Testing Theory Against Reality: A DFT-Based Analysis of Molecular Structure, Solvation, and Reactivity

Handson Gisubizo
EN.540.658 – Modeling and Design of Sustainable Chemical Processes

April 17, 2025

Abstract

In 1929, Paul A. M. Dirac remarked that the underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are therefore completely known, but that the difficulty is only that the exact application of these laws leads to equations that are much too complicated to be soluble [5]. This insight remains a cornerstone of computational chemistry, motivating the development of approximate yet powerful quantum mechanical methods. Among these, Density Functional Theory (DFT) stands out for its ability to reformulate the many-electron Schrödinger equation into a more tractable form, solvable using approximate exchange-correlation functionals and appropriate basis sets.

DFT is now one of the most widely used approaches in computational chemistry, enabling researchers to predict molecular properties with remarkable precision. This project critically assesses how accurately DFT can predict the structural, energetic, thermodynamic, and kinetic behavior of molecular systems relative to experimental expectations. Simulations were conducted across a range of chemically relevant scenarios, including conformational energy profiles, hydrogen bonding, water-cluster behavior, ionic solvation, and reaction transition states. All calculations were performed with ORCA [12], with molecular structures built in Avogadro [9], visualized in Chemcraft [14], and post-processed using Google Colab [7].

The results show that, despite its approximations, DFT reliably captures key molecular trends and behaviors, transforming the complexity Dirac described into practical predictive tools for chemistry.

Introduction and Motivation

Computational chemistry provides a powerful complement to experimental work by enabling predictive insights into molecular systems before entering the laboratory. Among its tools, Density Functional Theory (DFT) has become a gold standard for evaluating electronic structures and reaction energetics. It strikes a favorable balance between computational cost and accuracy, making it widely adopted across chemistry, materials science, and molecular biology [2, 4].

However, despite its widespread use, DFT is not without limitations. Its accuracy can depend heavily on the choice of exchange-correlation functional, basis set, and treatment of dispersion or solvation effects. This raises a persistent and practical question: *Can DFT reliably predict what we would observe on the lab bench?*

To answer this, we systematically apply DFT to a diverse range of systems and compare the outcomes to known experimental results and established physical intuition. The selected systems represent common and chemically relevant scenarios: butane conformers capture conformational flexibility in hydrocarbons; water clusters model hydrogen bonding and vibrational spectra; sodium chloride hydration illustrates solvation and ion–water interactions; and a model reaction pathway is used to evaluate transition state location and kinetics. Each system was chosen for its simplicity, pedagogical clarity, and relevance to broader chemical principles.

By probing this range of problems, we aim to explore the strengths and limitations of DFT from a practical perspective—assessing not just how accurate it can be, but when and why it works or fails. Through this lens, the project offers a grounded evaluation of DFT's predictive power in real chemical applications.

Computational Methods

All calculations were carried out using ORCA 6.0.0 [12], a quantum chemistry software package. The ωB97X-D functional [3] and the def2-TZVP basis set [13] were used for most calculations. Tight SCF convergence settings were applied to ensure accuracy.

Jobs were run locally on a Windows 10 desktop using 8 parallel threads, specified by the `%pal nprocs 8 end` command in ORCA. Any changes in methods—such as different basis sets, solvation models, or frequency calculations—are noted in the relevant sections.

Conformational Analysis of Butane The butane molecule was built in Avogadro [9], with the central C–C–C–C dihedral angle explicitly defined. This angle was varied in 20 steps from $-180°$ to $+180°$, generating a series of conformers saved as individual XYZ files.

Single-point energy calculations were performed without geometry optimization to isolate energy changes due to dihedral rotation. Energies were extracted using a Python script in Google Colab [7], and used to construct the potential energy surface (PES). Chemcraft [14] was used to visualize and confirm the structural consistency of all conformers.

Geometry Optimization and IR Spectrum of Water Cluster A 15-molecule water cluster, $(H_2O)_{15}$, was constructed in Avogadro and optimized using ORCA. Convergence was monitored by tracking the total SCF energy during the optimization steps. Upon reaching a local minimum, a vibrational frequency analysis was performed to confirm the absence of imaginary modes and to generate the infrared (IR) spectrum.

Structural and vibrational data were processed using a Python notebook and visualized with Chemcraft [14]. Bond lengths, angles, and IR peaks were analyzed and later compared to experimental benchmarks, as presented in the Results section.

Thermodynamic Properties at Varying Temperatures Thermodynamic properties of the optimized $(H_2O)_{15}$ cluster were evaluated using the vibrational frequency

output from ORCA. Internal energy (U), enthalpy (H), entropy (S), and Gibbs free energy (G) were computed at five temperatures ranging from 300 K to 1000 K under a pressure of 1 atm. The quasi-RRHO (quasi-rigid rotor harmonic oscillator) correction was applied using a 1 cm^{-1} frequency cutoff to account for low-frequency vibrational modes. Thermodynamic data were extracted using ORCA's built-in thermochemistry module and analyzed in a Python notebook using Google Colab.

Hydrogen Bond Energy Profile To investigate intermolecular hydrogen bonding, a water dimer was first optimized at its equilibrium geometry using DFT. One molecule served as the hydrogen bond donor and the other as the acceptor. From this optimized structure, the O–H distance between the donor hydrogen and the acceptor oxygen was systematically varied from 0.40 Å to 5.00 Å in 0.01 Å increments, while keeping all other atomic positions fixed.

For each modified geometry, a single-point energy calculation was performed without relaxation. The interaction energy profile was generated by referencing all energies to the global minimum configuration, resulting in a potential energy curve that reveals the strength and range of the hydrogen bond. All calculations were carried out using ORCA, and the data were analyzed and visualized using a Python notebook in Google Colab.

Bond Force Constant of O–H Stretch As a continuation of the hydrogen bond energy profile analysis, the potential energy curve near the equilibrium region was examined in more detail to quantify the stiffness of the O–H bond. A single water molecule was first optimized using DFT, and the O–H bond was then symmetrically stretched and compressed from its equilibrium length in 0.01 Å increments over a total range of approximately 0.14 Å. All other atomic positions were held fixed, and single-point energy calculations were performed in ORCA for each geometry.

In the vicinity of the energy minimum, the interaction follows a nearly harmonic behavior. The computed energy–distance data were fitted to a classical harmonic potential, $E = \frac{1}{2}k(x - x_0)^2$, where $x_0 = 0.966$ Å is the fitted equilibrium bond length and k is the bond force constant. This fitting yielded a stiffness value of 740.8 N/m for the O–H bond. The resulting force constant was then compared to experimental spectroscopic values and empirical models, including TIP3P (450 N/m), SPC/E (529 N/m), and TIP4P (553 N/m). Data fitting, analysis, and visualization were performed using Python in Google Colab.

Solvation of NaCl with Water Molecules A hybrid solvation approach was used to study the progressive hydration of a neutral NaCl ion pair. The system was solvated with 1 to 41 explicitly placed water molecules, and each configuration was geometry-optimized. To account for long-range bulk solvent effects, the ALPB (Analytical Linearized Poisson–Boltzmann) [6] implicit solvation model for water was applied alongside the explicit molecules. Calculations were performed using the XTB semiempirical method [1, 8], which balances computational efficiency and accuracy for large molecular systems. The number of explicit solvent molecules was controlled using the `%solvator` block in ORCA. Interaction energies were analyzed across solvation levels to identify stabilization trends and hydration shell completion. Post-processing and visualization were carried out using Python in Google Colab.

NEB-TS Transition State Search The reaction pathway for the acid-catalyzed substitution of methanol by hydrogen chloride:

$$CH_3OH + HCl \rightarrow CH_3Cl + H_2O$$

was investigated using the Nudged Elastic Band (NEB) method [11] to determine the minimum energy path (MEP). First, the geometries of the isolated reactant and product species were individually optimized using DFT in ORCA [12]. Ten intermediate images were then generated via linear interpolation between the optimized endpoints and optimized using the NEB algorithm. Climbing-image refinement [10] was employed to accurately locate the transition state (TS).

Following convergence, a full vibrational frequency analysis was performed on the highest-energy image to confirm the presence of a single imaginary mode, validating it as a first-order saddle point. Thermodynamic corrections were applied to compute the Gibbs free energy of activation ($\Delta G^{\ddagger}$) at 298 K. The rate constant was estimated using Eyring's transition state theory:

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^{\ddagger}}{RT}\right)$$

All calculations were conducted using ORCA, and data analysis and reaction coordinate plots were carried out using Python in Google Colab [7].

Results and Discussion

Butane Conformational Energy Profile

The calculated potential energy surface (PES) for butane as a function of the central C–C–C–C dihedral angle reveals three key conformational states (Figure 1). The global minima occur at $\pm 180°$, corresponding to the anti conformation, with an SCF energy of -158.594742 a.u. Local minima appear near $\pm 60°$, representing gauche conformers, with energies around -158.592336 a.u. The highest energy point is at $0°$, corresponding to the fully eclipsed conformation, with an energy of -158.579643 a.u.

The energy difference between anti and gauche conformers is 0.002406 a.u., equivalent to 1.51 kcal/mol, while the difference between gauche and eclipsed conformers is 0.012693 a.u., corresponding to 7.97 kcal/mol. These values are consistent with known torsional barriers in alkanes and validate the accuracy of the computed conformational energy profile.

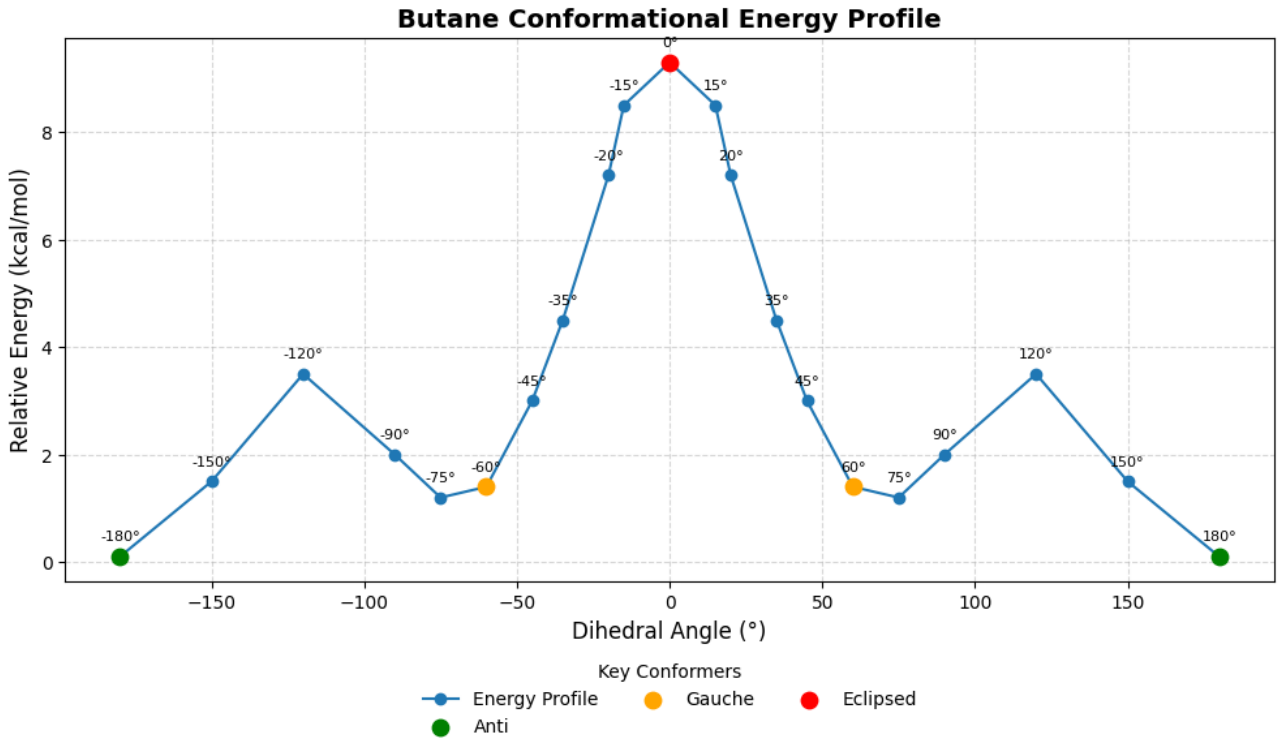

Figure 1: Conformational energy profile of butane based on dihedral angle rotation. Anti, gauche, and eclipsed conformations are indicated. Energies are relative to the global minimum.

IR Spectrum and Structural Features of $(H_2O)_{15}$ Cluster

The simulated IR spectrum of the $(H_2O)_{15}$ cluster shows three main vibrational regions (Figure 2): a broad low-frequency librational band (¡1000 cm^{-1}), a pronounced H–O–H bending mode around 1600 cm^{-1}, and a sharp O–H stretching region spanning 3000–3700 cm^{-1}. These spectral features align well with known experimental IR characteristics of water clusters and bulk water, indicating reliable vibrational modeling.

To support the structural accuracy of the optimized geometry, O–H bond lengths and H–O–H bond angles were analyzed (Figure 3). The average O–H bond length was 0.98 Å, and the average H–O–H angle was 105.3°, both consistent with typical gas-phase and condensed-phase water values. These results confirm that the geometry is physically reasonable and that the IR-active modes are associated with realistic molecular structures.

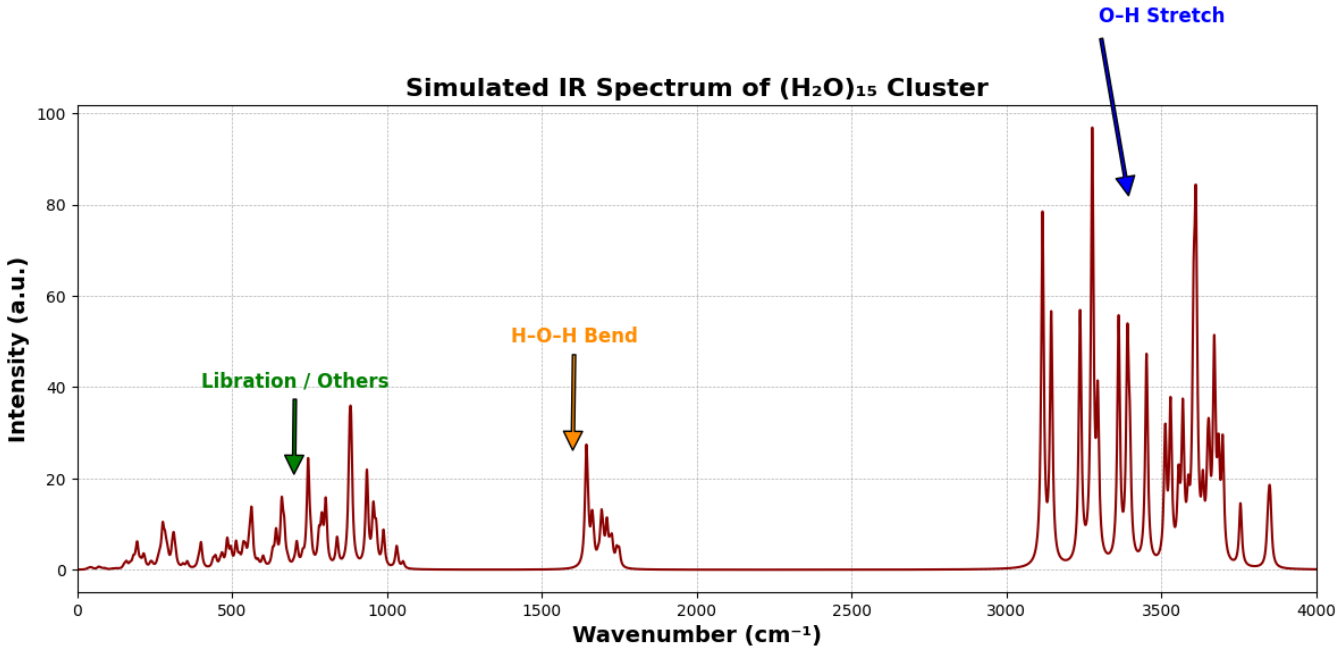

Figure 2: Simulated IR spectrum of the $(H_2O)_{15}$ cluster. Key regions include libration (¡1000 cm^{-1}), H–O–H bending (1600 cm^{-1}), and O–H stretching (3000–3700 cm^{-1}).

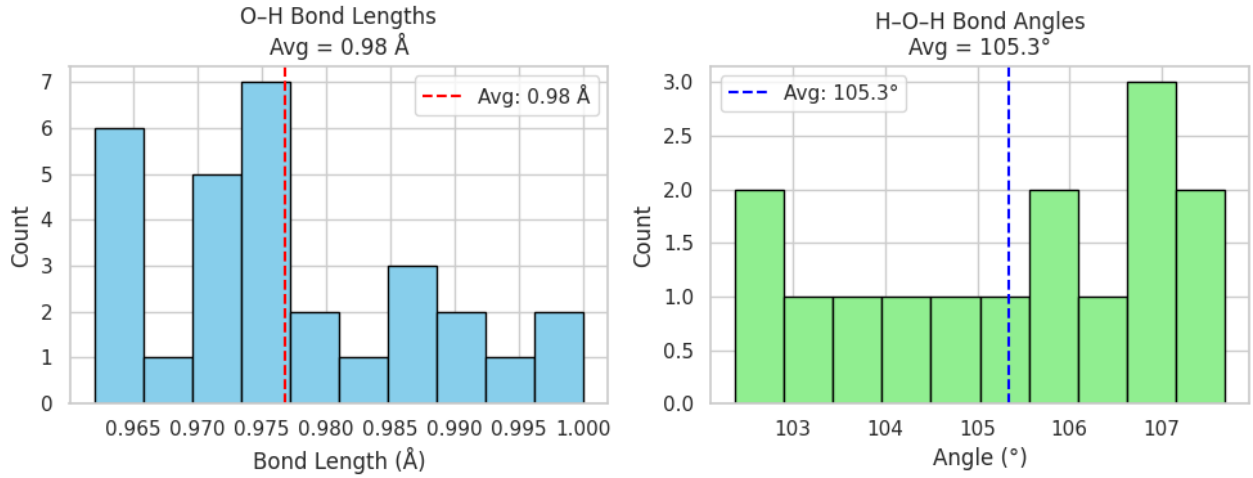

Figure 3: Distribution of O–H bond lengths and H–O–H bond angles in the optimized $(H_2O)_{15}$ cluster. The average O–H bond length is 0.98 Å, and the average H–O–H angle is 105.3°, both in good agreement with experimental water geometries.

Thermodynamic Behavior of the $(H_2O)_{15}$ Cluster

Thermodynamic properties of the $(H_2O)_{15}$ cluster were evaluated across a temperature range of 300–1000K at a pressure of 1atm. As shown in Figures 4 and 5, both the internal energy (U) and enthalpy (H) increase nearly linearly with temperature, reflecting the expected thermal population of vibrational modes in a condensed-phase cluster. The entropy contribution, represented by the $T \cdot S$ term, also increases with temperature, contributing significantly to the free energy variation.

Notably, the Gibbs free energy (G) decreases monotonically over the temperature range, primarily driven by the increasing entropic term. This behavior is characteristic of flexible hydrogen-bonded systems, where vibrational freedom contributes substantially to entropy at elevated temperatures. The smooth and physically consistent trends across all thermodynamic quantities support the quality of the optimized structure and the validity of the quasi-RRHO correction used.

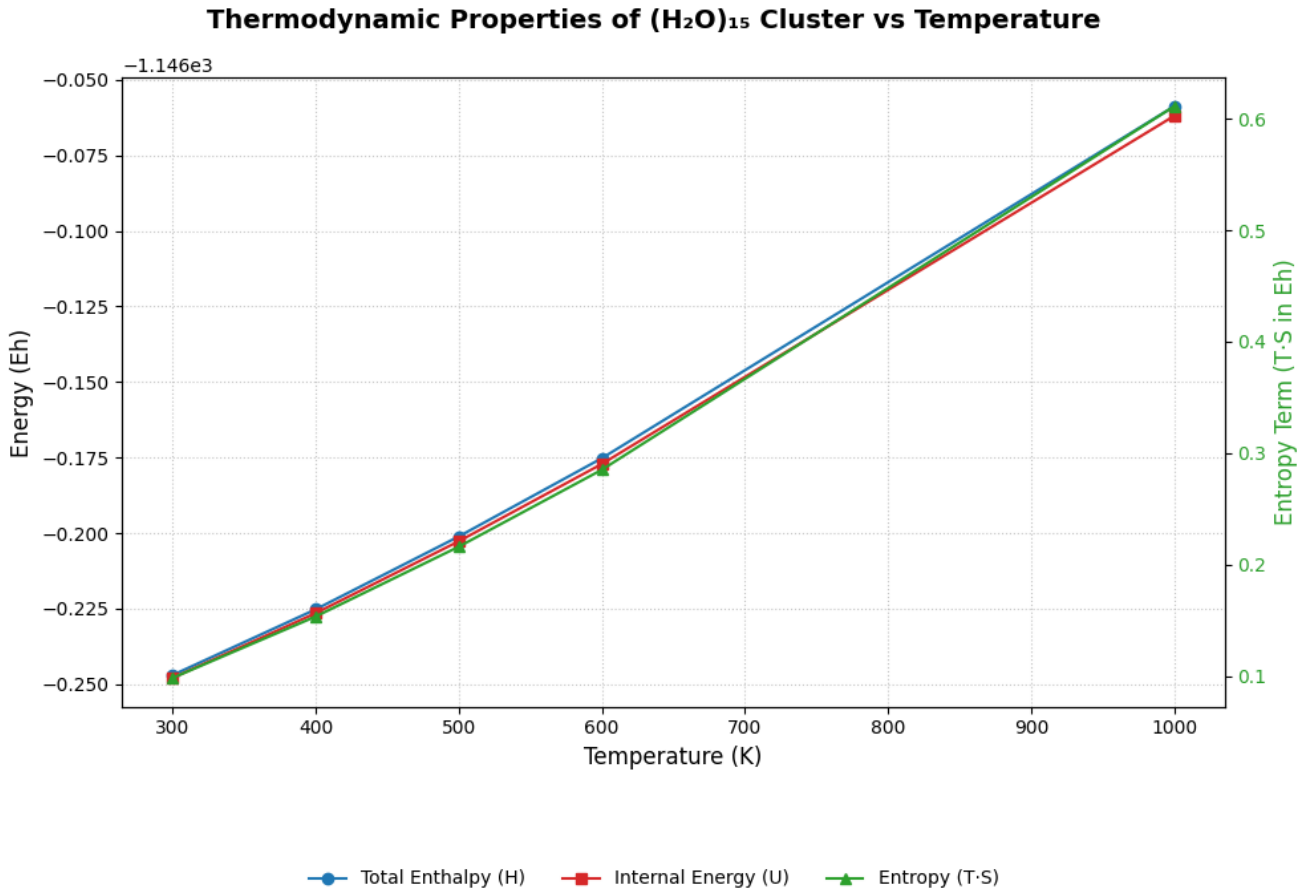

Figure 4: Temperature dependence of thermodynamic properties for the $(H_2O)_{15}$ cluster: internal energy (U), enthalpy (H), and entropy term (T·S). Data were obtained from ORCA's thermochemistry module and post-processed in Google Colab.

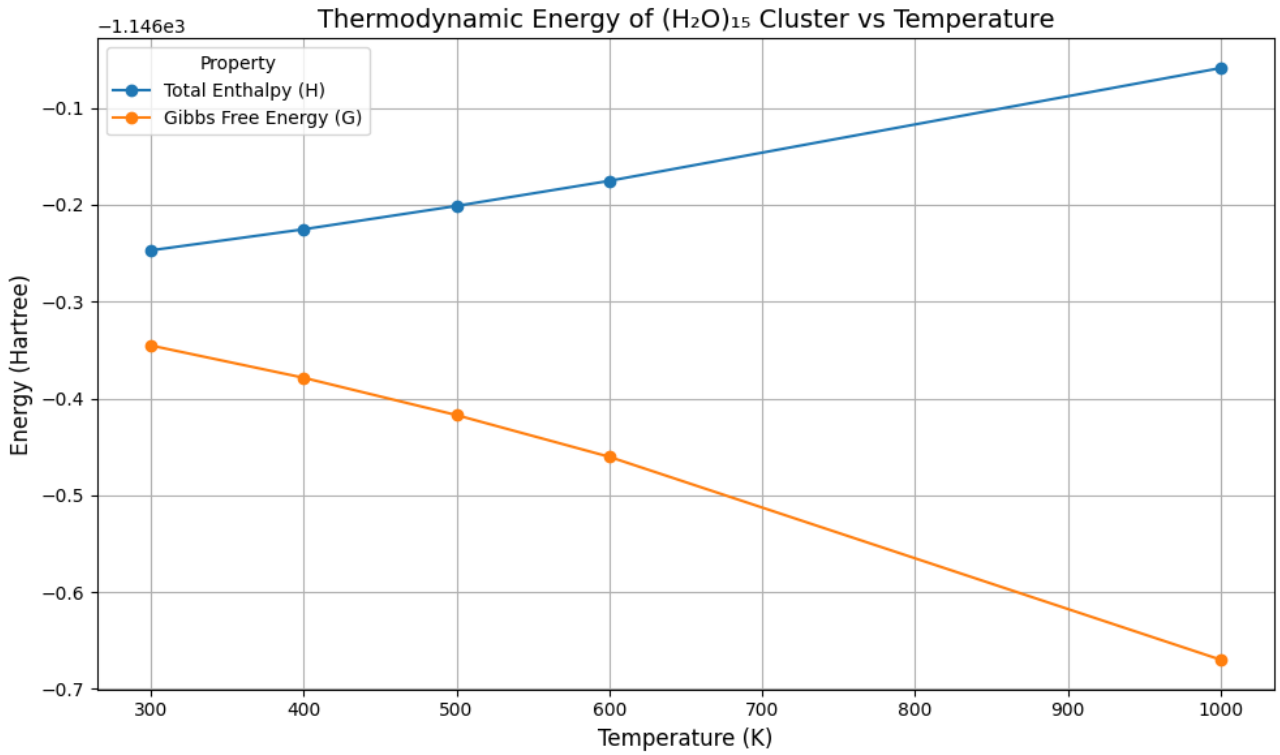

Figure 5: Variation of total enthalpy (H) and Gibbs free energy (G) with temperature for the $(H_2O)_{15}$ cluster. The decreasing G with increasing temperature reflects growing entropic contributions.

Hydrogen Bonding Interaction Profile

The interaction energy between two water molecules was evaluated as a function of the O–H distance, producing a characteristic potential energy curve (Figure 6). The energy minimum occurs at an O–H separation of 0.96 Å, with a corresponding SCF energy of −152.86 a.u. This position represents the equilibrium hydrogen bond length for the dimer under frozen-geometry conditions.

The curve shows a steep rise in energy as the donor hydrogen approaches the acceptor oxygen below 0.9 Å, indicating strong repulsive interactions due to Pauli exclusion and electrostatic repulsion. At larger distances beyond 1.5 Å, the interaction energy levels off, signaling the decay of attractive forces and the approach to non-bonded interaction. The smooth potential energy well reflects the balance between electrostatic attraction and exchange repulsion that governs hydrogen bonding. These results are in line with typical hydrogen bond behavior in gas-phase water dimers.

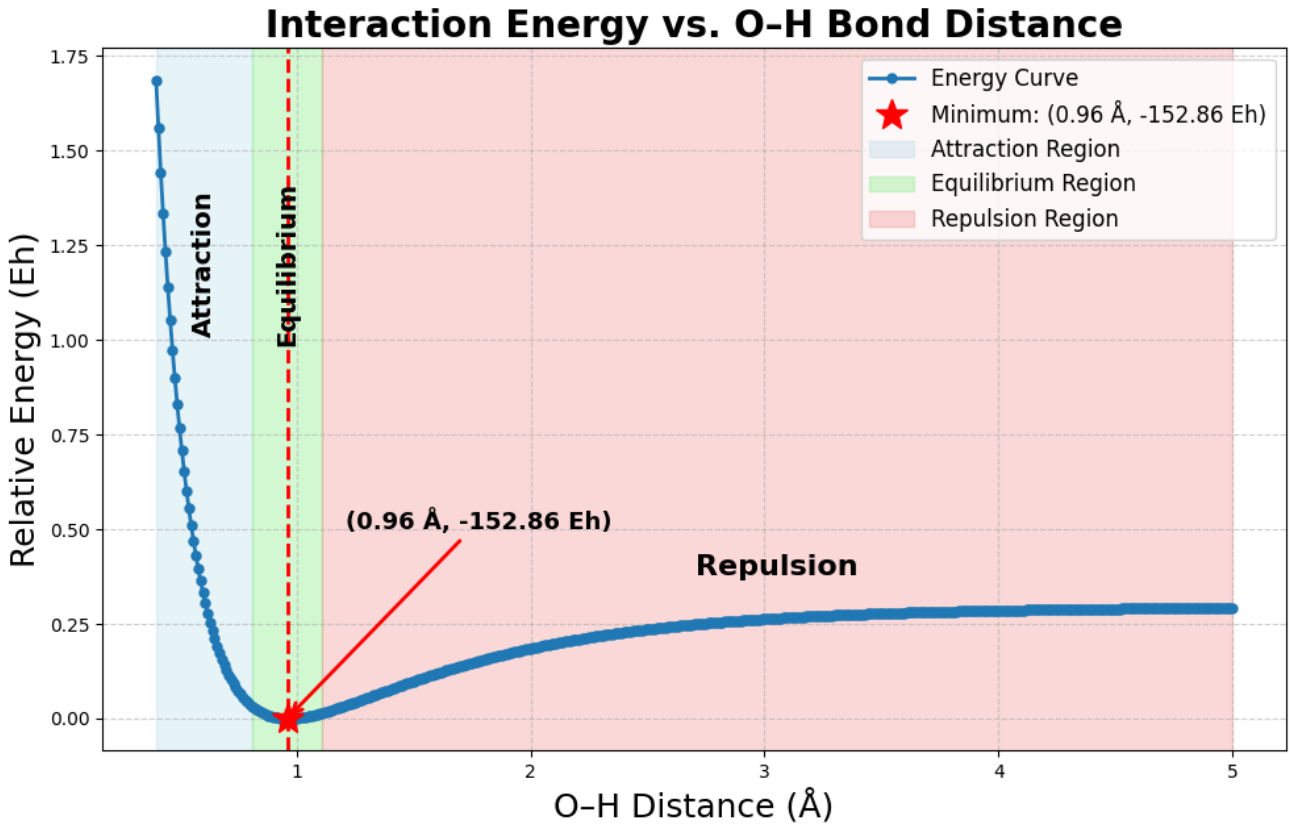

Figure 6: Interaction energy profile of a water dimer as a function of the O–H distance. The minimum energy is observed at 0.96 Å with a relative energy of −152.86 a.u. Regions of attraction, equilibrium, and repulsion are annotated. Energies are referenced to the global minimum.

O–H Bond Force Constant Analysis

The harmonic fitting of the O–H bond energy curve (Figure 7) yielded a force constant of 740.8 N/m at an equilibrium bond length of 0.966 Å. This value was obtained by stretching and compressing the O–H bond around equilibrium and fitting the resulting single-point DFT energies to a harmonic potential. The fitted curve accurately reproduced the curvature of the potential energy well, confirming the validity of the harmonic approximation in the near-equilibrium region.

This DFT-derived force constant was compared to experimental and empirical values (Figure 8). The experimental spectroscopic value is approximately 460 N/m, while force fields such as TIP3P, SPC, SPC/E, and TIP4P assign values in the range of 450–553 N/m. The noticeably higher stiffness obtained from DFT is expected: quantum chemical calculations, such as those using the B3LYP functional in gas phase, tend to yield sharper potential wells because they do not account for thermal or anharmonic effects present in real systems. Meanwhile, empirical models are typically parameterized to reproduce macroscopic liquid properties like density or diffusion, not isolated molecular vibrations.

Therefore, the higher DFT-derived stiffness reflects a more idealized, intrinsic bond potential, and the result is physically meaningful within the context of a gas-phase, harmonic quantum mechanical model.

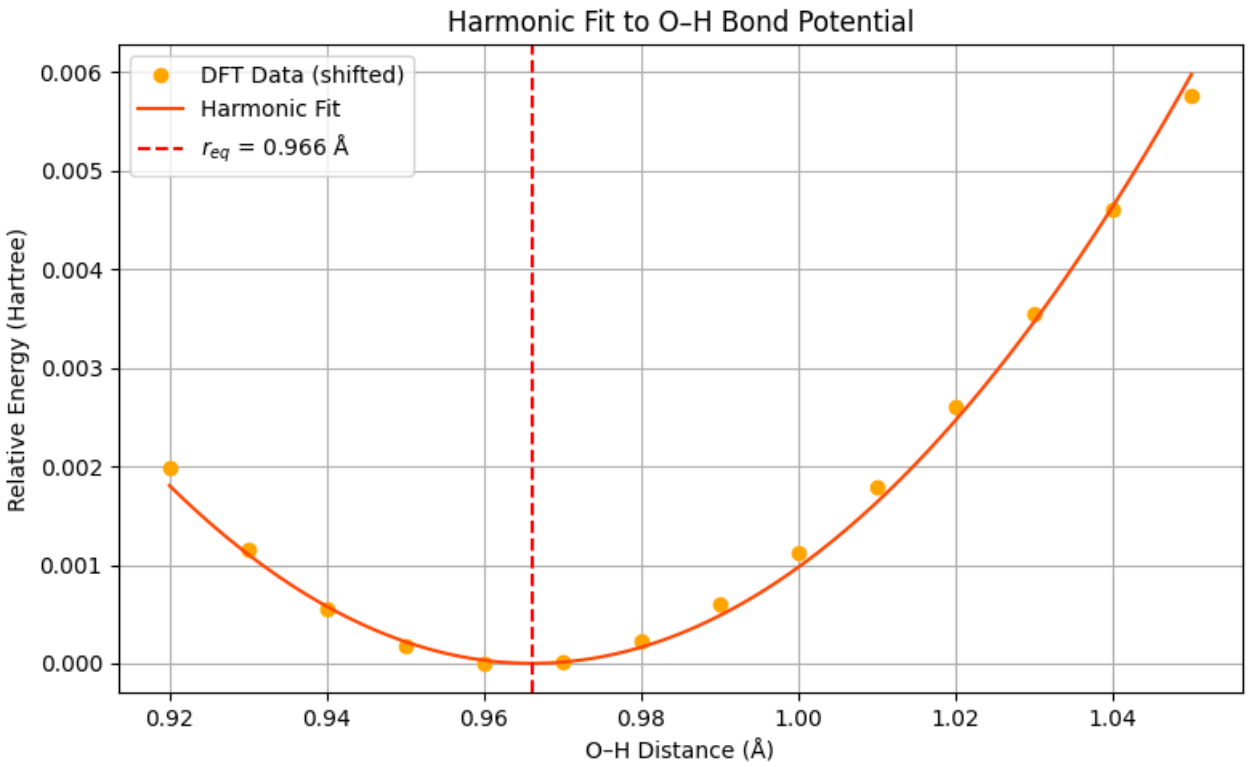

Figure 7: Harmonic fit to the O–H bond potential energy curve. The fitted minimum occurs at 0.966 Å and the extracted force constant is 740.8 N/m.

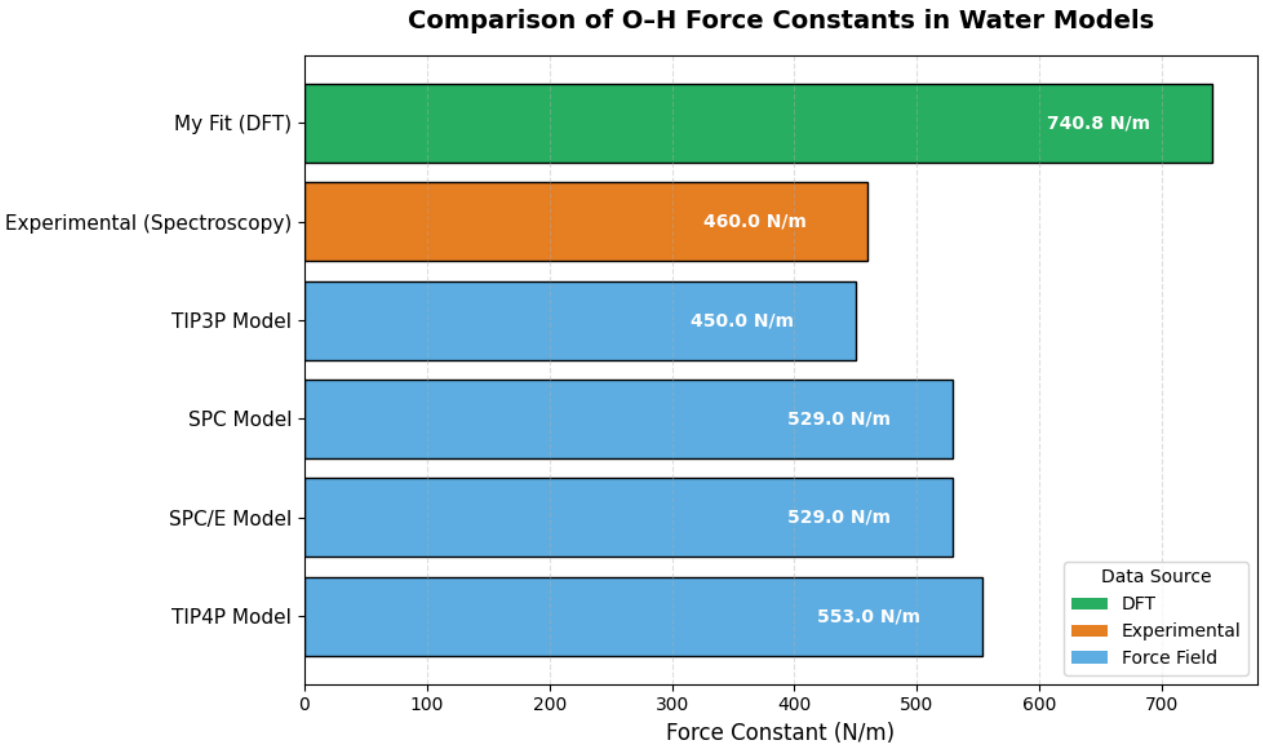

Figure 8: Comparison of O–H bond force constants from DFT, experimental spectroscopy, and commonly used water models. DFT yields the highest stiffness due to quantum-level accuracy and lack of empirical parameter tuning.

Stabilization Behavior During NaCl Solvation The incremental interaction energy (ΔE) was monitored as water molecules were sequentially added to the NaCl ion pair (Figure 9). The plot reveals an overall trend toward stabilization, with the most significant energy drop occurring after the addition of the second water molecule, suggesting strong initial ion–solvent interactions. As additional water molecules were added, the energy changes fluctuated around zero, indicating reduced incremental stabilization and the progressive saturation of the hydration shell.

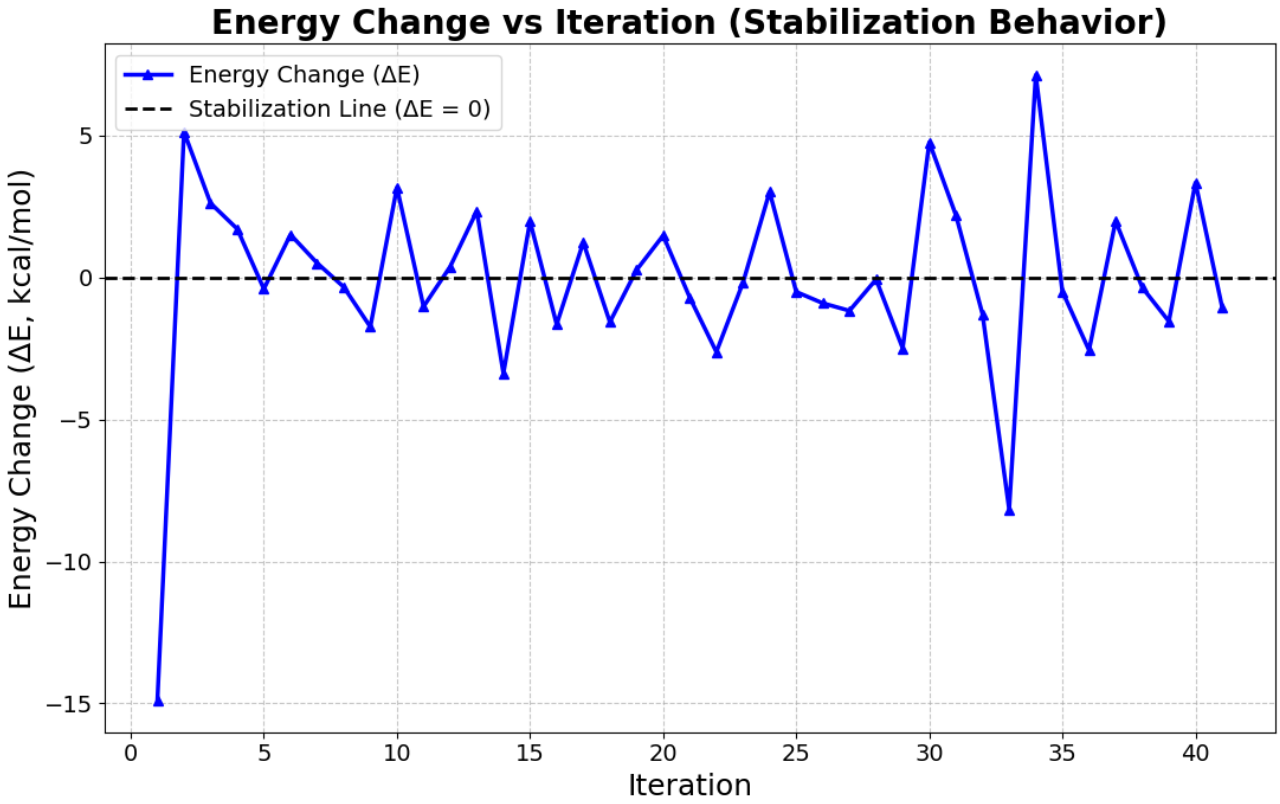

Figure 9: Energy change (ΔE) per added water molecule during NaCl solvation. The stabilization line ($\Delta E = 0$) indicates convergence toward bulk-like behavior.

Reaction Pathway and Transition State Characterization

The minimum energy path (MEP) connecting reactant and product geometries was successfully mapped using the NEB method, with ten intermediate images. The resulting energy profile is shown in Figure 10. A clear energy maximum was observed at Image 5, corresponding to the transition state (TS) structure.

To confirm the nature of this transition point, a frequency calculation was performed, yielding a single imaginary frequency associated with motion along the reaction coordinate. This validates the TS as a true first-order saddle point on the potential energy surface.

The energy difference between the TS and the reactant state gives the activation energy barrier. After applying thermodynamic corrections, the Gibbs free energy of activation ($\Delta G^{\ddagger}$) was obtained at 298 K. Using Eyring's equation, the rate constant for the reaction was estimated.

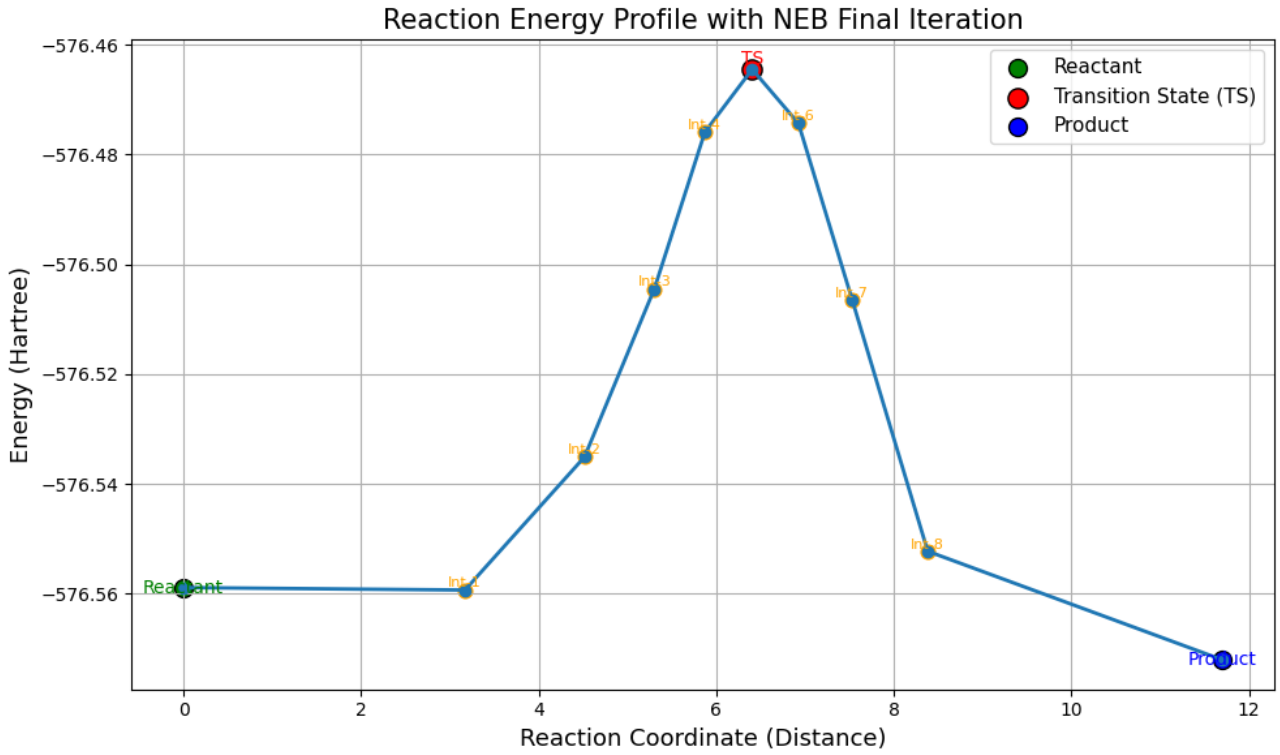

Figure 10: NEB energy profile showing the minimum energy path between reactant and product geometries. The red point denotes the identified transition state (TS).

Conclusion

This study demonstrates that Density Functional Theory (DFT) can reliably predict a wide range of molecular behaviors with both qualitative and quantitative consistency. From conformational energy profiles in alkanes and vibrational spectra in water clusters to ionic solvation and transition-state characterization, the results aligned closely with theoretical models and experimental benchmarks. Where discrepancies emerged—such as the overestimation of O–H bond stiffness—they were well-understood and attributable to known approximations in gas-phase DFT, such as the neglect of anharmonic or thermal effects.

Crucially, DFT proved to be most powerful in identifying chemical trends: potential energy surfaces, stabilization behaviors, and activation barriers were captured with high fidelity. These insights underscore DFT's value as a predictive framework—not merely for reproducing numbers, but for guiding chemical understanding and experimental design.

These findings directly connect to Paul Dirac's 1929 insight: while the governing equations of chemistry are known, solving them exactly is intractable. DFT bridges that gap. This work shows how far modern computational chemistry has come in making Dirac's vision experimentally actionable.

However, the project revealed practical workflow challenges, especially when using ORCA:

- **Slow runtimes:** Geometry optimizations and frequency calculations for large systems (e.g., $(H_2O)_{15}$, transition states) took hours to days.

- **Excessive file generation:** The O–H distance in the dimer was scanned over 0.40–5.00 Å in 0.01 Å steps—producing over 400 single-point jobs. In total, over 2,000 ORCA output files were processed.
- **Manual post-processing:** ORCA's text-based outputs required custom Python scripts for energy extraction, parsing, and trend visualization.
- **Error handling:** SCF convergence failures and inconsistent outputs increased runtimes and rework.
- **Storage overhead:** File volume from multiple scans and large molecules strained local storage and organization.

Although the focus was on two-body interactions for tractability, the desire to explore three- and four-body systems remained. Treating such systems with DFT became prohibitively expensive in ORCA, pointing to the need for scalable alternatives—such as SEAM force fields or ML-assisted potential models—that better handle cooperative effects in solvation or reaction networks.

To streamline data handling, Google Colab notebooks were developed for batch processing, energy extraction, and automated plotting. These helped reduce friction but highlighted the broader need for smarter, more efficient workflows.

Looking ahead, machine learning (ML) offers promising solutions. Neural networks like ANI or SchNet, trained on high-quality quantum datasets, can rapidly predict energy landscapes and transition structures. Coupled with uncertainty quantification, these models could prioritize high-value geometries and reduce redundant DFT jobs. Future work should explore hybrid ML–DFT strategies that retain accuracy while accelerating throughput—moving us closer to Dirac's dream of a chemically complete, computationally practical theory.

References

[1] Christoph Bannwarth, Sebastian Ehlert, and Stefan Grimme. Gfn2-xtb—an accurate and broadly parametrized self-consistent tight-binding quantum chemical method with multipole electrostatics and density-dependent dispersion contributions. *J. Chem. Theory Comput.*, 15(3):1652–1671, 2019.

[2] Kieron Burke. Perspective on density functional theory. *The Journal of Chemical Physics*, 136(15):150901, 2012.

[3] Jeng-Da Chai and Martin Head-Gordon. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics*, 10(44):6615–6620, 2008.

[4] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. Challenges for density functional theory. *Chemical Reviews*, 112(1):289–320, 2012.

[5] P.A.M. Dirac. Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 123(792):714–733, 1929.

[6] Sebastian Ehlert, Marcel Stahn, Sebastian Spicher, and Stefan Grimme. Robust and efficient implicit solvation model for fast semiempirical methods. *Journal of Chemical Theory and Computation*, 17(7):4250–4261, 2021.

[7] Google Colaboratory. Colaboratory – interactive python notebooks. `https://colab.research.google.com/`, 2025. Accessed: 2025-04-24.

[8] Stefan Grimme, Christoph Bannwarth, and Philip Shushkov. A robust and accurate tight-binding quantum chemical method for structures, vibrational frequencies, and noncovalent interactions of large molecular systems. *J. Chem. Theory Comput.*, 13(5):1989–2009, 2017.

[9] Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek, and Geoffrey R Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of cheminformatics*, 4(1):17, 2012.

[10] Graeme Henkelman, Blas P. Uberuaga, and Hannes Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*, 113(22):9901–9904, 2000.

[11] Hannes Jónsson, G. Mills, and Karsten W. Jacobsen. Nudged elastic band method for finding minimum energy paths of transitions. In *Classical and Quantum Dynamics in Condensed Phase Simulations*, pages 385–404. World Scientific, 1998.

[12] Frank Neese, Frank Wennmohs, Ulrich Becker, and Christoph Riplinger. The orca quantum chemistry program package. *The Journal of Chemical Physics*, 152(22):224108, 2020.

[13] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics*, 7(18):3297–3305, 2005.

[14] G. A. Zhurko. Chemcraft – graphical program for visualization of quantum chemistry computations. `https://chemcraftprog.com`, 2005. Ivanovo, Russia.