Joint Institute for Nuclear Research

Laboratory of Nuclear Problem

Final Report on the INTEREST Program

Quantum Chemistry in Action: Molecules, Spectra, and Crystals

Supervisor: Dr Polina Kobchikova

(*Frank Laboratory of Neutron Physics*)

Student: Aadhyatri Pathak

(Institute of Chemical Technology, Mumbai)

Participation period:

March 02 - April 19

JINR Wave 14, Summer Session 2026

Abstract

The current report deals with an explorative and theoretical approach to investigating computational approaches to analysing molecular structure and vibrations, employing sumatriptan as an example of flexible organic molecules. It covers theoretical background in the field of conformational analysis, density functional theory (DFT), and vibrational spectroscopy (primarily Raman spectroscopy as a method sensitive to changes in the molecular polarizability).

The basic steps of the process - preparation of the molecular structure, conformational sampling, molecular geometry optimization, and analysis of vibrational frequencies using semi-empirical approximation, DFT and Raman spectroscopy respectively - are discussed according to the existing literature and recommended best practices. Special attention is paid to the issue of the variety of conformations, impact of the electronic structure on vibrational modes, and theoretical aspects of interpretation of spectroscopic results.

The attempt to execute this process using appropriate software (CREST, ORCA) proved unsuccessful due to limitations of the time available for the research and problems with installation of the required software packages and their operation.

Despite this limitation, the current report lays the groundwork for more detailed investigations in terms of methodology and interpretation, allowing for computational implementation at the later stage of research.

Table of Contents:

1. Introduction

The development of computational chemistry has greatly changed how we study molecular systems. It offers a theoretical framework for understanding structural, electronic, and spectroscopic properties at the atomic level. Traditional experimental techniques like infrared (IR) and Raman spectroscopy provide valuable information about molecular vibrations and bonding. However, interpreting this data for complex, flexible molecules can be difficult. Computational methods help fill this gap by allowing the prediction and analysis of molecular properties based on quantum mechanical principles.

Among the various computational methods, density functional theory (DFT) has become one of the most popular due to its balance of computational cost and accuracy. DFT enables the calculation of molecular geometries, electronic distributions, and vibrational frequencies, making it especially useful for studying medium-sized organic molecules. When paired with suitable basis sets and dispersion corrections, DFT yields reliable insights into molecular behaviour that can support experimental observations.

Vibrational spectroscopy is essential for understanding molecular structure. IR spectroscopy relies on changes in dipole moments during molecular vibrations, while Raman spectroscopy depends on shifts in molecular polarizability. These two techniques offer complementary information and are often combined to create a complete picture of molecular vibrations. Raman spectroscopy is particularly sensitive to structural changes and is helpful for studying conformational effects and intermolecular interactions.

Sumatriptan is a pharmaceutical compound from the triptan class. It is a flexible organic molecule with multiple functional groups, such as an indole ring, sulphonamide group, and tertiary amine. These functional groups contribute to a complex vibrational spectrum. Moreover, the molecule can exist in various conformations due to its rotational flexibility, and its protonation state can further affect its electronic and vibrational properties.

A key challenge in studying these molecules computationally is their conformational diversity. At finite temperatures, molecules occupy multiple local minima on the potential energy surface instead of existing in a single fixed structure. Each conformer may contribute differently to the overall vibrational spectrum, making it necessary to consider a collection of structures rather than relying on just one optimized geometry. This adds complexity to both computational modelling and the interpretation of spectroscopic data.

Standard computational workflows usually involve conformational sampling using semiempirical methods, followed by geometry optimization and vibrational frequency calculations using DFT. In theory, the contributions from different conformers can be combined using statistical methods like Boltzmann weighting to estimate experimental spectra. However, applying these workflows in practice requires careful setup, computational resources, and validation at each step.

This report aims to develop a clear understanding of the computational methods used to analyse the vibrational properties of flexible molecules. The focus is on understanding the methodology, theoretical principles, and challenges involved in these studies rather than on presenting finished computational results. This approach seeks to build a solid foundation in computational chemistry and highlight the practical needs for successfully conducting these calculations in future work.

2. Methodology

The computational study of molecular structure and vibrational properties usually follows a multi-step process that includes conformational analysis, quantum chemical calculations, and spectroscopic interpretation. This study presents the methodology from both theoretical and conceptual viewpoints, based on established practices in computational chemistry.

The first step in this process involves preparing an initial molecular structure. This structure is often sourced from crystallographic databases or chemical structure repositories and serves as the foundation for further calculations. For molecules that can be protonated, additional structural forms might be considered to reflect different charge states. It is vital to prepare the molecular geometry accurately, as mistakes at this stage can greatly impact the results of subsequent calculations.

A key part of the computational analysis of flexible molecules is conformational sampling. Molecules like sumatriptan have several rotatable bonds, enabling them to take on multiple geometries with similar energies. Semiempirical methods, especially the GFN2-xTB approach, are commonly employed for this task due to their efficiency. Tools like CREST help explore the potential energy surface by generating a variety of conformers. This step is crucial because different conformers can show different vibrational properties.

After identifying a representative set of conformers, more sophisticated quantum chemical methods are usually applied for refinement. Density functional theory (DFT) is frequently used for this purpose. In DFT calculations, the electronic structure of the molecule is estimated based on electron density, allowing for an efficient yet accurate determination of molecular properties. Hybrid functionals like PBE0 are typically selected because they handle exchange interactions better, while basis sets such as def2-TZVP offer a detailed description of molecular orbitals. Dispersion corrections are added to account for weak intermolecular interactions that can affect conformational stability.

After geometry optimization, vibrational frequency calculations are conducted to find the normal modes of the molecule. These calculations are usually done within the harmonic approximation, where the potential energy surface is modelled as a quadratic function near the equilibrium geometry. The resulting frequencies match vibrational modes linked to specific atomic movements.

Interpreting vibrational spectra relies on the type of spectroscopy used. In infrared spectroscopy, vibrational modes are active if they involve a change in dipole moment. In contrast, Raman spectroscopy depends on changes in molecular polarizability. This difference means Raman spectroscopy is particularly sensitive to structural changes and provides additional insights compared to IR spectroscopy. Computational methods can predict both IR and Raman spectra, allowing for comparisons with experimental data.

For systems with multiple conformers, it is often necessary to consider their relative contributions to the overall spectrum. This is usually done using Boltzmann statistics, which weight each conformer according to its relative energy. Lower-energy conformers have a larger impact, while higher-energy ones contribute less. This statistical approach gives a more realistic representation of experimentally observed spectra, which reflect a mix of molecular configurations rather than a single structure.

While this workflow is standard in computational chemistry, successfully applying it requires careful choice of methods, appropriate parameter settings, and adequate computational resources. Each step must be validated to ensure the results are meaningful and consistent. In this work, the methodology was studied and conceptually understood, laying the groundwork for future practical application.

3. Experimental Work & Results

The goal of this section is to discuss what was learned during the study of the computational workflow and to reflect on the attempts to apply it in practice. Unlike a finished computational study, this work does not include fully validated numerical results. Instead, it focuses on analysing the workflow conceptually and identifying the challenges tied to its practical execution.

At the beginning of this project, we tried to set up the computational tools needed for conformational sampling and quantum chemical calculations. We installed and explored software packages like CREST and ORCA. We made efforts to understand their input structures, command-line execution, and output formats. This gave us basic familiarity with the workflow, including how to prepare molecular structures, initiate calculations, and obtain results.

However, moving from conceptual understanding to practical implementation turned out to be difficult. We faced challenges in configuring the computational environment, especially regarding software dependencies, parallel execution, and runtime limitations. As a result, we were unable to carry out complete and consistent calculations. Specifically, we could not generate reliable vibrational spectra or obtain a complete set of outputs for quantitative analysis.

Despite not having complete computational results, this study offers important insights into the nature of computational chemistry workflows. One key observation is that conformational sampling is not just an initial step; it is a crucial part of the entire process. Flexible molecules can take on many different geometries, and each geometry may affect their vibrational behaviour differently. This underscores the need to explore conformational space thoroughly before undertaking higher-level calculations.

Another significant aspect is the role of DFT in refining molecular structures and predicting vibrational properties. While the theoretical basis of DFT is well understood, its practical application needs careful selection of functionals, basis sets, and computational parameters. Even small changes in these choices can greatly influence the accuracy and reliability of the results. This highlights the necessity of a solid understanding of both theoretical principles and computational practices.

Interpreting vibrational spectra also poses challenges. Although the theoretical background for IR and Raman spectroscopy is clear, assigning specific vibrational modes to observed frequencies requires detailed analysis and often hinges on the quality of the computed data. Without reliable computational outputs, we cannot perform meaningful peak assignments or comparisons with experimental spectra.

The idea of combining contributions from multiple conformers using statistical approaches like Boltzmann weighting is theoretically sound. However, applying this in practice needs accurate energy calculations and consistent spectral data for each conformer. Without such data, this step remains more conceptual than computational in this work.

In summary, this study emphasizes that computational spectroscopy is not a simple or purely theoretical task. It requires careful coordination between theory, software implementation, and computational resources. The challenges faced during this project show how important consistent practice, time management, and systematic validation are when working with computational methods.

Although we did not obtain complete results, the effort to implement the workflow provided valuable learning experiences. It clarified the sequence of steps involved, the connections between different stages of the workflow, and the practical difficulties that must be overcome to achieve meaningful results. This understanding lays a necessary foundation for future work in computational chemistry.

4. Conclusion and Future Aspects

This report examines computational methods for analysing molecular structure and vibrational properties, focusing on flexible organic molecules like sumatriptan. It aims to clarify the theoretical principles and standard workflow, which includes conformational sampling, density functional theory (DFT), and vibrational spectroscopy.

Although the entire computational process could not be completed, the study clarified the sequence of steps involved in these analyses and how each stage contributes to determining molecular properties. It highlighted the significance of conformational diversity, electronic structure, and the interpretation of spectroscopic data. The report also explains the underlying theory of techniques like Raman spectroscopy and the statistical averaging of conformers.

The study stresses that computational chemistry involves more than just theoretical knowledge. It requires careful software setup, selecting appropriate parameters, and having enough computational resources. The challenges faced during this project, including issues with running calculations and getting reliable results, emphasize the importance of ongoing practice and time management when using computational tools.

From a learning viewpoint, this study introduced the practical aspects of computational chemistry workflows. It revealed how theoretical methods are applied in real research settings and identified the gaps between understanding concepts and putting them into practice. Recognizing these gaps is a crucial step toward building the skills needed for independent computational work.

Future efforts will concentrate on effectively implementing the computational workflow described in this report. This includes performing conformational sampling with tools like CREST, carrying out geometry optimization and vibrational frequency calculations using DFT techniques, and analysing the resulting spectra. With enough time and computational resources, it should be possible to generate reliable vibrational data and compare it with experimental findings.

Additionally, future improvements may involve including solvent effects, exploring different functionals and basis sets, and conducting more detailed conformational analysis. These actions would enhance the understanding of molecular behaviour.

Overall, this work lays the groundwork for understanding computational approaches to vibrational spectroscopy and points out the requirements for their successful use. It creates a foundation for future studies that combine theoretical knowledge and practical application to achieve meaningful and verifiable results.

5. References

1. Bursch, M., Mewes, J.-M., Hansen, A. and Grimme, S., 2022. Best-practice DFT protocols for basic molecular computational chemistry. *Angewandte Chemie International Edition*, 61(42), e202205735.
2. Koch, W. and Holthausen, M.C., 2015. *A Chemist's Guide to Density Functional Theory*. 2nd ed. Weinheim: Wiley-VCH.
3. Parr, R.G. and Yang, W., 1994. *Density-Functional Theory of Atoms and Molecules*. New York: Oxford University Press.
4. Neese, F., 2022. Software update: the ORCA program system-version 5.0. *WIREs Computational Molecular Science*, 12(5), e1606.
5. Grimme, S., Bannwarth, C. and Shushkov, P., 2017. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems. *Journal of Chemical Theory and Computation*, 13(5), pp.1989-2009.
6. Pracht, P., Bohle, F. and Grimme, S., 2020. Automated exploration of the low-energy chemical space with fast quantum chemical methods. *Physical Chemistry Chemical Physics*, 22, pp.7169-7192.
7. Barone, V., 2005. Vibrational zero-point energies and thermodynamic functions beyond the harmonic approximation. *Journal of Chemical Physics*, 122(1), 014108.
8. Jensen, F., 2017. *Introduction to Computational Chemistry*. 3rd ed. Chichester: Wiley.
9. Stuart, B., 2004. *Infrared Spectroscopy: Fundamentals and Applications*. Chichester: Wiley.
10. Scott, A.P. and Radom, L., 1996. Harmonic vibrational frequencies: an evaluation of Hartree-Fock, Møller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors. *Journal of Physical Chemistry*, 100(41), pp.16502-16513.