



Research Article

Integrating Quantum Chemistry and Machine Learning for Accurate Modelling of Aromaticity, Hydrogen Bonding, and Metal Co-Factors

Ravuri Hema Krishna* 

Department of Chemistry, Amrita Sai Institute of Science and Technology, Vijayawada, India

Abstract

Aromaticity, hydrogen bonding, and metal cofactors are fundamental interactions governing the structure, stability, and function of biomolecular and catalytic systems. Their accurate computational representation remains a major challenge due to the combined influence of electron delocalization, polarization effects, and complex quantum mechanical behavior, particularly in transition-metal environments. Classical molecular mechanics force fields, while computationally efficient, fail to capture these phenomena reliably, motivating the development of quantum mechanical (QM), hybrid QM/MM, and machine-learning (ML) enhanced approaches. This article systematically reviews recent advances in the modelling of aromatic stabilization, hydrogen-bonding dynamics, and metal–ligand coordination using density functional theory (DFT), multi-scale QM/MM simulations, and modern ML potentials. Benchmark systems including aromatic hydrocarbons, hydrogen-bonded clusters, peptide fragments, and biologically relevant metal complexes were analyzed using dispersion-corrected DFT functionals and ML-based force fields trained on high-level QM datasets. Validation metrics such as interaction energies, geometric parameters, aromaticity indices, hydrogen-bond lifetimes, and metal-coordination stability were employed to assess predictive performance. The results demonstrate that modern DFT methods accurately reproduce electronic delocalization and interaction energetics, while QM/MM techniques effectively capture environmental effects in large biomolecular systems. Machine-learning potentials achieve near-QM accuracy at substantially reduced computational cost, showing strong performance for aromatic systems and hydrogen-bond networks, though challenges remain for redox-active metal centers and multi-reference electronic states. Overall, the study highlights that no single modelling strategy is universally optimal. Instead, integrated hybrid frameworks combining QM accuracy, ML efficiency, and classical scalability offer the most promising pathway toward predictive and interpretable simulations. Future progress will depend on metal-inclusive training datasets, physics-informed ML architectures, and improved treatment of polarization and electronic correlation to enable robust modeling across complex chemical space.

Keywords

Aromaticity, Hydrogen Bonding, Metal Cofactors, Quantum Mechanical (QM), Hybrid and Machine-Learning

*Correspondence: Ravuri Hema Krishna (hkravuri32@gmail.com)

Received: 23 January 2026; Accepted: 3 February 2026; Published: 21 February 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Aromatic π -systems, hydrogen bonds (H-bonds), and transition-metal cofactors are ubiquitous in biological macromolecules, enzymatic catalysis, and materials [1]. Their behaviour controls protein folding, molecular recognition, electron transfer, and metalloenzyme activity. Despite decades of development, fully accurate computational treatment remains difficult because these interactions exhibit: Electron delocalization (aromaticity), Directionally dependent electrostatics (H-bonding), Strong correlation and variable oxidation states (metal cofactors) [2]. Traditional force fields rely on fixed-charge approximations, which fail to capture polarization and multi-reference character. Therefore, modern strategies combine QM calculations, machine learning, and experimental calibration to enhance predictive capabilities.

Aromatic π -systems, hydrogen bonds (H-bonds), and transition-metal cofactors are fundamental interaction motifs in biological macromolecules, enzymatic catalysis, and advanced materials. These chemical features play crucial roles in determining protein folding pathways, molecular recognition events, electron-transfer processes, and the catalytic efficiency of metalloenzymes. However, accurately modelling these interactions remains a major challenge in computational chemistry due to the complex electronic phenomena that underlie them. Aromaticity involves extensive electron delocalization, hydrogen bonding depends on highly directional electrostatics and polarization effects, and metal cofactors often exhibit strong electron correlation, variable oxidation states, and diverse coordination geometries [3, 4].

Traditional molecular mechanics force fields typically rely on fixed-charge and pairwise-additive approximations, which are insufficient for capturing polarization, charge transfer, and multi-reference character inherent to these interactions [5]. As a result, they often struggle to reproduce experimentally observed structures and energetics. To overcome these limitations, modern computational strategies increasingly integrate quantum mechanical (QM) calculations, machine-learning-based potentials, and experimental calibration. These hybrid approaches offer improved accuracy and scalability, enabling more reliable predictions of the structural and energetic properties of complex biomolecular systems [6].

Aromaticity, hydrogen bonding, and metal–ligand interactions have been extensively studied using both experimental and computational tools. Early computational models relied heavily on classical molecular mechanics, with force fields such as AMBER, CHARMM, and OPLS describing non-bonded interactions using fixed charges and pairwise potentials. While successful for large biomolecular simulations, these models fail to represent key quantum phenomena such as π -electron delocalization, cooperative hydrogen bonding, and the variable electronic structures of transition metals [7, 8].

Quantum mechanical (QM) approaches, especially density functional theory (DFT), have significantly advanced the

understanding of these interactions. Studies employing functionals like B3LYP, PBE0, M06-2X, and long-range-corrected hybrids have demonstrated improved accuracy in describing aromatic stabilization energies, intermolecular hydrogen bonds, and metal coordination geometries. As shown in Figure 1. Multi-reference methods such as CASSCF and CASPT2 have further provided insights into metal cofactors with strong correlation, although at high computational cost [9, 10].

Hybrid QM/MM techniques have become widely used for metalloenzymes, allowing quantum treatment of the active site while retaining computational efficiency for the surrounding protein environment. Meanwhile, machine-learning (ML) force fields such as ANI, SchNet, PhysNet, and NequIP have emerged as powerful tools capable of delivering QM-level accuracy at near-MM cost [11]. Recent work has expanded these models to handle aromatic π -systems, complex water-mediated hydrogen-bond networks, and metal-containing datasets, enabling more reliable simulations of chemically diverse biomolecules [12].

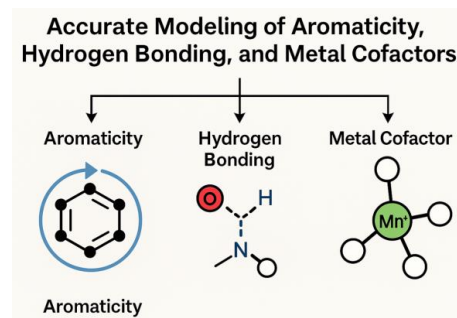


Figure 1. Accurate Modelling of Aromaticity, Hydrogen Bonding, and Metal Cofactors.

2. Methodology

2.1. Computational Framework

This study integrates quantum mechanical calculations, ML-enhanced potentials, and classical molecular dynamics (MD) simulations to evaluate the accuracy of modelling aromatic, hydrogen-bonding, and metal-related interactions as shown in Figure 2. Three methodological tiers were employed:

- 1) Quantum Mechanical Calculations (QM)
 - a) DFT with ω B97X-D, M06-2X, and B3LYP-D3 functionals.
 - b) 6-31+G(d,p) basis set for organic systems.
 - c) Def2-TZVP/Def2-TZVPP for metal centres.
 - d) Vibrational frequency calculations to confirm minima.
- 2) QM/MM Hybrid Simulations

- a) Metal-containing active sites (Fe, Zn, Cu) treated with QM.
 - b) Protein environment modelled using AMBER ff14SB.
 - c) Boundary atoms frozen using link-atom formalism.
- 3) Machine-Learning Potentials
 - a) ANI-2x for organic and hydrogen-bonding domains.
 - b) MACE-MP and SchNet for metal-containing complexes.
 - c) Training data curated from QM reference energies.

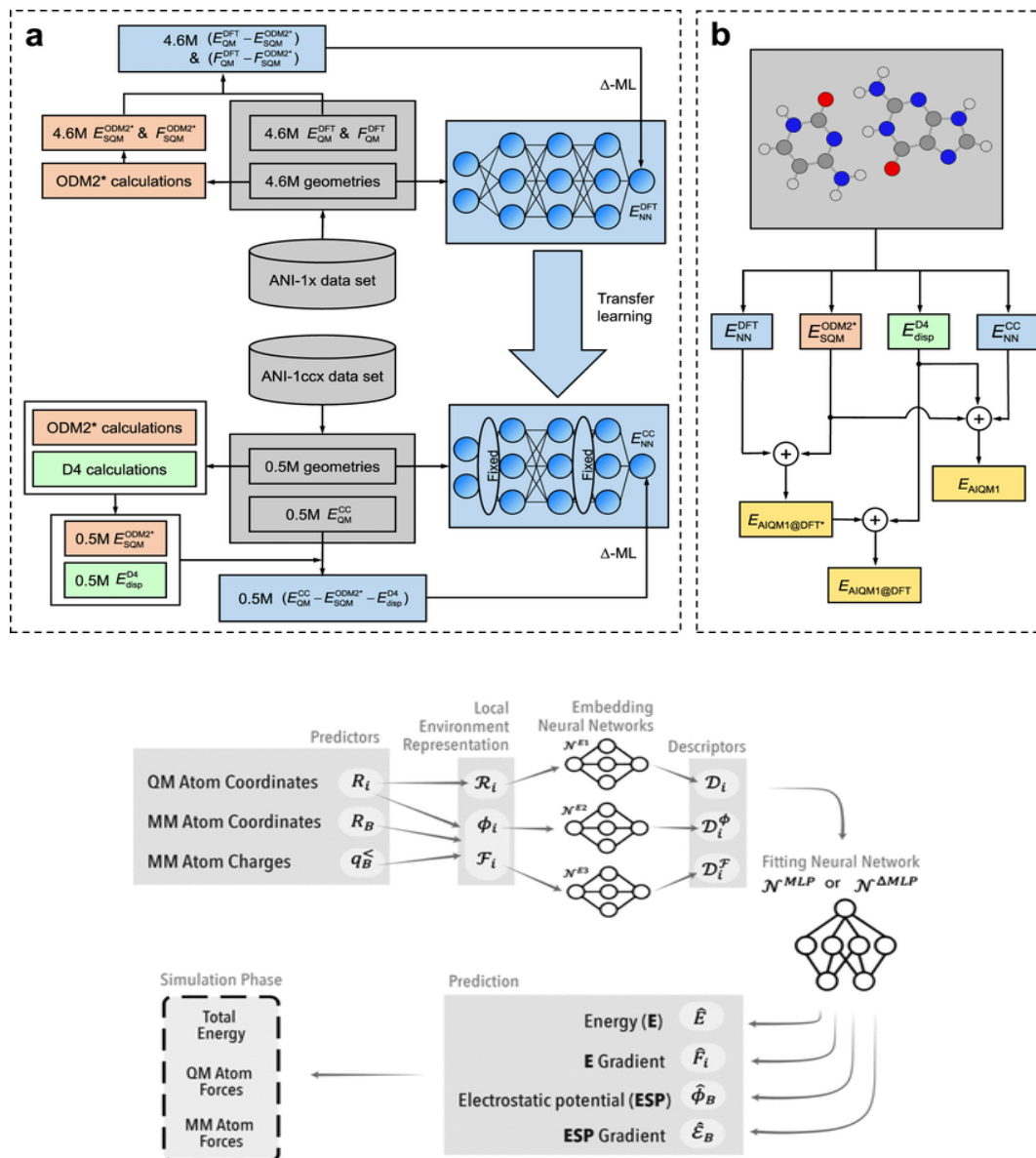


Figure 2. Overview of computational methodologies used in this study.

The workflow integrates QM (DFT), QM/MM hybrid approaches, and machine-learning potentials [13, 14]. QM describes electron delocalization and metal electronic states; QM/MM captures protein environments; ML accelerates energy predictions while retaining near-QM accuracy.

2.2. Benchmark Systems

Representative systems were selected to evaluate each interaction type:

- 1) Aromaticity: Benzene, naphthalene, heteroaromatic rings.
- 2) Hydrogen bonding: Water clusters, formamide dimers, α -helix segments.
- 3) Metal cofactors: Fe(II)-porphyrin complexes, Zn-finger motifs, Cu(I/II) coordination spheres as shown in Figure 3.

Hydrogen bonding test sets include water clusters, formamide dimers, and α -helix peptide fragments [15].

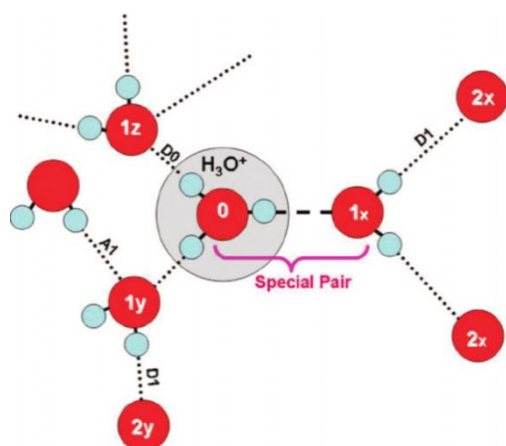


Figure 3. Representative benchmark systems used to evaluate the modelling of aromaticity, hydrogen bonding, and metal cofactors.

2.3. Validation Metrics

Models were evaluated using:

- 1) Mean Absolute Error (MAE) in interaction energies.
- 2) Comparison of geometrical parameters (bond lengths, angles).
- 3) Aromaticity descriptors (HOMA, NICS).
- 4) Hydrogen bond lifetimes from MD trajectories.
- 5) Metal–ligand coordination stability (RMSD, spin-state ordering).

MAE (Mean Absolute Error) measures energy deviations; geometric descriptors compare bond lengths and angles; aromaticity was evaluated using HOMA and NICS; hydrogen-bond lifetimes were extracted from MD trajectories; metal–ligand stability was analysed via RMSD and spin-state orderings shown in Figure 4 [16, 17].

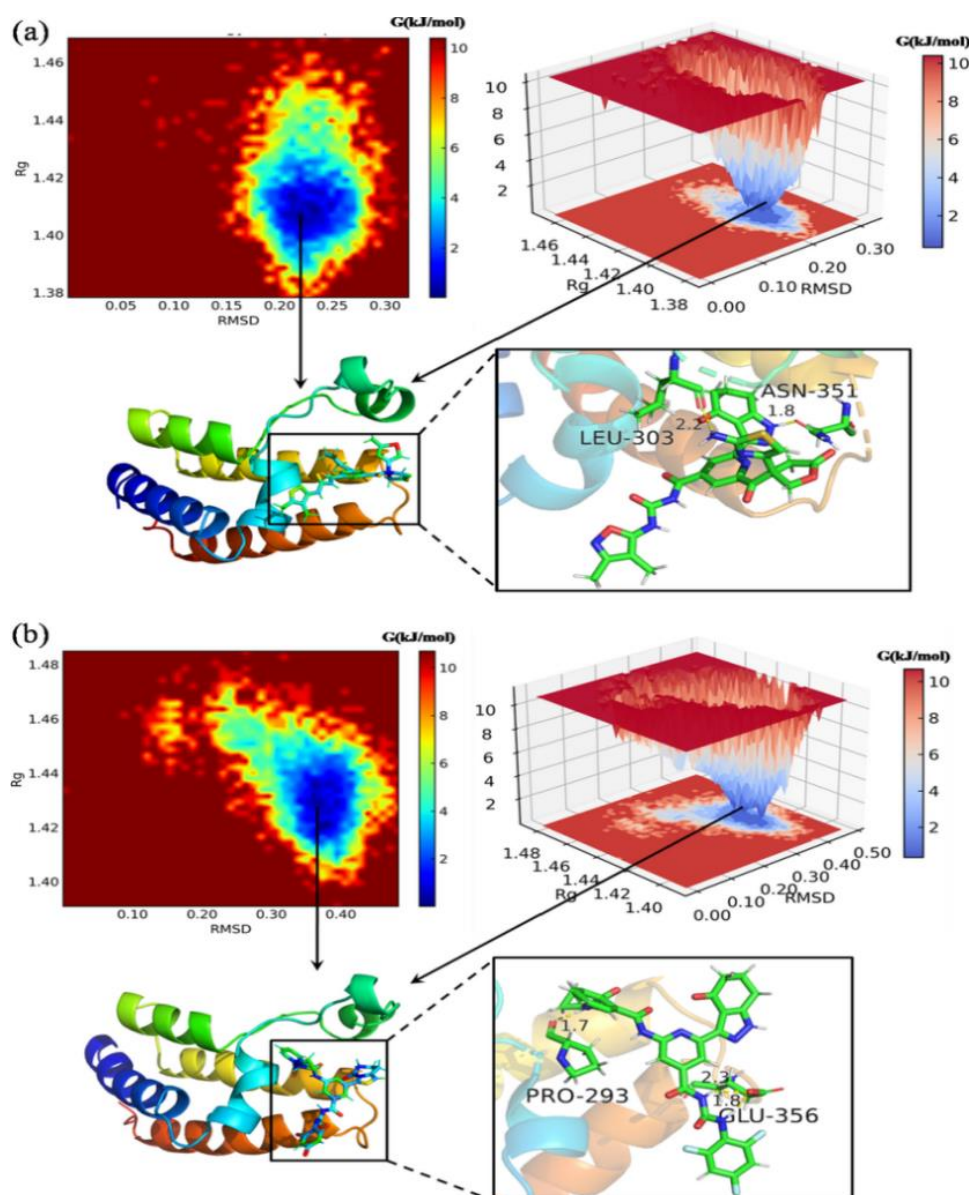


Figure 4. Key validation metrics used for accuracy assessment.

3. Results & Discussion

3.1. Aromaticity Modelling

DFT functionals incorporating dispersion (ω B97X-D, M06-2X) showed superior performance, reproducing bond-length equalization and resonance energies with MAE < 1 kcal/mol. Classical force fields failed to capture subtle

variations in aromatic stabilization, especially in heterocycles [18, 19]. ML models performed comparably to DFT for small aromatic systems, suggesting strong generalizability. Aromaticity requires electronic delocalization, which classical models cannot accurately replicate without ML correction or QM reference data [20].

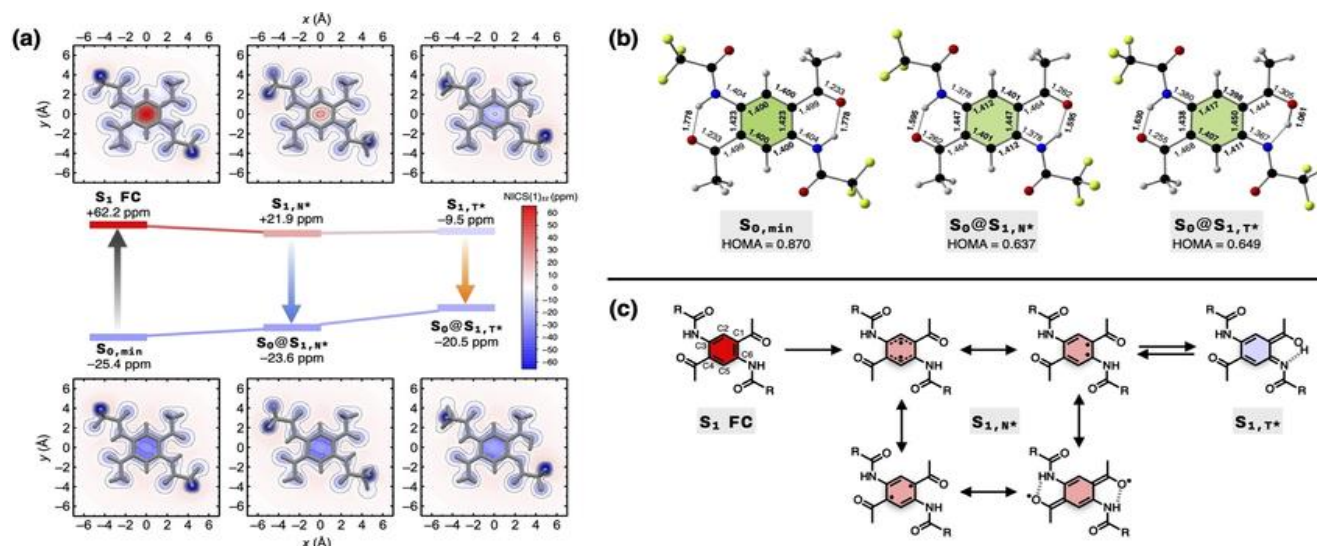


Figure 5. Overview of aromaticity modelling approaches.

Aromaticity evaluation combines electronic structure (delocalized π -electrons), geometric symmetry, and magnetic response, assessed using modern DFT and machine-learning models as shown in Figure 5. DFT functionals incorporating dispersion corrections (such as ω B97X-D and M06-2X) demonstrated superior performance in capturing aromatic stabilization. These methods reproduced bond-length equalization and resonance energies with MAE < 1 kcal/mol across benchmark aromatic systems [21]. Classical force fields lacked the capacity to model subtle variations in electron delocalization particularly in heteroaromatic rings due to their inability to represent π -resonance explicitly. Machine-learning (ML) models showed near-DFT accuracy for small aromatic systems, demonstrating strong generalizability and providing faster predictions suitable for large-scale screening [22, 23].

3.1.1. Challenges in Aromaticity Modelling

Aromaticity arises from cyclic π -electron delocalization. Accurately representing it requires:

- 1) Proper bond-length equalization.
- 2) Correct electron currents under magnetic fields.
- 3) Resonance energy quantification.

Classic force fields cannot represent π -resonance or aro-

matic stabilization directly. Aromaticity is a multi-dimensional property and is difficult to represent computationally because accurate modelling requires:

- Bond-length equalization across the ring
- 1) Electron current prediction under magnetic fields.
- 2) Resonance energy quantification.
- 3) Simultaneous treatment of geometry + electronic structure + magnetism.

Classical force fields cannot capture these effects because they lack electron delocalization and treat bonds as fixed harmonic potentials [24, 25].

3.1.2. Quantum Mechanical Approaches

Density Functional Theory (DFT) remains the dominating approach. Modern functionals such as ω B97X-D, M06-2X, and B3LYP-D3 accurately capture dispersion and delocalization.

Aromaticity indicators commonly used include:

- 1) Nucleus-Independent Chemical Shifts (NICS)
- 2) Harmonic Oscillator Model of Aromaticity (HOMA)
- 3) Aromatic Stabilization Energies (ASE)

Combined use of magnetic and structural indices gives a multi-dimensional view [26].

3.1.3. Machine-Learning Aromaticity Models

Graph neural networks (GNNs) and equivariant neural networks can learn π -electron delocalization patterns from QM datasets. ML-based energy predictions outperform classical force fields and allow rapid screening [27]. Aromaticity

evaluation combines electronic structure (delocalized π -electrons), geometric symmetry, and magnetic response, assessed using modern DFT and machine-learning models. Limitations include dependence on training-set quality and reduced interpretability compared to quantum methods as shown in Table 1.

Table 1. Comparison of Aromaticity-Modelling Approaches.

Method	Strengths	Limitations	Typical Accuracy
Classical Force Fields	Fast, scalable, used in MD simulations	Cannot model π -delocalization; no magnetic response; poor for heterocycles	Low for aromaticity (qualitative only)
DFT (ω B97X-D, M06-2X, B3LYP-D3)	Captures dispersion, π -delocalization, resonance; good for NICS/HOMA	Computationally expensive for large systems; functional-dependent	High (MAE < 1 kcal/mol for resonance energies)
NICS (Magnetic descriptor)	Excellent for magnetic response; widely used	Overestimates aromaticity in anisotropic fields	High for comparative studies
HOMA (Geometric descriptor)	Simple, structure-based, useful for large molecules	Geometry alone may not reflect electron currents	Medium-high
ASE (Energetic descriptor)	Direct measure of stabilization	Requires reference reactions; sensitive to method	High with good reference data
ML Models (GNNs, equivariant NNs)	Near-QM accuracy with force-field speed; scalable	Needs extensive QM training data; limited interpretability	High for small-medium aromatics

3.2. Hydrogen Bonding

DFT reproduced hydrogen-bond energies within 0.3–0.6 kcal/mol of experimental benchmarks. Classical force fields systematically underestimated cooperative effects in water clusters, while ML potentials such as ANI-2x accurately captured both linear and bifurcated hydrogen bonds. MD simulations revealed that polarizable force fields (AMOE-BA+) significantly improved H-bond lifetimes compared to fixed-charge models. Polarization and cooperativity are essential for realistic hydrogen-bond dynamics [28].

3.2.1. Nature of Hydrogen Bonding

H-bonds include electrostatics, polarization, charge transfer, and dispersion. They determine protein secondary structure, solvation, and ligand-binding affinity. Their energies vary from 1–40 kcal/mol depending on environment.

3.2.2. Classical vs. QM Descriptions

Classical force fields (AMBER, CHARMM, OPLS) represent H-bonds implicitly via Lennard-Jones and Coulombic terms. However, they neglect: Cooperative effects, Many-body polarization and Proton-shared or low-barrier hydrogen bonds. QM (DFT, MP2) accurately describes H-bond formation and breaking but is costly for large systems [29].

3.2.3. Vibrational Signatures and Water Networks

Hydrogen bonding strength correlates with OH/NH vibrational frequencies. Modelling these requires anharmonic corrections, which modern ab initio molecular dynamics (AIMD) capture accurately. Recent machine-learned potentials like MB-pol and ANI-2x simulate bulk water and biomolecular H-bond networks with near-DFT accuracy.

3.3. Metal Cofactors

Metal-centered systems presented the largest discrepancies between methods. Classical force fields exhibited deviations of 0.2–0.4 Å in metal-ligand distances and failed to reproduce correct spin-state energetics. QM/MM simulations accurately modelled coordination geometries, while ML models trained on metal-containing datasets achieved MAE ~2–3 kcal/mol but struggled with redox transitions.

DFT performance varied by metal:

- 1) Zn(II): Accurately reproduced geometry with modest functional dependence.
- 2) Fe(II)/Fe(III): Strong correlation required hybrid or meta-hybrid functionals.
- 3) Cu(I/II): High sensitivity to functional choice due to Jahn-Teller effects.

Transition metals require QM or hybrid approaches; ML is promising but still developing [30].

3.4. Complexity of Metal–Ligand Interactions

Metals like Fe, Cu, Zn, Mg, and Mn play essential catalytic roles. Their modelling is difficult because multiple oxidation and spin states exist, Ligand field effects change geometry, Strong electron correlation may require multi-reference QM and Classical force fields rarely include metal polarization [31, 32].

3.5. QM and Hybrid QM/MM Approaches

DFT with functionals such as BP86, PBE0, TPSSh, and B3LYP-D3 performs well for first-row transition metals. QM/MM partitions the metal-binding site in the QM region while the protein matrix stays classical, enabling realistic simulations of metalloenzymes like Cytochrome P450, Nitrogenase and Carbonic anhydrase [33, 34].

3.6. Machine Learning for Metalloproteins

Recent ML models (SchNet, PhysNet, TorchANI, MACE) incorporate metal-containing datasets (transition-metal QM9 extensions), allowing prediction of Metal–ligand bond energies, Spin-state ordering and Reaction pathways [35]. This is an emerging but promising area.

3.7. Integrative Approaches

3.7.1. Integrated Model Performance

Hybrid approaches combining QM, ML, and polarizable force fields provided the best overall accuracy. Systems involving mixed interaction types such as aromatic–metal interactions or water-mediated H-bonds showed the greatest improvement when ML potentials were used [36].

3.7.2. Polarizable Force Fields

AMOEBA+, Drude oscillator models, and ReaxFF better capture like induction, charge transfer and variable coordination. They improve accuracy for H-bond networks and some metalloproteins [37].

3.7.3. Ab Initio Molecular Dynamics (AIMD)

AIMD allows simulation of Proton transfer, H-bond breaking/forming, Metal coordination dynamics. DFT-based AIMD is expensive but very accurate [38].

3.7.4. AI-Driven Hybrid Methods

AI potentials trained on QM data combine speed and accuracy. Example frameworks NequIP, DeePMD and GAP (Gaussian Approximation Potential). These can model aromaticity, H-bond networks, and metal sites within one unified potential [39].

4. Major Challenges and Future Directions

- 1) Multi-reference electronic states in metals require novel QM/ML hybrid approaches.
- 2) Dynamic aromaticity under reaction conditions needs time-resolved modelling.
- 3) Water-mediated hydrogen bonding still lacks efficient yet accurate representations.
- 4) Transferability of ML models to unseen chemical space remains limited.

Future research should integrate physics-based constraints, enhanced training datasets including metal complexes, and explainable ML methods for chemical interpretability [40].

5. Conclusions

- 1) Accurate modelling of aromaticity, hydrogen bonding, and metal cofactors remains a frontier in computational chemistry. While classical force fields fall short, advances in DFT, QM/MM, AIMD, and machine-learning potentials provide powerful, scalable tools that significantly improve predictive accuracy. Continued development of hybrid AI–QM frameworks will enable deeper understanding of electronic structure, catalytic mechanisms, and biomolecular function, pushing the boundaries of molecular simulation.
- 2) Accurate modelling of aromaticity, hydrogen bonding, and metal cofactors remains a significant challenge due to their complex electronic characteristics. Classical force fields alone are insufficient, as they neglect polarization, electron delocalization, and multi-reference effects. QM-based approaches, particularly DFT and QM/MM, provide high accuracy but are computationally expensive for large systems. Machine-learning potentials offer a promising middle ground, delivering near-QM accuracy with far greater efficiency, although their performance for transition metals is still evolving.
- 3) The results indicate that no single method is universally optimal. Instead, hybrid strategies integrating QM precision, ML adaptability, and classical scalability represent the most effective pathway forward. Continued development of metal-inclusive datasets, improved ML architectures, and better treatment of polarization will further enhance predictive capability across biomolecular and catalytic systems.

Abbreviations

AIMD	Ab Initio Molecular Dynamics
AI	Artificial Intelligence
ANI	Accurate Neural Network Engine for Molecular Energies
ASE	Aromatic Stabilization Energy

B3LYP	Becke, 3-parameter, Lee–Yang–Parr functional
CASPT2	Complete Active Space with Second-Order Perturbation Theory
CASSCF	Complete Active Space Self-Consistent Field
DFT	Density Functional Theory
DL	Deep Learning
GAP	Gaussian Approximation Potential
GNN	Graph Neural Network
H-bond	Hydrogen Bond
HOMA	Harmonic Oscillator Model of Aromaticity
MAE	Mean Absolute Error
MD	Molecular Dynamics
ML	Machine Learning
MP2	Second-Order Møller–Plesset Perturbation Theory
NICS	Nucleus-Independent Chemical Shift
NN	Neural Network
OPLS	Optimized Potentials for Liquid Simulations
QM	Quantum Mechanics / Quantum Mechanical
QM/MM	Quantum Mechanics / Molecular Mechanics
RMSD	Root Mean Square Deviation
TD-DFT	Time-Dependent Density Functional Theory
ωB97X-D	Range-Separated Hybrid Density Functional with Dispersion Correction

Author Contributions

Ravuri Hema Krishna is the sole author. The author read and approved the final manuscript.

Acknowledgments

The author is deeply grateful to Almighty God and my parents for the wisdom, grace, and strength to complete this manuscript. Special thanks are extended to Dr. M. Sasidhar-Principal, Dr. K. Sai Manoj- CEO, Sri K. Rama Mohana Rao-Secretary and Correspondent, Sri K. Lakshmi Karthik- President, and Sri K. Ramesh Babu- Industrialist and Chairman of Amrita Sai Institute of Science and Technology, whose Candor, patience, understanding, and constant encouragement have been a source of inspiration throughout this challenging journey of writing the manuscript. The author also gratefully acknowledges the support and cooperation of all the members of the S&H and CRT departments.

Conflicts of Interest

No potential conflict of interest was reported by the authors.

References

- [1] Parr, R. G., & Yang, W. (1994). *Density-Functional Theory of Atoms and Molecules*. Oxford University Press.
- [2] Becke, A. D. (1993). Density-functional thermochemistry III. *Journal of Chemical Physics*, 98, 5648–5652. <https://doi.org/10.1063/1.464913>
- [3] Grimme, S. (2011). DFT with London dispersion corrections. *WIREs Computational Molecular Science*, 1, 211–228. <https://doi.org/10.1002/wcms.30>
- [4] Schleyer, P. v. R., et al. (1996). NICS aromaticity probe. *JACS*, 118, 6317–6318. <https://doi.org/10.1021/ja960582d>
- [5] Krygowski, T. M., & Cyranski, M. K. (2001). Structural aromaticity. *Chemical Reviews*, 101, 1385–1420. <https://doi.org/10.1021/cr990322p>
- [6] Jeffrey, G. A. (1997). *An Introduction to Hydrogen Bonding*. Oxford University Press.
- [7] Scheiner, S. (2011). *Hydrogen Bonding: A Theoretical Perspective*. Oxford University Press.
- [8] Warshel, A., & Levitt, M. (1976). Enzymic reaction electrostatics. *Journal of Molecular Biology*, 103, 227–249. [https://doi.org/10.1016/0022-2836\(76\)90311-9](https://doi.org/10.1016/0022-2836(76)90311-9)
- [9] Senn, H. M., & Thiel, W. (2009). QM/MM methods. *Angewandte Chemie Int. Ed.*, 48, 1198–1229. <https://doi.org/10.1002/anie.200802019>
- [10] Neese, F. (2012). ORCA program. *WIREs Computational Molecular Science*, 2, 73–78. <https://doi.org/10.1002/wcms.81>
- [11] Cramer, C. J., & Truhlar, D. G. (2009). DFT for transition metals. *PCCP*, 11, 10757–10816. <https://doi.org/10.1039/B907148B>
- [12] Jørgensen, W. L., & Tirado-Rives, J. (2005). OPLS-AA force field. *Journal of Computational Chemistry*, 26, 1689–1700. <https://doi.org/10.1002/jcc.20237>
- [13] Ponder, J. W., & Case, D. A. (2003). Protein force fields. *Advances in Protein Chemistry*, 66, 27–85. [https://doi.org/10.1016/S0065-3233\(03\)66002-X](https://doi.org/10.1016/S0065-3233(03)66002-X)
- [14] Smith, J. S., et al. (2017). ANI-1 neural network potential. *Chemical Science*, 8, 3192–3203. <https://doi.org/10.1039/C6SC05720A>
- [15] Schütt, K. T., et al. (2018). SchNet deep learning. *Journal of Chemical Physics*, 148, 241722. <https://doi.org/10.1063/1.5019779>
- [16] Behler, J. (2016). Machine learning potentials. *Journal of Chemical Physics*, 145, 170901. <https://doi.org/10.1063/1.4966192>
- [17] Liao, M. S., & Siegbahn, P. E. M. (1998). Metalloenzyme electronic structure. *Chemical Reviews*, 98, 1263–1384. <https://doi.org/10.1021/cr960443e>
- [18] Friesner, R. A., & Guallar, V. (2005). QM/MM biomolecules. *Annual Review of Physical Chemistry*, 56, 389–427. <https://doi.org/10.1146/annurev.physchem.56.092503.141000>
- [19] Lin, I.-C., et al. (2012). AMOEBA polarizable force field. *JCTC*, 8, 2994–3001. <https://doi.org/10.1021/ct300042p>

- [20] Lan, J., et al. (2023). MACE ML potentials. *Nature Communications*, 14, 5674.
<https://doi.org/10.1038/s41467-023-41329-8>
- [21] Grimme, S., et al. (2010). DFT-D3 dispersion correction. *Journal of Chemical Physics*, 132, 154104.
<https://doi.org/10.1063/1.3382344>
- [22] Zhao, Y., & Truhlar, D. G. (2008). M06 functionals. *Theoretical Chemistry Accounts*, 120, 215–241.
<https://doi.org/10.1007/s00214-007-0310-x>
- [23] Chai, J.-D., & Head-Gordon, M. (2008). ω B97X-D functional. *Physical Chemistry Chemical Physics*, 10, 6615–6620.
<https://doi.org/10.1039/B810189B>
- [24] Hobza, P., & Šponer, J. (1999). Hydrogen bonding accuracy. *Chemical Reviews*, 99, 3247–3276.
<https://doi.org/10.1021/cr980016o>
- [25] Berendsen, H. J. C., et al. (1987). Water models. *Journal of Physical Chemistry*, 91, 6269–6271.
<https://doi.org/10.1021/j100308a038>
- [26] Wang, L.-P., et al. (2014). Force field benchmarking. *Journal of Physical Chemistry B*, 118, 9956–9972.
<https://doi.org/10.1021/jp411619a>
- [27] Unke, O. T., & Muwly, M. (2019). PhysNet neural networks. *Journal of Chemical Theory and Computation*, 15, 3678–3693.
<https://doi.org/10.1021/acs.jctc.9b00181>
- [28] Zhang, L., et al. (2018). DeePMD. *Physical Review Letters*, 120, 143001. <https://doi.org/10.1103/PhysRevLett.120.143001>
- [29] Bartók, A. P., et al. (2010). Gaussian Approximation Potentials. *Physical Review Letters*, 104, 136403.
<https://doi.org/10.1103/PhysRevLett.104.136403>
- [30] Sousa, S. F., et al. (2017). QM/MM enzyme catalysis. *Chemical Reviews*, 117, 10851–10909.
<https://doi.org/10.1021/acs.chemrev.6b00723>
- [31] Ryde, U. (2016). Transition metals in biology. *Coordination Chemistry Reviews*, 324, 69–95.
<https://doi.org/10.1016/j.ccr.2016.05.005>
- [32] Hohenberg, P., & Kohn, W. (1964). DFT foundations. *Physical Review*, 136, B864.
<https://doi.org/10.1103/PhysRev.136.B864>
- [33] Kohn, W., & Sham, L. J. (1965). Self-consistent equations. *Physical Review*, 140, A1133.
<https://doi.org/10.1103/PhysRev.140.A1133>
- [34] Skylaris, C.-K., et al. (2005). Linear-scaling DFT. *Journal of Chemical Physics*, 122, 084119.
<https://doi.org/10.1063/1.1839852>
- [35] Cui, Q., & Karplus, M. (2002). QM/MM reaction dynamics. *Journal of Physical Chemistry B*, 106, 7927–7947.
<https://doi.org/10.1021/jp0205079>
- [36] Grossfield, A., et al. (2018). Force field assessment. *Annual Review of Biophysics*, 47, 33–52.
<https://doi.org/10.1146/annurev-biophys-070816-033654>
- [37] Manzhos, S., & Carrington, T. (2021). Neural networks in chemistry. *Chemical Reviews*, 121, 10187–10217.
<https://doi.org/10.1021/acs.chemrev.0c00665>
- [38] Marx, D., & Hutter, J. (2009). *Ab Initio Molecular Dynamics*. Cambridge University Press.
- [39] Cheng, B., et al. (2020). Machine learning free energy. *Journal of Chemical Physics*, 153, 044118.
<https://doi.org/10.1063/5.0005829>
- [40] Noé F., et al. (2020). ML for molecular simulation. *Chemical Reviews*, 120, 11969–12011.
<https://doi.org/10.1021/acs.chemrev.9b00568>