Received: 20 June 2025, Accepted: 22 June 2025, Published: 29 June 2025
Digital Object Identifier: https://doi.org/10.63503/j.ijssic.2025.158

Review Article

Quantum Machine Learning for Drug Discovery: A Systematic Review

Anisurrahman*, Bashir Alam, Muhammad Hamid

Department of Computer Engineering,

Jamia Millia Islamia University, New Delhi, India.

er.rahman2010@gmail.com, balam2@jmi.ac.in, hamidmd504@gmail.com

*Corresponding author: Anisurrahman, er.rahman2010@gmail.com

ABSTRACT

Quantum computing plays a significant role in simulating molecules and atoms and offers advantages in chemistry over classical computing. The potential of Quantum Machine Learning (QML) can be used in drug discovery, chemical reaction simulations, and Material design for pharmaceuticals. QML leverages quantum computing and advanced machine learning to accelerate the identification of drug candidates, predict molecular interactions, and optimize compounds. In this paper, we present a systematic review of the methods used for molecular property prediction and molecular generation using quantum machine learning. We have included the recent research, perspective, advantages, and challenges that must be addressed to achieve this task. The objective of this research is not only to discuss current strategies and methods used for drug discovery but also to promote interdisciplinary research in the field of quantum computing and chemistry for wellness.

Keywords: *Quantum Machine Learning, Drug Discovery, Molecular Simulation, Molecule Generation, Noisy intermediate-scale Quantum (NISQ).*

1. Introduction

Identifying a new therapeutic drug requires extensive calculations, deep analysis of data, and simulation of molecular structures so that biological interactions can be predicted and optimized. These tasks, particularly those involving quantum chemistry, are too intricate for classical approaches to manage. QML has the potential to improve the drug discovery process due to the power of quantum computing integrated into machine learning. We have witnessed that quantum computers have surpassed classical computers in solving many difficult problems, such as factorization of large prime numbers and searching in unstructured databases [1,2]. The implementation of quantum computing is being explored in agriculture, finance, communication, space, transportation, energy, and healthcare. Quantum computing uses quantum phenomena such as superposition, interference, and entanglement to process quantum information. In the early 1980s, Richard Feynman showed computation using quantum information [3], which enabled the development of quantum hardware. Today, many countries and organizations are developing quantum computers with different technologies such as superconducting qubits, trapped ion-based qubits, photonics-based qubits, and neutral atom qubits to surpass the classical computing limitations [4-8]. Companies like Google, IBM, and Rigetti are developing quantum hardware, and there are also software frameworks like Qiskit (IBM) and Cirq (Google) that help programmers run quantum algorithms.

Quantum computing uses qubits to process information. Unlike classical bits, qubits can simultaneously be in the state of 0 and 1, which enables quantum computing to process large amounts of data much faster than classical computers. Qubits are represented as a point on the Bloch sphere. A qubit is typically represented as a vector in a two-dimensional Hilbert space. In quantum mechanics, the state of a qubit can be written as a superposition of the basis states ($|0\rangle$ and $|1\rangle$)

$$|\psi\rangle=\alpha|0\rangle+\beta|1\rangle \quad (1)$$

Where $|\psi\rangle$ is the quantum state of the qubit, $|0\rangle$ and $|1\rangle$ are the basis states. α and β are complex numbers that represent the probability amplitudes of the qubit being in state $|0\rangle$ or $|1\rangle$, respectively. Entanglement is another crucial phenomenon of qubits. When two or more qubits become entangled, their quantum states are no longer independent. The state of one qubit can depend on the state of another, even across large distances. For example, consider two qubits in the entangled state:

$$|\Phi+\rangle=\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle) \quad (2)$$

This means that if we measure the first qubit and find it in state $|0\rangle$, we know that the second qubit will also be in state $|0\rangle$, and vice versa for $|1\rangle$.

Generally, drug discovery for any disease takes years and costs millions of dollars, because simulating molecules is typically hard on classical computers. When quantum computing meets machine learning, a new discipline emerges: quantum machine learning. Quantum machine learning methods, such as deep generative and discriminative GANs, CNNs, and VAE can be used to generate small drug molecules classify binding pockets in proteins, and generate large drug molecules [9]. Quantum machine learning is extensively used in chemistry to simulate the ions for finding the excited and ground states [10]. Quantum algorithms running on NISQ devices open avenues for the study of material design [11], protein folding [12], and chemical reaction dynamics [13]; these components are critical for drug discovery. The main contribution of this paper is that it provides a detailed systematic review of current methods and strategies that are being studied for drug design and discovery. It promotes interdisciplinary research in quantum computing and chemistry. This paper also highlighted the current challenges and future scope in drug design.

The paper is organized as follows. The first section is Introduction. In section 2, a comparison is made between quantum and classical simulations for drug discovery. In section 3, we have covered the literature review and previous work. Section 4 is a discussion about the approach and strategies used for this task. In section 5, current challenges and future scope are discussed. Finally, the conclusion is made in the last section.

2. Conventional ML vs Quantum ML approach

Conventional machine learning methods in drug discovery utilize deep neural networks, graph models, and generative models in conventional computing platforms. The methods can handle big data in the form of genetic sequences, clinical data, and molecular conformations. For example, deep learning platforms are conventionally utilized to predict drug-target interactions, chemically optimize structures, and determine molecular properties like toxicity and solubility [14]. Graph neural networks preserve structural information about activities like binding affinity prediction and virtual screening using the representation of molecules as graphs [15]. New drugs with targeted pharmacologic properties have

also been developed effectively using generative models such as generative adversarial networks (GANs) and variational autoencoders (VAEs). Though development has taken place, the conventional methods have not yet comprehensively accounted for the quantum component that defines molecular interactions. Figure 1 shows a deep learning-based drug discovery pipeline that demonstrates the main stages from data collection to molecule representation, model training, and prediction. Advanced neural architectures are used at each stage to ensure accurate and efficient compound screening.

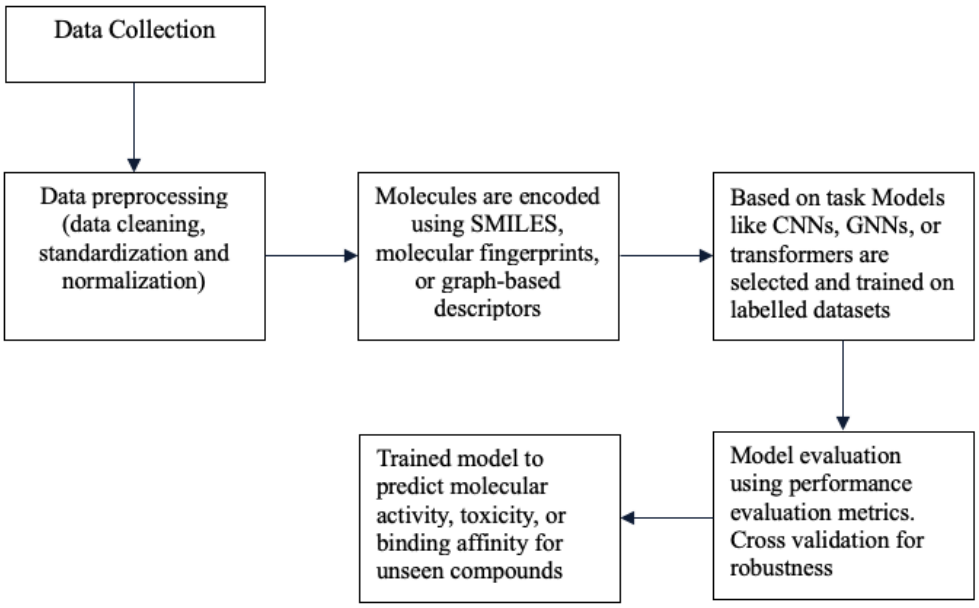

Figure 1. Deep learning-based workflow for drug discovery, from data preprocessing to compound screening

Quantum machine learning is seeking to solve some of these challenges by using the properties of quantum mechanics to enhance model processing power. Since molecules are quantum systems, quantum algorithms are better placed to simulate them than classical models. VQC and quantum kernel methods are employed to produce models that learn from small sets of data and make more accurate predictions [16]. Quantum generative models like quantum GANs and hybrid quantum-classical VAEs are under exploration to create new molecular structures by mapping structural properties to quantum states [17]. Hybrid methods that cycle quantum circuits and classical deep learning models are trying to leverage the strengths of the paradigms while avoiding the weaknesses of near-term quantum hardware. Table 1 provides a technical comparison of classical and quantum machine learning approaches in drug discovery.

While CML has come to a point of maturity and is being extensively used in real-world problems, QML holds tremendous promise in addressing problems that are challenging for traditional computers, especially in applications of complex molecular interactions. With the evolution of quantum hardware, quantum machine learning is set to become an efficient and effective tool.

Table 1: Comparison of classical and quantum machine learning approaches in drug discovery.

Aspects	Conventional ML	Quantum ML
Representation of data	Vectorized features, SMILES, molecular structures, graph structures	Quantum states using amplitude or angle encoding
Methods to be used	CNN, RNN, Transformer, VAE, GAN	Variational Quantum Circuit (VQC), Quantum Kernel, Hybrid Quantum-Classical Network
Optimization technique used	Gradient descent with backpropagation	Hybrid optimization (quantum circuits + classical optimizer)
Computation mechanism	Operates on classical bits, deterministic or stochastic model	Exploits quantum phenomena like superposition and entanglement
Tasks	Property prediction, de novo drug design, DTI prediction	Quantum-enhanced property prediction, quantum simulation of molecules, quantum generative design
Type of data	Large labelled datasets, extensive pre-processing	Potentially lower data requirements due to quantum advantage in expressivity
Issues	Scalability, interpretability, inability to model quantum interactions	Hardware limitations, noise, and the limited number of qubits.

In Figure 2 we show the workflow of quantum machine learning used in drug discovery. In the first step, we need to choose the relevant dataset such as proteins and drug-like molecules. Due to the limited number of available qubits, the dimensionality reduction technique is applied. Available data is in classical form so there is a need to convert this data into quantum states. Several encoding approaches such as amplitude or angle encoding schemes are used for this task. Quantum algorithms are applied that transform data and measurements are taken.

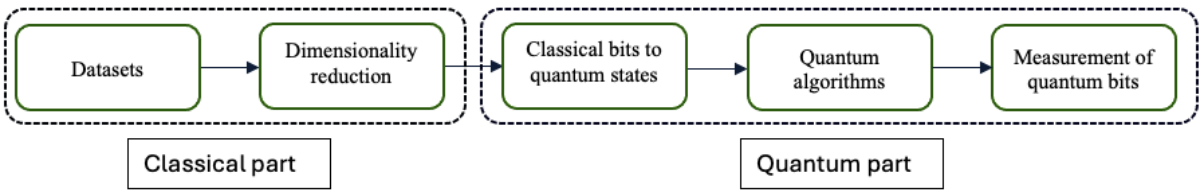

Figure 2. General process flow diagram for the hybrid quantum-classical algorithm.

3. Literature Review

Quantum computing has long been used more efficiently than traditional methods to simulate molecular systems. Aspuru-Guzik et al. [18] discussed the concept of a quantum simulator for the electronic structure problem. They showed the superiority of the quantum system over the classical approach in molecule modelling. In [19], the authors used a variational quantum eigen solver (VQE) to solve the small molecules problem, which illustrates the practical implementation of quantum algorithms in chemistry. There are three important tasks where the QML techniques play a crucial role in the field of drug discovery.

a. Molecular Property Identification: Identifying molecular properties, such as toxicity, solubility, and reactivity, is an important step for drug discovery, for which QML gives promising results. In [20], the authors applied a SMILES-based string kernel combined with a quantum support vector classifier to identify the ADME-Tox features of small molecules. They achieved ROC and AUC of 0.95 and 0.80 respectively, far better than the conventional counterpart techniques. Quantum variational ML models are widely used nowadays, but they face the problem of trainability because of vanishing gradients. To avoid this issue [21], they used Quantum Reservoir Computing (QRC) in gradient computation not required at the quantum hardware level.

a. Drug-Target Interaction (DTI) Prediction: DTI is a process in which the drug molecule binds to or affects a biological target (enzyme, receptor, or protein). Ruolan Chen et al. [22] published a research article using a quantum version of SVM combined with a Variational Quantum Classifier (VQC) for the DTI task. It shows enhanced accuracy on small datasets. In [16], the authors used a hybrid quantum-classical deep learning model for binding affinity prediction in drug discovery. This approach integrates a 3D spatial graphical CNN with an optimized quantum circuit. Quantum simulation results depict a 6% improvement in prediction performance and are more stable compared to their classical counterpart. The main aim is to find the dosage that maximizes the benefits and minimizes the serious side effects of drugs. CNN can help in the selection of a drug, but it requires a lot of data for training. So the objective is to find out the ML model that requires less data to train. A hybrid quantum neural network can be an option if the amount of training data is less. In [23], they proposed a hybrid approach of QCNN for drug effect prediction with 363 layers and 8 quantum bits applied to a cancer dataset. The obtained result surpasses the conventional ML techniques by 15%.

a. Molecular Generation and Design (de novo Drug Design): Sometimes it is necessary to generate new drug-like molecules. Generative models like GANs and VAEs are widely used to do this task but transforming them into huge chemical spaces is quite challenging. Quantum GANs and Quantum Boltzmann Machines (QBMs) are the most popular models to generate molecular structures more efficiently. Traditional GAN can identify drug candidates with the use of physical and chemical properties and show affinity to binding with the receptor for therapeutic effect. However, classical GANs struggle with training difficulties and are unable to explore specific areas of the chemical space. The enormous size of the area of search, which is made up of thousands of parameters, restricts the models' practical relevance. Even to create a small molecule with up to 9 dense atoms, a complete quantum GAN could require over 90 qubits. In [17], A hybrid quantum generator, which allows different numbers of qubits and quantum circuit layers, along with a classical discriminator, makes up the quantum GAN with a hybrid generator (QGAN-HG) model. Less than 20% of the initial parameters are needed for the QGAN-HG to learn molecular distributions as effectively as its classical version. In [24], a mixed quantum-classical framework for molecular design that blends deep generative models with quantum computation. The researchers created a compact discrete VAE with a latent layer that incorporates a Restricted Boltzmann Machine (RBM). By addressing the large and intricate structure space of drug-like molecules, this method seeks to improve the effectiveness of de novo drug discovery procedures.

4. Discussion

We are witnessing rapid growth in the number of available qubits, optimized quantum algorithms, and reduced error quantum hardware. It is widely believed that in the near-term future, we will achieve quantum supremacy in simulating molecules for drug design. From the previous studies, we have seen that Quantum Machine Learning simulation can be helpful in drug discovery and design. With the help of the expressive power of quantum circuits, we can simulate and predict the behaviour of complex molecular systems. Quantum machine learning-based simulation can significantly reduce the time and cost for drug discovery. QML can help in the optimization of chemical reactions and the study of molecular structure.

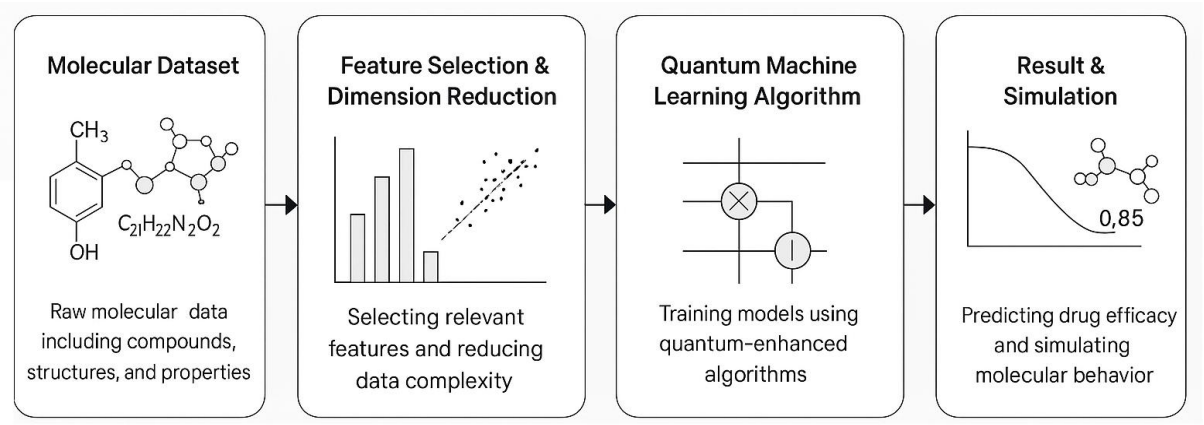

Figure 3. A detailed diagram of the process of Quantum Machine Learning (QML) for drug discovery

Fig.3 is a detailed explanation and working architecture of the procedure of the drug discovery pipeline. Quantum neural networks (QNN) and their newer variants such as QCNN and QLSTM have commonly been employed for drug discovery. QML helps in the early prediction of a molecule's potential for suitable drug candidates which is essential for reducing cost and time. The performance of the QML algorithm heavily depends on the selection of data embedding techniques [25,26]. For better performance, we must select suitable embedding methods depending on the nature of the data. Additionally, as the number of available qubits is less than the number of features in a dataset, to accommodate important features we have to apply dimension reduction policies such as PCA and Autoencoder as it also affects the performance of quantum algorithms.

5. Challenges and Future Directions

Advancements in quantum computing and its allied quantum machine learning offer significant progress in drug discovery. However, the use of QML in drug design still faces challenges that need to be entertained for efficiency.

Challenges:

- *Hardware:* A limited number of available qubits and short coherence time are restricting quantum machine learning-based simulation in drug design.
- *Noisy Qubits:* Today's available qubits are noisy, which causes significant errors in measuring qubits. To mitigate this challenge we must have more sophisticated error-correcting codes. The development of a more noise-resilient qubit system will lead to the modelling of more complex structures.
- *Data Embedding:* Drug discovery requires a large, high-quality, and curated dataset. For multi-dimensional data, more robust data embedding will be needed, such as Density Matrix Embedding Theory, to study complex molecular systems and characterize the energies of the proteins.
- *Quantum algorithms:* Many quantum algorithms such as VQE are computationally expensive and need to be optimized.

Future Scope:

- Methods to overcome non-trainability, barren plateau, and gradient computation must be explored. Curated datasets must be prepared and made available to the public.
- Quantum Generative Adversarial Networks (QGANs) and quantum variational autoencoders could design novel molecular structures with desired properties.
- Quantum Boltzmann machines can be explored for generating chemically viable drug candidates.

- Use of quantum computing to speed up molecular dynamics simulations, enabling real-time modelling of protein-ligand interactions.
- Exploration of quantum algorithms for free energy calculations is required which is crucial for drug binding affinity predictions.
- Inter-discipline research collaboration is required between experts from both fields. Open source drug discovery framework must be developed to enable innovations.

6. Conclusion

In recent years, we have seen steady growth in the use of machine learning based methods for drug discovery. Drug discovery using quantum machine learning is no longer a concept but an implementation. Though QML is in its early stage, we have seen a variety of applications in pharmaceutical and drug design, such as machine learning-based small molecule simulation and molecular force field generation. In this paper, we have summarized the study of quantum machine learning models such as QGANs, QLSTM, and VAE for drug discovery. In our study, we have found that a mostly hybrid approach is employed for this task. We conclude that the effectiveness of QML methods over classical ML for drug discovery is a topic of discussion, under which conditions QML outperforms its classical counterparts. We believe that our work will provide an overview of recent advancements in this field. We have highlighted the limitations that restrict breakthroughs in this field.

Funding source

None.

Conflict of Interest

The authors declare no potential conflict of interest.

References

[1] Shor, P. W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. SIAM Journal on Computing 1997, 26, 1484–1509, arXiv:quant-ph/9508027

[2] Grover, Lov K. "A fast quantum mechanical algorithm for database search." In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 212-219. 1996.

[3] R. P. Feynman, Quantum mechanical computers, Opt. News 11, 11 (1985).

[4] Krantz, Philip, Morten Kjaergaard, Fei Yan, Terry P. Orlando, Simon Gustavsson, and William D. Oliver. "A quantum engineer's guide to superconducting qubits." *Applied physics reviews* 6, no. 2 (2019).

[5] Bruzewicz, Colin D., John Chiaverini, Robert McConnell, and Jeremy M. Sage. "Trapped-ion quantum computing: Progress and challenges." *Applied physics reviews* 6, no. 2 (2019).

[6] Kok, Pieter, William J. Munro, Kae Nemoto, Timothy C. Ralph, Jonathan P. Dowling, and Gerard J. Milburn. "Linear optical quantum computing with photonic qubits." *Reviews of modern physics* 79, no. 1 (2007): 135-174.

[7] Graham, T. M., Y. Song, J. Scott, C. Poole, L. Phuttitarn, K. Jooya, P. Eichler et al. "Multi-qubit entanglement and algorithms on a neutral-atom quantum computer." *Nature* 604, no. 7906 (2022): 457-462.

[8] Hamid, Muhammad, Bashir Alam, and Om Pal. "Comparative Study of Quantum Computing Tools and Frameworks." In *International Conference on Innovation and Emerging Trends in Computing and Information Technologies*, pp. 87-104. Cham: Springer Nature Switzerland, 2024.

[9] Scott, Oliver B., Jing Gu, and AW Edith Chan. "Classification of protein-binding sites using a spherical convolutional neural network." *Journal of Chemical Information and Modeling* 62, no. 22 (2022): 5383-5396.

[10] Kawai, Hiroki, and Yuya O. Nakagawa. "Predicting excited states from ground state wavefunction by supervised quantum machine learning." *Machine Learning: Science and Technology* 1, no. 4 (2020): 045027.

[11] Lourenço, Maicon Pierre, Lizandra Barrios Herrera, Jiří Hostaš, Patrizia Calaminici, Andreas M. Köster, Alain Tchagang, and Dennis R. Salahub. "QMLMaterial─ A Quantum Machine Learning Software for Material Design and Discovery." *Journal of chemical theory and computation* 19, no. 17 (2023): 5999-6010.

[12] Casares, Pablo Antonio Moreno, Roberto Campos, and Miguel Angel Martin-Delgado. "QFold: quantum walks and deep learning to solve protein folding." *Quantum Science and Technology* 7, no. 2 (2022): 025013.

[13] Richings, Gareth W., and Scott Habershon. "Predicting molecular photochemistry using machine-learning-enhanced quantum dynamics simulations." *Accounts of Chemical Research* 55, no. 2 (2022): 209-220.

[14] Jin, W., Barzilay, R., & Jaakkola, T. (2018, July). Junction tree variational autoencoder for molecular graph generation. In *International conference on machine learning* (pp. 2323-2332). PMLR.

[15] Zitnik, M., Agrawal, M., & Leskovec, J. (2018). Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics*, *34*(13), i457-i466.

[16] Domingo, L., Chehimi, M., Banerjee, S., Yuxun, S. H., Konakanchi, S., Ogunfowora, L., ... & Johnson, C. (2024, September). A hybrid quantum-classical fusion neural network to improve protein-ligand binding affinity predictions for drug discovery. In *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)* (Vol. 2, pp. 126-131). IEEE.

[17] Li, J., Topaloglu, R. O., & Ghosh, S. (2021). Quantum generative models for small molecule drug discovery. *IEEE transactions on quantum engineering*, *2*, 1-8.

[18] Aspuru-Guzik, A., Dutoi, A. D., Love, P. J., & Head-Gordon, M. (2005). Simulated quantum computation of molecular energies. *Science*, *309*(5741), 1704-1707.

[19] Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., & Gambetta, J. M. (2017). Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *nature*, *549*(7671), 242-246.

[20] Bhatia, A. S., Saggi, M. K., & Kais, S. (2023). Quantum machine learning predicting ADME-Tox properties in drug discovery. *Journal of Chemical Information and Modeling*, *63*(21), 6476-6486.

[21] Beaulieu, D., Kornjaca, M., Krunic, Z., Stivaktakis, M., Ehmer, T., Wang, S. T., & Pham, A. (2024). Robust Quantum Reservoir Computing for Molecular Property Prediction. *arXiv preprint arXiv:2412.06758*.

[22] Chen, R., Liu, X., Jin, S., Lin, J., & Liu, J. (2018). Machine learning for drug-target interaction prediction. *Molecules*, *23*(9), 2208.

[23] Sagingalieva, A., Kordzanganeh, M., Kenbayev, N., Kosichkina, D., Tomashuk, T., & Melnikov, A. (2023). Hybrid quantum neural network for drug response prediction. *Cancers*, *15*(10), 2705.

[24] Gircha, A. I., Boev, A. S., Avchaciov, K., Fedichev, P. O., & Fedorov, A. K. (2023). Hybrid quantum-classical machine learning for generative chemistry and drug design. *Scientific Reports*, *13*(1), 8250.

[25] Li, J.; Alam, M.; Congzhou, M.S.; Wang, J.; Dokholyan, N.V.; Ghosh, S. Drug discovery approaches using quantum machine learning. In Proceedings of the 58th ACM/IEEE Design Automation Conference, San Francisco, CA, USA, 5–9 December 2021; pp. 1356–1359.

[26] Suzuki, T.; Katouda, M. Predicting toxicity by quantum machine learning. J. Phys. Commun. 2020, 4, 125012.