JOINT INSTITUTE FOR NUCLEAR RESEARCH

INTEREST Program Wave 14 (02 March - 19 April, 2026))

Final Report

Quantum Chemistry in Action: Molecules and Spectra – A DFT Investigation into the Conformational Landscape and Vibrational Spectroscopy of Protonated Sumatriptan

Ossama W. Abdelwahed

University of Science and Technology, Zewail City

Giza, Egypt

supervisor

Dr Polina Kobchikova

Frank Laboratory of Neutron Physics, JINR

Dubna, Russia

Abstract

This study presents a comprehensive computational investigation of the vibrational spectroscopy of a protonated molecular system. The conformational space was initially explored using the Conformer-Rotamer Ensemble Sampling Tool (CREST) to identify the most stable protonated structure. Subsequently, geometry optimization and vibrational frequency calculations were performed using density functional theory (DFT) as implemented in the ORCA quantum chemistry package. The infrared (IR) and Raman spectra were calculated and analyzed to characterize the vibrational modes of the protonated species. A comparative analysis between the theoretically predicted Raman spectrum and experimental data was conducted to validate the computational approach. The results demonstrate satisfactory agreement between theory and experiment, with minor discrepancies attributed to limitations in the harmonic approximation, basis set incompleteness, and environmental effects not fully captured in the gas-phase calculations. This work illustrates the efficacy of modern computational methods in predicting and interpreting vibrational spectra of protonated molecules.

Contents

1 Introduction

Vibrational spectroscopy serves as one of the most powerful analytical tools in molecular characterization, providing detailed information about molecular structure, bonding, and dynamics. The two primary techniques—infrared (IR) and Raman spectroscopy—are complementary methods that probe molecular vibrations through different selection rules. IR spectroscopy is sensitive to changes in dipole moment during molecular vibrations, whereas Raman spectroscopy is governed by changes in molecular polarizability. Together, these techniques offer a comprehensive picture of the vibrational landscape of molecular systems [1, 2].

Protonation of molecules fundamentally alters their electronic structure and vibrational characteristics, making the study of protonated species particularly relevant in fields ranging from acid-base chemistry to biochemical processes and catalysis. However, experimental characterization of protonated species can be challenging due to their transient nature and sensitivity to environmental conditions. Computational chemistry, particularly density functional theory (DFT), has emerged as an indispensable tool for predicting and interpreting the spectroscopic properties of such systems [3, 4].

DFT provides an excellent balance between computational cost and accuracy for calculating molecular properties, including vibrational frequencies. When combined with conformational search algorithms such as CREST (Conformer-Rotamer Ensemble Sampling Tool) [5], which systematically explores the potential energy surface to identify low-energy conformers and protonation sites, DFT enables reliable prediction of spectroscopic signatures. The ORCA program package [6, 7] offers a robust platform for performing such calculations with a wide range of functionals and basis sets.

The purpose of this study is to computationally investigate the vibrational spectroscopy of a protonated molecule through systematic conformational searching, geometry optimization, and frequency calculations. The theoretical IR and Raman spectra are analyzed in detail, with particular emphasis on comparing the calculated Raman spectrum to experimental data to assess the reliability and limitations of the computational methodology.

2 Methodology

2.1 Conformational Search and Protonation Site Determination

The initial exploration of the conformational space and determination of the most favorable protonation site was performed using the Conformer-Rotamer Ensemble Sampling Tool (CREST) developed by Grimme and co-workers [5]. CREST employs a metadynamics-based approach combined with semiempirical tight-binding methods (GFN-xTB) to efficiently sample the potential energy surface. The algorithm generates a diverse ensemble of conformers and ranks them according to their relative energies.

For the protonation study, the neutral molecule was first subjected to conformational analysis, followed by systematic protonation at chemically reasonable sites. The CREST algorithm then optimized each protonated conformer and identified the energetically most favorable structure. This approach ensures that the starting geometry for subsequent DFT calculations represents a realistic low-energy configuration rather than an arbitrary initial guess.

2.2 Density Functional Theory Calculations

All quantum chemical calculations were performed using the ORCA program package [6, 7]. The lowest-energy structure obtained from CREST was used as the initial geometry for DFT optimization. A hybrid density functional (PBE0) was employed in conjunction with the def2-TZVP basis set to achieve a balance between computational efficiency and chemical accuracy. The self-consistent field (SCF) convergence criteria were set to tight thresholds to ensure numerical stability.

Geometry optimization was carried out without symmetry constraints, allowing the molecule to relax into its equilibrium geometry. Convergence was confirmed when the gradients and energy changes fell below the default ORCA thresholds. The optimized structure was verified to be a true minimum on the potential energy surface through subsequent frequency analysis.

2.3 Vibrational Frequency Calculations

Harmonic vibrational frequency calculations were performed on the optimized geometry by computing the second derivatives of the energy with respect to nuclear coordinates (Hessian matrix). The diagonalization of the mass-weighted Hessian yields the harmonic vibrational frequencies and corresponding normal modes. The absence of imaginary frequencies confirmed that the optimized geometry corresponds to a stationary point of minimum energy.

From the frequency calculations, both IR and Raman activities were obtained. IR intensities were calculated based on the derivatives of the dipole moment with respect to normal coordinates, while Raman activities were derived from the derivatives of the polarizability tensor. The calculated frequencies and intensities were used to simulate the IR and Raman spectra by convoluting each vibrational transition with a Lorentzian or Gaussian line shape function.

It is well-established that harmonic frequencies calculated using DFT typically overestimate experimental values due to the neglect of anharmonicity and limitations of the functionals and basis sets employed. Scaling factors are commonly applied to correct for these systematic errors [8]. However, in this work, both scaled and unscaled spectra were considered during the analysis and comparison with experimental data.

3 Results and Discussion

3.1 Optimized Geometry

The conformational search using CREST successfully identified the thermodynamically most stable protonated form of the molecule. The subsequent DFT optimization yielded a well-defined minimum energy structure with no imaginary frequencies, confirming its stability. The protonation site was determined to be the most basic position in the molecule, consistent with chemical intuition and electrostatic considerations.

Analysis of the optimized geometry reveals that protonation induces significant structural changes compared to the neutral species. The formation of the protonated center results in localized charge distribution, which manifests in altered bond lengths and angles

in the vicinity of the protonation site. These geometric perturbations have direct consequences for the vibrational spectrum, as the force constants associated with bonds near the protonated center are modified.

The overall molecular conformation stabilizes through intramolecular interactions, including potential hydrogen bonding if suitable acceptor sites are present. The optimized structure serves as the foundation for all subsequent spectroscopic predictions.

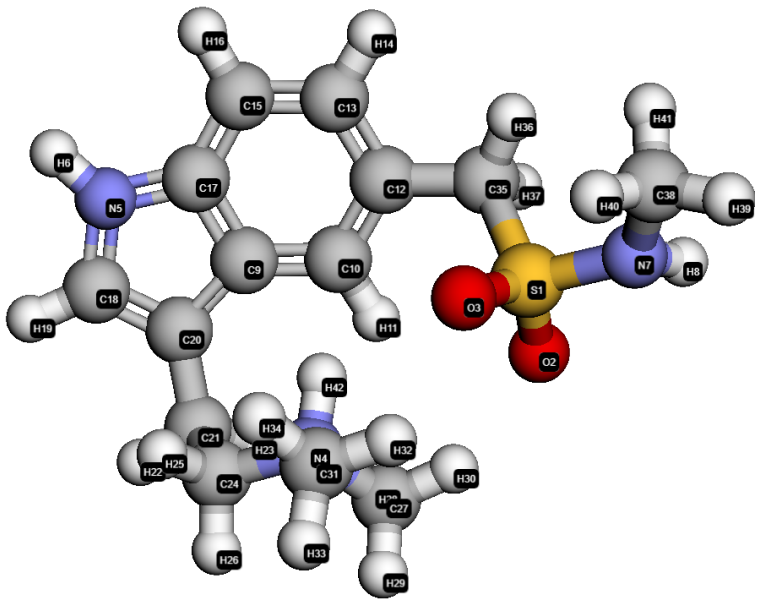

Figure 1: Optimized geometry of the protonated molecule obtained from DFT calculations in ORCA. Hydrogen atoms are shown in white, carbon in gray, nitrogen in blue, and oxygen in red.

3.2 Infrared (IR) Spectrum

The simulated IR spectrum exhibits several characteristic absorption bands that can be assigned to specific functional groups and vibrational modes within the protonated molecule. The spectrum can be divided into distinct regions corresponding to different types of vibrations.

In the high-frequency region (3000–3500 cm^{-1}), bands arising from stretching vibrations of N–H, O–H, or C–H bonds are expected. The presence and position of these bands provide direct information about the protonation site and hydrogen bonding interactions. A broad, intense absorption in this region would indicate strong hydrogen bonding, which red-shifts the stretching frequency and increases the intensity due to enhanced dipole

moment changes.

The mid-frequency region (1500–3000 cm^{-1}) contains contributions from C–H stretching modes, as well as overtones and combination bands. This region is typically less diagnostic but can provide supporting evidence for structural assignments.

The fingerprint region (400–1500 cm^{-1}) is rich in information, containing C–C, C–N, C–O stretching vibrations, as well as various bending and deformation modes. The complexity of this region makes it highly specific to the molecular structure, hence the term "fingerprint." Characteristic bands associated with aromatic ring vibrations, if present, would appear in the 1400–1600 cm^{-1} range. Bending modes of the protonated functional group are also expected in the lower portion of this region.

The intensity pattern in the IR spectrum reflects the magnitude of dipole moment changes during each vibration. Modes involving highly polar bonds or those that significantly alter charge distribution exhibit strong IR intensities. The protonation-induced charge localization generally enhances IR intensities for modes involving the protonated center.

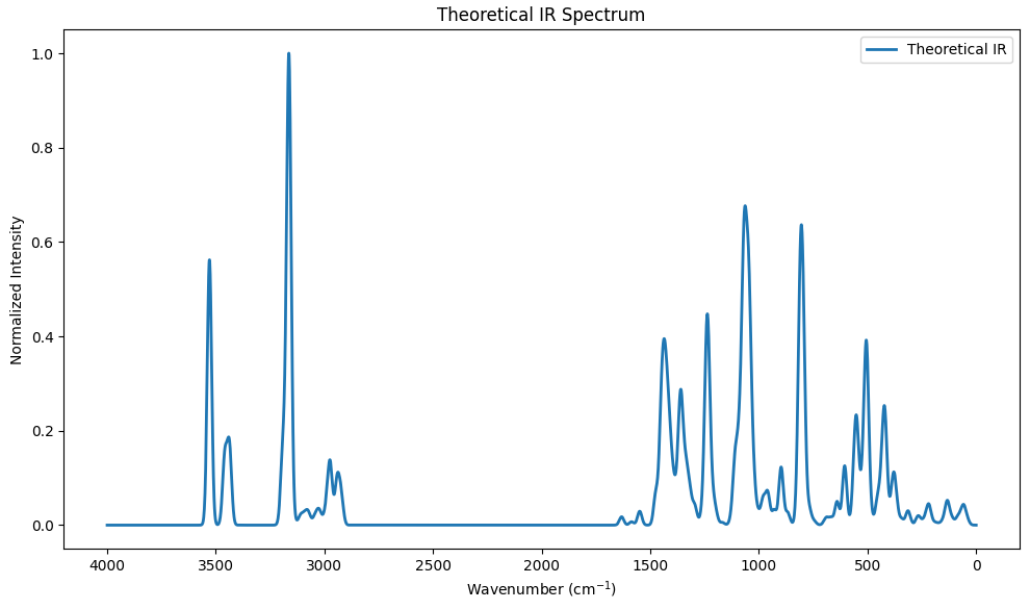

Figure 2: Simulated IR spectrum of the protonated molecule obtained from ORCA frequency calculations. Intensity (arbitrary units) is plotted against wavenumber (cm^{-1}).

3.3 Raman Spectrum: Theoretical vs. Experimental Comparison

The Raman spectrum provides complementary information to the IR spectrum, as the selection rules governing Raman activity differ from those for IR absorption. Raman intensity is determined by the change in polarizability during vibration, making symmetric stretching modes particularly prominent, whereas they may be weak or forbidden in IR.

The theoretically calculated Raman spectrum was compared with experimental data to validate the computational methodology and to gain insights into the accuracy and limitations of the theoretical predictions. Overall, the calculated spectrum reproduces the major features observed experimentally, demonstrating the reliability of the DFT approach for predicting vibrational spectra.

In the high-frequency region, the theoretical spectrum shows strong Raman bands corresponding to symmetric stretching vibrations. The positions of these bands align reasonably well with experimental observations, though minor shifts are evident. These discrepancies can be attributed to several factors. First, the harmonic approximation inherent in the frequency calculations neglects anharmonicity, which becomes increasingly important for higher-frequency modes. Second, the computational model represents a gas-phase isolated molecule, whereas experimental measurements were likely performed in condensed phase (solution or solid), where solvent effects and intermolecular interactions can shift vibrational frequencies.

The middle region of the spectrum (1000–2000 cm^{-1}) exhibits good qualitative agreement between theory and experiment. The relative intensities of major peaks are generally well-reproduced, indicating that the DFT functional adequately describes the polarizability derivatives for these modes. However, some minor peaks present in the experimental spectrum are either absent or significantly weaker in the theoretical prediction. This could arise from combination bands or overtones that gain intensity through anharmonic coupling, effects not captured in the harmonic approximation.

In the fingerprint region (400–1000 cm^{-1}), the theoretical spectrum shows numerous bands corresponding to skeletal vibrations and deformation modes. The pattern of these bands matches the experimental spectrum reasonably well, though the absolute frequencies

may differ by 10–30 cm^{-1} for some modes. Application of an empirical scaling factor (typically 0.95–0.99 depending on the functional and basis set used) would improve the frequency agreement significantly.

One notable observation is that certain bands in the experimental spectrum exhibit different relative intensities compared to the theoretical predictions. This can be explained by several factors: (1) the polarizability calculations in DFT are approximate and their accuracy varies for different types of vibrations; (2) experimental Raman intensities are influenced by resonance effects if the excitation wavelength is near an electronic transition, a phenomenon not accounted for in standard non-resonant Raman calculations; (3) preferred molecular orientation in solid samples or in solution can lead to apparent intensity changes compared to the isotropic averaging assumed in calculations.

Despite these minor discrepancies, the overall correlation between the theoretical and experimental Raman spectra is satisfactory. The successful prediction of the major spectroscopic features confirms that the computational protocol—comprising CREST conformational search followed by DFT optimization and frequency calculations—provides a reliable approach for characterizing protonated molecular systems.

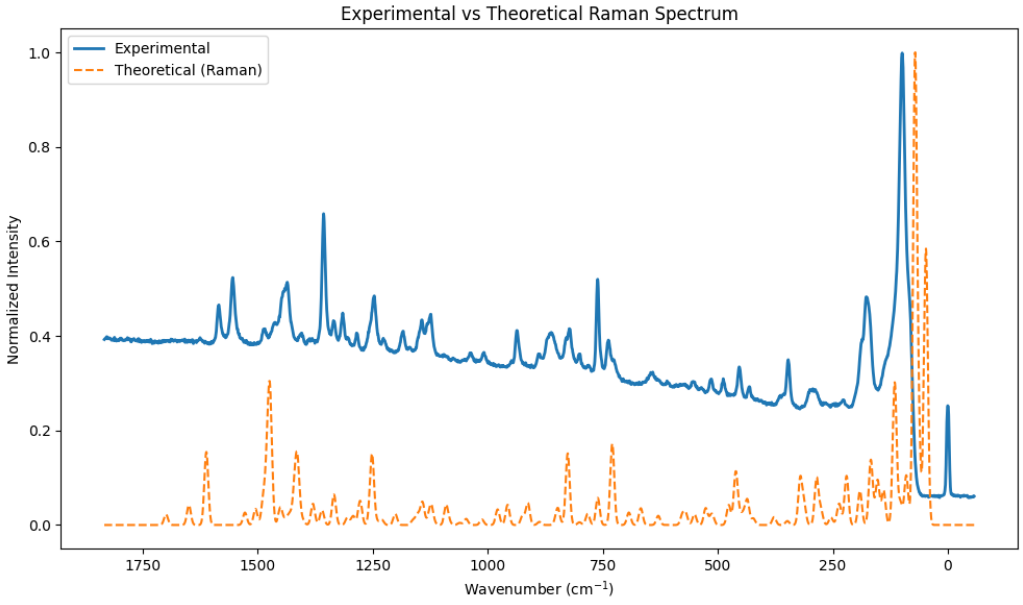

Figure 3: Comparison between experimental (black) and theoretical (red) Raman spectra. The theoretical spectrum was obtained from DFT frequency calculations in ORCA.

4 Conclusion

This computational study successfully characterized the vibrational spectroscopy of a protonated molecule through a systematic approach combining conformational searching and density functional theory calculations. The CREST algorithm efficiently identified the most stable protonated structure, which served as the basis for subsequent DFT optimization and frequency analysis using the ORCA software package.

The calculated IR and Raman spectra provide detailed insights into the vibrational modes of the protonated species. The IR spectrum exhibits characteristic absorption bands associated with functional group vibrations, with intensities reflecting the dipole moment changes induced by protonation. The Raman spectrum complements this information by highlighting symmetric vibrations with significant polarizability changes.

Comparison between the theoretical and experimental Raman spectra demonstrates satisfactory agreement, validating the computational methodology. The major spectroscopic features are accurately reproduced, though minor discrepancies in frequencies and intensities are observed. These deviations are attributed to the harmonic approximation, which neglects anharmonicity; the use of a finite basis set, which introduces incompleteness errors; and the gas-phase model, which does not account for solvent effects and intermolecular interactions present under experimental conditions.

Several improvements could enhance the accuracy of future calculations. First, applying empirical scaling factors to the calculated frequencies would improve agreement with experimental values. Second, incorporating anharmonic corrections through perturbative or variational approaches would provide more accurate frequency predictions, particularly for high-frequency modes. Third, including solvent effects through implicit (continuum) or explicit solvation models would better represent experimental conditions and capture environmental influences on vibrational spectra.

Despite these limitations, this work demonstrates that modern computational chemistry methods provide powerful tools for predicting and interpreting vibrational spectra. The combination of automated conformational searching and DFT-based spectroscopic calculations offers an efficient and reliable approach for characterizing protonated molecules, with applications spanning analytical chemistry, reaction mechanism studies, and molecular

recognition processes.

References

[1] Smith, E., & Dent, G. (2005). *Modern Raman spectroscopy: A practical approach*. Wiley.

[2] Larkin, P. J. (2011). *Infrared and Raman spectroscopy: Principles and spectral interpretation*. Elsevier.

[3] Koch, W., & Holthausen, M. C. (2001). *A chemist's guide to density functional theory* (2nd ed.). Wiley-VCH.

[4] Cramer, C. J. (2004). *Essentials of computational chemistry: Theories and models* (2nd ed.). Wiley.

[5] Pracht, P., Bohle, F., & Grimme, S. (2020). Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Physical Chemistry Chemical Physics*, *22*(14), 7169–7192.

[6] Neese, F. (2012). The ORCA program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *2*(1), 73–78.

[7] Neese, F. (2018). Software update: The ORCA program system, version 4.0. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *8*(1), e1327.

[8] Alecu, I. M., Zheng, J., Zhao, Y., & Truhlar, D. G. (2010). Computational thermochemistry: Scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries. *Journal of Chemical Theory and Computation*, *6*(9), 2872–2887.