



JOINT INSTITUTE FOR NUCLEAR RESEARCH
Frank laboratory of Neutron Physics

FINAL REPORT ON THE INTEREST PROGRAMME

*Quantum Chemistry in Action: Molecules,
Spectra, and Crystals*

Supervisor:

Dr. Polina Pavlovna Kobchikova

Student:

Valerya Mantsireva, Moscow
D. Mendeleev University of
Chemical Technology of Russia

Participation period:

October 20 – November 30,
Wave 13, 2025

Table of contents

Abstract.....	3
1 Introduction	4
2 Literature Review	5
2.1 Computational chemistry methods.....	5
2.1.1 Semi-empirical method GFN2-xTB	6
2.1.2 Density functional theory (DFT)	7
3 Experimental.....	8
3.1 Initial data.....	8
3.2 Software	8
3.3 Conformational search (CREST/GFN2-xTB).....	8
3.4 DFT Calculations: Geometry Optimization.....	9
3.5 Calculation of vibrational frequencies and Raman spectrum	10
4 Results and discussion.....	11
4.1 Conformational analysis of sumatriptan	11
4.1.1 Computational Details	12
4.2 Raman spectrum analysis.....	13
4.2.1 Computational Details	15
5 Conclusions	16
6 References	17

Abstract

This study is devoted to a two-level computational study of the sumatriptan molecule, a widely used triptan drug for the treatment of migraine. In the first stage, a conformational search for the protonated form of the molecule in the gas phase was performed using the semiempirical GFN2-xTB method in the CREST program, revealing an ensemble of 24 unique conformers in an energy window of 8 kcal/mol. In the second stage, the geometry of the global minimum was optimized using the density functional theory (DFT) at the PBE0/def2-TZVP level, after which its vibrational frequencies and theoretical Raman spectrum were calculated. The calculated spectrum was compared with the experimental spectrum of crystalline sumatriptan. The analysis revealed expected discrepancies due to differences in conditions: an isolated molecule in the harmonic approximation (calculation) versus a crystalline phase with intermolecular interactions (experiment).

1 Introduction

Migraine headaches are one of the most serious disorders. Triptan is the most common drug prescribed for migraines and will have the greatest effect when taken on time. Common triptans are such as: Imitrex (sumatriptan), Amerge (naratriptan), Zomig (zolmitriptan), Maxalt (rizatriptan) and so on [1,2]. The most well-known type of triptan used is the anti-migraine drug of Sumatriptan (Imitrex-Treximet), a selective agonist at serotonin 5-HT₁-like receptors. It is an effective treatment for acute migraine attacks and the injectable form has also shown efficacy in the treatment of cluster headaches [3]. The molecule is an interesting subject for fundamental research: it contains an indole heterocyclic ring, a sulfonamide group, and an aminoalkyl «tail», which determines its conformational flexibility and specific interactions with biological targets. Understanding the molecular structure, electronic properties, conformational preferences, and spectral characteristics of sumatriptan is critical for pharmaceuticals. This understanding underlies the design of dosage forms (e.g., to enhance bioavailability), quality control (identification and analysis using spectroscopy), and understanding the mechanisms of action at the molecular level.

Experimental structural analysis methods, such as X-ray diffraction (XRD) and vibrational spectroscopy (IR, Raman), often provide limited data. X-ray crystallography provides an accurate "snapshot" of a single, most stable conformation of a molecule in the crystal lattice. However, in solution or when interacting with a receptor, a molecule can exist as an ensemble of conformers. Experimental spectra, in turn, contain integrated information on all vibrational modes, but their interpretation and accurate band assignment without computational support can be complex and ambiguous. Thus, a gap arises that is effectively filled by modern computational chemistry.

The aim of this study is a two-level computational study of the sumatriptan molecule, linking its conformational properties with spectral characteristics.

2 Literature Review

2.1 Computational chemistry methods

Computational chemistry is a discipline based on solving the Schrödinger equation to predict the structures, energies, and properties of chemical systems. A characteristic of all methods is the balance between computational cost and the accuracy of results, which forms a hierarchy of approaches.

There are three main classes of methods, differing in the degree of use of empirical parameters and the fundamental nature of the approach (Figure 1).

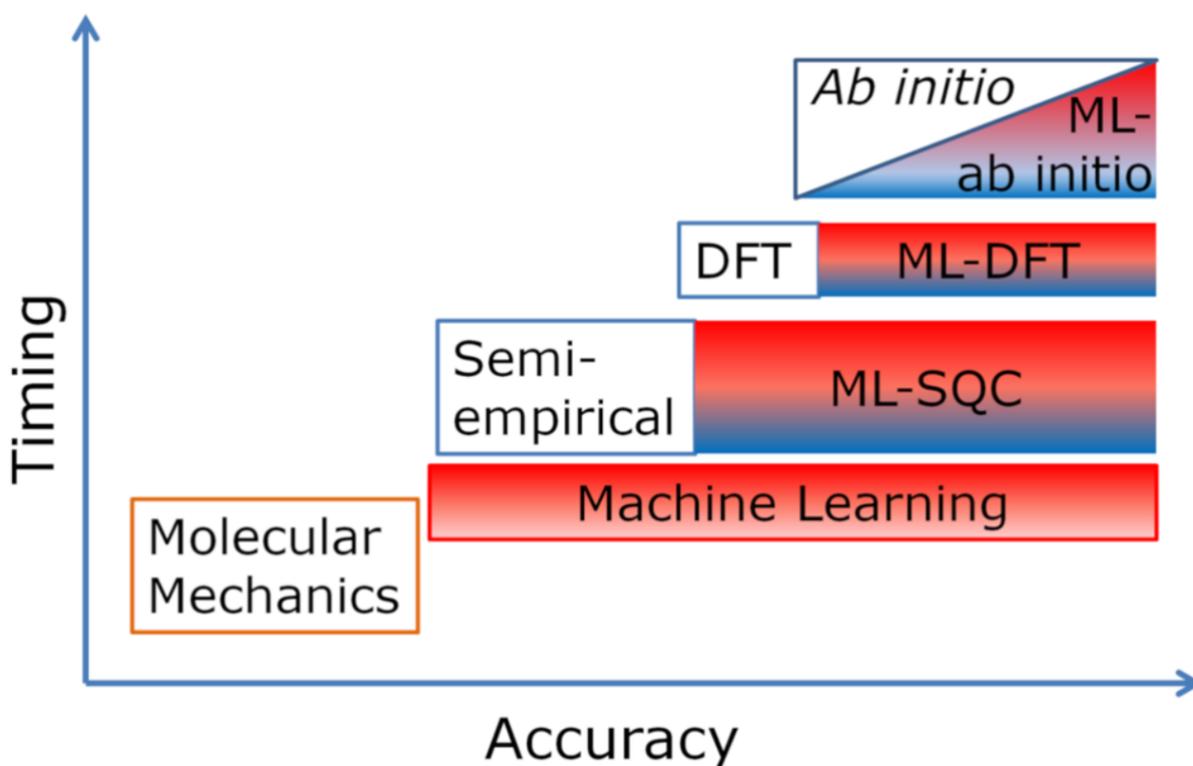


Figure 1. Schematic representation of various approximations with respect to accuracy and computational cost (timing) [4].

- 1 Molecular mechanics methods. Based on classical physics, they use parameterized potentials to describe interatomic interactions. They do not explicitly consider electrons and scale linearly with system size. Advantages: exceptional speed, the ability to simulate systems with tens and hundreds of thousands of atoms. Disadvantages: inability to describe the formation and breaking of chemical bonds, electronic excitations, and chemical reactions.

- Applications: molecular dynamics of biomacromolecules, polymer conformation studies, and preliminary analysis of large systems.
- 2 Semi-empirical quantum chemical methods. These are simplified versions of quantum chemical calculations in which many complex integrals are replaced by parameters selected based on experimental data or ab initio calculations. Advantages: significant calculation acceleration compared to ab initio methods. Disadvantages: accuracy depends on the quality of parameterization. Applications: screening conformational spaces, preliminary geometry optimization, calculations for medium-sized systems (up to 100-200 atoms).
 - 3 Ab initio methods. This is a group of approaches based on solving the equations of quantum mechanics without using empirical parameters. The advantage of first-principles calculations is the accurate description of atomic interactions, taking into account quantum effects. A disadvantage is the impossibility of calculating microscopic systems with a sufficiently large number of particles, such as atoms (rarely more than 100), in a reasonable amount of time.

2.1.1 Semi-empirical method GFN2-xTB

GFN2-xTB is a semiempirical method based on a self-consistent tight-binding scheme, expanded to include multipole electrostatic terms (up to quadrupole) and a self-consistent D4 dispersion correction dependent on atomic charges. The method is parameterized only at the elemental level, without pairwise parameters, and covers elements up to radon. It is positioned as a tool for calculating the structures, non-covalent interactions, and conformational energies of systems up to a thousand atoms, particularly for polar and hydrogen-bonded systems (Figure 2), where its predecessor, GFN-xTB, showed limitations. A key application of the method is the search for and ranking of rotamers and conformers for constructing thermostatic assemblies [5].

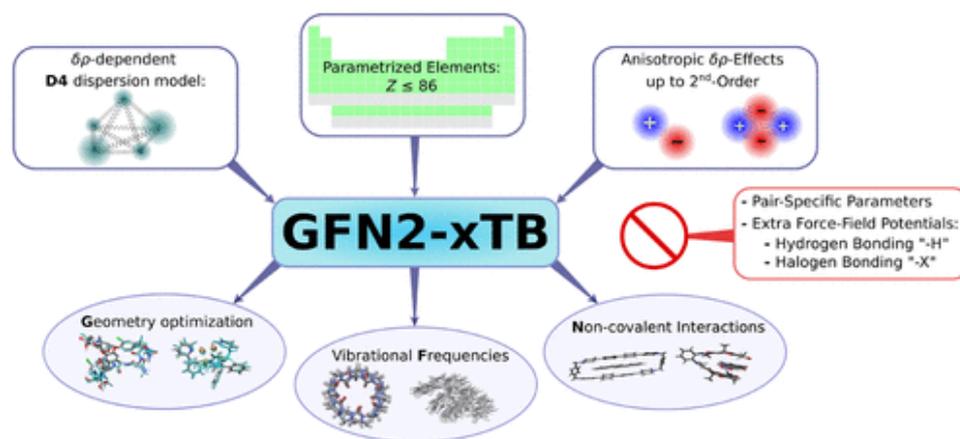


Figure 2. GFN2-xTB capabilities.

2.1.2 Density functional theory (DFT)

Density Functional Theory (DFT) is currently the dominant first-principles electron-structural method in computational chemistry [6]. Its key role stems from the tradeoff between computational efficiency and the accuracy of results for a wide range of chemical properties. Formally, DFT is an exact method, but in practice, approximate functionals of the exchange-correlation energy are used. Modern DFT implementations based on the Kohn-Sham approach allow one to solve the fermion many-electron problem for systems that are reasonably well described by a single-determinant wave function (systems with a single dominant electron configuration state). Such systems include the vast majority of diamagnetic molecules with closed shells in the ground state, making DFT a universal tool in organic and inorganic chemistry. Its application areas include determination and optimization of molecular and electronic structures, calculation of thermodynamic parameters, study of reactivity, prediction of spectral properties, investigation of non-covalent interactions (hydrogen bonds, stacking, van der Waals forces), etc. Application of the method requires a conscious choice of functional and basis set, as well as consideration of the physical conditions of the problem (solvent, temperature, conformational flexibility). Adherence to modern best-practice protocols, including multi-level approaches and composite methods, enables obtaining quantitatively accurate and chemically meaningful results for complex systems within an acceptable computation time.

3 Experimental

3.1 Initial data

Initial structure. The initial molecular geometry of protonated sumatriptan was presented in XYZ format.

Raman spectroscopy. Experimental Raman spectra of crystalline sumatriptan were used for comparison.

3.2 Software

CREST (version 3.0.2) and **XTB** (version 6.7.1) were used for conformational search and structure ensemble generation.

ORCA (version 6.0.0) was used for density functional theory (DFT) calculations, including geometry optimization and vibrational frequency calculations. CREST and ORCA output data were processed using **Bash** scripts.

Visualization of molecular structures, conformers, and geometric parameter analysis were performed using **Chemcraft**.

3.3 Conformational search (CREST/GFN2-xTB)

Calculations were performed in the gas phase for the sumatriptan cation (charge +1). The first step involved preliminary optimization of the initial structure:

```
xtb sum_plus_one.xyz --opt --chrg 1 --uhf 0 --gfn2
```

A conformational search was performed on the optimized structure (xtbopt.xyz) using CREST:

```
crest xtbopt.xyz --gfn2 --chrg 1 --ewin 8
```

Key parameters: GFN2-xTB method, +1 charge, energy window for structure conservation of 8 kcal/mol. This resulted in files with structure ensembles: *crest_rotamers.xyz*, *crest_conformers.xyz*, and the global minimum structure *crest_best.xyz*.

3.4 DFT Calculations: Geometry Optimization

ORCA input file (sum_plus_one.inp) for optimizing the original XYZ structure:

```
!PBE0 D4 DEF2-TZVP OPT  
*xyz 1 1  
S      0.57000    9.36060    3.27770  
O     -0.28000    8.28400    2.89120  
O      0.59070   10.50990    2.43820  
N      1.11110    2.70630    0.84050  
N      4.03520    7.28760   -1.69650  
N      0.18200    9.74900    4.78100  
C      3.14500    6.80570    0.29760  
C      2.70140    7.10230    1.57800  
H      2.34410    6.43160    2.11480  
C      2.78970    8.39760    2.05650  
C      3.34680    9.40530    1.24180  
H      3.41190   10.27140    1.57490  
C      3.79390    9.14580   -0.02710  
H      4.16550    9.81760   -0.55300  
C      3.67480    7.84480   -0.50290  
C      3.78900    5.95190   -1.66450  
H      3.97800    5.36210   -2.35760  
C      3.22990    5.59800   -0.47920  
C      2.89110    4.21840   -0.02390  
H      3.30460    4.06450    0.84070  
H      3.26930    3.58150   -0.65010  
C      1.40360    3.96570    0.09340  
H      0.98750    4.71270    0.55110  
H      1.01790    3.90730   -0.79400  
C     -0.35140    2.50860    1.01870  
H     -0.76520    2.38950    0.16070  
H     -0.72780    3.27810    1.45160  
H     -0.50580    1.73010    1.55990  
C      1.72840    1.50200    0.22870  
H      1.47570    0.72350    0.73050  
H      2.68410    1.59300    0.23730  
H      1.42280    1.41210   -0.67810  
C      2.23570    8.75030    3.41170  
H      2.79550    9.42660    3.82410  
H      2.24750    7.96390    3.97920  
C      0.48290   11.07580    5.28880  
H      1.38240   11.31280    5.05150
```

<i>H</i>	<i>0.39290</i>	<i>11.08030</i>	<i>6.24530</i>
<i>H</i>	<i>-0.12590</i>	<i>11.71170</i>	<i>4.90770</i>
<i>H</i>	<i>4.41510</i>	<i>7.65050</i>	<i>-2.31650</i>
<i>H</i>	<i>1.44100</i>	<i>2.85350</i>	<i>1.64770</i>
<i>H</i>	<i>-0.53380</i>	<i>9.44800</i>	<i>4.98800</i>

*

After successful calculation, an optimized structure was obtained.

3.5 Calculation of vibrational frequencies and Raman spectrum

The calculation was performed using optimized DFT geometry.

ORCA input file (raman.inp):

```
! NumFreq PBE0 def2-TZVP TightSCF D3BJ
```

```
%pal  
nprocs 16  
end
```

```
%elprop  
Polar 1  
end
```

```
%freq  
CentralDiff true  
Increment 0.02  
end
```

```
* xyzfile 1 1 sum_plus_one.xyz
```

Key parameters: calculation of the Hessian by numerical differentiation, calculation of polarizabilities to obtain Raman intensities.

4 Results and discussion

4.1 Conformational analysis of sumatriptan

A conformational search using CREST/GFN2-xTB revealed an ensemble of 24 unique conformers within an energy window of 8 kcal/mol. A Boltzmann population distribution at 298.15 K revealed that the lowest-energy structure (Figure 3) ($E = -60.98344$ a.u., population = 27.388%) is the most populated but not completely dominant, indicating significant conformational entropy of the system.

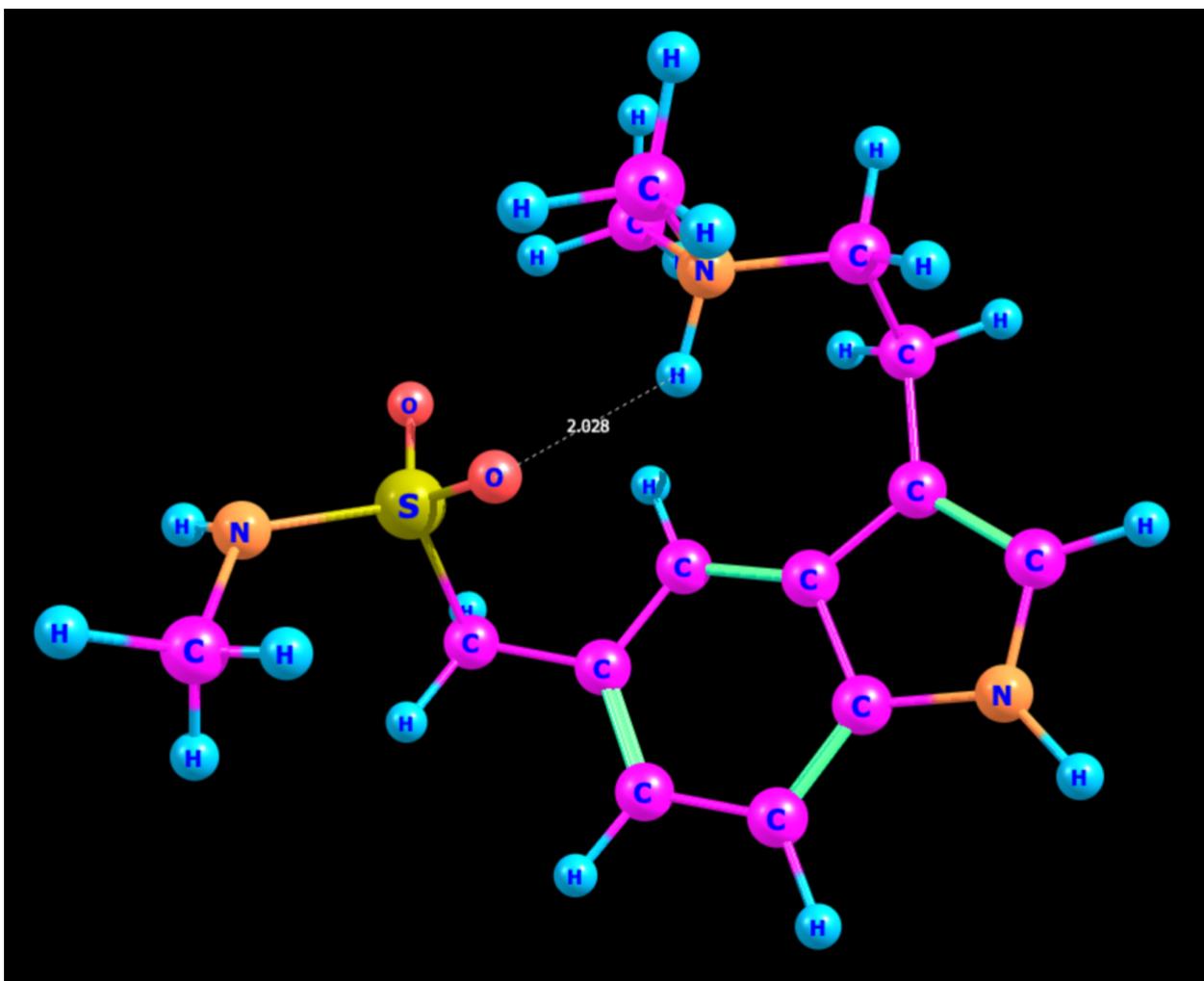


Figure 3. Global minimum.

To verify the true stability of the structures identified by the xTB method, vibrational frequencies were calculated at the DFT/PBE0 level for the lowest-energy conformer. Analysis of the Hessian revealed the presence of imaginary frequencies (negative eigenvalues) in the spectrum of this structure. This means that, within the

more accurate DFT method, this geometry is not a local minimum on the potential energy surface, but corresponds to a saddle point or is located in close proximity to one. Possible reasons for discrepancies:

- 1 Differences in accuracy between methods. The semi-empirical GFN2-xTB method, optimized for calculation speed, may produce a smoother or more biased potential energy surface (PES) than the highly accurate DFT method. The minimum found on the xTB PES may not coincide with the minimum on the DFT PES.
- 2 A putative role of hydrogen bonding. Visual analysis of the geometries obtained by CREST suggests that the relative stability of some conformers may be due to the possibility of forming an intramolecular hydrogen bond of the S=O \cdots H-N type between the sulfonamide and protonated amino groups. For example, in the structure indicated by CREST as the most stable, the O \cdots H distance is approximately 2.028 Å, consistent with a weak H-contact. However, it is impossible to assert that this interaction is the primary driving force of stabilization based solely on xTB screening data, as the energy differences between the conformers are comparable to the method's uncertainty.

4.1.1 Computational Details

As a preliminary step to efficiently search for low-energy conformational minima and optimize the geometry of protonated sumatriptan, we used the modern semiempirical extended tight-binding (xTB) method in the GFN2 (Geometries, Frequencies, Noncovalent Interactions — version 2) parameterization, implemented in the xtb program.

The calculation was performed with the following key parameters:

- `--gfn2`: Indicates the use of the GFN2-xTB parameterization. This version of the method is specifically parameterized to accurately reproduce molecular geometries, vibrational frequencies, and noncovalent interactions (van der

Waals, hydrogen bonds, π -stacking) for systems containing elements from H to Rn.

- --chrg 1: Sets the total charge of the system to +1, which corresponds to the protonated form of sumatriptan.
- --uhf 0: Specifies the number of unpaired electrons (spin multiplicity). A value of 0 indicates a closed shell (singlet state, all electrons are paired), which is typical for the valence configuration of this molecule in the ground electronic state.

4.2 Raman spectrum analysis

The theoretical Raman spectrum was calculated for the DFT-optimized structure and compared with the experimental spectrum of crystalline sumatriptan. To transform the discrete set of calculated harmonic modes into a continuous spectral profile comparable to the experiment, convolution with a Lorentzian function was used. The half-width (γ) was empirically selected and found to be 20 cm^{-1} . This value provides an optimal balance between resolving individual spectral features and achieving a smooth curve similar to the experimental record, where broadening is due to instrumental effects, anharmonicity of vibrations, and, in the case of a crystal, phonon interactions. To compensate for the systematic overestimation of harmonic frequencies inherent to the DFT/PBE0 method, a uniform scaling factor of 0.95 was applied to the calculated wavenumbers. The scaled theoretical spectrum was then normalized to its most intense band in the 500–1800 cm^{-1} region and overlaid with the similarly normalized experimental spectrum for visual comparison (Figure 4). This region was chosen for detailed analysis as it corresponds to the internal vibrational modes of the molecule, where a good visual correspondence between theory and experiment can be observed. In contrast, the low-frequency region ($< 400 \text{ cm}^{-1}$) is dominated by lattice (phonon) modes in the crystalline phase, which are not captured in the single-molecule calculation and is therefore omitted from detailed comparison.

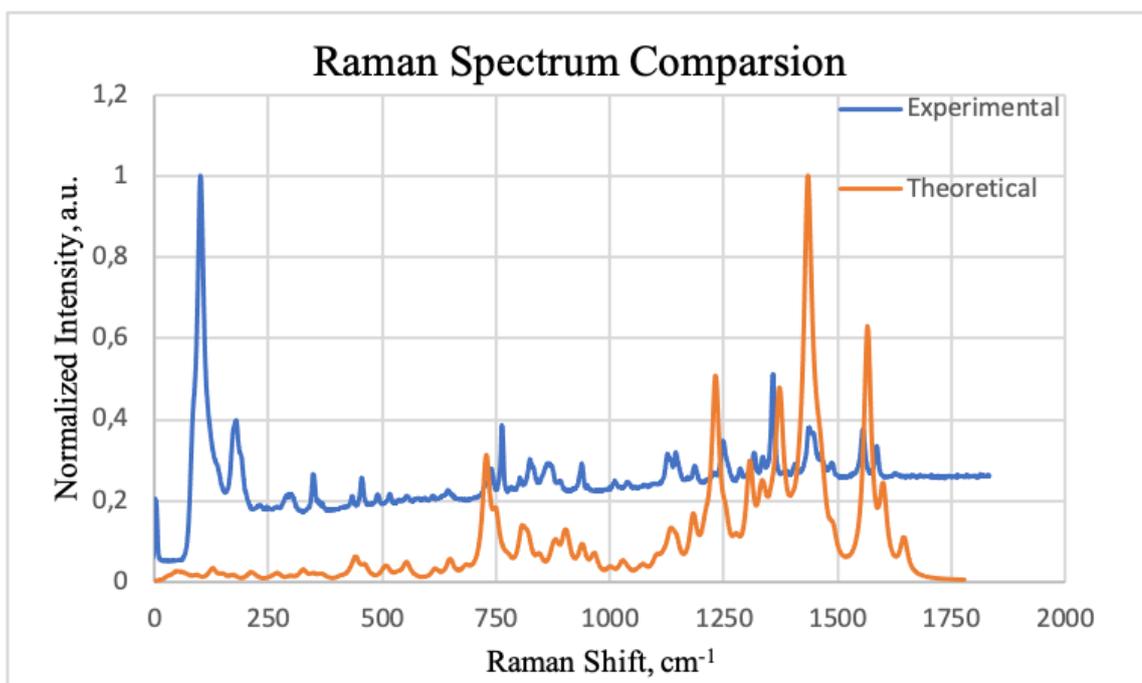


Figure 4. Raman Spectrum Comparison.

A comparison of the experimental and calculated Raman spectra of protonated sumatriptan reveals the expected discrepancies due to the different conditions. The theoretical spectrum is calculated for an isolated molecule in the gas phase using the harmonic approximation, while the experiment reflects the vibrational properties of the crystalline phase. In the crystal, molecules undergo intermolecular interactions and form a network of hydrogen bonds, which leads to systematic frequency shifts, predominantly downwards, as well as a redistribution of relative intensities. The harmonic approximation further underestimates the line widths and does not account for the anharmonic corrections characteristic of dense molecular packing. Low-frequency region shows the most pronounced qualitative differences. The gas-phase calculation yields only a few weak bands corresponding to skeletal deformations and torsions. The experimental spectrum, however, is rich in intense, broad features in this region. These are attributed to low-energy lattice (phonon) modes – collective vibrations of molecules within the crystal lattice, librations, and translations – which are entirely absent in the model of an isolated molecule. The stark contrast in this region serves as direct spectroscopic evidence of the significant influence of crystal packing on the dynamic properties of the solid. But despite this the theoretical

spectrum reproduces the key intramolecular vibrational features of sumatriptan observed in the experiment

4.2.1 Computational Details

Geometry Optimization (!PBE0 D4 DEF2-TZVP OPT):

- PBE0: Hybrid density functional.
- DEF2-TZVP: Qualitative triple zeta-valization basis set with polarization functions.

Calculating Frequencies and Polarizabilities (! NumFreq PBE0 def2-TZVP TightSCF D3BJ):

- NumFreq: Specifies that vibrational frequencies and properties are calculated numerically (via finite differences). This also allows for calculating polarizabilities and intensities for the Raman spectrum.
- TightSCF: Tightens the convergence criteria for the self-consistent field (SCF) procedure. This improves the accuracy of energy calculations and, consequently, derivatives (forces, frequencies), reducing numerical noise.
- D4 (or D3BJ in frequency-domain calculations): Modification to account for dispersion interactions.
- %freq: Block with settings for numerical frequency calculations.
- CentralDiff true: Use the central difference scheme to calculate second derivatives (Hessian).
- Increment 0.02: Nuclear displacement step in angstroms during numerical differentiation. A default value that provides a balance between accuracy and robustness.
- %elprop: Block for calculating electronic properties.
- Polar 1: Enables calculation of the molecular polarizability tensor (and its derivatives with respect to normal coordinates).

5 Conclusions

As part of the INTEREST program, a comprehensive computational study of the sumatriptan molecule was conducted using quantum chemistry methods.

- 1 Using the modern semiempirical GFN2-xTB method and the CREST program, an efficient conformational search was performed for the protonated form of sumatriptan. Twenty-four unique conformations were found within an energy window of 8 kcal/mol, highlighting the significant flexibility of the molecule. Boltzmann analysis revealed that the global minimum on the GFN2-xTB potential energy surface is not absolutely dominant (population ~27% at 298 K), indicating a significant contribution from conformational entropy.
- 2 Subsequent optimization of the geometry of the most stable conformer at a more accurate level of DFT theory (PBE0/def2-TZVP) revealed limitations of the semi-empirical approach: the structure determined at least by the GFN2-xTB method revealed the presence of imaginary frequencies when calculating frequencies at the DFT level. This indicates that, within the more accurate method, this geometry corresponds to a saddle point rather than a local minimum. This discrepancy illustrates the importance of verifying the results of rapid screening methods using more rigorous computational protocols.
- 3 For the DFT-optimized geometry, a theoretical Raman spectrum was calculated. After scaling the frequencies (factor 0.95) and broadening with a Lorentzian function, the spectrum was compared with the experimental Raman spectrum of crystalline sumatriptan. Despite the overall agreement in the positions of the main intramolecular bands, noticeable discrepancies in intensities and some peak shifts were observed. These differences arise naturally from the different physical conditions: the calculation was performed for an isolated molecule in the harmonic approximation, while the experiment reflects the influence of the crystal lattice, intermolecular hydrogen bonds, and anharmonic effects.

6 References

- 1 B. Bonthagarala, S. Nama, S. Kothamasu, M. Badamuddala, Sudarshan Donthiboina, Formulation and evaluation of sumatriptan succinate floating bilayered tablets, *Discovery* 17 (49) (2014) 23-35.
- 2 K.S. Lewis, Pediatric headache, *Semin. Pediatr. Neurol.* 17 (4) (2010) 224-229.
- 3 Perry, C.M., Markham, A. Sumatriptan. *Drugs* 55, 889–922 (1998).
- 4 Dral, P. O. Quantum Chemistry in the Age of Machine Learning. *The Journal of Physical Chemistry Letters* 2020, 11 (6), 2336–2347.
- 5 Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *Journal of Chemical Theory and Computation* 2019, 15 (3), 1652–1671.
- 6 Bursch, M.; Mewes, J.; Hansen, A.; Grimme, S. Best-Practice DFT Protocols for Basic Molecular Computational Chemistry**. *Angewandte Chemie International Edition* 2022, 61 (42).