Sudarshan Research Journal
Volume – 1, Issue - 5, September-2023
ISSN No: 2583-8792

COMPUTATIONAL CHEMISTRY: FROM ALGORITHMS TO APPLICATIONS

Himanshu Jain

Assistant Professor, Department of Chemistry, Desh Bhagat College, Bardwal, Dhuri, Punjab, India

ABSTRACT

The phrase "computational chemistry" is typically used when a mathematical technique is sufficiently developed to be automated for use on a full computer. Computational chemistry is the application of chemical, mathematical, and computing skills to the resolution of intriguing chemical problems, Information like molecular characteristics or simulated experimental outcomes are generated using computers. A qualitative or approximative quantitative computational system has been used to describe extremely basic parts of chemistry. The largest error computational chemists make is assuming the precision of any computed quantity. Similar to how not all spectra are properly resolved, a qualitative or approximative calculation can frequently provide helpful insight into chemistry if you understand what it tells you and what it does not. Since it is considerably simpler to depict a molecule on a computer screen than it is to manufacture, purify, and characterize a molecule in a lab, computational chemistry enables researchers to investigate a broad, diverse range of chemical space.

KEYWORDS- Approximative, Computational, Precision, Quantum, Quantitative.

I. INTRODUCTION

While non-computational quantum chemistry focuses with the formulation of analytical expressions for the properties of molecules and their reactions, computational theoretical chemistry is largely concerned with the numerical computation of molecular electronic structures and molecular interactions. The majority of computational chemistry theory and computer programs are based on concepts from statistical physics, thermodynamics, quantum and classical

mechanics. This is because they utilize mathematics to model the atoms and molecules. Despite the fact that computational results typically supplement the knowledge gained from chemical experiments and can occasionally predict previously unobserved chemical phenomena, the quantum many body problem cannot be solved analytically, much less in closed form, with the exception of relatively recent results concerning the hydrogen molecular ion.

Fig. 1: **Schrodinger's Solution For Hydrogenic Atom**

$$\frac{-\hbar^2}{2\mu}\frac{1}{r^2\sin\theta}\left[\sin\theta\frac{\partial}{\partial r}\left(r^2\frac{\partial\Psi}{\partial r}\right)+\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Psi}{\partial\theta}\right)+\frac{1}{\sin\theta}\frac{\partial^2\Psi}{\partial\phi^2}\right]$$
$$+U(r)\Psi(r,\theta,\phi)=E\,\Psi(r,\theta,\phi)$$

Source: **Schrodinger's Solution for Hydrogenic Atom From Donald.A.Mcquarrie**

Computational techniques are widely used for the designing of the new drugs and materials like Computer-aided drug discovery (CADD) algorithms are employed for the fast design of new drugs. Properties like Cross-Section for Collisions in other particles, spectroscopic quantities, calculating Thermodynamic Quantities, Probability of finding electrons, Transitions in between Translational, Rotational and Vibrational States can be executed easily using algorithms.

II. Methods Including Computational Chemistry

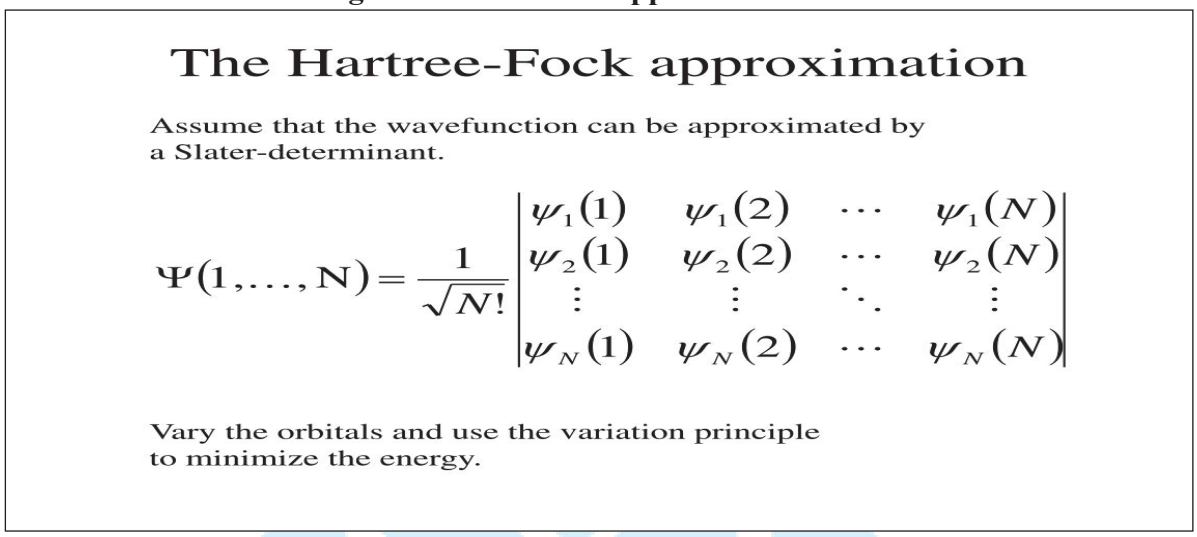
Computational ranges from high accuracy to low precision techniques and are explained as:

A. AB-Initio Method

Hartree-Fock calculations, whose primary approximation is known as the mean field approximation, are the most used kind of ab-initio computations. As a result, the computation includes the average effects of the columbic electron-electron repulsion even if it is not expressly taken into consideration. Due to the fact that this computation is variational, it is assumed that all of the estimated energies are greater than or equal to the actual energy. The basis set size employed determines how accurate the calculation is, although due to the mean field approximation, the energies from hartreefock calculations are always higher than the precise energies and tend to a limiting value known as the hartreefock limit as basis size increases.The precision of the computed results in the form used for the bases functions is also impacted by another problem. Of course, it is unknown what the single electronic molecule wave function actually looks like. The basis

function forms can give a more accurate or less accurate approximation to the exact numerical single electron solution of the HF equation. Atomic orbitals, which are nothing more than a linear collection of basis functions with coefficients derived from the relevant atomic hartreefock computations, are combined to generate molecular orbitals. Most Hartree-Fock calculations result in estimated energy that is more than the Hartree-Fock limit due to this approximation. A Hartreefock calculation is the first step in more complex calculations, after which correlations caused by electronic repulsion are corrected.

Fig. 2: **Hartree-Fock Approximation**



Source: https://en.wikipedia.org/wiki/Hartree%E2%80%93Fock_method

Quantum Monte Carlo (QMC) is another technique that prevents making the initial HF errors. Variational, diffusion, and green's functions are three types of QMC. These techniques utilize a numerical monte-carlo integration to evaluate integrals and operate on a wave function that is explicitly coupled. These calculations can take a long time, but they might produce incredibly accurate results.

Fig. 3: **Monte Carlo Algorithm**

Algorithm Monte Carlo with bounded variance

Require: A quantum algorithm A such that $Var(\nu(A)) \leq \sigma^2$ for some known σ, an accuracy ϵ such that $\epsilon < 4\sigma$.
Ensure: An estimate of $\mathbb{E}[\nu(A)]$.

1: Set $A' = A/\sigma$.
2: Run A' once and let $\widetilde{m}$ be the output.
3: Let B be the algorithm produced by executing A' and subtracting $\widetilde{m}$.
4: Apply algorithm 2 to algorithms $-B_{<0}/4$ and $B_{\geq 0}/4$ with accuracy $\epsilon/(32\sigma)$ and failure probability $1/9$, to produce estimates $\widetilde{\mu}^-$, $\widetilde{\mu}^+$ of $\mathbb{E}[\nu(-B_{<0}/4)]$ and $\mathbb{E}[\nu(B_{\geq 0}/4)]$, respectively.
5: Set $\widetilde{\mu} = \widetilde{m} - 4\widetilde{\mu}^- + 4\widetilde{\mu}^+$.
6: Output $\sigma\widetilde{\mu}$.

Source: https://quantumalgorithms.org/chap-montecarlo.html

The advantage of ab-initio approaches is that, after all approximations are reduced to a small enough size, they finally converge to the precise answer. This convergence is not monotonic, though. A more straightforward calculation may occasionally produce better results for a given property than a more complex calculation.

The drawback of ab-initio procedures is their high cost. These processes frequently need astronomical quantities of computer memory, CPU time, and disc space. A calculation that is twice as large will take sixteen times as long to perform using the HF approach since it scales as N, where N is the number of basis functions. Calculations involving correlations frequently scale very poorly.

In reality, only a few electrons in the molecule are required to produce incredibly precise solutions. Generally speaking, ab-intio calculations produce extremely good qualitative results, and as the size of the molecules in question decreases, they can produce increasingly precise quantitative answers.

B. Semi Empirical Calculations

The general framework of a semi-empirical calculation is the same as a hartreefock calculation. Certain details, such as the two electron integrals, are approximated or omitted within this framework. The procedure is parameterized by curve fitting in a few parameters or numbers to correct the inaccuracies produced by skipping certain portions of the calculation in order to provide the best possible agreement with experimental results.Semi-empirical calculations have the

advantage of being substantially faster than ab-initio computations. Semi-empirical computations have the drawback that their outcomes can be unpredictable. The results might be excellent if the molecule being studied is comparable to molecules in the data base used to parameterize the approach. The answers might be inaccurate if this molecule differs greatly from anything in the parameterization set. Both organic and inorganic chemistry have successfully used semi-empirical computations.

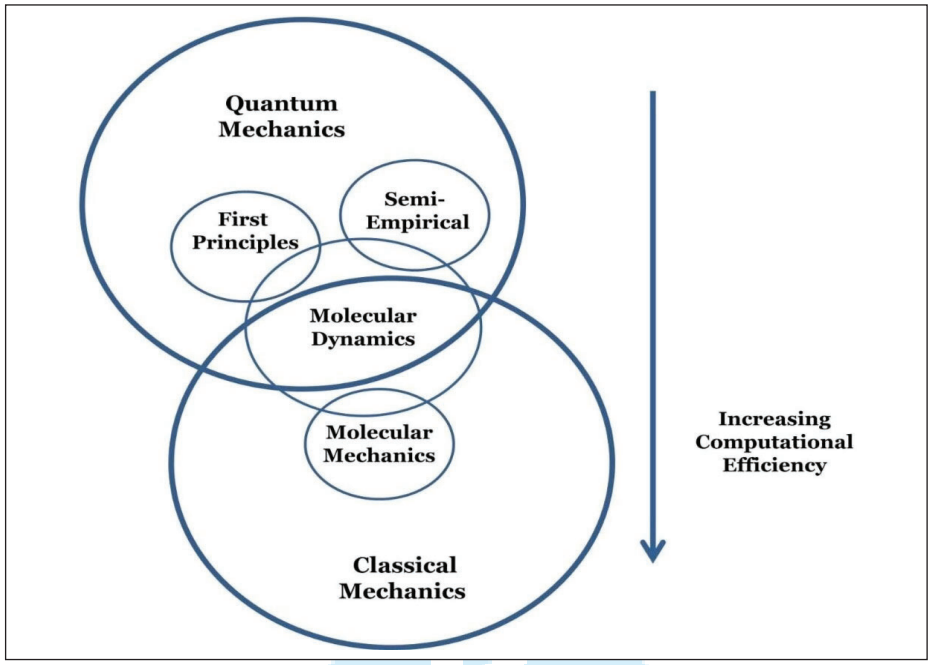
C. Molecular Mechanics Approach

Even if a molecule is too large to be treated semi-empirically, its behavior can still be modeled by avoiding quantum mechanics. This is accomplished by creating a straightforward expression for the potential energy function, which represents the molecular force field. The database of substances used to parameterize a molecular mechanics approach is essential to the method's success. A molecular mechanics approach may be parameterized against a particular class of molecules, such as proteins, as opposed to a collection of organic molecules for a semiempirical method. Only other proteins would be anticipated to be relevant to such a force field.

The advantage of molecular mechanics is that it makes it possible to simulate large molecules, such proteins and DNA segments, making it the main tool used by computational biochemists.

The drawback is that certain chemical characteristics, such electronic excited states, are not even characterized by the approach. The most potent and user-friendly visual interfaces are frequently found in molecular mechanics software packages, which are designed to operate with incredibly vast and complex systems.

Fig. 4: Molecular Dynamics Approach



Source: https://www.gloriabazargan.com/blog/comp-chem-methods

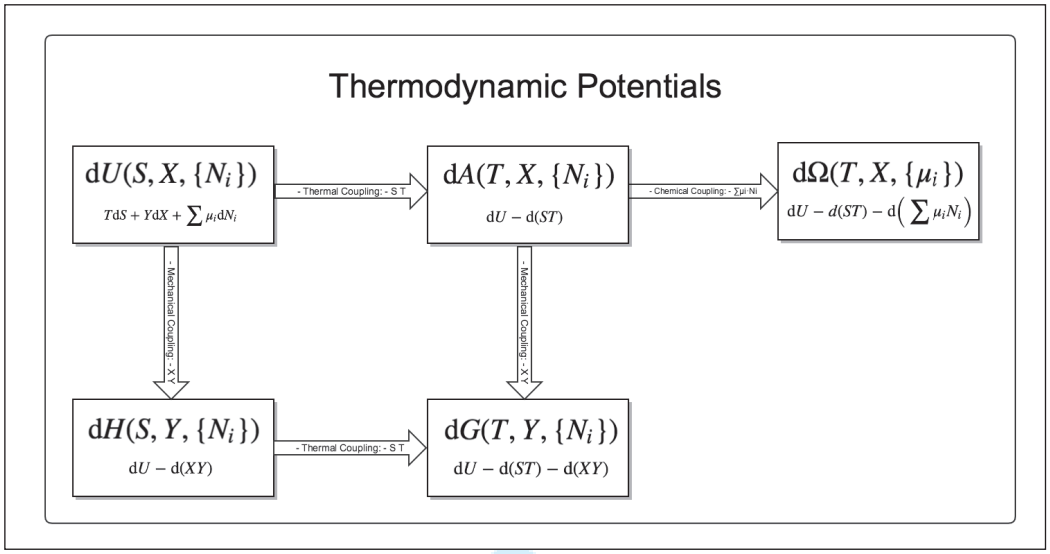
D. Molecular Dynamics Approach

The determination of parameters like diffusion coefficients or radical distribution functions for use in statistical mechanical treatments is made possible by the application of molecular dynamics to solvent/solute systems. A number of molecules are often assigned an initial position and velocity in a solvent/solute calculation. To achieve Equilibrium and provide a solid statistical description of the radial distribution function, new positions are generated based on this movement a short while later. This process is repeatedthousandsoftimes.

E. Thermodynamics And Statistical Mechanics

From a molecular description of a material, it is possible to mathematically extrapolate the thermodynamic properties of bulk materials. Calculations for statistical mechanics are frequently added at the end of ab-initio calculations for gas phase properties. In order to do a computational experiment for condensed phase properties, molecular dynamics calculations are frequently required. One of the most well-developed physical theories is thermodynamics, which provides a solid theoretical foundation for studying molecular systems.

Fig. 5: Thermodynamic Potential

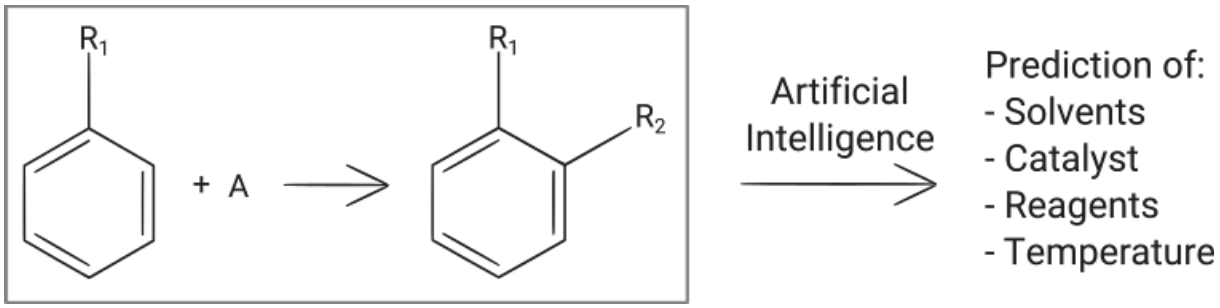


Source: https://statisticalphysics.leima.is/thermodynamics/summary.html

F. Artificial Intellegence (AI)

In recent years, techniques developed by computer scientists that are interested in artificial intelligence have mostly been used in the drug creation process. In most cases, a functional site has been found, and it is intended to develop a molecule with a structure that would interact with that site in order to impair its activity. Instead of having a chemist use a molecular mechanics software to test hundreds or thousands of possibilities, the molecular mechanics is integrated into an artificial intelligence computer that automatically tests a huge number of plausible alternatives. Artificial intelligence tends to increase limits for the chemists by solving many time consuming problems which can be completed experimentally.

Fig. 6: **Artifical Intelligence**



Source: Drawn through Chem Draw

III. APPLICABILITY

Pharmaceutical medications, packaging materials, batteries, and other breakthroughs are being driven by computational chemistry techniques and molecular modeling in a variety of industries. Computational chemistry has several uses, such as:

A. Designing Drug

For more than three decades, the generation of therapeutically significant small molecules has been greatly helped by computer-aided drug discovery and design techniques. These techniques can be roughly categorized as structure-based or ligand-based techniques. In theory, high-throughput screening and structure-based approaches are similar in that both target and ligand structure information are relevant. Pharmacophore, ligand design, and ligand docking are examples of structure-based approaches.

B. Chemoinformatics

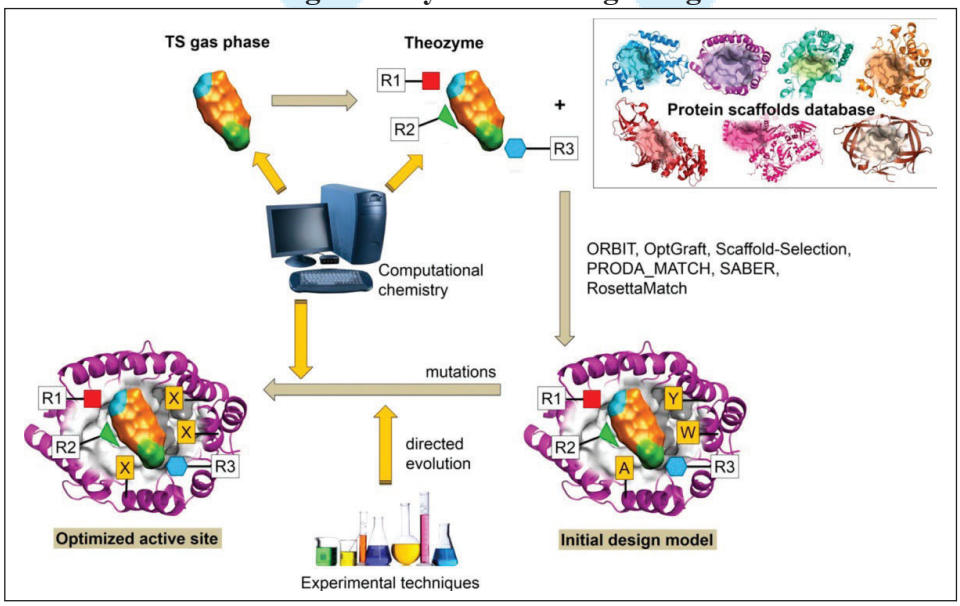
Chemoinformatics, also known as cheminformatics, is a term that has just lately been used to refer to a field that organizes and coordinates the use of computers in chemistry. This word is relatively new, despite the fact that computers have long helped chemists. Therefore, it is understandable that not all chemists are awed by this discovery. Actually, there are some disagreements on the necessity of the establishment of this relatively new discipline of chemistry.Utilizing readily available ligand resources including pharmacophore modeling, quantitative structure-activity relationship (QSAR), docking, and molecular dynamics (MD) simulations, chemoinformatics uses an integrated strategy to research and comprehend the function of chemical systems. The major methods for designing new ligands based on descriptor calculations and functional group

replacements are pharmacophore modeling and QSAR investigations. Using docking or virtual screening, the affinities of these recently created or well-known ligands for their respective targets can be predicted. However, it can be difficult to identify near-native binding and accurate grading, which requires worldwide discussion. However, to predict the dynamics of receptors or complexes, MD simulations are used to determine the atomic behavior of proteins and biomaterials in a system.As a result, this gives a general review of ligand databases for chemoinformatics, as well as of modern methods, tools for pharmacophore modeling, QSAR, docking, and MD simulations. The foundation of drug design is built on these chemoinformatics techniques, which can also help advance our understanding of chemical processes.

C. Enzyme Design

The "inside-out" method of enzyme engineering combines recent advances in computational chemistry and biology. Proteins were created to catalyze reactions in nature that had not before been accelerated. The success rate is still modest even if some of these proteins fold and serve as catalysts. In comparison to alternative protein engineering methods, the current technology's strengths and weaknesses are addressed. Although computational "inside-out" design can produce selective and catalytically active proteins on its own, natural enzymes still outperform these proteins in terms of kinetic performance. However, computational designs can be considerably enhanced in terms of binding, turnover, and thermal stability when integrated with directed evolution, molecular dynamics simulations, and crowd-sourced structure-prediction methods.

Fig. 6: Enzyme and Drug Design



Source: https://www.sciencedirect.com/science/article/abs/pii/S0223523421005547

D. Polymer Design

The field of polymer science and engineering benefits greatly from the use of molecular modeling and simulations. These computational methods allow for predictions and give explanations for macromolecular structure, dynamics, thermodynamics, and microscopic and macroscopic material qualities that have been experimentally seen. With recent improvements in computational power, polymer simulations can help with the design and discovery of in vitro macromolecular materials in a synergistic manner. Care must be taken to verify the validity and reproducibility of these simulations if significant results are to be achieved and this technology's expanding power is to be properly leveraged. With these factors in mind, we go over our philosophy for carefully creating or choosing the best models, running, and evaluating polymer simulations in this perspective.

E. Semiconductors

For a number of electrical devices, including organic photovoltaics and field effect transistors, organic semiconductors hold promise as an alternative to their inorganic counterparts because they can be solution-processed, and the possibility of roll-to-roll manufacturing makes this a more affordable option. Synthesizing novel materials and creating more effective devices have been the main focuses of the effort to create these technologies. Important advancements in computational chemistry have been made in the predictive models that are utilized to comprehend the basic characteristics of these semiconductors. Combining these initiatives holds the prospect of improving our knowledge of the requirements for more effective organic electronics.

IV. CONCLUSION

Computational chemistry, to put it simply, is a discipline of chemistry that produces data to supplement experimental data on the structures, characteristics, and reactions of substances. The calculations are based on semi-empirical structure-property relationships, theories of symbolic computations, and artificial intelligence, as well as quantum and classical physics, molecular dynamics, statistical theory, and thermodynamics.

In computational chemistry, various computations based on quantum and classical physics are used to model and predict molecular structures and properties. The effectiveness of computational chemistry is also being improved by advances in machine learning, which accelerate the calculation process. The time, money, and reagent resources required for synthesis, tests, and other experimental tasks are decreased using computational chemistry methods. By accelerating

complicated computations by multiple orders of magnitude, machine learning applications can significantly improve computational chemistry. Digital chemical design can quickly surpass wet lab design by carefully integrating machine learning with physics-based methods. Cost savings are a direct result of this time savings. Additionally, these techniques enable a deeper exploration of the chemical universe, increasing the possibility of discovering unique, unexpected compounds. The increased speed and breadth made possible by digital chemistry boosts the likelihood of acquiring intellectual property in the quick-paced field of molecular design, where the first to file a patent might mean the difference between success and the termination of a research program.

V. REFERENCES

I. Allouche, A. R. (2011). Gabedit—A graphical user interface for computational chemistry softwares. *Journal of computational chemistry*, *32*(1), 174-182.

II. Bernholdt, D. E., Apra, E., Früchtl, H. A., Guest, M. F., Harrison, R. J., Kendall, R. A., ... & Nieplocha, J. (1995). Parallel computational chemistry made easier: The development of NWChem. *International Journal of Quantum Chemistry*, *56*(S29), 475-483.

III. Cramer, C. J., & Bickelhaupt, F. M. (2003). Essentials of computational chemistry. *Angewandte Chemie-International Edition in English*, *42*(4), 381-381.

IV. Fernández, I., & Cossío, F. P. (2014). Applied computational chemistry. *Chemical Society Reviews*, *43*(14), 4906-4908.

V. Gloria B. (2018). An Overview of Computational Methods for Molecular Modelling. *International Reviews in Physical Chemistry, 37(1),1-82*

VI. Goh, G. B., Hodas, N. O., & Vishnu, A. (2017). Deep learning for computational chemistry. *Journal of computational chemistry*, *38*(16), 1291-1307.

VII. Grimme, S., & Schreiner, P. R. (2017). Computational Chemistry: The Fate of Current Methods and Future Challenges. *Angewandte Chemie*, *57*(16), 4170-4176.

VIII. Havel, T. F. (1998). Distance geometry: Theory, algorithms, and chemical applications. *Encyclopedia of Computational Chemistry*, *120*, 723-742.

IX. Keith, J. A., Vassilev-Galindo, V., Cheng, B., Chmiela, S., Gastegger, M., Müller, K. R., & Tkatchenko, A. (2021). Combining machine learning and computational chemistry for predictive insights into chemical systems. *Chemical reviews*, *121*(16), 9816-9872.

X. Kendall, R. A., Aprà, E., Bernholdt, D. E., Bylaska, E. J., Dupuis, M., Fann, G. I., ... & Wong, A. T. (2000). High performance computational chemistry: An overview of

NWChem a distributed parallel application. *Computer Physics Communications*, *128*(1-2), 260-283.

XI. Kowalski, K., Bair, R., Bauman, N. P., Boschen, J. S., Bylaska, E. J., Daily, J.,& Windus, T. L. (2021). From NWChem to NWChemEx: Evolving with the computational chemistry landscape. *Chemical reviews*, *121*(8), 4962-4998.

XII. Manly, C. J., Chandrasekhar, J., Ochterski, J. W., Hammer, J. D., & Warfield, B. B. (2008). Strategies and tactics for optimizing the Hit-to-Lead process and beyond—A computational chemistry perspective. *Drug Discovery Today*, *13*(3-4), 99-109.

XIII. Nieplocha, J., Foster, I., & Kendall, R. A. (1998). ChemIO: High performance parallel I/O for computational chemistry applications. *The International Journal of High Performance Computing Applications*, *12*(3), 345-363.

XIV. O'boyle, N. M., Tenderholt, A. L., & Langner, K. M. (2008). Cclib: a library for package-independent computational chemistry algorithms. *Journal of computational chemistry*, *29*(5), 839-845.

XV. Onodera, T., Morita, Y., Suzuki, A., Koyama, M., Tsuboi, H., Hatakeyama N.& Miyamoto, A. (2009). A computational chemistry study on friction of h-MoS2. Part I. Mechanism of single sheet lubrication. *The Journal of Physical Chemistry B*, *113*(52), 16526-16536.

XVI. Ren, X., Rinke, P., Joas, C., & Scheffler, M. (2012). Random-phase approximation and its applications in computational chemistry and materials science. *Journal of Materials Science*, *47*, 7447-7471.

XVII. Wei, L., Xu, M., Liu, Z., Jiang, C., Lin, X., Hu, Y., & Zhang, P. (2023). Hit Identification Driven by Combining Artificial Intelligence and Computational Chemistry Methods: A PI5P4HuK-β Case Study. *Journal of Chemical Information and Modeling*, *63*(16), 5341-5355.

XVIII. Zhang, X., Chung, L. W., & Wu, Y. D. (2016). New mechanistic insights on the selectivity of transition-metal-catalyzed organic reactions: the role of computational chemistry. *Accounts of chemical research*, *49*(6), 1302-1310.