https://doi.org/10.48047/AFJBS.6.11.2024.1851-1861

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

Demystifying Molecules: Unveiling the Power of AI in Computational Chemistry

Parinita Tripathy[1],Banchhanidhi Dash[2],Sanyogita Shahi[1*]

1. Kalinga University, Raipur, Chhattisgarh, India, 492101
2. KIIT University, Bhubaneswar, Odisha, India, 751024

*Corresponding Author: drsanyogitashahi@gmail.com

Volume 6, Issue 11, May 2024

Received: 15 Apr 2024

Accepted: 05 May 2024

Published: 25 May 2024

doi: 10.48047/AFJBS.6.11.2024.1851-1861

Abstract:
AI techniques, including machine learning algorithms, neural networks, and deep learning architectures, have significantly enhanced the accuracy and efficiency of molecular structure prediction, quantum chemistry calculations, and molecular dynamics simulations. In drug discovery, AI facilitates virtual screening, de novo drug design, and pharmacophore modeling, accelerating the identification and optimization of novel therapeutics. Furthermore, AI-driven approaches enable predictive synthesis planning, reaction outcome prediction, and property estimation with unprecedented precision, thereby streamlining synthetic route design and optimization of molecular properties for diverse applications. Methodological advancements encompass a spectrum of techniques, from traditional machine learning models like support vector machines and random forests to state-of-the-art deep learning models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and graph neural networks (GNNs). These methodologies are complemented by hybrid approaches and ensemble methods that leverage the strengths of different AI paradigms to address complex chemistry-related challenges effectively. Despite these achievements, challenges remain, including the need for high-quality and diverse datasets, interpretation of AI-driven predictions, and ensuring model generalizability beyond training data. Addressing these challenges is crucial for advancing the reliability and applicability of AI in computational chemistry. Looking ahead, the integration of AI with experimental chemistry holds promise for accelerating discovery processes and optimizing experimental design. Ethical considerations and societal impacts of AI applications in chemistry also warrant careful deliberation as the field continues to evolve.

Key Words: Computational Chemistry, Molecular Modelling, Quantum Chemistry, Pharmacophore Modelling, De Novo Design.

I. Introduction

Artificial Intelligence (AI) has ushered in a new era of innovation in computational chemistry, revolutionizing the field by augmenting traditional methods with unprecedented computational power and predictive capabilities. This introduction explores how AI has transformed the landscape of chemistry research, enabling more efficient exploration and understanding of molecular structures, reactions, and properties.

1. AI's Role in Transforming Computational Chemistry

Artificial Intelligence (AI) has emerged as a pivotal force in revolutionizing computational chemistry, fundamentally altering how chemical research is conducted and accelerating discovery processes. By leveraging AI techniques, computational chemists can harness vast amounts of data to predict molecular

structures, properties, reactions, and behaviours with unprecedented accuracy and efficiency. AI tools in computational chemistry encompass a wide array of methodologies, including machine learning algorithms, neural networks, and deep learning architectures. These techniques enable researchers to tackle complex problems that were once computationally prohibitive or required extensive empirical testing. For instance, AI-driven molecular modelling techniques facilitate the rapid optimization of molecular structures and the prediction of their properties, aiding in drug discovery and material science.

2. Historical Perspective and Evolution of AI Tools in Chemistry Research:

The integration of AI into chemistry research has evolved significantly over the decades. Initially, AI applications in chemistry focused on rule-based systems and expert systems for interpreting spectroscopic data and predicting chemical reactions. These early tools laid the foundation for more sophisticated AI approaches, such as statistical modelling and machine learning, which began to gain traction in the late 20th century. The advent of big data and computational resources in the 21st century has propelled AI to new heights in chemistry. Modern AI techniques can handle vast datasets generated from experiments, simulations, and literature mining, allowing for more accurate predictions and deeper insights into chemical phenomena. This evolution has led to breakthroughs in areas like drug design, materials discovery, catalysis, and environmental chemistry, among others. Overall, the historical development and ongoing evolution of AI tools in chemistry underscore their transformative impact on computational approaches, paving the way for innovative solutions to complex chemical challenges and opening new avenues for interdisciplinary research and collaboration.

II. Literature Review

1. Applications of AI in Molecular Modelling: Structure Prediction and Optimization:

AI has significantly advanced the field of structure prediction and optimization in computational chemistry. Traditional methods often relied on labor-intensive trial-and-error approaches or simplified models that may not accurately reflect real-world conditions. AI, particularly machine learning algorithms and neural networks has enabled more accurate predictions of molecular structures and their properties.

Machine learning models can analyze large datasets of molecular structures and properties, learning patterns, and correlations that help predict optimal configurations. This capability is crucial in drug discovery, where predicting the three-dimensional structure of a protein or optimizing the structure of a potential drug molecule can dramatically accelerate the development process. Moreover, AI-driven optimization algorithms can explore vast configuration spaces efficiently, guiding researchers towards stable and functionally relevant molecular structures. This approach not only saves time and resources but also expands the scope of what is computationally feasible in exploring molecular design.

2. Quantum Chemistry Calculations and Electronic Structure Prediction:

Quantum chemistry calculations, which provide detailed insights into electronic structure and molecular energetics, traditionally require substantial computational resources and expertise. AI techniques, including machine learning and neural networks, are increasingly being applied to enhance the efficiency and accuracy of these calculations. AI models can predict electronic properties and energies more rapidly than traditional quantum mechanical methods, often with comparable accuracy. This capability is particularly useful in studying complex molecular systems or reactions where a precise understanding of electronic structures is critical. Furthermore, AI-driven approaches can optimize computational workflows, such as selecting appropriate basis sets or improving convergence criteria, thereby streamlining the process of electronic structure prediction and making quantum chemistry more accessible to researchers across disciplines.

3. Molecular Dynamics Simulations and Protein-Ligand Interactions:

Molecular dynamics simulations play a vital role in understanding the dynamic behavior of molecules, such as proteins, and their interactions with ligands (small molecules or drugs). AI-enhanced molecular dynamics simulations leverage machine learning algorithms to improve the accuracy and efficiency of these simulations. AI models can predict molecular dynamics trajectories with greater fidelity, capturing subtle changes in molecular structure and dynamics over time. This capability is crucial for studying protein folding, binding mechanisms of drug molecules to their targets, and predicting the effects of mutations or environmental conditions on molecular behaviour. Moreover, AI can analyse vast amounts of simulation data to uncover hidden patterns or rare events that may influence biological function or drug efficacy. By integrating AI with molecular dynamics, researchers can gain deeper insights into complex biological processes and accelerate the discovery of new therapeutic agents.

III. Result and Discussion

1. AI in Drug Discovery and Design

1.1 Virtual Screening and Lead Identification:

Virtual screening is a computational technique used in drug discovery to identify potential drug candidates from vast libraries of molecules. Traditionally, this process relied on methods like rigid docking or predefined pharmacophore models, which could miss promising candidates due to their limitations. AI has revolutionized virtual screening by enabling more efficient and accurate methods. AI algorithms, such as machine learning models (including support vector machines and random forests) and deep neural networks, analyze molecular structures and their interactions with target proteins or receptors. These models can be trained on target-specific datasets of known bioactive compounds and their structures. By learning these features, the models can predict which molecules are likely to bind effectively to a specific target, prioritizing those with desired pharmacological activity. This in silico screening of millions of compounds accelerates drug discovery pipelines, reducing costs and minimizing extensive experimental testing.

1.2. De Novo Drug Design and Optimization:

De novo drug design involves creating entirely new molecules with desired therapeutic properties, guided by computational models and algorithms. This approach offers greater flexibility compared to traditional methods that modify existing compounds. However, navigating the vast chemical space efficiently to identify promising candidates remains a challenge. AI plays a crucial role in de novo drug design by generating novel chemical structures and optimizing them based on specified criteria, such as potency, selectivity, and pharmacokinetic properties. AI-driven approaches, including generative models like Variational Autoencoders (VAEs) and Generative Adversarial Networks (GANs), explore vast chemical space to propose candidate molecules with optimized properties. These models learn from large datasets of molecular structures and biological activities, enabling the design of compounds tailored to specific targets or disease mechanisms. Furthermore, AI can predict how modifications to molecular structures affect their interactions with biological targets, facilitating iterative optimization of drug candidates through integration with virtual screening techniques. This iterative process enhances the likelihood of identifying lead compounds with favourable drug-like properties early in the discovery phase, accelerating the drug discovery pipeline.

1.3.Pharmacophore Modelling and Target Prediction:

Pharmacophore modelling is a computational technique used to elucidate the essential structural features, such as hydrogen bond donors, acceptors, hydrophobic groups, aromatic rings, or charged groups, of a molecule required for binding to a biological target. AI-driven methods enhance pharmacophore modelling by automating feature selection and optimization, thereby improving the accuracy and reliability of predictions. AI algorithms can analyse molecular structures and biological data, including protein sequences, gene expression data, or known protein-ligand interactions, to identify common features associated with bioactivity across a dataset of known ligands. These features are then used to construct pharmacophore models that represent the spatial arrangement of key interactions between a molecule and its target. Moreover, AI facilitates target prediction by integrating these diverse biological and chemical data sources. This capability is crucial for understanding the polypharmacology (the ability of a drug to interact with multiple targets) of drug candidates and identifying off-target effects (unintended interactions with other targets) early in the drug discovery process, ultimately improving drug safety and efficacy.

1.4.AI for Reaction Prediction and Synthesis Planning

1.41.Predicting Reaction Outcomes and Designing Synthetic Pathways:

AI has transformed the prediction of chemical reaction outcomes and the design of synthetic pathways by leveraging machine learning algorithms, deep learning models, and knowledge graph-based approaches. Machine learning models trained on extensive datasets of reaction outcomes can predict the products of chemical reactions with high accuracy. These models learn from reaction conditions, reactants, catalysts, and other factors to anticipate which products are likely to form under specific circumstances. This capability accelerates the exploration of new chemical reactions and facilitates the discovery of novel synthetic routes.

Deep learning architectures, such as graph neural networks, can model molecular structures and reactions at a more detailed level, capturing subtle relationships between atoms and functional groups. This allows for more nuanced predictions of reaction outcomes and enables the design of synthetic pathways that optimize desired chemical transformations. AI-driven approaches also incorporate automated reaction planning algorithms that propose efficient sequences of chemical steps to synthesize target molecules. By integrating knowledge from reaction databases and chemical rules, these algorithms suggest feasible synthetic routes while considering factors such as yield, selectivity, and safety.

1.42.Retrosynthesis Analysis and Reaction Database Mining:

Retrosynthesis analysis involves working backwards from a target molecule to identify feasible starting materials and synthetic steps. AI-powered retrosynthesis tools automate this process by analyzing vast databases of known reactions and chemical transformations. AI algorithms employ techniques such as retrosynthetic planning using logic-based rules, template-based approaches, and machine learning-driven retrosynthesis prediction. These methods predict viable synthetic routes by matching the functional groups of the target molecule with those of available starting materials, considering reaction conditions and synthetic feasibility.

Furthermore, AI facilitates reaction database mining by extracting valuable insights from large repositories of experimental and theoretical reaction data. Natural language processing (NLP) techniques enable automated extraction and annotation of chemical information from scientific literature and patents, enriching reaction databases with structured data. By leveraging AI in retrosynthesis analysis and reaction database mining, researchers can streamline the process of designing synthetic routes, discover novel chemical transformations, and accelerate the development of new molecules with desired properties.

1.5.AI in Property Estimation and Optimization

1.51.Predicting Physical and Chemical Properties of Molecules:

AI has significantly enhanced the ability to predict the physical and chemical properties of molecules with high accuracy and efficiency. Traditional computational methods, such as density functional theory (DFT) and molecular dynamics simulations, can be resource-intensive and time-consuming. AI techniques, particularly machine learning (ML) and deep learning (DL) models offer a powerful alternative by leveraging large datasets to make rapid and accurate predictions. Machine learning models, trained on extensive datasets of molecular properties, can predict a wide range of properties, including boiling points, melting points, solubility, and reactivity. These models learn patterns and relationships within the data, enabling them to generalize predictions to new, unseen molecules. Deep learning architectures, such as graph neural networks (GNNs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs), are particularly effective in capturing the intricate relationships between molecular structure and properties. GNNs, for instance, represent molecules as graphs with atoms as nodes and bonds as edges, allowing the model to learn

from the connectivity and interactions within the molecule. AI-driven property prediction tools are widely used in various fields, including pharmaceuticals, materials science, and chemical engineering, to accelerate the design and discovery of molecules with desired properties.

1.52.Optimization of Molecular Properties for Specific Applications:

Optimizing molecular properties for specific applications is a critical aspect of molecular design and development. AI techniques facilitate this optimization by providing insights into how structural modifications impact properties and guiding the design of molecules with tailored characteristics. Generative models, such as variational autoencoders (VAEs) and generative adversarial networks (GANs), are used to explore chemical space and generate new molecular structures with optimized properties. These models can be trained to prioritize specific properties, such as drug-likeness, bioavailability, or material strength, enabling the design of molecules that meet application-specific criteria. Reinforcement learning (RL) approaches are also employed to iteratively improve molecular designs. In this framework, AI agents are trained to make modifications to molecular structures to maximize desired properties, receiving feedback in the form of a reward signal. This iterative process allows for the continuous optimization of molecular candidates until the desired property profile is achieved. Additionally, AI-driven multi-objective optimization techniques can simultaneously optimize multiple properties, balancing trade-offs between conflicting requirements. For example, in drug discovery, AI models can optimize compounds for both high potency and low toxicity, ensuring the development of safe and effective drugs.

2. AI Tools in Real-World Chemistry Problems

2.1 Drug Discovery and Development:

AI is revolutionizing drug discovery and development at every stage. BenevolentAI harnesses big data analysis to unearth new drug targets and predict candidate effectiveness. They repurposed Baricitinib for COVID-19 treatment, highlighting this approach's potential. Atomwise utilizes convolutional neural networks to power AtomNet, which rapidly identifies promising drug candidates by predicting how well molecules bind to protein targets. Insilico Medicine takes a bold leap forward with generative adversarial networks (GANs) to design entirely new drug molecules from scratch. Their AI-powered approach led to a fibrosis drug candidate in record time. Finally, Zymergen tackles drug production with AI. Their machine learning and automation platform optimizes microbe strains, leading to faster development of sustainable and efficient pharmaceutical manufacturing. In essence, AI is not only finding new drugs but also designing them and producing them better.

2.2. Case Studies from Academic Research

Machine learning (ML) is rapidly changing the landscape of scientific research. This passage highlights its impact on several key areas. In reaction prediction and synthesis planning, ML models like MIT's creation and IBM's RXN for Chemistry platform are aiding chemists. These tools analyze vast databases to predict

reaction outcomes and suggest efficient pathways, streamlining the design of new materials. Material discovery is another field benefiting from ML. Researchers at the University of Toronto used automated ML to discover promising materials for solar cells, while Stanford University's AI models identified materials with enhanced battery storage capabilities. Finally, DeepMind's AlphaFold represents a breakthrough in protein structure prediction. This program uses deep learning to predict protein structures with remarkable accuracy, furthering our understanding of biology and drug development. Overall, machine learning is proving to be a powerful tool for scientific discovery, prediction, and design.

3. Challenges and Limitations

The roadblocks for AI in this field are rooted in data. Training these models requires vast amounts of high-quality data, but obtaining it can be tricky. Validating data through experiments is expensive and time-consuming, and the pharmaceutical industry often keeps their data private. Even if data is available, its quality is crucial. Incomplete, biased, or inconsistent information can lead to inaccurate predictions and limit the model's ability to handle new situations. Another hurdle is interpretability. Many AI models, particularly complex ones, are opaque – we can't see how they arrive at their conclusions. This lack of transparency makes it difficult for chemists to trust the models and glean scientific insights. Finally, overfitting is a concern. This occurs when an AI model becomes overly specialized on the training data and struggles with new information. Small datasets are particularly susceptible to this issue. The consequence? Models that perform well on training data but fail in real-world applications. For AI to truly shine in computational chemistry, it needs to overcome these data-driven and interpretability challenges, with a focus on generalizability to handle the vast chemical landscape.

4. Future Directions and Emerging Trends

4.1. Enhanced Algorithms: This section highlights the development of more powerful and specialised AI tools for chemistry. Imagine AI with advanced neural network designs, allowing it to "learn" chemical data even better. Techniques like transfer learning will enable AI to adapt its knowledge from one area of chemistry to another, and reinforcement learning could allow AI to "experiment" virtually, optimizing processes and discovering new possibilities.

4.2.Integration of AI with Experimental Chemistry:

4.2.1. Closed-Loop Systems: This future envisions AI-powered labs! These systems would be self-contained, with robots performing synthesis, analyzing the results, and feeding the data back to the AI. The AI would then use this real-time information to refine the experiment, creating a continuous loop of learning and discovery.

4.2.2. Hybrid Approaches: Here, the focus is on combining the strengths of AI with traditional methods. Imagine AI suggesting promising avenues for research, then human chemists conducting experiments

to validate or disprove those suggestions. This collaboration would leverage the power of AI for creative exploration alongside the crucial role of human expertise in verification and refinement.

4.3.Ethical Considerations and Societal Impacts:

4.3.1. Bias and Fairness: Addressing biases in AI models to ensure fair and equitable outcomes in drug discovery and chemical research. Ensuring diverse and representative datasets to mitigate bias.

4.3.2. Transparency and Accountability: Developing guidelines and best practices for the transparent reporting and validation of AI models. Establishing standards for data sharing and model interpretability.

4.3.3. Societal Impacts: Considering the broader societal implications of AI-driven chemistry, such as the impact on employment, ethical use of AI in drug development, and potential misuse of AI technologies in chemical synthesis.

IV. Conclusion

AI's influence on computational chemistry is undeniable, with advancements in prediction accuracy leading to breakthroughs like efficiently designed drugs and streamlined synthesis pathways. Challenges persist, however, such as ensuring high-quality data, interpreting the inner workings of AI models, and avoiding overfitting. Despite these hurdles, the future looks bright. Researchers are developing improved algorithms, integrating AI with lab experiments for a more holistic approach, and addressing ethical considerations to ensure responsible development. By embracing these advancements, we can unlock the full potential of AI and accelerate scientific discovery in chemistry.

V. References

1. Abavisani, M., Khoshrou, A., Foroushan, S. K., & Sahebkar, A. (2024). Chatting with artificial intelligence to combat antibiotic resistance: Opportunities and challenges. *Current Research in Biotechnology*, 100197.

2. Baldi, P., & Nasr, R. (2020). When is chemical structure prediction accurate enough to be useful? The case of computational drug discovery and design. Journal of Chemical Information and Modeling, 60(4), 1960-1972.

3. Barzilay, R., & Jaakkola, T. (2019). Chemoinformatics and computational chemistry: Machine learning and AI in drug discovery. Drug Discovery Today, 24(4), 1002-1012.

4. Baum, Z. J., Yu, X., Ayala, P. Y., Zhao, Y., Watkins, S. P., & Zhou, Q. (2021). Artificial intelligence in chemistry: current trends and future directions. *Journal of Chemical Information and Modeling*, *61*(7), 3197-3212.

5. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798-1828.

6. Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., & Blaschke, T. (2018). The rise of deep learning in drug discovery. Drug Discovery Today, 23(6), 1241-1250.

7. Coley, C. W., Green, W. H., & Jensen, K. F. (2018). Machine learning in computer-aided synthesis planning. Accounts of Chemical Research, 51(5), 1281-1289

8. Dr Sanyogita Shahi, Dr Shirish Kumar Singh, Parinita Tripathy (2024)- "Wearable Device for measuring Biosignals from the body", Design No. 6333920, [Patent Country code- GB].

9. Dr Sanyogita Shahi, Dr Shirish Kumar Singh (2023)- "BIOSENSOR DEVICE FOR CANCER CELL DETECTION", Design No. 393041-002 [Patent Country code- IN]

10. Dr Sanyogita Shahi, Dr Shirish Kumar Singh, Dr. Vijaylaxmi Birdar, Dr. Ashtashil Vrushketu Bhambulkar, Mr. Honey Gaur (2023)- "BIOSENSOR DEVICE FOR CANCER CELL DETECTION", Design No. 6317655, [Patent Country code- GB].

11. Dr Sanyogita Shahi, Dr Shirish Kumar Singh (2023) - "WEARABLE DEVICE FOR MEASURING BIOSIGNALS FROM THE BODY", Design No. 399488-001, [Patent Country code- IN].

12. Dr Sanyogita Shahi, Dr Shirish Kumar Singh (2023) - "BIOSENSOR DEVICE FOR MEASURING BIOLOGICAL OR CHEMICAL REACTIONS", Design No. 383549-001, [Patent Country code- IN].

13. Dr Sanyogita Shahi, Dr Shirish Kumar Singh (2023) - "BIOSENSOR DEVICE FOR MEASURING BIOLOGICAL OR CHEMICAL REACTIONS", Design No. 6291366, [Patent Country code- GB].

14. Douguet, D. (2010). e-LEA3D: A computational-aided drug design web server. Nucleic Acids Research, 38(Web Server issue), W615-W621.

15. Duvenaud, D. K., Maclaurin, D., Aguilera-Iparraguirre, J., Gomez-Bombarelli, R., Hirzel, T., Aspuru-Guzik, A., & Adams, R. P. (2015). Convolutional networks on graphs for learning molecular fingerprints. Advances in Neural Information Processing Systems, 28, 2224-2232.

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

17. Gómez-Bombarelli, R., Aguilera-Iparraguirre, J., Hirzel, T. D., Duvenaud, D., Maclaurin, D., Blood-Forsythe, M. A., ... & Aspuru-Guzik, A. (2018). Design of efficient molecular organic light-emitting diodes by a high-throughput virtual screening and experimental approach. Nature Materials, 17(11), 977-985.

18. Guo, M., & Zhao, P. (2018). Advances in machine learning applications in chemoinformatics. Molecular Informatics, 37(9-10), 1700130.

19. Hamon, L., & O'Boyle, N. M. (2019). A Bayesian approach to the discovery of bioisosteres. Journal of Chemical Information and Modeling, 59(2), 760-768.

20. Han, R., Yoon, H., Kim, G., Lee, H., & Lee, Y. (2023). Revolutionizing medicinal chemistry: the application of artificial intelligence (AI) in early drug discovery. *Pharmaceuticals*, *16*(9), 1259.

21. Hartenfeller, M., & Schneider, G. (2011). De novo drug design. Wiley Interdisciplinary Reviews: Computational Molecular Science, 1(4), 742-759.

22. He, L., Bai, L., Dionysiou, D. D., Wei, Z., Spinney, R., Chu, C., ... & Xiao, R. (2021). Applications of computational chemistry, artificial intelligence, and machine learning in aquatic chemistry research. *Chemical Engineering Journal*, *426*, 131810.

23. Ivanenkov, Y. A., Polykovskiy, D., Bezrukov, D., Zagribelnyy, B., Aladinskiy, V., Kamya, P., ... & Zhavoronkov, A. (2023). Chemistry42: an AI-driven platform for molecular design and optimization. *Journal of Chemical Information and Modeling*, *63*(3), 695-701

24. Jiménez-Luna, J., Grisoni, F., & Schneider, G. (2020). Drug discovery with explainable artificial intelligence. Nature Machine Intelligence, 2(10), 573-584.

25. Jing, Y., Bian, Y., Hu, Z., Wang, L., Xie, X. Q. (2018). Deep learning for drug design: An artificial intelligence paradigm for drug discovery in the big data era. AAPS Journal, 20(3), 58.

26. Karthikeyan, A., & Priyakumar, U. D. (2022). Artificial intelligence: machine learning for chemical sciences. *Journal of Chemical Sciences*, *134*, 1-20.

27. Krenn, M., Häse, F., Nigam, A., Friederich, P., & Aspuru-Guzik, A. (2020). Self-driving laboratories: Machine learning for experimental design. Chemical Reviews, 120(12), 6981-7032.

28. Liu, G. Y., Yu, D., Fan, M. M., Zhang, X., Jin, Z. Y., Tang, C., & Liu, X. F. (2024). Antimicrobial resistance crisis: could artificial intelligence be the solution?. *Military Medical Research*, *11*(1), 7.

29. Lowe, D. (2012). Machine learning and chemical informatics: Promises and pitfalls. Drug Discovery Today, 17(17-18), 899-902.

30. Meredig, B., Agrawal, A., Kirklin, S., Saal, J. E., Doak, J. W., Thompson, A., ... & Wolverton, C. (2014). Combinatorial screening for new materials in unconstrained composition space with machine learning. Physical Review B, 89(9), 094104.

31. Popova, M., Isayev, O., & Tropsha, A. (2018). Deep reinforcement learning for de novo drug design. Science Advances, 4(7), eaap7885.

32. Pyzer-Knapp, E. O., Suh, C., Gómez-Bombarelli, R., Aguilera-Iparraguirre, J., & Aspuru-Guzik, A. (2015). What is high-throughput virtual screening? A perspective from organic materials discovery. Annual Review of Materials Research, 45, 195-216.

33. Rifaioglu, A. S., Atas, H., Martin, M. J., Cetin-Atalay, R., Dogan, T., & Atalay, V. (2019). Recent applications of deep learning and machine learning models in drug discovery: An overview. Computational and Structural Biotechnology Journal, 17, 207-218.

34. Schneider, G., & Fechner, U. (2005). Computer-based de novo design of drug-like molecules. Nature Reviews Drug Discovery, 4(8), 649-663.

35. Segler, M. H. S., Kogej, T., Tyrchan, C., & Waller, M. P. (2018). Generating focused molecule libraries for drug discovery with recurrent neural networks. ACS Central Science, 4(1), 120-131.

36. Sheridan, R. P. (2013). Time-split cross-validation as a method for estimating the goodness of prospective prediction. Journal of Chemical Information and Modeling, 53(4), 783-790.

37. Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. Nature, 529(7587), 484-489.

38. Skinnider, M. A., & Foster, L. J. (2020). Chemical space networks for predicting bioactivity. Nature Communications, 11(1), 2266.

39. Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., ... & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4), 688-702.

40. Tu, G., Fu, T., Zheng, G., Xu, B., Gou, R., Luo, D., ... & Xue, W. (2024). Computational Chemistry in Structure-Based Solute Carrier Transporter Drug Design: Recent Advances and Future Perspectives. *Journal of chemical information and modeling*.

41. Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., ... & Zhao, S. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery, 18(6), 463-477.

42. Walters, W. P., Murcko, M. A., & Murcko, M. A. (2002). Prediction of 'drug-likeness'. Advanced Drug Delivery Reviews, 54(3), 255-271.

43. Winter, R., Montanari, F., Steffen, A., Briem, H., Noé, F., & Clevert, D. A. (2019). Efficient multi-objective molecular optimization in a continuous latent space. Chemical Science, 10(6), 1692-1703.

44. Zhavoronkov, A., Aliper, A., & Zhavoronkov, A. (2019). Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. Molecular Pharmaceutics, 16(7), 3091-3098.