieved when the energy change was below 5^{-6} Eh, the maximum gradient must be lower than 3^{-4} Eh/bohr, and the RMS gradient must be below 1^{-4} Eh/bohr. The maximal displacement and RMS displacement were set to 4^{-3} bohr and 2^{-3} bohr, respectively. This approach allows the utilization of pressure-dependent geometry optimizations. Regarding the pressure-dependent optimizations, a step-by-step procedure was utilized, whereby in the first step, the previously calculated PCM optima were used as the starting point for EC-RISM geometry optimizations at ambient conditions. The MP2/6-311+G(d,p)/EC-RISM calculations with *Gaussian09 Rev E.01*.[274] with the GIAO method were also applied to the pressure-dependent EC-RISM optimized structures

4 Results

4.1 RISM-based IR Calculations using equilibrium and non-equilibrium solvation

In this chapter, a detailed analysis of how well EC-RISM can represent pressure-dependent frequency shifts will be performed. At first, different combinations of χ-functions and methods to calculate the electrostatic interactions between solvent and solute within 3D RISM will be tested on the different IR-calculation (equilibrium and non-equilibrium) setups using the B3LYP/6-311+G(d,p) level of theory. In the second part of this chapter, the MP2/6-311+G(d,p)/PSE-3 setup will be used to calculate the theoretically most accurate pressure-dependent IR frequencies. On the one hand, the diatomic anion cyanide, which shows only one normal mode, is used as a test system. The other system under investigation is the small osmolyte TMAO, which shows four significant IR-band in the spectrum (the corresponding modes are presented in Figure 5).

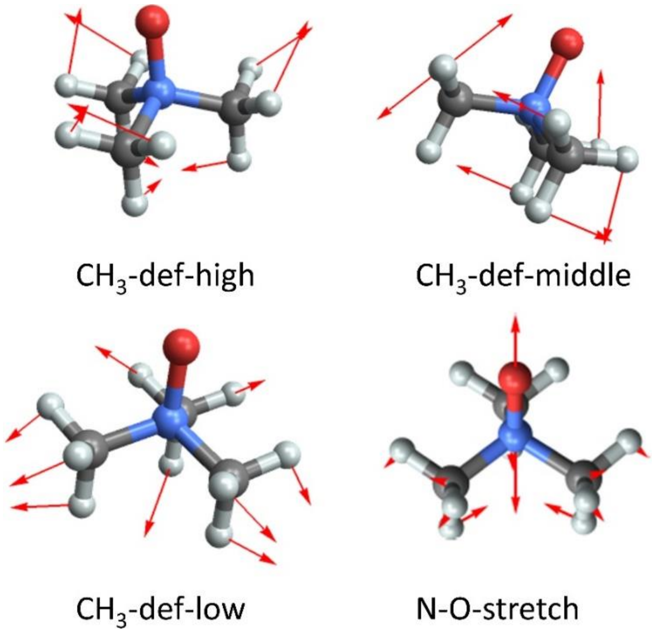

Figure 5 The four analyzed and experimental observable TMAO IR modes. An analogous picture can be found in ref[94] and ref[176])

4.1.1 Benchmarking pressure-dependent RISM-based IR-calculation settings

In the 2016 published paper by Imoto et al.,[94] pressure-dependent IR calculations computed with EC-RISM were presented for the first time. However, the absolute pressure-dependent EC-RISM wavenumbers were not presented at that time. Instead, the relative pressure-dependent change that could be evaluated using EC-RISM calculations was used to scale the wavenumbers calculated at ambient conditions using B3LYP/6-311+G(d,p)/PCM. This chapter will mainly show the advances which have led to a consistent improvement of the EC-RISM formalism and how they affect the calculation of IR frequencies.

First of all, it must be mentioned that in older EC-RISM calculations, an HF calculation was automatically performed during the iterations of the EC-RISM cycle as soon as a DFT functional was used. This is problematic because the final B3LYP calculation is then performed under the influence of an electrostatic solvent polarization, which was polarized with the HF electron density. Especially in the subsequent 3D RISM calculation, an inconsistency occurs since the B3LYP electron density now polarizes the solvent distribution. Therefore, one can no longer assume a complete convergence of the EC-RISM cycle. In the context of this thesis, all DFT calculations used the respective theory level in the iterations (unless explicitly mentioned).

Furthermore, it should be noted that the old IR calculations (in ref[94]) were performed[94] with χ-functions based on pressure-dependent g-functions from MD simulations.[18] These χ functions have proven to be particularly suitable for reproducing the pressure-dependent dipole moment changes of TMAO compared to AIMD dipole moment changes. However, in further investigations, especially regarding the calculation of pressure-dependent NMR chemical shifts,[26,27] the χ_{HNC} susceptibilities showed significantly better performances. Another possible issue is that the old calculations used the ESP-derived solute point charges to calculate the electrostatic interactions between solute and solvent in EC-RISM. However, during the SAMPL6[25] prediction challenge, a new efficient way to calculate the electrostatic interactions via the QC-derived electrostatic potential was developed (see 2.1.4). This is why four different combinations of electrostatic calculations and χ-functions will be compared in this part to benchmark the performances of EC-RISM with B3LYP/6-311+G(d,p). Furthermore, the two new approaches for calculating a non-equilibrium situation will be presented (see 3.1.2).

Table 1 IR-wavenumbers (in cm^{-1}) at 1 bar and 10 kbar of the four TMAO normal modes and the cyanide ion calculated with different EC-RISM settings utilizing the B3LYP/6-311+G(d,p) level of theory EQ represents the equilibrium approach, M1 is the first new non-equilibrium approach and M2 the second non-equilibrium approach. $\varphi/\chi_{\mathrm{SIM}}$ represents the combination of full electrostatic potential with χ_{SIM}. $\varphi\ /\chi_{HNC}$ is the combination of full electrostatic potential with χ_{HNC}, the q/χ_{SIM} model uses the solute point charges to compute the electrostatic interactions in conjunction with χ_{SIM}, and q/HNC represents χ_{HNC} with point charges. The above number in a cell represents the unscaled number, whereas the lower number is scaled by a scaling factor of 0.967.[202]. Raw data and structures can be found in SI_4.1. The displacement for the different energy components (E_{sol}, μ^{es}, and G_{sol}) plots can be found in SI_4.1.

Method	Mode	Pressure									
		1 bar					10 kbar				
EQ		$\varphi/\chi_{\mathrm{Sim}}$	$\varphi/\chi_{\mathrm{HNC}}$	q/χ_{Sim}	q/χ_{HNC}	Exp	$\varphi/\chi_{\mathrm{Sim}}$	$\varphi/\chi_{\mathrm{HNC}}$	q/χ_{Sim}	q/χ_{HNC}	Exp
	CH_3-def-high	1511.4/ 1461.5	1510.8/ 1460.9	1512.1/ 1462.1	1504.3/ 1454.5	1481.6	1510.0/ 1460.3	1510.5/ 1460.7	1527.1/ 1476.7	1506.2/ 1456.3	1483.1
	CH_3-def-middle	1490.9/ 1441.7	1490.6/ 1441.5	1493.3/ 1444.0	1488.9/ 1439.8	1465.6	1491.6/ 1442.4	1491.6/ 1442.4	1513.6/ 1463.7	1492.5/ 1443.2	1469.0
	CH_3-def-low	1433.4/ 1386.1	1435.6/ 1388.2	1450.7/ 1402.8	1451.2/ 1403.3	1404.1	1438.4/ 1391.0	1439.1/ 1391.6	1479.6/ 1430.8	1456.2/ 1408.1	1406.3
	NO-stretch	971.0/ 939.0	984.6/ 952.1	972.1/ 940.0	983.1/ 950.7	952.5	986.9/ 954.3	990.0/ 957.3	994.9/ 962.1	987.9/ 955.3	956.1
	CN^-	2135.9/ 2062.5	2127.9/ 2057.7	2140.4/ 2069.8	2134.6/ 2065.2	2079.3	2132.5/ 2062.1	2128.6/ 2058.4	2137.1/ 2050.4	2135.3/ 2064.8	2083.0
M1	CH_3-def-high	1510.1/ 1460.3	1510.7/ 1460.8	1510.0/ 1460.2	1510.8/ 1460.9	1481.6	1512.1/ 1462.2	1512.4/ 1462.5	1512.1/ 1462.2	1512.5/ 1462.6	1483.1
	CH_3-def-middle	1489.5/ 1440.3	1490.5/ 1441.3	1489.6/ 1440.4	1490.9/ 1441.7	1465.6	1491.5/ 1442.3	1491.8/ 1442.6	1491.7/ 1442.5	1492.2/ 1443.0	1469.0
	CH_3-def-low	1415.6/ 1368.9	1414.6/ 1367.9	1413.3/ 1366.7	1415.0/ 1368.3	1404.1	1415.7/ 1368.9	1416.2/ 1369.5	1416.1/ 1369.4	1416.8/ 1370.0	1406.3
	NO-stretch	890.3/ 860.9	896.5/ 866.9	892.2/ 862.8	899.9/ 870.2	952.5	898.4/ 868.8	900.0/ 870.3	901.9/ 872.1	903.7/ 873.9	956.1
	CN^-	2152.7/ 2081.7	2152.8/ 2081.8	2151.3/ 2080.3	2153.6/ 2080.6	2079.3	2152.7/ 2081.7	2152.6/ 2081.6	2151.0/ 2079.1	2153.4/ 2082.3	2083.0
M2	CH_3-def-high	1502.6 1451.5	1504.3/ 1453.2	1502.7/ 1451.6	1504.5/ 1453.3	1481.6	1503.9/ 1452.8	1504.4/ 1453.3	1504.0/ 1452.9	1504.6/ 1453.4	1483.1
	CH_3-def-middle	1485.6/ 1435.1	1488.3/ 1437.7	1485.8/ 1435.3	1488.7/ 1438.1	1465.6	1487.6/ 1437.0	1488.2/ 1437.6	1487.9/ 1437.3	1488.7/ 1438.1	1469.0
	CH_3-def-low	1411.8/ 1363.8	1414.7/ 1366.6	1412.1/ 1364.1	1415.1/ 1367.0	1404.1	1413.5/ 1365.4	1414.1/ 1366.0	1414.0/ 1365.9	1414.7/ 1366.6	1406.3
	NO-stretch	897.3/ 866.8	905.3/ 874.5	899.6/ 869.0	909.3/ 872,6	952.5	905.2/ 874.4	907.5/ 876.6	909.1/ 878.2	911.6/ 880.6	956.1
	CN^-	2149.5/ 2078.6	2149.9/ 2079.0	2150.1/ 2079.2	2150.7/ 2079.7	2079.3	2151.6/ 2080.6	2150.3/ 2079.3	2151.0/ 2080.0	2151.1/ 2080.1	2083.0

4.1.2 Benchmarking pressure-dependent RISM-based IR-calculation settings: The equilibrium ansatz

In the beginning, the focus will be set on the equilibrium method. In Table 1, the absolute wavenumbers at 1 bar and 10 kbar for the four different combinations and the four modes of TMAO are shown and were calculated with B3LYP/6-311+G(d,p). The experimental values (the measurements were performed by C. Rosin[94] and P. H. Schummel[27]) at ambient conditions are 1481.6 cm^{-1} for the CH_3-def-high mode, 1465.6 cm^{-1} for the CH_3-def-middle mode, 1404.1 cm^{-1} for the CH_3-def-low mode and 952.5 cm^{-1}. For the CH_3-def-high, the φ/χ_{Sim} variant leads to a frequency of 1514/1461.5 cm^{-1} (in the latter case, a precomputed vibrational scaling factor of 0.967 was applied (see computational chemistry comparison and benchmark database[202]). Only marginal differences occur when the φ/χ_{HNC} and q/χ_{Sim} methods are considered. The q/χ_{HNC} shows a smaller frequency with 1504.3/1454.3 cm^{-1}. If no precomputed scaling parameter is applied, the q/χ_{Sim} variants show the best agreement relative to the experimental value. In contrast, as soon as the scaling factor is applied, the other three variants show superior results; the relative difference to the experiment is 20 cm^{-1}.

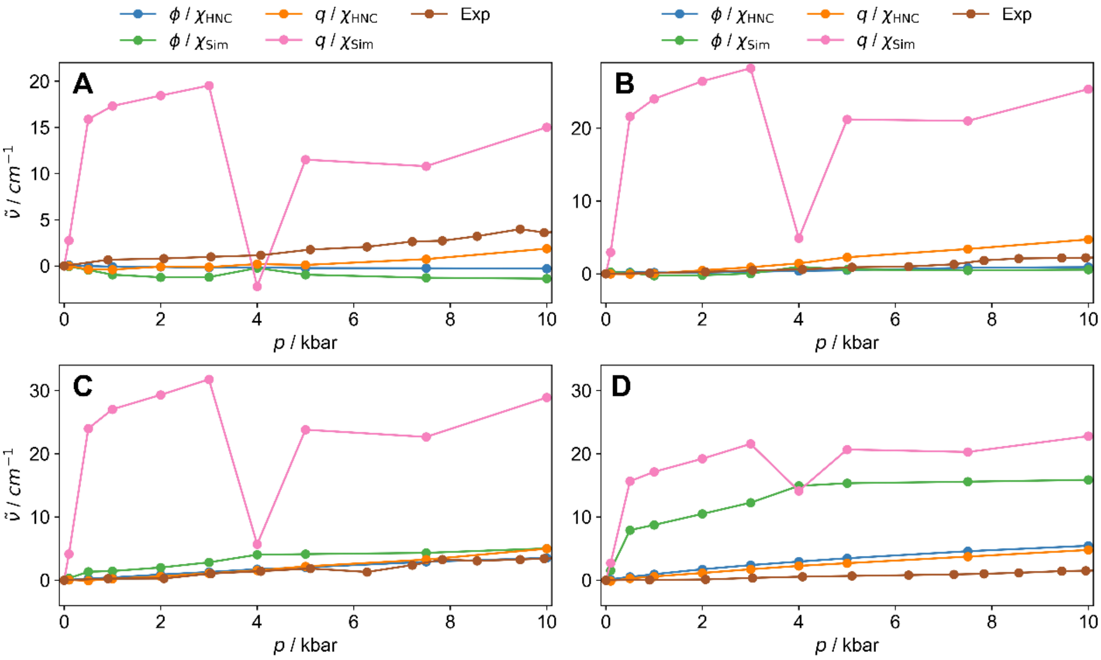

Figure 6 Pressure-induced frequency shifts of the four relevant TMAO modes calculated with the EQ approach using B3LYP/6-311+G(d,p)/PSE-3. Panel (A) shows the CH_3-def-high mode, in panel (B) CH_3-def-middle mode is depicted, (C) shows the CH_3-def-low mode, and in (D), the NO-stretch mode is presented. The green dots show the wavenumbers calculated using the exact electrostatic potential and χ_{HNC}; the black dots represent the calculation with the ESP-derived point charges and χ_{HNC}. The red dots represent the calculation with exact electrostatic potential and χ_{SIM}. The magenta points represent the experimentally measured values. Raw and calculated data can be found in SI_4.1.

In Figure 6, the pressure-dependent frequency shifts for the equilibrium method and the four TMAO-Modes are presented and compared to the experimentally obtained shifts. In Panel (A), the results for the CH_3-def-high mode are shown. A steady blue shift is observed in the experiment, leading to a difference relative to 1 bar of 1.6 cm^{1}. Something immediately noticeable are the pressure-dependent shifts calculated with the q/χ_{SIM} method. In direct comparison, the pressure-dependent changes are significantly larger than for the other variants, and no steady pressure trend is visible, and a very strong outlier can be observed at 4 kbar. The obtained wavenumber at 10 kbar increased nearly 15 cm^{-1} relative to the ambient value. The two variants with the exact electrostatic potential show the wrong pressure tendency, resulting in a pressure-dependent decrease. The best agreement relative to the experimental observation is obtained with the q/χ_{HNC} with an increase of 2 cm^{-1} at 10 kbar compared to the value at ambient conditions.

For the CH_3-def-middle (absolute values in Table 1), a wavenumber of 1490.6/1441.5 cm^{-1} was calculated with φ/χ_{HNC}, resulting in a difference relative to the experimental measurements of +25 cm^{-1} (unscaled) and -24.1 cm^{-1} (scaled frequency). The φ/χ_{SIM} variant (1490.9/1441.7 cm^{-1}) shows nearly no difference compared to the formerly discussed variant. The q/χ_{HNC} leads to slightly smaller wavenumbers, whereas the q/χ_{SIM} shows a higher frequency. In summary, the equilibrium method for the CH_3-def-middle mode produces a similar performance relative to the experiment as for the previously introduced mode.

Figure 6 Panel (B) depicts the pressure-dependent frequency shift of the CH_3-def-middle mode. The experimental reference shows a steady increase resulting in a difference of 3.4 cm^{-1}. All four approaches can represent the steady frequency increase due to pressurization. However, just as for the previously discussed mode, it can be observed that the q/χ_{SIM} method predicts a much too steep increase, which is also not continuous, with a strong outlier at 4 kbar. The φ/χ_{HNC} and φ/χ_{SIM} show an increase of 0.6 cm^{-1} and 0.9 cm^{-1} up to 10 kbar, respectively. The q/χ_{HNC} overestimates the pressure response; however, as for the CH_3-def-high mode, the pressure dependence can best be reproduced with this method.

The third methyl deformation mode is the CH_3-def-low mode (absolute values in Table 1). The differences between the various variants become much more pronounced for this vibrational mode. Using the φ/χ_{SIM} variant, a wavenumber of 1433.4/1386.1 cm^{-1} was obtained at ambient pressure. Without scaling, a difference of nearly 30 cm^{-1} to the experiment is observed, whereas, with the precomputed scaling factor, the offset reduces to –18 cm^{-1}. The φ/χ_{HNC} approach leads to a wavenumber of 1435.6/1388.2 cm^{-1} at 1bar. The CH_3-def-low mode strongly depends on the chosen electrostatic interaction method. The calculated wavenumbers with the two point-charge approaches are 1450.7/1402.8 cm^{-1} (q/χ_{SIM}) and 1451.2/1403.3 cm^{-1} (q/χ_{HNC}). The calculated numbers nearly match the experimentally obtained ones by scaling the frequencies.

The pressure dependence for the q/χ_{SIM} shows the same (Panel (C) of Figure 6) behavior for the CH_3-def-low mode as in the previous two cases, increasing nearly 30 cm^{-1} to 10 kbar. However, only a shift of 2.3 cm^{-1} was measured in the experiment. The other three variants show a good agreement with the experimental pressure progress. For q/χ_{HNC} and φ/χ_{SIM} an increase of 5 cm^{-1} (up to 10 kbar) can be observed, whereas the φ/χ_{HNC} shows with 3.5 cm^{-1} the best agreement with respect to the experiment.

The last TMAO mode that is discussed is the N-O stretch mode (absolute values in Table 1). In contrast to the CH_3-def-low mode, where a strong dependency on the electrostatic model was observed, a strong solvent susceptibility dependency is observed here. The φ/χ_{HNC} and q/χ_{HNC} nearly match the experimental value of 952.5 cm^{-1} if the values are scaled. For the other two methods, the calculated wavenumbers are 971.0/939.0 cm^{-1} (φ/χ_{SIM}) and 972.1/940.0 cm^{-1} (q/χ_{SIM}. The pressure dependence is depicted in Panel (D) of Figure 6 Panel (D). In the experiment, an increase of 3.6 cm^{-1} is obtained. Here, the two χ_{SIM} show a significantly too strong pressure-dependent increase. However, all variants overestimate the pressure dependence, whereas for the φ/χ_{HNC} 5.4 cm^{-1} and for the q/χ_{HNC} 4.8 cm^{-1}, respectively.

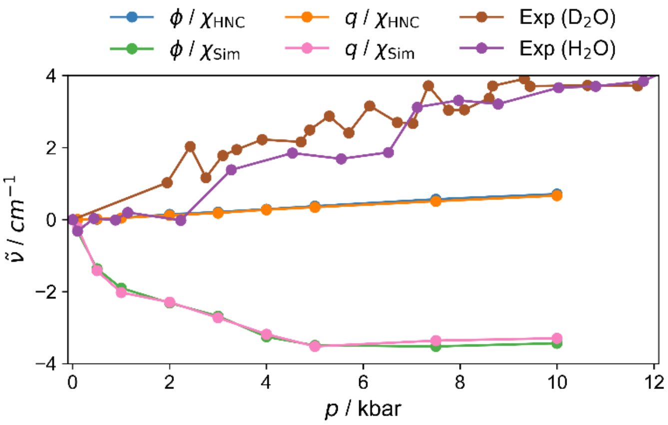

Figure 7 Pressure-induced wavenumber shifts of the cyanide anion calculated with B3LYP/6-311+G(d,p) and the EQ approach. The green dots were calculated using the exact electrostatic potential and χ_{HNC}; the black dots were calculated using the ESP-derived point charges and χ_{HNC}. The red dots represent the calculation with exact electrostatic potential and χ_{SIM}; the magenta points represent the experimentally measured values. Raw and calculated data can be found in SI_4.1.

For the cyanide anion, the experimentally measured wavenumbers are 2079.3 cm^{-1} in H_2O and 2080.4 cm^{-1} in D_2O at 1 bar (see Table 1). The unscaled wavenumbers are generally too high and range between 2127 cm-1 and 2140 cm^{-1}, whereby the full electrostatic potential calculations lead to slightly smaller values. By applying the scaling factor of 0.967, the relative differences to the experiment are significantly smaller, where the best result is obtained with the q/χ_{SIM} combination (2069.8 cm^{-1}). Under pressure, the cyanide mode wavenumber increases by around 4 cm^{-1} up to 10 kbar. In the calculations (see Figure 7), only the χ_{HNC} methods can correctly reproduce the pressure trend, whereas the two χ_{SIM} methods show a reduction in the wavenumbers. No clear differences between the two charge models are obtained. The predicted pressure trend is underestimated with χ_{HNC} methods; however, no clear outliers are observed. A significant wrong trend is observed for the two χ_{SIM} methods. A possible explanation for the very small differences between the electrostatic models could be that the point charges for the two-atom-containing molecule cyanide can accurately represent the exact electrostatic potential.

4.1.3 Benchmarking pressure-dependent EC-RISM calculation settings: The M1 non-equilibrium ansatz

In this part, the newly developed M1 non-equilibrium method is presented. The different combinations of electrostatic calculation and solvent susceptibilities are again analyzed. The corresponding data for the absolute wavenumbers are presented in Table 1. For the CH_3-def high mode and CH_3-def middle mode, the two variants with explicit electrostatics show similar results compared to the equilibrium approach at 1 bar, with fewer than 1% deviations. For q/χ_{SIM}, the absolute wavenumber at 1 bar decreases by 2.1 cm^{-1} compared to the equilibrium approach for the CH_3-def high mode, and for the CH_3-def middle mode, a decrease of 3.7 cm^{-1} is observed. The highest deviation for the CH_3-def high mode (6.5 cm^{-1}) relative to the equilibrium method is obtained with the q/χ_{HNC}. The latter variant leads to an increase of 2 cm^{-1} for the CH_3-def middle mode.

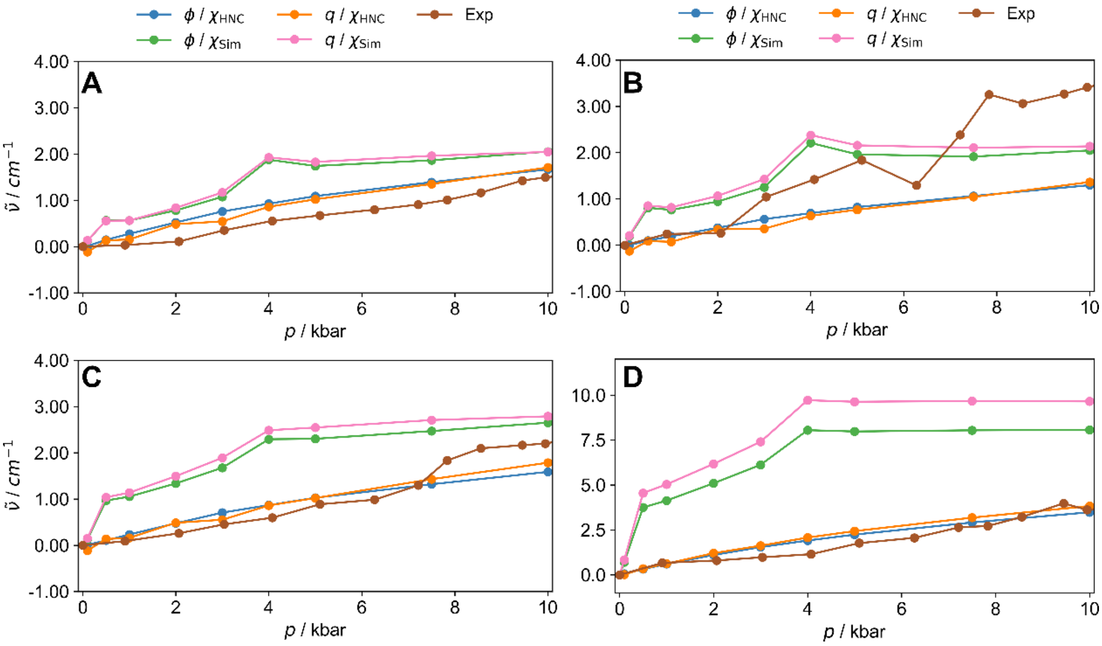

Figure 8 Pressure-induced frequency shifts of the four relevant TMAO modes calculated with the M1 non-equilibrium approach using B3LYP/6-311+G(d,p)/PSE-3. Panel (A) shows the CH_3-def-high mode, in panel (B) CH_3-def-middle mode is depicted, (C) shows the CH_3-def-low mode, and in (D), the NO-stretch mode is presented. The green dots were calculated using the exact electrostatic potential and χ_{HNC}; the black dots were calculated using the ESP-derived point charges and χ_{HNC}. The red dots represent the calculation with exact electrostatic potential and χ_{SIM}; the magenta points represent the experimentally measured values. Raw and calculated data can be found in SI_4.1.

Regarding the pressure dependence (see Figure 8 Panel (A)) for the CH_3-def high mode, it is noticeable that the fluctuations between the different model combinations are significantly lower compared to the equilibrium method. The experimentally measured pressure shift of 1.6 cm^{-1} at 10 kbar can be best described by the χ_{HNC} models, although the differences to the χ_{SIM} models are relatively small. The differences caused by the different electrostatic models are negligible in the case of the non-equilibrium method. However, the χ_{SIM} models again show an outlier at 4 kbar and, therefore, a non-monotonous curve.

For the CH_3-def middle variant, the differences between the EQ and M1 methods become more concise. A decrease of 21 cm^{-1} relative to the equilibrium approach is observed for the full electrostatic potential variants. The decrease is stronger for the point charge variants, with nearly 40 cm^{-1}. Without scaling of the wavenumbers, the M1 methods can represent the experimental value better than the EQ methods. All four variants accurately match the pressure trend (Figure 8 Panel (B)). However, the two χ_{HNC} variants go nearly hand in hand with the experimentally observed shifts up to a pressure of 7.5 kbar and then show a minor underestimation. The χ_{SIM} variants overestimate the pressure dependence at all pressures.

With the M1 method, the absolute wavenumbers for the N-O stretch mode are reproduced significantly worse than with the EQ method. The wavenumbers are about 80 cm^{-1} lower compared to the EQ results. In contrast to the other three modes, it can be observed for the N-O stretch mode that the differences between the individual electrostatic and susceptibility variants are more pronounced with a 10 cm^{-1} difference. The pressure dependence shows that the χ_{HNC} variants reproduce the experimental trend accurately, whereas the χ_{SIM} variants significantly overestimate the pressure trend.

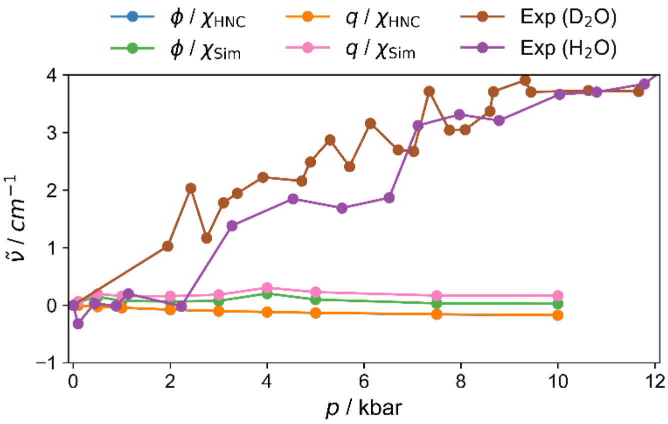

Figure 9 Pressure-induced wavenumber shifts of the cyanide anion, calculated with the M1 method. The green dots were calculated using the exact electrostatic potential and χ_{HNC}; the black dots were calculated using the ESP-derived point charges and χ_{HNC}. The red dots represent the calculation with exact electrostatic potential and χ_{SIM}; the magenta points represent the experimentally measured values. Raw and calculated data can be found in SI_4.1.

At ambient conditions, the results for the vibrational mode of the cyanide anion are in excellent agreement with the experimental results when the scaling factor is applied (see Table 1). The results range between 2079.7 cm^{-1} and 2080.6 cm^{-1}, leading to a deviation relative to the experiment of less than 1 cm^{-1}. In contrast to the TMAO modes, the pressure dependence (see Figure 8) is significantly underestimated for all four methods. The variants using the χ_{HNC} as solvent susceptibility show a pressure-induced decrease of the wavenumbers, which contradicts the experimental finding. Thus, cyanide is the first analyzed mode, where the M1 methods fail to predict the correct pressure trend.

4.1.4 Benchmarking pressure-dependent EC-RISM calculation settings: The M2 non-equilibrium ansatz

The final method used to calculate IR wavenumbers is the M2 non-equilibrium method. In contrast to the M1 method, only the intramolecular interactions are described in the context of a non-equilibrium approach, whereas a separate 1D RISM calculation is performed for each displaced structure to calculate the excess chemical potential; note that the partial charges of the minimum structure are used in these calculations.

For all three CH_3-def modes (see Table 1), lower wavenumbers are calculated under atmospheric pressure compared to the other two methods. Without applying the scaling factor, the M2 method leads to the best agreement with the experimental reference. The deviation for the CH_3-def-high relative to the experiment range between 22 cm^{-1} and 24 cm^{-1}. For CH_3-def-middle, similar results are obtained, whereas, for the CH_3-def-low mode, the unscaled wavenumbers show the smallest deviation from the experiment with 7 cm^{-1} to 11 cm$^{-1.}$ The worst agreement with the experiment is obtained using the M2 method with scaling. Noticeably, only the solvent susceptibility leads to differences between the individual methods, whereas the choice of the electrostatic model (with identical solvent susceptibility) leads to no relevant differences. The differences between χ_{HNC} and χ_{SIM} are 2 cm^{-1} (CH_3-def-high) and 3 cm^{-1} (CH_3-def-low). Regarding the NO-stretch mode, the absolute wavenumbers range from 897.3/866.8 cm^{-1} ($\varphi/\chi_{\mathrm{SIM}}$) up to 909.3/872.6 cm^{-1}. Therefore, the M2 non-equilibrium method yields similar magnitudes as the M1 method and thus also shows the largest deviation from the equilibrium method for the NO-stretch mode.

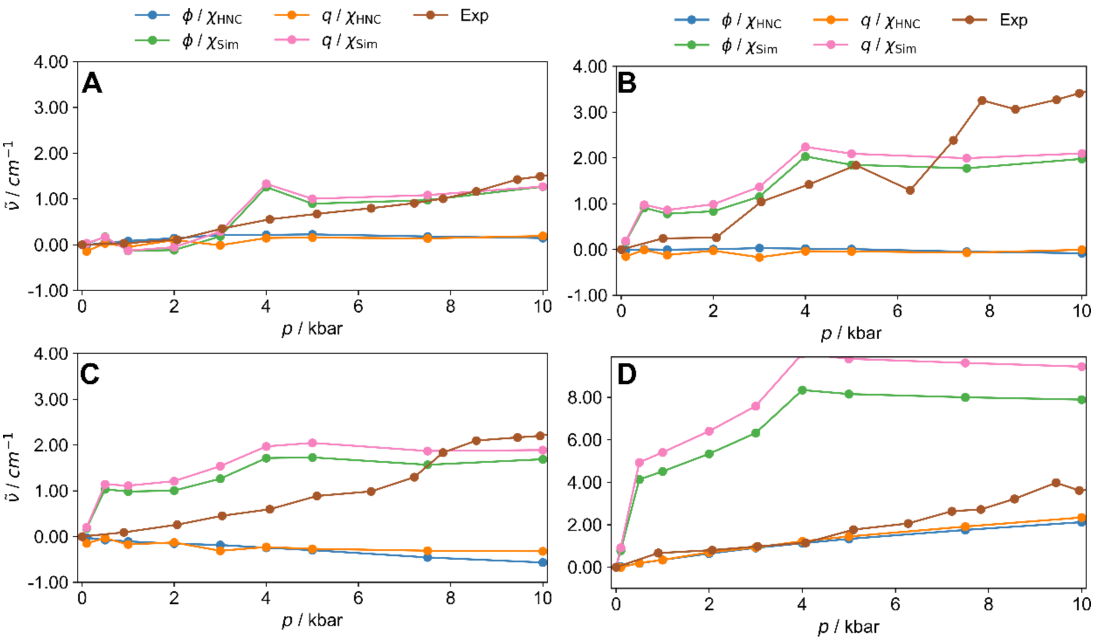

Figure 10 Pressure-induced frequency shifts of the four relevant TMAO modes calculated with the M2 non-equilibrium approach using B3LYP/6-311+G(d,p)/PSE-3. Panel (A) shows the CH_3-def-high mode, in panel (B) CH_3-def-middle mode is depicted, (C) shows the CH_3-def-low mode, and in (D), the NO-stretch mode is presented. The green dots were calculated using the exact electrostatic potential and χ_{HNC}; the black dots were calculated using the ESP-derived point charges and χ_{HNC}. The red dots represent the calculation with exact electrostatic potential and χ_{SIM}; the magenta points represent the experimentally measured values. Raw and calculated data can be found in SI_4.1.

The two χ_{SIM} variants best resemble the pressure shift (see Figure *10*) for the CH_3-def-high-mode. The difference between the two electrostatic models is negligible; however, the outlier at 4 kbar occurs again. The pressure rise to 10 kbar is hit almost exactly. In contrast, the two χ_{HNC} types fail to predict the correct pressure trend and show a decrease in the wavenumbers upon pressurization. A similar picture is obtained for the CH_3-def-middle-mode, where the χ_{SIM} methods can nearly represent the pressure change at 10 kbar, whereas the two other variants show a decline. The same result is observed for the CH_3-def-low-mode. In contrast, it can be observed again that the χ_{SIM} variants overestimate the pressure trend for the NO-stretch mode. The two χ_{HNC} approaches produce satisfying results in representing the pressure-dependent frequency shift.

In summary, the M2 method gives similar results to the M1 method at ambient pressure. However, the M2 method cannot reproduce the correct pressure trend for every combination of electrostatic model and solvent susceptibilities. Despite this, both χ_{SIM} variants, except for the obligatory outlier at 4 kbar, can correctly reproduce the pressure for the CH_3-def modes. Nonetheless, the M2 method cannot be considered suitable, and the M1 method should be used instead.

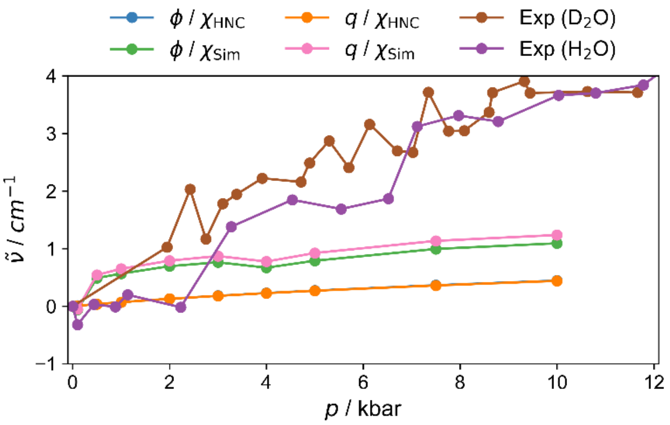

Figure 11 Pressure-induced wavenumber shifts of the cyanide anion, The green dots were calculated using the exact electrostatic potential and χ_{HNC}; the black dots were calculated using the ESP-derived point charges and χ_{HNC}. The red dots represent the calculation with exact electrostatic potential and χ_{SIM}; the magenta points represent the experimentally measured values. Raw data and calculated frequencies can be found in SI_4.1.

There are nearly no differences between the M2 and the M1 methods for the cyanide anion under ambient conditions (see Table 1), and the results are relatively satisfying. The difference to the experimentally obtained results is less than 1 cm^{-1} if the scaling is applied. In contrast to the M1 method, the M2 method can correctly reproduce the experimentally observed pressure trend. However, the absolute shifts are especially for the χ_{HNC} underestimated. A possible explanation for this could be that the cyanide mode is more sensitive to pressure-dependent changes in the equilibrium geometry, which are not included in the current approach.

4.1.5 Comparison between pressure-dependent IR calculations using MP2/6-311+G(d,p)/χ_{HNC} and B3LYP/6-311+G(d,p)/χ_{HNC}

Based on the results from the previous three chapters, the χ_{HNC} calculations are more reliable than the χ_{SIM} ones. Furthermore, a strong difference in the pressure trend compared to the experiment is observed when the point charge approach is used for the calculations with χ_{SIM}. This observation indicates that the combination of full electrostatic potential with χ_{HNC} is the most reliable. Nonetheless, the differences between φ/χ_{HNC} and $q/\ \chi_{HNC}$ are relatively small. On the other hand, the level of theory MP2/6-311+G(d,p)/χ_{HNC} proved to be the best EC-RISM variant in different benchmarks to calculate free energies in solution.[25,180] Furthermore, this variant uses χ_{HNC} with exact electrostatic potential, which produced the best results for the DFT calculations in most categories. In addition, Nicolas Tielker developed a pressure-dependent and PMV-dependent correction of the excess chemical potential. In this chapter, this level of theory shall now be applied to calculating the pressure-dependent IR problem. The raw data and structures can be found in SI_4.1.

In Table 2, the absolute wavenumbers calculated with MP2/6-311+G(d,p)/χ_{HNC} at 1 bar and 10 kbar for the three different EC-RISM variants are presented together with the experimental results and the results obtained from AIMD calculations (performed by S. Imoto[94] and see 2.4). Moreover, the results from the B3LYP/6-311+G(d,p)/χ_{HNC} are again shown in Table 2

Table 2 Absolute wavenumbers (in cm^{-1}) for the TMAO and cyanide IR modes calculated from different EC-RISM methods (here MP2/6-311+G(d,p)/χ_{HNC} and B3LYP/6-311+G(d,p)/χ_{HNC}) and (for TMAO) from AIMD compared to the experimental data for 1 bar and 10 kbar. Material for the MP2/6-311+G(d,p)/χ_{HNC} from ref[27] Raw and calculated data can be found in SI_4.1.

Mode	Exp.	M1/M1*	M2/M2*	EQ/EQ*	AIMD
1 bar		MP2/6-311+G(d,p)/χ_{HNC}			
CH$_3$-def-high	1481.6	1527.1/1450.6	1520.8/1444.8	1525.6/1449.3	1476
CH$_3$-def-middle	1465.6	1508.0/1432.6	1506.3/1430.7	1508.6/1433.1	1456
CH$_3$-def-low	1404.1	1430.7/1358.5	1430.6/1358.5	1451.9/1379.3	1394
NO-stretch	952.5	965.8/917.5	974.7/926.0	1060.0/1007.0	876
CN$^-$	2079.3/2080.4[a]	2028.3/1926.6	2025.5/1924.2	2033.6/1931.9	-
1 bar		B3LYP/6-311+G(d,p)/χ_{HNC}			
CH$_3$-def-high	1481.6	1510.7/1460.8	1504.3/1453.2	1510.8/1460.9	-
CH$_3$-def-middle	1465.6	1490.5/1441.3	1488.3/1437.7	1490.6/1441.5	-
CH$_3$-def-low	1404.1	1414.6/1367.9	1414.7/1366.6	1435.6/1388.2	-
NO-stretch	952.5	896.5/866.9	905.3/874,5	984.6/952.1	-
CN$^-$	2079.3/2080.4[a]	2150.8/2079,8	2150.7/2080.3	2127.9/2057.7	-
10 kbar		MP2/6-311+G(d,p)/χ_{HNC}			
CH$_3$-def-high	1483.1	1528.8/1452.4	1521.0/1445.0	1525.3/1449.1	1481
CH$_3$-def-middle	1469.0	1509.1/1433.6	1505.8/1430.5	1509.5/1434.0	1462
CH$_3$-def-low	1406.3	1432.4/1360.4	1430.2/1358.6	1455.5/1382.7	1402
NO-stretch	956.1	967.9/919.5	975.5/926.7	1064.1/1010.9	882
CN$^-$	2083.0/2084.1[b]	2028.0/1926.5	2025.8/1924.5	2034.9/1933.2	-
10 kbar		B3LYP/6-311+G(d,p)/χ_{HNC}			
CH$_3$-def-high	1483.1	1512.4/1462.5	1504.4/1453.3	1510.5/1460.7	-
CH$_3$-def-middle	1469.0	1491.8/1442.6	1488.2/1437.6	1491.6/1442.4	-
CH$_3$-def-low	1406.3	1416.2/1369.5	1414.1/1366.0	1439.1/1391.6	-
NO-stretch	956.1	900.0/870.3	907.5/876.6	990.0/957.3	-
CN$^-$	2083.0/2084.1[b]	2151.0/2080.0	2150.9/2079.9	2128.6/2058.4	-

Besides the fact that the correlation energy is calculated more correctly by MP2, there is a big difference between the two methods in calculating the electrostatic potential during the iterations. For MP2, the HF density is used during the iterations, whereas in the B3LYP calculation, the consistent DFT density is used. One observation that stands out directly is that the calculated absolute wavenumbers with MP2/6-311+G(d,p)/χ_{HNC} are mostly higher (except for the cyanide anion) than the corresponding results with B3LYP/-6311+G(d,p)/χ_{HNC} (see Table 2). A similar observation is also made in chapter 4.2.4, where the HFCC was calculated for the molecule HMI (see Table 5). Here, the HF density causes a stronger solvent polarization than the DFT density. The HFCC is a direct beacon for detecting the polarity of the solvent, and with the HF Density, higher HFCCs were calculated compared to the corresponding DFT Density. From the pressure-dependent results, an increase in wavenumbers was observed at higher pressure, with the polarity of the solvent increasing under pressure. Therefore, it can be concluded that higher wavenumbers are not solely due to the MP2 method but also due to the stronger polarization of the solvent caused by the HF electrostatics to polarize the solvent.

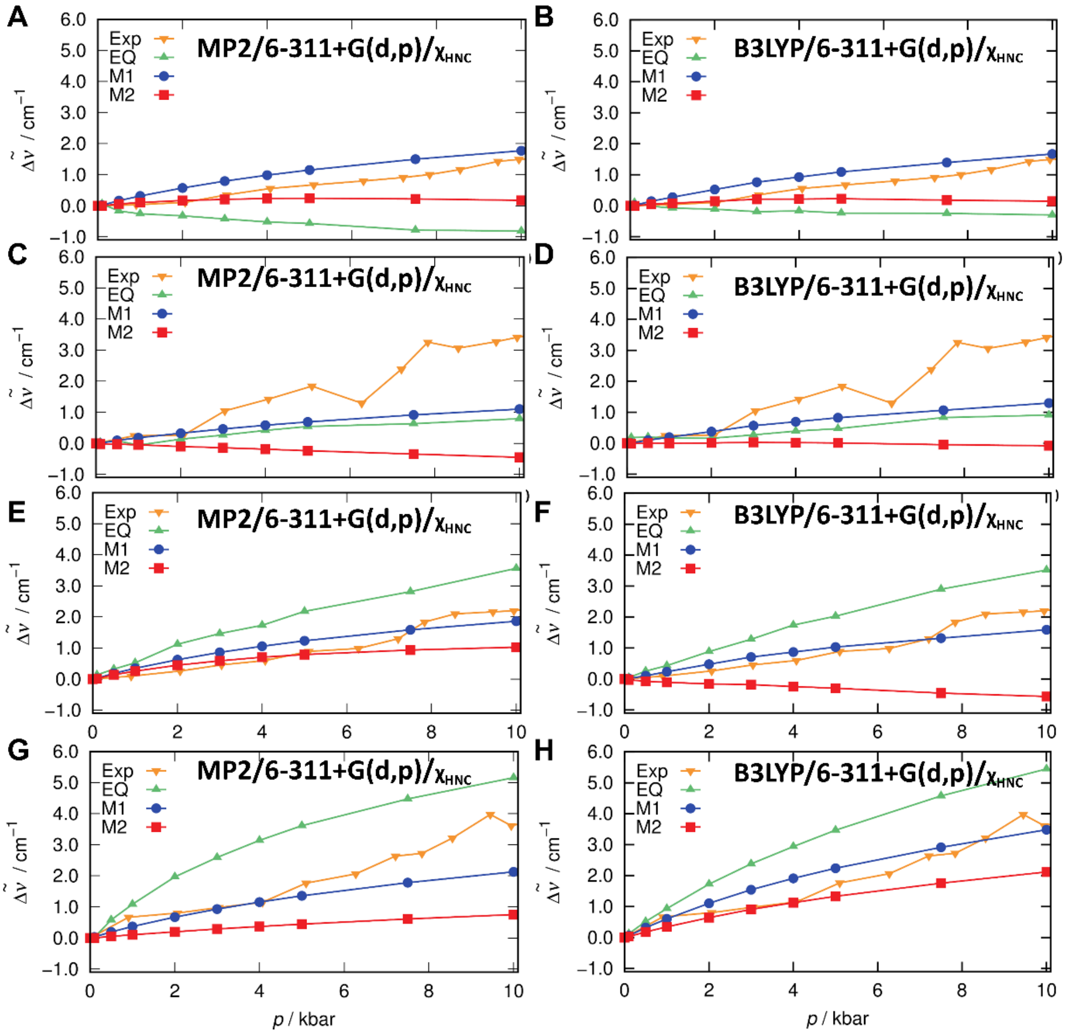

Figure 12 Pressure-dependent wavenumber shifts of the four TMAO-Modes calculated with MP2/6-311+G(d,p)/χ_{HNC} (left column) and B3LYP/6-311+G(d,p)/χ_{HNC}. (A) and (B) show the CH_3-def-high mode, in (C) and (D) the CH_3-def-middle mode, (E) and (F) CH_3-def-low, and (G) and (H) NO-stretch. The green triangles represent the equilibrium approach, the blue points the M1 non-equilibrium approach, and the red squares the M2 non-equilibrium approach. Material from the left column is published in ref[27]. Raw and calculated data can be found in SI_4.1. The displacement plots for the energy components of the MP2 calculations can be found in 6.1

For the CH_3 def-high mode, the best results for the absolute frequencies at 1 bar are obtained by applying the M1 (1527.1/1450.6 cm^{-1}) and EQ (1525.6/1449.3 cm^{-1}) methods. Since the scaling factor used for MP2/6-311+G(d,p) is 0.95[202], which is smaller than the 0.965[202] in the case of B3LYP/6-311+G(d,p), the scaled values are lower than the respective DFT values despite the stronger polarization. Regarding the pressure dependence, only the M1 method can model the correct behavior (see panels (A) and (B) of Figure 12). Both the EQ and M2 methods show a negative pressure trend, which agrees with the previous DFT calculations. The performances of the MP2 and DFT calculations are nearly identical, and both show a good representation of the experimentally observed increase.

The EQ and M1 methods show similar results at 1 bar (1508.0/1432.6 cm^{1}) and a similar trend in the pressure dependency for the CH_3-def-middle mode. The pressure increase is 1 cm^{-1} up to 10 kbar, which underestimates the pressure response by 2 cm^{-1}. The M2 method is not able to predict the correct pressure trend. Regarding the CH_3-def-low mode, the M1 method can accurately represent the pressure change, whereas the EQ mode shows a slight overestimation and the M2 method underestimates slightly.

For the CH_3-def-low mode and the N-O stretch mode, the differences between the EQ and the two NEQ methods are more pronounced (as previously seen for the DFT calculations). In the case of the former mode, without scaling the wavenumbers, the values at ambient conditions are overestimated. The 1 bar value of the stretch mode is significantly higher for the equilibrium approach than for the two NEQ methods. Regarding the pressure dependence, the EQ approach overestimates the effect slightly, whereas the M1 method shows a good agreement. Where for the CH_3-def-low mode, the differences between MP2 and B3LYP are negligible, for the N-O stretch, the B3LYP calculations with the M1 method can accurately represent the pressure trend, whereas the MP2 show a slightly smaller increase. For the EQ calculations, the B3LYP shows a stronger overestimation of the pressure-induced increase than the MP2 calculations.

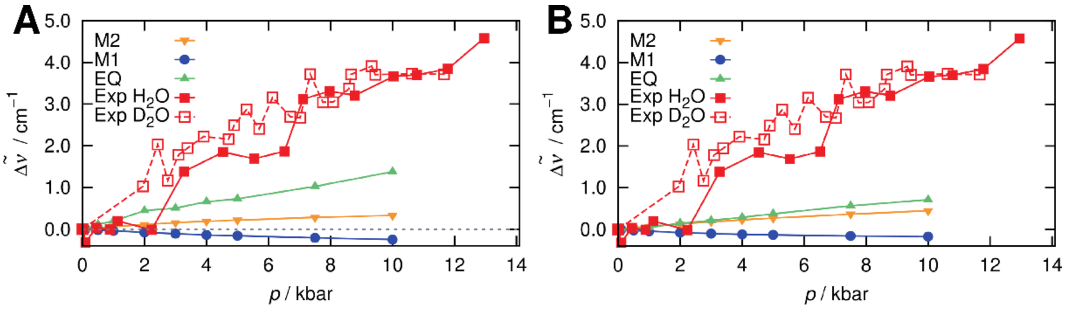

Figure 13 Pressure-induced wavenumber shifts of the cyanide anion calculated with MP2/6-311+G(d,p)/χ_{HNC} (Panel A) and B3LYP/6-311+G(d,p)/ χ_{HNC} (Panel B). The red squares represent the experimental data (open: in D_2O, filled in H_2O). The green triangles represent the equilibrium approach, the blue points the M1 non-equilibrium approach, and the red squares the M2 non-equilibrium approach. . Material from ref[27]. Raw and calculated data can be found in SI_4.1.

The calculated absolute wavenumbers by utilizing MP2/6-311+G(d,p)/χ_{HNC} are lower than the corresponding DFT calculations for the cyanide anion. This result contradicts the assumption that a higher solvent polarization is achieved with this level of theory. Since cyanide is a small molecule with only one vibrational mode, it could be possible that the small effect of pressure on the equilibrium distance significantly affects the frequency of the vibrational mode. Pressure-dependent geometry optimizations and numerical frequencies should be made to test this hypothesis as soon as they become available. The EQ and M2 methods can reproduce the pressure trend, whereas the M1 method fails to resemble the correct pressure trend. The best results are obtained with MP2 and the EQ methods; however, the increase is clearly underestimated.

Nonetheless, a possibility exists to attribute various degrees of equilibrium or non-equilibrium solvation contribution to different modes since the methodology used in this work does not account for intermediate states as, e.g., AIMD could. In this manner, for the N-O stretch mode, the equilibrium contribution could be stated to have a strong influence since it t resembles the experimentally obtained data in terms of absolute values at 1 bar (when the scaling is applied) for B3LYP/6-311+G(d,p), and it shows good results in representing the pressure dependence. Apparently, this could be a consequence of the large N-O dipole that interacts strongly with the surrounding water, causing more rapid water relaxation along the displacement coordinate. A similar picture can be observed for the cyanide anion, where the equilibrium approach best resembles the pressure-dependent mode. However, for the cyanide anion, the values at ambient conditions are better resolved by the two non-equilibrium methods. Therefore, no clear statement can be made, if the equilibrium or the non-equilibrium method can lead to better results both in terms of absolute value and pressure trend.

A similar observation was made by Reimann and Kaupp[173], who developed a new 3D RISM-SCF implementation, which allows for applying QC-geometry optimizations under the influence of the 3D RISM solvent. Based on these optimizations and numerical second derivatives, a consistent 3D RISM-SCF approach was introduced. This approach was termed the equilibrium approach since the full adaptation of the solvent is included. Additionally, a so-called frozen approach was presented, where the solvation potential of the equilibrium structure was unchanged used for the displaced structures. These two methods were benchmarked by comparing the calculated frequency shift (relative to gas-phase value) of tetramethyl urea (TMU) to the experimental reference. The results showed a significant difference in behavior between the equilibrium and the frozen approach. Especially for water, the experimental reference shift is -107 $\text{cm}^{-1,}$ and the corresponding values for the 3D RISM SCF model are -146 cm^{-1} and -67 cm^{-1} for the equilibrium and frozen methods, respectively. These results clarify that a sole description, either by an equilibrium state or a non-equilibrium situation, is insufficient to represent the real situation accurately.

4.1.6 Summary and outlook

Within the scope of this chapter, the numerical method developed by P. Kibies[94] for calculating pressure-dependent IR frequencies was modified and extended. The previously existing equilibrium method was subjected to a DFT benchmark in the first step. As model systems, the previously investigated TMAO was used, and on the other hand, the cyanide anion was considered, which is one of the simplest models to make pressure-dependent IR calculations in solution. A consistent EC-RISM variant was first used, where the B3LYP hybrid functional was applied in all iterations. Based on this, four combinations of coulomb potential calculation (φ and point charges) and solvent susceptibility were used (χ_{HNC} and χ_{SIM}).

In the equilibrium case, the absolute wavenumbers of the TMAO modes at ambient pressure are good compared to the experimental results. However, as soon as the pressure-dependent change was investigated, the combination of χ_{SIM} with point charges showed a massive overestimation of the frequency change. Instead, by utilizing the $\chi_{HNC,}$ the changes are significantly smaller. However, for some modes, the inverse pressure trend is calculated. To tackle this issue, a so-called non-equilibrium approach was introduced. This approach correctly reproduced the observed blue shift for every TMAO mode when the χ_{HNC} combined with exact electrostatics was used. The second benchmark system was the small cyanide anion, containing only one vibrational mode. Here, the performance is unsatisfactory for both the equilibrium and non-equilibrium methods. A possible reason could be that the 1 bar PCM geometries were used for all pressures, and the possible pressure-dependent geometry changes were not considered. While this effect is probably relatively small, it could strongly influence the wavenumbers for the small ion cyanide.

In the second part of this chapter, the MP2/6-311+G(d,p)/χ_{HNC}, considered the gold standard for EC-RISM calculations, was used to calculate the pressure-dependent wavenumbers for both test systems. Additionally, a pressure-dependent PMV correction of the excess chemical potential was utilized for the equilibrium method. For TMAO, the absolute wavenumbers at 1 bar increased compared to the B3LYP calculations. The non-equilibrium method fails to predict the experimentally observed pressure trend for the cyanide anion with the latter method. Overall, neither MP2/6-311+G(d,p)/χ_{HNC} nor B3LYP/6-311+G(d,p)/ χ_{HNC} shows superior results over the other in all observables. Especially in reproducing the pressure-dependent changes of the four TMAO modes, both methods showed similar results.

In summary, this chapter showed an improvement compared to the previously obtained pressure-dependent frequencies, mainly driven by the exact electrostatic potential to calculate Coulomb interactions and utilizing the χ_{HNC} to describe the solvent. However, at ambient conditions, no consistent picture regarding the quality of the wavenumbers compared to the experiment is gained. For some modes, the equilibrium approach shows better results and vice versa. However, it stands out that the equilibrium approach better resembles the pressure-dependent wavenumber increase of the N-O stretch mode and the vibrational mode of the cyanide anion. This result could be attributed to the relatively strong dipole moment that is involved in these displacements. Nevertheless, it should be noted here that the geometries considered are from a PCM optimization at 1 bar, and 1D RISM had to be used to calculate the excess chemical potential in the case of the non-equilibrium method. Moreover, in reality, the vibrational mode cannot be described by pure equilibrium or nonequilibrium solvation, but the solvent exhibits fast and slow relaxation processes.

While the absolute accuracy has room for improvement, EC-RISM can predict the pressure-dependent changes for TMAO relative good by applying either the non-equilibrium or equilibrium approach. This result is particularly relevant because Chapter 4.3 will examine the relative energetic pressure changes of peptides, for which are no direct experimental reference data available. Therefore, the performance of EC-RISM in calculating pressure trends must be evaluated beforehand and compared to experimental results.

As soon as geometry optimizations and numerical frequency calculations with EC-RISM are possible, these should be carried out pressure-dependently.

4.2 Calculations of EPR-Parameter in solution

After NMR calculations with the EC-RISM method were successfully performed in the past, this chapter aims to calculate EPR parameters with EC-RISM for the first time at ambient conditions. The nitroxide (3R,4S)-2,2,3,4,5,5-hexamethylimidazolidin-1-oxyl (HMI) was chosen as the model system (see Figure 14), which is, on the one hand, a relative rigid molecule, and on the other hand, it contains a protonatable group that allows the pH-dependent measurement of the HFCC. The strategy for calculating NMR parameters in previous publications primarily focused on investigating only a few or even only one minimum conformer. Furthermore, no explicit solvent molecules were taken into consideration. In this chapter, the minimum structures will be investigated first, and subsequently, calculations on thermally averaged structures originating from an *ab-initio* MD will be performed.

The overall goal is to explore how well EC-RISM can calculate EPR parameters with and without explicit water molecules. Therefore, a comparison between EC-RISM and the implicit solvent model CPCM on the geometry-optimized structures was performed, where the target parameter is the isotropic hyperfine coupling constant. In the second step, computations are performed with both EC-RISM and CPCM for the snapshots obtained from the trajectory; additionally, a comparison was made between vertical desolvation and the explicit consideration of explicit water molecules. Furthermore, in this chapter, for the first time, the DLPNO-CCSD method was used in combination with EC-RISM to obtain the highest possible accuracy compared to the experimental values. In the second part of the chapter, the full EPR spectrum will be calculated and compared to the experimentally obtained spectrum. The last part will deal with the pressure-dependent changes of the EPR parameters calculated with EC-RISM. Parts of this chapter were already published in ref[132].

4.2.1 HFCC parameters for the nitroxy nitrogen of optimized HMI structures

As a starting point, the results from the clustering should be analyzed. The clustering resulted in only one cluster; the corresponding optimized structure is shown on the left side of. Figure 14.

However, another optimized structure is shown in Figure 14, where an inversion of the pyramidal configuration occurred at the N-CH_3 group. This configurational change has not been observed during the AIMD, but it could be assumed that a bias is included when focusing on a single configuration of the N-CH_3 group. QC calculations were performed with different levels of theory and solvent models in water and methanol to estimate how strongly the second configuration is populated in solution. On the one hand, the energetic difference between the two configurations is calculated, and on the other hand, the calculated A^{iso} parameter for the nitrogen of the N-O motif is compared. The corresponding data is summarized in Table 3. L. Galazzo measured experimental reference values for the nitrogen HFCC of 44.87 +0.14 MHz and 43.00 MHz[293] in water and methanol, respectively.

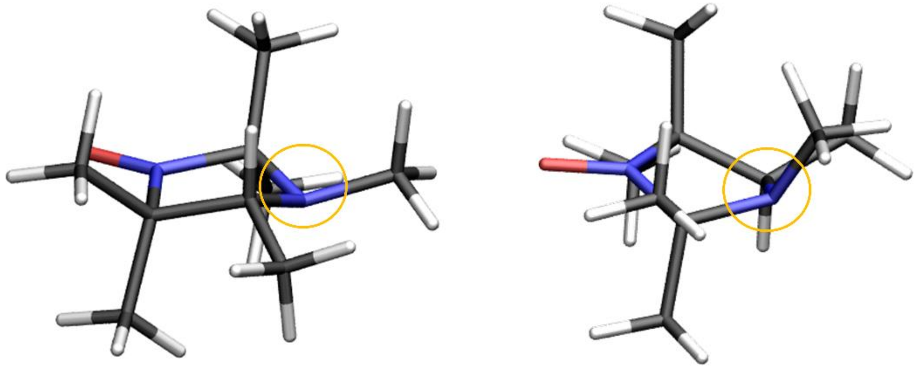

Figure 14 The two optimized structures of HMI with B3LYP-D3 def2-TZVPP/CPCM. The (3R,4S)-diastereomer (which is the one used in the AIMD) is depicted on the left, and the (3R,4R)-diastereomer is shown on the right.

Table 3: Nitrogen HFCCs of CPCM optimized geometries (in MHz) of (3R,4S)-HMI and (3R,4R) HMI calculated with different DFT methods for EC-RISM and CPCM in the solvents Water and Methanol. For calculations, the def2-TZVVP with decontracted s-function was used. The free energy difference (*(3R,4R) (3R,4S-*) between the two main diastereomers is shown in kcal/mol. The experimentally measured value for the HFCC is 44.87 MHz[132] in water and 43.00 MHz[293] in methanol. Optimized structures can be found in SI_4.2.

Solvent	Method	EC-RISM			CPCM		
		A^{iso} (3R,4S)	A^{iso}(3R,4R)	ΔG(4R-4S)	A^{iso} (*3R,4S*)	A^{iso} (*3R,4R*)	ΔG
H_2O	revPBE0	32.6	35.3	4.40	30.2	32.8	4.98
	revPBE0-D3	32.5	35.8	3.64	30.0	33.5	4.44
	B3LYP	31.7	34.7	4.07	29.8	32.6	4.90
	B3LYP-D3	31.6	34.9	3.48	29.7	32.9	4.52
MeOH	revPBE0	30.6	33.1	4.59	30.0	32.6	4.99
	revPBE0-D3	30.3	33.8	3.90	29.8	33.3	4.45
	B3LYP	29.8	32.7	4.54	29.6	32.5	4.90
	B3LYP-D3	29.8	33.0	4.00	29.5	32.8	4.52

Starting with the calculations in water, for EC-RISM, the free energy differences between the two diastereomers range from 3.48 kcal/mol (B3LYP-D3) to 4.40 kcal/mol (revPBE0). It is clearly noticeable that by adding the D3 correction, a lowering of the free energy difference between the two conformers is observed. For CPCM, even higher free energy differences are calculated, ranging from 4.44 kcal/mol (revPBE0-D3) to 4.90 kcal/mol (B3LYP). These values clearly indicate a very large population difference between the two diastereomers, so the (3R,4S)-diastereomer can certainly be assumed to be the main conformer. The calculated free energy differences between the two diastereomers for methanol are even larger than for the corresponding water calculations.

However, it should be noted that the results presented above for the free energy difference are based on DFT free energies alone without adding a PMV correction term. Especially for methanol, no PMV correction was not available at the time. However, the PMV correction has proven to be an indispensable instrument for calculating accurate free energies in solution with EC-RISM. Therefore, additional free energies in water were calculated with the MP2/6-311+G(d,p)/EC-RISM//B3LYP/def2-TZVPP/CPCM and revPBE0-D3/def2-TZVPP/EC-RISM//revPBE0-D3/def2-TZVPP/CPCM level of theories. The corresponding data is shown in Table 5. These calculations show that the PMV correction further increases the relative energy difference between the two diastereomers.

Table 4 Relative Gibbs energies G_{sol} (in kcal mol^{-1}) and contributions (for EC-RISM:[1] sum of intramolecular energy E_{sol}, excess chemical potential μ^{ex}, partial molar volume correction $c_V V_m$ with c_V = -0.1022 kcal mol^{-1} Å^{-3}, based on the V_m which was calculated via the total correlation function) from calculations on optimized solution structures. To correct the methanol 3D RISM calculations, the PC+ correciton[179] with a correction parameter c_{pc+}=-0.09696 kcal mol^{-1} Å^{-3} and a $\varkappa$=2.34204 $\cdot 10^9$ Pa^{-1} was used. Material from ref[132]

Method / Isomer	E_{sol}	μ^{ex}	$c_V V_m$/ $c_{pc+} V_m$/	G_{sol} (EC-RISM)
MP2/6-311+G(d,p)/EC-RISM//B3LYP/def2-TZVPP/CPCM (water)				
(3*R*,4*S*)	0	0	0	0
(3*R*,4*R*)	5.38	-1.42	0.47	4.43
revPBE0-D3/def2-TZVPP/EC-RISM//revPBE0-D3/def2-TZVPP/CPCM (water)				
(3*R*,4*S*)	0	0	0	0
(3*R*,4*R*)	4.95	-1.63	0.53	3.84
MP2/6-311+G(d,p)/EC-RISM//B3LYP/def2-TZVPP/CPCM (MeOH)				
(3*R*,4*S*)	0	0	0	0
(3*R*,4*R*)	5.34	-1.12	0.76	4.98
revPBE0-D3/def2-TZVPP/EC-RISM//revPBE0-D3/def2-TZVPP/CPCM (MeOH)				
(3*R*,4*S*)	0	0	0	0
(3*R*,4*R*)	4.96	-1.07	0.86	4.76

The obtained HFCCs, calculated with EC-RISM, for the main diastereomer range from 31.6 MHz to 32.6 MHz, whereby the revPBEO variants result in higher couplings than the B3LYP calculations. The minor diastereomer shows an increase of nearly 3 MHz in the A^{iso} for all variants. With CPCM, the calculated A^{iso} parameters are 2.4 MHz and 2.0 MHz lower for the revPBE0 and B3LYP calculations, respectively. Therefore, EC-RISM, compared to CPCM, is superior in reproducing the experimental results, but the absolute accuracy is rather sobering, with deviations of more than 12 MHz.

The experimental difference in HFCCs between methanol and water is approximately 2 MHz. The EC-RISM calculated HFCC in methanol ranges from 29.8 MHz (for the B3LYP variants) up to 30.6 MHz (for revPBE0 calculations). Hence, water and methanol differ from 1.8 MHz to 2.1 MHz. While the absolute accuracy is still dissatisfying, the relative difference between the two solvents can be accurately represented by EC-RISM.

In contrast, the differences between methanol and water for CPCM are marginal, with changes of 0.2 MHz. In this case, CPCM cannot predict the experimentally observed shift. Apart from that, it should be noted that for EC-RISM, a specific solvent susceptibility has to be generated, whereas in the case of CPCM, simply changing the dielectric constant and the corresponding refractive index is sufficient to describe a new solvent. To generate the solvent susceptibility, on the one hand, experimentally measured dielectric constants and densities are needed, and on the other hand, a suitable force field is required.

4.2.2 HFCCs parameters of nitroxy nitrogen of HMI calculated for 1000 snapshots with revPBE0/def2-TZVPP

After it was shown in the previous chapter that for the optimized structures, EC-RISM gives better results than CPCM relative to the experiment, the absolute accuracy was still quite far away from the experimental observations. One has to consider that only DFT calculations were used, and only one optimized structure was considered, which did not include explicit solvent molecules in the calculation. Better levels of theory will be discussed later in this chapter, whereas in this section, the influence of multiple structures and explicit solvent molecules on the A^{iso} parameter will be studied. Additionally, the results for the HFCC of the nitroxy oxygen of HMI can be found in chapter 6.7.

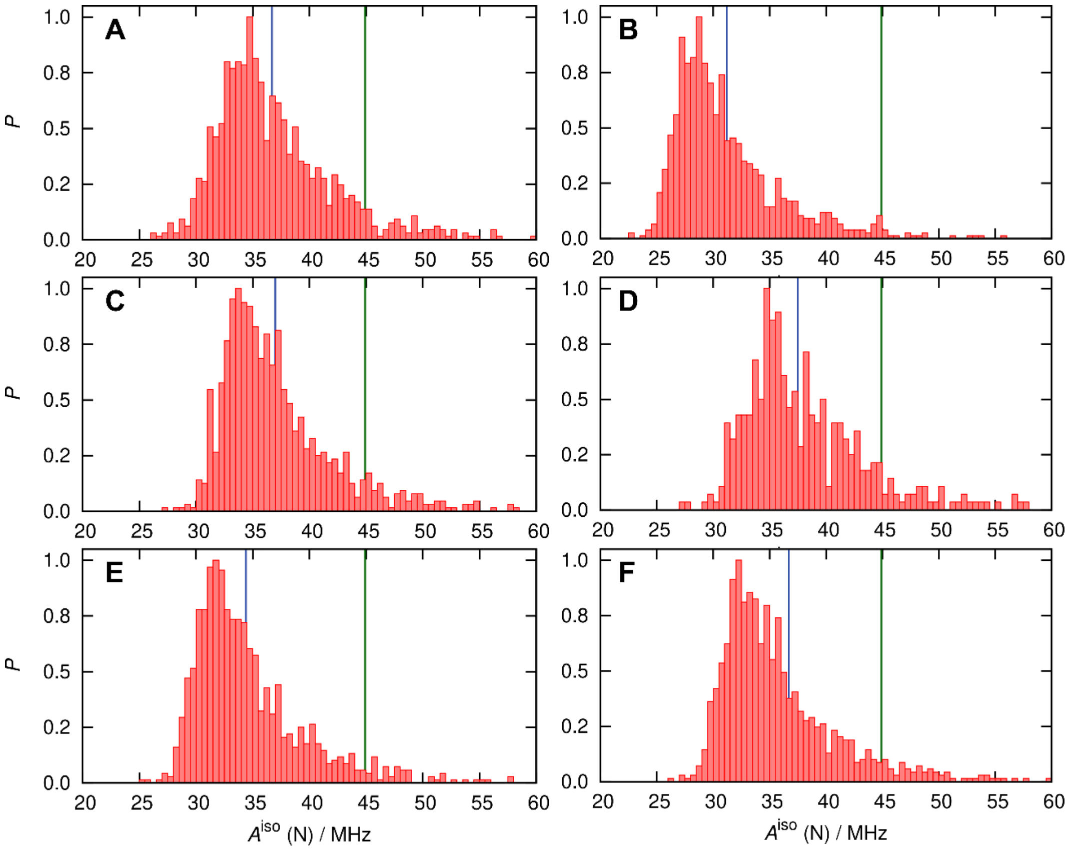

Figure 15 Normalized probability histograms of A^{iso} values of the nitroxy nitrogen of the EPR spin-probe HMI at ambient conditions calculated with revPBE0-D3/def2-TZVPP with decontrated s function on 1000 snapshots. Panel (A) shows the probability distribution (Mean: 36.7 MHz) calculated via the QM/MM approach (by V. A. Tran), including the two first water solvation shells around the N-O motif explicitly and approximating the remaining explicit water molecules via TIP3P point charges. Panel (B) shows vacuum calculations on the vertically desolvated structures (Mean: 31.2 MHz). Panel (C) shows the distribution calculated with EC-RISM/PSE-3 on the vertically desolvated structures (Mean: 37.0 MHz). In Panel (D), the distribution was calculated with EC-RISM/PSE-3 (Mean: 37.5 MHz) on the structures, explicitly including the two first solvation shells around the HMI oxygen. Panel (E) shows the distribution calculated with CPCM solvation on the vertically desolvated structures (Mean: 34.4 MHz). Panel (F) depicts the distribution calculated with CPCM solvation on the structures, containing the explicit solvent molecules of the first two solvation shells around the oxygen of HMI (Mean: 36.7 MHz). The blue line in each panel represents the average value of the corresponding distribution computing using the numerical A^{iso} data that underlie the respective histogram. The green vertical line in the panels represents the experimental value, and each distribution is separately normalized. Material from ref[132]. Raw data and structures for the EC-RISM and CPCM calculations can be found in SI_4.2._HMI,

A subset of 1000 structures, in equal temporal distances extracted every 0.2 ps from the 200 ps AIMD simulation, was used for the DFT calculations. In Figure 15, histograms of the probability distributions of the A^{iso} parameters for the subset calculated with revPBE0/def2-TZVPP with decontracted s-functions are shown. At this point, six different setups are compared with each other. Panel (A) shows a QM/MM approach in which the first two solvation shells around the nitroxyl oxygen atom are explicitly represented, and additionally, TIP3P charges approximate the remaining water molecules of the simulation box. These calculations were performed by V. A. Tran and serve as the benchmark for EC-RISM and CPCM calculations. For the QM/MM setup, a mean isotropic hyperfine coupling constant of 36.7 MHz (an overview of HFCC values is given in Table *5*) could be calculated, resulting in a difference of 7.5 MHz compared to the experimental value. By looking at the histograms, it can be seen that a A^{iso} value between 30 MHz and 45 MHz is obtained for most snapshots with a few outliers above 45 MHz.

Panel (B) of Figure 15 shows vacuum calculations on the vertically desolvated structures. Compared to the QM/MM method, a clear shift towards lower HFCCs can be seen. The mean isotropic HFCC is 31.2 MHz, which is 5.6 MHz lower than the QM/MM method, and, therefore, using a solvent model significantly improves the quality of the calculations.

Table 5: HFCC calculated on revPBE0-D3/def2-TZVPP/CPCM optimized structures (first three rows) and averaged HFCCs of the nitrogen of nitroxide group of HMI from different calculation settings. For all calculations, the def2-TZVPP with decontracted s-functions basis set was used. VD represents the vertical desolvated dataset, and SSS represents the subset, including the first two solvation shells. Raw data can be found in SI_4.2. Ros denoted with dagger‡ were calculated by V.A. Tran. Material was published in ref[132].

Method	A^{iso} / MHz
revPBE0/def2-TZVPP/CPCM Min	30.0
revPBE0/def2-TZVPP/EC-RISM Min	32.5
DLPNO-CCSD/def2-TZVPPP/EC-RISM Min	38.1
revPBE0-D3/def2-TZVPP/Vac‡	31.2
revPBE0-D3/def2-TZVPP/QM/MM SSS‡	36.7
revPBE0-D3/def2-TZVPP/CPCM VD	34.5±0.17
revPBE0-D3/def2-TZVPP/CPCM SSS	36.7±0.16
revPBE0-D3/def2-TZVPP/EC-RISM VD	37.0±0.17
revPBE0-D3/def2-TZVPP/EC-RISM SSS	37.5±0.16
DLPNO-CCSD/def2-TZVPP/EC-RISM VD(400)	42.7±0.26
DLPNO-CCSD/def2-TZVPP/QM/MM SSS‡ (400)	40.5
revPBE0/def2-TZVPP/EC-RISM VD(400) DFT*	36.8±0.27
revPBE0/def2-TZVPP/EC-RISM VD(400) HF*	37.7±0.27

Panels (C) and (D) of Figure 5 show revPBE0/def2-TZVPP/PSE-3 EC-RISM calculations, whereby Panel (C) is based on the vertically desolvated structures, and Panel (D) explicitly includes the first two solvation shells in the calculations. In the former case, a mean A^{iso} of 37.0 MHz is obtained, whereas for the latter one, an average A^{iso} of 37.5 MHz is obtained. With EC-RISM on the vertically desolvated structures, a similar value to the QM/MM method can be obtained and a deviation of 7.2 MHz relative to the experimental results. Thus, a massive improvement compared to vacuum calculations is obtained. If the explicit waters are included, the result can be improved further by 0.5 MHz.

The revPEB0/def2-TZVPP/CPCM calculations are shown in Panels (E) and (F) of Figure 15. Panel (E) shows the histogram for the desolvated structures, and Panel (F) again includes the first two solvation shells around the nitroxyl-O in the calculations. Interesting is the result for the former variant, where a mean isotropic hyperfine coupling constant of 34.4 MHz could be calculated. Especially in comparison to the calculations with EC-RISM on the vertically desolvated structures, a significantly worse result is obtained with a difference of -2.6 MHz. This result reproduces the previously observed results for the optimized structures very well, whereby with EC-RISM, about 2 MHz larger A^{iso} parameters were obtained compared to CPCM. However, using CPCM as a solvent model can significantly increase the accuracy compared to the vacuum-only calculations. Moreover, the influence of the explicit water in the CPCM case is more significant than EC-RISM. Thus, by adding the first two solvation shells, a mean HFCC of 36.7 MHz can be obtained; this precisely equals the mean value of the QM/MM calculations.

Three central conclusions can be derived from the calculations on the 1000 AIMD snapshots. First, it was shown that for the vertically desolvated structures (the computations that take the least time), EC-RISM produces significantly better results than CPCM. This result was also observed before for the geometry-optimized structures. The second finding is that EC-RISM gives approximately the same results for the vertically desolvated structures as in the case where the explicit waters were included. Therefore, the influence of the explicit quantum mechanical waters does not seem much stronger on the isotropic shielding constants than the effect obtained due to the polarization of the point charges. The third fundamental finding is that significantly better results are obtained by considering the whole trajectory rather than the geometry-optimized structure. This seems odd at first glance since HMI is a relatively rigid molecule, and there should not be large structural differences. Therefore, in the next chapter, the influence of structural parameters on the A^{iso} parameters will be investigated.

4.2.3 Dependence of HFCC parameters on structural parameters

This chapter aims to investigate the dependence of the HFCCs on structural properties. From the literature, it is well known that two major contributions to HFCC of the nitrogen in nitroxides exit, on the one hand, the surrounding solvent structure and, on the other hand, the nitrogen out-of-plane movement. The effect of the solvent polarity was already described previously; however, in this chapter, a more detailed analysis of the occurring number of hydrogen bonds and their influence on the HFCC shall be performed.

For this purpose, the two subsets are again compared with each other, although it should be noted that in the case of the vertically desolvated structures, there are no explicit hydrogen bonds. However, the surrounding water molecules influenced the associated vertically desolvated structures in the AIMD. The advantage of including the vertical desolvated subset in the comparison is that a certain baseline effect can be determined since the different structures show alone an effect on the HFCC, and the explicit waters do not influence them.

Table 6 Mean A^{iso} values in MHz of the nitroxy nitrogen of the EPR spin-probe HMI ((3R,4S)) at ambient conditions calculated with revPBE0-D3/def2-TZVPP with decontracted s function for the 1000 snapshots shown as a function of the number of hydrogen bonds around the oxygen atom of the nitroxy group. The corresponding distributions can be seen in Figure 16. The abbreviation SSS means that the first two solvation shells are considered explicitly. Raw data can be found in SI_4.1.

H-bonds	CPCM VD	CPCM SSS	EC-RISM VD	EC-RISM SSS	No. Snapshots
1	34.2±0.49	34.8±0.46	36.6±0.50	36.0±0.45	121
2	34.5±0.22	36.6±0.21	37.1±0.22	37.5±0.21	595
3	34.5±0.31	37.7±0.29	37.2±0.32	38.3±0.28	270
4	34.3 ±1.05	37.5±1.33	37.1±1.14	38.2±1.33	9
All	34.4±0.17	36.7±0.16	37.0±0.17	37.5±0.16	1000

Four structures were obtained from the 1000 snapshots from the AIMD that did not form determinable hydrogen bonds, and 121 snapshots contained one single hydrogen bond. In 595 snapshots, i.e., more than half of all snapshots, two hydrogen bonds could be observed, whereas 270 structures had three hydrogen bonds. Nine and one structures could be assigned to four and five hydrogen bonds. For the sake of clarity, the zero and five hydrogen bond structures are not considered in this chapter.

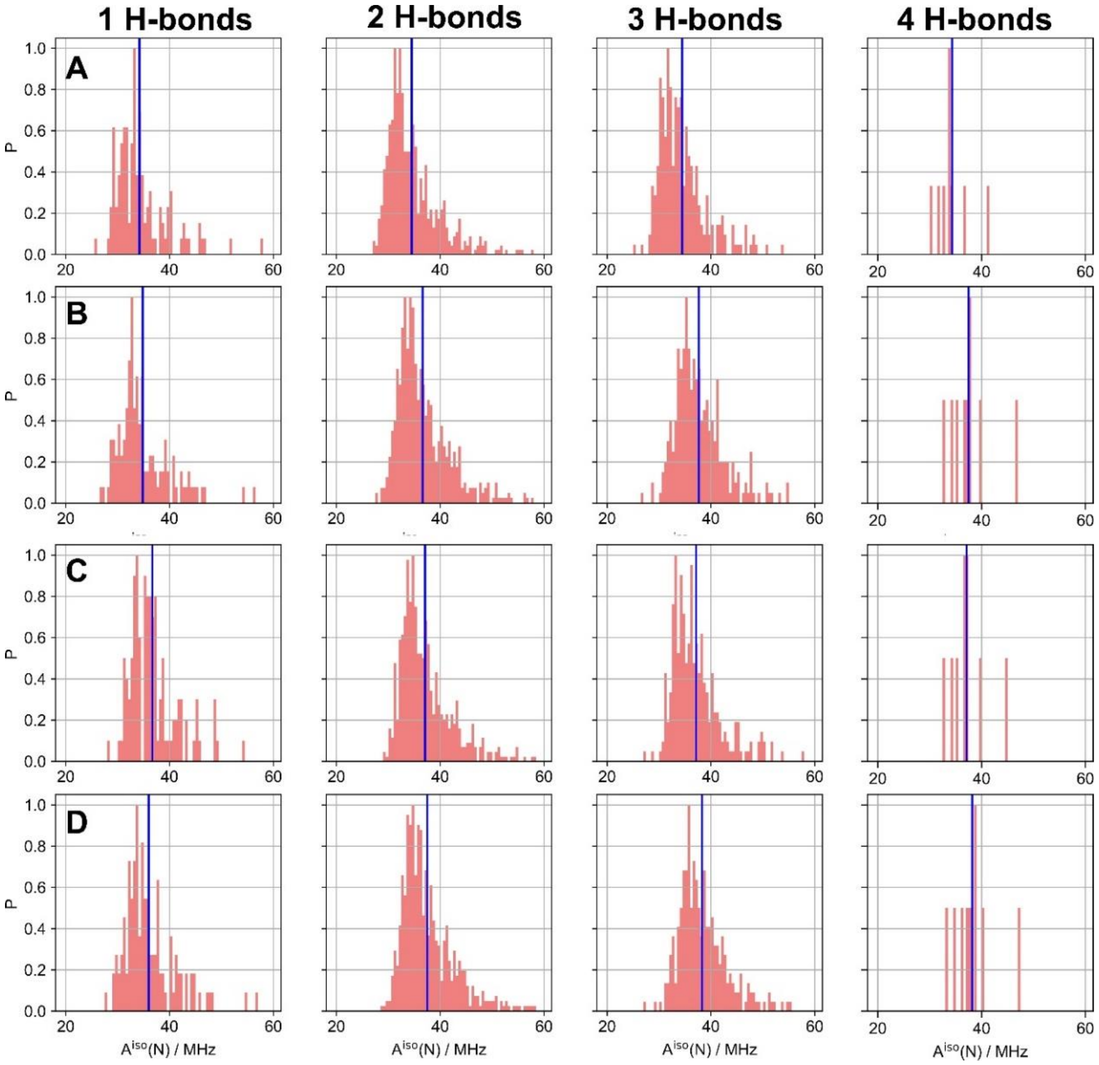

Figure 16 Probability histograms of A^{iso} values of the nitroxy nitrogen of the EPR spin-probe HMI at ambient conditions calculated with revPBE0-D3/def2-TZVPP with decontrated s function on 1000 snapshots containing a subset, divided into the dependence of the number of hydrogen bonds to the oxygen of the nitroxy group. In the histograms, there are 121 underlying snapshots for the 1 H-bond, 595 snapshots for the 2 H-bonds, 270 snapshots for the 3 H-bonds, and 9 snapshots for the 4 H-bonds. Panel (A) shows the distribution calculated with CPCM solvation on the vertically desolvated structures. Panel (B) depicts the distribution calculated with CPCM solvation on the structures that contain the explicit solvent molecules of the first two solvation shells around the oxygen of HMI. Panel (C) shows the distribution calculated with EC-RISM/PSE-3 on the vertically desolvated structures. In Panel (D), the distribution was calculated with EC-RISM/PSE-3 on the structures, explicitly including the two first solvation shells around the HMI oxygen. The corresponding average values are summarized in Table 6. Raw data can be found in SI_4.2.

In the case of CPCM, the vertically desolvated structures (Figure 16 Panel (A)), show average values of 34.2 MHz for one hydrogen bond, 34.5 MHz for two and three hydrogen bonds, respectively, and 34.3 MHz for four hydrogen bonds. Thus, only a marginal structural baseline effect is obtained due to the presence of hydrogen bonds in the simulation. As soon as the hydrogen bonds are explicitly included, a clear dependence on the number of hydrogen bonds becomes apparent. When one hydrogen bond is formed, the average HFCC is 34.8 MHz, whereas as soon as two explicit hydrogen bonds are present, the HFCC increases by 1.8 MHz to 36.8 MHz. If a structural baseline effect of 0.3 MHz is assumed, the increase is 1.5 MHz, solely induced by the presence of the hydrogen bonds. The presence of a third hydrogen bond leads to a further increase of 1.1 MHz. For all hydrogen bond configurations, as soon as explicit hydrogen bonds are included in the calculation, higher A^{iso} parameters are obtained. A different picture is obtained for EC-RISM-based calculation. Here, for the vertically desolvated subset, the average values are 36.6 MHz (1 H-bond), 37.1 MHz (2 H-bonds), 37.2 MHz (3 H-bonds), and 37.1 MHz for the structures with four H-bonds. The baseline structural effect is 0.5 MHz, and hence slightly higher than for CPCM. The major difference occurs by looking at the explicitly solvated structures; here, for the case with one single explicit hydrogen bond, the mean HFCC is 36 MHz and thus 0.6 MHz smaller than for the vertically desolvated structures under the influence of pure EC-RISM solvation. This means that the EC-RISM charge distribution exerts a stronger polarization than the explicit water configurations, which form only one hydrogen bond. When two explicit hydrogen bonds are present in combination with EC-RISM, the surrounding explicit environment generates a stronger polarization than the pure EC-RISM calculations on the vertically desolvated structures, with a difference between these two methods of 0.4 MHz. The effect becomes even stronger with three explicit hydrogen bonds, with a mean HFCC of 38.3 MHz.

Another point that stands out is the fact that for all the sub-datasets analyzed in this chapter, the mean A^{iso} is significantly larger than the comparable value for the optimized structures. Consequently, a relatively strong structural effect must influence the A^{iso} parameters. From chapter 3.2, we know that in the framework of the relatively simple VB theory, the A^{iso} parameter depends very strongly on the spin densities located at the nitrogen and the oxygen. In the past, several studies have been conducted to assess the influence of the nitrogen out-of-plane movement on the HFCC.[263,278–280] The corresponding relevant improper dihedral of HMI is depicted in Figure 17.

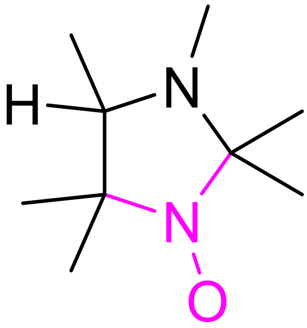

Figure 17 Two-dimensional representation of HMI. The magenta-colored atoms show the CNOC-improper dihedral, reflecting the out-of-plane movement.

In Figure 18 Panel (C), the nitrogen out-of-plane movement distribution is shown for the 1000 snapshots of the trajectory. The blue line indicates the corresponding value of the optimized structure, which is -0.65°. The value of the improper dihedral ranges from -20° to 20°, with a nearly symmetric distribution. A small maximum can be found at 5°; however, as mentioned in chapter 4.2.1, only one optimized structure was found after optimizations in CPCM solvation. Most of the structures can be found in a range between -5° to 5°. In panels (A) and (B) of Figure 18, the distribution of HFCCs versus the underlying improper dihedral is presented for revPBE0/def2-TZVPP/CPCM (Panel (A)) and revPEB0/def2-TZVPP/EC-RISM (Panel (B)). The plots clearly show that a stronger out-of-plane motion results in a greater HFCC. An increase of the A^{iso} parameters is observed for both positive and negative deflections, respectively. For an improper dihedral of 20°, there is almost a doubling compared to the values for the planar geometry.

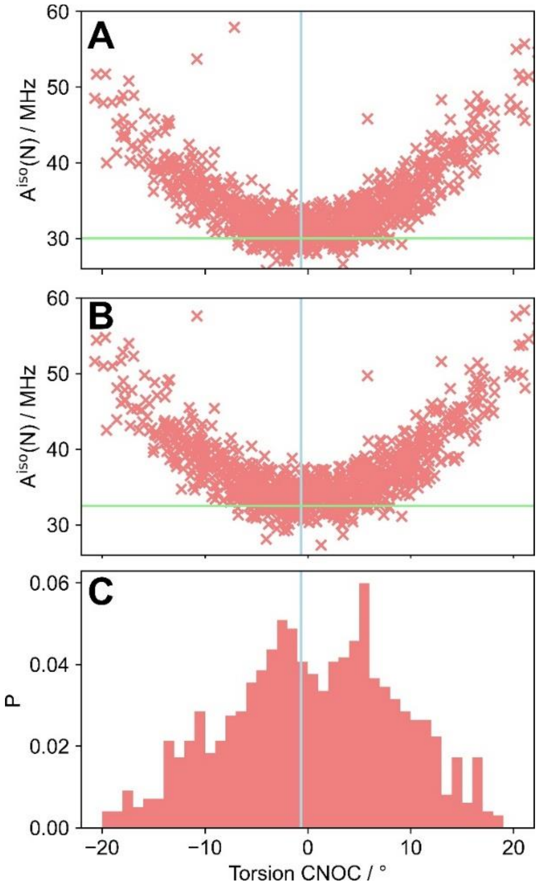

Figure 18: Panel (A) shows the distribution of A^{iso} parameters with respect to the improper torsion NCOC calculated with revPBE0/def2-TZVPP/CPCM with decontracted s-functions. In panel (B), the distribution of A^{iso} parameters with respect to the improper torsion NCOC calculated with revPBE0/def2-TZVPP/EC-RISM with decontracted s-functions is depicted. The blue line shows the CNOC improper torsion value for the optimized structure using revPBE0-D3/def2-TZVPP. The horizontal green lines show the A^{iso} value calculated for the optimized structures with the corresponding level of theory (see Table 3). Panel (C) shows the histogram of the CNOC improper torsion, which is shown in Figure 17. Raw data can be found in SI_4.2.

The nitrogen HFCC increases significantly with the pyramidalization of the NO group. This effect can also be explained by applying the VB model introduced in chapter 3.2. As long as the nitroxide moiety is planar, the direct contribution to the singly occupied molecular orbital (since the SOMO is basically the π^* and no s-orbital contribution is present) vanishes.[263] As soon as a deviation of the planarity comes into place, the s atomic orbitals contribute to the SOMO and, therefore, increase the spin density at the nitrogen and the HFCC. The optimized structure of HMI has almost complete planar geometry, but due to thermal fluctuations, there are slightly pyramidal deflections and, therefore, a larger HFCC over the entire trajectory compared to the optimized value is observed. The left resonance structure from Figure 3 is favored in the planar configuration, whereas a stronger out-of-plane movement favors the right resonance structure.

In summary, the explicit hydrogen bonds influence the HFCC, but their effect is not the most substantial. From the AIMD, structures with two explicit hydrogen bonds could be extracted most frequently, and structures with one and three hydrogen bonds are the second most common. The number of hydrogen bonds increases the HFCCs because the polarity of the solvent is increased in close proximity to the nitroxide group. However, EC-RISM calculations on the vertically desolvated structures can reproduce this effect quite well. Therefore, at least for the HMI nitroxide, it can be concluded that the inclusion of explicit waters is not necessary to characterize the effect of the solvent on the HFCC. This is particularly relevant when considering larger systems, such as protein fragments, where the reduction of explicit atoms within the QC calculations is essential. However, to obtain accurate results, it is crucial to include the fluctuations of the out-of-plane motion of the NO moiety. Thus, a thermal averaging of different structures must be performed to improve absolute accuracy. Therefore, one or two optimized structures are insufficient, and some kind of molecular dynamic simulation (in the best case AIMD) or Monte-Carlo simulation must be utilized for larger systems.

4.2.4 HFCC parameters of nitroxy nitrogen of HMI calculated for 400 snapshots with DLPNO-CSSD/def2-TZVPP

The HFCCs calculated with revPBE0/def2-TZVPP deviate from the experimental value by about 7-8 MHz. One reason for the still relatively large deviation is the level of theory utilized here. Hybrid functional calculations are not the best way to describe core-level spin polarization, and in particular, highly correlated wavefunction approaches have shown excellent agreement with experimental reference data. The following chapter performs calculations to more accurately depict these effects with the domain-based pair natural orbital coupled-cluster singles and doubles (DLPNO-CCSD) method.[129–131]

These are the first attempts to couple the DLPNO-CCSD method with the EC-RISM formalism. However, it must be noted here that, like in the case of EC-RISM computations with the theory level MP2, the iterations are performed using the HF formalism, i.e. concretely, that in the last iteration, in which the DLPNO-CCSD calculation takes place, the polarizing solvent, itself was polarized by HF electron density. A consistent EC-RISM cycle in which a DLPNO-CCSD calculation is performed in each iteration is not practicable with current resources.

The focus is on the revPBE0-D3/def2-TZVPP/CPCM optimized structure since this one was generated with the same level of theory used in the AIMD and vertical desolvated subset containing 400 snapshots. The hyperfine coupling constant is 38.1 MHz and, therefore, 5.6 MHz higher than the value obtained with revPBE0-D3/def2-TZVPP. Still, a gap of 6.3 MHz to the experimental benchmark value is observed. Nonetheless, by choosing the better theory level, a massive improvement could be obtained; however, by choosing the DLPNO-CCSD theory level; however, the computational time and requirement at the computational resources are significantly increased. Thus, the complete trajectory of 1000 snapshots was not used to validate the effect of thermal averaging, but only 400 snapshots were included in the consideration, with a time interval of 500 fs between two snapshots.

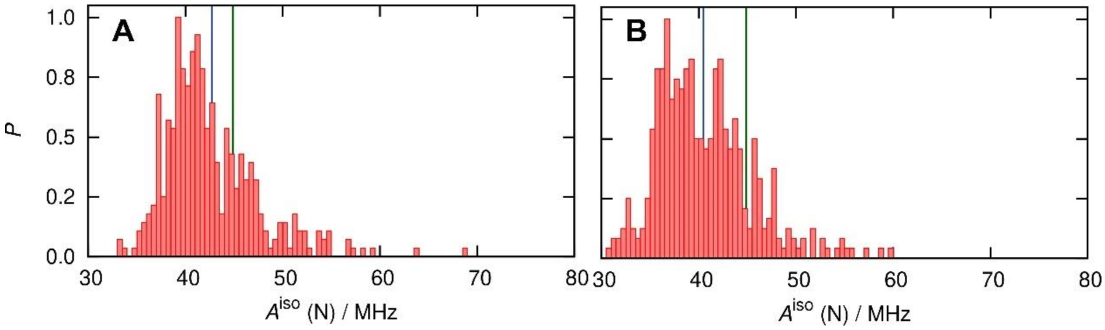

Figure 19 Probability distributions of A^{iso} values of nitroxy nitrogen calculated with DLPNO-CCSD/def2-TZVPP(s-decontracted) on the subset containing 400 structures. In Panel (A), the EC-RISM calculations on the VD dataset are shown (Mean 42.7 MHz). Panel (B) shows the probability distribution (Mean: 40.5 MHz) calculated via QM/MM approach, including the two first water solvation shells around the N-O motif explicitly and approximating the remaining explicit water molecules via TIP3P point charges. Raw and calculated data can be found in SI_4.2. This material was already published in ref[132]

The corresponding probability distributions are depicted in Figure 19, whereby in Panel (A), the DLPNO-CCSD/def2-TZVPP/EC-RISM distribution on the vertically desolvated structures is shown. In Panel (B), the probability distribution obtained by the QM/MM approach, where the two first solvation shells are explicitly considered, and the remaining waters are approximated by TIP3P point charges, is shown. For the latter one, an average value of 40.5 MHz for the A^{iso} is obtained (see Table *5*). This is an improvement of 3.8 MHz compared to the corresponding calculations with revPEB0-D3/def2-TZVPP. The difference to the experimental benchmark value is only 4.3 MHz. The average value of the HFCC with DLPNO-CCSD/def2-TZVPP/EC-RISM on the vertically desolvated structures is 42.7 MHz. Therefore, this new approach results in an increase of 5.7 MHz compared to the revPBE0-D3/def2-TZVPP calculations on the vertically desolvated subset. The 5.7 MHz increase is in good agreement with the 5.6 MHz increase obtained for the optimized structure, and the difference from the experimentally obtained value is only 2.1 MHz.

Something that stands out is that the difference between EC-RISM for the vertically desolvated structures and the QM/MM method in the case of DLPNO-CCSD/def2-TZVPP is 2.2 MHz, whereas, in the case of revPBE0-D3/def2-TZVPP, the difference is only 0.3 MHz.

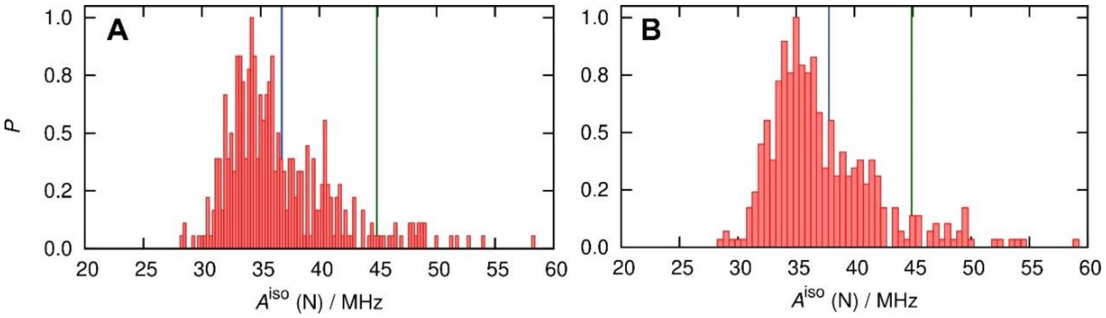

Figure 20 Probability distributions of A^{iso} values of nitroxy nitrogen calculated with revPBE0-D3 functional and s-decontracted def2-TZVPP basis set using the AIMD ensemble of vertically desolvated spin probe configurations. Panel (A) shows the consistent DFT electron density to calculate the exact electrostatic potential (average value 36.8 MHz), (B) calculated with HF electron density, obtained from DLPNO-CCSD/def2-TZVPP calculations, to calculate the exact electrostatic potential (average value 37.8 MHz). The blue line in each figure represents the average value of the corresponding distribution. The green vertical line represents the experimental value, and each distribution is separately normalized. Raw data can be found in SI_4.1. This material was already published in ref[132]

One possible reason could be the fact that in the case of revPBE0/def2-TZVPP, the consistent DFT density was used to calculate the electrostatic potential, yet, for DLPNO-CCSD, the HF density was used. To investigate this hypothesis, two different revPEB0-D3/def2-TZVPP calculations were performed for the 400 vertically desolvated structures. On the one hand, the consistent route was chosen, and on the other hand, the HF density extracted from the DLPNO-CCSD calculations was used. The corresponding probability distributions are shown in Figure 20. For the approach with consistent revPBE0-D3/def2-TZVPP, an average HFCC of 36.8 MHz is obtained. This result agrees with the 37.0 MHz obtained from the 1000 snapshots calculated with revPBE0-D3/def2-TZVPP. Thus, the subset containing the 400 snapshots is reliable for calculating the HFCCs. However, if the solvent charge distribution is used, which was polarized by the HF-ESP, an average A^{iso} value of 37.8 MHz. Thus, the HF-density leads to a solvent distribution that causes a stronger solute polarization. Due to this effect, the difference between the QM/MM approach with DLPNO-CCSD/def2-TZVPP and DLPNO-CCSD/def2-TZVPP can be explained. In addition, two striking outliers appear in the distribution of the DLPNO-CCSD/def2-TZVPP/EC-RISM calculations, for which there are no explainable and recognizable reasons. Nonetheless, the inconsistency resulting from the different electrostatics is smaller than the effect generated from the improved electronic structure at the DLPNO-CCSD level.

The difference relative to the experimentally obtained HFCC with DLPNO-CCSD/def2-TZVPP/EC-RISM is relatively small with 2.1 MHz. Nevertheless, adding triple excitation effects into the response density (DLPNO-CCSD(T)) could further improve the results.

4.2.5 HFCC parameters for the nitroxy nitrogen calculated for the protonated HMI

This chapter will consider the influence of pH on the HFCC of HMI. It is well known that nitroxides with a protonatable group near the spin motif show a pH-dependent change of the A^{iso}. For example, the N-H group can be protonated in HMI, and the measured $\mathrm{p}K_{\mathrm{a}}$ is 4.37. The experimental HFCC value shown in the previous chapters was measured at a pH of 10, and at a pH of 2, an HFCC of 41.2 MHz was measured by L. Galazzo[293], resulting in a pH-dependent shift of -3.6 MHz. The same workflow shown above for HMI was repeated under acidic conditions to calculate the theoretical values. In the following chapters of this thesis, the protonated HMI will be abbreviated as HHMI.

In Table 7, a detailed comparison of the HFCCs of HMI and HHMI is presented, starting with the optimized structures on the revPBE0-D3/def2-TZVPP/CPCM level of theory. Here, a A^{iso} of 27.5 MHz can be obtained, leading to a pH-dependent change of -2.5 MHz relative to the unprotonated HMI, where the calculated HFCC was 30.0 MHz. Therefore an underestimation of 1.1 MHz relative to the experimentally obtained pH-dependent change is given. For revPBE0/def2-TZVPP with EC-RISM on the one optimized structure a A^{iso}-parameter of 28.9 MHz was calculated for HHMI. Compared to the HFCC of 32.5 MHz mentioned previously for HMI, a pH-dependent change of 3.6 MHz results perfectly matches the experimentally measured one. Also, in absolute terms, the EC-RISM results show a smaller deviation from the experiment than the CPCM calculations; however, like in the case of HMI, the absolute performance is poor. Better performance is again achieved with the DLPNO-CCSD/def2-TZVVP ansatz, with 32.3 MHz. However, the pH-dependent change with 5.8 MHz is significantly greater than the experimentally measured one.

Table 7 Summarized A^{iso} parameters in MHz for HMI and HHMI. On the one hand, results are shown for the revPBE0-D3/def2-TZVPP/CPCM optimized structure, calculated with revPBE0/def2-TZVPP/CPCM, revPBE0/def2-TZVPP/EC-RISM, and DLPNO-CCSD/def2-TZVPP/EC-RISM. On the other hand, mean A^{iso} parameters are shown that were calculated on the snapshots obtained from the AIMD. Rows with a star * were calculated by V. A. Tran[294], and raw data can be found in SI_4.2.

Method	$\boldsymbol{A}^{\mathbf{iso}}$(HMI)	$\boldsymbol{A}^{\mathbf{iso}}$(HHMI)	$\Delta\boldsymbol{A}^{\mathbf{iso}}$
Based on revPBE0-D3/def2-TZVPP optimized geometry			
revPBE0/def2-TZVPP/CPCM	30.0	27.5	2.5
revPBE0/def2-TZVPP/EC-RISM	32.5	28.9	3.6
DLPNO-CCSD/def2TZVPP/EC-RISM	38.1	32.3	5.8
Based on the full trajectory			
revPBE0/def2-TZVPP/QM/MM*	36.7	33.3	3.4
revPBE0/def2-TZVPP/CPCM/VD	34.5±0.17	32.0±0.17	2.5
revPBE0/def2-TZVPP/EC-RISM/VD	37.0±0.16	33.5±0.17	3.5
revPBE0/def2-TZVPP/CPCM/SSS	36.7±0.17	33.6±0.18	3.1
revPBE0/def2-TZVPP/EC-RISM/SSS	37.5±0.16	34.4±0.17	3.1
DLPNO-CCSD/def2-TZVPP/QM/MM*	40.5	36.1	4.4
DLPNO-CCSD/def2-TZVPP/EC-RISM	42.7±0.26	36.9±0.27	5.8
Experimental	44.8	41.2	3.6

B. Sharma performed another 200 ps long AIMD using the protonated HHMI as a solute, leading again into two subsets, one with 1000 snapshots with a timestep of 0.2 ps and the other subset with 400 structures and a time interval of 0.5 ps.

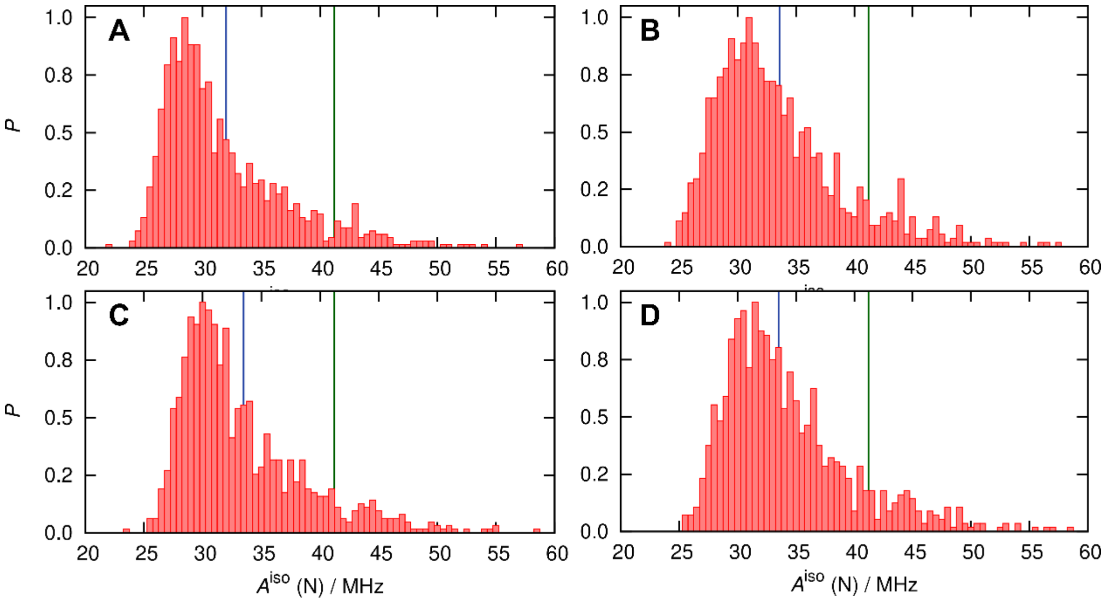

Figure 21 Normalized probability histograms of A^{iso} values of the nitroxy nitrogen of the EPR spin-probe HHMI (protonated HMI) at ambient conditions calculated with revPBE0-D3/def2-TZVPP with decontrated s function on 1000 snapshots containing subset. Panel (A) shows the distribution calculated with CPCM solvation on the vertically desolvated structures (Mean: 34.4 MHz). Panel (B) depicts the distribution calculated with CPCM solvation on the structures, containing the explicit solvent molecules of the first two solvation shells around the oxygen of HMI (Mean: 33.6 MHz). Panel (C) shows the distribution calculated with EC-RISM/PSE-3 on the vertically desolvated structures (Mean: 33.5 MHz). In panel (D), the distribution was calculated with EC-RISM/PSE-3 (Mean: 34.4 MHz) on the structures, explicitly including the two first solvation shells around the HMI oxygen. The blue line in each panel represents the average value of the corresponding distribution computing using the numerical A^{iso} data that underlie the respective histogram. The green vertical line in the panels represents the experimental value, and each distribution is separately normalized. Raw and calculated data can be found in SI_4.2.

With revPBE0/def2-TZVPP/CPCM on the 1000 vertically desolvated structures, a mean HFCC value of 32.0 MHz is observed. Thus, a pH-dependent change of 2.5 MHz is achieved, which is exactly the value obtained for the single optimized structure. By applying revPBE0/def2-TZVPP/EC-RISM, a mean A^{iso} of 33.5 MHz is calculated, leading to a pH-dependent change of 3.5 MHz. Therefore, EC-RISM on the vertically desolvated structures can reproduce the pH-dependent change more accurately than CPCM.

By explicitly adding the two first solvation shells around the N-O motif the A^{iso} parameter increases again. For revPBE0/def2-TZVPP/CPCM, an increase of 1.6 MHz relative to the vertically desolvated structures is calculated. For HHMI, the influence of adding the first two solvation shells is less strong compared to the unprotonated HMI, where an increase of 2.2 MHz was obtained. Due to the smaller increase, the pH-dependent change is 3.1 MHz, reproducing the experiment better than the vertically desolvated subset. By using the revPBE0/def2-TZVPP/EC-RISM level of theory on the solvated subset, an A^{iso} parameter of 34.4 MHz is calculated. The resulting pH-dependent change is 3.1 MHz, matching the revPBE0/def2-TZVPP/CPCM calculated value on the second solvation shell data set. For the revPBE0/def2-TZVPP/QM/MM (by V. A. Tran[294]) calculations, a value of 33.3 MHz was obtained. As for the unprotonated HMI, this value is lower than the EC-RISM calculations. The pH-dependent change is 3.4 MHz. also showing a very good agreement with the experimentally measured value.

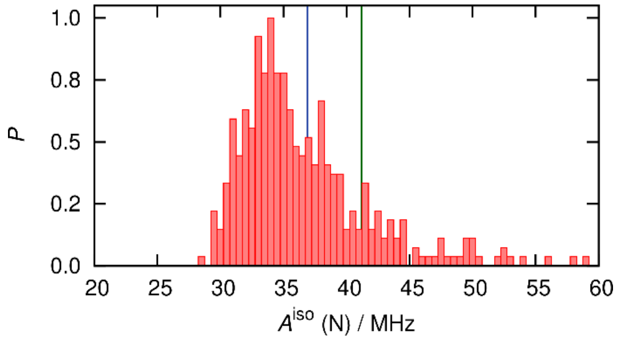

Figure 22 Normalized probability histograms of A^{iso} values of the nitroxy nitrogen of the EPR spin-probe HHMI (protonated HMI) at ambient conditions calculated with DLPNO-CCSD/def2-TZVPP with decontrated s function on 1000 snapshots containing subset. The calculated mean value is 36.9 MHz. Raw and calculated data can be found in SI_4.2.

Utilizing the DLPNO-CCSD/def2-TZVPP/EC-RISM (see Figure 22) on the subset containing 400 vertically desolvated snapshots, a mean HFCC of 36.9 MHz is calculated. Similar to the optimized structures, the calculated pH-dependent change (5.8 MHz) is much greater than the experimentally measured one and the corresponding pH-dependent changes that were obtained with the DFT methods. However, with DLPNO-CCSD/def2-TZVPP/EC-RISM, the absolute value shows the best agreement with the experiment. The DLPNO-CCSD/def2-TZVPP/QM/MM[294] approach leads to a value of 36.1 MHz, resulting in a pH-dependent change of 4.4 MHz. As for the EC-RISM computations, an overestimation is observed for the DLPNO-CCSD/def2-TZVPP-based calculations. A clear reason for this result could not be found.

4.2.6 The complete EPR W-Band spectra of HMI and HHMI

The previous chapters dealt with the isotropic hyperfine coupling constants obtained from X-band experiments at room temperature. However, EPR spectra can also be recorded at higher frequencies, e.g., at 94 GHz, in which case one speaks of a W-band spectrum (see section 3.2). These spectra are normally obtained at very low temperatures (the solution is frozen), whereby the movements of the molecules are restrained. In the W-band spectra at low temperatures, the anisotropic elements of the *g*-tensors and *A*-tensors can be extracted. In this chapter, the W-band spectra of HMI and HHMI were calculated theoretically by applying the same workflow used for calculating the HFCCs, and the spectra will be compared to the experimentally measured spectra at pH 10 and pH 2. The spectra are determined via the spin Hamiltonian from the equation (179), where one additional parameter is the intrinsic linewidth of the peaks. The *g*- and *A*-tensors were extracted from the QC calculations, and the *g*-tensors were calculated with the GIAO method. For another publication, *g*-tensor calculations, where the gauge origin is chosen to be the spin density center[194], were evaluated; however, at the current time, these works were not finished.

First, the W-band spectra for the revPBE0-def2TZVPP/CPCM optimized structure are examined. The *g*-tensors are presented in Table 8, and the corresponding *A*-tensors are shown in Table 9. Within this chapter, DLPNO-CCSD/def2-TZVPP/EC-RISM spectra are presented, whereby it is important to notice that these spectra are a combination of revPBE0/def2-TZVPP calculations for the *g*-tensors and only the *A*-tensors were calculated with DLPNO/def2-TZVPP. For simplicity and clarity, in tables that show *g*-tensor elements, the revPBE0/def2-TZVPP calculations for the *g*-tensors of DLPNO-CCSD spectra are denoted as DLPNO-CCSD.

Table 8 *g*-tensors calculated for the revPBE0-D3/def2-TZVPP/CPCM optimized structures of the unprotonated HMI and the protonated HHMI using revPBE0/def2-TZVPP/Vac, revPBE0/def2-TZVPP/CPCM, revPBE0-D3/def2-TZVPP/EC-RISM, and DLPNO-CCSD/def2-TZVPP/EC-RISM. The experimental values were measured at a frequency of 93.933 MHz and were performed by L. Galazzo[293] and M. Teucher[295]. Raw and calculated data can be found in SI_4.2.

Method	HMI			HHMI			Δ(HMI-HHMI)		
	g_{xx}	g_{yy}	g_{zz}	g_{xx}	g_{yy}	g_{zz}	Δg_{xx}	Δg_{yy}	Δg_{zz}
revPBE0/VAC	2.0086	2.0059	2.0021	2.0092	2.0059	2.021	-0.0004	0.0000	0
revPBE0/CPCM	2.0082	2.0058	2.0021	2.0086	2.0058	2.0021	-0.0005	0.0000	0
revPBE0/EC-RISM	2.0078	2.0057	2.0021	2.0084	2.0058	2.0021	-0.0006	-0.0001	0
DLPNO-CCSD/EC-RISM	*2.0078*	*2.0057*	*2.0021*	*2.0084*	*2.0058*	*2.0021*	*-0.0006*	*-0.0001*	*0*
Experimental	2.0083	2.0060	2.0023	2.0090	2.0061	2.0023	-0.0007	-0.0001	0
				Δ(Experimental - Theory)					
revPBE0/VAC	0.0003	0.0001	0.0002	-0.0002	0.0002	0.0002	-0.0003	-0.0001	0
revPBE0/CPCM	0.0002	0.0002	0.0002	0.0004	0.0003	0.0002	-0.0002	-0.0001	0
revPBE0/EC-RISM	0.0005	0.0003	0.0002	0.0006	0.0003	0.0002	-0.0001	0.0000	0
DLPNO-CCSD/EC-RISM	0.0005	0.0003	0.0002	0.0006	0.0003	0.0002	-0.0001	0.0000	0

The calculated *g*-tensors show that the g_{zz} and g_{yy} components are almost identical, with 2.0058 for the g_{yy} component and 2.0021 for the g_{zz} component. The experimentally measured g_{zz}-value is 2.0023, which fits the g_e value and is expected according to Eq. (182) and the simple LCAO model. According to Eq (180), things get much more interesting when one looks at the g_{xx} component, which mostly depends on the local solvency environment. In the experiment, the g_{xx} components are 2.0083 and 2.0090 for HMI and HHMI, respectively. Based on the LCAO model, the g_{xx} should decrease with increasing polarity and an increasing number of hydrogen bonds. The revPBE0-D3/def2-TZVPP/VAC level of theory leads to the highest g_{xx} components with 2.0086 and 2.0092 for HMI and HHMI, respectively. These results are greater than the experimental values, leading to differences of +0.0003 (HMI) and +0.0002 (HHMI).

Using the revPBE0/def2-TZVPP/CPCM, the g_{xx} decreased to 2.0082 for HMI and 2.0086 for HHMI. This result supports the LCAO model that with higher polarization, the g_{xx} decreases, leading to deviations relative to the experiment of

0.0002 (HMI) and 0.0004 (HHMI). By utilizing the revPBE0-D3/def2-TZVVP/EC-RISM model, a further decrease of g_{xx} is observed. The corresponding values are 2.0078 and 2.0084 for HMI and HHMI, respectively. The absolute discrepancy relative to the experiment increases, with errors of 0.0005 (HMI) and 0.0006 (HHMI). It should be noted again that for DLPNO-CCSD/def2-TZVPP, revPBE0/def2-TZVPP must be used to calculate the g-tensors. However, small deviations from the revPBE0-D3/def2-TZVVP/EC-RISM calculation are possible since a different subset (only 400 structures with 500 fs distance) is considered. While the absolute accuracy with all methods has errors, the relative difference between HMI and HHMI is very well reproduced.

Table 9 A-tensors calculated for the revPBE0-D3/def2-TZVPP/CPCM optimized structures of the unprotonated HMI and the protonated HHMI using revPBE0/def2-TZVPP/Vac, revPBE0/def2-TZVPP/CPCM, revPBE0-D3/def2-TZVPP/EC-RISM, and DLPNO-CCSD/def2-TZVPP/EC-RISM. All values are in MHz. The experimental values were measured at a frequency of 93.933 MHz and were performed by L. Galazzo[293] and M. Teucher[295]. Raw and calculated data can be found in SI_4.2.

Method	HMI			HHMI			Δ(HMI-HHMI)		
	A_{xx}	A_{yy}	A_{zz}	A_{xx}	A_{yy}	A_{zz}	ΔA_{xx}	ΔA_{yy}	ΔA_{zz}
revPBE0/VAC	0.2	0.64	78.65	0.02	-0.82	70.57	0.18	1,46	8.09
revPBE0/CPCM	1.62	1.52	86.89	1.2	0.74	80.67	0.42	0.78	6.22
revPBE0/EC-RISM	2.55	2.65	92.00	1.8	1.48	83.56	0.75	1.17	8.44
DLPNO-CCSD/EC-RISM	7.8	7.8	99.1	5.3	5.31	86.33	2.5	2.49	12.77
Experimental	14.0	14.0	100.3	13.0	13.0	92.2	1	1	8.1
Δ(Experimental - Theory)									
revPBE0/Vac	13.8	13.34	21.85	7.68	13.82	21.93	0.82	-0.46	0.09
revPBE0/CPCM	12.38	12.48	13.61	11.8	12.26	11.83	0.58	0.22	1.78
revPBE0/EC-RISM	11.45	11.35	8.5	11.2	11.52	8.94	0.25	-0.17	-0.44
DLPNO-CCSD/EC-RISM	6.2	6.2	1.4	7.7	7.69	6.17	-1.5	-1.49	-4.77

Regarding the A tensors, the focus within this chapter shall be set on the A_{zz} component since the other two components cannot be sufficiently resolved in the experiment. The experimentally obtained A_{zz} couplings are 100.5 MHz and 92.5 MHz for HMI and HHMI, respectively; therefore, a pH-dependent change of 8 MHz results. The lowest calculated A_{zz} coupling is obtained using the revPBE0-D3/def2-TZVPP/Vac approach with 78.7 MHz (HMI) and 70.57 MHz (HHMI). However, vacuum calculations can best reflect the pH-dependent change of the A_{zz} component with 8.1 MHz.

Significant improvement is achieved by applying the revPBE0-D3/def2-TZVPP/CPCM method, leading for HMI to an A_{zz} coupling of 86.9 MHz, and for the protonated HHMI, 80.7 MHZ is obtained. The differences relative to the experimental data are still relatively large, with 13.6 MHz and 11.6 MHz for HMI and HHMI, respectively. In contrast to all other methods, the CPCM variant underestimates the pH-dependent change with a value of 6.44 MHz.

Better absolute values are obtained by applying the revPBE0-D3/def2-

TZVVP/EC-RISM approach. For HMI, the difference to the experimental reference A_{zz} coupling value can be decreased to 8.5 MHz, with an absolute value of 92 MHz. An A_{zz} coupling of 83.6 MHz is calculated for the protonated state, resulting in a deviation from the experiment of 8.9 MHz. The pH-dependent change shows with 8.44 MHz a small deviation from the experimentally obtained pH-dependent change. These results fit the previous results for the A^{iso} parameter, where EC-RISM causes a stronger polarity and leads to higher values.

The best results are obtained with DLPNO-CCD/def2-TZVPP/EC-RISM ansatz, where an A_{zz} component of 99.1 MHz is calculated for HMI. Thus, the deviation from the experiment is only 1.4 MHz. Slightly worse results are obtained for HMHI, with a deviation relative to the experiment of 6.2 MHz. This result is in good agreement with the previous chapter's HFCC calculation. Thus, this level of theory shows the biggest discrepancy for the pH-dependent change relative to the experimental value, with a calculated difference of 12.8 MHz.

The results for the g-tensors of the complete trajectory are enlisted in Table 10. As previously shown for the single optimized structures, the g_{yy} and g_{zz} values are identical for all calculations and show the same deviations from the experiment. The best agreement with the experimental reference data for the g_{xx} is again achieved by applying the revPBE0-D3/def2-TZVPP/CPCM approach on the vertically desolvated dataset. For HMI, a value of 2.0084 is obtained, leading to a deviation relative to the experimental value of only -0.0001. 2.0088 is the g_{xx} component for the protonated state, resulting in a slightly higher error than the experiment of 0.0002. The corresponding pH-dependent change is -0.0004 and shows a deviation of -0.0003 to the experimentally measured pH-dependent change. Compared to the single optimized structures, changes of 0.0002 are observed.

Table 10 Averaged g-tensors calculated for the unprotonated HMI and protonated HHMI. revPBE0 is representative of revPBE0/def2-TZVPP with decontracted s-functions. These calculations were performed on the 1000 snapshots containing subset. VD stands for vertical desolvated, and SSS for second solvation shell included. DLPNO-CCSD EC-RISM VD was calculated with revPBE0/def2-TZVPP/EC-RISM (since no g-tensor calculations were possible with DLPNOP-CCSD) on the subset containing the 400 vertically desolvated structures. The experimental values were measured at a frequency of 93.933 MHz and were performed by L. Galazzo[293] and M. Teucher[295]. Raw and calculated data can be found in SI_4.2.

Method	HMI			HHMI			Δ (HMI-HHMI)		
	g_{xx}	g_{yy}	g_{zz}	g_{xx}	g_{yy}	g_{zz}	Δg_{xx}	Δg_{yy}	Δg_{zz}
revPBE0 CPCM VD	2.0084	2.0050	2.0021	2.0088	2.0058	2.0021	-0.0004	0.0000	0.0000
revPBE0 CPCM SSS	2.0078	2.0057	2.0021	2.0083^5	2.0058	2.0021	-0.0005	-0.0001	0.0000
revPBE0 EC-RISM VD	2.0079	2.0058	2.0021	2.0085	2.0058	2.0021	-0.0006	0.0000	0.0000
revPBE0 EC-RISM SSS	2.0077	2.0057	2.0021	2.0083	2.0058	2.0021	-0.0006	-0.0001	0.0000
*DLPNO-CCSD EC-RISM**	*2.0079*	*2.0057*	*2.0021*	*2.0085*	*2.0058*	*2.0021*	-0.0006	-0.0001	0.0000
Experimental	2.0083	2.0060	2.0023	2.0090	2.0061	2.0023	-0.0007	-0.0001	0.0000
Δ (Experimental - Theory)									
revPBE0 CPCM VD	-0.0001	0.0002	0.0002	0.0002	0.0003	0.0002	-0.0003	-0.0001	0.0000
revPBE0 CPCM SSS	0.0005	0.0003	0.0002	0.0007	0.0003	0.0002	-0.0002	0.0000	0.0000
revPBE0 EC-RISM VD	0.0004	0.0002	0.0002	0.0005	0.0003	0.0002	-0.0001	-0.0001	0.0000
revPBE0 EC-RISM SSS	0.0006	0.0003	0.0002	0.0007	0.0003	0.0002	-0.0001	0.0000	0.0000
DLPNO-CCSD EC-RISM	0.0004	0.0003	0.0002	0.0005	0.0003	0.0002	-0.0001	0.0000	0.0000

As soon as the explicit waters are added to the revPBE0-D3/def-TZVPP/CPCM calculations, the g_{xx} drops significantly. In the case of HMI, the effect of adding the explicit waters is a decrease of -0.0006. The resulting error relative to the experiment increases to 0.0005 and, thus, is significantly worse than the calculations on the single optimized structure and the vertically desolvated dataset. A similar picture is observed for the low pH region, with a g_{xx} of 2.0083 and a decrease of -0.0005 relative to the result of the vertically desolvated structures. However, a slight improvement is seen in the pH-dependent change with -0.0005, leading to a very small deviation relative to the experiment.

Considering the revPBE0-D3/def2-TZVPP/EC-RISM calculations on the vertically desolvated subset, the obtained g_{xx} are 2.0079 and 2.0085 for HMI and HHMI, respectively. The values are marginally higher than for CPCM on the solvated structures. Compared to the single optimized structure, a change of 0.0001 is observed. The calculated pH-dependent change is -0.0006, and the deviation to the experimental delta is only 0.0001.

For the DLPNO-CCSD/def2-TZVPP/EC-RISM spectra (by applying revPBE0-D3/def2-TZVPP/EC-RISM for g-tensors) calculations, the smaller subset with 400 vertically desolvated structures was used, and no differences relative to the larger subset can be observed. By adding the two first solvation shells explicitly into the revPBE0-D3/def2-TZVPP/EC-RISM calculations, the obtained g_{xx} components are 2.0077 (HMI) and 2.0083 (HHMI). Here one can see that by adding the explicit waters, the g_x becomes smaller, which is in good agreement with the LCAO model. However, the absolute accuracy relative to the experiment decreases by adding the explicit waters. This result could further indicate that the underlying QC DFT calculations have an intrinsic error in calculating the *g*-tensor. The calculated pH-dependent change with -0.0006 is still very good and nearly in excellent agreement with the experiment.

Overall, the results for the g_{xx} show, in absolute terms, a deviation from the experiment, especially a systematic worsening if theoretically more reliable systems are investigated. Nonetheless, the difference between the unprotonated and protonated state can be accurately represented.

In Table 11, the *A*-tensors are enlisted and were calculated for the whole trajectory. The focus will again be solely on the A_{zz} component, and similar results compared to HFCC calculations are expected. First, for all different levels of theory combinations, an increase of the A_{zz} compared to the value of the single optimized structure is observed. These results are again due to the fluctuations in the nitrogen out-of-plane movement.

The revPBE0-D3/def2-TZVPP/CPCM calculations on the 1000 vertically desolvated snapshots lead to a mean A_{zz} of 90.6 MHz for HMI and 84.3 MHz for HHMI. Compared to the single optimized structure, an increase of 3.6 MHz is observed, and the deviations to the experimental value are reduced to 9.9 MHz and 8.2 MHz for HMI and HHMI, respectively. The pH-dependent change with 6.3 MHz is nearly identical to the optimized structure.

For HMI and revPBE0-D3/def2TZVPP/CPCM level of theory, the A_{zz} component increases by 5.2 MHz to 95.8 MHz by adding the second solvation shell into the calculations. The corresponding increase for HHMI is 3.9 MHz, and the absolute value is 88.2 MHz. These results are significantly better than the results for the vertically desolvated structures, with a halving of the relative error. The pH-dependent change shows with 7.6 MHz only a deviation of 0.4 MHz compared to the observation in the experiment.

Table 11 Averaged *A*-tensors (all values in MHz) calculated for the unprotonated HMI and protonated HHMI. revPBE0 is representative for revPBE0/def2-TZVPP with decontracted s-functions. The line revPBE0 CPCM VD describes the *A*-tensors calculated with revPBE0/def2-TZVPP/CPCM on the 1000 snapshots subset using the vertically desolvated structures. The line revPBE0 CPCM SSS was calculated with revPBE0/def2-TZVPP/CPCM on the 1000 snapshots subset using structures including the two first solvations shells around the N-O motif. revPBE0 EC-RISM VD refers to calculations with revPBE0/def2-TZVPP/EC-RISM on the 1000 snapshot subset containing the vertically desolvated structures. revPBE0 EC-RISM SSS was calculated with revPBE0/def2-TZVPP/EC-RISM utilizing the 1000 snapshots dataset with the first two solvation shells around the N-O group included. DLPNO-CCSD/def2-TZVPP/EC-RISM on the subset containing the 400 vertically desolvated structures. All values are in MHz. Raw and calculated data can be found in SI_4.2.

Method	HMI			HHMI			Δ(HMI-HHMI)		
	A_{xx}	A_{yy}	A_{zz}	A_{xx}	A_{yy}	A_{zz}	ΔA_{xx}	ΔA_{yy}	ΔA_{zz}
revPBE0 CPCM VD	6.4	6.4	90.6	6.0	5.6	84.3	0.4	0.8	6.3
	±0.2	±0.2	±0.2	±0.2	±0.2	±0.1			
revPBE0 CPCM SSS	7.0	7.2	95.8	6.5	6.3	88.2	0.5	0.9	7.6
	±0.2	±0.2	±0.2	±0.2	±0.2	±0.2			
revPBE0 EC-RISM VD	7.4	7.7	96.1	6.6	6.4	87.4	0.8	1.3	8.7
	±0.2	±0.2	±0.2	±0.2	±0.2	±0.2			
revPBE0 EC-RISM SSS	7.5	7.8	97.3	6.9	6.8	89.5	0.6	1	7.8
	±0.2	±0.2	±0.2	±0.2	±0.2	±0.2			
DLPNO-CCSD EC-RISM VD	12.3	12.8	102.8	10.2	10.4	90.3	2.1	2.4	12.5
	±0.3	±0.3	±0.3	±0.3	±0.3	±0.2			
Experimental	14.0	14.0	100.3	13.0	13.0	92.2	1	1	8.1
Δ(Experimental - Theory)									
revPBE0 CPCM VD	7.6	7.6	9.9	7	7.4	8.2	0.6	0.2	1.7
revPBE0 CPCM SSS	7	6.8	4.7	6.5	6.7	4.3	0.5	0.1	0.4
revPBE0 EC-RISM VD	6.6	6.3	4.4	6.4	6.6	5.1	0.2	-0.3	-0.7
revPBE0 EC-RISM SSS	6.5	6.2	3.2	6.1	6.2	3	0.4	0	0.2
DLPNO-CCSD EC-RISM VD	1.7	1.2	-2.3	2.8	2.6	2.2	-1.1	-1.4	-4.5

Applying the EC-RISM solvation model with the revPBE0-D3/def2-TZVPP level of theory for HMI on the vertically desolvated subset, an A_{zz} coupling constant of 96.1 MHz is obtained. The increase compared to the single optimized structure is 4.1 MHz, and the difference relative to the experiment is reduced to 4.4 MHz; therefore, a small improvement is obtained compared to the CPCM calculations on the solvated structures. The calculated value for the low pH region is 87.4 MHz, leading to a deviation of 5.1 MHz relative to the experiment. Also, the coupling constant is slightly smaller than the one obtained for the previously mentioned CPCM calculations, and a slightly higher pH-dependent change than the experiment is obtained with 8.7 MHz. All in all, EC-RISM can deliver the same performance on the vertically desolvated structures as CPCM calculations, in which the first two solvation shells were included.

A further improvement for the DFT-based calculations is achieved by adding the two first solvation shells to the revPBE0-D3/def2-TZVPP/EC-RISM calculation. The calculated A_{zz} coupling constants are 97.3 MHZ and 89.5 MHz for HMI and HHMI, respectively. These results lead to a pH-dependent change of 7.8 MHz, which best agrees with the experimental value (deviation of 0.2 MHz to the experiment).

Higher couplings are observed for DLPNO-CCSD/def2-TZVPP/EC-RISM calculations on the 400 vertically desolvated structures. The A_{zz} component is 102.8 MHz in the high pH region and, therefore, even greater than the experimental reference value, resulting in a deviation of -2. MHz. As already seen for the HFCC, the calculated pH-dependent change for DLPNO-CCSD in combination with EC-RISM strongly overestimates the experimental trend by 4 MHz. The calculated A_{zz} component of HHMI is 90.3 MHz.

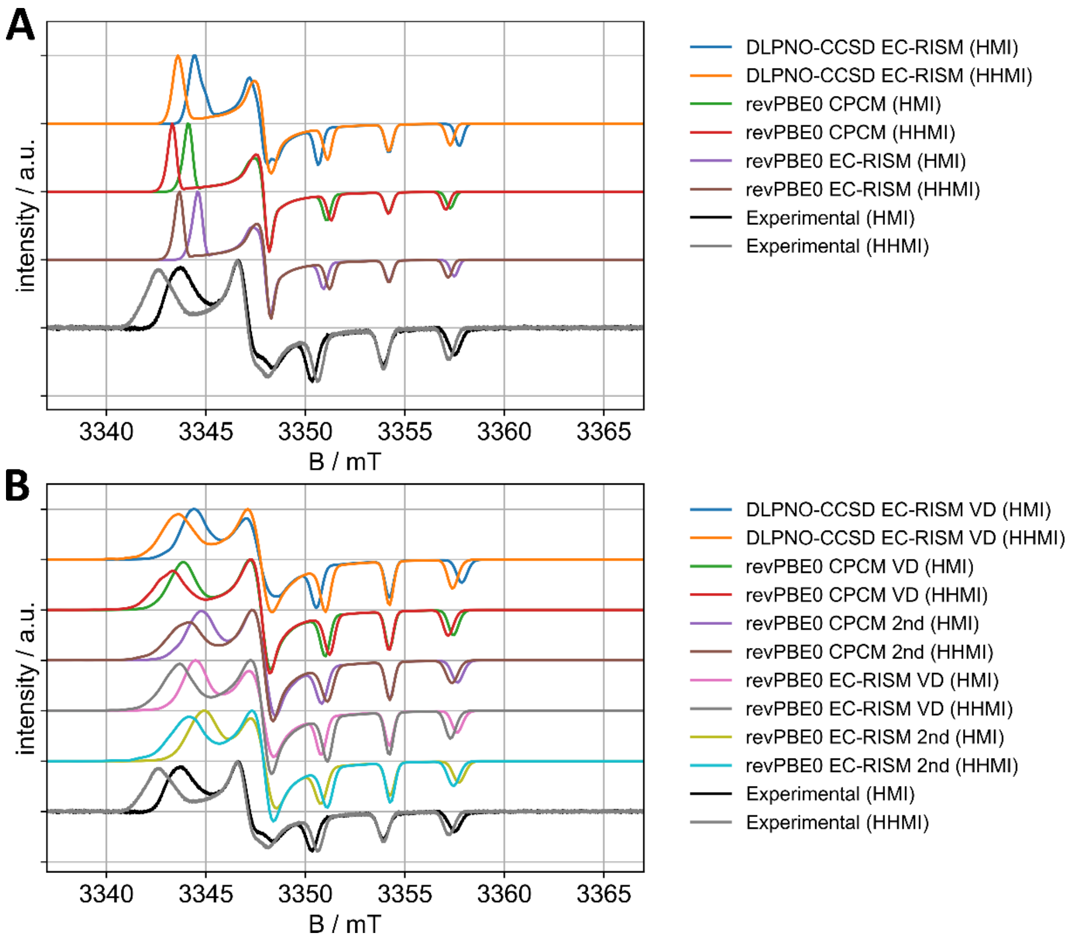

Figure 23 Panel (A) shows EPR W-band spectra (linewidth=0.45 mT) calculated for the revPBE0-D3/def2-TZVPP/CPCM optimized structures of the unprotonated HMI and the protonated HHMI using revPBE0/def2-TZVPP/CPCM, revPBE0-D3/def2-TZVPP/EC-RISM, and DLPNO-CCSD/def2-TZVPP/EC-RISM (the corresponding *g*-tensors were calculated using revPBE0-D3/def2-TZVPP/EC-RISM). Panel (B) shows the calculation of the DLPNO-CCSD/def2-TZVPP/EC-RISM spectra on the 400 VD subset, revPBE0/def2-TZVPP/CPCM on the VD (1000 snapshots) subset, revPBE0/def2-TZVPP/CPCM on the second solvation shell included (1000 snapshots) subset, revPBE0/def2-TZPP/EC-RISM on the 1000 vertically desolvated structures and revPEB0/def2-TZVPP/EC-RISM on the 1000 snapshots, including the two first solvation shells around the N-O motif. At first, the individual spectrum was calculated for all snapshots, and subsequently, the spectra were averaged. Experimental and calculated W-Band spectra of HMI (pH2) and HHMI (pH10) are shown. The experimental values were measured at a frequency of 93.933 MHz and were performed by L. Galazzo[293] and M. Teucher[295]. Raw and calculated data can be found in SI_4.2.

The graphical summary of the results for the *A*-tensors and *g*-tensors is shown in the simulated W-band spectra in Figure 23, whereby panel (A) shows the spectra for the single optimized structures, and in panel (B), the averaged spectra over the whole trajectory are presented. The different trends for the g_{xx} and A_{zz} components can be seen. With more polarizing methods, the g_{xx} value decreases but at the same time, the A_{zz} coupling increases. In this manner, the g_{xx} deviates stronger from the experimental value, whereas the A_{zz} shows a better agreement by applying a higher level of theory. This result indicates a systematic error in the QC calculation of the *g* tensors. For further validations, spectra calculations for the entire trajectory were performed. First, one sees that the optimized structures' spectra significantly underestimate the spectra's width. If, for every snapshot, the spectrum is calculated and subsequently averaged, the theoretical calculations better represent the intrinsic width of the experimental spectrum. Furthermore, one can see that the CPCM calculations on the vertically desolvated structures provide the best agreement in the g_{xx} compared to the experiment. All other methods, which in theory should give better results, show larger errors in calculating the g-tensors. In contrast to this result, the opposite picture is observed with the couplings, where the A_{zz} principal component increases with a better theory level, reducing the error relative to the experiment.

4.2.7 Pressure-dependent EPR-Spectra of HMI and HHMI

As mentioned in chapter 4.1, higher pressure leads to a stronger solvent polarization. Consequently, a constant increase with rising pressure is expected for the HFCC. This chapter will analyze the influence of pressure on HFCC and the full EPR spectrum by applying pressure-dependent EC-RISM calculations. As the level of theory, the revPBE0/def2-TZVPP with decontracted s-functions was utilized in combination with pressure-dependent χ_{HNC} functions. Since no experimental results are available at the time of writing, the presented results are exploratory and should give a possible prediction.

The absolute results for the pressure-dependent A^{iso} parameters, g_{xx}, and A_{zz} are summarized in Table 12.

Table 12 Pressure-dependent nitrogen HFCC (in MHz), g_{xx} components, and nitrogen A_{zz} (in MHz) of HMI and HHMI at ten different pressures calculated with revPBE0/def2-TZVPP/EC-RISM.

Pressure	HMI			HHMI			*Δ*(HMI-HHMI)		
	HFCC	g_{xx}	A_{zz}	HFCC	g_{xx}	A_{zz}	*Δ* HFCC	Δg_{xx}	ΔA_{zz}
1 bar	32.40	2.00776	92.00	28.94	2.00838	83.56	3.46	-0.00062	8.44
100 bar	32.41	2.00776	92.01	28.95	2.00838	83.58	3.46	-0.00062	8.43
500 bar	32.43	2.00775	92.08	29.00	2.00837	83.68	3.43	-0.00062	8.4
1 kbar	32.47	2.00775	92.15	29.05	2.00836	83.80	3.42	-0.00061	8.35
2 kbar	32.52	2.00774	92.28	29.14	2.00835	84.01	3.38	-0.00061	8.27
3 kbar	32.57	2.00773	92.39	29.21	2.00833	84.19	3.36	-0.0006	8.2
4 kbar	32.61	2.00772	92.49	29.28	2.00832	84.36	3.33	-0.0006	8.13
5 kbar	32.65	2.00772	92.58	29.35	2.00831	84.50	3.3	-0.00059	8.08
7.5 kbar	32.74	2.00770	92.78	29.48	2.00829	84.82	3.26	-0.00059	7.96
10 kbar	32.81	2.00769	92.94	29.59	2.00827	85.08	3.22	-0.00058	7.86

As expected, a pressure-dependent increase is observed for the HFCC, with a slightly steeper increase for the protonated HHMI than for the unprotonated version. In panel (A) of Figure 24, the relative changes for both species are graphically depicted. For HMI the A^{iso} parameter increases from 32.40 MHz at ambient pressure up to 32.81 MHz at 10 kbar. The protonated HHMI increases by 0.66 MHz from ambient pressure up to 10 kbar. By looking at the pressure-dependent curve, it can be seen that a quadratic fit of the form:

$$A^{\mathrm{iso}}(p) = a \cdot (p - p_0)^2 + b \cdot (p - p_0) + c, \tag{188}$$

can be applied. The corresponding parameters are summarized in the appendix in Table *24*. This particular form was already applied for the pressure-dependent trend of NMR chemical shifts. [26,27]

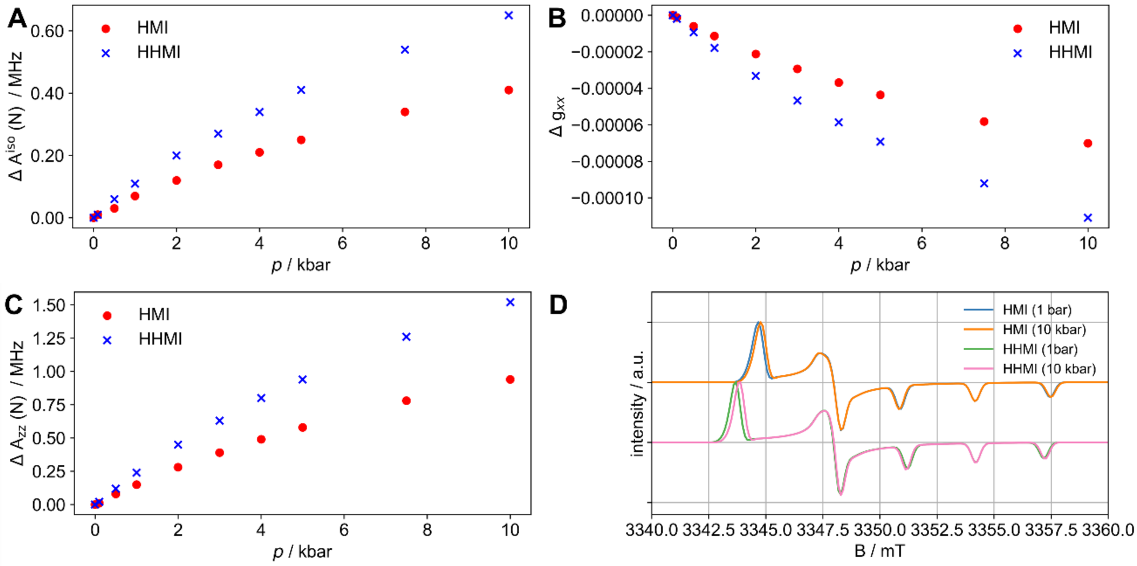

Figure 24 Relative changes of different EPR-Parameters of HMI and HHMI due to pressurization. In panel (A), the change of the HFCCs is shown; in panel (B), the g_{xx} principal components are presented; and in panel (C), the A_{zz} components are depicted. The blue and red points represent HMI and HHMI, respectively. Additionally, in panel (D), the complete EPR spectra are shown for HMI and HHMI at 1 bar and 10 kbar, respectively. All calculations were performed with revPBE0/def2-TZVPP/EC-RISM.

The g_{xx} principal component shows a steady decrease upon pressurization (see panel (B) in Figure 24), and this effect was already seen for the chemical shielding constants in NMR calculations. Since the distance between solute and solvent decreases upon pressurization, the delocalization of the c_{ny}^2 orbital into the H_2O orbitals increases, and therefore the absolute value of the g_{xx} decreases (see eq (180)). Again, the protonated HHMI shows a stronger pressure dependence than HMI. For the latter, a change of -0.00007 is calculated going from the ambient condition into the very high-pressure realm. The comparable calculated decrease for HHMI is -0.00009. A reasonable equation format to describe the pressure-dependency of the g_{xx} would be a quadratic fit (in Table 24, the parameters are enlisted), too.

The same things that have already been noted for the HFCC apply to the A_{zz}. The pressure effect is stronger on HHM (see panel (C) in Figure 24), and a steady increase is observed. For HMI and HHMI, the changes from 1 bar to 10 kbar are 0.94 MHz and 1.52 MHz, respectively. A quadratic fit also seems to be a reasonable choice to describe the trajectory. Panel (D) in Figure 24 shows the pressure-dependent W-band spectra at 1 bar and 10 kbar to complete the overall picture.

Overall the calculated pressure-dependent changes are relatively small. For example, in the experiment, a change of smaller than 1 MHz can barely be detected in an EPR experiment.[132] Therefore, it may be difficult to resolve the pressure differences sufficiently. However, if one wants to study, e.g., the conformational changes of proteins under high hydrostatic pressures[296], the changes can be attributed to the conformational changes and not to the solvent-mediated pressure effect on the EPR parameters.

4.2.8 Summary and outlook

The goal of this chapter was to make the first attempts to calculate EPR parameters using EC-RISM at ambient conditions to benchmark the methodology, and in addition, some perspective exploratory pressure-dependent calculations were performed. The molecule HMI was chosen as a model system within the RESOLV cluster of excellence, characterized by its relative rigidity and protonatable N-H group. In the first part, the performances of EC-RISM and CPCM in combination with different DFT levels of theory were evaluated, showing that EC-RISM can significantly better reproduce the A^{iso} parameter, but the overall performance was rather sobering. EC-RISM retains structural heterogeneity in the proximity of the solute, whereas CPCM omits this important environmental feature.

To perform a more detailed analysis and get more insights into the underlying characteristics of the solute-solvent interactions, B. Sharma provided an AIMD trajectory using the revPBE0-D3 functional in combination with a def2-TZVPP basis set. This trajectory was divided into two subsets: the so-called vertically desolvated subset, where only the solute structures were retained, and all surrounding waters were removed. All waters except the two first solvation shells around the N-O motif were removed in the second subset. CPCM and EC-RISM calculations were performed for the two subsets, and significant improvements were observed compared to the single optimized structure. These improvements are due to the displacements in the C-N-O-C improper angle, which are only reflected in the complete trajectory. The second important finding is that EC-RISM performs better on both subsets than CPCM. Especially in the case of the vertically desolvated subset, EC-RISM significantly outperforms the CPCM calculations. EC-RISM on the vertically desolvated subset can even reproduce better HFCC than CPCM with the two first solvation shells. Further improvements were obtained by the first implementation of the DLPNO-CCSD within EC-RISM, leading to A^{iso} values that show the best agreement from all studied methods compared to the experiment.

The HFCC of HMI in the protonated state was investigated next. A pH-dependent change of 3.6 MHz was measured in the experiment when going from the low pH-region into the high pH region. The best absolute results are again obtained using the DLPNO-CCSD/def2-TZVVP/EC-RISM method. However, the calculated pH-dependent change is too big compared to the experiment with this approach. In contrast, utilizing the revPBE0/def2-TZVPP combination with CPCM or EC-[297]RSIM solvation leads to a very good agreement in the pH-dependent change compared to the experiment. Furthermore, EC-RISM shows a significantly better performance regarding the absolute values than CPCM.

Usually, the isotropic hyperfine coupling constant is measured in an X-band experiment at ambient conditions. By freezing the sample, it is possible to make a W-band experiment, where detailed information about the g- and A-tensors can be extracted. In this thesis, the focus was set on the g_{xx} principal component of the g-tensor and the A_{zz} of the A-tensor. In principle, for the latter one, the same conclusions hold that were made for the HFCC. The best results relative to the experiment are obtained with DLPNO-CCSD/def2-TZVPP/EC-RISM, and for the DFT calculations, EC-RISM performs better than CPCM with comparable settings. For the g_{xx}, a different picture is observed, where the revPBE0-D3/def2-TZVPP/CPCM calculations show the best agreement compared to the experiment. Furthermore, the performance worsens for theoretically better models, indicating a systematic error at the QC level. In addition, it is to be noted that for the g-tensor calculations, no DLPNO-CCSD calculations could be applied.

Overall, however, the results for the EC-RISM calculations are very satisfactory and can be applied to other models in the future. For example, a spin label attached to a protein or a small peptide fragment could be investigated under ambient or extreme conditions, and an MD simulation should generate the underlying conformational ensemble. The first work was already done here, L. Jauer[297] used the Stendardo force field[118] to generate a force field-based ensemble to reproduce the HFCC obtained from the AIMD ensemble. The first results were very promising; however, further QC-based postprocessing needs to be done.

4.3 Pressure dependence of the peptide backbone

The aim of this chapter is, on the one hand, to evaluate how strongly pressure affects the conformational equilibria of the three typical protein backbone model systems (see chapter1.2) and, on the other hand, whether the magnitude of the effects makes it necessary to perform a pressure-dependent reparameterization of currently used protein force fields. In the first section, the pressure-dependent NMR shifts of NMA are presented. Based on these results, the NMR chemical shifts of the main conformers of Ac-Gly-NHMe and Ac-Ala-NHMe calculated for B3LYP/6-311+G(d,p)/PCM geometries are discussed. Subsequently, the effect of EC-RISM-based pressure-dependent geometry optimizations on the dipeptides compared to the PCM geometries is discussed. In the following section, the pressure-dependent change of the *cis-trans* equilibrium of NMA is discussed experimentally and by various theoretical methods. Based on these results, the second section evaluates the pressure dependence of the minima of the more complex systems, Ac-Gly-NHMe and Ac-Ala-NHMe. A detailed discussion is performed, and the results are compared with various reference data from the literature. The influence of pressure on the complete Ramachandran plane of Ac-Gly-NHMe and Ac-Ala-NHMe is also analyzed.

4.3.1 **Calculations of NMR chemical shifts of *cis*/*trans*-NMA**

In the publication of Frach et al.[26], the pressure dependence of the chemical shifts of ^{1}H, ^{13}C, and ^{15}N of the peptide backbone model NMA was investigated, and an extensive discussion about reasonable referencing was conducted. The chosen reference standard DSS showed only a slight change in the shielding constants upon pressurization. It was demonstrated that EC-RISM could predict the correct experimental pressure trend for all peptide backbone atoms. However, these results were generated with a previous based on point-charges EC-RISM version (see chapter 4.1), and in the past years, many improvements within the EC-RISM method were developed and published[25], including optimized electrostatic and empirical free energy models. These new improvements were applied to the fully flexible water standard DSS and the ^{15}N standard NH_3.[27] To benchmark the performance of the EC-RISM improvements and reliability of NMR standards, NMR parameters for the small osmolyte TMAO were calculated. After these auspicious results, the question arises of how the methodology affects the NMR parameters of *cis*/*trans*-NMA. The chemical shifts of NMA are an excellent example to evaluate the influence of pressure on two different conformers since both conformers are distinguishable in the experiment. The corresponding calculations were already performed by L. Eberlein and published in his dissertation.[28] At this point, the results will be reused and should represent the solvent-induced baseline for chemical shifts and later adapted to the chemical shifts of Ac-Gly-NHMe and Ac-Ala-NHMe.

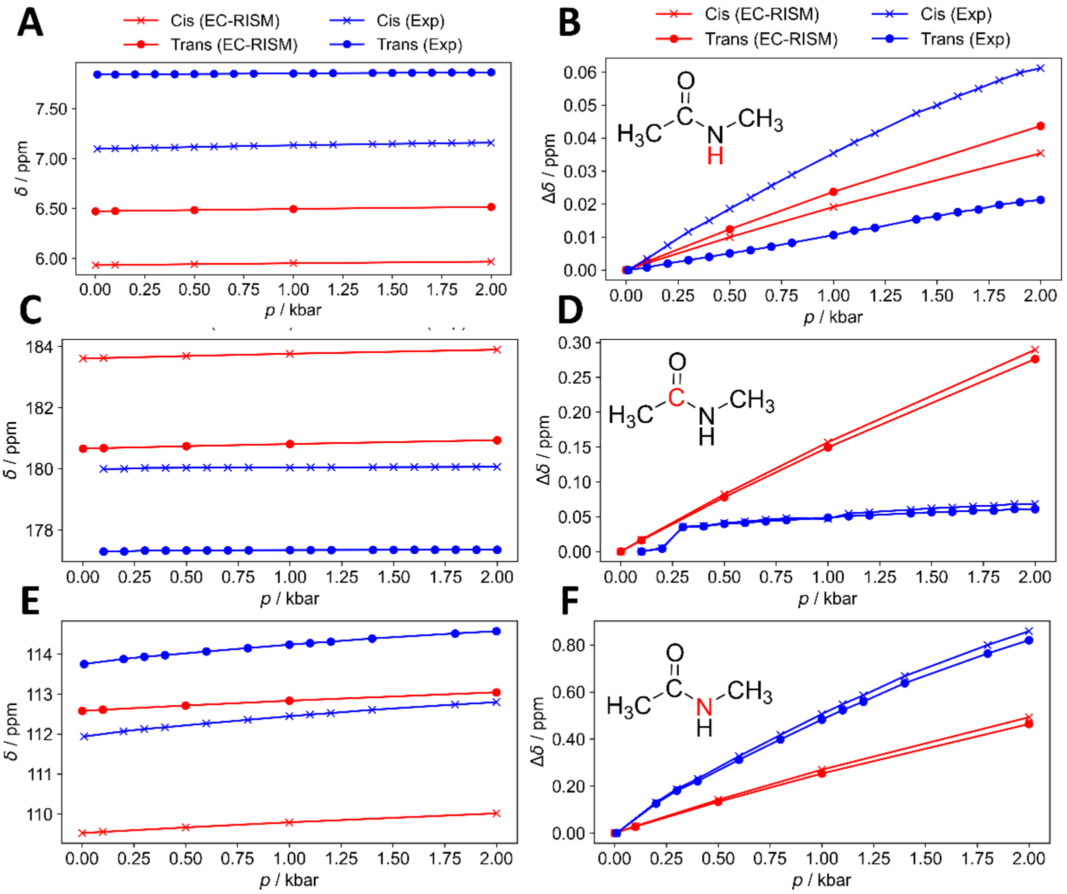

Figure 25 Pressure-dependent absolute shifts (left side) and relative shifts to the 1 bar value of *trans*- and *cis*-NMA peptide bond atoms ^{1}H(A & B), ^{13}C (C & D), and ^{15}N (E & F). The red stars represent the *cis*-conformer calculated, and the upward green triangles calculations with MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p). Here direct referencing was applied, and the corresponding reference shielding constants are collected in Table 43. The raw chemical shielding constants can be found in SI_4.3. Similar plots can be found in ref[28]

Figure 25 shows the chemical shifts, relative to their value at ambient pressure conditions, of the ^{1}H, ^{13}C, and ^{15}N nuclei experimentally and theoretically for the two NMA conformers. Direct referencing is applied throughout this work, which means that the calculated pressure-dependent chemical shielding constants (see Table 43) for the reference nuclei are used, and it showed the best performance in calculating pressure-dependent shift changes of TMAO.[27,28] In the past, performing a quadratic fit to the experimental and calculated data was useful for describing the pressure change, and the corresponding parameters can be taken from Table 13. Here the absolute chemical shifts are listed, too. By looking at the pressure-dependent change of the chemical shifts, one can see that for both *trans*- and *cis*-NMA, each nucleus shows an increase of the chemical shifts; this observation agrees with previously made experiences.[27] The best result is obtained for the ^{1}H nucleus and the worst for the ^{13}C nucleus of the C-O group; these results agree with those observed for TMAO. Concerning the ^{1}H nucleus, the pressure-dependent change in the chemical shift is stronger for the *cis* conformer than for the *trans* conformer within the experiment. EC-RISM cannot accurately represent this trend; the order is reversed.

On the one hand, the pressure dependence of *cis*-NMA is overestimated by EC-RISM, and on the other hand, the pressure dependence of *trans*-NMA is underestimated by EC-RISM. The overall changes of the ^{1}H chemical shifts are rather small, with changes between 0.03 ppm and 0.06 ppm up to a pressure of 2 kbar. One possible reason for underestimating the *trans*-NMA ^{1}H shifts could be a dimerization of NMA. In this manner, the direct interaction of the ^{1}H nucleus with the surrounding water would be blocked, and therefore the main interaction causing the pressure-dependent changes are prevented.

Table 13 Linear (B_1, in ppm kbar^{-1}) and quadratic (B_2, in ppm kbar^{-2}) coefficients from fitting experimental and calculated *cis*- and *trans*-NMA chemical shifts to the following form: $\Delta\delta(p)= B_1p+B_2p^2$. The 1 bar chemical shifts (δ_0 in ppm) are shown. Experimental data were measured by M-Beck Erlach.[298] The EC-RISM calculations were performed with MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM. An analogous representation can be found in "Eberlein, L. J. A combined computational and NMR-spectroscopic approach for tautomer elucidation under extreme conditions towards investigating the robustness of genetic codes, 2021." Raw shielding constants can be found in SI_4.3.

Nucleus	δ_0(1 bar).	B_1	B_2
^{1}H			
cis-NMA (EC-RISM)	5.93	0.0177	-0.00062
cis-NMA (Exp)	7.09	0.0410	-0.00490
trans-NMA (EC-RISM)	6.47	0.0257	-0.00161
trans-NMA (Exp)	7.84	0.0109	0.00054
^{13}C			
cis-NMA (EC-RISM)	183.61	0.1662	-0.01051
cis-NMA (Exp)	179.99	0.0708	-0.02007
trans-NMA (EC-RISM)	180.66	0.1574	-0.00668
trans-NMA (Exp)	177.29	0.0696	-0.02152
^{15}N			
cis-NMA (EC-RISM)	109.52	0.2871	-0.02003
cis-NMA (Exp)	111.95	0.5745	-0.07545
trans-NMA (EC-RISM)	112.59	0.2517	-0.01059
trans-NMA (Exp)	113.76	0.5467	-0.07122

The picture for the ^{13}C nucleus looks much more sobering; the theoretically calculated changes in the chemical shifts at 2 kbar are about six times larger than those measured experimentally. In contrast to the ^{1}H nucleus, the experimental measured pressure-dependent trend between *cis* and *trans* NMA can be reproduced correctly.

The ^{15}N nucleus shows the opposite picture. Here EC-RISM underestimates the pressure-dependent change of the chemical shift for both conformers. The experiment leads to an increase of 0.8 ppm to 2 kbar, whereas EC-RISM only shows a shift of 0.5 ppm.

The calculated absolute chemical shifts are shown in Table 13. The measured experimental chemical shift of the ^{1}H nucleus for the *trans* conformer is 7.84 ppm. The chemical shift calculated in this work is 6.47 ppm, resulting in a difference of 1.4 ppm. The calculated chemical shift was 6.8 ppm in the previous publication, and the resulting difference between the experiment and theory was 1 ppm. Concerning this observable, the results have worsened, but it should be noted that the level of theory used in this work is not optimized for calculating chemical shifts. With a deviation of 1.1 ppm, the chemical shift of the *cis*-NMA ^{1}H of the *cis*-NMA is much better represented. A possible explanation could be, on the one hand, dimerization of the *trans*-NMA or a stronger interaction with explicit waters that are missing in these calculations.

The chemical shifts for the ^{13}C nucleus of the *trans*-NMA are 177.3 ppm and 180.66 ppm for the calculation and experiment, respectively. The resulting difference is 3.3 ppm. Regarding the ^{15}N nucleus, a significant improvement can be obtained compared to published results. This finding can be attributed to choosing a plausible NMR standard; a detailed description of this standard can be found in Ref[27]. For *trans*-NMA, the experimental value is 113.8 ppm, the chemical shift calculated in this work is 112.6 ppm, and the difference between the two is only 1.2 ppm. 122 ppm was the chemical shift in our previous publication.[26]

At this point, it can be stated that EC-RISM can calculate the correct pressure-dependent tendency for all nuclei, and especially for the important ^{1}H nucleus, the chemical shifts are of a similar order of magnitude as the experimental data. The greatest inaccuracies occur in the calculation of the pressure-dependent ^{13}C chemical shifts. The best agreement regarding the absolute chemical shift at 1 bar can be obtained for the ^{15}N nucleus. A possible explanation could be that the methyl group of the DSS was chosen as the standard for ^{1}H and ^{13}C, whereas for ^{15}N, ammonia is used as the standard, and a better chemical overlap exists here. The settings for calculating chemical shifts evaluated in this chapter were applied to the minima of Ac-Gly-NHMe and Ac-Ala-NHMe, and are presented in the next section.

4.3.2 NMR spectroscopic parameters of Ac-Gly/Ala-NHMe under high-pressure conditions

After the intensive look at the pressure dependence of the chemical shifts of the peptide bond of NMA in the previous section, the next step is to examine the chemical shifts of the peptide bond atoms of the model molecules Ac-Gly-NHMe and Ac-Ala-NHMe (The remaining shifts of the capping groups and α-atoms can be found for the two main conformers in Figure 61 and Figure 62*)*. The focus will be first on the two primary minima of Ac-Ala-NHMe known from the literature and calculated with EC-RISM, i.e., α_R and P_{II} (details for the populations following chapter 4.3.5). Additionally, the results obtained for the other conformers are presented to give a complete picture. Figure 26 shows the chemical shifts of the two N-H protons and their pressure-dependent changes relative to their value at ambient pressure. In Figure 26, panels (A) and (C), which describe the proton that is bound near the acetyl group, it can be seen that all conformers have a smaller absolute chemical shift compared to the experimental value, and the pressure-dependent change is much lower compared to the experimental changes. The chemical shifts of the two main conformers and α_R are 6.6 ppm and 6.5 ppm (see Table 14 for ^{1}H chemical shifts), respectively, whereas the experimental reference value is 8.3 ppm. The energetically weighted average chemical shift over all conformers is 6.6 ppm. Therefore, it results in a deviation of 1.7 ppm, similar to the results obtained for *trans*-NMA. On the positive side, the pressure-dependent trend can be reproduced correctly, even if the calculated changes are only half of those observed experimentally. In the experiment, the chemical shift changes by 0.11 ppm up to a pressure of 3 kbar, whereas the calculated shifts change between 0.04-0.06 ppm. The differences between the individual conformers are relatively small, and the P_{II} conformer best reflects the pressure dependence for this nucleus.

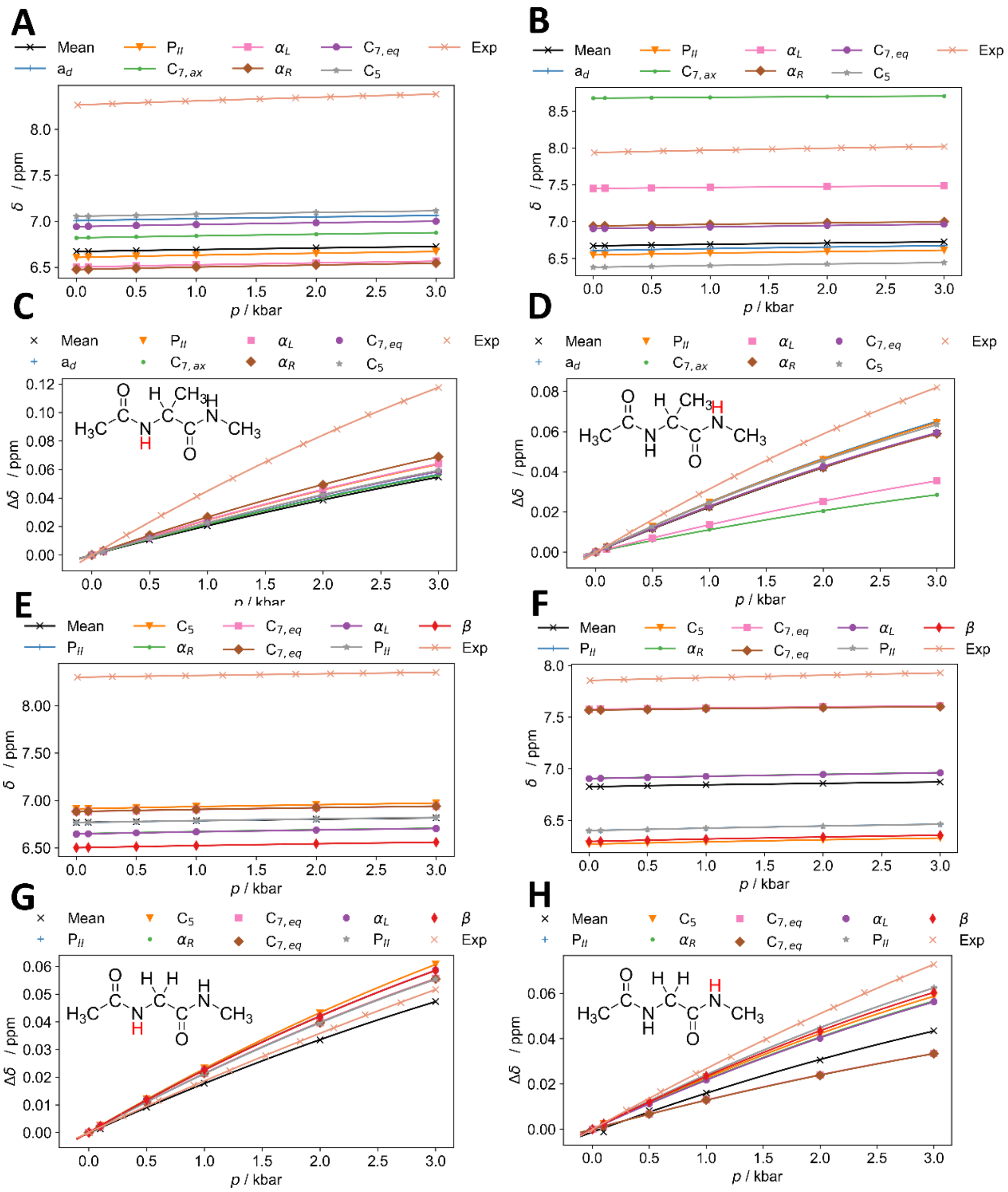

Figure 26 Pressure-dependent NMR shifts of the seven and eight Ac-Ala-NHMe and Ac-Gly-NHMe minima compared to the experimental value for the ^{1}H atoms of the two peptide bonds. A and B represent the absolute chemical shifts of the two amide hydrogens atoms of Ac-Ala-NHMe, whereas C and D represent the pressure-dependent changes in chemical shifts relative to 1 bar. In E and F, the absolute values for the corresponding nuclei of Ac-Gly-NHMe are shown, and in G and H, the difference relative to ambient pressure is presented. The theoretical calculations were performed with MP2/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Table 14: Calculated pressure-dependent chemical shifts for the ^{1}H amide protons for the seven and eight minima of Ac-Ala-NHMe and Ac-Gly-NHMe, respectively. The calculations were performed with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Conformer	1 bar	100 bar	500 bar	1 kbar	2 kbar	3 kbar
Ac-Ala-NHMe: ^{1}H of Ala N-H						
P_{II}	6.61	6.61	6.62	6.63	6.65	6.67
α_R	6.48	6.48	6.49	6.50	6.52	6.54
C_5	7.05	7.06	7.07	7.08	7.10	7.11
$C_{7,eq}$	6.50	6.50	6.51	6.53	6.55	6.57
α_L	6.94	6.94	6.95	6.96	6.98	7.00
a_d	7.01	7.01	7.02	7.03	7.05	7.06
$C_{7,ax}$	6.82	6.82	6.83	6.84	6.86	6.88
Mean	6.58	6.59	6.60	6.61	6.63	6.64
Exp	8.26	8.27	8.29	8.31	8.35	8.38
Ac-Ala-NHMe: ^{1}H of NHMe N-H						
P_{II}	6.55	6.55	6.56	6.57	6.59	6.61
α_R	6.94	6.94	6.95	6.96	6.98	7.00
C_5	6.38	6.38	6.39	6.40	6.43	6.44
$C_{7,eq}$	7.45	7.45	7.46	7.46	7.48	7.49
α_L	6.90	6.91	6.92	6.93	6.95	6.96
a_d	6.61	6.61	6.62	6.63	6.65	6.67
$C_{7,ax}$	8.68	8.68	8.68	8.69	8.70	8.70
Mean	6.75	6.75	6.76	6.77	6.79	6.81
Exp	7.94	7.94	7.95	7.97	8.00	8,02
Ac-Gly-NHMe: ^{1}H of Gly N-H						
α_R	6.65	6.65	6.66	6.67	6.69	6.71
α_L	6.65	6.65	6.66	6.67	6.69	6.71
P_{II}	6.77	6.77	6.78	6.79	6.81	6.82
P_{II}	6.77	6.77	6.78	6.79	6.81	6.82
β	6.50	6.50	6.51	6.52	6.54	6.56
C_5	6.91	6.91	6.92	6.94	6.96	6.97
$C_{7,eq}$	6.88	6.89	6.90	6.91	6.92	6.94
$C_{7,eq}$	6.88	6.89	6.89	6.90	6.92	6.94
Mean	6.69	6.69	6.70	6.71	6.73	6.74
Exp	8.30	8.30	8.31	8.32	8.34	8.35
Ac-Gly-NHMe: ^{1}H of NHMe N-H						
α_R	6.91	6.91	6.92	6.93	6.95	6.96
α_L	6.90	6.91	6.92	6.93	6.95	6.96
P_{II}	6.40	6.40	6.41	6.43	6.45	6.46
P_{II}	6.40	6.40	6.41	6.43	6.45	6.46
β	6.30	6.30	6.31	6.32	6.34	6.36
C_5	6.27	6.27	6.28	6.29	6.31	6.33
$C_{7,eq}$	7.57	7.57	7.58	7.58	7.59	7.60
$C_{7,eq}$	7.58	7.58	7.58	7.59	7.60	7.61
Mean	6.75	6.75	6.76	6.77	6.78	6.79
Exp	7.86	7.86	7.87	7.88	7.91	7.93

In the experiment, a chemical shift of 7.9 ppm (see Table 14) was obtained for the N-H nucleus, which belongs to the NHMe group (see Figure 26, panels B and D). The calculated values for the main conformers are 6.6 ppm for the P_{II} (lower blue triangle) and 6.9 ppm for α_R (yellow dots), respectively. The calculated energetically averaged value is 6.8 ppm. The value of the α_R conformer for this nucleus is the closest to the experimental value regarding absolute chemical shifts. The experimental pressure-dependent trend for this proton is slightly smaller than for the other N-H proton, with a 0.08 ppm increase up to 3 kbar. The calculated changes are approximately 0.06 ppm up to 3 kbar, with no significant difference between the two main conformers and the averaged calculated value.

Regarding the ^{13}C atoms of the two carbonyl groups, EC-RISM overestimates the experimental chemical shifts for all conformers (see Figure 27 A and B and Table 15) and, therefore, the energetically averaged mean. The experimentally measured chemical shift for the ^{13}C atom attached to the acetyl group is 177 ppm (Figure 27, panel A and Table 15). For the ^{13}C carbonyl atom attached to alanine (Figure 27 B), the experimental chemical shift is 178.5 ppm. The following chemical shifts for the two main conformers were calculated for the ^{13}C atom of the acetyl carbonyl group: P_{II} 179.3 ppm, α_R 179.8 ppm, and the mean value is 179.4 ppm. The difference to the experimental value is 2.4 ppm, which is also in line with the results for NMA.

For the second ^{13}C carbonyl atom (attached to the alanine fragment), the calculated chemical shifts are 181.8 ppm and 181.0 ppm for the P_{II} and α conformers, respectively. The averaged calculated value is 181.3 ppm, leading to an error of 2.8 ppm with respect to the experiment. For both ^{13}C nuclei of the carbonyl groups, as previously examined for NMA, only a small experimental pressure-dependent change can be observed compared to the EC-RISM calculations. Especially for the alanine ^{13}C carbonyl atom (Figure 27 D), almost no pressure-dependent change can be detected, whereas, for EC-RISM, changes of about 0.35 ppm at 3 kbar are calculated. The picture looks much better for the acetyl ^{13}C carbonyl atom, as EC-RISM only shows about twice the pressure-dependent increase compared to the experiment. The chemical shift changes up to 0.23 ppm at 3 kbar in the experiment, whereas about 0.4 ppm was calculated for the averaged value. A statement about the quality of the calculations is difficult to make since the differences between experiment and theory are relatively large.

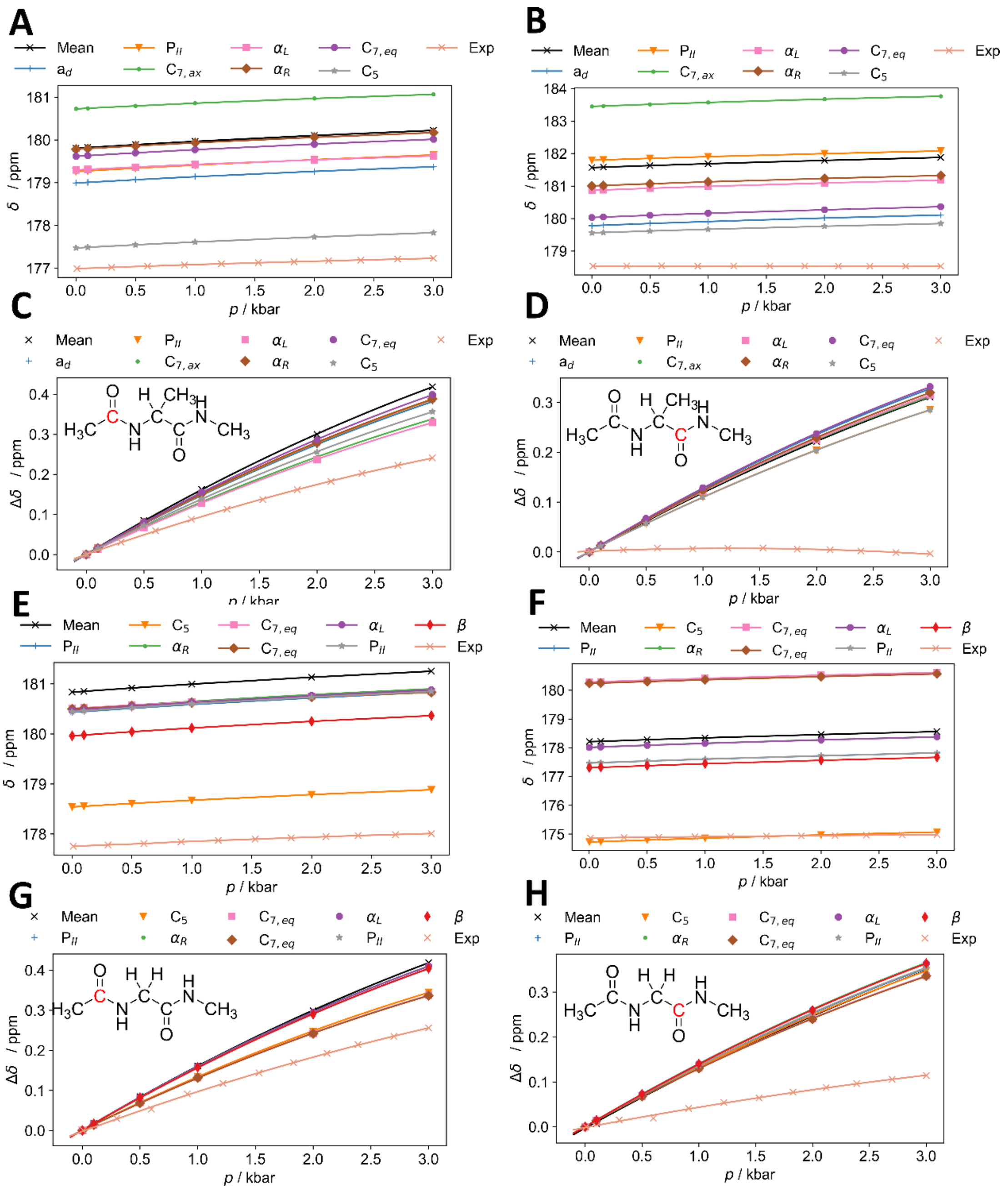

Figure 27 Pressure-dependent NMR shifts of the seven and eight Ac-Ala-NHMe and Ac-Gly-NHMe minima compared to the experimental value for the ^{13}C carbonyl atoms of the two peptide bonds. A and B represent the absolute chemical shifts of the two amide hydrogens atoms of Ac-Ala-NHMe, whereas C and D represent the pressure-dependent changes in chemical shifts relative to 1 bar. In E and F, the absolute values for the corresponding nuclei of Ac-Gly-NHMe are shown, and in G and H, the difference relative to ambient pressure. The theoretical calculations were performed with MP2/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Table 15 Calculated pressure-dependent chemical shifts for the ^{13}C amide carbonyl nuclei for the seven and eight minima of Ac-Ala-NHMe and Ac-Gly-NHMe, respectively. The calculations were performed with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Conformer	1 bar	100 bar	500 bar	1 kbar	2 kbar	3 kbar
Ac-Ala-NHMe: ^{13}C of Acetyl C-O						
P_{II}	179.26	179.28	179.34	179.41	179.54	179.65
α_R	179.78	179.80	179.86	179.93	180.06	180.17
C_5	177.47	177.49	177.55	177.61	177.73	177.83
$C_{7,eq}$	179.30	179.31	179.36	179.43	179.53	179.63
α_L	179.62	179.63	179.70	179.77	179.90	180.01
a_d	178.99	179.01	179.07	179.14	179.26	179.37
$C_{7,ax}$	180.73	180.74	180.79	180.86	180.97	181.06
Mean	179.39	179.40	179.47	179.54	179.68	179.79
Exp	176.99	177.00	177.04	177.08	177.16	177.23
Ac-Ala-NHMe: ^{13}C of Ala C-O						
P_{II}	181.79	181.80	181.85	181.90	182.00	182.08
α_R	181.00	181.02	181.07	181.13	181.23	181.32
C_5	179.56	179.57	179.62	179.67	179.76	179.84
$C_{7,eq}$	180.87	180.88	180.93	180.99	181.09	181.18
α_L	180.03	180.04	180.10	180.16	180.27	180.36
a_d	179.78	179.79	179.85	179.91	180.01	180.11
$C_{7,ax}$	183.45	183.46	183.51	183.57	183.67	183.76
Mean	181.25	181.27	181.31	181.37	181.47	181.55
Exp	178.53	178.53	178.53	178.54	178.53	178.52
Ac-Gly-NHMe: ^{13}C of Acetyl C-O						
α_R	180.49	180.51	180.57	180.65	180.79	180.90
α_L	180.47	180.49	180.55	180.63	180.76	180.88
P_{II}	180.43	180.45	180.51	180.59	180.72	180.83
P_{II}	180.44	180.46	180.53	180.60	180.73	180.85
β	179.96	179.98	180.04	180.12	180.25	180.37
C_5	178.54	178.55	178.61	178.67	178.79	178.88
$C_{7,eq}$	180.51	180.52	180.58	180.64	180.75	180.85
$C_{7,eq}$	180.49	180.51	180.56	180.63	180.74	180.83
Mean	180.42	180.44	180.50	180.58	180.71	180.83
Exp	177.75	177.76	177.80	177.85	177.94	178.01
Ac-Gly-NHMe: ^{13}C of Gly C-O						
α_R	178.01	178.03	178.09	178.16	178.28	178.38
α_L	178.01	178.03	178.08	178.15	178.27	178.37
P_{II}	177.47	177.48	177.54	177.60	177.72	177.82
P_{II}	177.47	177.48	177.54	177.61	177.72	177.82
β	177.30	177.32	177.38	177.44	177.56	177.67
C_5	174.72	174.73	174.79	174.85	174.97	175.07
$C_{7,eq}$	180.23	180.25	180.30	180.36	180.47	180.57
$C_{7,eq}$	180.28	180.30	180.35	180.41	180.52	180.62
Mean	177.90	177.91	177.96	178.02	178.13	178.22
Exp	174.86	174.87	174.88	174.91	174.94	174.98

The two remaining nuclei of the peptide bond of Ac-Ala-NHMe, which will be discussed, are the two ^{15}N nuclei. The corresponding theoretically calculated and experimentally measured chemical shifts are shown in Figure 28 and Table 16. Here one can observe a remarkably contrary picture. On the one hand, the calculated chemical shifts for the ^{15}N nucleus of the alanine fragment (Figure 28 A) deviate greatly from the experiment. On the other side, for the ^{15}N nucleus, which can be assigned to the NHMe group (Figure 28 B), the calculated chemical shifts for the two main conformers lie precisely between the experimental value. The experimental chemical shift (see Table 16) for the alanine ^{15}N nucleus is 110.1 ppm, whereas the calculated values are 130.0 ppm and 129.3 ppm for P_{II} and α_R, respectively. The energetically estimated average value is 129.7 ppm. This large discrepancy between the experiment and theory of nearly 20 ppm is difficult to rationalize physically, especially since for NMA, as shown in the previous part, and for Ac-Gly-NHMe, which is discussed in detail below, the results for the ^{15}N nuclei of peptide bond are much more reliable.

The experimental chemical shift at ambient pressure for the ^{15}N nucleus of the NHMe group (see Table 16) is 107.4 ppm. The calculated values for the two main conformers are 107.5 ppm (P_{II}) and 104.6 ppm (α_R) at a pressure of 1bar; the calculated mean value is 106.2 ppm, resulting in an error of 1.2 ppm relative to the experiment. The pressure dependence of the chemical shifts can be represented particularly well by EC-RISM for the alanine fragments ^{15}N nucleus (Figure 28 C), where the calculated average increase is 0.4 ppm. The corresponding experimental change is 0.35 ppm at 3 kbar relative to ambient pressure. The pressure dependence for the ^{15}N nucleus of the NHMe fragment is much worse resembled than for the previously discussed nucleus. While a chemical shift change of 1.4 ppm is experimentally observed, the calculated chemical shifts can only reflect a maximum change of 0.8 ppm.

Table 16 Calculated pressure-dependent chemical shifts for the ^{15} N amide nitrogen nuclei for the seven and eight minima of Ac-Ala-NHMe and Ac-Gly-NHMe, respectively. The calculations were performed with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Conformer	1 bar	100 bar	500 bar	1 kbar	2 kbar	3 kbar
Ac-Ala-NHMe: ^{15}N of Ala-N-H						
P_{II}	130.03	130.05	130.12	130.21	130.36	130.50
α_R	129.33	129.36	129.46	129.58	129.79	129.97
C_5	127.73	127.76	127.85	127.96	128.15	128.32
$C_{7,eq}$	135.11	135.14	135.23	135.33	135.51	135.67
α_L	124.93	124.96	125.08	125.22	125.47	125.69
a_d	125.70	125.72	125.82	125.94	126.14	126.32
$C_{7,ax}$	126.14	126.16	126.26	126.37	126.57	126.75
Mean	129.68	129.70	129.78	129.87	130.03	130.18
Exp	110.13	110.16	110.27	110.41	110.65	110.85
Ac-Ala-NHMe: ^{15}N of Ala-NHMe						
P_{II}	107.46	107.50	107.63	107.79	108.07	108.31
α_R	104.54	104.58	104.71	104.86	105.12	105.35
C_5	105.42	105.45	105.57	105.70	105.94	106.15
$C_{7,eq}$	111.49	111.52	111.63	111.77	112.00	112.20
α_L	102.49	102.52	102.65	102.79	103.05	103.26
a_d	107.74	107.77	107.90	108.05	108.32	108.55
$C_{7,ax}$	111.69	111.72	111.84	111.98	112.23	112.44
Mean	106.22	106.24	106.36	106.49	106.72	106.92
Exp	107.38	107.43	107.65	1007.9	108.37	108.78
Ac-Gly-NHMe: ^{15}N of Gly-N-H						
α_R	113.53	113.56	113.67	113.80	114.04	114.25
α_L	113.53	113.56	113.67	113.80	114.04	114.25
P_{II}	112.73	112.75	112.84	112.95	113.14	113.31
P_{II}	112.75	112.78	112.87	112.98	113.17	113.33
β	113.77	113.80	113.89	114.00	114.19	114.36
C_5	112.51	112.54	112.64	112.76	112.98	113.16
$C_{7,eq}$	118.79	118.82	118.92	119.04	119.25	119.44
$C_{7,eq}$	118.79	118.81	118.92	119.04	119.25	119.44
Mean	113.57	113.60	113.68	113.79	113.98	114.16
Exp	114.83	114.87	115.02	115.2	115.54	115.83
Ac-Gly-NHMe: ^{15}N of Gly-NHMe						
α_R	105.91	105.94	106.07	106.20	106.44	106.65
α_L	105.90	105.93	106.05	106.19	106.42	106.63
P_{II}	104.87	104.91	105.02	105.16	105.40	105.60
P_{II}	104.87	104.90	105.02	105.15	105.39	105.59
β	102.82	102.85	102.96	103.08	103.30	103.49
C_5	103.14	103.17	103.28	103.40	103.62	103.80
$C_{7,eq}$	112.05	112.08	112.20	112.34	112.58	112.79
$C_{7,eq}$	112.00	112.03	112.15	112.29	112.53	112.73
Mean	105.72	105.75	105.85	105.96	106.17	106.35
Exp	108.29	108.34	108.55	108.79	109.24	109.63

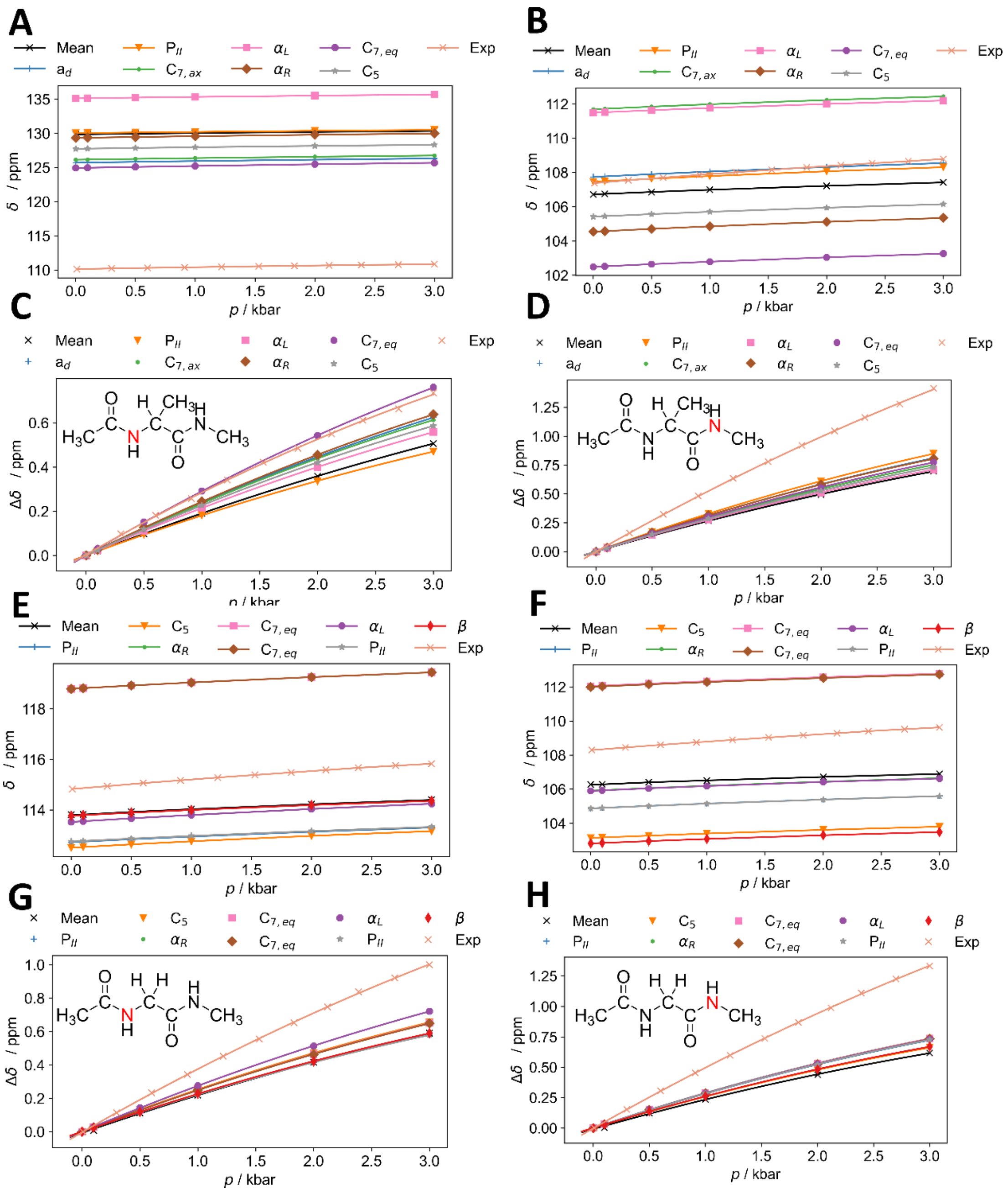

Figure 28 Pressure-dependent NMR shifts of the seven and eight Ac-Ala-NHMe and Ac-Gly-NHMe minima compared to the experimental value for the ^{15}N atoms of the two peptide bonds. A and B represent the absolute chemical shifts of the two amide hydrogens atoms of Ac-Ala-NHMe, whereas C and D represent the pressure-dependent changes in chemical shifts relative to 1 bar. In E and F, the absolute values for the corresponding nuclei of Ac-Gly-NHMe are shown, and in G and H, the difference relative to ambient pressure is presented. The theoretical calculations were performed with MP2/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

In the following step, the same analysis is made for Ac-Gly-NHMe. The focus here is also on two main conformers (α and P_{II}), whereby only one is examined in the case of mirror-symmetric conformers (the difference between two mirror-symmetric conformers is very small). The pressure-dependent chemical shifts of ^1H N-H nuclei are plotted in Figure 26. In panel A, the N-H proton, which belongs to the glycine motif, is depicted, and just as seen for Ac-Ala-NHMe, EC-RISM underestimates the experimental chemical shift. At ambient pressure, the measured shift is 8.3 ppm, whereas for the α-helical conformers (see Table 29 and the orange and cyan circles in), 6.7 ppm was obtained, and for the P_{II} conformers (red crosses and brown asterisks), the calculated value is 6.8 ppm. The calculations accurately reflect the experimentally observed pressure dependency for this nucleus, resulting in an increase of 0.05 ppm up to 3 kbar.

The measured chemical shift of the ^{1}H proton of the NHMe group is 7.9 ppm. By using the EC-RISM method, chemical shifts of 7 ppm and 6.5 ppm (P_{II}) were calculated for the α-helical and P_{II} conformers, respectively. Furthermore, a mean chemical shift averaged over the eight minima of 6.8 ppm is calculated. The overall pressure dependence of the eight minima underestimated the changes compared to the experiment. In the latter, the chemical shift change is 0.07 ppm at 3kbar, and the calculated values range between 0.04 ppm and 0.05 ppm.

The next focus will be on the two carbonyl ^{13}C nuclei of Ac-Gly-NHMe (see Figure 34), where the same observation is made as before for NMA and Ac-Ala-NHMe; EC-RISM consistently overestimates the chemical shifts compared to the experimental data. The measured chemical shift for the carbonyl atom, which can be assigned to the acetyl group, is 177.8 ppm at ambient pressure. The α-helical structures show a calculated chemical shift of 180.5 ppm, whereas the chemical shifts of the P_{II} conformers are 180.4 ppm. Since the differences between the two main conformers are minor, the average chemical shift over the data set is 180.4 ppm. The pressure dependence of the chemical shifts is overestimated by 0.2 ppm at 3 kbar, which is exactly twice the observed change in the experiment (0.2 ppm up to 3 kbar).

In the experimental spectrum at ambient pressure, the peak at 174.9 ppm (see Table 15) can be assigned to the ^{13}C carbonyl atom belonging to the glycine fragment. For the α-helix conformations, a chemical shift of 178.0 ppm was calculated, resulting in a deviation of 3.1 ppm relative to the experiment. In the range of 177.5 ppm, the chemical shifts of the P_{II} conformers (2.7 ppm deviation from the experiment) are located at ambient pressure. The corresponding value of the average is 177.9 ppm. Due to pressurization, the experimentally measured chemical shift increases by 0.1 ppm at 3 kbar, whereas in the EC-RISM calculations, a change between 0.3 ppm and 0.35 ppm was calculated for all conformers.

To conclude this chapter on the chemical shifts of the peptide bond nuclei, the focus is on the chemical shifts of the two ^{15}N nuclei of Ac-Gly-NHMe (see Figure 28 panels E-H). While a massive overestimation of the chemical shift by EC-RISM was observed for the ^{15}N nucleus, which could be assigned to the alanine fragment of Ac-Ala-NHMe, the calculated chemical shifts at 1 bar for the equivalent nucleus of Ac-Gly-NHMe are much closer to the measured value of 114.8 ppm (see Table 29). For the α-helical conformers, chemical shifts of 113.5 ppm could be calculated, and thus, a relatively small deviation of 1.3 ppm is obtained. A slightly larger deviation of 2.1 ppm was calculated for the P_{II} conformers (112.7 ppm). As a result of the pressure increase, the chemical shift in the experiment increases by 1.0 ppm. All EC-RISM calculations underestimate the experimental change, with the α-helical conformer showing the best agreement with the experiment with a 0.7 ppm increase at 3 kbar.

An experimental chemical shift of 108.3 ppm was measured for the ^{15}N nucleus of the NHMe group. The best agreement with the experiment for this specific nucleus can be found for the α-helix conformers with 105.9 ppm, a deviation of 2.4 ppm. Furthermore, a chemical shift of 104.8 ppm was calculated for the P$_{II}$ conformer. As with the other ^{15}N nuclei discussed, the calculated change in chemical shift due to pressure increase is significantly less than the measured pressure-dependent chemical shift. An increase of 1.3 ppm at 3 kbar can be measured in the experiment, whereas the calculations can only represent a maximum of 0.6 ppm.

In summary, many observations already made for TMAO and NMA can also be made for the chemical shifts of Ac-Gly-NHMe and Ac-Ala-NHMe. The chemical shifts of the ^{13}C carbonyl atoms calculated by EC-RISM are significantly higher compared to the experimental values. In the same way, the pressure-dependent change for the ^{13}C carbonyl atoms is small to virtually non-existent, and EC-RISM consistently calculates higher pressure-dependent changes. The absolute calculated values for the ^{1}H are lower than the experimental observation; however, the calculated pressure-dependent change is quite accurately reproduced. Nonetheless, an error could be attributed to small geometrical changes and explicit interactions with water molecules. A possible way to get more accurate results would be to perform an AIMD simulation analogous to the EPR parameter calculations in 4.2.2 and use a relatively large number of snapshots for the calculations. Additionally, the chemical shifts of the amide protons were used for a population fit (more details at the end of the next chapter).

4.3.3 Pressure-dependence of Ac-Gly/Ala-NHMe based on pressure-dependent EC-RISM optimized minima compared to PCM-optimized minima

This section is dedicated to provide a more detailed look at pressure-dependent EC-RISM-based geometry optimizations of Ac-Gly-NHMe and Ac-Ala-NHMe at different pressures. The minima obtained in chapter 4.3.5 were taken and reoptimized with EC-RISM at ambient conditions. These new minima were then optimized at 1 kbar, and this procedure was repeated up to a pressure of 5 kbar, with an increment of 1 kbar. The corresponding φ and ψ dihedral angles are presented in Table 17. The conformational changes are very small, e.g., the backbone dihedrals show marginal changes in the second decimal place. Compared to the B3LYP/6-311+G(d,p)/PCM optimizations, the changes due to the solvent model are greater than the pressure-induced changes.

Additionally, the pressure-dependent changes of N-H and C-O bonds of the amide groups are presented in Table 30 of the appendix, and here only the changes occur in the fourth decimal place. For example, the N-H bonds of the P_{II} show stretching of 0.0006 Å However, it can be seen that both bonds become larger under high-pressure conditions. This result could be explained by the hypothesis that the external water moves closer to the amide bonds under high pressure, thus exerting a pulling effect on the amide proton or the oxygen of the carbonyl group.

Next, the new optimized conformers were used for pressure-dependent NMR-shift calculations with MP2/6-311+G(d,p), analogous to 4.3.2. Since no reference calculations for pressure-dependent geometry-optimized structures of DSS and NH_3 were available at the time of writing, the reference shielding constants of the PCM-optimized structures were used for reference (calculated with MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM by L. Eberlein[28] and shown in Table 43). This practice will definitely produce an error, but when the shielding constants of the methyl protons are compared to the shielding constants of the amide protons, it can be seen that the effect of pressure on the amide protons (see SI 4.3 and Figure 61 and Figure 62) are an order of magnitude greater.

In Figure 29 A-D, the pressure-dependent shifts and shift changes of the amide protons of the two main conformers of Ac-Ala-NHMe are depicted. The previously shown value for the B3LYP/6-311+G(d,p)/PCM minimized structures and the B3LYP/6-311+G(d,p)/EC-RISM ones are compared. In panels A and B, the results for the amide proton next to the acetyl group are shown, and it can be seen that the absolute values for the EC-RISM optimized structures show a smaller deviation compared to PCM-optimized ones from the experiment. However, the overall deviation is still relatively high, with 1.25 ppm and about 0.8 ppm for the two amide-protons at ambient conditions. Nonetheless, the pressure-dependent changes are better resembled by the EC-RISM optimized structures, with slightly higher changes. This effect could be due to the longer N-H bonds and the resulting stronger interactions with the oxygen atoms of water, which lead to an electron-withdrawing effect.

Table 17: Pressure-dependent φ and ψ values of the seven minima of Ac-Ala-NHMe and eight minima of Ac-Gly-NHMe calculated with B3LYP/6-311+G(d,p)//B3LYP/6-311+G(d,p)/EC-RISM in comparison to B3LYP/6-311+G(d,p)/PCM optimized structures. The optimized structures can be found in SI_4.3.

φ				Ac-Ala-NHMe			
Conformer	PCM	1 bar	1 kbar	2 kbar	3 kbar	4 kbar	5 kbar
P_{II}	-90.34	-90.40	-90.40	-90.39	-90.39	-90.37	-90.36
α_R	-75.54	-74.42	-74.64	-74.63	-74.59	-74.61	-74.62
C_5	-152.15	-152.36	-152.35	-152.35	-152.35	-152.35	-152.34
$C_{7,eq}$	-85.40	-85.72	-85.71	-85.71	-85.71	-86.79	-86.76
α_L	64.64	63.73	63.73	63.73	63.72	63.77	63.77
a_d	59.12	58.90	58.90	58.90	58.90	58.92	58.92
$C_{7,ax}$	73.49	73.28	73.27	73.26	73.26	73.25	73.25
ψ							
P_{II}	143.02	146.36	146.48	146.46	146.46	146.45	146.45
α_R	-11.70	-10.90	-10.91	-10.91	-10.92	-10.93	-10.93
C_5	154.70	153.27	153.26	153.24	153.23	153.21	153.19
$C_{7,eq}$	71.42	71.81	71.79	71.78	71.78	72.34	72.33
α_L	33.36	32.51	32.51	32.51	32.51	32.64	32.63
a_d	-140.47	-141.42	-141.42	-141.41	-141.42	-141.48	-141.49
$C_{7,ax}$	-53.55	71.81	71.79	71.78	71.78	72.34	72.33
φ				Ac-GlyNHMe			
α_R	-98.26	-100.42	-100.39	-100.40	-100.22	-100.10	-100.02
α_L	98.91	100.74	99.70	100.21	100.31	100.33	100.57
P_{II}	78.07	-80.14	-80.29	-80.47	-80.30	-80.31	-80.29
P_{II}	-77.95	81.18	81.16	80.86	80.88	80.79	80.79
β	98.22	116.88	118.23	118.12	118.23	118.41	119.67
C_5	179.88	174.87	174.98	174.96	174.95	174.95	174.95
$C_{7,eq}$	-83.11	84.75	84.75	84.70	84.75	84.71	84.70
$C_{7,eq}$	83.18	-84.56	-84.49	-84.50	-84.32	-84.52	-84.43
ψ							
α_R	-1.25	-0.09	0.21	0.47	0.46	0.27	0.28
α_L	1.02	0.83	0.80	0.64	0.69	0.63	0.55
P_{II}	-157.85	155.77	156.08	156.46	156.34	156.38	156.50
P_{II}	157.83	-156.23	-156.29	-156.41	-156.51	-156.67	-156.72
β	163.59	154.58	154.46	154.13	153.94	153.80	151.79
C_5	-179.71	-176.51	-176.05	-175.89	-175.84	-175.81	-175.76
$C_{7,eq}$	60.99	-66.64	-66.69	-66.68	-66.72	-66.71	-66.69
$C_{7,eq}$	-61.29	67.98	67.94	67.97	67.56	67.92	67.89

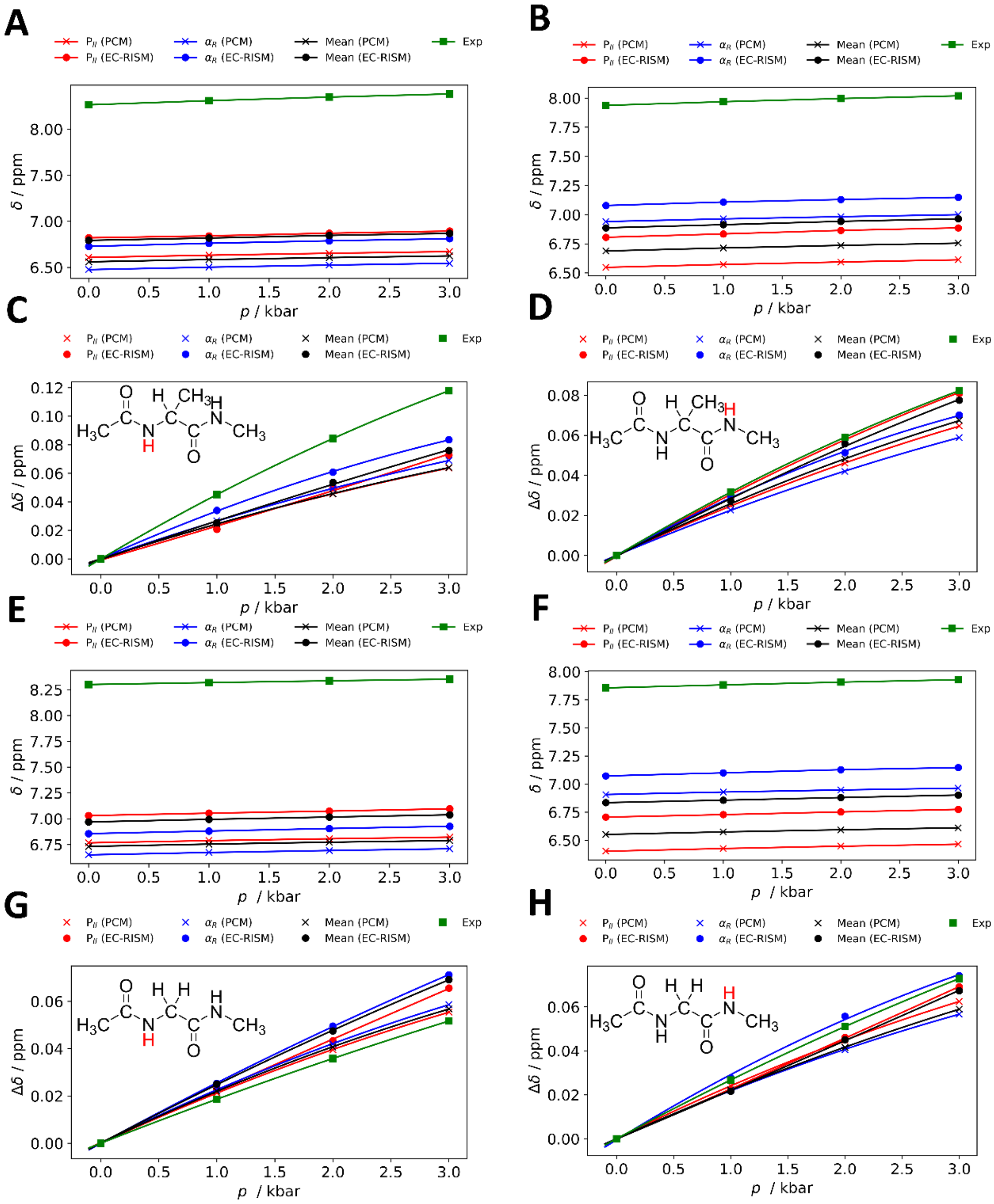

Figure 29 Pressure-dependent NMR-Shifts of the two main conformers of Ac-Ala-NHMe and Ac-Gly-NHMe compared to the experimental value for the ^{1}H peptide bond nuclei calculated with MP2/6-311+G(d,p)/EC-RISM. A and B show the absolute shifts, and C and D are the corresponding Hpressure-dependent changes for Ac-Ala-NHMe. E and F show the absolute shifts, and G and D are the corresponding pressure-dependent changes for Ac-Gly-NHMe. The red lines show the P_{II} conformer, and the blue lines show the α conformers. The green line represents the experimental shifts. The x denotes conformers optimized with B3LYP/6-311+G(d,p)/PCM(1 bar), and the dot points represent the conformers optimized with B3LYP/6-311+G(d,p)/EC-RISM(p-dep). Raw shielding constants can be found in SI_4.3.

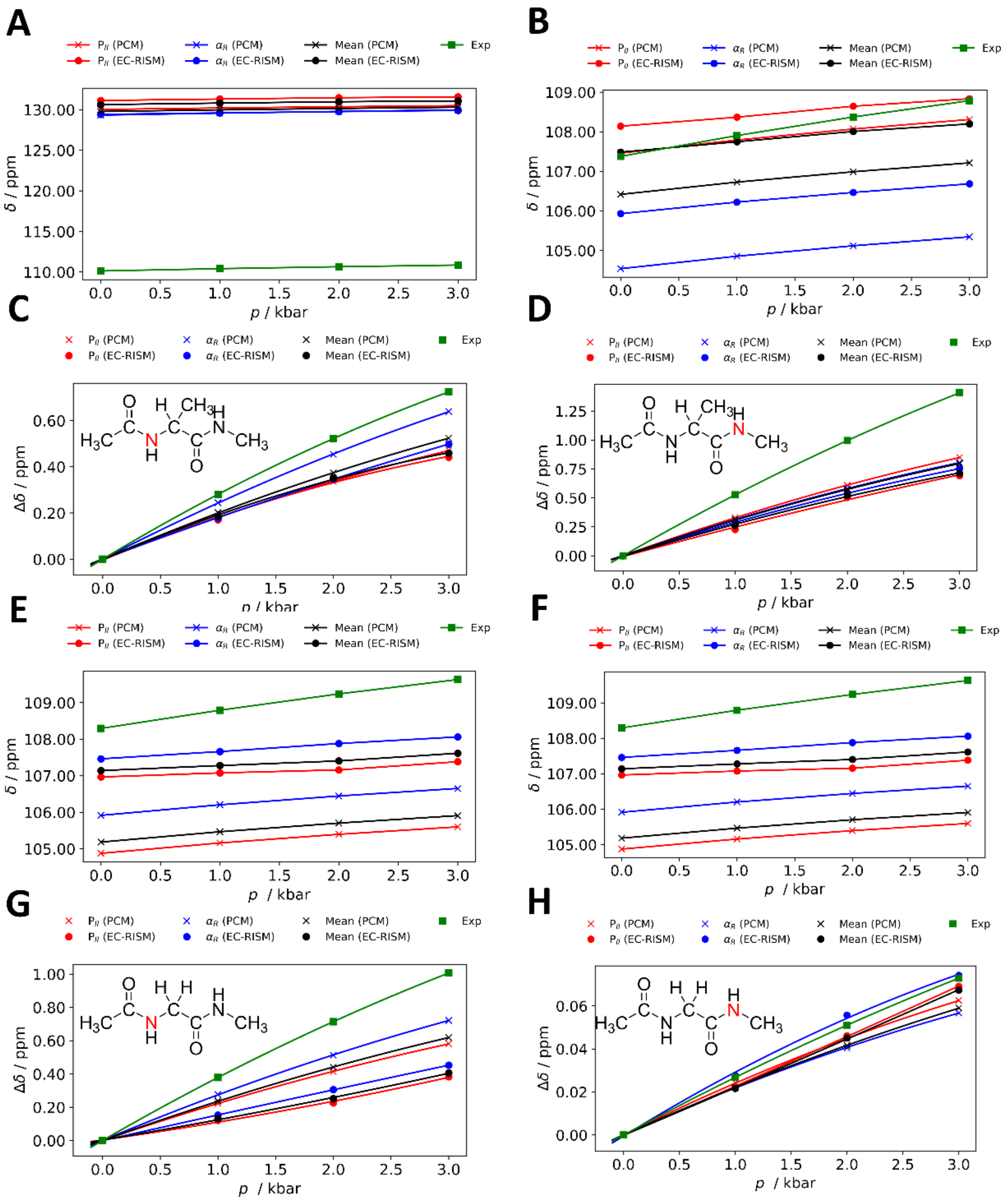

Figure 30 Pressure-dependent NMR-Shifts of the two Ac-Ala-NHMe and Ac-Gly-NHMe main conformers compared to the experimental value for the ^{15}N peptide bond nuclei calculated with MP2/6-311+G(d,p). A and B show the absolute shifts, and C and D are the corresponding Hpressure-dependent changes for Ac-Ala-NHMe. E and F show the absolute shifts, and G and D are the corresponding pressure-dependent changes for Ac-Gly-NHMe. The red lines show the P_{II} conformer, and the blue lines show the α conformers. The green line represents the experimental shifts. The x denotes conformers optimized with B3LYP/6-311+G(d,p)/PCM(1 bar), and the dot points represent the conformers optimized with B3LYP/6-311+G(d,p)/EC-RISM(p-dep). Raw shielding constants can be found in SI_4.3.

For the ^{15}N and ^{13}C nuclei of the amide bond, the chemical shift plots are depicted in Figure 30 and Figure 31, respectively. The calculated chemical shifts of the nitrogen bound next to the acetyl group (panels A and B) are even greater than the previously calculated ones for the PCM-optimized structures. The calculated pressure-dependence is for the α_R conformer significantly smaller, and for the P_{II}, only a marginal change is observed. Similar results are obtained for the other ^{15}N nucleus., whereby the change in the absolute numbers is higher for the α_R conformer. The fit between the experiment and calculations is excellent for the mean value of the EC-RISM optimized structures. The pressure-dependent changes are smaller than the corresponding values from the PCM-optimized structures. However, the question regarding correct referencing remains and could be a huge source of error.

For the ^{13}C shifts, similar observations as for the ^{15}N shifts are obtained. The absolute values are higher due to the EC-RISM optimization, which further worsens relative to the experiment. However, the effect of the EC-RISM optimizations on the chemical shifts of the ^{13}C methyl group (see SI 4.3) nuclei is more intense than the ^{1}H changes. Thus, the error by using the PCM-optimized DSS shielding constants could be a larger source of error. On the other side, the pressure-dependent change is becoming smaller due to the pressure-dependent EC-RISM optimizations; therefore, the error relative to the experiment decreases. However, the pressure change of the ^{13}C carbonyl atom, which can be linked to the glycine fragment, is still massively overestimated.

Regarding the conformational changes of Ac-Gly-NHMe, the same results are observed as for Ac-Ala-NHMe. The changes in the backbone dihedral angles are mostly less than 1° compared to the 1 bar value. The C-O and N-H bonds (Table *31*) slightly increase upon pressurization and are in the same range as the results for Ac-Ala-NHMe. The pressure-dependent chemical shifts of the α- and P_{II} conformers and the corresponding average are depicted in Figure 29, Figure 30, and Figure 31(^{1}H, ^{15}N, and ^{13}C). Again, a comparison between MP2/6-311+G(d,p)/EC-RISM// B3LYP/6-311+G(d,p)/PCM and MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM relative to the experiment is made. The absolute chemical shifts of the ^{1}H nuclei increase at ambient conditions and minimize the error relative to the experimental observation. The increase compared to PCM-optimized structures is in the range of 0.2 to 0.35 ppm. The pressure dependence is also enhanced, which is slightly overestimated for the hydrogen bonded to the glycine fragment (see Figure 29 panel G). The pressure-dependent change of the other ^{1}H nucleus is exactly reproduced by the α_R conformer (see Figure 29 panel H).

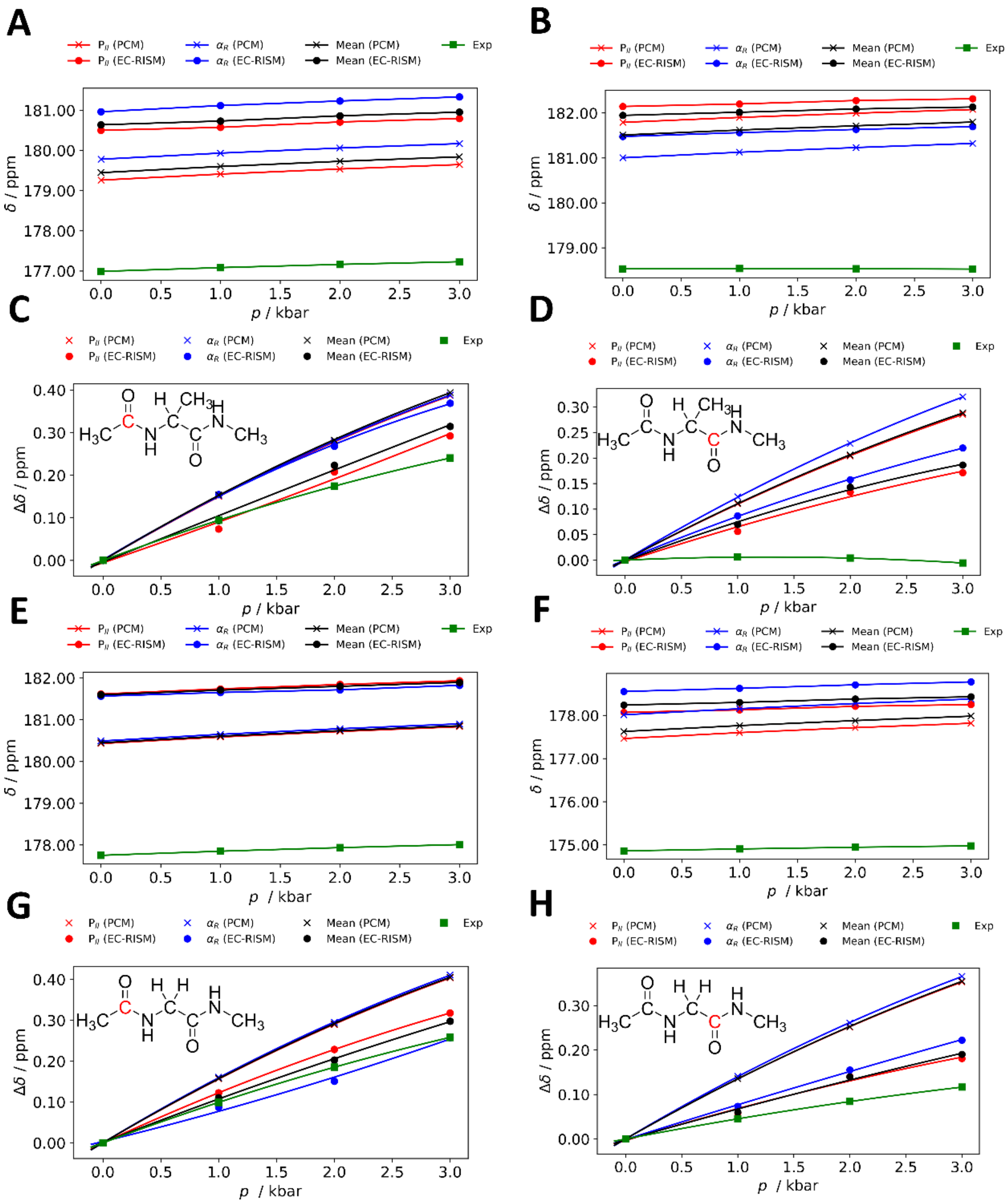

Figure 31 Pressure-dependent NMR-Shifts of the two Ac-Gly-NHMe main conformers compared to the experimental value for the ^{15}N peptide bond nuclei calculated with MP2/6-311+G(d,p). A and B show the absolute shifts, and C and D are the corresponding Hpressure-dependent changes for Ac-Ala-NHMe. E and F show the absolute shifts, and G and D are the corresponding pressure-dependent changes for Ac-Gly-NHMe. The red lines show the P$_{II}$ conformer, and the blue lines show the α conformers. The green line represents the experimental shifts. The + denotes conformers optimized with B3LYP/6-311+G(d,p)/PCM(1 bar), and the dot points represent the conformers optimized with B3LYP/6-311+G(d,p)/EC-RISM(p-dep). Raw and calculated data can be found in SI_4.3.

The difference between the two minima is negligible for the ^{13}C nucleus of the acetyl group. Nonetheless, an increase of 1 ppm compared to the PCM-optimized structures (see Figure 31 E and F) is observed if the geometry is optimized under EC-RISM solvation. The same effect is present for the ^{13}C nucleus of the glycine fragment, leading to worse results than the experimental reference. Nevertheless, the pressure-dependent changes are significantly smaller when the EC-RISM optimized structures are considered, and thus, the theory slightly better resolves the experimental observation.

The mean absolute shifts (Figure 30) calculated for the EC-RISM optimized structures for the ^{15}N nuclei resolve the experiment better at ambient conditions. However, for the glycine ^{15}N (Figure 30 E and G), the effect is not consistent for both conformers, whereby the α-conformer shows a small decrease, and the P_{II} has a significantly larger chemical shift. The chemical shift increases consistently for both conformers of the other ^{15}N nucleus. Regarding the pressure-dependent changes, a smaller shift difference is observed, which agrees with the results for Ac-Ala-NHMe.

In this chapter, pressure-dependent geometry optimizations of the two dipeptides were performed. It was shown that the effect of pressure on the geometries is relatively small; nonetheless, some interesting trends could be observed, like the shortening of C-O and N-H bonds. The effect on the relative populations of the conformers is small and in the same range as for the pressure-dependent MP2/6-311+G(d,p)/EC-RISM calculations on the PCM-optimized structures. In the case of Ac-Ala-NHMe, the P_{II} conformer has a higher occupation probability when the EC-RISM-optimized structures are considered compared to the PCM-optimized ones. For Ac-Gly-NHMe, only marginal trends were observed. Regarding the NMR chemical shifts, the amide protons can be better represented by EC-RISM-optimized structures for 1 bar and high pressure. In the calculation of ^{13}C carbonyl chemical shifts at ambient conditions, the EC-RISM optimized structures showed slightly higher deviation from the experiment compared to the PCM-optimized ones; nonetheless, the chemical shift change upon pressurization is better resembled by the pressure-dependent optimized structures. For the ^{15}N, the accuracy is better for the EC-RISM-optimized structures at ambient conditions; however, no clear picture is obtained for the pressure-dependent change.

4.3.4 Benchmarking: equilibrium of *cis* and *trans* NMA under high pressure

NMA is the simplest model to reflect an element of the peptide backbone, which can be present in two different states (excluding possible dimerizations)[299], the *trans* and the *cis* state (see Figure 1). The *trans*-NMA conformer is more populated than the *cis*-conformer, where the relative energy difference between the two conformers is approximately 2.3 kcal mol^{-1} in aqueous solutions at ambient pressure.[57] In Table 18, the population and the relative free energy of the *cis*-conformer for different EC-RISM calculations compared to experimental results and ff19SB-RISM calculations are enlisted. For EC-RISM, MP2/6-311+G(d,p) and DLPNO-CCSD/def2-TZVPP on B3LYP/6-311+G(d,p)/PCM and B3LYP/6-311+G(d,p)/EC-RISM optimized structures were performed. Additionally, thermal corrections calculated with B3LYP/6-311+G(d,p) were added to the EC-RISM free energies.

Table 18: Relative populations and free energies of *cis*- NMA compared to *trans*-NMA at different pressures, and all values are in %. For rows with a degree symbol °, the structures were optimized by applying ff19SB/ALPB. The experimental results were obtained by NMR-spectroscopy and were measured by M. Beck-Erlach.[298] In parenthesis, the term PMV signals that the pressure-dependent change of ΔG was solely computed by the integral over the pressure-dependent change of the partial molar volume according to eq (185). For the other RISM-based calculations, the normal pressure-dependent PMV correction was applied to the excess chemical potential (for details, see [27]). Thermal corrections were calculated with B3LYP/6-311+G(d,p)/PCM. Raw and calculated data can be found in SI_4.3.

Method		1 bar	500 bar	1 kbar	2 kbar
	Optimizations	Populations			
Experimental (NMR, M. Beck-Erlach)[298]		1.69	1.77	1.87	2.00
Experimental (IR Ataka[57])		2.02	-	-	-
MP2/6-311+G(d,p)	B3LYP/6-311+G(d,p)/PCM	2.16	2.22	2.30	2.41
MP2/6-311+G(d,p) (PMV)	B3LYP/6-311+G(d,p)/PCM	2.16	2.21	2.26	2.35
MP2/6-311+G(d,p)/Tcorr	B3LYP/6-311+G(d,p)/PCM	1.97	2.04	2.09	2.20
MP2/6-311+G(d,p) (PMV)/Tcorr	B3LYP/6-311+G(d,p)/PCM	1.97	2.02	2.06	2.14
DLPNO-CCSD/def2-TZVPP[‡]	B3LYP/6-311+G(d,p)/PCM	1.91	1.96	2.01	2.10
DLPNO-CCSD/def2-TZVPP[‡] (PMV)	B3LYP/6-311+G(d,p)/PCM	1.91	1.94	1.97	2.02
DLPNO-CCSD/def2-TZVPP[‡]/Tcorr.	B3LYP/6-311+G(d,p)/PCM	1.74	1.79	1.83	1.90
DLPNO-CCSD/def2-TZVPP[‡] (PMV)/Tcorr	B3LYP/6-311+G(d,p)/PCM	1.74	1.77	1.79	1.84
MP2/6-311+G(d,p)	B3LYP/6-311+G(d,p)/EC-RISM	2.46	2.57	2.67	2.85
MP2/6-311+G(d,p) (PMV)	B3LYP/6-311+G(d,p)/EC-RISM	2.46	2.54	2.60	2.73
MP2/6-311+G(d,p)/Tcorr	B3LYP/6-311+G(d,p)/EC-RISM	2.24	2.32	2.43	2.60
MP2/6-311+G(d,p) (PMV)/Tcorr	B3LYP/6-311+G(d,p)/EC-RISM	2.24	2.31	2.37	2.50
DLPNO-CCSD/def2-TZVPP[‡]	B3LYP/6-311+G(d,p)/EC-RISM	2.14	2.22	2.26	2.41
DLPNO-CCSD/def2-TZVPP[‡] (PMV)	B3LYP/6-311+G(d,p)/EC-RISM	2.14	2.19	2.23	2.31
DLPNO-CCSD/def2-TZVPP[‡]/Tcorr.	B3LYP/6-311+G(d,p)/EC-RISM	1.95	2.02	2.06	2.19
DLPNO-CCSD/def2-TZVPP[‡] (PMV)/Tcorr	B3LYP/6-311+G(d,p)/EC-RISM	1.95	1.99	2.03	2.11
ff19SB +Δ RISM° (p)	ff19SB/ALPB	0.23	0.23	0.24	0.24
ff19SB +Δ RISM° (p) (PMV)	ff19SB/ALPB	0.23	0.23	0.24	0.24
	ΔG /kcal mol^{-1}				
Experimental (NMR, M. Beck-Erlach)[298]		2.40	2.38	2.35	2.31
Experimental (IR Ataka[57])		2.30	-	-	-
MP2/6-311+G(d,p)	B3LYP/6-311+G(d,p)/PCM	2.26	2.24	2.22	2.19
MP2/6-311+G(d,p)[‡](PMV)	B3LYP/6-311+G(d,p)/PCM	2.26	2.24	2.23	2.21
MP2/6-311+G(d,p) /Tcorr	B3LYP/6-311+G(d,p)/PCM	2.31	2.29	2.28	2.25
MP2/6-311+G(d,p) (PMV) /Tcorr	B3LYP/6-311+G(d,p)/PCM	2.31	2.30	2.29	2.27
DLPNO-CCSD/def2-TZVPP[‡]	B3LYP/6-311+G(d,p)/PCM	2.33	2.32	2.30	2.28
DLPNO-CCSD/def2-TZVPP[‡] (PMV)	B3LYP/6-311+G(d,p)/PCM	2.33	2.32	2.32	2.30
DLPNO-CCSD/def2-TZVPP[‡] /Tcorr	B3LYP/6-311+G(d,p)/PCM	2.39	2.37	2.36	2.34
DLPNO-CCSD/def2-TZVPP[‡] (PMV) /Tcorr	B3LYP/6-311+G(d,p)/PCM	2.39	2.38	2.37	2.36
MP2/6-311+G(d,p)*	B3LYP/6-311+G(d,p)/EC-RISM	2.18	2.15	2.13	2.09
MP2/6-311+G(d,p)* (PMV)	B3LYP/6-311+G(d,p)/EC-RISM	2.18	2.16	2.14	2.11
MP2/6-311+G(d,p)*/Tcorr	B3LYP/6-311+G(d,p)/EC-RISM	2.24	2.21	2.19	2.15
MP2/6-311+G(d,p)* (PMV)/Tcorr	B3LYP/6-311+G(d,p)/EC-RISM	2.24	2.22	2.20	2.17
DLPNO-CCSD/def2-TZVPP[‡]	B3LYP/6-311+G(d,p)/EC-RISM	2.27	2.24	2.23	2.19
DLPNO-CCSD/def2-TZVPP[‡] (PMV)	B3LYP/6-311+G(d,p)/EC-RISM	2.27	2.25	2.24	2.22
DLPNO-CCSD/def2-TZVPP[‡]/Tcorr.	B3LYP/6-311+G(d,p)/EC-RISM	2.32	2.30	2.29	2.25
DLPNO-CCSD/def2-TZVPP[‡] (PMV)/Tcorr	B3LYP/6-311+G(d,p)/EC-RISM	2.32	2.31	2.30	2.27
ff19SB +RISM° (p)	ff19SB/ALPB	3.60	3.59	3.58	3.57
ff19SB +Δ RISM° (p) (PMV)	ff19SB/ALPB	3.60	3.59	3.59	3.58

At ambient conditions, the experiment yields a population of 1.69% for the *cis*-conformer, which is a little smaller than the data from the literature, where, e.g., Ataka determined a population via IR measurements of roughly 2%.[57] Due to pressurization, a steady increase of the *cis*-conformer is observed, resulting in a population of 2.0% at 2 kbar. Four theoretical approaches will be examined to calculate the *cis*/*trans* equilibrium. The first is the well-established combination MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM with pressure-dependent PMV-correction of the excess chemical potential.[27] The calculated population of the *cis*-conformer is 2.16% and, therefore, slightly higher than the experimental value. However, the difference is small and can be translated into energy differences within the margin of error of EC-RISM calculations. The pressure-dependent increase can then be correctly reproduced by EC-RISM, leading to a population of the *cis*-conformer of 2.41 % at 2 kbar. The second MP2/6-311+G(d,p) approach, shown in Table 18, does not use the pressure-dependent PMV correction; instead, the excess chemical potential change is solely calculated by Eq (185). The differences to the former model are small, and the population of the *cis*-conformer is 2.35 % at 2 kbar.

Further EC-RISM calculations were performed using the DLPNO-CCSD/def2-TZVPP/EC-RISM//B3LYP/6-311+G(d,p)/PCM level of theory.[129–131] However, it should be noted that at the current time, no explicit PMV correction is available for this level of theory. In order to perform a pressure-dependent PMV-correction of the excess chemical potential, the parameters developed originally for MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM were reused. For ambient conditions, the resulting occupation probability of the *cis*-conformer is 1.8 %, and it shows a slightly better performance relative to the experiment compared to the MP2/6-311+G(d,p). Regarding the pressure trend, the two DLPNO-CCSD/def2-TZVPP variants show similar behavior as the MP2/6-311+G(d,p) calculations.

The pressure-dependent MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM results are shown in the sixth row. Here, the absolute value at ambient conditions is 2.5 %, slightly higher than the values of the other EC-RISM calculations based on the PCM-optimized structures.

These results underline that high-level EC-RISM calculations can accurately represent pressure-dependent solvation-caused effects, and it provides a reliable basis to adopt the EC-RISM calculations to more complex systems in the next part of this chapter. Additionally, thermal corrections based on B3LYP/6-311+G(d,p)/PCM calculations were added to the free energy values. The thermal corrections are the same for all methods and were only calculated for ambient conditions. Due to the addition, the *cis*-conformer gets a 0.03 kcal/mol penalty.

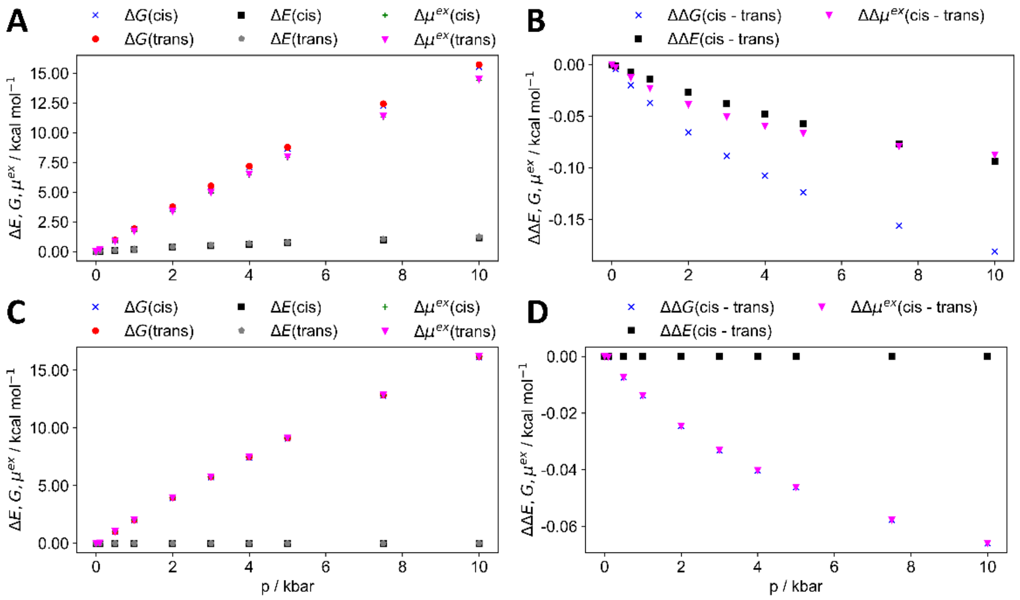

Figure 32 Pressure-dependent energy components (G, E, and μ^{ex}) calculated for *cis* and *trans*-NMA. Panel (A) shows the relative changes of the three energy types calculated with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM. Panel (B) depicts the corresponding ΔΔ differences between *cis* and *trans* NMA. Panel (C) shows the relative pressure-dependent changes of the three main energy components calculated with ff19SB+RISM. In Panel (D), the relative ΔΔ differences between the two conformers are shown. The corresponding data is shown in Table 19

Nonetheless, this thesis aims to evaluate how well the current force fields can represent pressure-dependent changes without explicit reparameterization. Therefore, the Amber version ff19SB[35] was chosen as the force field. To create a comparable basis to the EC-RISM calculations, the ff19SB was used to calculate the intramolecular energy. The pressure dependence is represented by 3D RISM, whereby the point charges and Lennard-Jones parameters polarize the solvent. At 1 bar, the population of the *cis*-conformer is only 0.26 %, which is only a tenth of the obtained value from MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-11+G(d,p)/PCM. Although the pressure change can be reproduced correctly, the absolute change is a third of the experimentally observed change and of the EC-RISM calculations.

Table 19 Pressure-dependent energy components of *cis*- and *trans*-NMA relative to the 1 bar value for MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM (with dagger‡) and ff19SB +RISM//ff19SB/ALPB, denoted by the star *. Since no pressure-dependent change of the intramolecular energy is obtained for the force field calculations, the corresponding

Pressure	$\Delta G_{\mathrm{trans}}$‡	ΔG_{cis}‡	$\Delta E_{\mathrm{trans}}$‡	ΔE_{cis}‡	$\Delta\mu_{\mathrm{trans}}$‡	$\Delta\mu_{\mathrm{cis}}$‡	$\Delta G.\mu_{\mathrm{trans}}$*	$\Delta G.\mu_{\mathrm{cis}}$*
1 bar	0.	0.	0.	0.	0.0	0.0	0.	0.
100 bar	0.19	0.19	0.02	0.02	0.17	0.16	0.02	0.02
500 bar	0.97	0.95	0.11	0.1	0.86	0.85	1.03	1.02
1 kbar	1.93	1.89	0.21	0.2	1.72	1.69	2.03	2.02
2 kbar	3.77	3.70	0.39	0.36	3.38	3.34	3.95	3.93
3 kbar	5.52	5.43	0.54	0.5	4.98	4.93	5.77	5.74
4 kbar	7.19	7.08	0.67	0.62	6.52	6.46	7.49	7.45
5 kbar	8.76	8.65	0.79	0.73	7.98	7.92	9.11	9.07
7.5 kbar	12.42	12.27	1.04	0.96	11.38	11.31	12.85	12.79
10 kbar	15.71	15.53	1.24	1.15	14.47	14.38	16.2	16.13

The three single energy components (G, E, and μ^{ex}, see Table 19) will be analyzed to understand the underlying characteristics. The corresponding pressure-dependent profiles are depicted in Figure 32. The pressure-dependent changes of the three components for the EC-RISM calculations are shown in panel (A). The absolute change of the free energy is mainly driven by the excess chemical potential, whereby the increase is roughly 15 kcal/mol at a pressure of 10 kbar. The increase of the intramolecular energy is significantly smaller, with 1.1 kcal/mol at 10 kbar. Panel (C) shows the corresponding terms for the ff19SB+RISM calculations. In contrast to the EC-RISM calculations, no solute polarization is obtained in the force field calculations since the ff19SB is an unpolarizable force field. Therefore, no pressure-dependent change of intramolecular energy is observed. Thus one can conclude that pressure-dependent electronic polarization due to the influence of the solvent is important. Similar results were obtained for TMAO in ref[18], where electronic polarization was necessary to reproduce AIMD-based spatial distribution functions. For the excess chemical potential, about 16 kcal/mol increase up to 10 kbar can be observed. This result fits well with the overall free energy change obtained for the EC-RISM calculations.

However, to answer the question if the ff19SB force field is suitable to use in MD simulations under high-pressure conditions, the more important quantity is the relative change of the energy terms between *cis*- and *trans*-NMA. In Figure 32, panels (B) (EC-RISM) and (D) (ff19SB+RISM), the energy differences between *cis*- and *trans*-NMA are plotted to address this question. In the case of EC-RISM, the observed changes are roughly -0.07 kcal/mol for both intramolecular energy and excess chemical potential. Thus, the relative change for the free energy is -0.15 kcal/mol at 10 kbar. In the case of ff19SB, only the chemical excess potential changes about -0.07 kcal/mol at 10 kbar are obtained. Therefore, ff19SB underestimated the pressure change provided by EC-RISM with 0.08 kcal/mol. However, these relative pressure-dependent changes are so small that the absolute error introduced by utilizing the ff19SB force field is significantly larger than the pressure effect.

4.3.5 Pressure-dependent free energy landscapes of Ac-Gly/Ala-NHMe

Pressure-dependent EC-RISM-based calculations successfully represent the populations of *cis* and *trans*-NMA compared to experimentally obtained NMR reference data. With that knowledge, the more complicated systems Ac-Gly-NHMe and Ac-Ala-NHMe shall now be analyzed. Due to the significantly higher conformational complexity, an exhaustive dihedral scan of the φ and ψ angles was chosen as the starting point, resulting in 400 different angle pair combinations. Subsequently, each angle combination was quantum-mechanically optimized in geometry, resulting in a finite number of minima for the two. For Ac-Gly-NHMe and Ac-Ala-NHMe, eight and seven distinguishable minima were found, respectively (see Figure 33.) This chapter will give a detailed discussion about the minimum population. For pressure-dependent EC-RISM calculations, different setups will be compared. On the one hand, the structures based on B3LYP/6-311+G(d,p)/PCM optimization will be studied with MP2/6-311+G(d,p) at different pressures. Here, the pressure-dependent PMV correction of the excess chemical potential was applied, with and without the addition of B3LYP/6-311+G(d,p)/PCM thermal corrections. Additionally, the change of ΔG was calculated via eq (185).

On the other hand, the same procedures were applied to the new B3LYP/6-311+G(d,p)/EC-RISM(p) optimized structures. Furthermore, pressure-dependent NMR calculations were performed for the minima and based on these calculations, pressure-dependent populations were fitted to resemble the experimental observation. S. Kast and S. Maste[300] co-developed the fit procedure, and S. Maste performed the fits, whereby he first measured the pressure-dependent chemical shielding constants of some reference substances to remove the intrinsic errors of the EC-RISM calculations in the chemical shifts. The corresponding formula is shown in the appendix in chapter 6.8. The chemical shifts obtained after the correction and the corresponding mean value after fitting are presented in the appendix in Figure 69 and Figure 70

In the literature, some references are available regarding the populations of dipeptide conformations. Takekiyo[69,70] previously studied the pressure dependence of dipeptide conformations by applying Raman spectroscopy. of the skeletal stretching region. Four conformations could be found for Ac-Ala-NHMe in the spectrum, whereby only the $\mathrm{P_{II}}$ conformation and the α_R conformation could be clearly assigned. The assignment was based on extensive DFT studies from Han[301] and Deng.[302] The peak of the $\mathrm{P_{II}}$ conformer has the largest intensity at ambient conditions, whereas the α_R has the second highest intensity. By increasing the pressure up to 2.5 kbar, the peak associated with the α_R conformer vanishes. To describe the relative pressure-dependence between the different conformers, the partial molar volume difference ΔV relative to the $\mathrm{P_{II}}$ conformer was determined via:

$$\Delta V = -RT\left\{\frac{\partial \ln\left(\frac{I_j}{I_i}\right)}{\partial p}\right\}, \tag{189}$$

where I is the integrated intensity of the conformers i and j, the corresponding results for the volume difference from Takekiyo[69] can be found in Table 21. Here, the assumption was made that the ratio of scattering cross sections of the individual conformers is independent of pressure. They observed a strong change in the population of the α_R for Ac-Ala-NHme upon pressurization, and at a pressure of 1.6 kbar, the α_R peak was no longer visible. In another study, Takekiyo made the same measurements for the Ac-Gly-NHMe dipeptide.[70] In contrast to the Ac-Ala-NHMe experiment, the α peak remained visible at high pressure, and even the population increased relative to the $\mathrm{P_{II}}$ conformer. It doesn't seem easy to explain why for the Ac-Ala-NHMe, the α_R conformer vanished, and for the glycine dipeptide, it increased, contradicting also calculated results discussed below. For Ac-Ala-NHMe, the populations at 1 bar can be extracted from the given publication (taken from Grdadolnik[64] and summarized in Table 20), and they are 69% for the $\mathrm{P_{II}}$, 17% for the α_R, 6% for $\mathrm{C_5}$, and 8% for $\mathrm{C_{7,\,eq}}$. Another combined IR and Raman study[303] measured the populations obtained from the IR amide III band and the Raman skeletal vibration. The averaged populations over the two measurements (see Table 20) are 68% for $\mathrm{P_{II}}$, 15% for α_R, and 17% for $\mathrm{C_5}$. In both measurements, the $\mathrm{P_{II}}$ conformer was clearly dominant, yet the populations of α_R (IR: 11%, Raman: 18%) and $\mathrm{C_5}$ (IR: 29%, Raman: 6%) differ significantly depending on the chosen method. Grdadolnik *et al.* have not found a clear reason; however, they suggested that the Raman scattering coefficients have different values for various conformers.[64] The group performed another study, where the amide III region was investigated with IR as well as Raman spectroscopy (see Table 20). Here a consistent picture between IR and Raman was obtained, whereby the measured populations were 60% for $\mathrm{P_{II}}$, 11% for α_R, and 29% for $\mathrm{C_5}$.[64] Here, the populations of Ac-Gly-NHMe at ambient conditions were also measured, and in contrast to all other amino acids, the α_R conformer is with 66% the dominant observed conformer; for the $\mathrm{P_{II}}$ a population of 22% and for the $\mathrm{C_5}$ of 12%. Gaigeot[56,86] utilized an extensive DFT-MD study in liquid water to reproduce the experimental IR spectrum. They found that $\mathrm{P_{II}}/\beta$ (these two conformers were examined as one) and the α_R have nearly no energetic difference, and the interconversion barrier between these two states is 2 kcal/mol (these results could be challenged by extended statistics, as the authors explained). However, to resolve the experimental spectrum, the population ratios needed to be changed to a 2:1 ratio favoring the $\mathrm{P_{II}}/\beta$ conformers.

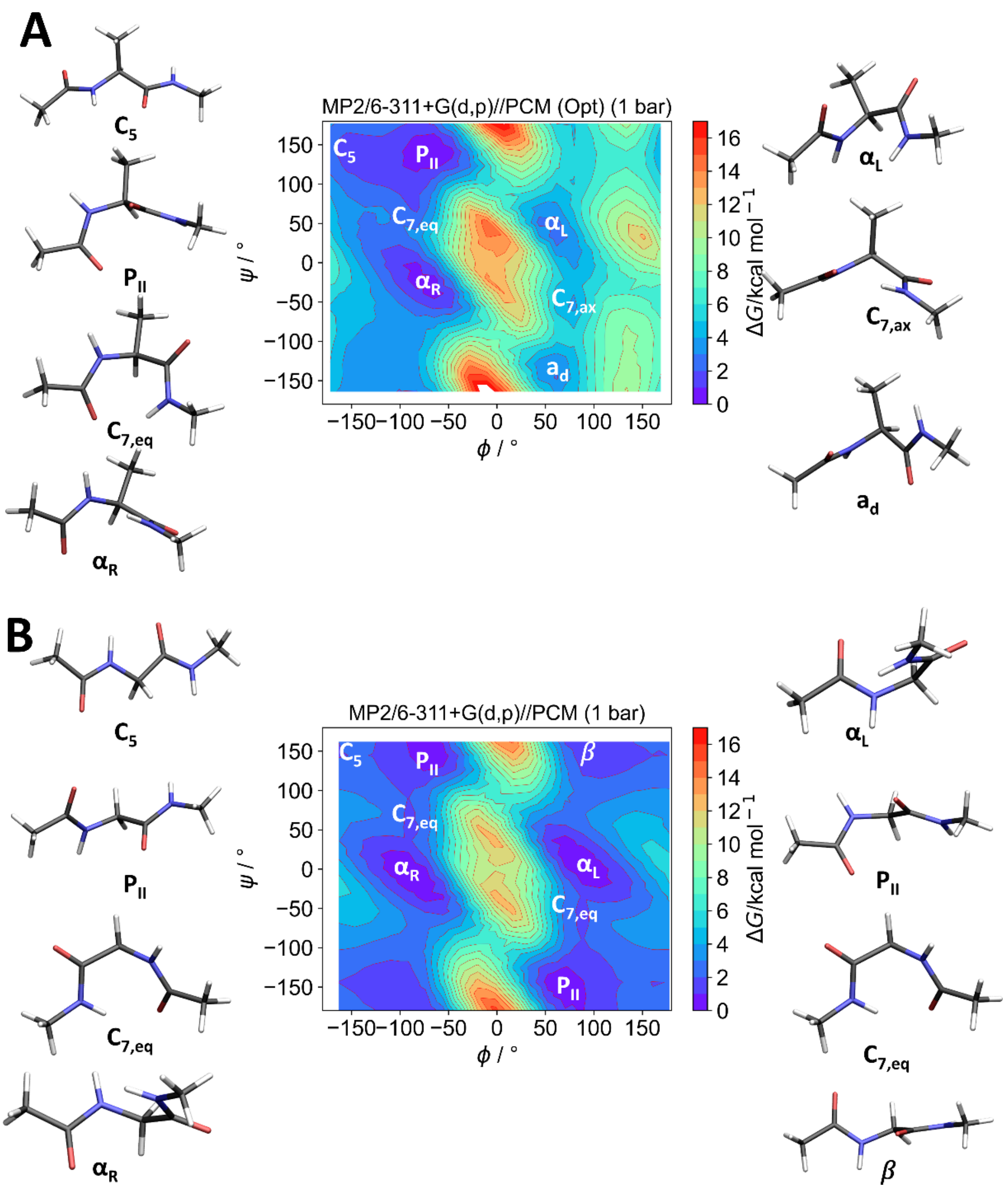

Figure 33 Conformational free energy landscapes of Ac-Ala-NHMe (A) and Ac-Gly-NHMe calculated with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM (1 bar). The positions of the optimized minima in the Ramachandran plane[304] are shown. Calculated and raw data can be found in SI_4.3.

In the following, the EC-RISM results will be presented. Regarding Ac-Ala-NHMe, a more compact picture (in terms of relevant conformers) is observed than for Ac-Gly-NHMe, since no symmetry plane in the Ramachandran plane due to the CH_3 group at C_α (see Figure 33) is present. Two clear minima can be isolated for the MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM calculations (the descriptions of used methods can be found in 3.3.2), the P_{II} conformer with a population probability of 46% without thermal corrections and 59% with thermal corrections at 1 bar and α_R with a population probability of 42% and 33%, respectively (see Table 20). Thus, practically only these two conformers are present at ambient pressure in an aqueous solution. Similar results are obtained for the MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/EC-RISM with thermal corrections, whereby the P_{II} conformer has a probability of 64.75% and the α_R of 28%. Here, it is important to mention that the added thermal corrections are the same B3LYP/6-311+G(d,p)/PCM since no frequency calculations are available for EC-RISM calculations. The next most favored states are the C_5 and the C_{7eq} conformations, with 6 % and 4 % probability at 1 bar without thermal corrections, respectively. Only marginal populations show the conformations α_L, a_d, and $C_{7,\,ax}$. The calculated results are in good agreement with the Raman skeletal stretching measurement of Takekijo[69] and Grdadolnik[303]. However, compared to the combined IR and Raman amide III experiments from Grdadolnik[64], the population of the α_R is overestimated, and of C_5 is underestimated.

On the other hand, three distinct minima can be found for B3LYP/6-311+G(d,p)/PCM//B3LYP/6-311+G(d,p)/PCM at ambient pressure (Table 20). The α_R conformer is the most preferred conformer with a population probability of 37 %. In contrast to EC-RISM, the ranking of the conformers P_{II} and C_5 is reversed in the PCM calculations. The C_5 conformer has a population probability of 35 % and the P_{II} conformer of 18 %. The C_{7eq} conformer also has a noteworthy population probability of 8.8 %. These results agree with a recent study on the conformational behavior of Ac-Ala-NHMe, published by Cornaich et al.[66]. In this study, the theoretical level of theory B3LYP-D3/aug-cc-pVDZ was used with the IEF-PCM solvent model resulting in population probabilities of 16.4% for the P_{II} conformer, 32 % for the α_R conformer, and 36.8% for the C_5 conformer. The resulting differences between PCMs and EC-RISM may be best explained by the lack of resembling hydrogen bonds within PCMs. In a previous NMR study,[305], it was shown that the P_{II} conformer is best able to reflect experimental dipolar couplings, while α_R conformations may also have substantial populations (or roughly 1 kcal/mol higher energy). The main conclusions from this study were that the P_{II} conformer is the most favorable conformer due to the hydrogen bond pattern with the surrounding solvent molecules. Several past theoretical calculations support these results, including CHARMM22 and TIP3P[306] calculations and density functional theory calculations with explicit solvent molecules.[301]

Table 20: Overview of populations and free energies of the seven and eight minima of Ac-Ala-NHMe and Ac-Gly-NHMe, respectively. Experimental and theoretical reference values from the literature are given. EC-RISM calculations were performed with MP2/6-311+G(d,p) on B3LYP/6-311+G(d,p)/PCM- or B3LYP/6-311+G(d,p)/EC-RISM-optimized structures. The pressure-dependent correction of the excess chemical potential was used. PMV signals that ΔG was solely computed by eq (*185*). Raw and further calculated data can be found in SI_4.3. S. Maste performed the NMR fits.[300] Metadynamics were performed by C. Hölzl.[307]

Reference data									Opt: B3LYP/6-311+G(,p)/PCM:												Opt: B3LYP/6-311+G(,p)/EC-RISM:									
Ac-Ala-NHMe										w. o. Tcorr			w. Tcorr			PMV w. Tcorr			NMR fit[300]		w. Tcorr			PMV w. Tcorr			NMR fit[300]		Metadynamics	
Conf	φ/ °.	ψ / °	PCM[66]	Av[303]	IR[303]	RS[303]	RS[69]	IR/RS[64]	PCM	1 bar	1 kbar	5 kbar	1 bar‡	1 kbar‡	5 kbar‡	1 bar‡	1 kbar‡	5 kbar‡	1 bar	3 kbar	1 bar‡	1 kbar‡	5 kbar‡	1 bar‡	1 kbar‡	5 kbar‡	1 bar	3 kbar	1 bar	10kbar
P_{II}	-76	143	16.8	68	60	76	69	60	18.11	45.71	45.51	44.74	58.95	58.46	57.81	58.95	58.64	57.47	65.04	63.07	64.75	65.81	65.27	64.75	64.40	63.27	51.04	48.74	75.90	37.99
α_R	-90	-12	32.0	15	11	18	17	11	37.12	41.52	42.25	43.90	32.76	33.60	34.70	32.76	33.18	34.57	34.96	36.93	27.34	26.81	27.77	27.34	27.78	29.12	48.96	51.26	14.83	28.15
C_5	-152	155	36.8	17	29	6	6	29	35.39	5.59	5.41	5.02	5.91	5.68	5.31	5.91	5.85	5.74	-	-	5.84	5.45	5.05	5.84	5.80	5.69	-	-	8.31	27.75
$C_{7,eq}$	-85	71	-	-	-	-	8	-	8.82	4.22	3.86	3.21	1.19	1.06	0.90	1.19	1.14	1.02	-	-	0.91	0.79	0.68	0.91	0.86	0.74	-	-	-	-
α_L	64	33	0.5	-	-	-	-	-	0.34	2.00	2.01	2.11	0.91	0.91	0.96	0.91	0.91	0.91	-	-	0.82	0.79	0.84	0.82	0.82	0.82	-	-	0.82	5.58
a_d	59	-140	-	-	-	-	-	-	0.04	0.75	0.78	0.88	0.25	0.26	0.29	0.25	0.25	0.26	-	-	0.32	0.32	0.37	0.32	0.33	0.34	-	-	0.13	0.53
$C_{7,ax}$	73	-53	-	-	-	-	-	-	0.16	0.21	0.19	0.15	0.03	0.03	0.02	0.03	0.03	0.02	-	-	0.02	0.02	0.01	0.02	0.02	0.02	-	-	-	-
ΔG / kcal mol^{-1}																														
P_{II}	-76	143	0.48	0.0	0.0	0.0	0.0	0.0	0.0	0,0	0.0	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.0
α_R	-90	-12	0.08	0.90	1.00	0.85	0.83	1.00	-0.42	0.06	0.04	0.01	0.35	0.34	0.30	0.35	0.34	0.30	0.37	0.32	0.51	0.53	0.51	0.51	0.50	0.46	0.02	-0.03	0.97	0.18
C_5	-152	155	0.00	0.82	0.43	1.50	1.44	0.43	-0.40	1.24	1.26	1.30	1.36	1.38	1.41	1.36	1.37	1.37	-	-	1.43	1.48	1.52	1.43	1.43	1.43	-	-	1.31	0.19
$C_{7,eq}$	-85	71	-	-	-	-	1.28	-	0.43	1.41	1.46	1.56	2.31	2.36	2.46	2.31	2.34	2.39	-	-	2.53	2.62	2.70	2.53	2.56	2.63	-	-	-	-
α_L	64	33	2.55	-	-	-	-	-	2.34	1.85	1.85	1.81	2.47	2.47	2.43	2.47	2.47	2.45	-	-	2.59	2.62	2.58	2.59	2.59	2.58	-	-	2.68	1.14
a_d	59	-140	-	-	-	-	-	-	3.63	2.43	2.41	2.33	3.23	3.21	3.13	3.23	3.22	3.19	-	-	3.14	3.15	3.06	3.14	3.13	3.09	-	-	3.79	2.53
$C_{7,ax}$	73	-53	-	-	-	-	-	-	2.79	3.19	3.25	3.38	4.51	4.57	4.70	4.51	4.54	4.62	-	-	4.70	4.81	4.97	4.70	4.74	4.85	-	-	-	-
Ac-Gly-NHMe																														
Conf	φ/ °.	ψ / °	PCM[67]	--	-	-	--	Exp[64]	PCM	1 bar	1 kbar	5 kbar	1 bar‡	1 kbar‡	5kbar ‡	1 bar‡	1 kbar‡	5 kbar‡	1 bar	3 kbar	1 bar‡	1 kbar‡	5kbar ‡	1 bar‡	1 kbar‡	5 kbar‡	1 bar	3 kbar	1 bar	10kbar
α_R	-98	-1	54.1	--	-	-	-	66	24.42	29.72	30.07	30.32	31.40	31.57	31.49	31.40	31.43	31.34	1.53	0.00	29.83	29.72	30.50	29.83	29.91	30.04	36.59	40.78	13.76	14.84
α_L	98	1	-	--	-	-	-	-	22.96	28.06	28.44	28.82	27.92	28.11	28.18	27.92	27.97	27.99	-	-	26.65	27.98	26.95	26.65	26.76	26.96	-	-	5.94	7.95
P_{II}	78	-154	-	--	-	-	-	22	4.22	15.24	15.22	15.47	13.24	13.13	13.20	13.85	13.81	13.82	98.47	100	16.22	15.67	15.90	16.22	16.07	15.66	63.41	59.22	25.71	20.50
P_{II}	-78	158	-	-	-	-	-	-	4.42	15.23	15.20	15.43	13.85	13.73	13.79	13.24	13.20	13.21	-	-	15.67	15.40	15.85	15.67	15.53	15.17	-	-	48.25	49.76
β	98	164	-	--	-	-	-	12	7.2	5.43	5.54	5.82	10.43	10.57	10.97	10.43	10.53	10.86	-	-	8.23	8.06	8.10	8.23	8.45	9.22	-	-	12.07	6.94
C_5	179	-179	37.6	--	-	-	-	-	33.31	0.96	0.90	0.80	2.04	1.91	1.69	2.04	2.00	1.90	-	-	2.41	2.30	2.03	2.41	2.35	2.21	-	-	-	-
$C_{7,eq}$	-83	61	5.6	-	-	-	-	-	1.71	2.69	2.31	1.66	0.56	0.48	0.34	0.56	0.53	0.44	-	-	0.50	0.44	0.34	0.50	0.47	0.38	-	-	-	-
$C_{7,eq}$	83	-61	-	-	-	-	-	-	1.71	2.68	2.31	1.67	0.56	0.48	0.35	0.56	0.52	0.43	-	-	0.49	0.43	0.33	0.49	0.45	0.37	-	-	-	-
ΔG / kcal mol^{-1}																														
α_R	-98	-1	0.0	-	-	-	-	-	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.0	0.0	0.0	0.0
α_L	98	1	-	-	-	-	-	-	0.04	0.03	0.03	0.03	0.07	0.07	0.07	0.07	0.07	0.07	-	-	0.07	0.04	0.07	0.07	0.07	0.06	-	-	0.49	0.37
P_{II}	78	-154	-	-	-	-	-	-	1.04	0.40	0.40	0.40	0.49	0.49	0.49	0.51	0.51	0.51	-2.46	-8,74	0.36	0.38	0.39	0.36	0.37	0.39	-0.33	-0.22	-0.37	-0.19
P_{II}	-78	158	-	-	-	-	-	-	1.01	0.40	0.40	0.40	0.51	0.52	0.51	0.49	0.49	0.48	-	-	0.38	0.39	0.39	0.38	0.39	0.40	-	-	-0.74	-0.72
β	98	164	-	-	-	-	-	-	0.73	1.01	1.00	0.98	0.65	0.65	0.62	0.65	0.65	0.63	-	-	0.76	0.77	0.79	0.76	0.75	0.70	-	-	0.46	0.45
C_5	179	-179	0.22	-	-	-	-	-	-1.18	2.04	2.08	2.15	1.62	1.66	1.73	1.62	1.63	1.66	-	-	1.49	1.52	1.61	1.49	1.51	1.55	-	-	-	-
$C_{7,eq}$	-83	61	1.35	-	-	-	-	-	1.57	1.43	1.52	1.72	2.39	2.48	2.68	2.38	2.42	2.53	-	-	2.42	2.50	2.67	2.42	2.46	2.59	-	-		
$C_{7,eq}$	83	-61	-	-	-	-	-	-	1.57	1.42	1.52	1.72	2.38	2.48	2.67	2.39	2.43	2.53	-	-	2.43	2.51	2.68	2.43	2.48	2.61	-	-	-	-

The relative pressure-dependent change of the minimum occupation probabilities of Ac-Ala-NHMe is minimal, although the order of the occupation probabilities changes slightly for PCM-optimized structures when no thermal corrections are added. A steady decrease in the probabilities can be observed for the P_{II} conformer, whereas, for the α_R conformer, a steady increase is observable. For the EC-RISM optimized structures with thermal corrections, the corresponding populations are 65.27% and 27.77% for the P_{II} and α_R conformers at 5 kbar. In contrast to the PCM-optimized structures, the pressure-dependent changes for the EC-RISM-optimized structures are not continuous. This fact indicates that the optimized B3LYP/6-3111+G(d,p) structures are not necessarily the same as possible MP2/6-311+G(d,p) optimizations would provide. Currently, geometry optimizations with EC-RISM and MP2/6-311+G(d,p) are not feasible, which could remedy these artifacts.

Nonetheless, the order of magnitude of the pressure-dependent tendencies is similar to the results for the PCM-optimized structures. Due to pressurization, the relative occupation of the α-conformer increases compared to the P_{II} conformer. The conformers C_5 and α_L also show a decreasing trend under high pressure, whereas the population probability of the C_{7eq} conformer increases slightly but is also far from being significantly occupied. To compare these results with already published ones, the Takeiyo et al. study[69,70] will be focused at. The results obtained here contradict the EC-RISM results that the α_R conformer is more likely to be occupied with higher pressure.

Table 21 Partial molar volumes of the four different conformers of Ac-Ala-NHMe and Ac-Gly-NHMe, taken from Ref[69,70] and calculated with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM at 1 bar and 10 kbar. All values are in cm³/mol. Raw data for all conformers can be found in SI_4.3.

Conf	ΔV_m(Raman) [69,70]	V_m (1D RISM) [69,70]	ΔV_m(1D RISM)	V_m (EC.RISM 1 bar)	ΔV_m (EC.RISM 1 bar)	V_m (EC-RISM 10 kbar)	ΔV_m (EC.RISM 10 kbar
			Ac-Ala-NHMe				
α_R	1.1	118.5	-1.2	116.47	-0.46	109.58	-0.29
P_{II}	0	119.7	0.0	116.93	0.0	109.87	0
C_5	0.1	117.9	-1.8	117.06	0.13	109.76	-0.11
$C_{7,eq}$	-2.4	116.3	-2.4	118.04	1.11	110.05	0.18
			Ac-Gly-NHMe				
α	-0.4	110.4	0.4	99.05	-0.13	94.68	0.14
P_{II}	0.0	110.0	0.0	99.18	0.0	94.54	0.0
C_5	0.5	110.1	0.1	99.71	0.53	94.62	0.06
$C_{7,eq}$	-2.2	108.0	-2.0	100.89	1.61	95.21	0.65

As mentioned above, the peak assigned to the α_R conformer decreases upon pressurization and vanishes at 1.6 kbar. It is observed that another peak grows under pressure, and the C_{7eq} conformer was assigned to it. These statements were supported by theoretical calculations of the partial molar volume of the respective conformers using the RISM-KB (1D RISM) theory with the amber ff99SB force field. The corresponding values are shown in Table 21, and the calculated PMVs from EC-RISM calculations at ambient pressure and high-pressure conditions.

Takeyo et al. showed that the $C_{7,eq}$ conformer has the smallest partial molar volume with 116.3 cm^3/mol, whereas the P_{II} conformer has the largest partial molar volume with 119.7 cm^3/mol. These results do not fit the experimental observation, whereby the P_{II} conformer has the second smallest PMV. In the experiment, the α_R has a higher PMV than the P_{II}; however, these results can be neither confirmed with 1D RISM nor EC-RISM.

Furthermore, these values do not match the PMVs with those calculated by EC-RISM, which is expected to be more accurate. The α_R conformer has the lowest partial molar volume with 110.59 cm^3/mol. To measure the experimental volume differences, a pressure-independent PMV was assumed, but the EC-RISM model includes the changes in the PMV upon pressurization. Over the whole pressure range, the α_R conformer has the smallest partial molar volume and hence should be preferable under high pressure. In this thesis, the PMVs were calculated with 3D RISM, whereas Takekiyo used 1D RISM. Another possible reason for the deviations could be in the underlying conformers, e.g., the Takekiyo structures were taken from average results from different simulations. For example, the φ,ψ values were always rounded to the tenth. It is unclear if the structures were optimized and if the optimizations were made in a vacuum or solvent. This procedure could lead to huge differences if unoptimized structures were used for the calculations.

Another way to determine the population of the conformers is to make a fit in a way that the fitted populations are multiplied with the theoretically calculated chemical shifts of the corresponding conformers to represent the experimentally measured chemical shifts accurately. S. Maste performed the corresponding fits, and the focus was only on the two main conformers P_{II} and α_R. For the former, in the case of the PCM-optimized structures, a population of 65.04% is obtained, and the α_R has a population of 34.96%. This result agrees with the literature data that the P_{II} is the main conformer. Upon pressurization, only marginal changes occur; nonetheless, the same trend is observed for the single point calculations, that the population of α_R increases. Similar results are obtained for the EC-RISM optimized structures, whereby the two conformers are nearly identical, populated with 51.04% and 48.96% for P_{II} and α_R, respectively. At a pressure of 3 kbar, the ratio changes to 51.26 for α_R and 48.96% for P_{II}. This outcome can be attributed to the fact that in Figure 29 C, the α_R can better resemble the experimental observation, whereas, in Figure 29 D, the P_{II} conformer is closer to the experiment. Thus, it seems reasonable that the fit process tries to balance the two conformers. The pressure-dependent NMR fit results also contradict the pressure-dependent Raman measurements.[69,70]

Although Ac-Gly-NHMe has one methyl group less than Ac-Ala-NHMe, the conformational landscape is more diverse (see Figure 33). Due to the achiral structure of Ac-Gly-NHMe, a symmetry plane can be observed with respect to the inversion of the dihedral angles φ and ψ. The two main conformations obtained by EC-RISM on the PCM-optimized structures without the addition of thermal corrections are the α-helical right and left-directed structures, with occupation probabilities of 30% and 28%, respectively. In contrast to Ac-Ala-NHMe, the effect of adding thermal corrections is marginal on the population. The next preferred conformers can be assigned to the P_{II} or beta-sheet domain, with about a 15% probability. The C_{7eq} conformations have only marginal probabilities of around 2.6%. The conformer with the lowest probability by far is the C_5 conformer, with only a 0.9% probability. Furthermore, in Table 20, the probabilities for the eight minima of Ac-Gly-NHMe optimized with B3LYP/6-311+G(d,p)/EC-RISM are presented. The relative order of the minima for the MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM is identical to the PCM optimized structures. The differences between the PCM-optimized populations and the EC-RISM ones are much smaller than the results for Ac-Ala-NHMe. The populations are The relative occupation number of α conformers increases upon pressurization. However, a steady trend is only observed for the B3LYP/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM calculations (see Table 32), whereas the MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM calculations do not show a steady trend, again indicating, that the B3LYP/6-311+G(d,p)/EC-RISM optimized structures are not the ideal ones for pressure-dependent MP2/-6311+G(d,p) calculations. Due to the EC-RISM geometry optimizations, the relative population of the P_{II} conformers at ambient conditions increases, which was already observed for Ac-Ala-NHMe.

Compared to Ac-Ala-NHMe, experimental reference populations are sparse; however, the EC-RISM results can be compared to the combined amide III IR and Raman measurements of Grdadolnik[64]. Here, a population of 66% was determined for the α-helical region, which agrees very well with the EC-RISM calculated values, which are roughly 60% occupation probability combined for both α conformation. In the experiment, 22% and 12% were determined for the P_{II} and β, respectively. With EC-RISM, the calculated populations are 27% for the two P_{II} conformers and 11% for β.

In the B3LYP/6-311+G(d,p)/PCM//B3LYP/6-311+G(d,p)/PCM calculations, the C_5 conformer with 33% is the conformer with the lowest energy. The two α-helical structures have probabilities of about 23 %. According to PCM, the P_{II} conformations are much less likely to be occupied, with only a 4.2% probability. The two C_{7eq} conformations are also energetically higher in the PCM model with a 1.73% occupation probability. A comparison with a publication by Cormanich et al. can be made in order to assess the PCM results in concrete terms.[67]. This publication found three main conformers for Ac-Gly-NHMe, calculated with the level of theory B3LYP-D3/aug-cc-pVDZ and the IEF-PCM solvent model. The main conformer is in the α-helical region with a population of 51.1 %, although it should be noted that only one helix conformer was evaluated here. The other two conformers are $C_{7,\,eq}$ with 5.6 %, and C_5 with 37.6 %. The results for the C_5 and α conformers agree with the PCM results.

Concerning the pressure-dependent changes in the population probabilities of the main conformers of the Ac-Gly-NHMe, no relevant change can be observed. In contrast to Ac-Ala-NHMe, where a slight increase of the $C_{7\,eq}$ conformer was observed, this conformer becomes less favored in the case of Ac-Gly-NHMe under pressure. These results contradict the study of Takekijo[70], where again, a combination of Raman measurements and 1D RISM theory calculations was used as the method of choice. The corresponding PMVs between the conformers can be found in Table 21. In the experiment, the $C_{7\,eq}$ conformer also has the lowest partial molar volume, supported by 1D RISM calculations. Furthermore, the order between the α_R and the P_{II} conformer is reversed, with the α_R conformer having a smaller volume and, thus, will be more populated upon pressurization. This result agrees with the EC-RISM calculations on the PCM-optimized structures, where a stronger population of the α conformers is observed (see Table 21). In the EC-RISM calculations, the $C_{7,eq}$ conformer has the largest partial molar volume, whereas α and P_{II} conformers have the smallest partial molar volume at ambient pressure. Again a possible explanation for the large differences between 1D RISM and EC-RISM can be in the underlying structures that were used for the PMV calculations.

The NMR-fit-based populations (see Table 20) significantly differ from the experimental reference results and the EC-RISM calculations for the PCM-optimized structures. Here, P_{II} is clearly the main conformer, with a population of 98.4%, and the α conformer is only occupied with a 1.6% probability. Upon pressurization, the result looks even worse; at 3 kbar, the P_{II} is effectively the only conformer populated. This effect is due to the fit procedure, where only the two ^1H amide protons were used. In Figure 29, panels G and H, we can see that for both amide nuclei, the P_{II} conformer is closer to the experiment; therefore, it is plausible that only this conformer is probably occupied. The picture looks considerably better when the EC-RISM optimized structures are set into the focus. The P_{II} conformer is still the main conformer with 63.41%, and a decrease due to pressure is observed, with a probability of 59.22% at 3 kbar. This pressure-dependence change agrees with the pure EC-RISM calculations and the observation from Takekiyo.[70]

So far, only the effect of pressure on the minima was discussed; however, pressure can affect the complete so-called Ramachandran plane. To discuss this, the complete Ramachandran plane of the dipeptide Ac-Gly-NHMe (different Ramchandran planes of Ac-Ala-NHMe can be found in Figure 57 and Figure 58) should be extensively investigated with EC-RISM, FF-RISM, metadynamics- and AIMD simulations. For the EC-RISM calculations, the level of theories MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM and RPBE-D3/TZVPP//B3LYP/6-311+G(d,p) were used. The latter was used to get an appropriate comparable basis to the AIMD simulations. These were performed by S. Körning[308] and used the RPBE-D3 functional with the TZVPP basis set. Additionally, in the appendix (see Figure 56), DLPNO-CCSD/def2-TZVPP//B3LYP/6-311+G(d,p)/PCM calculations are presented. For the latter two methods, the pressure-dependent PMV correction originally developed for MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM was used (parameters originally published in ref [27]). For the FF-RISM and metadynamics simulations, the ff14SB force field was utilized, whereby, in the case of the metadynamics simulations, the SPC/E water force field was chosen. C. Hölzl[307] performed these simulations. This chapter focuses on the pressure 1 bar and 10 kbar only. The results for two further force fields, the CHARMM36m, and the ff19SB, can be found in the appendix (see Figure 56). To get an impression of the effect of high pressure on the geometries, B3LYP/6-311+G(d,p)/EC-RISM geometry optimizations were performed, and MP2/6-311+G(d,p) singlepoint calculations were made for these structures.

Figure 34 depicts the Ramachandran free energy surfaces for Ac-Gly-NHMe calculated with various methods at 1 bar and 10 kbar. The EC-RISM calculations at ambient conditions are shown in panels A and C, and the biggest differences occur in the region around the minima, where the MP2/6-311+G(d,p) calculations show a narrower picture compared to the DFT calculations. Larger differences can be observed when the QC calculations are compared with the force-field-based (here, the calculations were performed with a combination of the ff14SB amber force field and 3D RISM) ones. On the one side, the EC-RISM calculations show relative broad minima in the alpha-helical region (around ϕ: ±100,ψ:±0) and P_{II} region (ϕ: ±100,ψ:±0), whereas, on the other side, the force field-based computations show significant narrower minima.

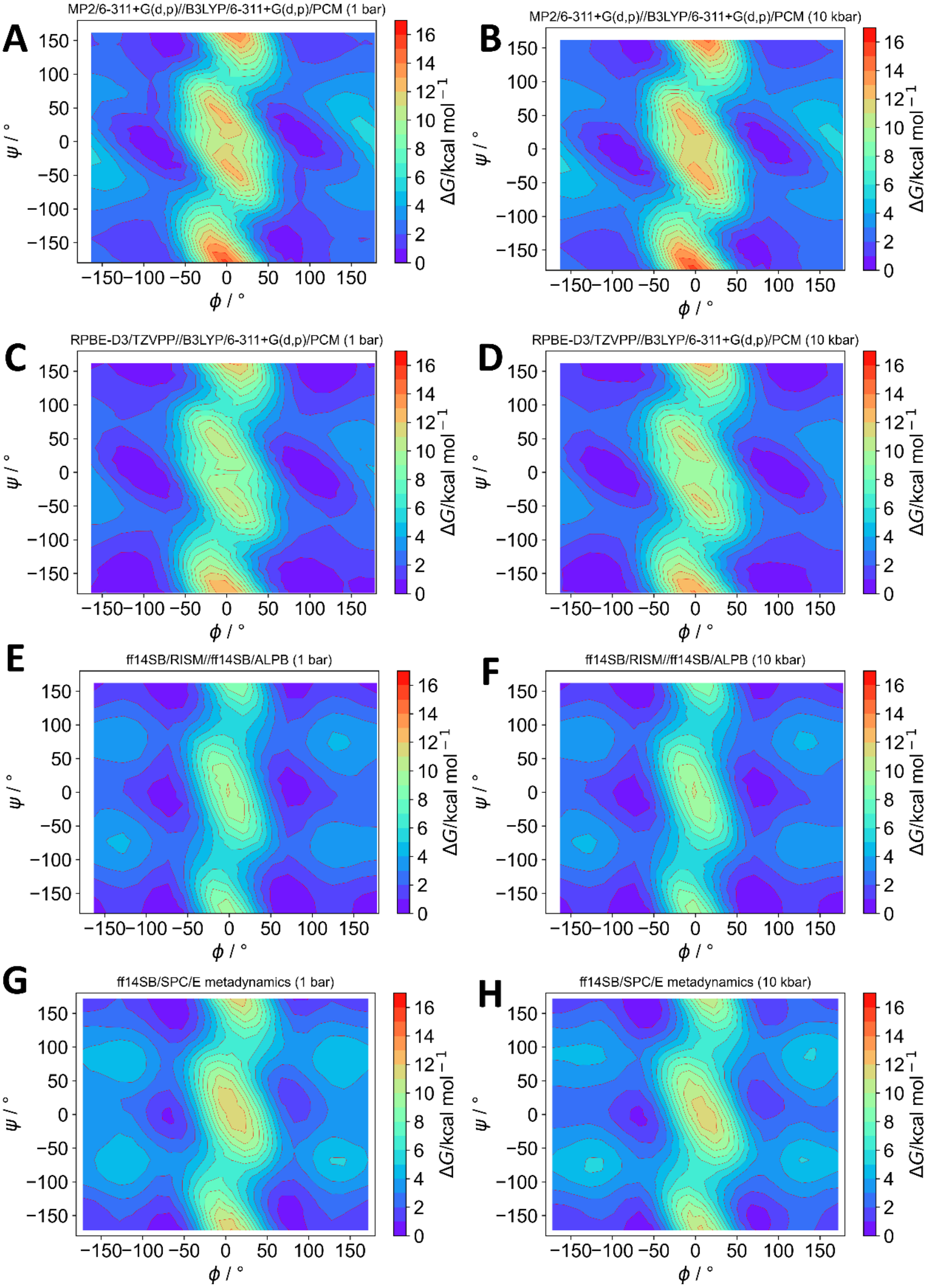

Figure 34: Pressure-dependent Ramachandran plane of Ac-Gly-NHMe. Panel A and B were calculated using MP2/6-311+G(d,p)//B3LYP-63111+G(d,p)/PCM at 1 bar (already shown in Figure 34 B) and 10 kbar, respectively. Panel C and D show the RPBE-D3/TZVPP//B3LYP/6-311+G(d,p) calculations at 1 bar and 10 kbar, respectively. The results obtained from the ff14SB/RISM calculations are presented in panels E (1 bar) and F (10 kbar). Panels G and H show the results obtained from the metadynamics simulations using ff14SB with SPC/E. The color scale ranges from 0 kcal/mol to 16 kcal/mol, and the iso levels are 1 kcal/mol. Raw and calculated data can be found in SI_4.3. Metadynamics simulations were performed by C. Hölzl. [307]

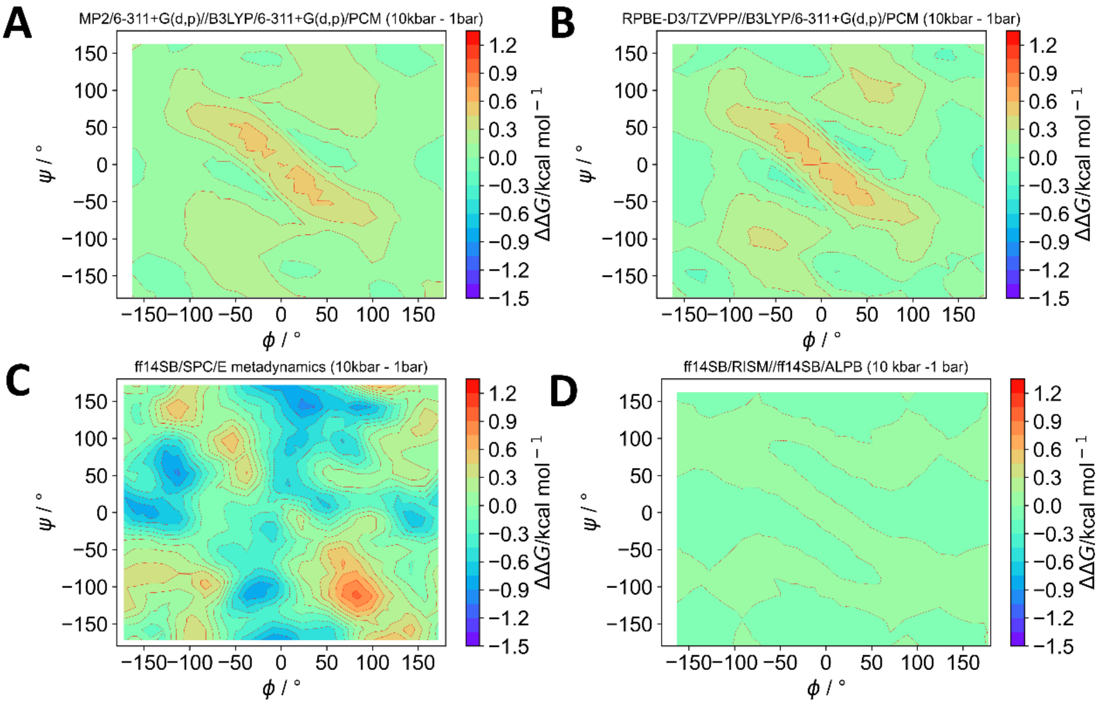

Figure 35 Difference free energy Ramachandran surfaces between 10 kbar and 1 bar of Ac-Gly-NHMe calculated with different methods. Panel A shows the free energy difference calculated with MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p). In panel B, the corresponding RPBE-D3/TZVPP calculations are depicted. C shows the ff14SB- based metadynamics results with SPCE/E performed by C. Hölzl. D shows the ff14SB + RISM calculations. Raw data can be found in SI_4.3

In Figure 35, the relative free energy change between 10 kbar and 1 bar for the four methods is presented. The two EC-RISM-based approaches are shown in panels A and B, and no significant difference is visible. The changes are in the range of -0.2 to 0.5 kcal/mol, whereby the biggest changes do not occur in the minima region, but instead, the largest changes can be observed between the minima. Nonetheless, the absolute numbers and the effect of pressure on the solvated free energy are relatively small. For the ff14SB+RISM combination, the differences are significantly smaller, ranging from -0.1 to 0.1 kcal/mol. This result is similar to the results obtained for the *cis-trans* equilibrium of NMA. The pressure-dependent changes in the metadynamics simulation (Panel (C)) are much more diverse. However, it should be noted that explicit waters (SPC/E) were present in the metadynamics simulation, and these affect the geometries even if identical torsion angle combinations are considered. In metadynamics simulations, the effect of orthogonal degrees of freedom, like rotational and vibrational modes, is considered. On the other hand, in the EC-RISM calculations, these contributions were only optimized. Furthermore, statistical errors can occur in the simulation if some regions are not long enough sampled. Nevertheless, the pressure-dependent changes are still relatively small and in a similar order of magnitude as the results for the EC-RISM calculations.

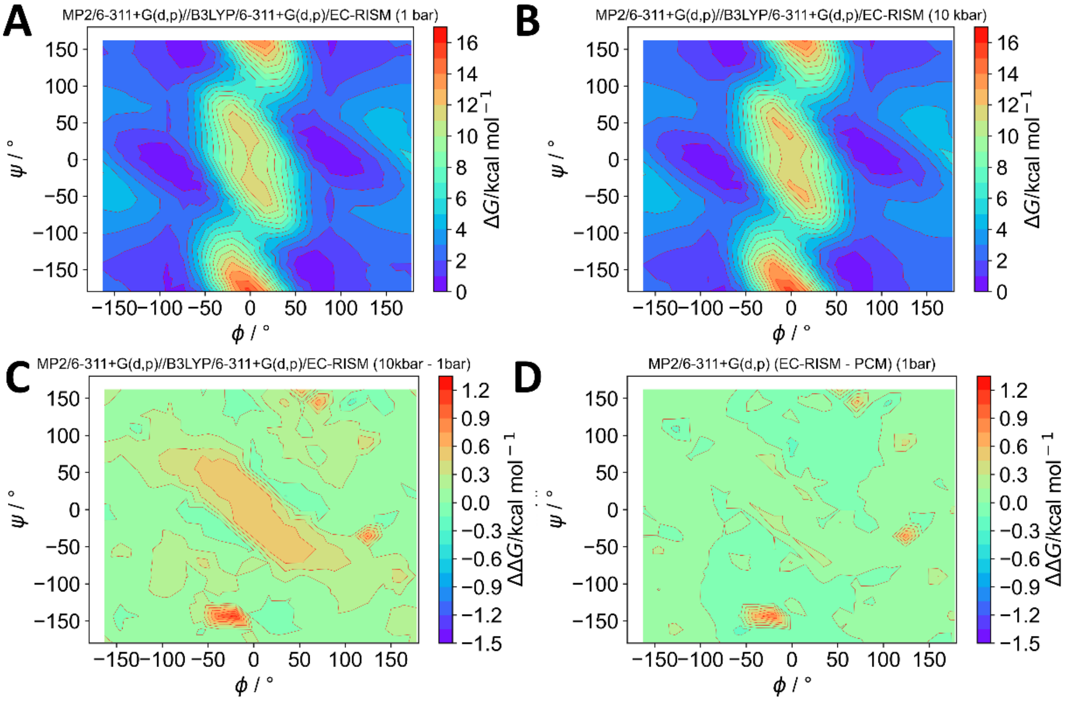

Figure 36: Panels A and B show the Ramchandran surfaces of Ac-Gly-NHMe at 1 bar and 10 kbar, respectively. These calculations were done with MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/EC-RISM(p). In Panel C, the difference between panels B and A is depicted. Panel D shows the difference between Panel C of this figure and Panel A of Figure 35 (Pressure difference between 10 kbar and 1 bar calculated with MP2/6-311+G(d,p). Raw data can be found in SI_4.3.

In the case of EC-RISM, only calculations on rigid 1 bar geometries based on B3LYP/6-311+G(d,p)/PCM optimization have been shown so far. In Figure 36, the results from MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM calculations are shown (Panel (A) and (B)). The corresponding difference between the two surfaces is depicted in Panel (C). In Panel (D). the changes between Panel (A) of Figure 35 and Panel (D) are depicted.

The difference between the calculations on the PCM-optimized structures and pressure-dependently optimized EC-RISM ones is very small. With the bare eye, no significant differences can be observed for the surfaces at ambient and high-pressure conditions, respectively.

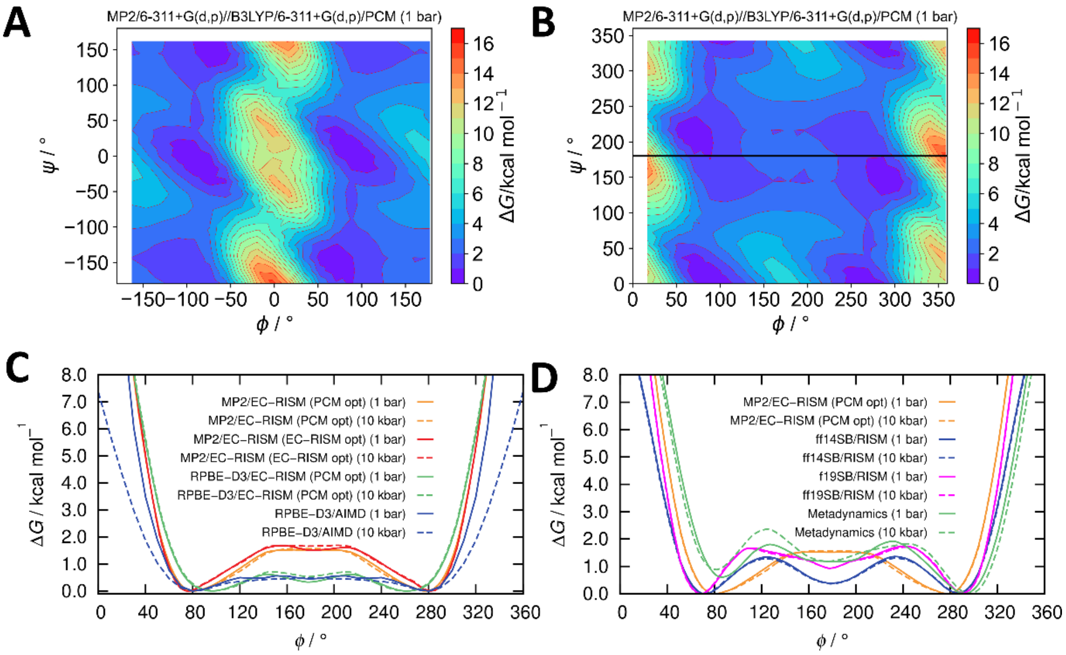

Figure 37: Panel A shows the original Ramachandran surface from Panel A of Figure 34 calculated with MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM. Panel B shows the identical surface; however, to the negative φ and ψ the number 360 were added. In this manner the huge barriers are moved at the edges of the surface. In Panels C and D, a specific slice (at ψ=180°, see black line in panel B) is shown. In Panel C the QC-based calculations are shown, where the orange lines refer to MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM calculations at 1 bar (solid) and 10 kbar (dashed). MP2/-6311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM(p) results are depicted through the red lines (1 bar:solid, 10 kbar:dashed). The green lines show RPBE-D3/TZVPP/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM calculations, and the blue lines show data extracted from multiple AIMD simulations using RPBE-D3/TZVPP at 1 bar (solid) and 10 (kbar), respectively (performed by S. Körning)[308]. In Panel D, the force-field-based methods are summarized. The orange line is the same as Panel C and should provide some orientation. The blue lines were calculated with the ff14SB+RISM combination. The ff19SB force field with RISM solvation was utilized for the magenta lines. The green lines show the results obtained from the ff14SB/SPCE/E metadynamics simulation.[307] Raw data can be found in SI_4.3.

It is difficult to find significant differences between the methods by looking alone at the surfaces. Of course, it can be seen that the force field-based surfaces differ significantly from the QC-based surfaces. In the following, one specific slice through the Ramachandran plane should be picked out and analyzed to get a better look. The slice of choice is located at a constant ψ dihedral value of 180°, and the results are summarized in Figure 37. Here, Panels (A) and (B) show the Ramchandran surface calculated with MP2/6-311+G(d,p) in its normal form (Panel (A)) and in a form where the negative φ and ψ were shifted into the positive range, by adding a value of 360°. This change is made to move the high barriers within the hyperplane to the edges of the slice. In Panel (C), the results for QC-based methods are shown. On the one hand, there are the MP2/6-311+G(d,p)/EC-RISM calculations based on the PCM and EC-RISM optimized minima, respectively. And on the other hand, DFT-based calculations using the RPBE-D3/TZVPP level of theory for EC-RISM calculations and an AIMD simulation.

For all QC calculations, the local minima are located at 80° and 280°, and between these two minima, a barrier is observed, with a plateau between 130°

and 220°. The main difference between the RPBE-D3/TZVPP and the MP2/6-311+G(d,p) calculations lies in the height of the barrier, which is for the former around 1 kcal mol^{1-}and for the latter one twice as high. However, the pressure-dependent changes from ambient conditions to extremely high pressure are very small for all methods. In the case of the EC-RISM calculations, rigid structures were used, and orthogonal degrees of freedom (like vibrational and rotational contributions) were neglected, whereas, in the case of the AIMD simulation, the degrees of freedom were taken into account. Since AIMD simulations are relatively time-consuming and the complete free energy surface can not be sampled, multiple AIMD simulations were done and subsequently combined. The AIMD simulations were started from the minima positions, and in the end, a histogram was generated, from which the free energy changes along the slice were calculated.

The force-field-based methods are shown in Panel (D) of Figure 47. Similar to the QC results, two local minima are present at 80° and 280°. However, in the force field, a third local minimum is observed in the region of 180°. The pressure dependence is marginal due to the missing polarization for the ff14SB+RISM and the ff19SB+RISM variants. Nonetheless, the error introduced due to the force-field design is significantly larger than the small error observed in the pressure dependence. In the metadynamics simulation, larger differences are observed, and the clear lowest minimum is located at 280°, and no symmetric curvature along the slice is observed. Also, the effect of pressure is larger here. This effect can be seen from the calculated populations for the main minima (see Table 20), which were determined by interpolating the Ramachandran plane and subsequently by performing minimizations in the proximity of the PCM-optimized φ and ψ values. For example, the population of the P_{II} of Ac-Ala-NHMe is 75.9% at 1 bar and drops to 38% at 10 kbar. The effect is also clearly visible in the difference free energy Ramachandran plane (see Figure 58) This result is in accordance with the EC-RISM results but far too strong. An explanation for this could be a lack of sufficient sampling in the corresponding regions.

In summary, for both Ac-Ala-NHMe and Ac-Gly-NHMe, the minima calculated by EC-RISM are consistent with the consensus opinion from the literature at ambient pressure. The P_{II} conformation is the main conformer for Ac-Ala-NHMe, whereas the C_5 and α_R are also significantly populated. In the case of EC-RISM, α_R is higher populated, which agrees with experimental Raman studies for Ac-Ala-NHMe. However, the experimental references show no clear picture of whether the α_R or the C_5 is the higher populated one. For Ac-Gly-NHMe, EC-RISM can predict the ratios obtained in experiments at ambient conditions. The pressure dependence of the populations is marginal, whereby high pressure favors α_R conformation due to the smaller PMV for both dipeptides. However, in the literature, a pressure-dependent Raman study for Ac-Ala-NHMe showed a strong decrease in the α_R conformer population. However, a similar study of the pressure dependence of Ac-Gly-NHMe showed an increase in αR conformer populations due to pressurization.

Furthermore, it can be stated that with EC-RISM and AIMD, the calculated effect of pressure on the Ramachandran planes is marginal, even in a 10 kbar regime. Slight changes are observed, which are mostly located in the barrier regions between the minima and not directly at the different minimum conformations. With a combination of force field + RISM solvation, the correct pressure-dependent tendency can be reproduced, whereby the effects are one magnitude smaller due to the missing polarization of intramolecular interactions. Metadynamics simulations show larger pressure-dependent changes, possibly due to statistical errors.

4.3.6 Summary and outlook

This chapter aimed to investigate the influence of pressure on three typical protein backbone models, NMA, Ac-Gly-NHMe, and Ac-Ala-NHMe. The first two sections were dedicated to pressure-dependent NMR calculations on the amide nuclei of the three model systems. For all systems, experimental reference data were available. The absolute deviations of the calculated values were the smallest for the amide protons, and the deviations for the ^{13}C and ^{15}N were larger. An inconsistent picture is observed regarding the pressure-dependent changes for the ^{1}H nuclei of NMA. On the one hand, the pressure dependence for *trans*-NMA is overestimated, whereas, on the other hand, the pressure-dependent increase of all other ^{1}H amide nuclei (including the two dipeptides) is underestimated. Here, a possible dimerization of *trans*-NMA could be the problem. Nonetheless, for all nuclei, EC-RISM could reproduce the correct pressure-trend.

Regarding the pressure-dependent EC-RISM optimizations, the main effects on the geometries are an elongating of the C-O and N-H amide bonds, with changes of less than 0.001 Å up to a pressure of 5 kbar. However, the effect of these small changes can be observed in the NMR calculations. The EC-RISM optimized structures better represent the pressure-dependent changes for the 1H chemical shifts. Also, the absolute values at ambient conditions are better resolved; however, the pressure-dependent shielding constants of the PCM-optimized DSS at 1 bar were used for reference.

Next, the relative population changes of the two main conformers (*cis* and *trans*) of NMA were compared to NMR-based experimental reference data. A very good agreement between experiment and theory is observed, with a slight increase in the *cis* conformer's relative population. Additionally, a version of the amber force field was combined with pressure-dependent 3D RISM calculations. These calculations showed that the error introduced by the force field at ambient conditions is significantly larger than the error produced due to the missing pressure-dependent polarization.

In the last part, the main conformers (which are based on PCM geometry optimizations at ambient conditions) of Ac-Ala-NHMe and Ac-Gly-NHMe were analyzed and compared to experimental reference data. The α and P_{II} conformers were extracted for both molecules as the most populated ones, and these results agree with literature data from theoretical calculations in solution. Regarding the pressure dependence, relatively small changes were obtained., favoring the α conformer. The contrary picture was observed in an experimental Raman study for Ac-Ala-NHMe; however, in a similar study of Ac.-Gly-NHMe, the EC-RISM results were confirmed. New pressure-dependent EC-RISM geometry optimizations were performed for both dipeptides. Here only small pressure-dependent changes were observed, and the changes in the relative populations have the same magnitude as the observed ones for the PCM-optimized structures. Additionally, a population fit to best represent the pressure-dependent chemical shifts between experiment and theory was used to combine theoretical studies with experimental ones. For Ac-Ala-NHMe, the free energy differences between the main conformers were very small, and only marginal change upon pressurization was observed. For Ac-Gly-NHMe, the overlap between fit and EC-RISM results is not good, especially for the PCM-optimized structures.

Furthermore, a comparison of the complete Ramachandran plane of Ac-Gly-NHMe with various methods at 1 bar and 10 kbar was performed. Again, only small pressure-dependent changes were observed, with less than 1 kcal/mol difference between the two pressures. A significantly larger distinguishment is observed when QC calculations are compared with force field one. Here, the error introduced due to the force field is significantly larger than the error observed due to the missing polarization of the force field. Therefore it can be concluded that the general priority should be to develop better force fields at ambient conditions before an adaption to the small pressure-dependent changes is attempted.

The accuracy of pressure-dependent NMR calculations should be further targeted in the future. One huge potential source of error is the choice of the correct referencing. Pressure-dependent EC-RISM optimized DSS structures are unavailable, but these could lead to further improvement. In the context of this work, the placement of explicit water molecules around the peptides was neglected. Nonetheless, these should greatly impact the chemical shielding constants and increase the ^{1}H chemical shifts. An optimal way would be to use snapshots from an AIMD simulation like it was done in calculating the spectroscopic parameters of HMI.

5 Overall summary and outlook

This work aimed to apply the RISM formalism in combination with the QC description of the solute to different spectroscopic observables at high pressure. The spectroscopy variants IR, EPR, and NMR were the targets of this thesis. Furthermore, the pressure dependence of the peptide backbone was analyzed.

First, the pressure-dependent IR wavenumbers of the small osmolyte TMAO and the cyanide anion with equilibrium and non-equilibrium solvation were calculated. Compared to previous IR calculations[94], the performance in terms of absolute wavenumbers and pressure-dependent changes was strongly improved. This effect can mostly be attributed to the usage of solvent susceptibilities generated by applying 1D RISM HNC calculations and using exact electrostatic potentials in calculating the Coulomb interactions. Regarding the effect of equilibrium and non-equilibrium description, no clear trend was observed in calculations of absolute frequencies at ambient conditions. However, only the non-equilibrium method could correctly predict all pressure trends for TMAO, whereas it failed to predict the right trend for the cyanide anion. Nonetheless, it was observed that for the cyanide anion and the N-O stretch mode of TMAO, the equilibrium mode shows the best agreement with the experimental observation. The application of EC-RISM-based numerical frequency calculations under high pressure could be a new way to better understand the effect of pressure on the IR frequencies.

In the second part, for the first time, the EPR parameters of the relative rigid spin probe HMI were calculated with EC-RISM at ambient conditions and as prospective prediction at high pressure. For the calculation of the isotropic HFCC, revPBE0/def2-TZVPP calculations were performed for EC-RISM, CPCM, and a QM/MM method on snapshots of an AIMD. Two subsets were used, one where all structures were vertically desolvated and another where the explicit solvent molecules of the two first solvations shells around the N-O motif were retained. EC-RISM calculations using the DFT level of theory on the vertical desolvated data set showed significantly better results than the corresponding CPCM calculations. On the second solvation shell data set, EC-RISM performed slightly better than the QM/MM and CPCM methods. The results could be further improved by applying the DLPNO-CCSD/def2-TZVPP method, whereby EC-RISM had the smallest error compared to the experiment. Moreover, the change of the HFCC due to the protonation of HMI was investigated. Here, the DFT-based variant could accurately reproduce the observed pH-dependent change, whereas the DLPNO-CCSD methods overestimated the change. Furthermore, the W-band spectra were calculated for both unprotonated and protonated HMI. Here, for the fine-splitting in the A_{zz} region, the same trend was observed for the HFCC, where DLPNO-CCSD calculations showed the best performance. Only DFT-based methods were available for g-tensor calculations, and an intrinsic underestimation of g-values was observed here.

The last chapter was dedicated to the influence of pressure on the protein backbone. As a first benchmark, the populations of *cis*- and *trans*-NMA at ambient and high-pressure conditions were calculated and compared to experimental NMR measurements. Here, a very good agreement between theory and experiment was observed, especially in reproducing the pressure trend. In the following, the main conformers for the dipeptides Ac-Ala-NHMe and Ac-Gly-NHMe were determined. For Ac-Gly-NHMe, a very good agreement with experimental reference data was obtained, whereby in the case of Ac-Ala-NHMe, the literature is not unanimous what is the order for the three main conformers. However, EC-RISM results were still in a plausible range compared to the experiments. Under pressure, only marginal changes were observed, whereby the α-helical conformers become stronger populated due to smaller PMV. Moreover, EC-RISM-based geometry optimizations were performed in this chapter. Here, the observed pressure-induced changes were very small; however, a relatively strong effect on the chemical shifts compared to PCM-optimized structures was observed. In contrast, for the EC-RISM optimized structures, a much better agreement was achieved. A so-called NMR population fit was also made based on the calculated chemical shifts obtained from PCM- and EC-RISM-optimized structures. For Ac-Ala-NHMe, the results agree relatively well with the pure EC-RISM calculations, whereas for the PCM-optimized structures of Ac-Gly-NHMe large deviations were obtained. For analyzing the question if a force field parameterization becomes necessary for high-pressure MD simulations, the Ramachandran planes of Ac-Gly-NHMe, calculated with forcefield-based methods and QC methods, were studied at ambient pressure and an extreme pressure of 10 kbar. Here, it was shown that the errors from the force fields at ambient conditions are larger than the error in neglecting the polarization effect upon pressurization.

The thesis showed that a description of the solvent via RISM and QC description of the solute in the form of the EC-RISM could be used to accurately calculate different spectroscopic observables in solution. Especially when non-ambient conditions are the subject of investigation, EC-RISM can show its strength since it can be easily adapted to such conditions. Future work should address pressure-dependent NMR calculations based on a conformational ensemble obtained from an AIMD, which would be the best, or, as an alternative classical MD simulations could be used. High potential is grounded in the novel EC-RISM geometry optimizations, which could also be applied to other problems like pK_a calculations.

6 Appendix

6.1 Exemplary pressure-dependent distortions plots of TMAO calculated with MP2/6-311+G(d,p)

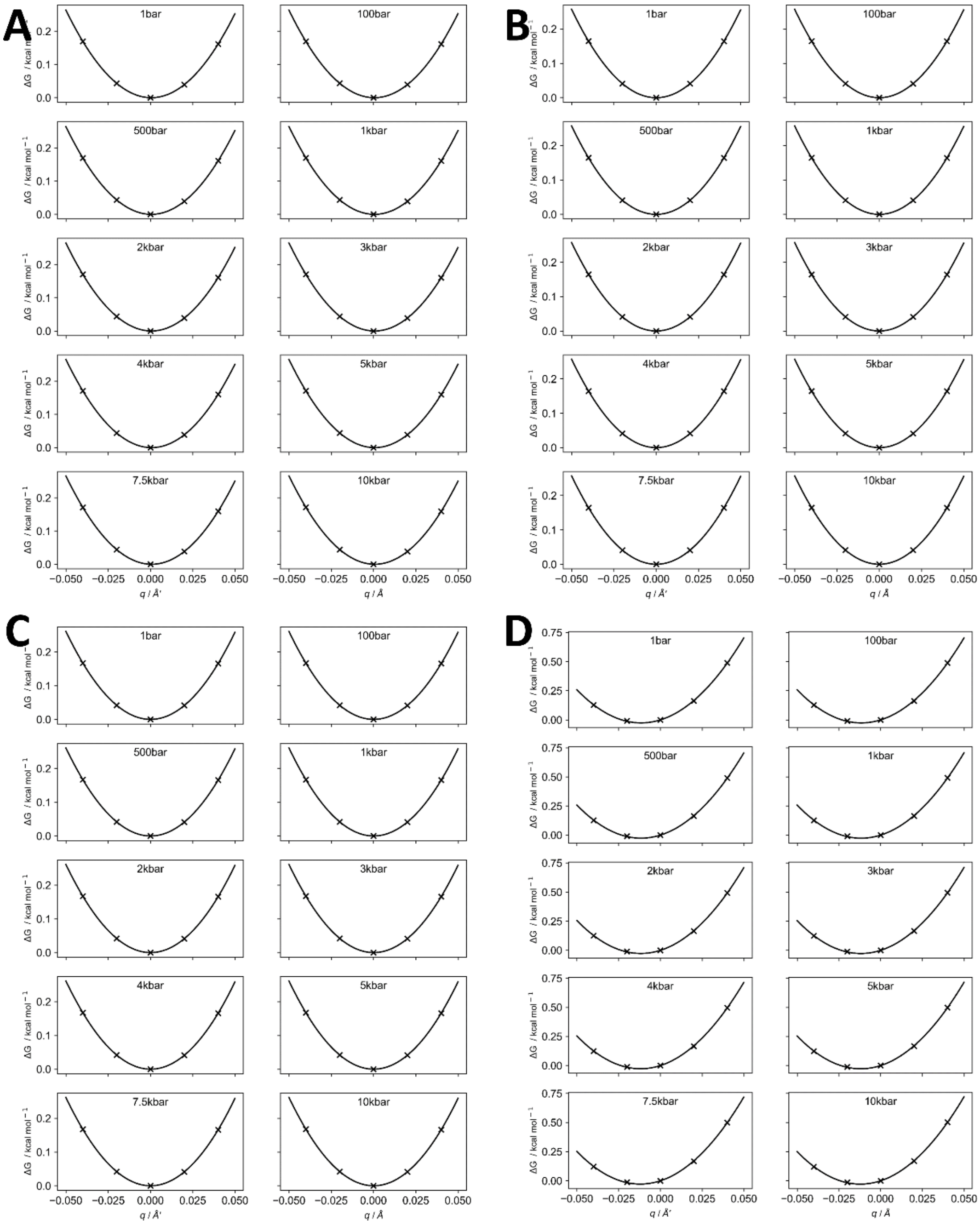

Figure 38: Pressure-dependent distortion plots for the Gibbs free energies of the CH_3-def-high mode (A), CH_3-def-middle mode (B), CH_3-def-low mode (C), and NO-stretch mode (D) of TMAO calculated with MP2/6-311+G(d,p)/χ_{HNC} and EQ method. Raw data can be found in SI_4.1.

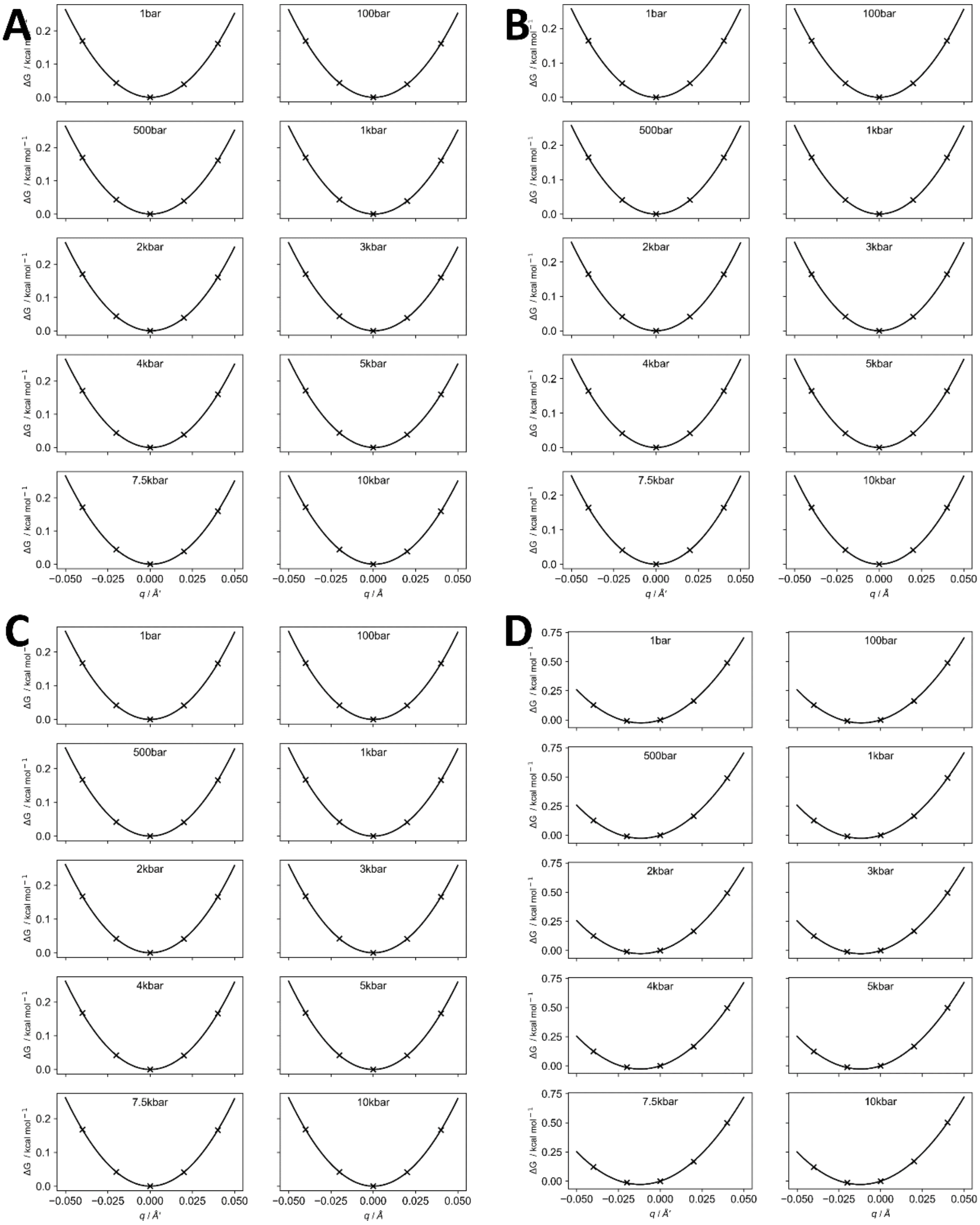

Figure 39: Pressure-dependent distortion plots for the intramolecular energies of the CH_3-def-high mode (A), CH_3-def-middle mode (B), CH_3-def-low mode (C), and NO-stretch mode (D) of TMAO calculated with MP2/6-311+G(d,p)/χ_{HNC} and EQ method. Raw data can be found in SI_4.1.

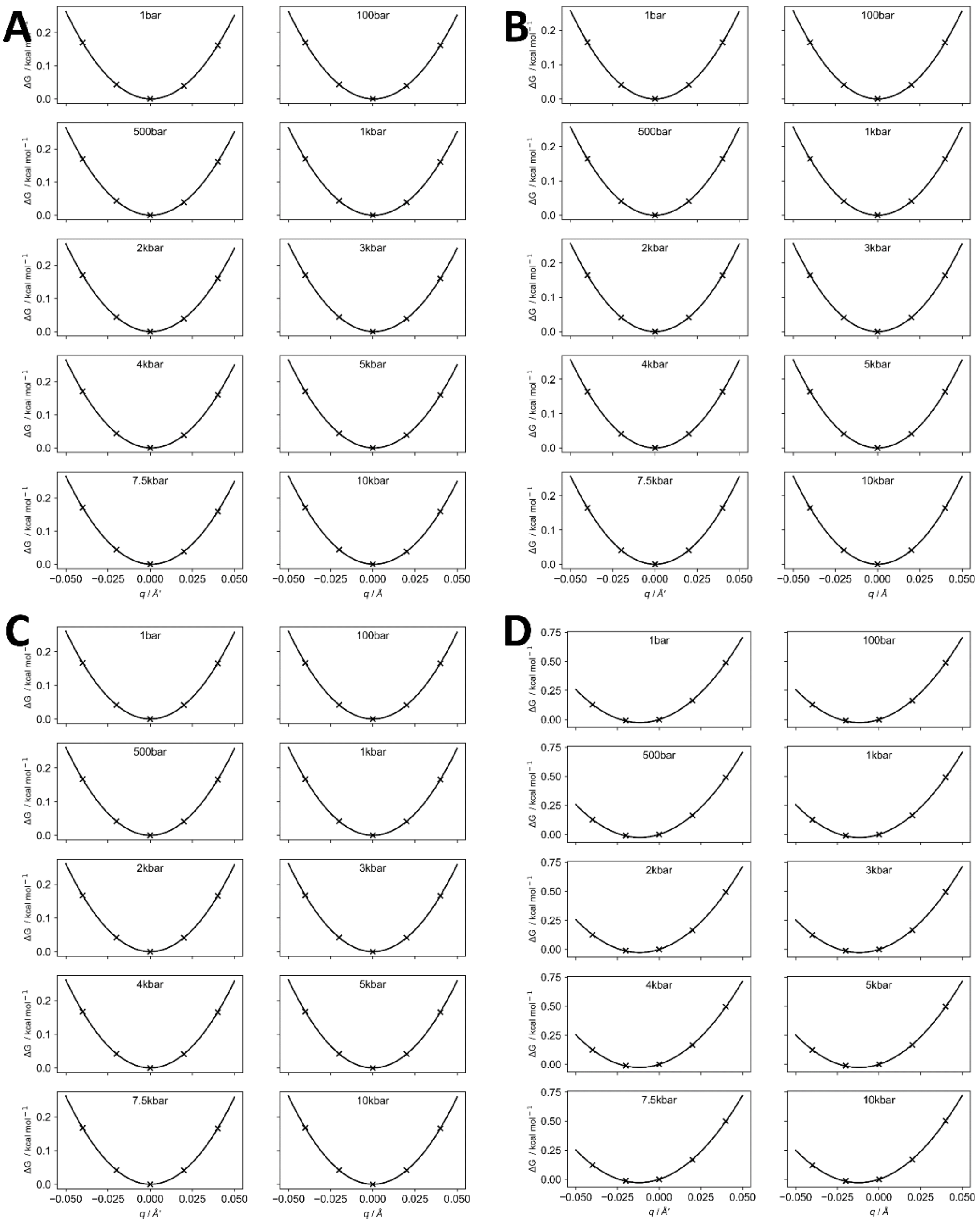

Figure 40: Pressure-dependent distortion plots for the μ^{ex} of the CH_3-def-high mode (A), CH_3-def-middle mode (B), CH_3-def-low mode (C), and NO-stretch mode (D) of TMAO calculated with MP2/6-311+G(d,p)/χ_{HNC} and EQ method. Raw data can be found in SI_4.1.

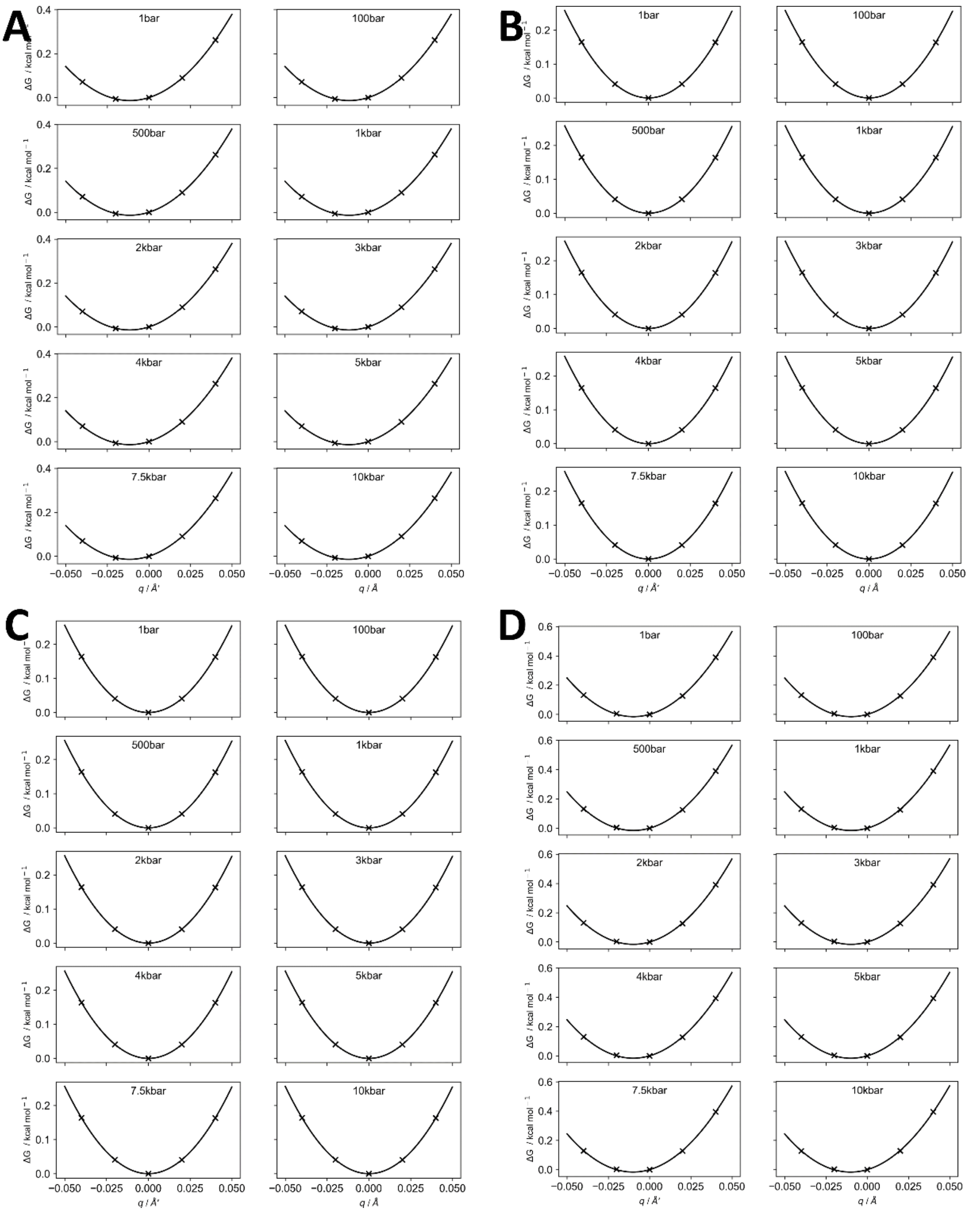

Figure 41: Pressure-dependent distortion plots for the intramolecular energies of the CH_3-def-high mode (A), CH_3-def-middle mode (B), CH_3-def-low mode (C), and NO-stretch mode (D) of TMAO calculated with MP2/6-311+G(d,p)/χ_{HNC} for the two non-equilibrium methods. Raw data can be found in SI_4.1.

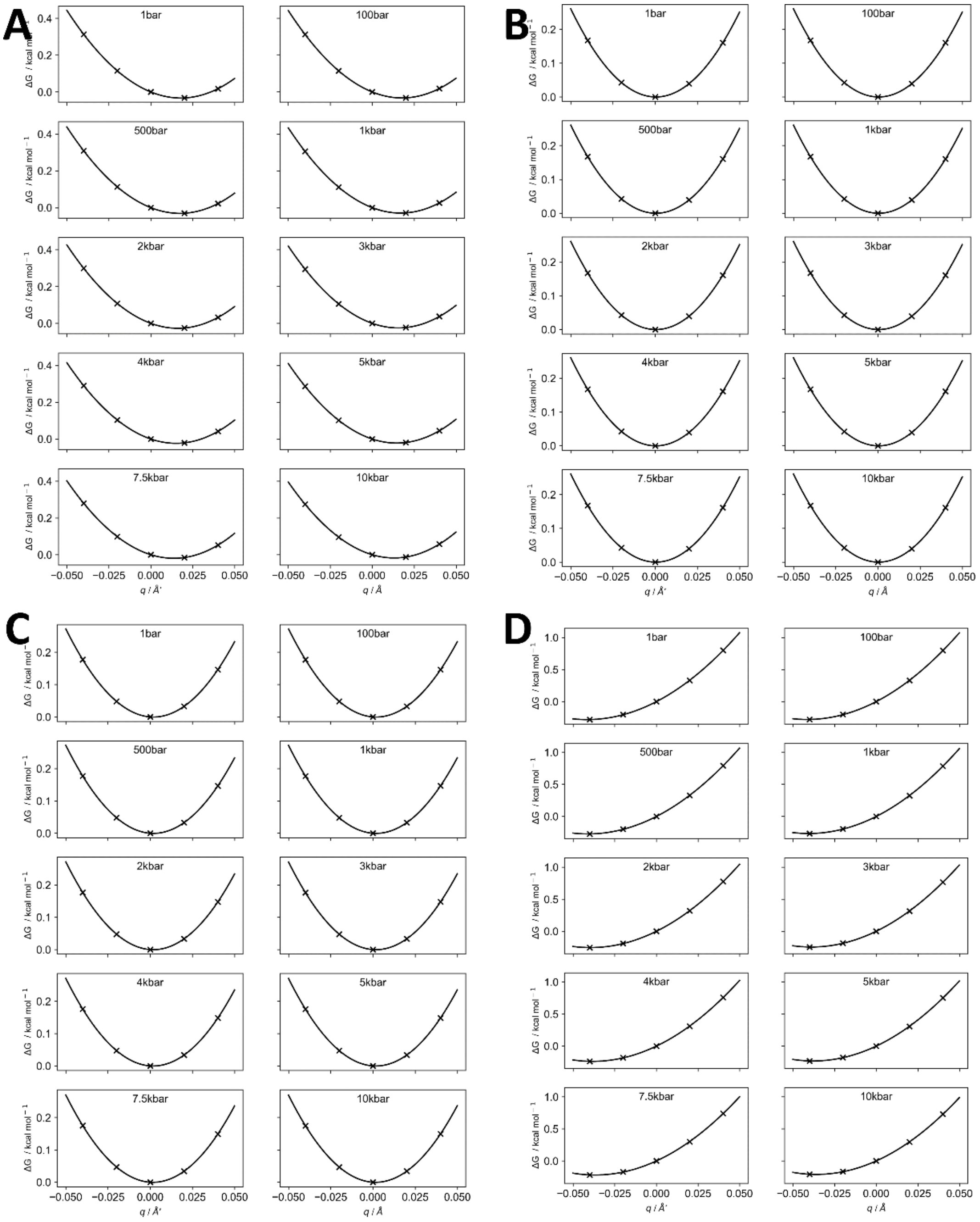

Figure 42: Pressure-dependent distortion plots for the Gibbs free energies of the CH_3-def-high mode (A), CH_3-def-middle mode (B), CH_3-def-low mode (C), and NO-stretch mode (D) of TMAO calculated with MP2/6-311+G(d,p)/χ_{HNC} for the M1 non-equilibrium method. Raw data can be found in SI_4.1.

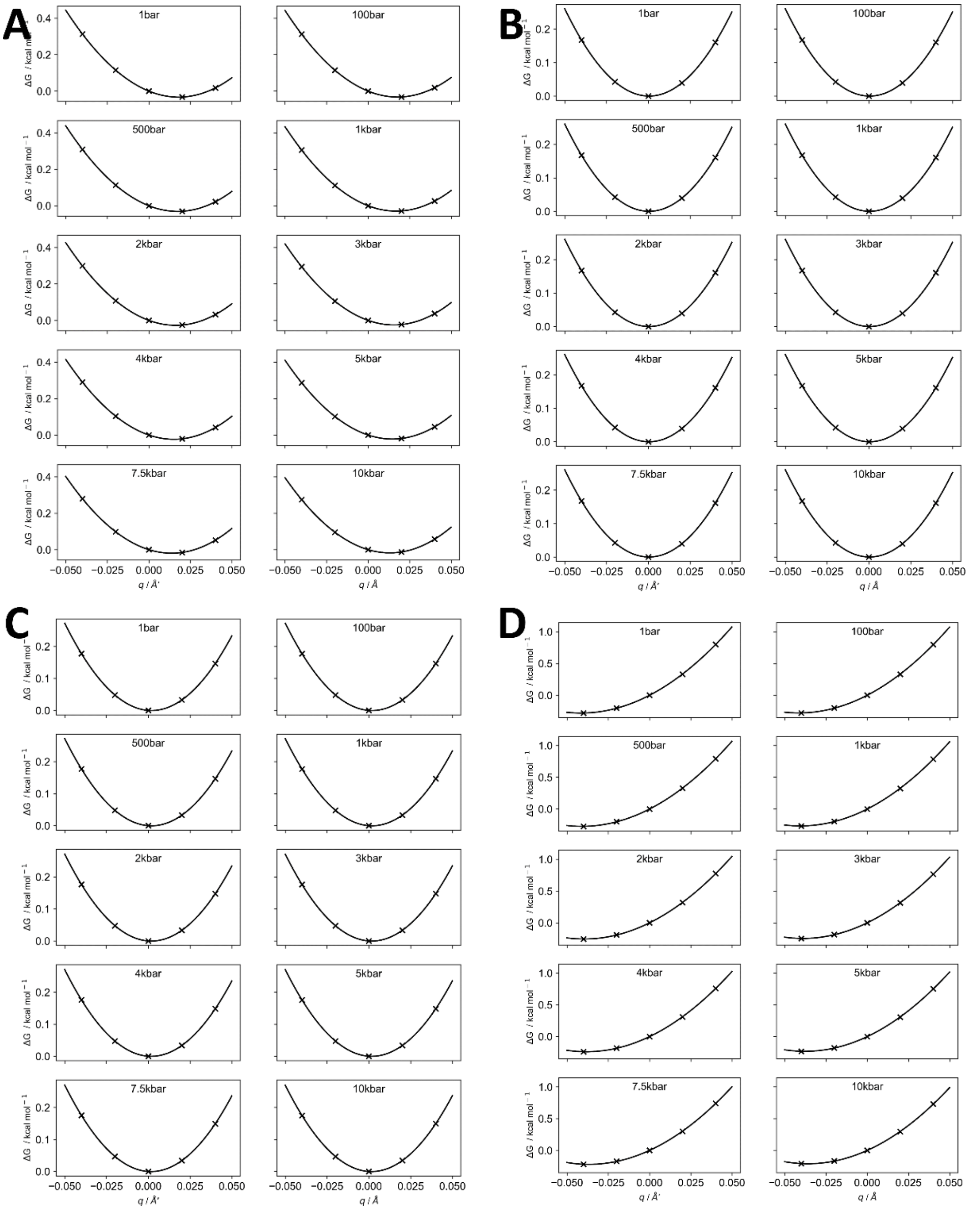

Figure 43: Pressure-dependent distortion plots for the μ^{ex} of the CH_3-def-high mode (A), CH_3-def-middle mode (B), CH_3-def-low mode (C), and NO-stretch mode (D) of TMAO calculated with MP2/6-311+G(d,p)/χ_{HNC} for the M1 non-equilibrium method. Raw data can be found in SI_4.1.

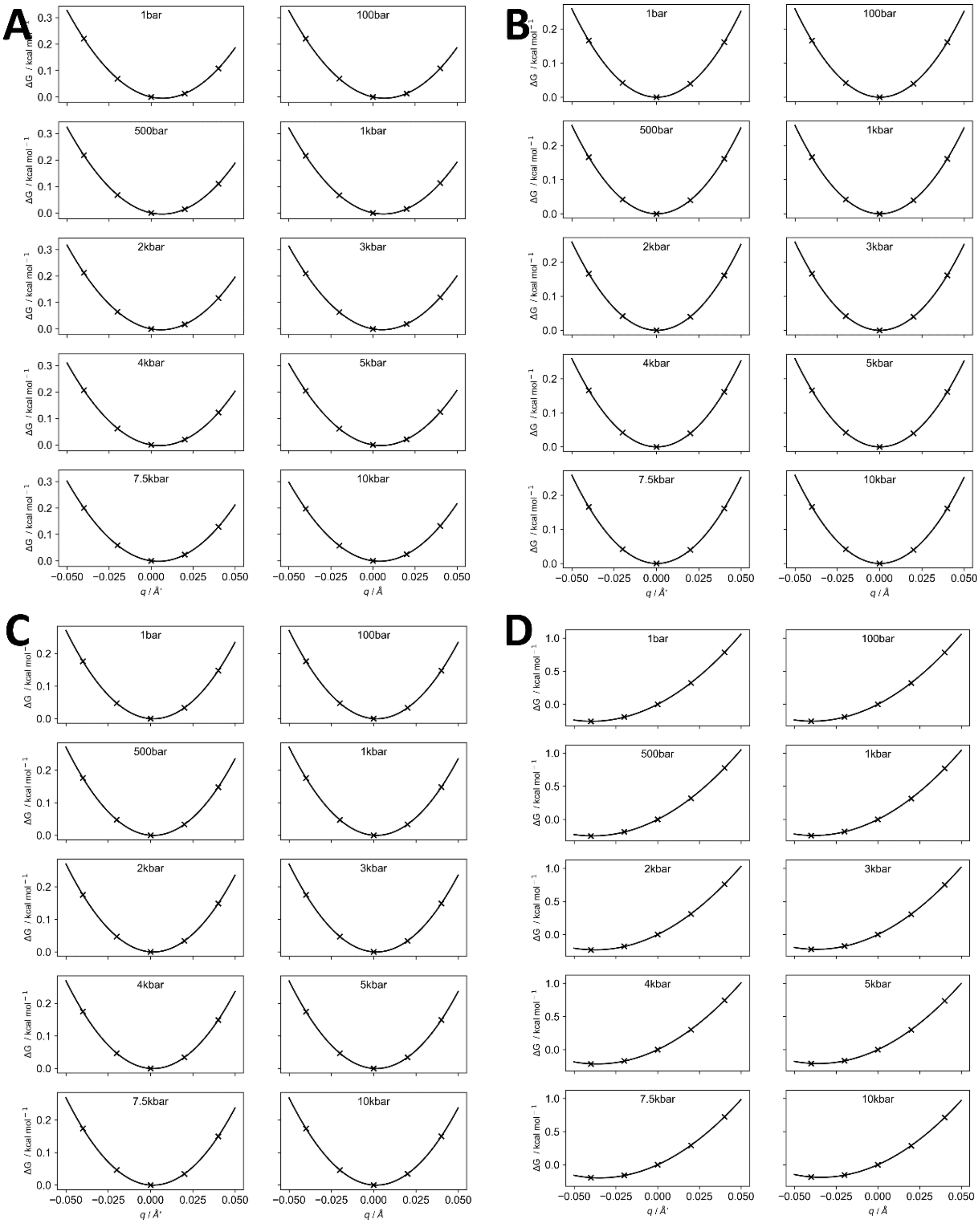

Figure 44: Pressure-dependent distortion plots for the Gibbs free energies of the CH_3-def-high mode (A), CH_3-def-middle mode (B), CH_3-def-low mode (C), and NO-stretch mode (D) of TMAO calculated with MP2/6-311+G(d,p)/χ_{HNC} for the M2 non-equilibrium method. Raw data can be found in SI_4.1.

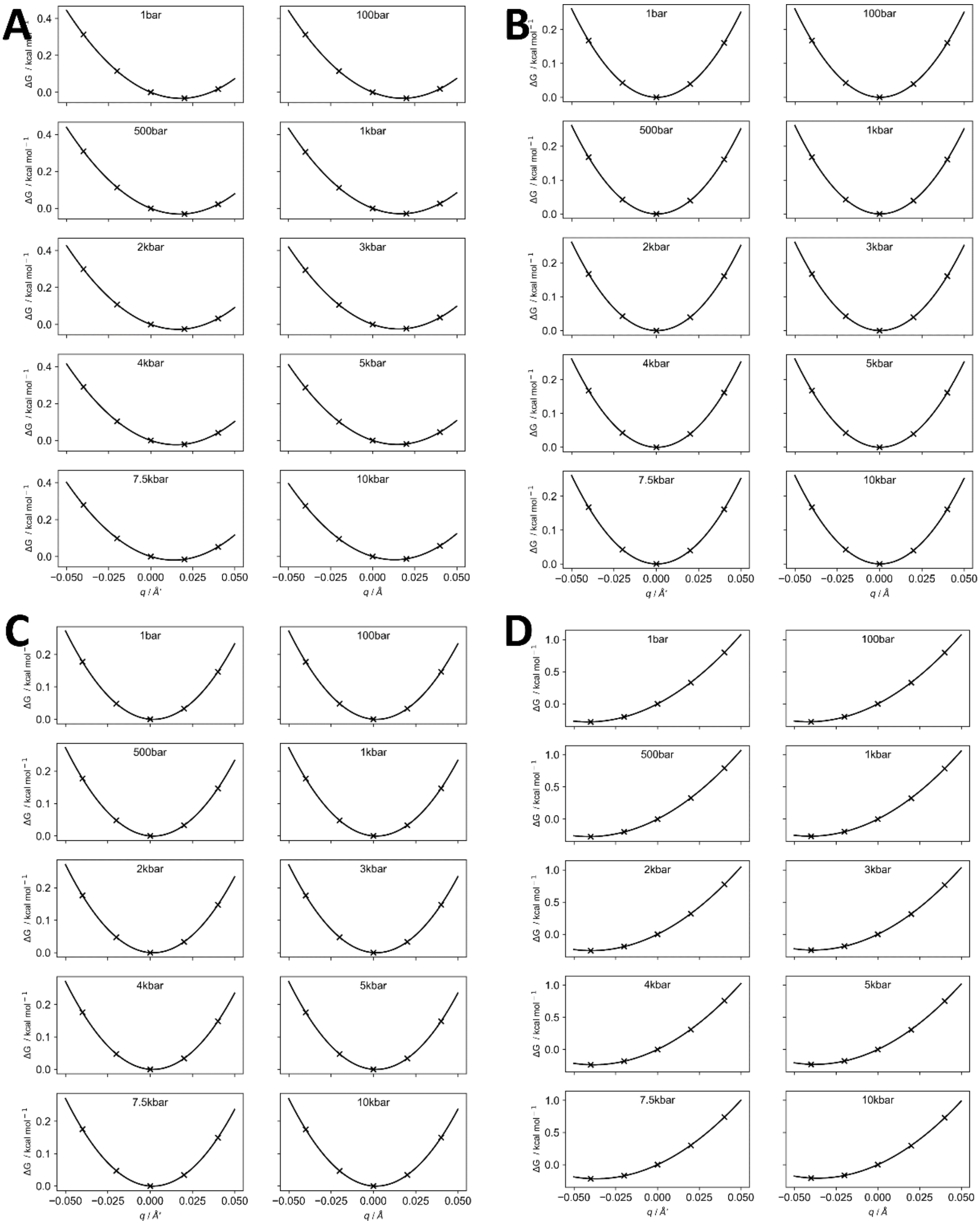

Figure 45: Pressure-dependent distortion plots for the μ^{ex} of the CH_3-def-high mode (A), CH_3-def-middle mode (B), CH_3-def-low mode (C), and NO-stretch mode (D) of TMAO calculated with MP2/6-311+G(d,p)/χ_{HNC} for the M2 non-equilibrium method. Raw data can be found in SI_4.1.

6.2 High-Pressure principal components of HMI and HHMI

Table 22: Pressure-dependent principal components of HMI calculated with revPBE0/def2-TZVVP and decontracted s-functions. The coupling tensors are in MHz.

Pressure	A_{zz}	A_{yy}	A_{xx}	g_{zz}	g_{yy}	g_{xx}
1 bar	92	2.65	2.55	2.002126	2.005708	2.00776
100 bar	92.01	2.65	2.55	2.002126	2.005707	2.007759
500 bar	92.08	2.66	2.56	2.002126	2.005706	2.007754
1 kbar	92.15	2.68	2.57	2.002126	2.005705	2.007749
2 kbar	92.28	2.7	2.58	2.002126	2.005703	2.007739
3 kbar	92.39	2.72	2.6	2.002126	2.005701	2.007731
4 kbar	92.49	2.73	2.61	2.002127	2.005699	2.007723
5 kbar	92.58	2.75	2.62	2.002127	2.005697	2.007717
7.5 kbar	92.78	2.78	2.65	2.002127	2.005694	2.007702
10 kbar	92.94	2.81	2.67	2.002127	2.005691	2.00769

Table 23: Pressure-dependent principal components of HHMI calculated with revPBE0/def2-TZVVP and decontracted s-functions. The coupling tensors are in MHz.

Pressure	A_{zz}	A_{yy}	A_{xx}	g_{zz}	g_{yy}	g_{xx}
1 bar	83.56	1.48	1.79	2.002111	2.005751	2.008381
100 bar	83.58	1.49	1.79	2.002111	2.005751	2.008379
500 bar	83.68	1.5	1.8	2.002111	2.00575	2.008371
1 kbar	83.8	1.52	1.82	2.002111	2.005748	2.008363
2 kbar	84.01	1.56	1.84	2.002111	2.005745	2.008347
3 kbar	84.19	1.58	1.86	2.002111	2.005743	2.008334
4 kbar	84.36	1.61	1.88	2.002112	2.00574	2.008322
5 kbar	84.5	1.63	1.9	2.002112	2.005738	2.008311
7.5 kbar	84.82	1.68	1.94	2.002112	2.005734	2.008288
10 kbar	85.08	1.72	1.97	2.002112	2.00573	2.00827

Table 24: Linear (a, in kbar^{-1} (g_{xx}) or MHz·kbar^{-1} (HFCC and A_{zz})), quadratic (b in, in kbar^{-2} (g_{xx}) or MHz·kbar^{-2} (HFCC and A_{zz})), and offset (c in MHz for HFCC and A_{zz}) coefficients from fitting calculated HMI and HHMI g_{xx}, HFCC, and A_{zz}, to the following form: $\boldsymbol{\Delta}\delta(p)= B_1p+B_2p^2$. The 1 bar chemical shifts (δ_0 in ppm) are shown.

Parameters	a	b	c	a	b	c
		HMI			HHMI	
g_{xx}	$-6.98\cdot10^{-7}\cdot$	$-1.05\cdot10^{-5}$	$3.61\cdot10^{-7}$	$-9.6410^{-7}\cdot$	$-1.05\cdot10^{-5}$	$5.80\cdot10^{-7}$
HFCC	0.0047	0.0594	-0.002	0.0079	0.0974	-0.00037
A_{zz}	0.0071	0.1397	-0.005	0.0092	0.2273	-0.00774

6.3 Histograms of g- and A-tensors of HMI and HHMI

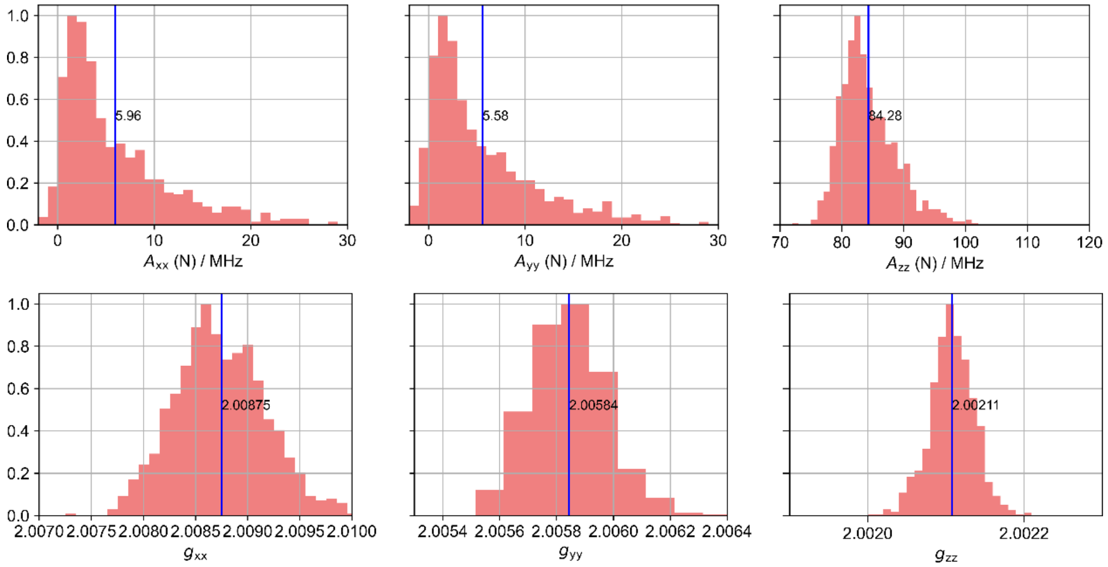

Figure 46 Normalized A- and g-tensor distributions for HHMI calculated with revPBE0/def2-TZVVP/CPCM on the subset containing 1000 vertically desolvated structures.

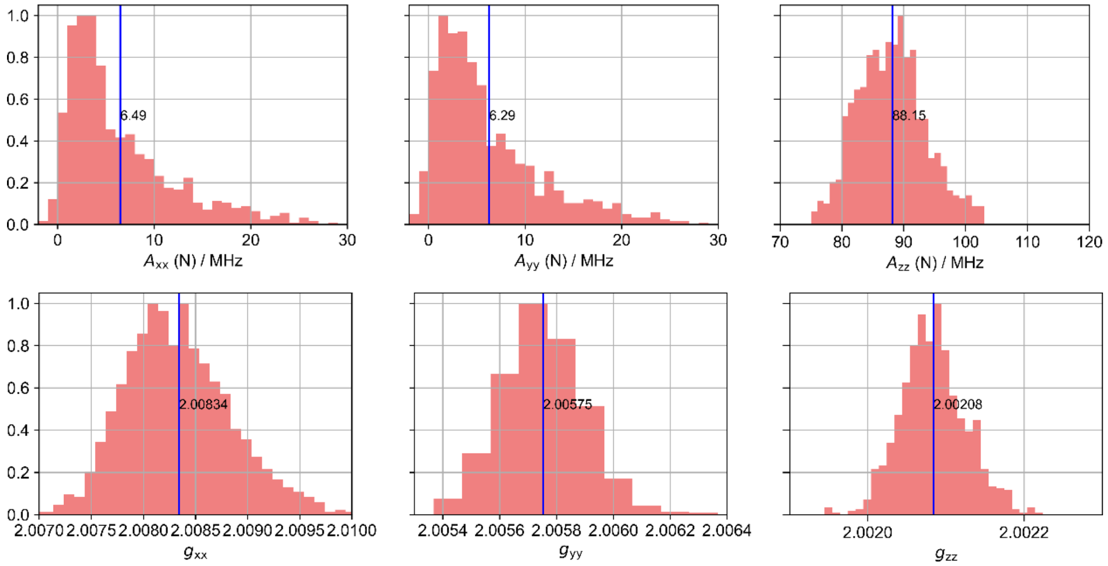

Figure 47 Normalized A- and g-tensor distributions for HHMI calculated with revPBE0/def2-TZVVP/CPCM on the subset containing 1000 structures with explicitly the second solvation shell included around the N-O motif.

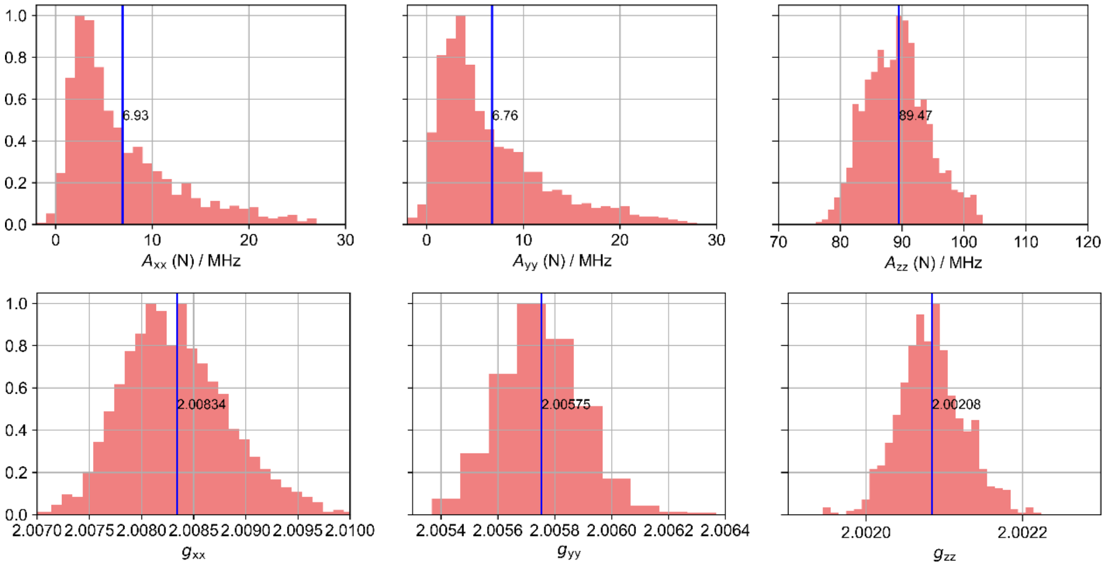

Figure 48 Normalized *A*- and *g*-tensor distributions for HHMI calculated with revPBE0/def2-TZVVP/EC-RISM on the subset containing 1000 vertically desolvated structures.

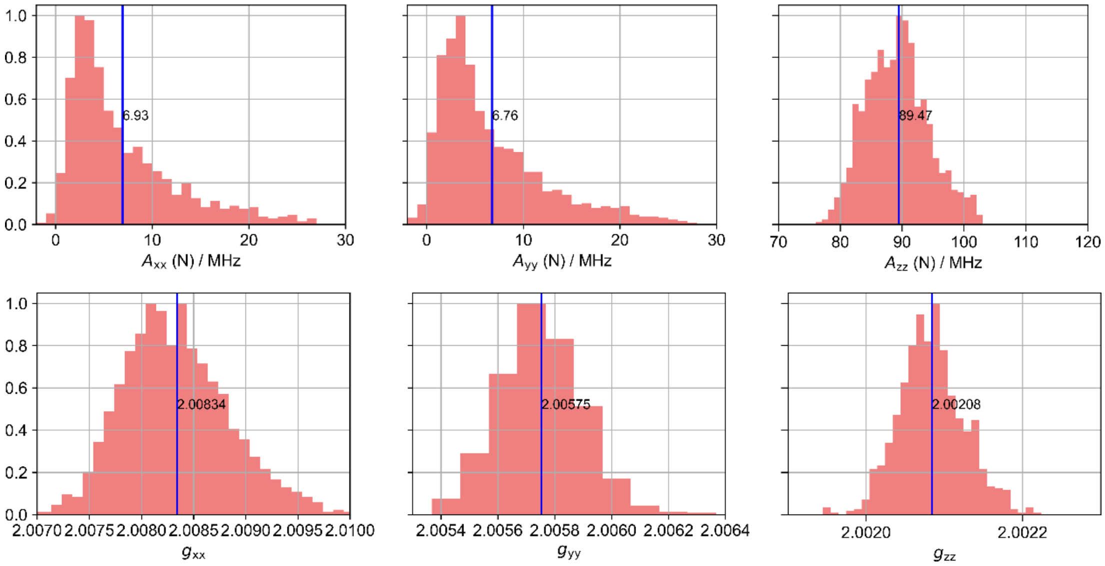

Figure 49 Normalized *A*- and *g*-tensor distributions for HHMI calculated with revPBE0/def2-TZVVP/EC-RISM on the subset containing 1000 structures with explicitly the second solvation shell included around the N-O motif.

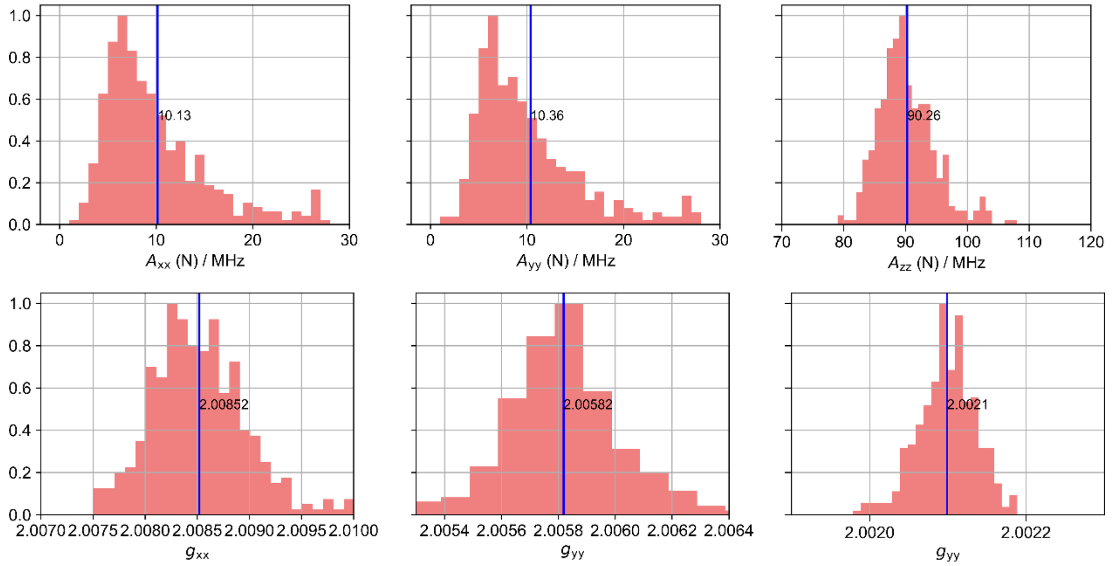

Figure 50 Normalized *A*- and *g*-tensor distributions for HHMI calculated with DLPNO-CCSD/def2-TZVVP/EC-RISM on the subset containing 400 vertically desolvated structures.

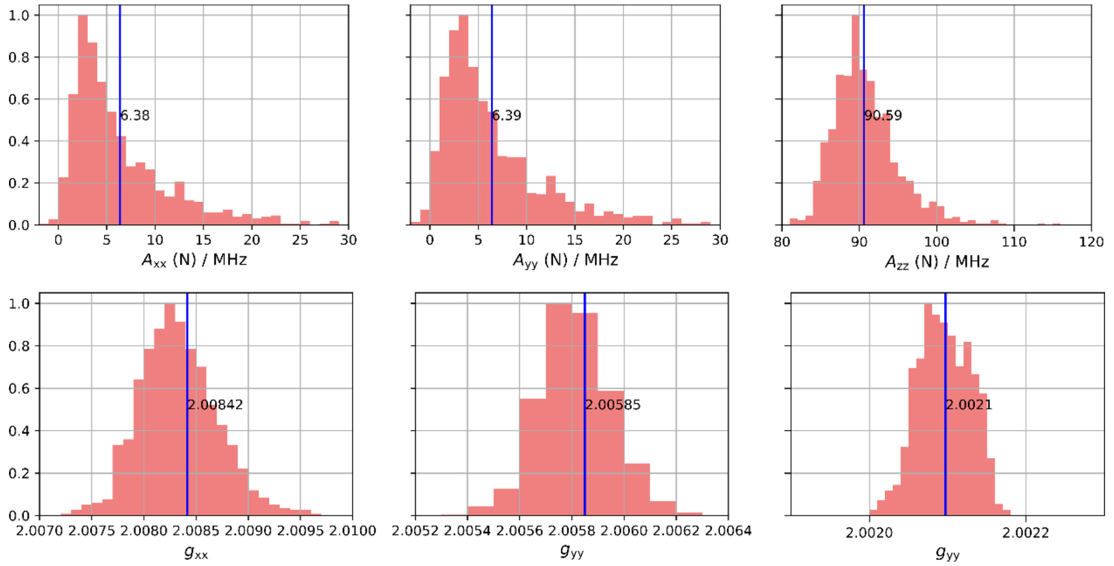

Figure 51 Normalized *A*- and *g*-tensor distributions for HMI calculated with revPBE0/def2-TZVVP/CPCM on the subset containing 1000 vertically desolvated structures.

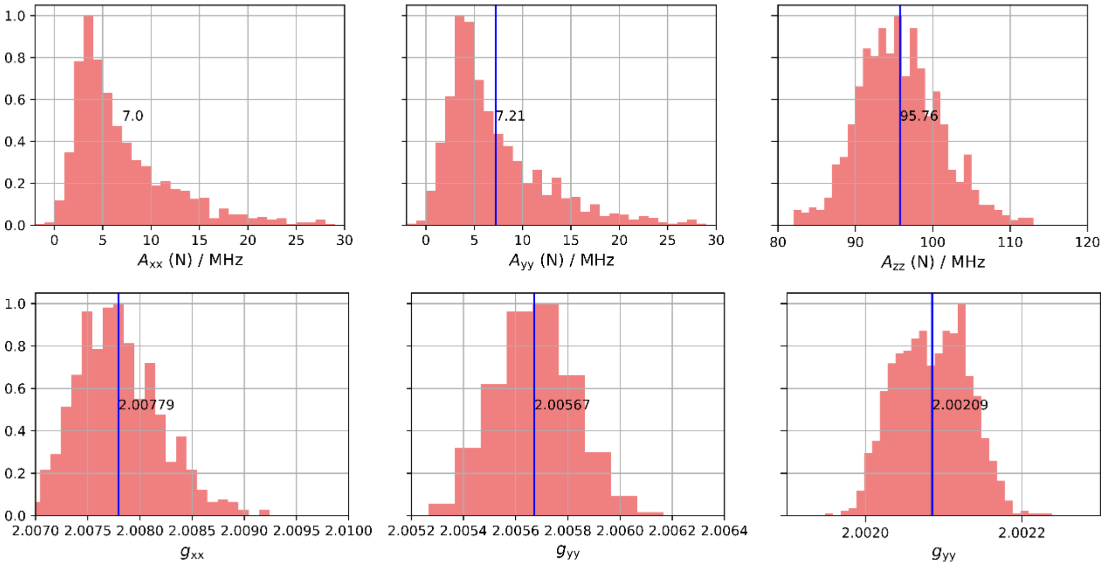

Figure 52 Normalized *A*- and *g*-tensor distributions for HMI calculated with revPBE0/def2-TZVVP/CPCM on the subset containing 1000 structures with explicitly the second solvation shell included around the N-O motif.

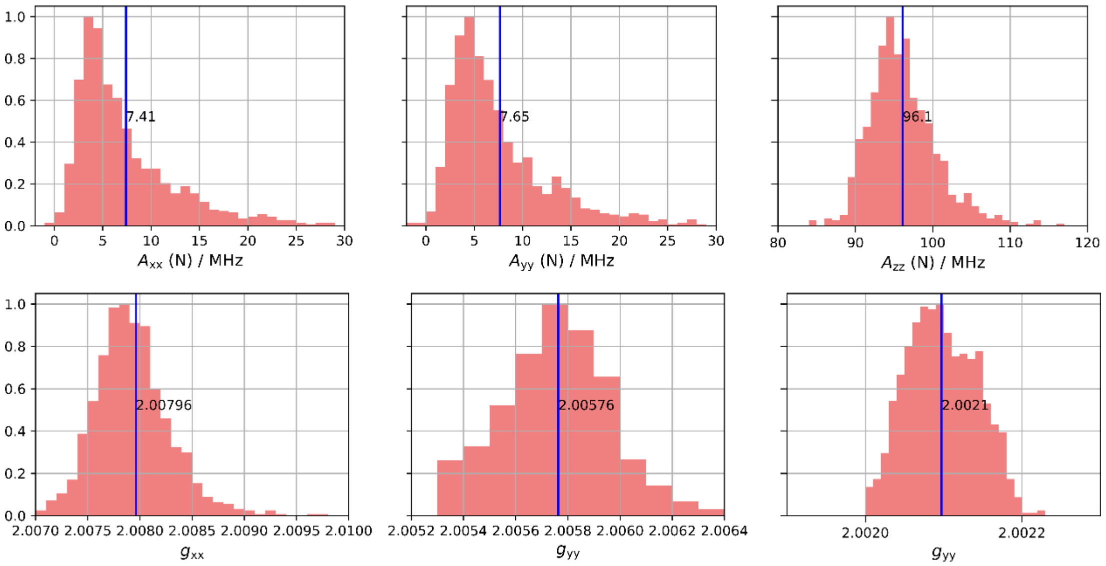

Figure 53 Normalized *A*- and *g*-tensor distributions for HMI calculated with revPBE0/def2-TZVVP/EC-RISM on the subset containing 1000 vertically desolvated structures.

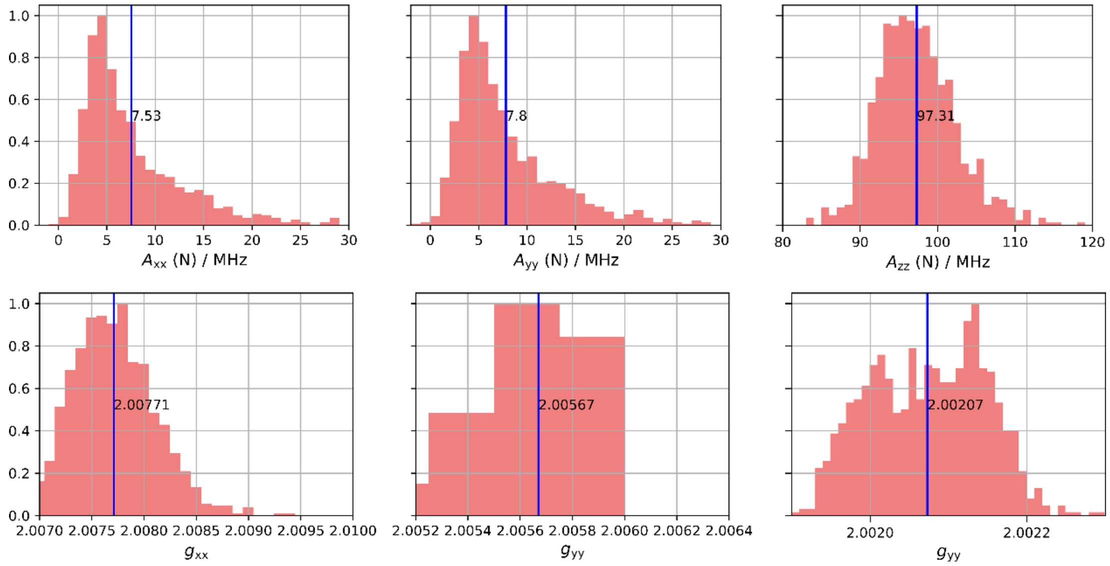

Figure 54 Normalized *A*- and *g*-tensor distributions for HMI calculated with revPBE0/def2-TZVVP/EC-RISM on the subset containing 1000 structures with explicitly the second solvation shell included around the N-O motif.

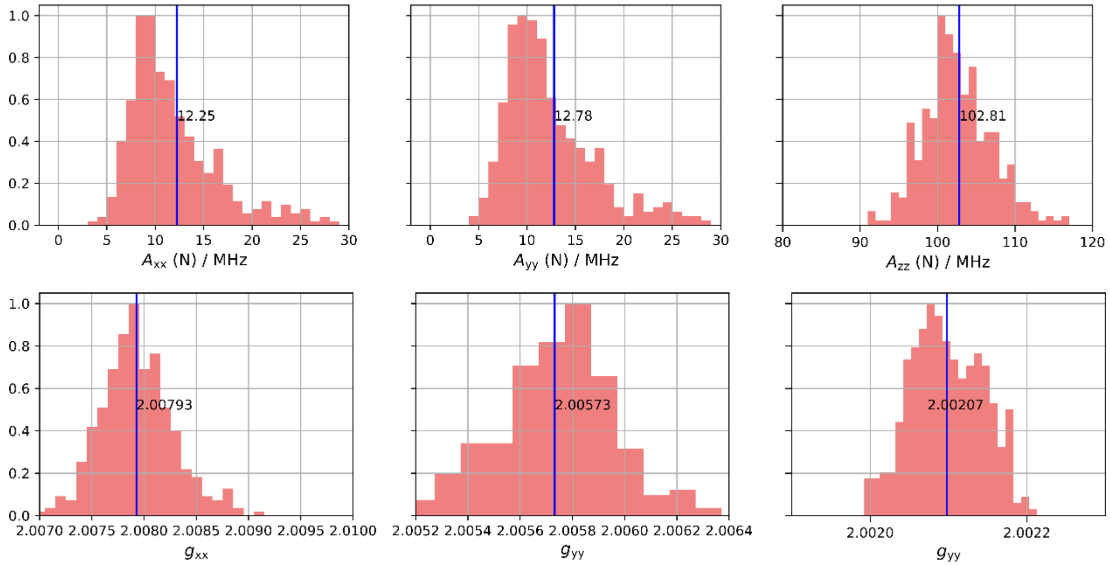

Figure 55 Normalized *A*- and *g*-tensor distributions for HMI calculated with DLPNO-CCSD/def2-TZVVP/EC-RISM on the subset containing 400 vertically desolvated structures.

6.4 Ramachandran free energy surfaces of Ac-Gly/Ala-NHMe

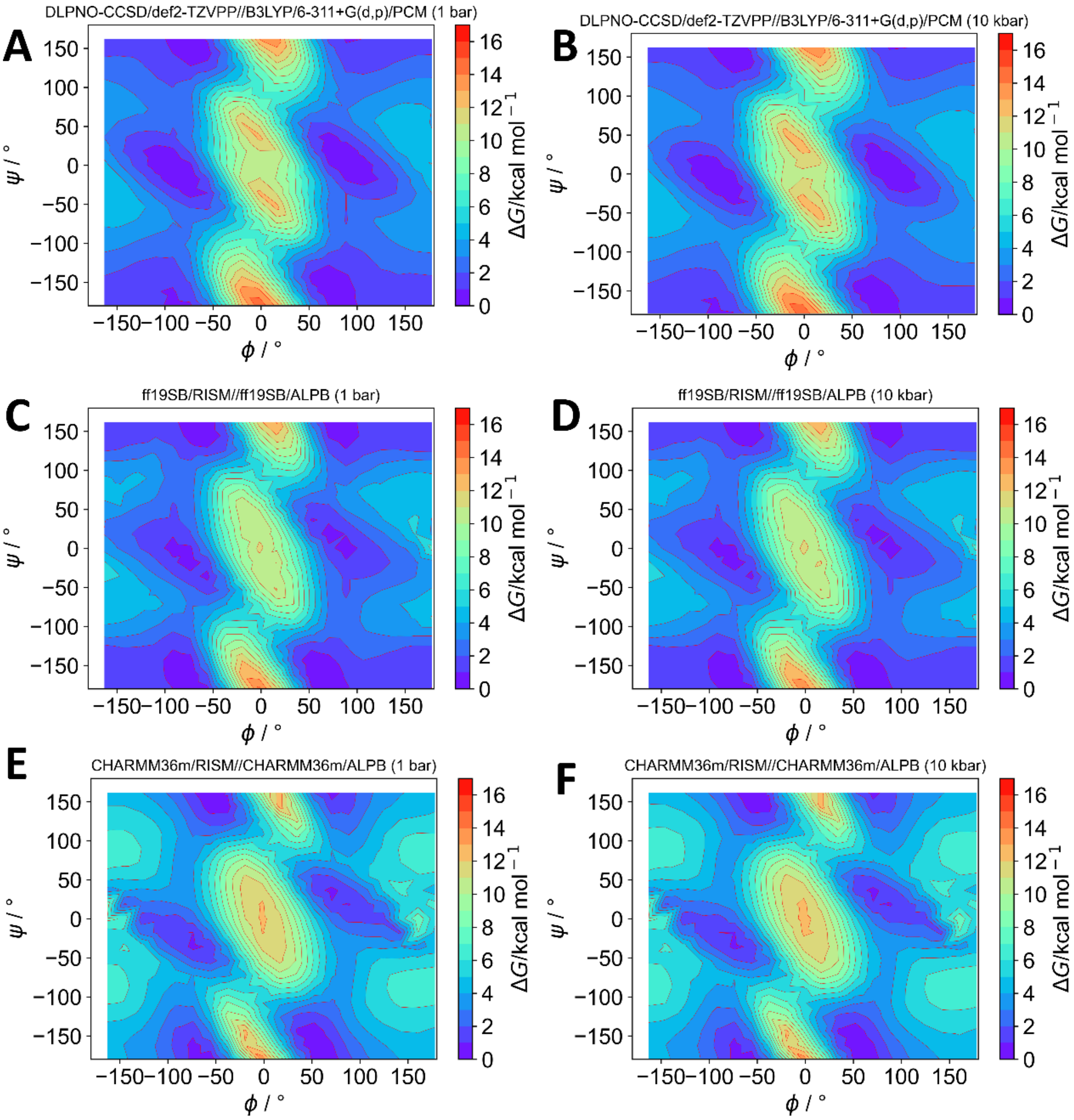

Figure 56 Pressure-dependent Ramachandran plane of Ac-Gly-NHMe. Panel A and B were calculated using DLPNO-CCSD/def2-TZVPP/PSE-3//B3LYP-63111+G(d,p) at 1 bar and 10 kbar, respectively. Panel C and D show the ff19SB/RISM calculations at 1 bar and 10 kbar, respectively. The results obtained from the CHARMM/36m/RISM calculations are presented in panels E (1 bar) and F (10 kbar). The color scale ranges from 0 kcal/mol to 16 kcal/mol, and the iso levels are 1 kcal/mol.

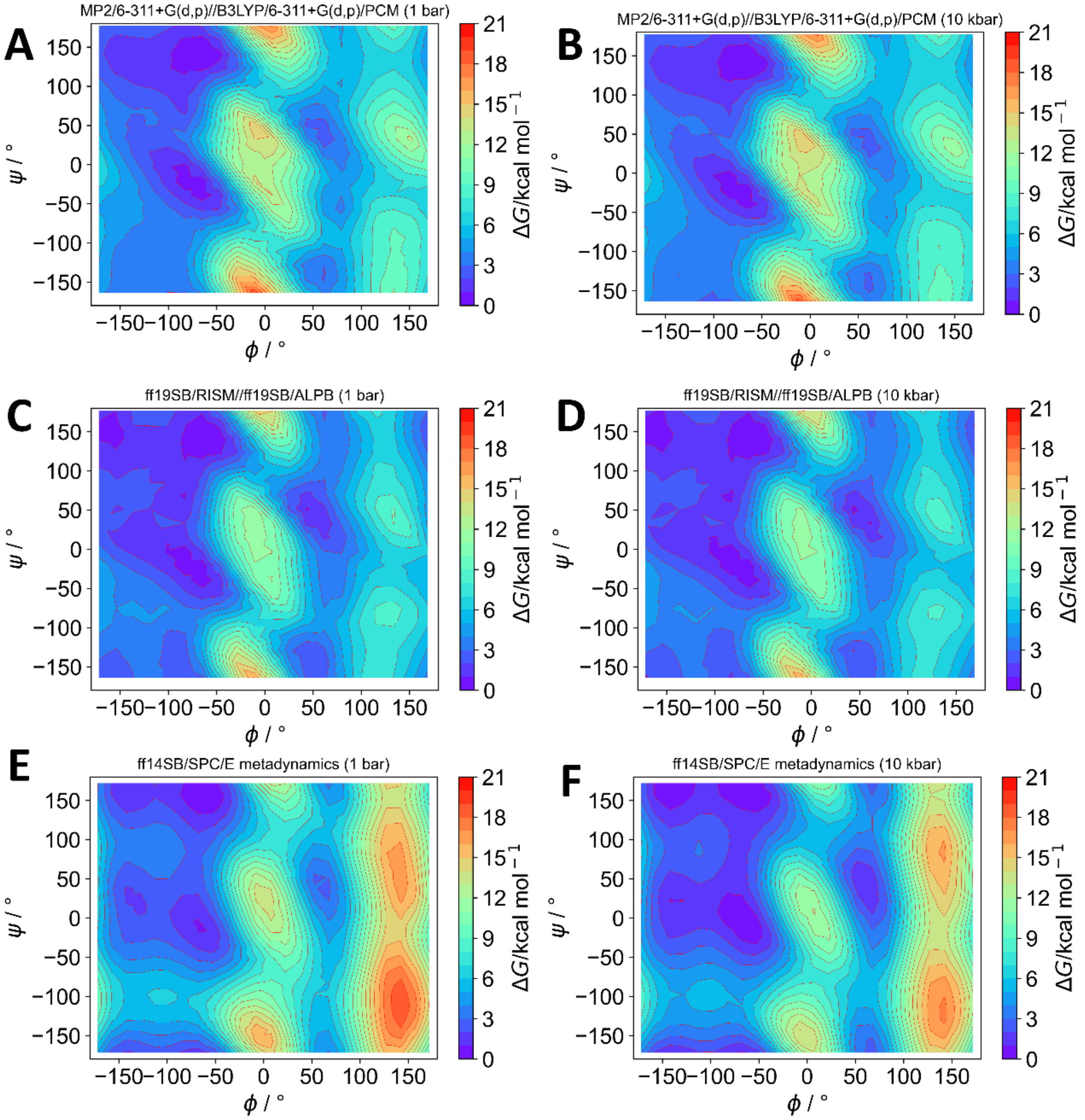

Figure 57 Pressure-dependent Ramachandran plane of Ac-Gly-NHMe. Panel A and B were calculated using MP2/6-311+G(d,p)//B3LYP-63111+G(d,p)/PCM at 1 bar and 10 kbar, respectively. Panel C and D show the ff19SB/RISM calculations at 1 bar and 10 kbar, respectively. The results obtained from the ff14SB/SPC/E metadynamics calculations are presented in panels E (1 bar) and F (10 kbar). The color scale ranges from 0 kcal/mol to 22kcal/mol, and the iso levels are 1 kcal/mol.

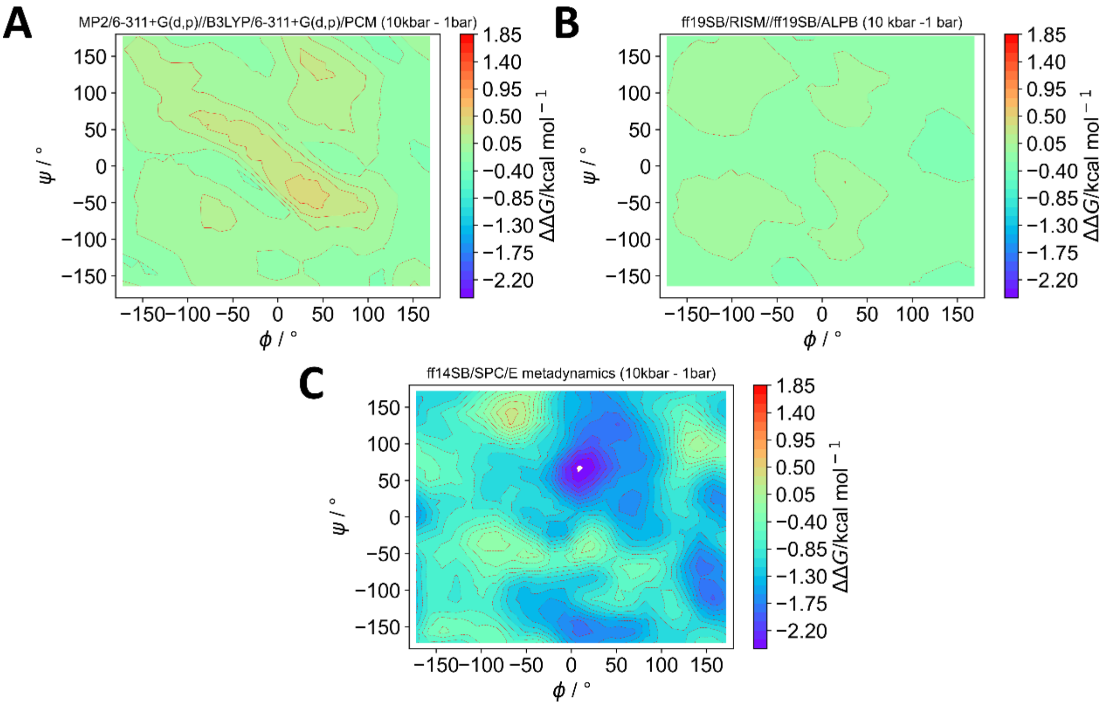

Figure 58 Difference-free energy Ramachandran surfaces between 10 kbar and 1 bar of Ac-Ala-NHMe calculated with different methods. Panel A shows the free energy difference calculated with MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p). In panel B, the corresponding RPBE-D3/TZVPP calculations are depicted.

6.5 Pressure-dependent chemical shifts of the peptide backbone

Table 25 Calculated pressure-dependent chemical shifts for the ^{1}H amide protons for the seven and eight minima of Ac-Ala-NHMe and Ac-Gly-NHMe, respectively. The calculations were performed with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Conformer	1 bar	1 kbar	2 kbar	3 kbar
Ac-Ala-NHMe				
P_{II}	6.82	6.84	6.87	6.89
α_R	6.73	6.76	6.79	6.81
C_5	7.18	7.21	7.23	7.25
$C_{7,eq}$	6.72	6.75	6.78	6.80
α_L	7.18	7.21	7.23	7.25
a_d	7.21	7.24	7.27	7.29
$C_{7,ax}$	7.02	7.05	7.07	7.09
Exp	8.26	8.31	8.35	8.38
P_{II}	6.81	6.83	6.86	6.89
α_R	7.08	7.11	7.13	7.15
C_5	6.67	6.70	6.72	6.74
$C_{7,eq}$	7.47	7.49	7.50	7.52
α_L	7.03	7.05	7.08	7.10
a_d	6.84	6.88	6.90	6.92
$C_{7,ax}$	8.84	8.86	8.87	8.89
Exp	7.94	7.97	8.00	8,02
Ac-Gly-NHMe				
α_R	6.85	6.88	6.90	6.93
α_L	6.87	6.89	6.91	6.93
P_{II}	7.03	7.05	7.07	7.10
P_{II}	7.02	7.05	7.08	7.09
β	6.70	6.72	6.75	6.76
C_5	7.12	7.14	7.16	7.19
$C_{7,eq}$	7.12	7.14	7.16	7.18
$C_{7,eq}$	7.05	7.08	7.11	7.13
Exp	8.30	8.32	8.34	8.35
α_R	7.07	7.10	7.13	7.15
α_L	7.06	7.10	7.12	7.14
P_{II}	6.71	6.73	6.75	6.77
P_{II}	6.70	6.72	6.75	6.77
β	6.66	6.70	6.72	6.75
C_5	7.39	7.41	7.44	7.45
$C_{7,eq}$	7.39	7.43	7.45	7.47
$C_{7,eq}$	6.51	6.55	6.57	6.59
Exp	7.86	7.88	7.91	7.93

Table 26 Calculated pressure-dependent chemical shifts for the ^{13}C amide protons for the seven and eight minima of Ac-Ala-NHMe and Ac-Gly-NHMe, respectively. The calculations were performed with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Conformer	1 bar	1 kbar	2 kbar	3 kbar
Ac-Ala-NHMe				
P_{II}	180.50	180.57	180.71	180.79
α_R	180.96	181.12	181.23	181.33
C_5	179.13	179.26	179.37	179.45
$C_{7,eq}$	180.05	180.16	180.24	180.33
α_L	180.56	180.72	180.82	180.92
a_d	180.06	180.20	180.31	180.41
$C_{7,ax}$	181.58	181.70	181.79	181.87
Exp	180.50	180.57	180.71	180.79
P_{II}	182.15	182.20	182.28	182.32
α_R	181.48	181.56	181.63	181.70
C_5	179.94	180.03	180.09	180.14
$C_{7,eq}$	181.17	181.26	181.34	181.41
α_L	180.37	180.47	180.54	180.60
a_d	180.29	180.39	180.47	180.54
$C_{7,ax}$	184.17	184.26	184.34	184.40
Exp	182.15	182.20	182.28	182.32
Ac-Gly-NHMe				
α_R	181.57	181.65	181.72	181.82
α_L	181.58	181.62	181.73	181.84
P_{II}	181.62	181.74	181.84	181.93
P_{II}	181.62	181.72	181.83	181.93
β	180.93	181.08	181.16	181.26
C_5	181.00	181.08	181.15	181.22
$C_{7,eq}$	180.98	181.06	181.14	181.23
$C_{7,eq}$	180.26	180.37	180.45	180.51
Exp	177.75	177.76	177.80	177.85
α_R	178.55	178.63	178.71	178.78
α_L	178.62	178.63	178.73	178.81
P_{II}	178.07	178.13	178.21	178.25
P_{II}	178.04	178.11	178.16	178.23
β	178.18	178.28	178.36	178.43
C_5	180.31	180.39	180.47	180.53
$C_{7,eq}$	180.16	180.26	180.33	180.44
$C_{7,eq}$	175.07	175.19	175.26	175.32
Exp	174.86	174.87	174.88	174.91

Table 27 Calculated pressure-dependent chemical shifts for the ^{13}C amide protons for the seven and eight minima of Ac-Ala-NHMe and Ac-Gly-NHMe, respectively. The calculations were performed with MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Conformer	1 bar	1 kbar	2 kbar	3 kbar
Ac-Ala-NHMe				
P_{II}	131.14	131.31	131.48	131.58
α_R	129.42	129.60	129.77	129.92
C_5	128.51	128.68	128.83	128.96
$C_{7,eq}$	135.59	135.75	135.90	136.01
α_L	125.12	125.35	125.53	125.69
a_d	126.99	127.17	127.34	127.49
$C_{7,ax}$	126.32	126.48	126.65	126.79
Exp	110.13	110.16	110.27	110.41
P_{II}	108.14	108.37	108.64	108.83
α_R	105.93	106.22	106.47	106.68
C_5	106.83	107.08	107.27	107.45
$C_{7,eq}$	112.46	112.70	112.90	113.08
α_L	103.74	104.01	104.24	104.43
a_d	108.65	108.94	109.19	109.40
$C_{7,ax}$	112.99	113.25	113.49	113.70
Exp	107.38	107.43	107.65	1007.9
Ac-Gly-NHMe				
α_R	113.55	113.71	113.86	114.00
α_L	113.61	113.85	113.92	114.08
P_{II}	113.50	113.62	113.73	113.88
P_{II}	113.43	113.58	113.74	113.87
β	115.59	115.69	115.88	116.05
C_5	118.89	119.01	119.15	119.28
$C_{7,eq}$	118.72	118.84	118.96	119.16
$C_{7,eq}$	113.12	113.28	113.44	113.59
Exp	114.83	114.87	115.02	115.2
α_R	107.46	107.66	107.88	108.06
α_L	107.46	107.65	107.88	108.09
P_{II}	106.96	107.08	107.16	107.38
P_{II}	106.78	106.96	107.15	107.29
β	107.30	107.62	107.87	108.10
C_5	112.59	112.75	112.95	113.11
$C_{7,eq}$	112.84	113.07	113.27	113.40
$C_{7,eq}$	103.97	104.20	104.37	104.51
Exp	108.29	108.34	108.55	108.79

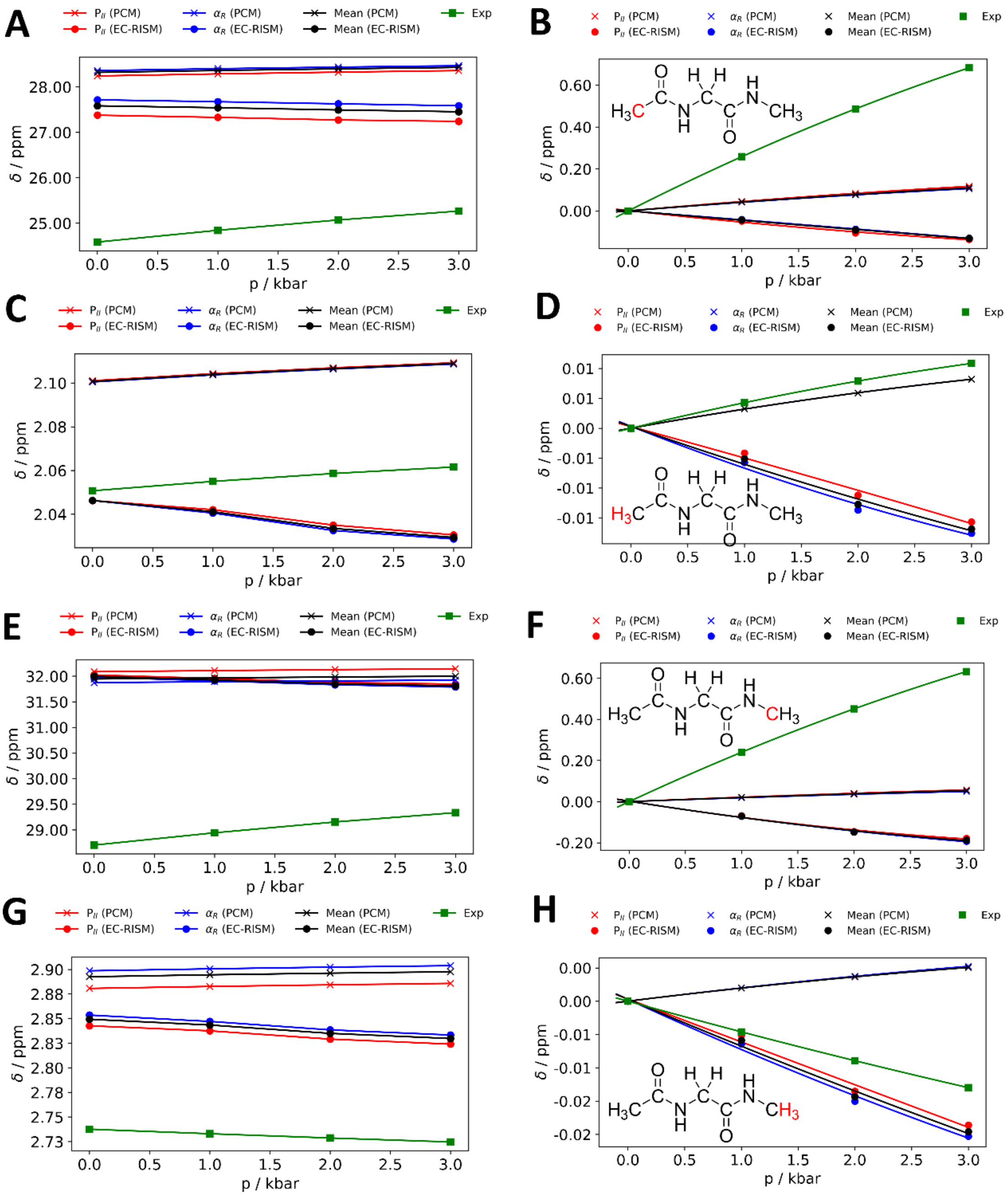

Figure 59 Pressure-dependent NMR-Shifts of the two Ac-Gly-NHMe main conformers compared to the experimental value for the ^{1}H and ^{13}C methyl peptide bond nuclei calculated with MP2/6-311+G(d,p). A, C, E, and G show the absolute shifts, whereas B, D, F, and H correspond to pressure-dependent changes. The red lines show the P$_{II}$ conformer, and the blue lines show the α conformers. The green line represents the experimental shifts. The + denotes conformers optimized with B3LYP/6-311+G(d,p)/PCM(1 bar), and the dot points represent the conformers optimized with B3LYP/6-311+G(d,p)/EC-RISM(p-dep).

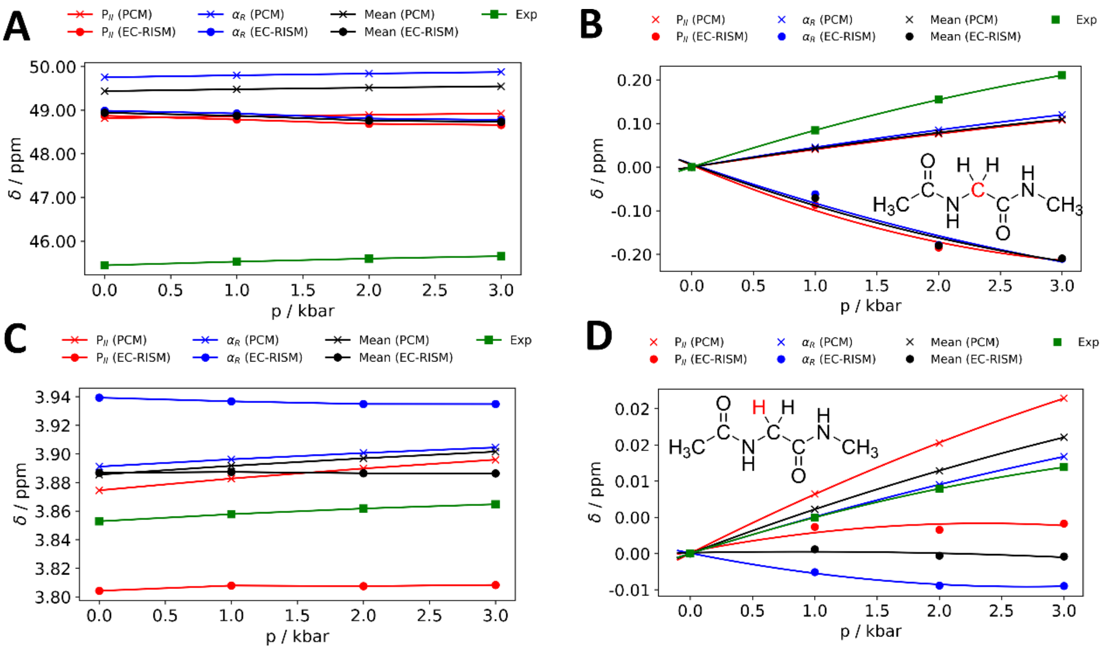

Figure 60 Pressure-dependent NMR-Shifts of the two Ac-Gly-NHMe main conformers compared to the experimental value for the ^{1}H and ^{13}C Cα peptide bond nuclei calculated with MP2/6-311+G(d,p). A and C show the absolute shifts, whereas B and D correspond to pressure-dependent changes. The red lines show the P_{II} conformer, and the blue lines show the α conformers. The green line represents the experimental shifts. The + denotes conformers optimized with B3LYP/6-311+G(d,p)/PCM(1 bar), and the dot points represent the conformers optimized with B3LYP/6-311+G(d,p)/EC-RISM(p-dep).

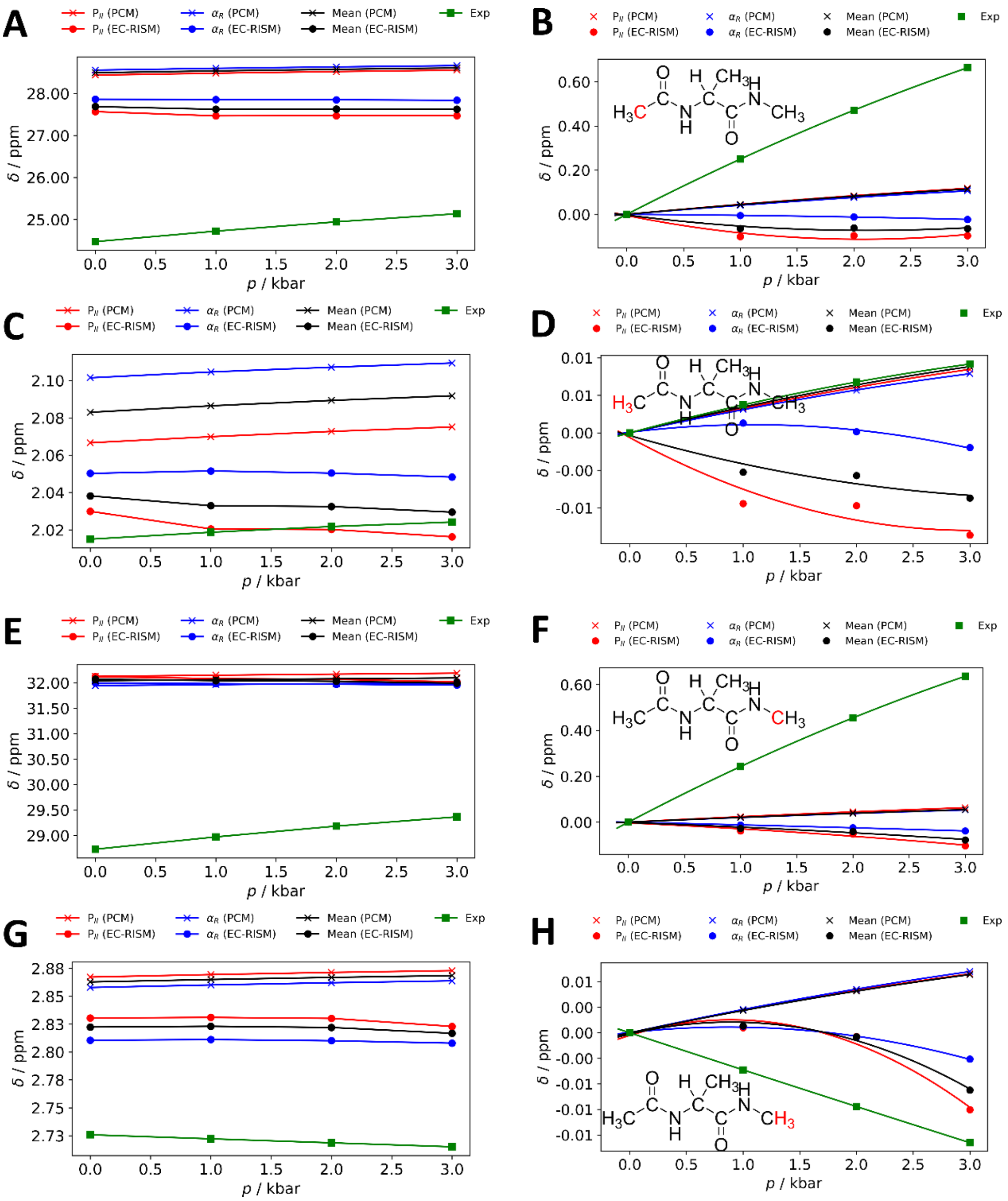

Figure 61 Pressure-dependent NMR-Shifts of the two Ac-Gly-NHMe main conformers compared to the experimental value for the ^1H and ^{13}C methyl peptide bond nuclei calculated with MP2/6-311+G(d,p). A, C, E, and G show the absolute shifts, whereas B, D, F, and H correspond to pressure-dependent changes. The red lines show the P_{II} conformer, and the blue lines show the α conformers. The green line represents the experimental shifts. The + denotes conformers optimized with B3LYP/6-311+G(d,p)/PCM(1 bar), and the dot points represent the conformers optimized with B3LYP/6-311+G(d,p)/EC-RISM(p-dep).

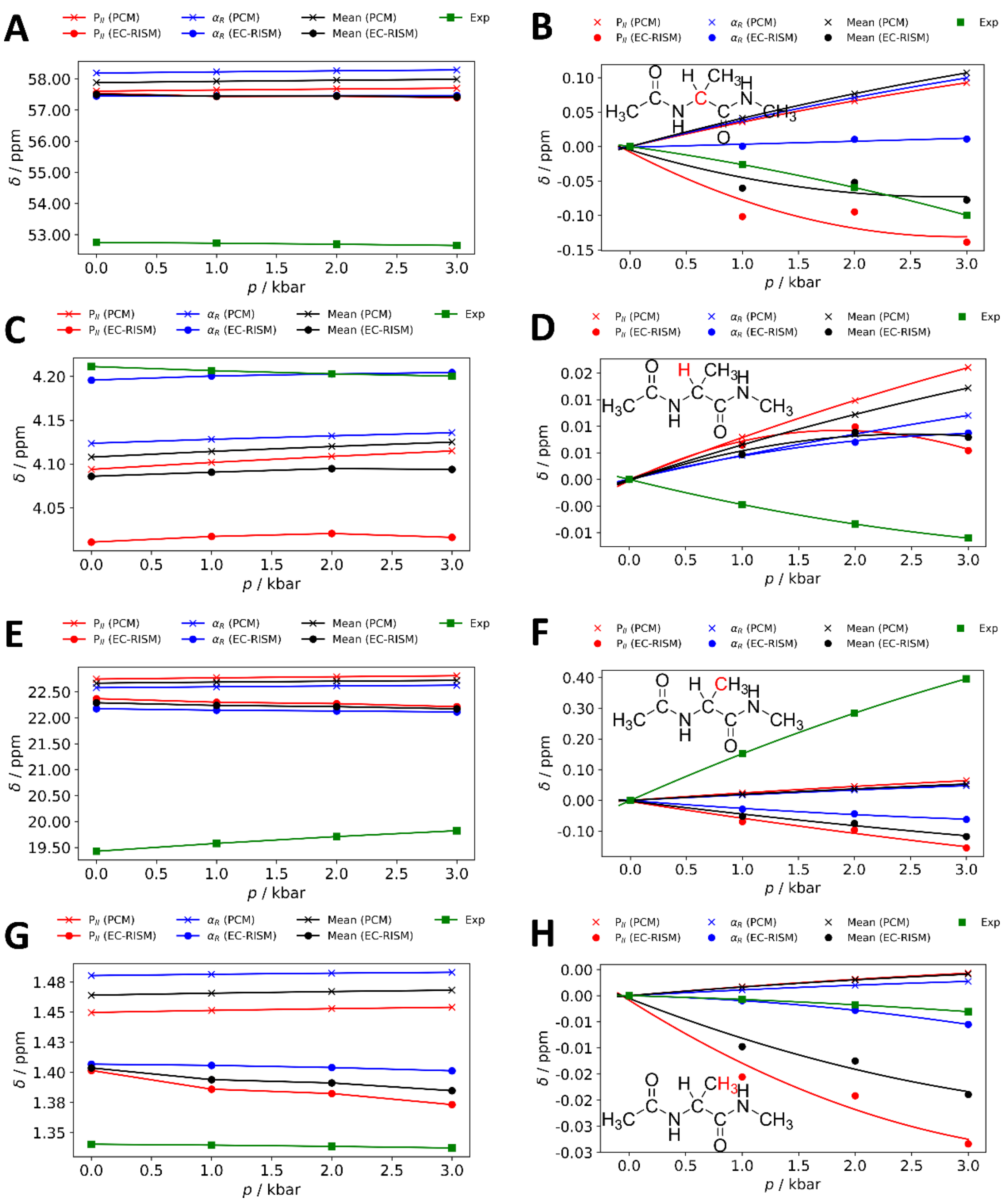

Figure 62 Pressure-dependent NMR-Shifts of the two Ac-Gly-NHMe main conformers compared to the experimental value for the ^{1}H and ^{13}C C$_\alpha$ peptide bond nuclei calculated with MP2/6-311+G(d,p). A, C, E, and G show the absolute shifts, whereas B, D, F, and H correspond to pressure-dependent changes. The red lines show the P$_{II}$ conformer, and the blue lines show the α conformers. The green line represents the experimental shifts. The + denotes conformers optimized with B3LYP/6-311+G(d,p)/PCM(1 bar), and the dot points represent the conformers optimized with B3LYP/6-311+G(d,p)/EC-RISM(p-dep).

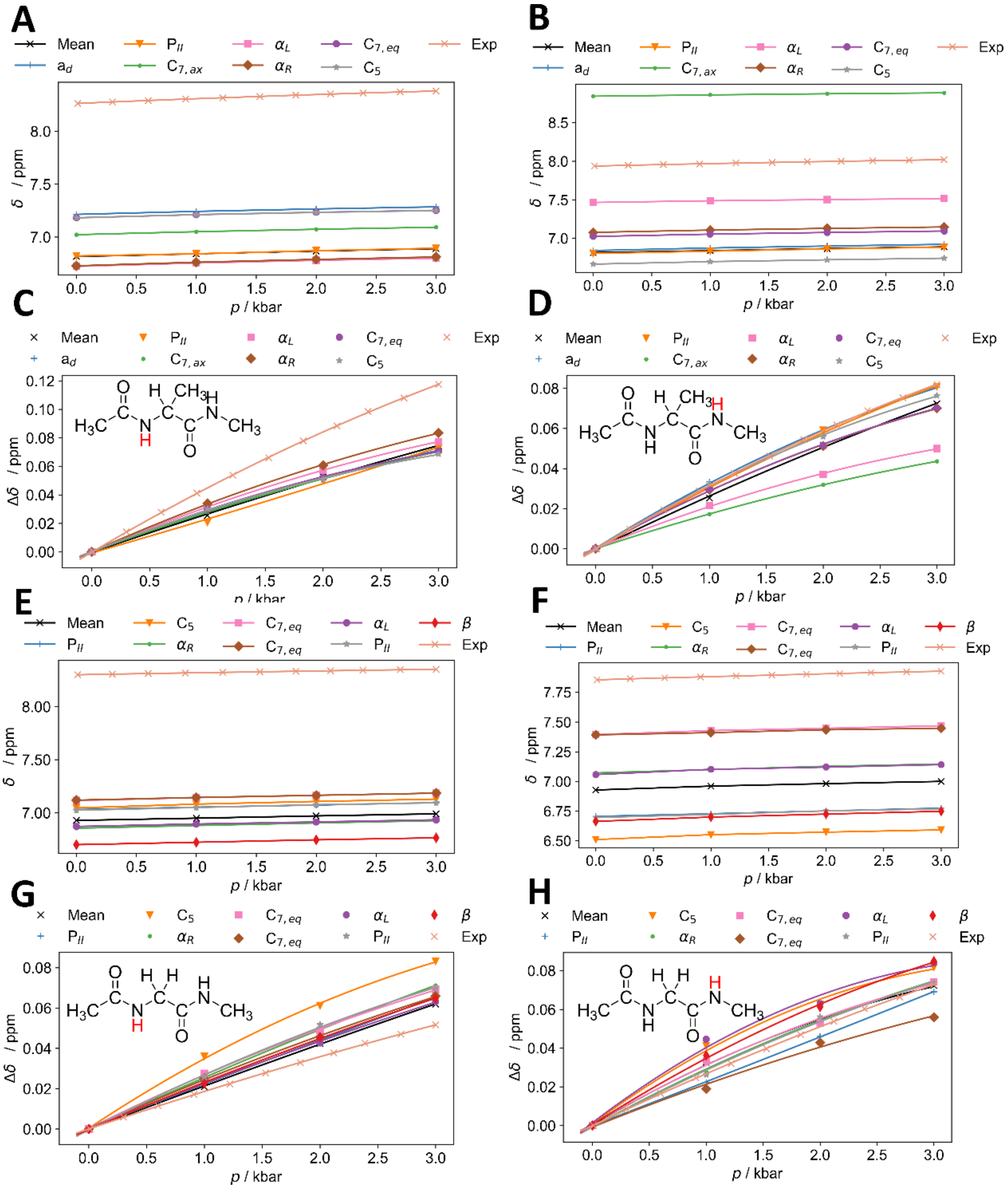

Figure 63 Pressure-dependent NMR shifts of the seven and eight Ac-Ala-NHMe and Ac-Gly-NHMe minima compared to the experimental value for the ^{13}C carbonyl atoms of the two peptide bonds. A and B represent the absolute chemical shifts of the two amide hydrogens atoms of Ac-Ala-NHMe, whereas C and D represent the pressure-dependent changes in chemical shifts relative to 1 bar. In E and F, the absolute values for the corresponding nuclei of Ac-Gly-NHMe are shown, and in G and H, the difference relative to ambient pressure. The theoretical calculations were performed with MP2/6-311+G(d,p)/EC-RISM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

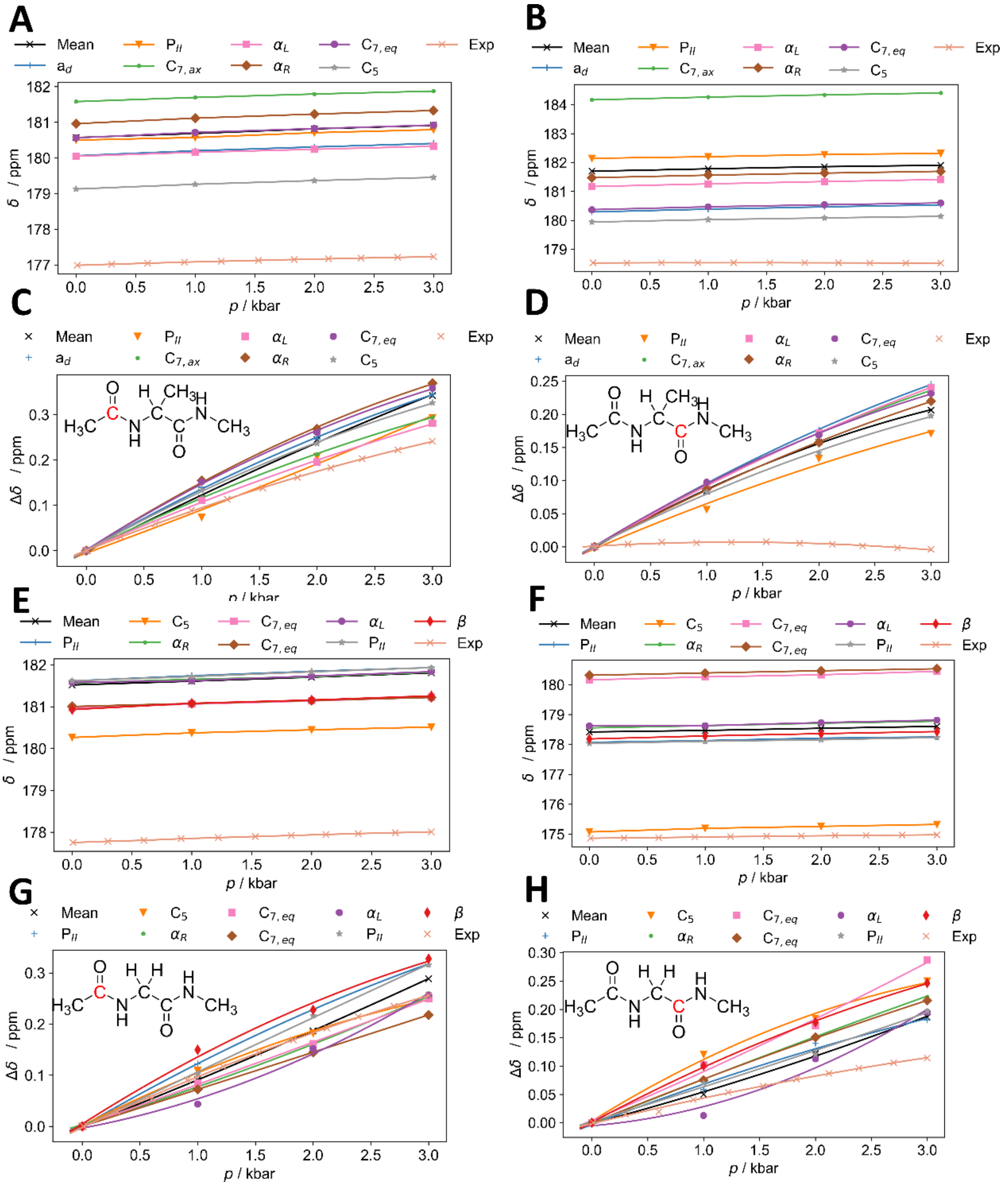

Figure 64 Pressure-dependent NMR shifts of the seven and eight Ac-Ala-NHMe and Ac-Gly-NHMe minima compared to the experimental value for the ^{13}C carbonyl atoms of the two peptide bonds. A and B represent the absolute chemical shifts of the two amide hydrogens atoms of Ac-Ala-NHMe, whereas C and D represent the pressure-dependent changes in chemical shifts relative to 1 bar. In E and F, the absolute values for the corresponding nuclei of Ac-Gly-NHMe are shown, and in G and H, the difference relative to ambient pressure. The theoretical calculations were performed with MP2/6-311+G(d,p)/EC-RISM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

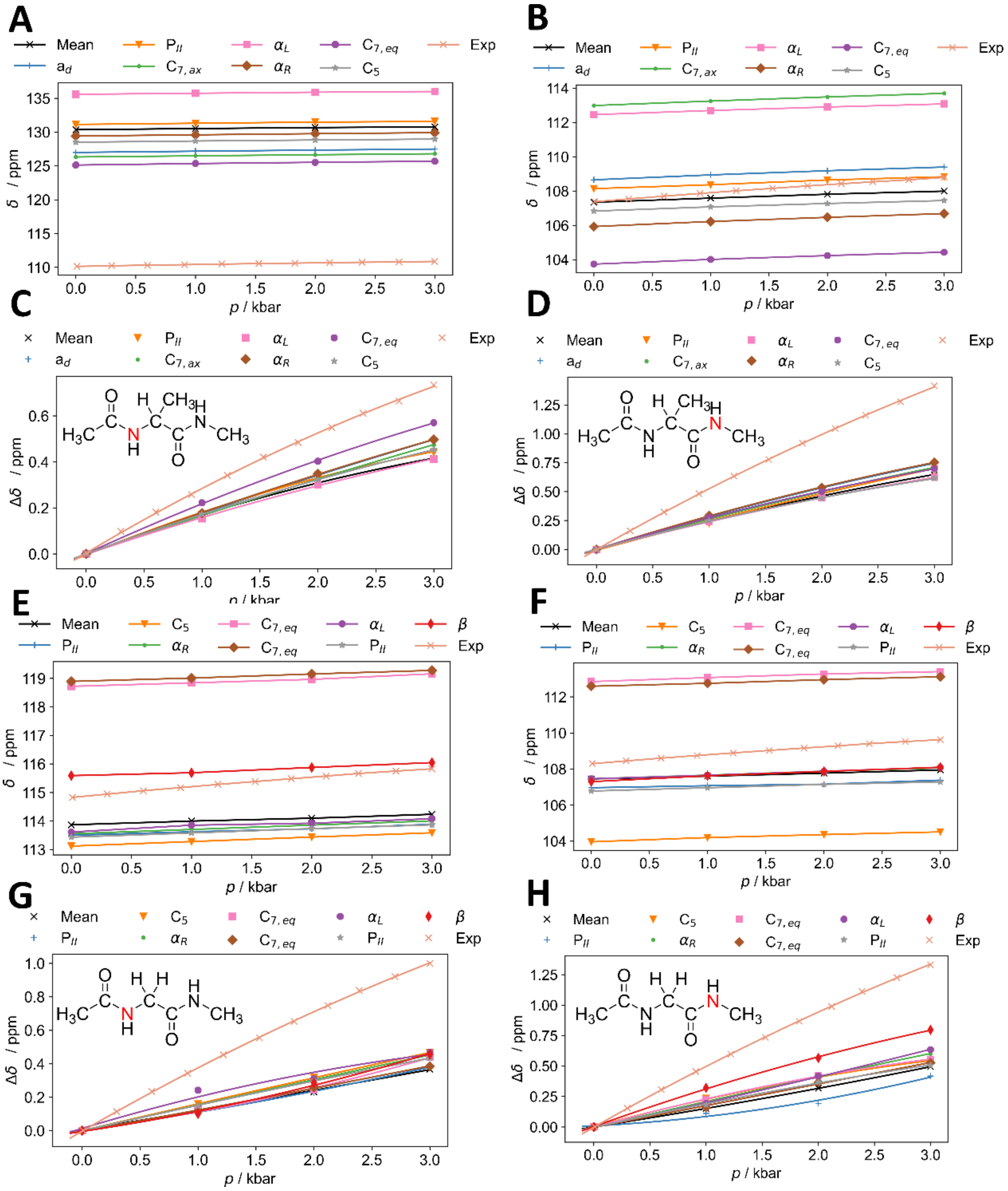

Figure 65 Pressure-dependent NMR shifts of the seven and eight Ac-Ala-NHMe and Ac-Gly-NHMe minima compared to the experimental value for the ^{13}C carbonyl atoms of the two peptide bonds. A and B represent the absolute chemical shifts of the two amide hydrogens atoms of Ac-Ala-NHMe, whereas C and D represent the pressure-dependent changes in chemical shifts relative to 1 bar. In E and F, the absolute values for the corresponding nuclei of Ac-Gly-NHMe are shown, and in G and H, the difference relative to ambient pressure is presented. The theoretical calculations were performed with MP2/6-311+G(d,p)/EC-RISM. The corresponding reference shielding constants can be found in Table 43. Raw chemical shielding constants can be found in SI_4.3. Experimental values were measured by M. Beck Erlach[298].

Table 28 Linear (B_1, in ppm kbar^{-1}) and quadratic (B_2, in ppm kbar^{-2}) coefficients from fitting experimental and calculated chemical shifts (MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM) of the seven minimum conformers of Ac-Ala-NHMe to the following form: $\delta(p) = \delta_0 + B_1 p + B_2 p^2$. Also, the 1 bar chemical shifts (δ_0 in ppm) are shown. The letters A and B refer to the corresponding subfigures in Figure 26, Figure 27, and Figure 28. Experimental values were measured by M. Beck Erlach[298].

Nucleus	δ_0.	B_1	B_2	δ_0.	B_1	B_2
^1H	A	A	A	B	B	B
Exp	8.26374	0.04798	-0.00290	7.93701	0.03374	-0.00210
P_{II}	6.69526	0.02283	-0.00127	6.63016	0.02273	-0.00132
α_R	6.57317	0.02453	-0.00142	7.01358	0.02226	-0.00134
C_5	7.12644	0.02299	-0.00127	6.47638	0.02315	-0.00136
α_L	6.60058	0.02241	-0.00127	7.48794	0.01397	-0.00081
$C_{7,eq}$	7.02875	0.02105	-0.00122	6.98228	0.02248	-0.00132
a_d	7.08987	0.02059	-0.00119	6.69339	0.02528	-0.00149
$C_{7,ax}$	6.90324	0.02003	-0.00115	8.71119	0.01114	-0.00071
Mean	6.67458	0.02150	-0.00111	6.82467	0.02028	-0.00067
^{13}C	A	A	A	B	B	B
Exp	176.98931	0.10144	-0.00711	178.53068	0.00961	-0.00383
P_{II}	179.67637	0.16376	-0.01001	182.09015	0.12033	-0.00684
α_R	180.17884	0.16323	-0.01020	181.31879	0.13402	-0.00775
C_5	177.83741	0.15047	-0.00927	179.83551	0.12137	-0.00683
α_L	179.67690	0.14050	-0.00877	181.15593	0.13350	-0.00782
$C_{7,eq}$	180.02568	0.16911	-0.01041	180.34530	0.13965	-0.00822
a_d	179.39570	0.16144	-0.00980	180.10683	0.13691	-0.00801
$C_{7,ax}$	181.09666	0.14580	-0.00886	183.76433	0.13116	-0.00789
Mean	179.78788	0.14886	-0.00526	181.56664	0.11227	-0.00372
^{15}N	A	A	A	B	B	B
Exp	110.12974	0.30025	-0.01961	107.37677	0.55512	-0.02858
P_{II}	130.18457	0.18924	-0.01160	108.04669	0.34232	-0.02174
α_R	129.54492	0.25431	-0.01475	105.11784	0.31906	-0.02107
C_5	128.05768	0.23359	-0.01457	105.90772	0.29408	-0.01948
α_L	135.38387	0.22563	-0.01371	112.05921	0.28771	-0.01878
$C_{7,eq}$	125.27499	0.30200	-0.01763	103.10617	0.30904	-0.01954
a_d	125.86924	0.24792	-0.01450	108.35225	0.32255	-0.02020
$C_{7,ax}$	126.42804	0.24604	-0.01485	112.28392	0.30127	-0.01993
Mean	129.87973	0.17909	-0.00560	106.80060	0.25113	-0.00856

Table 29 Linear (B_1, in ppm kbar^{-1}) and quadratic (B_2, in ppm kbar^{-2}) coefficients from fitting experimental and calculated chemical shifts (MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM) of the eight minima conformers of Ac-Gly-NHMe to the following form: $\delta(p)= B_1p+B_2p^2$. Also, the 1 bar chemical shifts (δ_0 in ppm) are shown. The letters A and B refer to the corresponding subfigures in Figure 26, Figure 27, and Figure 28. Experimental values were measured by M. Beck Erlach[298].

Nucleus	δ_0.	B_1	B_2	δ_0.	B_1	B_2
^{1}H	A	A	A	B	B	B
Exp	8.29980	0.01930	-0.00069	7.85553	0.02800	--0.00123
α_R	6.73001	0.02075	-0.00121	6.98158	0.02142	-0.00132
α_L	6.72725	0.02088	-0.00124	6.97898	0.02127	-0.00129
P_{II}	6.84027	0.01956	-0.00112	6.49097	0.02264	-0.00135
P_{II}	6.84126	0.01951	-0.00111	6.49057	0.02264	-0.00135
P_{II}	6.58197	0.02080	-0.00122	6.37849	0.02137	-0.00127
$C_{7,eq}$	6.96461	0.01974	-0.00116	7.60453	0.01313	-0.00077
$C_{7,eq}$	6.96370	0.01973	-0.00113	7.61237	0.01312	-0.00078
C_5	6.97686	0.02384	-0.00130	6.35547	0.02030	-0.00118
Mean	6.76976	0.01695	-0.00054	6.82884	0.01605	-0.00055
^{13}C	A	A	A	B	B	B
Exp	177.75101	0.10575	-0.00655	174.86069	0.04840	-0.00318
α_R	180.91498	0.17290	-0.01079	178.34786	0.15253	-0.00876
α_L	180.89169	0.17285	-0.01076	178.33451	0.15228	-0.00871
P_{II}	180.82311	0.16884	-0.01021	177.75601	0.14812	-0.00828
P_{II}	180.84264	0.16911	-0.01025	177.76538	0.14862	-0.00831
P_{II}	180.38256	0.16959	-0.01022	177.61058	0.15136	-0.00851
$C_{7,eq}$	180.85378	0.14455	-0.00894	180.51209	0.14158	-0.00831
$C_{7,eq}$	180.87157	0.14476	-0.00894	180.56715	0.14159	-0.00836
C_5	178.88563	0.14495	-0.00899	175.00236	0.14598	-0.00816
Mean	180.84278	0.14949	-0.00518	178.21565	0.12529	-0.00396
^{15}N	A	A	A	B	B	B
Exp	114.82558	0.40046	-0.02163	108.29265	0.52527	-0.02634
α_R	113.78701	0.28735	-0.01651	106.50342	0.29314	-0.01998
α_L	113.78709	0.28776	-0.01649	106.46660	0.29114	-0.01985
P_{II}	112.92306	0.22749	-0.01371	105.38239	0.29251	-0.01925
P_{II}	112.95394	0.22691	-0.01369	105.38488	0.29349	-0.01932
P_{II}	114.00139	0.23502	-0.01421	103.34730	0.27075	-0.01796
$C_{7,eq}$	119.05126	0.26019	-0.01546	112.54257	0.29526	-0.01949
$C_{7,eq}$	119.05858	0.26112	-0.01554	112.60256	0.29508	-0.01951
C_5	112.87144	0.26231	-0.01641	103.60936	0.26618	-0.01773
Mean	113.81459	0.21232	-0.00663	106.28694	0.22376	-0.00794

6.6 Pressure-dependent parameters of Ac-Gly/Ala-NHMe based on B3LYP/6-311+G(d,p)/EC-RISM optimizations

Table 30 Pressure-dependent bond lengths of the two C-O amide bonds and two N-H amide bonds of the seven minima of Ac-Ala-NHMe. Calculations were performed with B3LYP/6-311+G(d,p)/EC-RISM. Structures can be found in SI_4.3.

C-O 1						
Conformer	1 bar	1 kbar	2 kbar	3 kbar	4 kbar	5 kbar
a_d	1.2452	1.2455	1.2457	1.2458	1.2459	1.2460
P_{II}	1.2461	1.2463	1.2466	1.2468	1.2468	1.2470
$C_{7,ax}$	1.2482	1.2484	1.2486	1.2487	1.2488	1.2489
$C_{7,eq}$	1.2482	1.2485	1.2486	1.2487	1.2485	1.2486
α_R	1.2438	1.2443	1.2445	1.2447	1.2448	1.2450
α_L	1.2436	1.2440	1.2442	1.2444	1.2444	1.2446
C_5	1.2489	1.2493	1.2494	1.2496	1.2497	1.2498
C-O.-2						
a_d	1.2461	1.2465	1.2467	1.2468	1.2469	1.2470
P_{II}	1.2443	1.2446	1.2449	1.2450	1.2452	1.2453
$C_{7,ax}$	1.2491	1.2494	1.2496	1.2497	1.2499	1.2500
$C_{7,eq}$	1.2440	1.2443	1.2445	1.2447	1.2447	1.2450
α_R	1.2473	1.2477	1.2479	1.2480	1.2481	1.2483
α_L	1.2445	1.2449	1.2450	1.2452	1.2452	1.2453
C_5	1.2440	1.2443	1.2445	1.2446	1.2447	1.2448
a_d	1.2461	1.2465	1.2467	1.2468	1.2469	1.2470
N-H 1						
a_d	1.0136	1.0139	1.0141	1.0142	1.0141	1.0142
P_{II}	1.0148	1.0147	1.0150	1.0152	1.0153	1.0154
$C_{7,ax}$	1.0140	1.0142	1.0144	1.0146	1.0147	1.0148
$C_{7,eq}$	1.0155	1.0158	1.0159	1.0160	1.0157	1.0159
α_R	1.0158	1.0161	1.0162	1.0164	1.0163	1.0164
α_L	1.0142	1.0145	1.0146	1.0147	1.0145	1.0147
C_5	1.0148	1.0151	1.0151	1.0152	1.0153	1.0153
N-H2						
a_d	1.0115	1.0118	1.0119	1.0120	1.0119	1.0120
P_{IO}	1.0146	1.0147	1.0149	1.0150	1.0151	1.0152
$C_{7,ax}$	1.0179	1.0180	1.0181	1.0182	1.0182	1.0182
$C_{7,eq}$	1.0147	1.0149	1.0150	1.0150	1.0145	1.0147
α_R	1.0134	1.0137	1.0138	1.0139	1.0139	1.0140
α_L	1.0132	1.0134	1.0136	1.0137	1.0141	1.0139
C_5	1.0148	1.0151	1.0152	1.0153	1.0153	1.0154

Table 31 Pressure-dependent bond lengths of the two C-O amide bonds and two N-H amide bonds of the seven minima of Ac-Ala-NHMe. Calculations were performed with B3LYP/6-311+G(d,p)/EC-RISM. Structures can be found in SI_4.3.

Conformer	1 bar	1 kbar	2 kbar	3 kbar	4 kbar	5 kbar
C-O 1						
P_{II}	1.2453	1.2455	1.2458	1.2459	1.2459	1.2462
C_5	1.2498	1.2501	1.2502	1.2503	1.2504	1.2505
α_R	1.2430	1.2431	1.2433	1.2436	1.2438	1.2440
$C_{7,eq}$	1.2473	1.2475	1.2476	1.2478	1.2479	1.2480
$C_{7,eq}$	1.2474	1.2475	1.2476	1.2478	1.2478	1.2479
α_L	1.2430	1.2433	1.2434	1.2437	1.2438	1.2440
P_{II}	1.2452	1.2453	1.2457	1.2458	1.2460	1.2461
β	1.2456	1.2459	1.2459	1.2461	1.2463	1.2465
C-O.-2						
P_{II}	1.2437	1.2439	1.2441	1.2443	1.2443	1.2445
C_5	1.2421	1.2424	1.2424	1.2425	1.2425	1.2426
α_R	1.2478	1.2479	1.2483	1.2485	1.2486	1.2488
$C_{7,eq}$	1.2464	1.2466	1.2469	1.2470	1.2472	1.2473
$C_{7,eq}$	1.2464	1.2466	1.2467	1.2470	1.2470	1.2471
α_L	1.2480	1.2481	1.2482	1.2486	1.2486	1.2488
P_{II}	1.2437	1.2438	1.2441	1.2442	1.2444	1.2444
β	1.2436	1.2441	1.2442	1.2444	1.2445	1.2448
P_{II}	1.2437	1.2439	1.2441	1.2443	1.2443	1.2445
N-H 1						
P_{II}	1.0137	1.0138	1.0139	1.0140	1.0140	1.0140
C_5	1.0148	1.0149	1.0150	1.0151	1.0151	1.0152
α_R	1.0150	1.0151	1.0152	1.0153	1.0153	1.0154
$C_{7,eq}$	1.0144	1.0145	1.0146	1.0146	1.0147	1.0148
$C_{7,eq}$	1.0144	1.0145	1.0146	1.0146	1.0147	1.0147
α_L	1.0149	1.0151	1.0151	1.0152	1.0153	1.0153
P_{II}	1.0137	1.0138	1.0139	1.0140	1.0141	1.0141
N-H2						
P_{II}	1.0146	1.0147	1.0148	1.0149	1.0150	1.0150
C_5	1.0150	1.0151	1.0151	1.0152	1.0153	1.0153
α_R	1.0137	1.0139	1.0140	1.0141	1.0141	1.0142
$C_{7,eq}$	1.0147	1.0147	1.0147	1.0148	1.0148	1.0148
$C_{7,eq}$	1.0147	1.0147	1.0147	1.0148	1.0148	1.0148
α_L	1.0137	1.0138	1.0140	1.0140	1.0142	1.0142
P_{II}	1.0146	1.0147	1.0148	1.0149	1.0150	1.0150
β	1.0146	1.0147	1.0148	1.0149	1.0149	1.0149

Table 32 Pressure-dependent populations of Ac-Ala-NHMe for the seven minima conformers, calculated with B3LYP/6-311+G(d,p)//B3LYP/6-311+G(d,p)/EC-RISM and MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p) at T=298.15 K. Additionally, the MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM populations are shown.. The population values are in %. The pressure-dependent PMV correction was utilized, and no thermal corrections were added. Raw data can be found in SI_4.3.

Conformer	1 bar	1 kbar	2 kbar	3 kbar	4 kbar	5 kbar
MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM						
P_{II}	46.53	46.17	45.84	45.55	45.29	45.06
α_R	41.17	42.12	42.84	43.39	43.84	44.20
C_5	5.55	5.33	5.17	5.05	4.95	4.87
$C_{7,eq}$	3.98	3.60	3.35	3.17	3.04	2.94
α_L	1.84	1.84	1.86	1.87	1.90	1.92
a_d	0.73	0.76	0.79	0.82	0.84	0.87
$C_{7,ax}$	0.20	0.18	0.16	0.15	0.14	0.14
B3LYP/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM						
P_{II}	38.58	37.92	37.74	37.48	37.50	37.34
α_R	47.12	48.32	48.99	49.61	49.84	50.24
C_5	11.25	10.92	10.59	10.34	10.15	9.96
$C_{7,eq}$	2.27	2.05	1.89	1.77	1.71	1.65
α_L	0.48	0.49	0.49	0.50	0.50	0.50
a_d	0.23	0.24	0.24	0.25	0.25	0.26
$C_{7,ax}$	0.07	0.06	0.06	0.05	0.05	0.05
MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM						
P_{II}	52.31	53.51	53.22	52.94	52.69	52.41
α_R	35.78	35.75	36.53	37.15	37.41	37.84
C_5	5.82	5.33	5.11	4.94	4.79	4.69
$C_{7,eq}$	3.19	2.60	2.29	2.08	2.13	2.01
α_L	1.73	1.66	1.66	1.67	1.71	1.74
a_d	1.01	1.02	1.07	1.12	1.18	1.23
$C_{7,ax}$	0.17	0.13	0.11	0.10	0.09	0.08

Table 33 Pressure-dependent populations of Ac-Gly-NHMe for the eight minima conformers, calculated with B3LYP/6-311+G(d,p)//B3LYP/6-311+G(d,p)/EC-RISM at T=298.15 K The population values are in %. The pressure-dependent PMV correction was utilized.

Conformer	1 bar	1 kbar	2 kbar	3 kbar	4 kbar	5 kbar
MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/PCM						
α_R	30.08	30.34	30.58	30.63	30.64	30.64
α_L	28.48	28.75	29.03	29.10	29.13	29.13
P_{II}	15.01	15.04	15.11	15.16	15.22	15.22
P_{II}	14.98	15.00	15.06	15.10	15.16	15.16
β	5.27	5.34	5.42	5.45	5.48	5.48
C_5	0.94	0.89	0.83	0.81	0.80	0.80
$C_{7,eq}$	2.63	2.33	1.99	1.88	1.79	1.79
$C_{7,eq}$	2.61	2.31	1.97	1.86	1.78	1.78
B3LYP/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM						
α_R	29.69	30.07	29.89	29.95	30.03	30.20
α_L	28.87	28.85	29.44	29.50	29.66	29.80
P_{II}	13.39	13.35	13.23	13.16	13.21	13.11
P_{II}	13.24	13.24	12.96	13.00	12.75	13.01
β	9.05	9.14	9.38	9.49	9.61	9.24
C_5	4.44	4.20	4.05	3.93	3.84	3.77
$C_{7,eq}$	0.67	0.59	0.53	0.49	0.46	0.44
$C_{7,eq}$	0.65	0.57	0.52	0.48	0.45	0.43
MP2/6-311+G(d,p)/EC-RISM//B3LYP/6-311+G(d,p)/EC-RISM						
α_R	28.33	28.29	28.57	28.77	28.76	29.00
α_L	26.88	28.42	27.84	27.75	27.79	27.37
P_{II}	18.03	17.59	17.52	17.88	17.85	18.27
P_{II}	16.72	16.58	17.30	17.13	17.37	17.41
β	4.27	4.31	4.50	4.61	4.68	4.58
C_5	1.12	1.04	0.99	0.94	0.90	0.88
$C_{7,eq}$	2.34	1.89	1.65	1.47	1.33	1.25
$C_{7,eq}$	2.30	1.88	1.63	1.46	1.31	1.24

6.7 HFCC parameters of the nitroxy oxygen of HMI

Table 34 Summarized A^{iso} parameters in MHz for HMI and HHMI. On the one hand, results are shown for the revPBE0-D3/def2-TZVPP/CPCM optimized structure, calculated with revPBE0/def2-TZVPP/CPCM, revPBE0/def2-TZVPP/EC-RISM, and DLPNO-CCSD/def2-TZVPP/EC-RISM. On the other hand, mean A^{iso} parameters are shown that were calculated on the snapshots obtained from the AIMD.

Method	A^{iso}(HMI)	A^{iso}(HHMI)	ΔA^{iso}
Based on revPBE0-D3/def2-TZVPP optimized geometry			
revPBE0/def2-TZVPP/CPCM	-39.4	-40.9	1.5
revPBE0/def2-TZVPP/EC-RISM	-37.6	-40.3	2.7
DLPNO-CCSD/def2TZVPP/EC-RISM	-47.9	-54.2	6.3
Based on the full trajectory			
revPBE0/def2-TZVPP/CPCM/VD	-39.2±0.03	-40.8±0.02	1.6
revPBE0/def2-TZVPP/EC-RISM/VD	-37.4±0.05	-40.0±0.02	2.6
revPBE0/def2-TZVPP/CPCM/SSS	-37.3±0.07	-39.8±0.02	2.5
revPBE0/def2-TZVPP/EC-RISM/SSS	-36.6±0.06	-39.2±0.05	2.6
DLPNO-CCSD/def2-TZVPP/EC-RISM	-47.8±0.06	-53.7±0.05	5.9

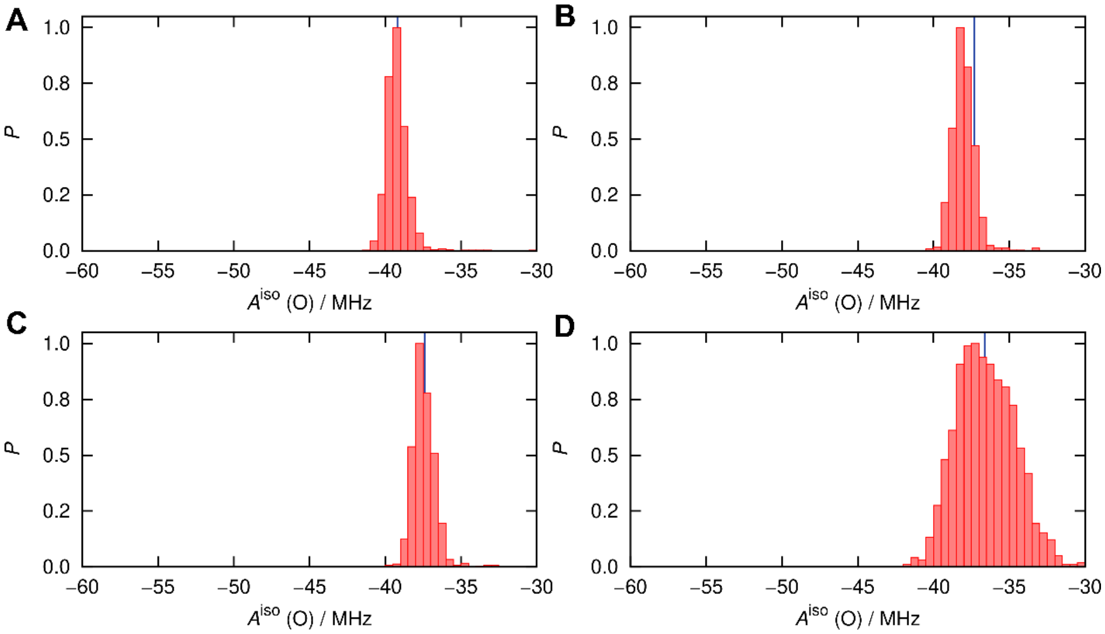

Figure 66 Normalized probability histograms of nitroxy oxygen A^{iso} values of the EPR spin-probe HMI at ambient conditions calculated with revPBE0-D3/def2-TZVPP with decontrated s function on 1000 snapshots. Panel (A) shows the calculations on the vertical desolvated set calculated with revPBE0/def2-TZVPP/CPCM. In Panel (B), the CPCM calculations on the subset with the second solvation shell are included. The two lower panels show the calculations with revPBE0/def2-TZVPP/EC-RISM on the VD set (C) and the SSS set (D). The blue lines in the panels depict the corresponding averaged value and be taken from Table 34. Material from ref [132]

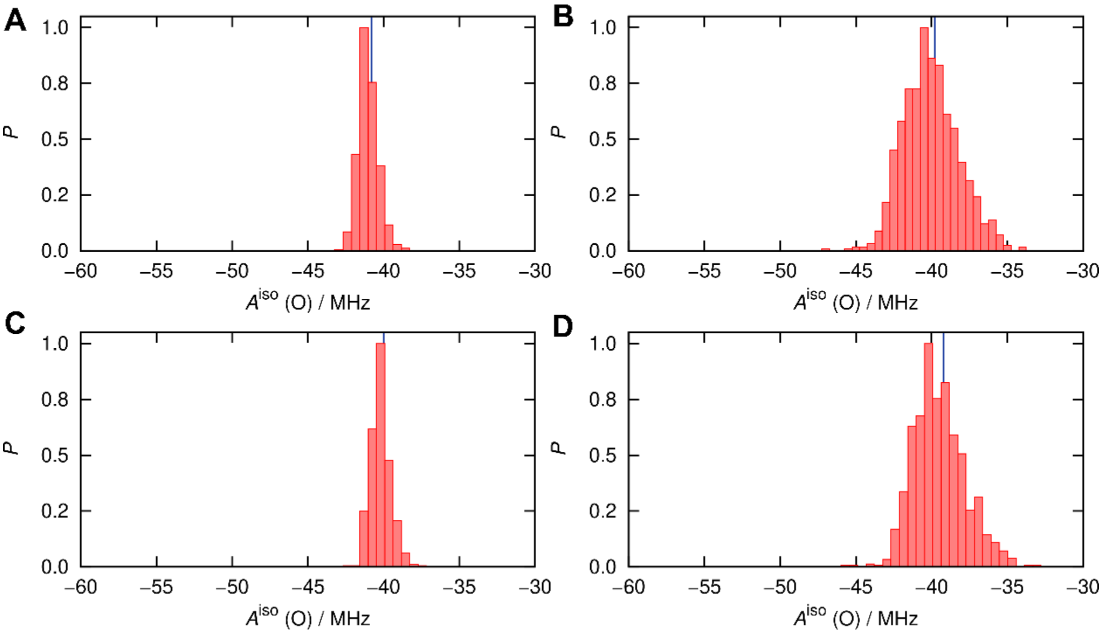

Figure 67 Normalized probability histograms of nitroxy oxygen A^{iso} values of the EPR spin-probe HMI at ambient conditions calculated with revPBE0-D3/def2-TZVPP with decontrated s function on 1000 snapshots. Panel (A) shows the calculations on the vertical desolvated set calculated with revPBE0/def2-TZVPP/CPCM. In Panel (B), the CPCM calculations on the subset with the second solvation shell are included. The two lower panels show the calculations with revPBE0/def2-TZVPP/EC-RISM on the VD set (C) and the SSS set (D). The blue lines in the panels depict the corresponding averaged value and be taken from Table 34.

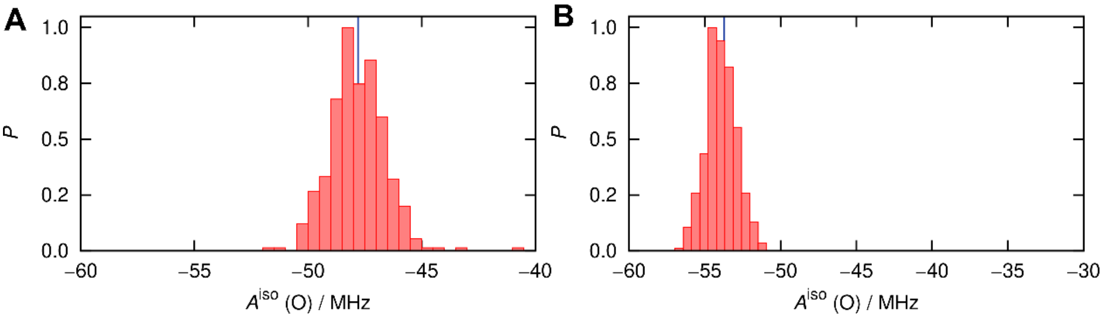

Figure 68 Normalized probability histograms of nitroxy oxygen A^{iso} values of the EPR spin-probe HMI (A) and HMHI (B) at ambient conditions calculated with revPBE0-D3/def2-TZVPP with decontrated s function on 400 snapshots. The blue lines in the panels depict the corresponding averaged value and be taken from Table 34. Panel (A) is taken from ref[132].

6.8 NMR fit results

To apply the NMR-based population fit, first, the calculated EC-RISM-based proton shifts need to be corrected by the following equation (co-developed by S. Kast and S. Maste):

$$\delta_{\text{corr}}(p) = \delta_{\text{orig}}(p) + a \cdot (p - p_0)^2 + b(p - p_0) + \delta_0. \qquad (190)$$

Here, δ_{corr} is the corrected chemical shift, a is a scaling parameter for a quadratic term, b is the linear parameter, and δ_0 is the correction of the ambient condition value. To determine the corresponding parameters, S. Maste measured the pressure-dependent chemical shifts of a dataset containing small amides and made the corresponding EC-RISM calculations. For the population fit, an average chemical shift based on the corrected chemical shifts is defined in the following way:

$$\delta_{\text{corr}}^{\text{mean}}(p) = \frac{1}{1+\exp\left(-\beta \Delta G(p_0) + s(p-p_0) + t(p-p_0)\right)} \delta_{\text{corr}}^{i} \\ +1 - \frac{1}{1+\exp\left(-\beta \Delta G(p_0) + s(p-p_0) + t(p-p_0)\right)} \delta_{\text{corr}}^{j}, \tag{191}$$

whereby ΔG is the free energy difference at 1 bar, s is a linear parameter, and t is the quadratic fit parameter. The loss function was minimized so that the quadratic difference between experimental and calculated shifts became minimal. The parameters obtained from the fit can be found in Table 35.

Table 35 Fitparamter obtained from NMR population fit performed by S. Maste.[300] ΔG_0 is in kcal mol^{-1}, s in kcal mol^{-1} bar^{-1}-and t in kcal mol^{-1} bar^{-2}.

Molecule/Method	ΔG_0	s	t
Ac-Ala-NHMe/ MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM	0.621	-0.035	0.0020
Ac-Gly-NHMe/ MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM	-4.164	0.220	-1.441
Ac-Ala-NHMe/ MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/EC-RISM	0.042	-0.347	0.106
Ac-Gly-NHMe/MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/EC-RISM	0.550	-0.133	0.025

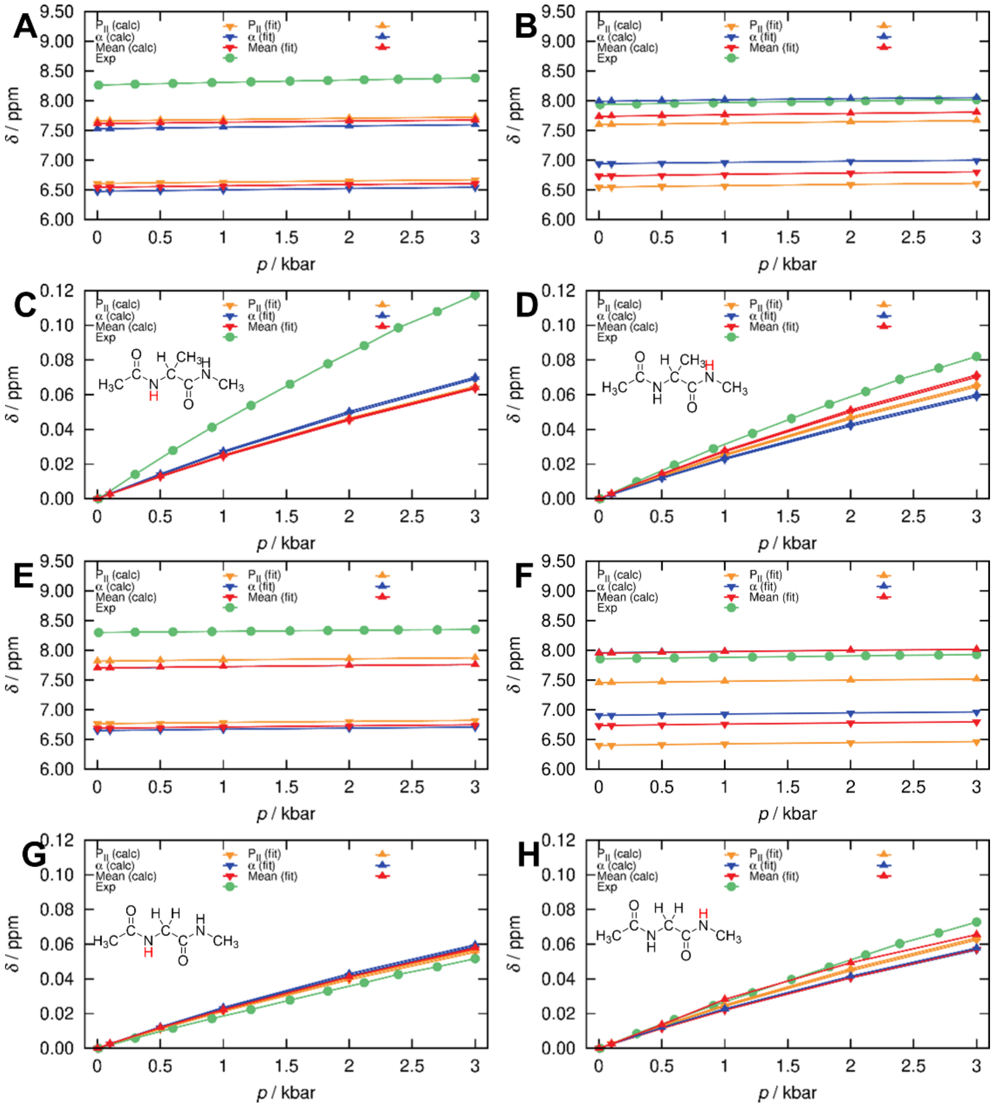

Figure 69 Pressure-dependent chemical shifts of Ac-Ala-NHMe and Ac-Gly-NHMe for the two main conformers P_{II} and α. In panels A and B, the absolute chemical shifts of Ac-Ala-NHMe are shown, and the corresponding pressure-dependent changes are depicted in panels C and D. The absolute chemical shifts of Ac-Gly-NHMe are depicted in panels E and F, and the pressure changes in G and H. The downward triangles show the pure MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/PCM results, whereas for the upward triangles, the chemical shifts correction scheme is applied. The NMR fits were provided by S. Maste.[300]

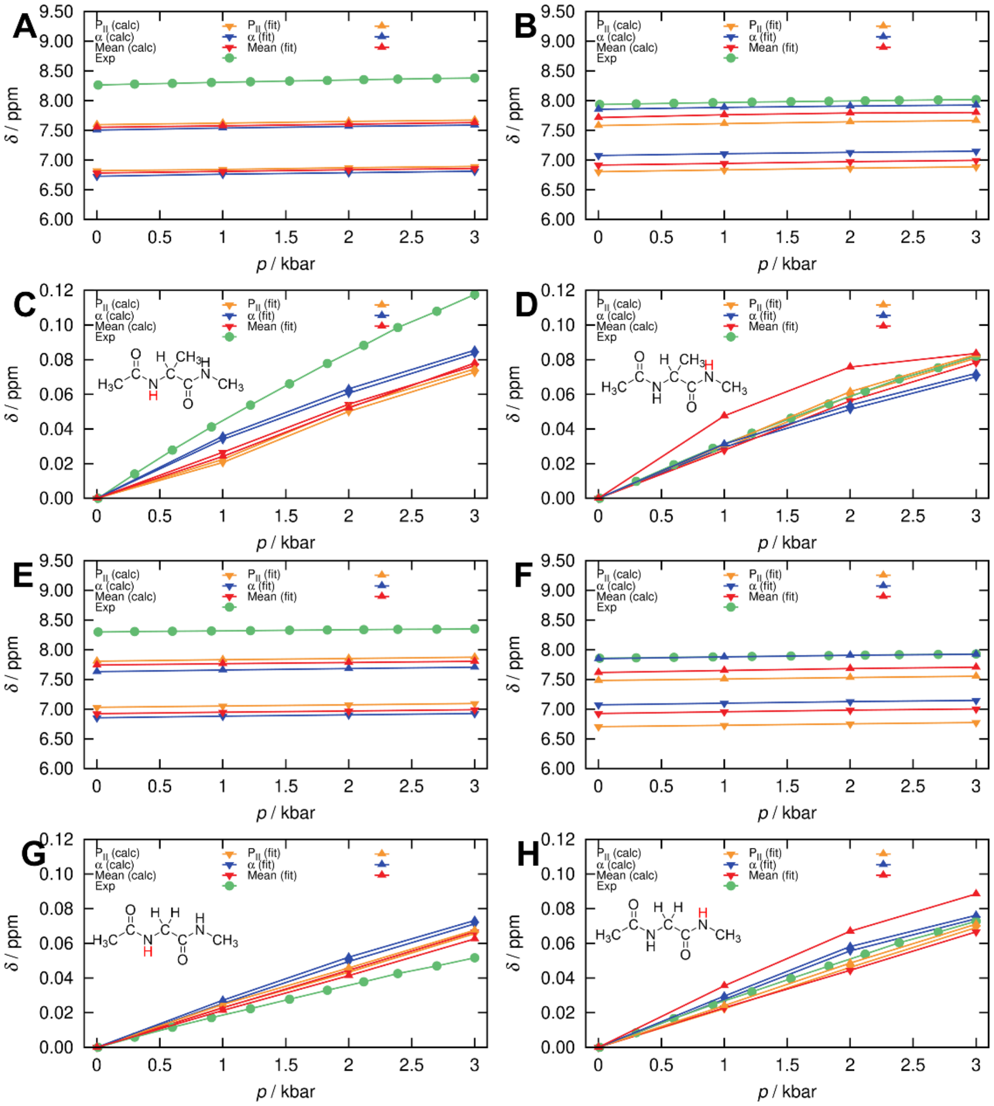

Figure 70 Pressure-dependent chemical shifts of Ac-Ala-NHMe and Ac-Gly-NHMe for the two main conformers P$_{II}$ and α. In panels A and B, the absolute chemical shifts of Ac-Ala-NHMe are shown, and the corresponding pressure-dependent changes are depicted in panels C and D. The absolute chemical shifts of Ac-Gly-NHMe are depicted in panels E and F, and the pressure changes in G and H. The downward triangles show the pure MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p)/EC-RISM results, whereas for the upward triangles, the chemical shifts correction scheme is applied. The NMR fits were provided by S. Maste.[300]

6.9 Lennard-Jones parameters of the solutes for EC-RISM calculations

Table 36 Lennard-Jones parameters of atoms used for TMAO calculations. The numbers in the brackets describe the corresponding GAFF[162]-type atoms.

Atom	σ (Å)	ε (kJ/mol)
N(4)	2.926	1.181109
C(3)	3.707	0.469665
O	3.266	1.06022
H(C)	2.130	0.128532

Table 37 Lennard-Jones parameters of atoms used for HMI and HHMI in EC-RISM calculations. The numbers in the brackets describe the corresponding GAFF-type atoms.

Atom	σ (Å)	ε (kJ/mol)
N(3)	3.245	1.181109
C(3)	3.399	0.760078
O	2.959922	1.459017
H(C)	2.649533	0.109079
H(1)	2.471353	0.109079

6.10 Force field parameters for FF+RISM calculations

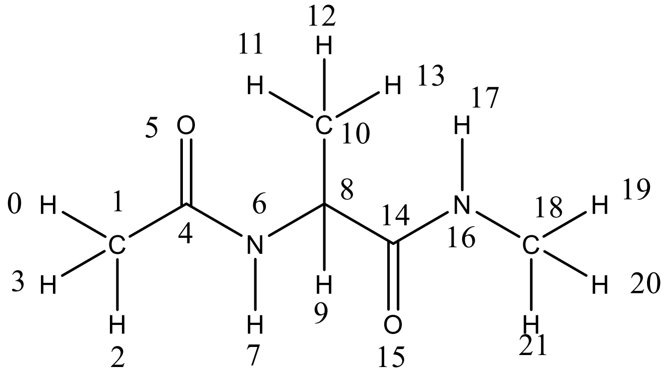

Figure 71: Assignment of atoms numbers to the corresponding atoms.

Table 38 Partial charges and Lennard-Jones parameters for the ff19SB+RISM calculations of Ac-Ala-NHMe.

Atom	Type	*q*	σ (Å)	*ε* (kJ/mol)
1	HC	0.1123	2.649533	0.109079
2	CT	-0.3662	3.39967	0.760078
3	HC	0.1123	2.649533	0.109079
4	HC	0.1123	2.649533	0.109079
5	C	0.5972	3.39967	0.597502
6	O	-0.5679	2.959922	1.459017
7	N	-0.4157	3.249999	1.181109
8	H	0.2719	1.069078	0.109079
9	CX	0.0337	3.39967	0.760078
10	H1	0.0823	2.471353	0.109079
11	CT	-0.1825	3.39967	0.760078
12	HC	0.0603	2.649533	0.109079
13	HC	0.0603	2.649533	0.109079
14	HC	0.0603	2.649533	0.109079
15	C	0.5973	3.39967	0.597502
16	O	-0.5679	2.959922	1.459017
17	N	-0.4157	3.249999	1.181109
18	H	0.2719	1.069078	0.109079
19	CT	-0.149	3.39967	0.760078
20	H1	0.0976	2.471353	0.109079
21	H1	0.0976	2.471353	0.109079
22	H1	0.0976	2.471353	0.109079

Table 39 Partial charges and Lennard-Jones parameters for the ff14SB+RISM calculations of Ac-Ala-NHMe.

Atom	Type	*q*	σ (Å)	ε (kJ/mol)
1	HC	0.1123	2.649533	0.109079
2	CT	-0.3662	3.39967	0.760078
3	HC	0.1123	2.649533	0.109079
4	HC	0.1123	2.649533	0.109079
5	C	0.5972	3.39967	0.597502
6	O	-0.5679	2.959922	1.459017
7	N	-0.4157	3.249999	1.181109
8	H	0.2719	1.069078	0.109079
9	CX	0.0337	3.39967	0.760078
10	H1	0.0823	2.471353	0.109079
11	CT	-0.1825	3.39967	0.760078
12	HC	0.0603	2.649533	0.109079
13	HC	0.0603	2.649533	0.109079
14	HC	0.0603	2.649533	0.109079
15	C	0.5973	3.39967	0.597502
16	O	-0.5679	2.959922	1.459017
17	N	-0.4157	3.249999	1.181109
18	H	0.2719	1.069078	0.109079
19	CT	-0.149	3.39967	0.760078
20	H1	0.0976	2.471353	0.109079
21	H1	0.0976	2.471353	0.109079
22	H1	0.0976	2.471353	0.109079

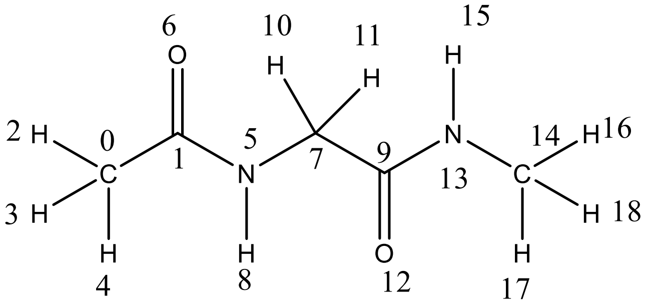

Figure 72: Assignment of atoms numbers to the corresponding chemical shifts and shielding constants.

Table 40 Partial charges and Lennard-Jones parameters for the charm/36m+RISM calculations of Ac-Gly-NHMe.

Atom	Type	*q*	σ (Å)	*ε* (kJ/mol)
1	CAY	-0.27	3.634867	0.541921
2	HY1	0.09	2.387609	0.166745
3	HY2	0.09	2.387609	0.166745
4	HY3	0.09	2.387609	0.166745
5	CY	0.51	3.563595	0.764247
6	OY	-0.51	3.029056	0.833724
7	N	-0.47	3.296325	1.38954
8	HN	0.31	0.400014	0.319594
9	CA	-0.02	3.581413	0.389071
10	HA1	0.09	2.387609	0.194536
11	HA2	0.09	2.387609	0.194536
12	C	0.51	3.563595	0.764247
13	O	-0.51	3.029056	0.833724
14	NT	-0.47	3.296325	1.38954
15	HNT	0.31	0.400014	0.319594
16	CAT	-0.11	3.634867	0.541921
17	HT1	0.09	2.387609	0.166745
18	HT2	0.09	2.387609	0.166745
19	HT3	0.09	2.387609	0.166745

Table 41 Partial charges and Lennard-Jones parameters for the ff14SB+RISM calculations of Ac-Gly-NHMe.

Atom	Type	*q*	σ (Å)	*ε* (kJ/mol)
1	HC	0.1123	2.649533	0.109079
2	CT	-0.3662	3.39967	0.760078
3	HC	0.1123	2.649533	0.109079
4	HC	0.1123	2.649533	0.109079
5	C	0.5972	3.39967	0.597502
6	O	-0.5679	2.959922	1.459017
7	N	-0.4157	3.249999	1.181109
8	H	0.2719	1.069078	0.109079
9	XC	-0.0252	3.39967	0.760078
10	H1	0.0698	2.471353	0.109079
11	H1	0.0698	2.471353	0.109079
12	C	0.5973	3.39967	0.597502
13	O	-0.5679	2.959922	1.459017
14	N	-0.4157	3.249999	1.181109
15	H	0.2719	1.069078	0.109079
16	CT	-0.149	3.39967	0.760078
17	H1	0.0976	2.471353	0.109079
18	H1	0.0976	2.471353	0.109079
19	H1	0.0976	2.471353	0.109079

Table 42 Partial charges and Lennard-Jones parameters for the ff19SB+RISM calculations of Ac-Gly-NHMe.

Atom	Type	*q*	σ (Å)	*ε* (kJ/mol)
1	HC	0.1123	2.649533	0.109079
2	CT	-0.3662	3.39967	0.760078
3	HC	0.1123	2.649533	0.109079
4	HC	0.1123	2.649533	0.109079
5	C	0.5972	3.39967	0.597502
6	O	-0.5679	2.959922	1.459017
7	N	-0.4157	3.249999	1.181109
8	H	0.2719	1.069078	0.109079
9	XC	-0.0252	3.39967	0.760078
10	H1	0.0698	2.471353	0.109079
11	H1	0.0698	2.471353	0.109079
12	C	0.5973	3.39967	0.597502
13	O	-0.5679	2.959922	1.459017
14	N	-0.4157	3.249999	1.181109
15	H	0.2719	1.069078	0.109079
16	CT	-0.149	3.39967	0.760078
17	H1	0.0976	2.471353	0.109079
18	H1	0.0976	2.471353	0.109079
19	H1	0.0976	2.471353	0.109079

6.11 Reference chemical shielding constants of DSS and NH_3 at high pressure

Table 43 Reference chemical shielding constants for pressure-dependent NMR calculations. The ^{1}H and ^{13}C shielding constants were calculated with MP2/6-311+G(d,p)//B3LYP/6-311+G(d,p) on DSS conformations.

Pressure	^{1}H	^{13}C	^{15}N
1 bar	31.8997	200.934	258.687
100 bar	31.8998	200.937	258.6975
500 bar	31.9002	200.949	258.7387
1 kbar	31.9007	200.963	258.7871
2 kbar	31.9015	200.987	258.8753
3 kbar	31.9023	201.008	258.9543
4 kbar	31.9030	201.027	259.0262
5 kbar	31.9036	201.045	259.0922
7.5 kbar	31.9050	201.082	259.2373
10 kbar	31.9062	201.115	259.3618

7 References

1. Inagaki, F., Hinrichs, K.-U., Kubo, Y., Bowles, M. W., Heuer, V. B., Hong, W.-L., Hoshino, T., Ijiri, A., Imachi, H., Ito, M., Kaneko, M., Lever, M. A., Lin, Y.-S., Methé, B. A., Morita, S., Morono, Y., Tanikawa, W., Bihan, M., Bowden, S. A., Elvert, M., Glombitza, C., Gross, D., Harrington, G. J., Hori, T., Li, K., Limmer, D., Liu, C.-H., Murayama, M., Ohkouchi, N., Ono, S., Park, Y.-S., Phillips, S. C., Prieto-Mollar, X., Purkey, M., Riedinger, N., Sanada, Y., Sauvage, J., Snyder, G., Susilawati, R., Takano, Y., Tasumi, E., Terada, T., Tomaru, H., Trembath-Reichert, E., Wang, D. T. & Yamada, Y. DEEP BIOSPHERE. Exploring deep microbial life in coal-bearing sediment down to ~2.5 km below the ocean floor. *Science* **349,** 420–424; 10.1126/science.aaa6882 (2015).

2. Panick, G., Malessa, R., Winter, R., Rapp, G., Frye, K. J. & Royer, C. A. Structural characterization of the pressure-denatured state and unfolding/refolding kinetics of staphylococcal nuclease by synchrotron small-angle X-ray scattering and Fourier-transform infrared spectroscopy. *J. Mol. Biol.* **275,** 389–402; 10.1006/jmbi.1997.1454 (1998).

3. Prigozhin, M. B., Liu, Y., Wirth, A. J., Kapoor, S., Winter, R., Schulten, K. & Gruebele, M. Misplaced helix slows down ultrafast pressure-jump protein folding. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 8087–8092; 10.1073/pnas.1219163110 (2013).

4. Rouget, J.-B., Schroer, M. A., Jeworrek, C., Pühse, M., Saldana, J.-L., Bessin, Y., Tolan, M., Barrick, D., Winter, R. & Royer, C. A. Unique features of the folding landscape of a repeat protein revealed by pressure perturbation. *Biophys. J.* **98,** 2712–2721; 10.1016/j.bpj.2010.02.044 (2010).

5. Roche, J., Caro, J. A., Norberto, D. R., Barthe, P., Roumestand, C., Schlessman, J. L., Garcia, A. E., García-Moreno, B. E. & Royer, C. A. Cavities determine the pressure unfolding of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 6945–6950; 10.1073/pnas.1200915109 (2012).

6. Akasaka, K., Kitahara, R. & Kamatari, Y. O. Exploring the folding energy landscape with pressure. *Arch. Biochem. Biophys.* **531,** 110–115; 10.1016/j.abb.2012.11.016 (2013).

7. Cavini, I. A., Munte, C. E., Erlach, M. B., van Groen, T., Kadish, I., Zhang, T., Ziehm, T., Nagel-Steger, L., Kutzsche, J., Kremer, W., Willbold, D. & Kalbitzer, H. R. Inhibition of amyloid Aβ aggregation by high pressures or specific d-enantiomeric peptides. *Chem. Commun. (Camb)* **54,** 3294–3297; 10.1039/c8cc01458b (2018).

8. Patterson, M. F., Quinn, M., Simpson, R. & Gilmour, A. Sensitivity of Vegetative Pathogens to High Hydrostatic Pressure Treatment in Phosphate-Buffered Saline and Foods. *J. Food. Prot.* **58,** 524–529; 10.4315/0362-028X-58.5.524 (1995).

9. Oey, I., Lille, M., van Loey, A. & Hendrickx, M. Effect of high-pressure processing on colour, texture and flavour of fruit- and vegetable-based food products: a review. *Trends Food Sci. Technol.* **19,** 320–328; 10.1016/j.tifs.2008.04.001 (2008).

10. Trzesniak, D., Lins, R. D. & van Gunsteren, W. F. Protein under pressure: molecular dynamics simulation of the arc repressor. *Proteins* **65,** 136–144; 10.1002/prot.21034 (2006).

11. Paci, E. High pressure simulations of biomolecules. *BBA, Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1595,** 185–200; 10.1016/S0167-4838(01)00343-0 (2002).

12. Paschek, D. & García, A. E. Reversible temperature and pressure denaturation of a protein fragment: a replica exchange molecular dynamics simulation study. *Phys. Rev. Lett.* **93,** 238105; 10.1103/PhysRevLett.93.238105 (2004).

13. Day, R. & García, A. E. Water penetration in the low and high pressure native states of ubiquitin. *Proteins* **70,** 1175–1184; 10.1002/prot.21562 (2008).

14. Paschek, D., Hempel, S. & García, A. E. Computing the stability diagram of the Trp-cage miniprotein. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 17754–17759; 10.1073/pnas.0804775105 (2008).

15. Sarupria, S., Ghosh, T., García, A. E. & Garde, S. Studying pressure denaturation of a protein by molecular dynamics simulations. *Proteins* **78,** 1641–1651; 10.1002/prot.22680 (2010).

16. Sarma, R. & Paul, S. Trimethylamine-N-oxide's effect on polypeptide solvation at high pressure: a molecular dynamics simulation study. *J. Phys. Chem. B.* **117,** 9056–9066; 10.1021/jp405202j (2013).

17. Hata, H., Nishiyama, M. & Kitao, A. Molecular dynamics simulation of proteins under high pressure: Structure, function and thermodynamics. *Biochim. Biophys. Acta. Gen. Subj.* **1864,** 129395; 10.1016/j.bbagen.2019.07.004 (2020).

18. Hölzl, C., Kibies, P., Imoto, S., Frach, R., Suladze, S., Winter, R., Marx, D., Horinek, D. & Kast, S. M. Design principles for high-pressure force fields: Aqueous TMAO solutions from ambient to kilobar pressures. *J. Chem. Phys.* **144,** 144104; 10.1063/1.4944991 (2016).

19. Hölzl, C., Kibies, P., Imoto, S., Noetzel, J., Knierbein, M., Salmen, P., Paulus, M., Nase, J., Held, C., Sadowski, G., Marx, D., Kast, S. M. & Horinek, D. Structure and thermodynamics of aqueous urea solutions from ambient to kilobar pressures: From thermodynamic modeling, experiments, and first principles simulations to an accurate force field description. *Biophys. Chem.* **254,** 106260; 10.1016/j.bpc.2019.106260 (2019).

20. Tomasi, J., Mennucci, B. & Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **105,** 2999–

3093; 10.1021/cr9904009 (2005).

21. Cammi, R. A new extension of the polarizable continuum model: Toward a quantum chemical description of chemical reactions at extreme high pressure. *J .Comput. Chem.* **36,** 2246–2259; 10.1002/jcc.24206 (2015).

22. Cammi, R., Cappelli, C., Mennucci, B. & Tomasi, J. Calculation and analysis of the harmonic vibrational frequencies in molecules at extreme pressure: methodology and diborane as a test case. *J. Chem. Phys.* **137,** 154112; 10.1063/1.4757285 (2012).

23. Chen, B., Hoffmann, R. & Cammi, R. The Effect of Pressure on Organic Reactions in Fluids-a New Theoretical Perspective. *Angew. Chem. Int. Ed. Engl.* **56,** 11126–11142; 10.1002/anie.201705427 (2017).

24. Kloss, T., Heil, J. & Kast, S. M. Quantum chemistry in solution by combining 3D integral equation theory with a cluster embedding approach. *J. Phys. Chem. B.* **112,** 4337–4343; 10.1021/jp710680m (2008).

25. Tielker, N., Eberlein, L., Güssregen, S. & Kast, S. M. The SAMPL6 challenge on predicting aqueous pKa values from EC-RISM theory. *J. Comput. Aided. Mol. Des.* **32,** 1151–1163; 10.1007/s10822-018-0140-z (2018).

26. Frach, R., Kibies, P., Bottcher, S., Pongratz, T., Strohfeldt, S., Kurrmann, S., Koehler, J., Hofmann, M., Kremer, W., Kalbitzer, H. R., Reiser, O., Horinek, D. & Kast, S. M. The Chemical Shift Baseline for High-Pressure NMR Spectra of Proteins. *Angew. Chem. Int. Ed. Engl.* **55,** 8757–8760; 10.1002/anie.201602054 (2016).

27. Pongratz, T., Kibies, P., Eberlein, L., Tielker, N., Hölzl, C., Imoto, S., Beck Erlach, M., Kurrmann, S., Schummel, P. H., Hofmann, M., Reiser, O., Winter, R., Kremer, W., Kalbitzer, H. R., Marx, D., Horinek, D. & Kast, S. M. Pressure-dependent electronic structure calculations using integral equation-based solvation models. *Biophys. Chem.* **257,** 106258; 10.1016/j.bpc.2019.106258 (2020).

28. Eberlein, L. J. A combined computational and NMR-spectroscopic approach for tautomer elucidation under extreme conditions towards investigating the robustness of genetic codes, 2021.

29. Munte, C. E., Karl, M., Kauter, W., Eberlein, L., Pham, T.-V., Erlach, M. B., Kast, S. M., Kremer, W. & Kalbitzer, H. R. High pressure response of 1H NMR chemical shifts of purine nucleotides. *Biophys. Chem.* **254,** 106261; 10.1016/j.bpc.2019.106261 (2019).

30. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **117,** 5179–5197; 10.1021/ja00124a002 (1995).

31. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A. & Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65,** 712–725; 10.1002/prot.21123 (2006).

32. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E. & Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory. Comput.* **11,** 3696–3713; 10.1021/acs.jctc.5b00255 (2015).

33. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S. & Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106,** 765–784; 10.1021/ja00315a051 (1984).

34. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. An all atom force field for simulations of proteins and nucleic acids. *J. Comput. Chem.* **7,** 230–252; 10.1002/jcc.540070216 (1986).

35. Tian, C., Kasavajhala, K., Belfon, K. A. A., Raguette, L., Huang, H., Migues, A. N., Bickel, J., Wang, Y., Pincay, J., Wu, Q. & Simmerling, C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **16,** 528–552; 10.1021/acs.jctc.9b00591 (2020).

36. Song, D., Luo, R. & Chen, H.-F. The IDP-Specific Force Field ff14IDPSFF Improves the Conformer Sampling of Intrinsically Disordered Proteins. *J. Chem. Inf. Model.* **57,** 1166–1178; 10.1021/acs.jcim.7b00135 (2017).

37. Song, D., Wang, W., Ye, W., Ji, D., Luo, R. & Chen, H.-F. ff14IDPs force field improving the conformation sampling of intrinsically disordered proteins. *Chem. Biol. Drug. Des.* **89,** 5–15; 10.1111/cbdd.12832 (2017).

38. Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M. & Mackerell, A. D., JR. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J. Chem. Theory. Comput.* **8,** 3257–3273; 10.1021/ct300400x (2012).

39. Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., Groot, B. L. de, Grubmüller, H. & MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14,** 71–73; 10.1038/nmeth.4067 (2017).

40. MacKerell, A. D., Feig, M. & Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25,** 1400–1415; 10.1002/jcc.20065 (2004).

41. Oostenbrink, C., Villa, A., Mark, A. E. & van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of

hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **25,** 1656–1676; 10.1002/jcc.20090 (2004).

42. Besler, B. H., Merz, K. M. & Kollman, P. A. Atomic charges derived from semiempirical methods. *J. Comput. Chem.* **11,** 431–439; 10.1002/jcc.540110404 (1990).

43. Singh, U. C. & Kollman, P. A. An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **5,** 129–145; 10.1002/jcc.540050204 (1984).

44. Chirlian, L. E. & Francl, M. M. Atomic charges derived from electrostatic potentials. A detailed study. *J. Comput. Chem.* **8,** 894–905; 10.1002/jcc.540080616 (1987).

45. Breneman, C. M. & Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **11,** 361–373; 10.1002/jcc.540110311 (1990).

46. Kang, Y. K. Ab initio MO and density functional studies on trans and cis conformers of N-methylacetamide. *J. Mol. Struct. Theochem* **546,** 183–193; 10.1016/S0166-1280(01)00445-6 (2001).

47. Jorgensen, W. L. & Gao, J. Cis-trans energy difference for the peptide bond in the gas phase and in aqueous solution. *J. Am. Chem. Soc.* **110,** 4212–4216; 10.1021/ja00221a020 (1988).

48. Radzicka, A., Pedersen, L. & Wolfenden, R. Influences of solvent water on protein folding: free energies of solvation of cis and trans peptides are nearly identical. *Biochemistry* **27,** 4538–4541; 10.1021/bi00412a047 (1988).

49. Song, S., Asher, S. A., Krimm, S. & Shaw, K. D. Ultraviolet resonance Raman studies of trans and cis peptides: photochemical consequences of the twisted .pi.* excited state. *J. Am. Chem. Soc.* **113,** 1155–1163; 10.1021/ja00004a015 (1991).

50. Luque, F. J. & Orozco, M. Theoretical study of N-methylacetamide in vacuum and aqueous solution: implications for the peptide bond isomerization. *J. Org. Chem.* **58,** 6397–6405; 10.1021/jo00075a039 (1993).

51. Han, W.-G. & Suhai, S. Density Functional Studies on N -Methylacetamide–Water Complexes. *J. Phys. Chem.* **100,** 3942–3949; 10.1021/jp952250l (1996).

52. Yu, H. A., Karplus, M. & Pettitt, B. M. Aqueous solvation of N-methylacetamide conformers: comparison of simulations and integral equation theories. *J. Am. Chem. Soc.* **113,** 2425–2434; 10.1021/ja00007a012 (1991).

53. Morgantini, P.-Y. & Kollman, P. A. Solvation Free Energies of Amides and Amines: Disagreement between Free Energy Calculations and Experiment. *J. Am. Chem. Soc.* **117,** 6057–6063; 10.1021/ja00127a019 (1995).

54. Ding, Y., Bernardo, D. N., Krogh-Jespersen, K. & Levy, R. M. Solvation Free Energies of Small Amides and Amines from Molecular Dynamics/Free Energy Perturbation Simulations Using Pairwise Additive and Many-Body Polarizable Potentials. *J. Phys. Chem.* **99,** 11575–11583; 10.1021/j100029a040 (1995).

55. Gao, J. & Freindorf, M. Hybrid ab Initio QM/MM Simulation of N -Methylacetamide in Aqueous Solution. *J. Phys. Chem. A* **101,** 3182–3188; 10.1021/jp970041q (1997).

56. Gaigeot, M. P., Vuilleumier, R., Sprik, M. & Borgis, D. Infrared Spectroscopy of N-Methylacetamide Revisited by ab Initio Molecular Dynamics Simulations. *J. Chem. Theory. Comput.* **1,** 772–789; 10.1021/ct050029z (2005).

57. Ataka, S., Takeuchi, H. & Tasumi, M. Infrared studies of the less stable cis form of N-methylformmaide and N-methylacetamide in low-temperature nitrogen matrices and vibrational analyses of the trans and cis forms of these molecules. *J. Mol. Struct.* **113,** 147–160; 10.1016/0022-2860(84)80140-4 (1984).

58. Nandi, A., Qu, C. & Bowman, J. M. Full and fragmented permutationally invariant polynomial potential energy surfaces for trans and cis N-methyl acetamide and isomerization saddle points. *J. Chem. Phys.* **151,** 84306; 10.1063/1.5119348 (2019).

59. Bisetty, K., Catalan, J. G., Kruger, H. G. & Perez, J. J. Conformational analysis of small peptides of the type Ac–X–NHMe, where X=Gly, Ala, Aib and Cage. *J. Mol. Struct. Theochem* **731,** 127–137; 10.1016/j.theochem.2005.04.037 (2005).

60. Bohr, H. G., Jalkanen, K. J., Elstner, M., Frimand, K. & Suhai, S. A comparative study of MP2, B3LYP, RHF and SCC-DFTB force fields in predicting the vibrational spectra of N-acetyl-L-alanine-N'-methyl amide. VA and VCD spectra. *Chem. Phys. Lett.* **246,** 13–36; 10.1016/S0301-0104(99)00176-7 (1999).

61. Elstner, M., Jalkanen, K. J., Knapp-Mohammady, M., Frauenheim, T. & Suhai, S. Energetics and structure of glycine and alanine based model peptides. Approximate SCC-DFTB, AM1 and PM3 methods in comparison with DFT, HF and MP2 calculations. *Chem. Phys. Lett.* **263,** 203–219; 10.1016/S0301-0104(00)00375-X (2001).

62. Anderson, A. G. & Hermans, J. Microfolding: conformational probability map for the alanine dipeptide in water from molecular dynamics simulations. *Proteins* **3,** 262–265; 10.1002/prot.340030408 (1988).

63. Avbelj, F., Grdadolnik, S. G., Grdadolnik, J. & Baldwin, R. L. Intrinsic backbone preferences are fully present in blocked amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 1272–1277; 10.1073/pnas.0510420103 (2006).

64. Grdadolnik, J., Mohacek-Grosev, V., Baldwin, R. L. & Avbelj, F. Populations of the three major backbone conformations in 19 amino acid dipeptides. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 1794–1798; 10.1073/pnas.1017317108 (2011).

65. Avignon, M., Garrigou-Lagrange, C. & Bothorel, P. Conformational analysis of dipeptides in aqueous solution. II. Molecular structure of glycine and alanine dipeptides by depolarized Rayleigh scattering and laser Raman spectroscopy. *Biopolymers* **12,** 1651–1669; 10.1002/bip.1973.360120716 (1973).

66. Cormanich, R. A., Bühl, M. & Rittner, R. Understanding the conformational behaviour of Ac-Ala-NHMe in different media. A joint NMR and DFT study. *Org. Biomol. Chem* **13,** 9206–9213; 10.1039/C5OB01296A (2015).

67. Cormanich, R. A., Rittner, R. & Bühl, M. Conformational preferences of Ac-Gly-NHMe in solution. *RSC Adv* **5,** 13052–13060; 10.1039/C4RA16472E (2015).

68. Grouleff, J. & Jensen, F. Searching Peptide Conformational Space. *J. Chem. Theory. Comput.* **7,** 1783–1790; 10.1021/ct200161e (2011).

69. Takekiyo, T., Imai, T., Kato, M. & Taniguchi, Y. Temperature and pressure effects on conformational equilibria of alanine dipeptide in aqueous solution. *Biopolymers* **73,** 283–290; 10.1002/bip.10548 (2004).

70. Takekiyo, T., Imai, T., Kato, M. & Taniguchi, Y. Understanding high pressure stability of helical conformation of oligopeptides and helix bundle protein high pressure FT-IR and RISM theoretical studies. *Biochim. Biophys. Acta.* **1764,** 355–363; 10.1016/j.bbapap.2005.11.013 (2006).

71. Cortis, C. M., Langlois, J.-M., Beachy, M. D. & Friesner, R. A. Quantum mechanical geometry optimization in solution using a finite element continuum electrostatics method. *J. Chem. Phys.* **105,** 5472–5484; 10.1063/1.472388 (1996).

72. Beachy, M. D., Chasman, D., Murphy, R. B., Halgren, T. A. & Friesner, R. A. Accurate ab Initio Quantum Chemical Determination of the Relative Energetics of Peptide Conformations and Assessment of Empirical Force Fields. *J. Am. Chem. Soc.* **119,** 5908–5920; 10.1021/ja962310g (1997).

73. Wang, Z.-X. & Duan, Y. Solvation effects on alanine dipeptide. A MP2/cc-pVTZ//MP2/6-31G** study of (Phi, Psi) energy maps and conformers in the gas phase, ether, and water. *J. Comput. Chem.* **25,** 1699–1716; 10.1002/jcc.20092 (2004).

74. Dudek, M. J. & Ponder, J. W. Accurate modeling of the intramolecular electrostatic energy of proteins. *J. Comput. Chem.* **16,** 791–816; 10.1002/jcc.540160702 (1995).

75. Brooks, C. & Case, D. A. Simulations of peptide conformational dynamics and thermodynamics. *Chem. Rev.* **93,** 2487–2502; 10.1021/cr00023a008 (1993).

76. Cheam, T. C. & Krimm, S. Ab initio force fields of alanine dipeptide in C5 and C7 conformations. *J. Mol. Struct. Theochem* **188,** 15–43; 10.1016/0166-1280(89)85023-7 (1989).

77. Madison, V. & Kopple, K. D. Solvent-dependent conformational distributions of some dipeptides. *J. Am. Chem. Soc.* **102,** 4855–4863; 10.1021/ja00535a001 (1980).

78. Head-Gordon, T., Head-Gordon, M., Frisch, M. J., Brooks, C. L. & Pople, J. A. Theoretical study of blocked glycine and alanine peptide analogs. *J. Am. Chem. Soc.* **113,** 5989–5997; 10.1021/ja00016a010 (1991).

79. Boehm, H. J. & Brode, S. Ab initio SCF calculations on low-energy conformers of N-acetyl-N'-methylalaninamide and N-acetyl-N'-methylglycinamide. *J. Am. Chem. Soc.* **113,** 7129–7135; 10.1021/ja00019a007 (1991).

80. Rommel-Möhle, K. & Hofmann, H.-J. Conformation dynamics in peptides. Quantum chemical calculations and molecular dynamics simulations on N-acetylalanyl-N'-methylamide. *J. Mol. Struct. Theochem* **285,** 211–219; 10.1016/0166-1280(93)87036-D (1993).

81. Scarsdale, J. N., van Alsenoy, C., Klimkowski, V. J., Schaefer, L. & Momany, F. A. Ab initio studies of molecular geometries. 27. Optimized molecular structures and conformational analysis of N.alpha.-acetyl-N-methylalaninamide and comparison with peptide crystal data and empirical calculations. *J. Am. Chem. Soc.* **105,** 3438–3445; 10.1021/ja00349a010 (1983).

82. Schäfer, L., van Alsenoy, C. & Scarsdale, J. N. Abinitio studies of structural features not easily amenable to experiment. 23. Molecular structures and conformational analysis of the dipeptide N -acetyl- N ′-methyl glycyl amide and the significance of local geometries for peptide structures. *J. Chem. Phys.* **76,** 1439–1444; 10.1063/1.443103 (1982).

83. Pettitt, B. M. & Karplus, M. The potential of mean force surface for the alanine dipeptide in aqueous solution. A theoretical approach. *Chem. Phys. Lett.* **121,** 194–201; 10.1016/0009-2614(85)85509-3 (1985).

84. Lovell, S. C., Davis, I. W., Arendall, W. B., Bakker, P. I. W. de, Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* **50,** 437–450; 10.1002/prot.10286 (2003).

85. Karplus, M. Contact Electron-Spin Coupling of Nuclear Magnetic Moments. *J. Phys. Chem.* **30,** 11–15;

10.1063/1.1729860 (1959).

86. Gaigeot, M.-P. Infrared spectroscopy of the alanine dipeptide analog in liquid water with DFT-MD. Direct evidence for P(II)/beta conformations. *Phys. Chem. Chem. Phys.* **12,** 10198–10209; 10.1039/C003485A (2010).

87. Schweitzer-Stenner, R., Eker, F., Huang, Q., Griebenow, K., Mroz, P. A. & Kozlowski, P. M. Structure Analysis of Dipeptides in Water by Exploring and Utilizing the Structural Sensitivity of Amide III by Polarized Visible Raman, FTIR–Spectroscopy and DFT Based Normal Coordinate Analysis. *J. Phys. Chem. B.* **106,** 4294–4304; 10.1021/jp0137118 (2002).

88. Jordan, T. & Spiro, T. G. Enhancement of Cα hydrogen vibrations in the resonance Raman spectra of amides. *J. Raman Spectrosc.* **25,** 537–543; 10.1002/jrs.1250250715 (1994).

89. Atkins, P. W. & Paula, J. de. *Physikalische Chemie* (Wiley-VCH, Place of publication not identified, 2014).

90. Atkins, P. W. & Friedman, R. *Molecular quantum mechanics.* 5th ed. (Oxford University Press, Oxford, New York, 2011).

91. Griffiths, P. R. & Haseth, J. A. de. *Fourier transform infrared spectrometry.* 2nd ed. (Wiley-Interscience, Hoboken, NJ, 2007).

92. Réhault, J., Borrego-Varillas, R., Oriana, A., Manzoni, C., Hauri, C. P., Helbing, J. & Cerullo, G. Fourier transform spectroscopy in the vibrational fingerprint region with a birefringent interferometer. *Opt. Express* **25,** 4403–4413; 10.1364/OE.25.004403 (2017).

93. Winter, R., Noll, F. & Czeslik, C. *Methoden der biophysikalischen Chemie.* 2nd ed. (Vieweg + Teubner, Wiesbaden, 2011).

94. Imoto, S., Kibies, P., Rosin, C., Winter, R., Kast, S. M. & Marx, D. Toward Extreme Biophysics: Deciphering the Infrared Response of Biomolecular Solutions at High Pressures. *Angew. Chem. Int. Ed. Engl.* **55,** 9534–9538; 10.1002/anie.201602757 (2016).

95. Friebolin, H. *Ein- und zweidimensionale NMR-Spektroskopie. Eine Einführung.* 5th ed. (Wiley-VCH, Weinheim, 2013).

96. Akasaka, K. Probing conformational fluctuation of proteins by pressure perturbation. *Chem. Rev.* **106,** 1814–1835; 10.1021/cr040440z (2006).

97. Akasaka, K. Exploring the entire conformational space of proteins by high-pressure NMR. *Pure Appl. Chem.* **75,** 927–936; 10.1351/pac200375070927 (2003).

98. Kalbitzer, H. R., Spoerner, M., Ganser, P., Hozsa, C. & Kremer, W. Fundamental link between folding states and functional states of proteins. *J. Am. Chem. Soc.* **131,** 16714–16719; 10.1021/ja904314q (2009).

99. Akasaka, K. & Li, H. Low-lying excited states of proteins revealed from nonlinear pressure shifts in 1H and 15N NMR. *Biochemistry* **40,** 8665–8671; 10.1021/bi010312u (2001).

100. Goldfarb, D. & Stoll, S. (eds.). *EPR spectroscopy. Fundamentals and methods* (Wiley, Chichester, West Sussex, 2018).

101. Kaupp, M., Bühl, M. & Malkin, V. G. *Calculation of NMR and EPR Parameters* (Wiley, 2004).

102. Hoffmann, A. K. & Henderson, A. T. A NEW STABLE FREE RADICAL: DI-t-BUTYLNITROXIDE. *J. Am. Chem. Soc.* **83,** 4671–4672; 10.1021/ja01483a048 (1961).

103. Forrester, A. R. & Thomson, R. H. Stable Nitroxide Radicals. *Nature* **203,** 74–75; 10.1038/203074a0 (1964).

104. Subczynski, W. K., Widomska, J. & Feix, J. B. Physical properties of lipid bilayers from EPR spin labeling and their influence on chemical reactions in a membrane environment. *Free Radic. Biol. Med.* **46,** 707–718; 10.1016/j.freeradbiomed.2008.11.024 (2009).

105. Stone, T. J., Buckman, T., Nordio, P. L. & McConnell, H. M. Spin-labeled biomolecules. *Proc. Natl. Acad. Sci. U. S. A.* **54,** 1010–1017; 10.1073/pnas.54.4.1010 (1965).

106. Hubbell, W. L. & Altenbach, C. Investigation of structure and dynamics in membrane proteins using site-directed spin labeling. *Curr. Opin. Struct. Biol.* **4,** 566–573; 10.1016/S0959-440X(94)90219-4 (1994).

107. Altenbach, C., Marti, T., Khorana, H. G. & Hubbell, W. L. Transmembrane protein structure: spin labeling of bacteriorhodopsin mutants. *Science* **248,** 1088–1092; 10.1126/science.2160734 (1990).

108. Altenbach, C., Flitsch, S. L., Khorana, H. G. & Hubbell, W. L. Structural studies on transmembrane proteins. 2. Spin labeling of bacteriorhodopsin mutants at unique cysteines. *Biochemistry* **28,** 7806–7812; 10.1021/bi00445a042 (1989).

109. Greenhalgh, D. A., Altenbach, C., Hubbell, W. L. & Khorana, H. G. Locations of Arg-82, Asp-85, and Asp-96 in helix C of bacteriorhodopsin relative to the aqueous boundaries. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 8626–8630; 10.1073/pnas.88.19.8626 (1991).

110. Shin, Y. K. & Hubbell, W. L. Determination of electrostatic potentials at biological interfaces using electron-electron

double resonance. *Biophys. J.* **61,** 1443–1453; 10.1016/S0006-3495(92)81950-1 (1992).

111. Anthony-Cahill, S. J., Benfield, P. A., Fairman, R., Wasserman, Z. R., Brenner, S. L., Stafford, W. F., Altenbach, C., Hubbell, W. L. & DeGrado, W. F. Molecular characterization of helix-loop-helix peptides. *Science* **255,** 979–983; 10.1126/science.1312255 (1992).

112. McCoy, J. & Hubbell, W. L. High-pressure EPR reveals conformational equilibria and volumetric properties of spin-labeled proteins. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 1331–1336; 10.1073/pnas.1017877108 (2011).

113. Plato, M., Steinhoff, H.-J., Wegener, C., Törring, J. T., Savisky, A. & Möbius, K. Molecular orbital study of polarity and hydrogen bonding effects on the g and hyperfine tensors of site directed NO spin labelled bacteriorhodopsin. *Mol. Phys.* **100,** 3711–3721; 10.1080/00268970210166246 (2002).

114. Margita, K., Voinov, M. A. & Smirnov, A. I. Effect of Solution Ionic Strength on the pKa of the Nitroxide pH EPR Probe 2,2,3,4,5,5-Hexamethylimidazolidin-1-oxyl. *Cell Biochem. Biophys.* **75,** 185–193; 10.1007/s12013-017-0780-y (2017).

115. Voinov, M. A. & Smirnov, A. I. Ionizable Nitroxides for Studying Local Electrostatic Properties of Lipid Bilayers and Protein Systems by EPR. *Methods Enzymol.* **564,** 191–217; 10.1016/bs.mie.2015.08.007 (2015).

116. Khramtsov, V. V., Marsh, D., Weiner, L. & Reznikov, V. A. The application of pH-sensitive spin labels to studies of surface potential and polarity of phospholipid membranes and proteins. *Biochim. Biophys. Acta - Biomembr.* **1104,** 317–324; 10.1016/0005-2736(92)90046-O (1992).

117. Barone, V., Bencini, A., Cossi, M., Di Matteo, A., Mattesini, M. & Totti, F. Assessment of a Combined QM/MM Approach for the Study of Large Nitroxide Systems in Vacuo and in Condensed Phases. *J. Am. Chem. Soc.* **120,** 7069–7078; 10.1021/ja981079z (1998).

118. Stendardo, E., Pedone, A., Cimino, P., Cristina Menziani, M., Crescenzi, O. & Barone, V. Extension of the AMBER force-field for the study of large nitroxides in condensed phases: an ab initio parameterization. *Phys. Chem. Chem. Phys.* **12,** 11697–11709; 10.1039/C001481H (2010).

119. Asher, J. R., Doltsinis, N. L. & Kaupp, M. Ab initio molecular dynamics simulations and g-tensor calculations of aqueous benzosemiquinone radical anion: effects of regular and "T-stacked" hydrogen bonds. *J. Am. Chem. Soc.* **126,** 9854–9861; 10.1021/ja0485053? (2004).

120. Giovannini, T., Lafiosca, P., Chandramouli, B., Barone, V. & Cappelli, C. Effective yet reliable computation of hyperfine coupling constants in solution by a QM/MM approach: Interplay between electrostatics and non-electrostatic effects. *J. Chem. Phys.* **150,** 124102; 10.1063/1.5080810 (2019).

121. Giovannini, T., Lafiosca, P. & Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. *J. Chem. Theory Comput.* **13,** 4854–4870; 10.1021/acs.jctc.7b00776 (2017).

122. Hedegård, E. D., Kongsted, J. & Sauer, S. P. A. Optimized Basis Sets for Calculation of Electron Paramagnetic Resonance Hyperfine Coupling Constants: aug-cc-pVTZ-J for the 3d Atoms Sc-Zn. *J. Chem. Theory Comput.* **7,** 4077–4087; 10.1021/ct200587k (2011).

123. Hameka, H. F. & Turner, A. G. Calculation of ESR coupling constants by means of the unrestricted Hartree-Fock method. *J. Magn. Reson.* **64,** 66–75; 10.1016/0022-2364(85)90031-9 (1985).

124. Chipman, D. M. The spin polarization model for hyperfine coupling constants. *Theoret. Chim. Acta* **82,** 93–115; 10.1007/BF01113132 (1992).

125. Ishii & Shimizu. Calculation of hyperfine coupling constants of radicals by density-functional theory. *Phys. Rev. A.* **48,** 1691–1694; 10.1103/PhysRevA.48.1691 (1993).

126. Barone, V., Adamo, C. & Russo, N. Density functional approach to the structures and EPR parameters of open shell systems. The case of fluorovinyl radicals. *Chem. Phys. Lett.* **212,** 5–11; 10.1016/0009-2614(93)87099-O (1993).

127. Kossmann, S., Kirchner, B. & Neese, F. Performance of modern density functional theory for the prediction of hyperfine structure: meta-GGA and double hybrid functionals. *Mol. Phys.* **105,** 2049–2071; 10.1080/00268970701604655 (2007).

128. Improta, R. & Barone, V. Interplay of electronic, environmental, and vibrational effects in determining the hyperfine coupling constants of organic free radicals. *Chem. Rev.* **104,** 1231–1254; 10.1021/cr960085f (2004).

129. Guo, Y., Riplinger, C., Becker, U., Liakos, D. G., Minenkov, Y., Cavallo, L. & Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method DLPNO-CCSD(T). *J. Chem. Phys.* **148,** 11101; 10.1063/1.5011798 (2018).

130. Riplinger, C. & Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **138,** 34106; 10.1063/1.4773581 (2013).

131. Riplinger, C., Sandhoefer, B., Hansen, A. & Neese, F. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *J. Chem. Phys.* **139,** 134101; 10.1063/1.4821834 (2013).

132. Sharma, B., Tran, V. A., Pongratz, T., Galazzo, L., Zhurko, I., Bordignon, E., Kast, S. M., Neese, F. & Marx, D. A Joint Venture of Ab Initio Molecular Dynamics, Coupled Cluster Electronic Structure Methods, and Liquid-State Theory to Compute Accurate Isotropic Hyperfine Constants of Nitroxide Probes in Water. *J. Chem. Theory Comput.* **17,** 6366–6386; 10.1021/acs.jctc.1c00582 (2021).

133. Saitow, M. & Neese, F. Accurate spin-densities based on the domain-based local pair-natural orbital coupled-cluster theory. *J. Chem. Phys.* **149,** 34104; 10.1063/1.5027114 (2018).

134. Miertuš, S., Scrocco, E. & Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys. Lett.* **55,** 117–129; 10.1016/0301-0104(81)85090-2 (1981).

135. Klamt, A. & Schüürmann, G. COSMO. A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2,* 799–805; 10.1039/P29930000799 (1993).

136. Andzelm, J., Kölmel, C. & Klamt, A. Incorporation of solvent effects into density functional calculations of molecular energies and geometries. *J. Chem. Phys.* **103,** 9312–9320; 10.1063/1.469990 (1995).

137. Barone, V. & Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **102,** 1995–2001; 10.1021/jp9716997 (1998).

138. Mennucci, B. Polarizable continuum model. *WIREs Comput. Mol. Sci.* **2,** 386–404; 10.1002/wcms.1086 (2012).

139. Tomasi, J., Mennucci, B. & Cancès, E. The IEF version of the PCM solvation method: an overview of a new method addressed to study molecular solutes at the QM ab initio level. *J. Mol. Struct. Theochem* **464,** 211–226; 10.1016/S0166-1280(98)00553-3 (1999).

140. Cammi, R., Verdolino, V., Mennucci, B. & Tomasi, J. Towards the elaboration of a QM method to describe molecular solutes under the effect of a very high pressure. *Chem. Phys. Lett.* **344,** 135–141; 10.1016/j.chemphys.2007.12.010 (2008).

141. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **136,** B864-B871; 10.1103/PhysRev.136.B864 (1964).

142. Evans, R. The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids. *Adv. Phys.* **28,** 143–200; 10.1080/00018737900101365 (1979).

143. Hansen, J.-P. & McDonald, I. R. *Theory of simple liquids. With applications to soft matter.* 4th ed. (Academic Press, Amsterdam, Boston, 2013).

144. Mermin, N. D. Thermal Properties of the Inhomogeneous Electron Gas. *Phys. Rev.* **137,** A1441-A1443; 10.1103/PhysRev.137.A1441 (1965).

145. Ornstein, L. S. & Zernike, F. Accidental deviations of density and opalescence at the critical point of a single substance. *Proc. Akad. Sci,* **17,** 793 (1914).

146. Blum, L. & Torruella, A. J. Invariant Expansion for Two-Body Correlations: Thermodynamic Functions, Scattering, and the Ornstein—Zernike Equation. *J. Chem. Phys.* **56,** 303–310; 10.1063/1.1676864 (1972).

147. Andersen, H. C. & Chandler, D. Optimized Cluster Expansions for Classical Fluids. I. General Theory and Variational Formulation of the Mean Spherical Model and Hard Sphere Percus-Yevick Equations. *J. Chem. Phys.* **57,** 1918–1929; 10.1063/1.1678512 (1972).

148. Andersen, H. C., Chandler, D. & Weeks, J. D. Optimized Cluster Expansions for Classical Fluids. III. Applications to Ionic Solutions and Simple Liquids. *J. Chem. Phys.* **57,** 2626–2631; 10.1063/1.1678641 (1972).

149. Lowden, L. J. & Chandler, D. Theory of intermolecular pair correlations for molecular liquids. Applications to the liquids carbon tetrachloride, carbon disulfide, carbon diselenide, and benzene. *J. Chem. Phys.* **61,** 5228–5241; 10.1063/1.1681868 (1974).

150. Hirata, F. *Molecular Theory of Solvation* (Springer Netherlands, Dordrecht, 2004).

151. Kast, S. M., Friedemann Schmidt, K. & Schilling, B. Integral equation theory for correcting truncation errors in molecular simulations. *Chem. Phys. Lett.* **367,** 398–404; 10.1016/S0009-2614(02)01754-2 (2003).

152. Morita, T. & Hiroike, K. Integral Equation for Pair Distribution Function. *Prog. Theor. Phys.* **23,** 385–387; 10.1143/PTP.23.385 (1960).

153. Singer, S. J. & Chandler, D. Free energy functions in the extended RISM approximation. *Mol. Phys.* **55,** 621–625; 10.1080/00268978500101591 (2006).

154. Schmidt, K. F. & Kast, S. M. Hybrid Integral Equation/Monte Carlo Approach to Complexation Thermodynamics. *J. Phys. Chem. B.* **106,** 6289–6297; 10.1021/jp020495%2B (2002).

155. Sato, H., Hirata, F. & Kato, S. Analytical energy gradient for the reference interaction site model multiconfigurational self-consistent-field method: Application to 1,2-difluoroethylene in aqueous solution. *J. Chem. Phys.* **105,** 1546–

1551; 10.1063/1.472015 (1996).

156. Beglov, D. & Roux, B. An Integral Equation To Describe the Solvation of Polar Molecules in Liquid Water. *J. Phys. Chem. B* **101,** 7821–7826; 10.1021/jp971083h (1997).

157. Kovalenko, A. & Hirata, F. Three-dimensional density profiles of water in contact with a solute of arbitrary shape. A RISM approach. *Chem. Phys. Lett.* **290,** 237–244; 10.1016/S0009-2614(98)00471-0 (1998).

158. Ratkova, E. L., Palmer, D. S. & Fedorov, M. V. Solvation thermodynamics of organic molecules by the molecular integral equation theory: approaching chemical accuracy. *Chem. Rev.* **115,** 6312–6356; 10.1021/cr5000283 (2015).

159. Kast, S. M. & Kloss, T. Closed-form expressions of the chemical potential for integral equation closures with certain bridge functions. *J. Chem. Phys.* **129,** 236101; 10.1063/1.3041709 (2008).

160. Kovalenko, A. & Hirata, F. Hydration free energy of hydrophobic solutes studied by a reference interaction site model with a repulsive bridge correction and a thermodynamic perturbation method. *J. Chem. Phys.* **113,** 2793–2805; 10.1063/1.1305885 (2000).

161. Lorentz, H. A. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. *Ann. Phys.* **248,** 127–136; 10.1002/andp.18812480110 (1881).

162. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J .Comput. Chem.* **25,** 1157–1174; 10.1002/jcc.20035 (2004).

163. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25,** 247–260; 10.1016/j.jmgm.2005.12.005 (2006).

164. Ten-no, S., Hirata, F. & Kato, S. A hybrid approach for the solvent effect on the electronic structure of a solute based on the RISM and Hartree-Fock equations. *Chem. Phys. Lett.* **214,** 391–396; 10.1016/0009-2614(93)85655-8 (1993).

165. Ten-no, S., Hirata, F. & Kato, S. Reference interaction site model self-consistent field study for solvation effect on carbonyl compounds in aqueous solution. *J. Chem. Phys.* **100,** 7443–7453; 10.1063/1.466888 (1994).

166. Kovalenko, A. & Hirata, F. Self-consistent description of a metal–water interface by the Kohn–Sham density functional theory and the three-dimensional reference interaction site model. *J. Chem. Phys.* **110,** 10095–10112; 10.1063/1.478883 (1999).

167. Sato, H., Kovalenko, A. & Hirata, F. Self-consistent field, ab initio molecular orbital and three-dimensional reference interaction site model study for solvation effect on carbon monoxide in aqueous solution. *J. Chem. Phys.* **112,** 9463–9468; 10.1063/1.481564 (2000).

168. te Velde, G., Bickelhaupt, F. M., Baerends, E. J., Fonseca Guerra, C., van Gisbergen, S. J. A., Snijders, J. G. & Ziegler, T. Chemistry with ADF. *J .Comput. Chem.* **22,** 931–967; 10.1002/jcc.1056 (2001).

169. Gusarov, S., Ziegler, T. & Kovalenko, A. Self-consistent combination of the three-dimensional RISM theory of molecular solvation with analytical gradients and the Amsterdam density functional package. *J. Phys. Chem. A* **110,** 6083–6090; 10.1021/jp054344t (2006).

170. Casanova, D., Gusarov, S., Kovalenko, A. & Ziegler, T. Evaluation of the SCF Combination of KS-DFT and 3D-RISM-KH; Solvation Effect on Conformational Equilibria, Tautomerization Energies, and Activation Barriers. *J. Chem. Theory Comput.* **3,** 458–476; 10.1021/ct6001785 (2007).

171. Minezawa, N. & Kato, S. Efficient implementation of three-dimensional reference interaction site model self-consistent-field method: application to solvatochromic shift calculations. *J. Chem. Phys.* **126,** 54511; 10.1063/1.2431809 (2007).

172. Aono, S. & Sakaki, S. Evaluation procedure of electrostatic potential in 3D-RISM-SCF method and its application to hydrolyses of cis- and transplatin complexes. *J. Phys. Chem. B.* **116,** 13045–13062; 10.1021/jp307879j (2012).

173. Reimann, M. & Kaupp, M. Evaluation of an Efficient 3D-RISM-SCF Implementation as a Tool for Computational Spectroscopy in Solution. *J. Phys. Chem. A* **124,** 7439–7452; 10.1021/acs.jpca.0c06322 (2020).

174. Hoffgaard, F., Heil, J. & Kast, S. M. Three-Dimensional RISM Integral Equation Theory for Polarizable Solute Models. *J. Chem. Theory. Comput.* **9,** 4718–4726; 10.1021/ct400699q (2013).

175. Ganyecz, Á. & Kállay, M. Implementation and Optimization of the Embedded Cluster Reference Interaction Site Model with Atomic Charges. *J. Phys. Chem. A* **126,** 2417–2429; 10.1021/acs.jpca.1c07904 (2022).

176. Kibies, P. J. Integral equation-based calculations of the electronic structure of small molecules under high pressure, 2019.

177. Palmer, D. S., Frolov, A. I., Ratkova, E. L. & Fedorov, M. V. Toward a universal model to calculate the solvation thermodynamics of druglike molecules: the importance of new experimental databases. *Mol. Pharm.* **8,** 1423–1429; 10.1021/mp200119r (2011).

178. Palmer, D. S., Frolov, A. I., Ratkova, E. L. & Fedorov, M. V. Towards a universal method for calculating hydration free

energies: a 3D reference interaction site model with partial molar volume correction. *J. Phys. Condens. Matter.* **22,** 492101; 10.1088/0953-8984/22/49/492101 (2010).

179. Sergiievskyi, V., Jeanmairet, G., Levesque, M. & Borgis, D. Solvation free-energy pressure corrections in the three dimensional reference interaction site model. *J. Chem. Phys.* **143,** 184116; 10.1063/1.4935065 (2015).

180. Tielker, N., Tomazic, D., Heil, J., Kloss, T., Ehrhart, S., Gussregen, S., Schmidt, K. F. & Kast, S. M. The SAMPL5 challenge for embedded-cluster integral equation theory: solvation free energies, aqueous pK a, and cyclohexane-water log D. *J. Comput. Aided. Mol. Des.* **30,** 1035–1044; 10.1007/s10822-016-9939-7 (2016).

181. Tomazic, D. Optimizing free energy functionals in integral equation theories. Technische Universität Dortmund, 2016.

182. Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Comput. Mol. Sci.* **12**; 10.1002/wcms.1606 (2022).

183. Wilson, E. B., Decius, J. C. & Cross, P. C. *Molecular vibrations. The theory of infrared and raman vibrational spectra* (Dover Publications, New York, post 2006], 1980).

184. Pople, J. A., Krishnan, R., Schlegel, H. B. & Binkley, J. S. Derivative studies in hartree-fock and møller-plesset theories. *Int. J. Quantum Chem.* **16,** 225–241; 10.1002/qua.560160825 (1979).

185. Ochterski, J. W. Vibrational analysis in gaussian (1999).

186. Helgaker, T., Jaszuński, M. & Ruud, K. Ab Initio Methods for the Calculation of NMR Shielding and Indirect Spinminus signSpin Coupling Constants. *Chem. Rev.* **99,** 293–352; 10.1021/cr960017t (1999).

187. Prigogine, I. & Rice, S. A. *Advances in Chemical Physics* (John Wiley & Sons, Inc, Hoboken, NJ, USA, 2002).

188. Ramsey, N. F. A Molecular Beam Resonance Method with Separated Oscillating Fields. *Phys. Rev.* **78,** 695–699; 10.1103/PhysRev.78.695 (1950).

189. Ramsey, N. F. Magnetic Shielding of Nuclei in Molecules. *Phys. Rev.* **78,** 699–703; 10.1103/PhysRev.78.699 (1950).

190. Kutzelnigg, W. Origin and meaning of the Fermi contact interaction. *Theoret. Chim. Acta* **73,** 173–200; 10.1007/BF00528203 (1988).

191. Jensen, F. *Introduction to computational chemistry.* 2nd ed. (Wiley, Chichester, 2007).

192. Jakobsen, P. & Jensen, F. Probing basis set requirements for calculating hyperfine coupling constants. *J. Chem. Phys.* **151,** 174107; 10.1063/1.5128286 (2019).

193. Epstein, S. T. Gauge Invariance of the Hartree-Fock Approximation. *J. Chem. Phys.* **42,** 2897–2898; 10.1063/1.1703259 (1965).

194. Glasbrenner, M., Vogler, S. & Ochsenfeld, C. Gauge-origin dependence in electronic g-tensor calculations. *J. Chem. Phys.* **148,** 214101; 10.1063/1.5028454 (2018).

195. Schindler, M. & Kutzelnigg, W. Theory of magnetic susceptibilities and NMR chemical shifts in terms of localized quantities. II. Application to some simple molecules. *J. Chem. Phys.* **76,** 1919–1933; 10.1063/1.443165 (1982).

196. Hansen, A. E. & Bouman, T. D. Localized orbital/local origin method for calculation and analysis of NMR shieldings. Applications to 13 C shielding tensors. *J. Chem. Phys.* **82,** 5035–5047; 10.1063/1.448625 (1985).

197. London, F. Théorie quantique des courants interatomiques dans les combinaisons aromatiques. *J. Phys. Radium* **8,** 397–409; 10.1051/jphysrad:01937008010039700 (1937).

198. Ditchfield, R. Self-consistent perturbation theory of diamagnetism. *Mol. Phys.* **27,** 789–807; 10.1080/00268977400100711 (1974).

199. Helgaker, T. & Jo/rgensen, P. An electronic Hamiltonian for origin independent calculations of magnetic properties. *J. Chem. Phys.* **95,** 2595–2601; 10.1063/1.460912 (1991).

200. Hall, G. G. Gauge invariant gaussian orbitals and the ab initio calculation of diamagnetic susceptibility for molecules. *Int. J. Quantum Chem.* **7,** 15–25; 10.1002/qua.560070104 (1973).

201. Simmie, J. M. & Somers, K. P. Snakes on the Rungs of Jacob's Ladder: Anomalous Vibrational Spectra from Double-Hybrid DFT Methods. *J. Phys. Chem. A* **124,** 6899–6902; 10.1021/acs.jpca.0c05120 (2020).

202. Vibrational scaling factors. Available at https://cccbdb.nist.gov/vibnotes.asp. 08.09.2022.

203. Sylvetsky, N., Peterson, K. A., Karton, A. & Martin, J. M. L. Toward a W4-F12 approach: Can explicitly correlated and orbital-based ab initio CCSD(T) limits be reconciled? *J. Chem. Phys.* **144,** 214101; 10.1063/1.4952410 (2016).

204. Thorpe, J. H., Lopez, C. A., Nguyen, T. L., Baraban, J. H., Bross, D. H., Ruscic, B. & Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. IV. A modified recipe for computational efficiency. *J. Chem. Phys.* **150,** 224102; 10.1063/1.5095937 (2019).

205. Martin, J. M. L. & Santra, G. Empirical Double-Hybrid Density Functional Theory: A 'Third Way' in Between WFT and DFT. *Isr. J. Chem.* **60,** 787–804; 10.1002/ijch.201900114 (2020).

206. Helgaker, T., Coriani, S., Jørgensen, P., Kristensen, K., Olsen, J. & Ruud, K. Recent advances in wave function-based methods of molecular-property calculations. *Chem. Rev.* **112,** 543–631; 10.1021/cr2002239 (2012).

207. Grayce, C. J. & Harris, R. A. Magnetic-field density-functional theory. *Phys. Rev. A.* **50,** 3089–3095; 10.1103/PhysRevA.50.3089 (1994).

208. Vignale, G. & Rasolt, M. Density-functional theory in strong magnetic fields. *Phys. Rev. Lett.* **59,** 2360–2363; 10.1103/PhysRevLett.59.2360 (1987).

209. Reimann, S., Ekström, U., Stopkowicz, S., Teale, A. M., Borgoo, A. & Helgaker, T. The importance of current contributions to shielding constants in density-functional theory. *Phys. Chem. Chem. Phys.* **17,** 18834–18842; 10.1039/C5CP02682B (2015).

210. Gauss, J. Calculation of NMR chemical shifts at second-order many-body perturbation theory using gauge-including atomic orbitals. *Chem. Phys. Lett.* **191,** 614–620; 10.1016/0009-2614(92)85598-5 (1992).

211. Gauss, J. Effects of electron correlation in the calculation of nuclear magnetic resonance chemical shifts. *J. Chem. Phys.* **99,** 3629–3643; 10.1063/1.466161 (1993).

212. Gauss, J. GIAO-MBPT(3) and GIAO-SDQ-MBPT(4) calculations of nuclear magnetic shielding constants. *Chem. Phys. Lett.* **229,** 198–203; 10.1016/0009-2614(94)01031-5 (1994).

213. Gauss, J. & Stanton, J. F. Gauge-invariant calculation of nuclear magnetic shielding constants at the coupled–cluster singles and doubles level. *J. Chem. Phys.* **102,** 251–253; 10.1063/1.469397 (1995).

214. Gauss, J. & Stanton, J. F. Coupled-cluster calculations of nuclear magnetic resonance chemical shifts. *J. Chem. Phys.* **103,** 3561–3577; 10.1063/1.470240 (1995).

215. Gauss, J. & Stanton, J. F. Perturbative treatment of triple excitations in coupled-cluster calculations of nuclear magnetic shielding constants. *J. Chem. Phys.* **104,** 2574–2583; 10.1063/1.471005 (1996).

216. Flaig, D., Maurer, M., Hanni, M., Braunger, K., Kick, L., Thubauville, M. & Ochsenfeld, C. Benchmarking Hydrogen and Carbon NMR Chemical Shifts at HF, DFT, and MP2 Levels. *J. Chem. Theory Comput.* **10,** 572–578; 10.1021/ct400780f (2014).

217. Stoychev, G. L., Auer, A. A. & Neese, F. Efficient and Accurate Prediction of Nuclear Magnetic Resonance Shielding Tensors with Double-Hybrid Density Functional Theory. *J. Chem. Theory Comput.* **14,** 4756–4771; 10.1021/acs.jctc.8b00624 (2018).

218. Vignale & Rasolt. Current- and spin-density-functional theory for inhomogeneous electronic systems in strong magnetic fields. *Phys. Rev. B* **37,** 10685–10696; 10.1103/PhysRevB.37.10685 (1988).

219. Malkin, V. G., Malkina, O. L. & Salahub, D. R. Calculations of NMR shielding constants beyond uncoupled density functional theory. IGLO approach. *Chem. Phys. Lett.* **204,** 87–95; 10.1016/0009-2614(93)85609-R (1993).

220. Furness, J. W., Verbeke, J., Tellgren, E. I., Stopkowicz, S., Ekström, U., Helgaker, T. & Teale, A. M. Current Density Functional Theory Using Meta-Generalized Gradient Exchange-Correlation Functionals. *J. Chem. Theory Comput.* **11,** 4169–4181; 10.1021/acs.jctc.5b00535 (2015).

221. Teale, A. M., Lutnæs, O. B., Helgaker, T., Tozer, D. J. & Gauss, J. Benchmarking density-functional theory calculations of NMR shielding constants and spin-rotation constants using accurate coupled-cluster calculations. *J. Chem. Phys.* **138,** 24111; 10.1063/1.4773016 (2013).

222. Zhang, Y., Wu, A., Xu, X. & Yan, Y. OPBE: A promising density functional for the calculation of nuclear shielding constants. *Chem. Phys. Lett.* **421,** 383–388; 10.1016/j.cplett.2006.01.095 (2006).

223. Zhao, Y. & Truhlar, D. G. Improved description of nuclear magnetic resonance chemical shielding constants using the M06-L meta-generalized-gradient-approximation density functional. *J. Phys. Chem. A* **112,** 6794–6799; 10.1021/jp804583d (2008).

224. Stoychev, G. L., Auer, A. A., Gauss, J. & Neese, F. DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings. *J. Chem. Phys.* **154,** 164110; 10.1063/5.0047125 (2021).

225. Perdew, J. P. & Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conf. Proc.,* 1–20; 10.1063/1.1390175 (2001).

226. Jensen, F. Basis Set Convergence of Nuclear Magnetic Shielding Constants Calculated by Density Functional Methods. *J. Chem. Theory Comput.* **4,** 719–727; 10.1021/ct800013z (2008).

227. Jensen, F. The Basis Set Convergence of Spin-Spin Coupling Constants Calculated by Density Functional Methods. *J. Chem. Theory Comput.* **2,** 1360–1369; 10.1021/ct600166u (2006).

228. Jensen, F. Segmented contracted basis sets optimized for nuclear magnetic shielding. *J. Chem. Theory Comput.* **11,**

132–138; 10.1021/ct5009526 (2015).

229. Pople, J. A., Beveridge, D. L. & Dobosh, P. A. Approximate Self-Consistent Molecular-Orbital Theory. V. Intermediate Neglect of Differential Overlap. *J. Chem. Phys.* **47,** 2026–2033; 10.1063/1.1712233 (1967).

230. Beveridge, D. L. & McIver, J. W. INDO Molecular Orbital Study of Hyperfine Tensors: Theory, Methodology, and Applications to CH, CH 3 , and Radicaloid Derivatives of Malonic Acid. *J. Chem. Phys.* **54,** 4681–4690; 10.1063/1.1674742 (1971).

231. Chipman, D. M. Comment on ab initio calculation of spin densities in hydrocarbon radicals. *J. Chem. Phys.* **78,** 4785–4786; 10.1063/1.445283 (1983).

232. Chipman, D. M. Abinitio calculation of spin densities in hydrocarbon radicals. *J. Chem. Phys.* **71,** 761–768; 10.1063/1.438364 (1979).

233. Bauschlicher, C. W. Theoretical study of the nitrogen-atom hyperfine coupling constant. II. *J. Chem. Phys.* **92,** 518–521; 10.1063/1.458600 (1990).

234. Ishii, M., Morihashi, K. & Kikuchi, O. Ab initio calculations of g values of free radicals by finite perturbation theory. *J. Mol. Struct. Theochem* **235,** 39–49; 10.1016/0166-1280(91)85083-J (1991).

235. Hermosilla, L., Calle, P., La García de Vega, J. M. & Sieiro, C. Density functional theory predictions of isotropic hyperfine coupling constants. *J. Phys. Chem. A* **109,** 1114–1124; 10.1021/jp0466901 (2005).

236. Schreckenbach, G. & Ziegler, T. Calculation of the G-Tensor of Electron Paramagnetic Resonance Spectroscopy Using Gauge-Including Atomic Orbitals and Density Functional Theory. *J. Phys. Chem. A* **101,** 3388–3399; 10.1021/jp963060t (1997).

237. Schreckenbach, G. & Ziegler, T. Density functional calculations of NMR chemical shifts and ESR g-tensors. *Theoret. Chim. Acta* **99,** 71–82; 10.1007/s002140050306 (1998).

238. Patchkovskii, S. & Ziegler, T. Calculation of the EPR g-Tensors of High-Spin Radicals with Density Functional Theory. *J. Phys. Chem. A* **105,** 5490–5497; 10.1021/jp010457a (2001).

239. Malkina, O. L., Vaara, J., Schimmelpfennig, B., Munzarová, M., Malkin, V. G. & Kaupp, M. Density Functional Calculations of Electronic g -Tensors Using Spin−Orbit Pseudopotentials and Mean-Field All-Electron Spin−Orbit Operators. *J. Am. Chem. Soc.* **122,** 9206–9218; 10.1021/ja000984s (2000).

240. Puzzarini, C. & Barone, V. Toward spectroscopic accuracy for organic free radicals: Molecular structure, vibrational spectrum, and magnetic properties of F(2)NO. *J. Chem. Phys.* **129,** 84306; 10.1063/1.2969820 (2008).

241. Barone, V., Fusè, M., Pinto, S. M. V. & Tasinato, N. A Computational Journey across Nitroxide Radicals: From Structure to Spectroscopic Properties and Beyond. *Molecules* **26**; 10.3390/molecules26237404 (2021).

242. Cammi, R., Mennucci, B. & Tomasi, J. Nuclear magnetic shieldings in solution: Gauge invariant atomic orbital calculation using the polarizable continuum model. *J. Chem. Phys.* **110,** 7627–7638; 10.1063/1.478674 (1999).

243. Cammi, R. The Hartree–Fock calculation of the magnetic properties of molecular solutes. *J. Chem. Phys.* **109,** 3185–3196; 10.1063/1.476910 (1998).

244. Mennucci, B., Cammi, R. & Tomasi, J. Analytical free energy second derivatives with respect to nuclear coordinates: Complete formulation for electrostatic continuum solvation models. *J. Chem. Phys.* **110,** 6858–6870; 10.1063/1.478591 (1999).

245. Adamo, C., Heitzmann, M., Meilleur, F., Rega, N., Scalmani, G., Grand, A., Cadet, J. & Barone, V. Interplay of intrinsic and environmental effects on the magnetic properties of free radicals issuing from H-atom addition to cytosine. *J. Am. Chem. Soc.* **123,** 7113–7117; 10.1021/ja004284z (2001).

246. Rega, N., Cossi, M. & Barone, V. Development and validation of reliable quantum mechanical approaches for the study of free radicals in solution. *J. Chem. Phys.* **105,** 11060–11067; 10.1063/1.472906 (1996).

247. Rega, N., Cossi, M. & Barone, V. Intrinsic and Environmental Effects in the Structure and Magnetic Properties of Glycine Radical in Aqueous Solution. *J. Am. Chem. Soc.* **119,** 12962–12967; 10.1021/ja971838y (1997).

248. Frach, R. & Kast, S. M. Solvation effects on chemical shifts by embedded cluster integral equation theory. *J. Phys. Chem. A* **118,** 11620–11628; 10.1021/jp5084407 (2014).

249. Manalo, M. N. & Dios, A. C. de. Anab initio study of solvent polarity and hydrogen bonding effects on the nitrogen NMR shieldings ofN,N-dimethylacetamidine. *Magn. Reson. Chem.* **40,** 781–785; 10.1002/mrc.1095 (2002).

250. Klamt, A. The COSMO and COSMO-RS solvation models. *WIREs Comput. Mol. Sci.* **1,** 699–709; 10.1002/wcms.56 (2011).

251. Sinnecker, S., Rajendran, A., Klamt, A., Diedenhofen, M. & Neese, F. Calculation of solvent shifts on electronic g-tensors with the conductor-like screening model (COSMO) and its self-consistent generalization to real solvents (direct COSMO-RS). *J. Phys. Chem. A* **110,** 2235–2245; 10.1021/jp056016z (2006).

252. Klamt, A. & Diedenhofen, M. Calculation of solvation free energies with DCOSMO-RS. *J. Phys. Chem. A* **119,** 5439–5445; 10.1021/jp511158y (2015).

253. Nishiyama, K., Watanabe, Y., Yoshida, N. & Hirata, F. Solvent dependence of Stokes shift for organic solute-solvent systems: A comparative study by spectroscopy and reference interaction-site model-self-consistent-field theory. *J. Chem. Phys.* **139,** 94503; 10.1063/1.4819268 (2013).

254. Tanaka, Y., Yoshida, N. & Nakano, H. Solvent effect on excited states of merocyanines: A theoretical study using the RISM–SCF method. *Chem. Phys. Lett.* **583,** 69–73; 10.1016/j.cplett.2013.08.004 (2013).

255. Yokogawa, D. Time-dependent density functional theory (TD-DFT) coupled with reference interaction site model self-consistent field explicitly including spatial electron density distribution (RISM-SCF-SEDD). *J. Chem. Phys.* **145,** 94101; 10.1063/1.4962062 (2016).

256. Hirano, K., Nakano, H., Nakao, Y., Sato, H. & Sakaki, S. Photo absorption of p-coumaric acid in aqueous solution: RISM-SCF-SEDD theory approach. *J .Comput. Chem.* **38,** 1567–1573; 10.1002/jcc.24784 (2017).

257. Suda, K., Sarinastiti, A., Arifin, Kimura, Y. & Yokogawa, D. Understanding Structural Changes through Excited-State Intramolecular Proton Transfer in 4'-N,N-Diethylamino-3-hydroxyflavone (DEAHF) in Solution Based on Quantum Chemical Calculations. *J. Phys. Chem. B.* **123,** 9872–9881; 10.1021/acs.jpcb.9b07549 (2019).

258. Kaminski, J. W., Gusarov, S., Wesolowski, T. A. & Kovalenko, A. Modeling solvatochromic shifts using the orbital-free embedding potential at statistically mechanically averaged solvent density. *J. Phys. Chem. A* **114,** 6082–6096; 10.1021/jp100158h (2010).

259. Yang, C., Watanabe, Y., Yoshida, N. & Nakano, H. Three-Dimensional Reference Interaction Site Model Self-Consistent Field Study on the Coordination Structure and Excitation Spectra of Cu(II)-Water Complexes in Aqueous Solution. *J. Phys. Chem. A* **123,** 3344–3354; 10.1021/acs.jpca.9b01364 (2019).

260. Improta, R. & Barone, V. Interplay of electronic, environmental, and vibrational effects in determining the hyperfine coupling constants of organic free radicals. *Chem. Rev.* **104,** 1231–1254; 10.1021/cr960085f (2004).

261. Pavone, M., Cimino, P., Angelis, F. de & Barone, V. Interplay of stereoelectronic and enviromental effects in tuning the structural and magnetic properties of a prototypical spin probe: further insights from a first principle dynamical approach. *J. Am. Chem. Soc.* **128,** 4338–4347; 10.1021/ja0574872 (2006).

262. Houriez, C., Ferré, N., Siri, D. & Masella, M. Further insights into the environmental effects on the computed hyperfine coupling constants of nitroxides in aqueous solution. *J. Phys. Chem. B.* **113,** 15047–15056; 10.1021/jp906828v (2009).

263. Neugebauer, J., Louwerse, M. J., Belanzoni, P., Wesolowski, T. A. & Baerends, E. J. Modeling solvent effects on electron-spin-resonance hyperfine couplings by frozen-density embedding. *J. Chem. Phys.* **123,** 114101; 10.1063/1.2033749 (2005).

264. Improta, R., Di Matteo, A. & Barone, V. Effective modeling of intrinsic and environmental effects on the structure and electron plaramagnetic resonance parameters of nitroxides by an integrated quantum mechanical/molecular mechanics/polarizable continuum model approach. *Theoret. Chim. Acta* **104,** 273–279; 10.1007/s002140000122 (2000).

265. Houriez, C., Ferré, N., Masella, M. & Siri, D. Prediction of nitroxide hyperfine coupling constants in solution from combined nanosecond scale simulations and quantum computations. *J. Chem. Phys.* **128,** 244504; 10.1063/1.2939121 (2008).

266. Sezer, D., Freed, J. H. & Roux, B. Parametrization, molecular dynamics simulation, and calculation of electron spin resonance spectra of a nitroxide spin label on a polyalanine alpha-helix. *J. Phys. Chem. B* **112,** 5755–5767; 10.1021/jp711375x (2008).

267. Mazurek, A. H., Szeleszczuk, Ł. & Pisklak, D. M. A Review on Combination of Ab Initio Molecular Dynamics and NMR Parameters Calculations. *Int. J. Mol. Sci.* **22**; 10.3390/ijms22094378 (2021).

268. Dračínský, M., Möller, H. M. & Exner, T. E. Conformational Sampling by Ab Initio Molecular Dynamics Simulations Improves NMR Chemical Shift Predictions. *J. Chem. Theory Comput.* **9,** 3806–3815; 10.1021/ct400282h (2013).

269. Flaig, D., Beer, M. & Ochsenfeld, C. Convergence of Electronic Structure with the Size of the QM Region: Example of QM/MM NMR Shieldings. *J. Chem. Theory Comput.* **8,** 2260–2271; 10.1021/ct300036s (2012).

270. Mathias, G. & Baer, M. D. Generalized Normal Coordinates for the Vibrational Analysis of Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **7,** 2028–2039; 10.1021/ct2001304 (2011).

271. Mathias, G., Ivanov, S. D., Witt, A., Baer, M. D. & Marx, D. Infrared Spectroscopy of Fluxional Molecules from (ab Initio) Molecular Dynamics: Resolving Large-Amplitude Motion, Multiple Conformations, and Permutational Symmetries. *J. Chem. Theory Comput.* **8,** 224–234; 10.1021/ct2006665 (2012).

272. Baer, M., Marx, D. & Mathias, G. Theoretical messenger spectroscopy of microsolvated hydronium and Zundel cations. *Angew. Chem. Int. Ed. Engl.* **49,** 7346–7349; 10.1002/anie.201001672 (2010).

273. Mathias, G. & Marx, D. Structures and spectral signatures of protonated water networks in bacteriorhodopsin. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 6980–6985; 10.1073/pnas.0609229104 (2007).

274. 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, J. M., 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* (Gaussian Inc., Wallingford CT, 2009).

275. Heil, J. Effiziente Fluidphasentheorie für Protonierungsprozesse in komplexen Systemen. Technische Universität Dortmund, 2016.

276. Kast, K. M., Brickmann, J., Kast, S. M. & Berry, R. S. Binary Phases of Aliphatic N -Oxides and Water. Force Field Development and Molecular Dynamics Simulation. *J. Phys. Chem. A* **107,** 5342–5351; 10.1021/jp027336a (2003).

277. Lee, M. W. & Meuwly, M. On the role of nonbonded interactions in vibrational energy relaxation of cyanide in water. *J. Phys. Chem. A* **115,** 5053–5061; 10.1021/jp202503m (2011).

278. McConnell, H. M. & Chesnut, D. B. Theory of Isotropic Hyperfine Interactions in π-Electron Radicals. *J. Chem. Phys.* **28,** 107–117; 10.1063/1.1744052 (1958).

279. McConnell, H. M. Electron Densities in Semiquinones by Paramagnetic Resonance. *J. Chem. Phys.* **24,** 632; 10.1063/1.1742580 (1956).

280. McConnell, H. M. Indirect Hyperfine Interactions in the Paramagnetic Resonance Spectra of Aromatic Free Radicals. *J. Chem. Phys.* **24,** 764–766; 10.1063/1.1742605 (1956).

281. Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Comput. Mol. Sci.* **8,** e1327; 10.1002/wcms.1327 (2018).

282. Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2,** 73–78; 10.1002/wcms.81 (2012).

283. Stoll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178,** 42–55; 10.1016/j.jmr.2005.08.013 (2006).

284. Luong, T. Q., Kapoor, S. & Winter, R. Pressure-A Gateway to Fundamental Insights into Protein Solvation, Dynamics, and Function. *Chemphyschem* **16,** 3555–3571; 10.1002/cphc.201500669 (2015).

285. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B. & Fox, D. J. *Gaussian 16* (Gaussian Inc., Wellington CT, 2016).

286. D.A. Case, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman. *AmberTools21* (San Francisco, 2021).

287. Wickstrom, L., Okur, A. & Simmerling, C. Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophys. J.* **97,** 853–856; 10.1016/j.bpj.2009.04.063 (2009).

288. Sigalov, G., Scheffel, P. & Onufriev, A. Incorporating variable dielectric environments into the generalized Born model. *J. Chem. Phys.* **122,** 94511; 10.1063/1.1857811 (2005).

289. Sigalov, G., Fenley, A. & Onufriev, A. Analytical electrostatics for biomolecules: beyond the generalized Born approximation. *J. Chem. Phys.* **124,** 124902; 10.1063/1.2177251 (2006).

290. Perkyns, J. S. & Pettitt, B. M. A dielectrically consistent interaction site theory for solvent—electrolyte mixtures. *Chem. Phys. Lett.* **190,** 626–630; 10.1016/0009-2614(92)85201-K (1992).

291. Perkyns, J. & Pettitt, B. M. A site–site theory for finite concentration saline solutions. *J. Chem. Phys.* **97,** 7656–7666; 10.1063/1.463485 (1992).

292. Floriano, W. B. & Nascimento, M. A. C. Dielectric constant and density of water as a function of pressure at constant

temperature. *Braz. J. Phys.* **34,** 38–41; 10.1590/S0103-97332004000100006 (2004).

293. Galazzo, L. Unpublished results.

294. Tran, V. A., personal communication.

295. Teucher, M. Unpublished results.

296. Schultz, K. M. & Klug, C. S. High-pressure EPR spectroscopy studies of the E. coli lipopolysaccharide transport proteins LptA and LptC. *Appl. Magn. Reson.* **48,** 1341–1353; 10.1007/s00723-017-0948-z (2017).

297. Jauer, L. Bachelor thesis. TU Dortmund University, 2022.

298. Beck Erlach, M. Unpublished results.

299. Forsting, T., Gottschalk, H. C., Hartwig, B., Mons, M. & Suhm, M. A. Correcting the record: the dimers and trimers of trans-N-methylacetamide. *Phys. Chem. Chem. Phys.* **19,** 10727–10737; 10.1039/C6CP07989J (2017).

300. Maste, S., personal communication.

301. Han, W.-G., Jalkanen, K. J., Elstner, M. & Suhai, S. Theoretical Study of Aqueous N -Acetyl- l -alanine N ‘-Methylamide. Structures and Raman, VCD, and ROA Spectra. *J. Phys. Chem. B* **102,** 2587–2602; 10.1021/jp972299m (1998).

302. Deng, Z., Polavarapu, P. L., Ford, S. J., Hecht, L., Barron, L. D., Ewig, C. S. & Jalkanen, K. Solution-Phase Conformations of N -Acetyl- N ‘-methyl- l -alaninamide from Vibrational Raman Optical Activity. *J. Phys. Chem.* **100,** 2025–2034; 10.1021/jp951865f (1996).

303. Grdadolnik, J., Grdadolnik, S. G. & Avbelj, F. Determination of conformational preferences of dipeptides using vibrational spectroscopy. *J. Phys. Chem. B* **112,** 2712–2718; 10.1021/jp7096313 (2008).

304. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7,** 95–99; 10.1016/s0022-2836(63)80023-6 (1963).

305. Poon, C.-D., Samulski, E. T., Weise, C. F. & Weisshaar, J. C. Do Bridging Water Molecules Dictate the Structure of a Model Dipeptide in Aqueous Solution? *J. Am. Chem. Soc.* **122,** 5642–5643; 10.1021/ja993953%2B (2000).

306. Smith, P. E. The alanine dipeptide free energy surface in solution. *J. Chem. Phys.* **111,** 5568–5579; 10.1063/1.479860 (1999).

307. Hölzl, C., personal communication.

308. Körning, S., personal communication.