Advances in Energy and Materials Research https://jaem.qom.ac.ir/

Computational Approaches in Natural Product Research: Advances, Challenges, and Future Directions

Hasan Zandi[✉1] and Reza Safari[2]

1. Department of Chemistry, Faculty of Science, University of Qom, Qom, Iran. Email: h.zandi@qom.ac.ir
2. Department of Chemistry, Faculty of Science, University of Qom, Qom, Iran. Email: r.safari@qom.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 25 Dec 2023
Received in revised form 29 Feb 2024
Accepted 12 Apr 2024
Published online 24 Jun 2024

Keywords:
Natural products, Computational methods, Structural characterization, Activity prediction, Drug discovery.

ABSTRACT

Natural products offer immense potential for drug discovery, but their structural complexity and diverse bioactivities pose significant challenges. This review highlights the pivotal role of computational methods in addressing these challenges. We explore techniques for structural characterization, including DFT, molecular dynamics, and computational spectroscopy, which provide detailed insights into molecular properties and enable accurate structure elucidation. For activity prediction, molecular docking and QSAR modeling are discussed, emphasizing their utility in virtual screening and lead optimization. The integration of computational and experimental approaches is crucial for efficient drug discovery, with high-throughput virtual screening emerging as a powerful strategy. Despite advancements, challenges such as predicting complex structures and accurately estimating activity remain. Future directions include incorporating multi-omics data, exploring vast chemical spaces, and developing atomic-scale computational methods like QTAIM for a deeper understanding of molecular properties. By combining computational and experimental expertise, we can unlock the full potential of natural products for therapeutic and other applications.

Cite this article: Zandi, H. & Safari, R. (2024). Computational Approaches in Natural Product Research: Advances, Challenges, and Future Directions, *Advances in Energy and Materials Research*, 1 (2), 1-6. https://doi.org/10.22091/jaem.2024.11133.1016

DOI: https://doi.org/10.22091/jaem.2024.11133.1016
Publisher: University of Qom.

1. Introduction

Natural products, derived from plants, microorganisms, and other biological sources, have been the foundation of traditional medicine and continue to inspire modern drug discovery [1, 2]. The complex structures and diverse bioactivities of these compounds present both opportunities and challenges for researchers [3]. In recent years, computational tools have emerged as powerful allies in natural product research, enabling scientists to explore these molecules at unprecedented levels of detail [4].

Natural products encompass a wide range of compounds, from simple secondary metabolites to complex macromolecules [5]. This structural and functional diversity makes them invaluable sources for new drug leads, bioactive materials, and industrial applications [6]. However, this same diversity poses significant challenges in terms of characterization, activity prediction, and optimization [7].

Computational methods, with their ability to model and predict molecular properties, play a vital role in overcoming these challenges [8]. These methods not only aid in better understanding the structure and function of natural products but also accelerate the process of discovering and developing new compounds [9].

In this mini review, we provide a comprehensive examination of advanced computational methods in natural product research. We first explore the methods used for structural characterization, then delve into techniques for activity prediction. We also discuss how these methods interface with experimental techniques and examine current challenges and future prospects.

2. Computational Methods for Structural Characterization

2.1. Density Functional Theory (DFT)

Density Functional Theory has become the method of choice for investigating the electronic structure of natural products (Fig. 1) [10]. It offers a balance between accuracy and computational cost, making it suitable for molecules of varying sizes. Figure 1 provides an overview of various computational chemistry methods used to determine molecular electronic structures. It covers methods from past, present, and future predictions, illustrating the system sizes computable within a day using a single-core CPU. Methods include force-field (FF), semi-empirical (SMO), Hartree-Fock (HF), configuration interaction (CI), and modern approximations like DMRG and stochastic Monte-Carlo. The predicted capabilities for 2043 are based on conservative estimates of computing power growth. These estimates depend significantly on molecular structure and computational resources.

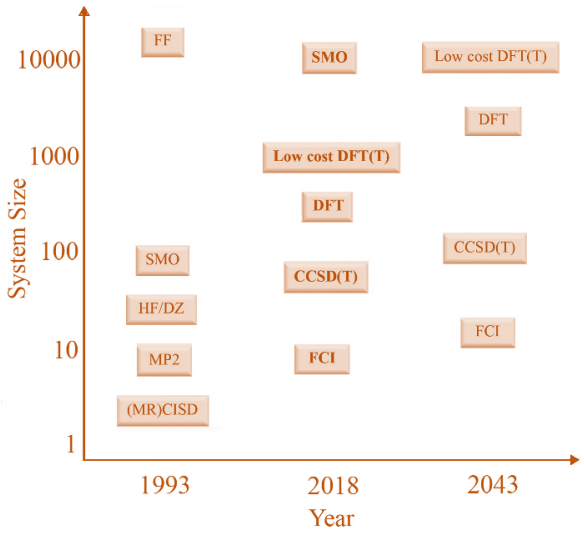

Figure 1: Molecular Structures and Electronic Features

Advantages

- Provides detailed information on molecular geometry and electronic properties
- Can predict spectroscopic data (NMR, IR, UV-Vis) with high accuracy
- Enables the study of transition states and reaction mechanisms

Limitations

- Computationally intensive for large molecules
- Choice of functional can significantly affect results
- May require additional corrections to describe weak interactions like van der Waals forces

Case Study

A recent study used DFT to elucidate the structure-activity relationship of curcumin derivatives, leading to the design of more potent anti-inflammatory compounds [11]. Researchers were able to predict the effect of small structural changes on electronic distribution and, consequently, biological activity.

2.2. Molecular Dynamics (MD) Simulations

MD simulations offer insights into the dynamic behavior of natural products, especially in biological environments [12].

Advantages

- Can simulate interactions with solvents and biomolecules
- Provides information on conformational changes and flexibility
- Allows for the study of temperature and pressure effects on structure and function

Limitations

- Accuracy depends on the force field used
- Limited to relatively short time scales
- Requires significant computational resources for large systems

Case Study

MD simulations revealed the mechanism by which epigallocatechin gallate (EGCG) from green tea interacts with lipid membranes, explaining its potential health benefits [13]. This study showed how EGCG can penetrate and alter the structure of cell membranes, which may explain some of its anti-cancer effects.

2.3. Computational Spectroscopy

Computational spectroscopy, often based on DFT calculations, is a powerful tool for predicting and interpreting experimental spectra of natural products [14].

Advantages

- Aids in structure determination of unknown compounds
- Predicts NMR, IR, and Raman spectra
- Facilitates the interpretation of experimental data

Limitations

- Accuracy depends on the chosen theoretical level and basis set
- Calculations can be very time-consuming for large molecules
- May require empirical corrections to improve accuracy

Case Study

In a recent study, computational NMR spectroscopy was used to confirm the structure of a novel alkaloid isolated from a plant species [15]. Comparison of computed and experimental spectra allowed researchers to determine the correct structure from several possible isomers.

3. Computational Methods for Activity Prediction

3.1. Molecular Docking

Docking simulations predict the binding modes and affinities of natural products to biological targets [16].

Advantages

- Rapid screening of large compound libraries
- Can guide structure-based drug design
- Provides insights into molecular mechanisms of biological activity

Limitations

- Simplified treatment of protein flexibility and solvation effects
- May miss non-conventional binding modes
- Accuracy of scoring functions can be limited

Case Study

Virtual screening of a natural product library using molecular docking led to the discovery of novel inhibitors of the SARS-CoV-2 main protease [17]. This study demonstrated how computational methods can accelerate the drug discovery process in emergency situations like the COVID-19 pandemic.

3.2. Quantitative Structure-Activity Relationship (QSAR)

QSAR models relate molecular descriptors to biological activity, enabling activity prediction for novel compounds [18].

Advantages

- Can predict activities for untested compounds
- Useful for lead optimization
- Integrates diverse chemical and biological data

Limitations

- Requires high-quality experimental data for model building
- May not capture complex, non-linear relationships
- Limited extrapolation to compounds very different from the training set

Case Study

A QSAR model developed for antimalarial compounds from traditional Chinese medicine helped identify promising leads for further development [19]. The model incorporated both structural features and predicted pharmacokinetic properties to improve its predictive power.

4. Integration of Computational and Experimental Approaches

While computational methods offer numerous advantages, their true power emerges when combined with experimental techniques [20]. This integration can lead to more efficient and effective research strategies in natural product discovery and development.

4.1. Synergistic Workflows

- Computational predictions can guide experimental design, reducing the number of compounds that need to be synthesized and tested [21].
- Experimental data can validate and refine computational models, improving their accuracy and predictive power [22].
- Iterative cycles of computation and experimentation can accelerate the drug discovery process [23].

4.2. High-Throughput Virtual Screening

Combining computational screening with high-throughput experimental assays allows for rapid identification of promising natural product leads [24].

5. Challenges and Future Directions

5.1. Dealing with the complexity of natural products

Many natural products have complex structures with multiple stereogenic centers. Developing methods to accurately predict the 3D structures and properties of these compounds remains a challenge [25].

5.2. Improving the accuracy of activity predictions

While current methods can often identify active compounds, accurately predicting potency and selectivity remains difficult. Machine learning approaches, particularly deep learning, show promise in this area [26].

5.3. Integrating multi-omics data

As we gather more data on the systems-level effects of natural products, developing computational methods to integrate and interpret this information will be crucial [27].

5.4. Exploring the chemical space of natural products

Computational methods can help us explore the vast chemical space of potential natural products, guiding the discovery of new compounds with desired properties [28-31].

6. Computational method at *atomic-scale*: QTAIM theory

The intramolecular (at *atomic scale*) charge and energy transfer in molecular systems (such as nanoelectronic devices) can be studied using quantum theory of atoms in molecule (QTAIM), which is a generalization of quantum mechanics to proper open systems, describing open systems in terms of the topology of the electron density $\rho(r)$ [32-35]. The QTAIM can thus be utilized to determine chemical behavior and reactivity of the quantum systems. The partitioning of the molecular space into *atomic basins* (Ω) can be used to partition the overall electronic properties into atomic contributions systematically. Also, based on the QTAIM, the atomic electronic energy, $E_{elec}(\Omega)$, is given by

$$E_{elec}(\Omega) = V_{elec}(\Omega) + K_{elec}(\Omega) \quad (1)$$

where $V_{elec}(\Omega)$ and $K_{elec}(\Omega)$ are the total *atomic* potential and kinetic energies, respectively. In addition, the QTAIM remains equally valid in the presence of external electric field (EF), and thus QTAIM allows also to determine the extent of transferability of the field-induced atomic properties and to obtain an understanding of the physical factors governing their values in given situations [35-36].

7. Conclusion

Computational approaches (at molecular or atomic scale) have become indispensable tools in natural product research, offering insights that complement and extend experimental techniques. As these methods continue to evolve, they promise to accelerate the discovery and development of novel compounds for various applications, from drug discovery to materials science. The future of natural product research lies in the seamless integration of computational and experimental approaches, leveraging the strengths of each to unlock the full potential of nature's chemical diversity.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgment

There is no acknowledgment for this manuscript.

References

[1] Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* 2020, *83* (3), 770-803. https://doi.org/10.1021/acs.jnatprod.9b01285

[2] Mark S. Butler, The Role of Natural Product Chemistry in Drug Discovery. *Journal of Natural Products*. 2004, https://doi.org/10.1021/np040106y.

[3] Stratton, C. F.; Newman, D. J.; Tan, D. S. Cheminformatic comparison of approved drugs

from natural product versus synthetic origins. *Bioorg. Med. Chem. Lett.* 2015, *25* (21), 4802-4807. https://doi.org/10.1016/j.bmcl.2015.09.019

[4] Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat. Chem.* 2016, *8*(6), 531-541. https://doi.org/10.1038/nchem.2479

[5] Katz, L.; Baltz, R. H. Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* 2016, *43* (2-3), 155-176. https://doi.org/10.1007/s10295-015-1682-x

[6] Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* 2015, *14*(2), 111-129. https://doi.org/10.1038/nrd4510

[7] Saldívar-González, F. I.; García-Rodríguez, J.; Naveja, J. J.; Sánchez-Cruz, N.; Medina-Franco, J. L. Computational approaches to identify natural products for drug discovery. *Expert Opin. Drug Discov.* 2019, *14*(9), 913-927. https://doi.org/10.1080/17460441.2019.1622106

[8] Medina-Franco, J. L. Advances in computational approaches for drug discovery based on natural products. *Rev. Latinoam. Quím.* 2013, *41*(2), 95-110. https://doi.org/10.2174/092986712799320511

[9] Rodrigues, T. Harnessing the potential of natural products in drug discovery from a cheminformatics vantage point. *Org. Biomol. Chem.* 2017, *15*(44), 9275-9282. https://doi.org/10.1039/C7OB02193C

[10] Grimme, S.; Schreiner, P. R. Computational chemistry: the fate of current methods and future challenges. *Angew. Chem., Int. Ed.* 2018, *57* (16), 4170-4176. https://doi.org/10.1002/anie.201709943.

[11] Priyadarsini, K. I. The chemistry of curcumin: from extraction to therapeutic agent. *Molecules* 2014, *19* (12), 20091–20112. https://doi.org/10.3390/molecules191220091

[12] Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpí, J. L. Molecular dynamics simulations: advances and applications. Adv. Appl. Bioinform. Chem. 2015, 8, 37-47. https://doi.org/10.2147/AABC.S70333.

[13] Sirk, T. W.; Wang, Y.; McCormick, S.; Pogorzelski, W. Molecular dynamics study of the effect of EGCG on the structure of phospholipid membranes. *J. Phys. Chem. B* 2008, *112* (30), 9176-9184. https://doi.org/10.1021/jp802250e

[14] Tantillo, D. J. Walking in the woods with quantum chemistry—applications of quantum chemical calculations in natural products research. *Nat. Prod. Rep.* 2013, *30* (8), 1079-1086. https://doi.org/10.1039/C3NP70028C.

[15] Lodewyk, M. W.; Siebert, M. R.; Tantillo, D. J. Computational prediction of ^{1}H and ^{13}C chemical shifts: a useful tool for natural product, mechanistic, and synthetic organic chemistry. *Chem. Rev.* 2012, *112* (3), 1839-1862. https://doi.org/10.1021/cr200106v

[16] Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput.-Aided Drug Des.* 2011, *7* (2), 146-157. https://doi.org/10.2174/157340911795677602.

[17] Wu, C., Liu, X., Wang, Z., Xu, J., Yu, D., Chen, J., Zhang, L., Li, Y., Yang, J., & Wu, J. Analysis of therapeutic targets for SARS-CoV-2 and discovery of potential drugs by computational methods. *Acta Pharmaceutica Sinica B*, 2020 10(5), 766-788. https://doi.org/10.1016/j.apsb.2020.02.008

[18] Cherkasov, A.; Caffrey, D. R.; Nair, V.; Horvath, D.; Ritchie, T. J.; Tetko, I. V.; Varnek, A.; Wang, S.; Zholtkevich, A.; Zinevich, Y. QSAR modeling: where have you been? Where are you going to? *J. Med. Chem.* 2014, *57* (12), 4977-5010. https://doi.org/10.1021/jm500290n

[19] Xu, Y.; Zhang, Z.; Xu, H.; Zhang, Y.; Wu, L. QSAR models for the prediction of antimalarial activities of artemisinin analogues using calculated molecular descriptors. *Chemometrics Intell. Lab. Syst.* 2017, *162*, 99-108. https://doi.org/10.1016/j.chemolab.2017.01.006

[20] Medina-Franco, J. L.; Forero, S.; Pérez-García, S.; López-Vallejo, F.; Alí-Benali, M.; N. S.; Hernández-Gallego, J.; Gómez-Mendoza, J. S.; Martínez-Mayorga, K. Integrating virtual screening in the drug discovery process. *Drug Discov. Today* 2021, *26* (3), 886-901. https://doi.org/10.1016/j.drudis.2020.12.018

[21] Bielska, E.; Tymecka, D.; Wrzesiński, T.; Sołtys, M.; Tykarski, J.; Płotka-Wasylka, J.; Namieśnik, J. Virtual screening strategies in drug design–methods and applications. *J. Biotechnol.* 2011, *92* (3), 249-264. https://doi.org/10.1016/j.jbiotec.2011.01.023

[22] Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W. Computational methods in drug discovery. *Pharmacol. Rev.* 2014, *66* (1), 334-395. https://doi.org/10.1124/pr.112.007360

[23] Jorgensen, W. L. The many roles of computation in drug discovery. *Science* 2004, *303* (5665), 1813-1818. https://doi.org/10.1126/science.1090937

[24] Pereira, D. A.; Williams, J. A. Origin and evolution of high throughput screening. *Br. J. Pharmacol.* 2007, *152* (1), 53-61. https://doi.org/10.1038/sj.bjp.0707376

[25] Chen, Y.; Li, L.; Yang, M.; Xu, W.; Wang, L. Advances in high-throughput experimental methods for natural products drug discovery. *Nat. Prod. Rep.* 2021, *38* (7), 1298-1315. https://doi.org/10.1039/D1NP00041F

[26] Vamathevan, J.; Clark, D. E.; Czodrowski, P.; Dunham, N.; Gaulton, A.; Hersey, A.; Kettunen, M.; McGovern, K.; Overington, J. P.; Paolini, G. V.; et al. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* 2019, *18* (6), 463-477. https://doi.org/10.1038/s41573-019-0024-5

[27] Wishart, D. S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* 2016, *15* (7), 473-484. https://doi.org/10.1038/nrd.2016.47

[28] Reymond, J. L. The chemical space project. *Acc. Chem. Res.* 2015, *48* (3), 722-730. https://doi.org/10.1021/ar500158w

[29] Schneider, G. Automating drug discovery. *Nat. Rev. Drug Discov.* 2018, *17* (2), 97-113. https://doi.org/10.1038/nrd.2017.213

[30] Baell, J.; Walters, M. A. Chemistry: Chemical con artists foil drug discovery. *Nature* 2014, *513* (7519), 481. https://doi.org/10.1038/513481a

[31] Zandi, H.; Harismah, K. Computer-Based Tools for Structural Characterizations and Activity Specifications of Natural Products: *A Quick Review*. *Lab-in-Silico* 2021, 2 (1), 50-54. https://doi.org/10.22034/labinsilico2102105

[32] Bader, R. F. W., Atoms in Molecules, Oxford University, U.K., 1995.

[33] Popelier, P., Atoms in Molecules, Pearson, U.K., 2000.

[34] Matta, C. F., Quantum Biochemistry, Wiley, Weinheim, 2010.

[35] Safari, R. and Sabzyan, H., Local energy dissipation/transition in field effect molecular nanoelectronic systems: a quantum mechanical methodology. *Communications in Theoretical Physics*, 2019, 71(4), 441. https://doi.org/10.1088/0253-6102/71/4/441

[36] Safari, R. and Sabzyan, H., Detailed mapping of intramolecular energy transfer in field-effect single-molecule nanoelectronic devices. Journal of the Iranian Chemical Society, 2024, 11, .1513-1532. https://doi.org/10.1007/s13738-024-0241-5