ISSN (ELECTRONIC): 2815-0562

Jain Neha, Kumar Dhiraj, Singh Supriya, Joshi Sarita, Joshi Ashish, Yadav Sachin Kumar. (2024). Integration of Artificial Intelligence In Computer-Aided Drug Design: Advancing Drug Discovery And Development Processes. International Journal of Medical Science. 4(2), 2024, 15-27.

INTERNATIONAL JOURNAL OF MEDICAL SCIENCE

journal homepage: https://www.ijmsci.org/

INTEGRATION OF ARTIFICIAL INTELLIGENCE IN COMPUTER-AIDED DRUG DESIGN: ADVANCING DRUG DISCOVERY AND DEVELOPMENT PROCESSES

Neha Jain*[1], Dhiraj Kumar[2], Supriya Singh[3], Sarita Joshi[4], Ashish Joshi[5], Sachin Kumar Yadav[6]

*[1]Research Scholar/Associate Professor, SOP Raffles University, Neemrana, Rajasthan/ Sunder Deep Pharmacy College, Ghaziabad, India.

[2,3,6] Assistant Professor, Sunder Deep Pharmacy College, Dasna, Ghaziabad (U.P.)–201003, India.

[4]Research Scholar, Dr. K. N. M.I.P.E.R, Modinagar, Ghaziabad, India.

[5]Assistant Professor, ABESIT College of Pharmacy, Ghaziabad, India.

How to Cite the Article: *Jain Neha, Kumar Dhiraj, Singh Supriya, Joshi Sarita, Joshi Ashish, Yadav Sachin Kumar. (2024). Integration of Artificial Intelligence In Computer-Aided Drug Design: Advancing Drug Discovery And Development Processes. International Journal of Medical Science. 4(2), 2024, 15-27.*

DOI: https://doi.org/10.56815/IJMSCI.V4I2.2024/15-27

Article History: 22-10-2024; Revision: 21-11-2024; Acceptance: 28-12-2024;

Keywords	***Abstract***
AIDD, CADD, Molecular Models, Spatial Symmetry Preserving Networks, Hybrid De Novo Designs, Data Sharing, Model Development.	Artificial intelligence (AI) is also called computer-aided drug discovery (CADD). The initial development of AI in pharmaceutical discoveries has arisen from the applications of AI in Medicinal chemistry. The basic areas where AI is much involved are: (1) quantitative structure activity relationship, (2) structure based modeling, (3) de novo molecular design, and (4) predictions of chemical synthesis. The widespread adoption of AI, particularly in deep learning, is driven by advancements across various scientific disciplines, improvements in computing hardware and software, and other contributing factors. Advanced methodologies like message-interlinking models, spatial-symmetry-preserving networks, hybrid de novo designs, and other innovative machine learning techniques provide effective solutions to complex problems. The sharing of data and the advancement of models are crucial in exploitation of Artificial Intelligence for drug discovery and development. At present, The AI methods are synonyms for molecular modeling methods. So, it appears that the AIDD

	(Artificial Intelligence based Drug Discovery) offers the methods to discover a drug are totally dependent on molecular techniques. The advancement of data-intensive biomedical research tools and technologies, including DNA sequencing, imaging, and tools designed to track and record a patient's health information, has driven the need for developers to incorporate Artificial Intelligence. Artificial Intelligence offers a broad spectrum of statistical methods to process and analyze the large range of data generated by these assays. Moreover, these technologies have uncovered tremendous differences among individuals at genetic, biochemical, physiological, exposure, and communicative levels, particularly proportionate to disease progression and treatment response.As found the benefits of AI in drug development, AI can be used in assistance of gene therapy, which are currently not available as tools in healthcare. With AI, the possibilities of combining pharmacology and gene therapy would provide satisfactory results.In this review, we will provide an introduction to potential uses of AI within the drug designing and development process, in particular compared to conventional methods for carrying out these tasks and highlighting the pros and cons of AI. We will also focus on the early stages of discovery of new drug compounds and preclinical drug development.

[1] INTRODUCTION

The life of human beings is constantly changing and the final aim of humans is to magnify these changes for their uses and not to harm. This special change has to be seen in the field of medicines and pharmaceuticals. The pharmaceutical industry is the only discipline, which is focused on the detection, development, & marketization of newly developed molecules day by day. The main aim to discover new drugs for the same disorders is to compensate for the old one's side effects and aftereffects. So, pharmacists and scientists discover or create chemical compounds and mixtures to ease physiological and psychic suffering. From the last so many years, the creation of pharmaceutical products has been guided by a regulatory system that ensures the quality of the end product. (1)

In today's scenario, the pharmaceutical industry is urged to develop mechanics in the field of drug development, to create suitable medications for humans. Creating complex medications for the safe usage of humans and incorporating them for therapeutic use is a great challenge for technological resources. (2)

Today, scientists working on the "one size fits all" principle; anyhow, various departments of pharmaceuticals need new and novel approaches to develop these complexions of medicines. Improvements in genomics and molecular modeling help enable these procedures. The tasks of AI is improving gradually and providing a platform for clinical examinations and interns training. Participation of health care providers can also be useful in the development of these technologies maximize the possibilities of artificial intelligence in healthcare. (3)

There are some ways through which AI can be used in the pharmaceutical industry. Some of them are as follows:

- Evaluation of the intensity of the condition and forecasting the efficacy of the treatment prior to its application.
- Prevention of complications during treatment.
- Determination of the chemicals and drug usage during treatment.

- Creation of innovative medications and tools to enhance the effectiveness and safety of medical treatments.

There are various techniques to generate data in pharmaceutical research such as DNA sequencing, visual examination techniques, vital signs monitoring systems, etc. Despite the fact that a wide variety of numerical methods have been available to analyze the findings, there is a need for Artificial intelligence to accommodate such kind of massive data. (4)

Furthermore, the use of data-heavy technologies has shown that people can differ in their genetic makeup, bio and physicochemical processes, physical functions, and cognitive behaviors, particularly in how they develop diseases and respond to treatments. It is suggesting that there is an urge to shift to tailor or personalized medicines which can be avail to individuals. However, to reveal appropriate targets and strategies, there is a need for data-intensive assays, and then AI can help to precede such assays. So, AI is important in the development of tailored medicines and implementations of new personalized health products for their complete utility. In terms of personalized medicines, pharmaceutical sciences are guided by so many interrelated concepts. The most important threads are:

- Possession of unique individual features like genetic, behavioral, biochemical, and physiological to express the need for personalized medicines.
- Misuse of data-generating assays such as DNA sequencing, proteomics, molecular modeling, imaging strategy, and wireless health monitoring devices.
- Big data research programs in which massive amounts of data were generated from different sources to identify the patterns that would normally not be identified were analyzed independently.
- Artificial intelligence, which is used for algorithms, based on machine learning, deep learning, neural network, and hybrid techniques, can be used to find relevancy in massive data sets.

These four subjects are closely connected. For example, tailoring medical treatment to an individual necessitates a thorough comprehension of the patient's health status and unique circumstances. Achieving this level of understanding correlates with the utilization of advanced techniques that produce large summation of data, such as DNA sequencing or detailed imaging procedures. (5)

[2] COMPUTER AIDED MOLECULAR DESIGN (CAMD)

There are two important categories found of CAMD: (i) Structure based drug design (SBDD) and, (ii) ligand based drug design (LBDD). SBDD is based at the stereographic structure of the target and principal sites to find ligand-target interactions, while, LBDD is used when target is unknown for the three dimensional structures. (6)

Drug Design (SBDD) and Ligand-Based Drug Design (LBDD):

Aspect	SBDD (Structure-Based Drug Design)	BDD (Ligand-Based Drug Design)
Approach	Utilizes the 3D structure of the target (e.g., protein, enzyme) to design drugs.	Relies on the knowledge of ligands (active molecules) that bind to the target.
Input Data	Crystal structure or modeled 3D	A set of known active and inactive

Aspect	SBDD (Structure-Based Drug Design)	BDD (Ligand-Based Drug Design)
Required	structure of the target (e.g., from X-ray crystallography or NMR).	ligands with activity data.
Focus	Focuses on understanding the interaction between the target and the ligand.	Focuses on identifying structural features of ligands responsible for activity.
Key Techniques	Docking, molecular dynamics simulations, and binding affinity predictions.	QSAR (Quantitative Structure-Activity Relationship), pharmacophore modeling.
Strength	Enables rational drug design by visualizing binding interactions.	Useful when the target structure is unknown but active ligands are available.
Limitations	Requires accurate target structure; limited by target flexibility.	Limited by the quality and quantity of ligand data; may not work if ligand diversity is low.
Applications	Identification of novel binding sites, fragment-based drug design.	Derivation of SAR (Structure-Activity Relationships), virtual screening based on ligand similarity.
Dependency on Target	Strongly dependent on the availability and accuracy of the target structure.	Independent of the target structure; relies on ligand information.
Computational Tools	AutoDock, Glide, MOE, GROMACS.	Discovery Studio, ChemOffice, Schrödinger's Phase.

A. Structure Based Drug Design

The frequently employed plan of action for drug design and refinement within structure-based drug design (SBDD) include molecular dynamics (MD) simulations, docking studies, and analysis of target-ligand interactions to assess structural alterations in the biological target. Leveraging SBDD approaches, several approved drugs, including Imatinib, Indinavir, Nilotinib, and Lifitegrast, have been successfully developed. The SBDD process typically necessitates steps such as target modeling, identifying binding pockets, preparing compound libraries, conducting docking experiments with scoring evaluations, and performing MD simulations. (7)

(I) Target Preparation

In structural biology, there is structure of target proteins available. Computational biology approaches like homology modeling, Alpha-fold and ab-initio protein structures can predict to the binding at target structures based on sequencing.

Homology modeling involves choosing a specific template structure to construct the target structure. Alpha Fold predicts three-dimensional protein structures based on their amino acid sequences. Meanwhile, ab initio protein structure prediction is used when no template structure

is available in the Protein Data Bank (PDB). This approach, often referred to as global optimization, aims to determine the tertiary structure of a given target with the least energy, relying solely on its primary structure. (8)

(II) Ligand Binding Site Information

Identifying binding sites is a crucial step in molecular docking. These sites on target proteins can be determined through site-directed mutagenesis or by analyzing co-crystallized protein-ligand structures. When no prior knowledge of the binding pocket is available, a blind search is necessary to predict potential binding sites. In this blind docking approach, the entire protein surface is explored to identify possible binding modes. This process often involves multiple trials to achieve an optimal ligand-protein complex. (9, 10)

(III) Compound Library Preparation

Compounds chosen for drug development are typically sourced from various compound libraries, such as Enamine's REAL library (which contains around 1.4 billion compounds, tailored to specific requirements), ZINC (offering 750 million compounds ready for docking), MCULE (with 122 million synthetic compounds), PubChem (featuring 112 million bioactive compounds), DrugBank (providing 14,528 approved drug molecules), ChEMBL (holding roughly 2.2 million bioactive molecules), and ChemDB (which includes around 5 million commercially accessible molecules). These compounds are then filtered using criteria such as Lipinski's "Rule of Five," Veber's guidelines, ADMET properties, and other specific factors to determine their potential activity in physiological systems. (11, 12, 13)

(IV) Molecular Docking & Scoring

Currently, a variety of tools are utilized for molecular docking, such as Autodock, AutoDock Vina, CDOCKER, GLIDE, DOCK6, GOLD, FLEXX, and SwissDock. When considering the flexibility of both the ligand and target, molecular docking can be classified into the following categories: (a) rigid docking, where both the ligand and target are fixed in shape; (b) semi-flexible docking, where the ligand is flexible but the target remains fixed; and (c) flexible docking, where both the ligand and target are flexible. The precision of molecular docking largely relies on the binding affinity and the modes of interaction between the ligand and target. To assess binding affinity, several scoring functions are available, including physics-based, empirical, deep learning, and machine learning approaches. Among the prominent deep learning techniques are EquiBind, GNINA, and DiffDock, which are designed to predict the binding modes between a ligand and a specific protein target.. (14, 15, 16)

(V) MD Simulations

MD simulations may enhance the malleability of the targeted protein and generate confirmations for the binding at target protein. MD simulations can be used for the scoring (obtained from the molecular docking) and lead development. With the help of MD simulations researchers can assess the binding efficacy of the compound with the targeted molecules. MD simulations can also be helpful to find ligand-target interactions, which can provide further informations for the development of peptides. (17, 18, 19)

B. Ligand Based Drug Design

LBDD starts with a single compound or compound with a mixture against a specific target protein. During the process of drug discovery (for new categories, especially) there is lack of availability of target binding sites. For this, LBDD would be applied, which starts with a single

compound against the specific target protein structure. Now, if the compound is similar as its chemical structure is resemble with any priory molecule, and then it will show the physiological activities similar as the prior available molecule. According to SAR, the physicochemical and biological properties can be improved by designing appropriate analogues like removal of any toxic group and adjuvant of any other group. This practice is generally used in pharmacophore modeling and QSAR, and the finalized molecule shows better activity then formerly available molecule in its efficacy and cause lesser side effects. (20, 21)

(I) Ligand Structures and Bioactivity

Structure oriented generation of drugs produces novel drugs that binds with the specific target proteins. These target proteins consist of receptors, enzymes, and both structural and functional proteins found at the binding sites (such as GABA, glycine, glutamate, etc.). The creation of a new molecule follows the "fragment strategy," which involves constructing an initial chemical scaffold that fits within the binding site of the target protein. A pre-trained model generates new molecules by adding, removing, inserting, replacing, and linking fragments accordingly. The availability of structural information for both the target proteins and the molecular components, along with structure-guided design, enhances the efficacy and tolerability of the molecule at the binding site of the target protein.

(II) Pharmacophore & Qsar / Qspr Based Modeling:

QSAR/QSPR modeling has been proven effective in predicting biological activity and pharmacokinetic parameters, including absorption, distribution, metabolism, excretion, and toxicity (ADMET). (6-11)

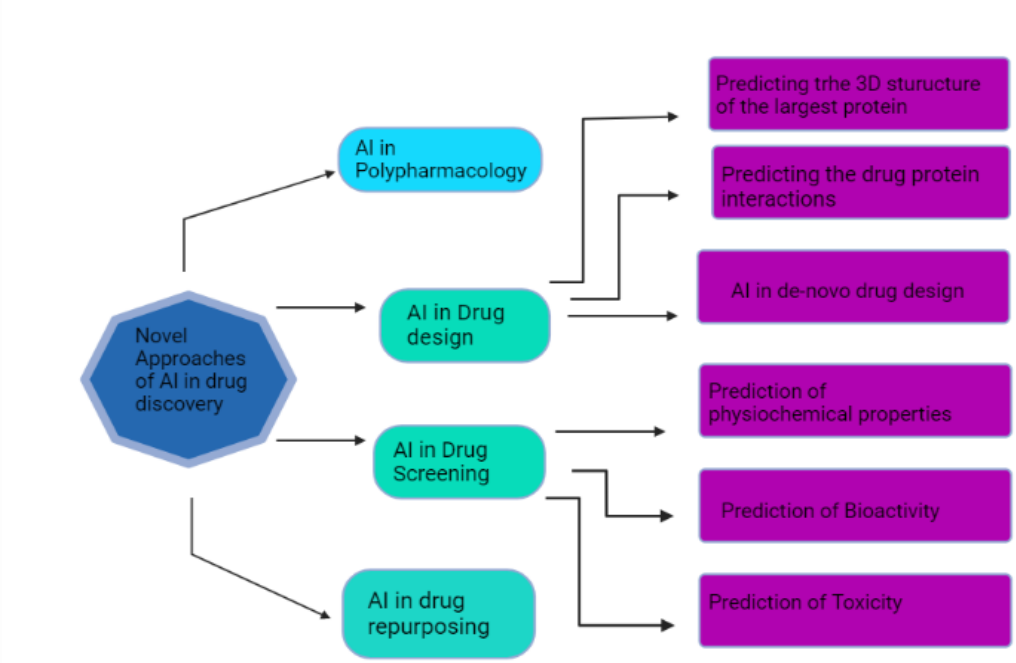
For ligand based QSAR/QSPR modeling, the basic structure of molecules is converted into the coding numbers, known as molecular descriptors. The aim to provide a molecular descriptor is to capture other chemical moieties with the same structure. Ligand based QSAR/QSPR modeling transitions toward machine learning methods (universally accepted), like support vector machines (SVM) and gradient boosting methods (GBM), which is based on the more potential relationships between the chemical structure and its physical, chemical and biological belongings. (12)

Artificial neural networks, which fall under the category of machine learning or artificial intelligence, encompass various network structures such as deep and adaptive networks, self-organizing maps, auto-encoders, and methods for analyzing sequential time series, among others. Deep learning methods offer various advantages, such as, automatic feature extraction, graphing and presentation of molecular structures. Therefore, deep learning perform well planned modeling task for descriptors, for example, modeling of peptides, macrocycles, and proteolysis targeting chimers. (13)

(III) Ligand Based Virtual Screening

Drug designing through AI is a creative process through which the molecule with ligand binding properties can be created. Meanwhile, a molecule binds with a particular protein so that it can modify physiological functions related with that particular protein. This relation shows the adverse and physiological effects of the particular molecule at the same time. Ligands bindings with machine learning can handle the complex relationship between the input and output variables for high dimensional data. So, machine learning has been used for molecular structural properties prediction and chemical moieties carried by it and its drug-like properties. (14, 15, 16, 17)

ISSN (ELECTRONIC): 2815-0562
Jain Neha, Kumar Dhiraj, Singh Supriya, Joshi Sarita, Joshi Ashish, Yadav Sachin Kumar. (2024). Integration of Artificial Intelligence In Computer-Aided Drug Design: Advancing Drug Discovery And Development Processes. International Journal of Medical Science. 4(2), 2024, 15-27.



[3] DE NOVO DRUG DESIGN BY ARTIFICIAL INTELLIGENCE

De novo drug design is the generation of a novel molecular moiety with willing pharmacological properties, considered as a challenging computer assisted task in drug discovery, due to the presence of combinatorial explosion (exponential growth rate at which most programs grow) (18), these approaches related with De novo molecule designs cab either be ligand-based, structure based or a mixture of both. Ligand based methods have two categories: (i) Rule based: Which is based on construction rules of building blocks, and (ii) Rule free: which do not refer any building blocks? (19)

AI Technique	**Description**	**Examples/Applications**
Generative Adversarial Networks (GANs)	GANs generate novel molecular structures by training a generator and discriminator in tandem.	**MolGAN**: Generates molecules with desired properties.
Variational Autoencoders (VAEs)	VAEs encode molecular data into a latent space and decode it to design new molecules.	**Junction Tree VAE**: Generates chemically valid molecules.
Reinforcement Learning (RL)	RL optimizes molecule generation by rewarding desired properties (e.g., binding affinity).	**REINVENT**: Optimizes molecules for specific objectives.
GraphNeural Networks (GNNs)	GNNs model molecular graphs to predict properties and design structures.	Used for **structure-property predictions** and generation.

AI Technique	Description	Examples/Applications
Transformer Models	Sequence-based models learn chemical syntax (SMILES) for molecule generation.	**ChemBERTa, MolGPT**: Learn from chemical datasets.

[4] AUTOMATED SYNTHESIS PLANNING

The major category of organic compounds can be synthesized with the intense reactions. However, fully automated planning to synthesize a new molecule is a challenge for expertise. Retrosynthetic planning with increased computational capabilities and development of structure based algorithms for deep learning have resulted in the rejuvenation of AI for synthetic molecular chemistry. (20) This is rule based learning, which aim to suggest reaction mechanism encoding and skeletal building of the molecule which is to be synthesized. (21)

[5] ARTIFICIAL INTELLIGENCE IN DRUG DESIGNING

Applications of Artificial Intelligence (AI) in various stages of Drug Designing:

AI Technique	Application	Stage in Drug Design	Examples/Tools
Machine Learning (ML)	Predicting drug-target interactions, toxicity, and ADMET properties.	Lead identification and optimization.	DeepChem, scikit-learn, ChemProp.
Deep Learning (DL)	Generating novel molecular structures, QSAR modeling.	De novo drug design, hit-to-lead optimization.	ChemGAN, DeepDrug, AlphaFold.
Natural Language Processing (NLP)	Mining scientific literature and patents for drug discovery insights.	Data curation and hypothesis generation.	IBM Watson, PubMedNLP, SciBite.
Generative Models (e.g., GANs)	Designing novel molecules with desired properties.	De novo drug design.	MolGAN, Graph-to-Graph.
Reinforcement Learning (RL)	Optimizing molecular designs for specific objectives.	Lead optimization.	REINVENT, DeepRL.
Predictive Analytics	Forecasting clinical trial outcomes, patient responses.	Clinical trial design.	IBM Watson Health, SAS Predictive Analytics.
AI in Docking	Accelerating virtual screening by predicting binding affinities.	Structure-based drug design (SBDD).	Glide AI, AtomNet.
Graph Neural Networks (GNNs)	Modeling molecular interactions, predicting properties of molecules.	Drug-target interaction prediction.	D-MPNN, TorchDrug.
Bioinformatics Integration	Analyzing omics data for target identification.	Target identification and validation.	TensorFlow, Keras in genomic analysis pipelines.

AI Technique	Application	Stage in Drug Design	Examples/Tools
AI-Driven High-Throughput Screening	Screening millions of compounds virtually to identify leads.	Hit identification.	Schrodinger's Maestro, AutoDock with AI enhancements.
AI for ADMET Prediction	Predicting absorption, distribution, metabolism, excretion, and toxicity.	Safety evaluation and optimization.	pkCSM, SwissADME, ADMETlab.
AI in Precision Medicine	Designing drugs tailored to individual genetic profiles.	Personalized drug development.	DeepVariant, GATK, personalized AI platforms.

Modern drug designing and discovery is based on target identification. There are various approaches such as structural biology, molecular biology, cell biology genomics, proteomics and bioinformatics to identify the target and its pathogenesis. Computational chemistry techniques such as MM (molecular modeling), QM (quantum mechanics), MD (molecular dynamics), etc. are widely used to explore pathogenic mechanism and drug resistance. (22, 23)

[6] FUTURE SCOPE

The primary benefit of AI in the pharmaceutical industry is its ability to reduce costs and enhance the effectiveness of designed molecules. In-depth research has shown that dynamic learning can create highly accurate AI models while using only half or even less data compared to traditional AI and data subsampling methods. Although the exact cause of this increased efficiency is not completely understood, it seems that reducing repetition and bias, as well as acquiring more meaningful data to navigate decision boundaries, are key factors contributing to improved performance. As a result, excluding the expected technical overhead of implementing dynamic learning efforts, screening costs can be reduced by up to 90%. Machine learning techniques are capable of handling complex analyses of large, diverse, and high-dimensional data sets without manual input, which has proven valuable in various business applications. The integration of machine learning, particularly deep learning, with human expertise and experience may offer the best approach for managing extensive data repositories. The powerful data-mining abilities of AI technology have brought a new significance to computer-assisted drug design, especially when it involves multiple clinical factors, improving the speed of drug development. With the continued collection of clinical data and refinement of AI algorithms, AI technology is expected to revolutionize many aspects of drug discovery and development, ultimately becoming the standard method for computer-assisted drug planning. The combination of automation and technology resulting from these integrated advancements should lead to breakthroughs in drug development by enabling more accurate analysis of large and complex datasets. This progress will be critical in shortening drug development timelines, reducing costs, and increasing success rates, which is the ultimate objective of implementing AI in this field.

[7] SUMMARY

In conclusion, several factors influence the successful incorporation of AI and machine learning into drug discovery, development, and the pharmaceutical sector, particularly in areas like

polypharmacology, drug design, screening, and repurposing. Technological progress, especially in AI, will always be essential for cutting down research, development, and production time and costs, while enhancing efficiency. This literature review demonstrates that AI and machine learning have the potential to boost both the efficiency and precision of drug discovery and development. These technologies not only streamline processes but, in some instances, replace the need for clinical trials through simulations, allowing researchers to explore molecules more thoroughly without trials, thus cutting costs and addressing ethical concerns. While integrating AI and machine learning could eventually transform drug development, obstacles such as cleaning unstructured data and occasional limitations in computing devices may hinder progress. Overcoming these challenges will enable broader application and refinement of AI and machine learning, marking the beginning of a new era in the pharmaceutical industry.

[8] AUTHOR(S) CONTRIBUTION

The authors agreed to have no connections or engagements with any group or body that provides financial and non-financial assistance for the topics and resources covered in the article.

[9] LIMITATIONS

The size of the sample was very small.
The study was completely conducted on senior citizens.

[10] RECOMMENDATIONS

Needs to conduct in Tai-chi exercise to assess the physical problems in old age people.
Comparison research may be done to discover changes in adults and old age
Recommend to do this study as qualitative research.

[11] ACKNOWLEDGEMENT

Individuals / resources participated in the work are acknowledged properly.

[12] SOURCES OF FUNDING

The authors received no financial aid to support the study.

[13] PLAGIARISM POLICY

The authors declare that any kind of violation of plagiarism, copyright, and ethical matters will be handled by all authors. Journalists and editors are not liable for the aforesaid matters.

[14] CONFLICT OF INTEREST

The authors declared that no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

[15] PROTECTION OF RESEARCH PARTICIPANTS

This study do not involve any such criteria or condition.

REFERENCES

[1] Patel V., Shah M., Artificial intelligence and machine learning in drug discovery and Development, Intelligent Medicine 2 (2022) 134–140.

[2] Rantanen J, Khinast J., "The future of pharmaceutical manufacturing sciences". J Pharm Sci 2015;104 (11):3612–38. doi:10.1002/jps.24594.

[3] Greenhill AT, Edmunds BR. A primer of artificial intelligence in medicine. Tech Innov Gastrointest Endosc 2020;22 (2):85–9. doi:10.1016/j.tgie.2019.150642.

[4] Schork J.N., "Artificial Intelligence and Personalized Medicines", Cancer Treat Res. 2019; 178: 265–283. doi:10.1007/978-3-030-16391-4_11.

[5] Russell S and Norvig P, Artificial Intelligence: A Modern Approach (3rd Edition). 3rd ed 2009, Carmel, Indiana: Pearson.

[6] Huang, H.-J.; Yu, H.W.; Chen, C.-Y.; Hsu, C.-H.; Chen, H.-Y.; Lee, K.-J.; Tsai, F.-J.; Chen, C.Y.-C. Current developments of computer-aided drug design. J. Taiwan Instit. Chem. Eng. 2010, 41, 623–635.

[7] Macalino, S.J.Y.; Basith, S.; Clavio, N.A.B.; Chang, H.; Kang, S.; Choi, S. Evolution of in silico strategies for protein-protein interaction drug discovery. Molecules 2018, 23, 1963.

[8] Hetényi, C.; van der Spoel, D. Blind docking of drug-sized compounds to proteins with up to a thousand residues. FEBS Lett. 2006, 580, 1447–1450.

[9] Hetényi, C.; van der Spoel, D. Blind docking of drug-sized compounds to proteins with up to a thousand residues. FEBS Lett. 2006, 580, 1447–1450.

[10] Hassan, N.M.; Alhossary, A.A.; Mu, Y.; Kwoh, C.-K. Protein-ligand blind docking using QuickVina-W with inter-process spatio-temporal integration. Sci. Rep. 2017, 7, 15451

[11] Kim, S.; Thiessen, P.A.; Bolton, E.E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B.A. PubChem substance and compound databases. Nucleic Acids Res. 2016, 44, D1202–D1213.

[12] Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B.A.; Thiessen, P.A.; Yu, B. PubChem 2019 update: Improved access to chemical data. Nucleic Acids Res. 2019, 47, D1102–D1109.

[13] Wishart, D.S.; Knox, C.; Guo, A.C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. DrugBank: A knowledgebase for drugs, drug actions and drug targets. Nucleic Acids Res. 2008, 36, D901–D906.

[14] Coley, C.W.; Green, W.H.; Jensen, K.F. Machine learning in computer-aided synthesis planning. Acc. Chem. Res. 2018, 51, 1281–1289.

[15] Luo, Y.; Yan, K.; Ji, S. Graphdf: A discrete flow model for molecular graph generation. In Proceedings of the 38th International Conference on Machine Learning, Online, 18–24 July 2021; pp. 7192–7203.

[16] Jin, W.; Barzilay, R.; Jaakkola, T. Hierarchical generation of molecular graphs using structural motifs. In Proceedings of the 37th International Conference on MachineLearning, Online, 13–18 July 2020; pp. 4839–4848.

[17] Jin, W.; Barzilay, R.; Jaakkola, T. Multi-objective molecule generation using interpretable substructures. In Proceedings of the 37th International Conference on MachineLearning, Online, 13–18 July 2020; pp. 4849–4859.

[18] McNutt, A.T.; Francoeur, P.; Aggarwal, R.; et. al. 1.0: Molecular docking with deep learning. J. Cheminformatics 2021, 13, 43.

[19] Corso, G.; Stärk, H.; Jing, B.; Barzilay, R.; Jaakkola, T. DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking. arXiv 2022, arXiv:2210.01776.

[20] Stark, H.; Ganea, O.; Pattanaik, L.; Barzilay, R.; Jaakkola, T. Equibind: Geometric deep learning for drug binding structure prediction. In Proceedings of the International Conference on Machine Learning, Baltimore, MD, USA, 17–23 July 2022; pp. 20503–20521.

[21] Salsbury, F.R., Jr. Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. Curr. Opin. Pharmacol. 2010, 10, 738–744.

[22] Ferreira, L.G.; Dos Santos, R.N.; Oliva, G.; Andricopulo, A.D. Molecular docking and structure-based drug design strategies. Molecules 2015, 20, 13384–13421.

[23] Sledz, P.; Caflisch, A. Protein structure-based drug design: From docking to molecular dynamics. Curr. Opin. Struct. Biol. 2018, 48, 93–102.

[24] Gurung, A.B.; Ali, M.A.; Lee, J.; Farah, M.A.; Al-Anazi, K.M. An Updated Review of Computer-Aided Drug Design and Its Application to COVID-19. BioMed Res. Int. 2021, 2021, 8853056.

[25] Philip, P.; Anshuman, D.; Anil, K.S. Computer-aided drug design: Integration of structure-based and ligand-based approaches in drug design. Curr. Comput. Aided Drug Des. 2007, 3, 133–148.

[26] Hansch C, Maloney PP, Fujita T, et al. Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. Nature. 1962;194 (4824):178–180.

[27] Goller A, Kuhnke L, Montanari F, et al. Bayer's in silico ADMET platform: a journey of machine learning over the past two decades. Drug Discov Today. 2020;25(9):1702–1709.

[28] Winiwarter S, Ahlberg E, Watson E, et al. In silico ADME in drug design – enhancing the impact. ADMET DMPK. 2018;6(1):15–33.

[29] Beresford AP, Segall M, Tarbit MH. In silico prediction of ADME properties: are we making progress? Curr. Opin. Drug Discov. Devel. 2004;7(1):36–42.

[30] Norinder U, Bergstrom CAS. Prediction of ADMET properties. Chem Med Chem. 2006;1(9):920–937.

[31] Beck B, Geppert T. Industrial applications of in silico ADMET. J. Mol. Model. 2014;20(7). DOI:10.1007/s00894-014-2322-5.

[32] Chen H, Engkvist O, Wang Y, et al., The rise of deep learning in drug discovery. Drug Discov Today. 23(6): 1241–1250. 2018.. (••) An excellent introduction to the basics of deep learning in the context of drug discovery pipelines.

[33] Henninot A, Collins JC, Nuss JM. The current state of peptide drug discovery: back to the future? J. Med. Chem. 2018;61(4):1382–1414.

[34] Ruddigkeit L, Van Deursen R, Blum LC, et al. Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17. J. Chem. Inf. Model. 2012;52(11):2864–2875.

[35] Christ CD, Zentgraf M, Kriegl JM. Mining electronic laboratory notebooks: analysis, retrosynthesis, and reaction based enumeration. J. Chem. Inf. Model. 2012;52(7):1745–1756.

[36] Jimenez-Luna, J., Grisoni, F., Weskamp, N., & Schneider, G. (2021). Artificial intelligence in drug discovery: recent advances and future perspectives. Expert Opinion on Drug Discovery, 16(9), 949-959.

[37] Boomsma, W.; Nielsen, S.V.; Lindorff-Larsen, K.; Hartmann-Petersen, R.; Ellgaard, L. Bioinformatics analysis identifies several intrinsically disordered human E3 ubiquitin-protein ligases. PeerJ 2016, 4, e1725. [PubMed]