International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 12 | December - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Drug Discovery using Quantum Simulation

Revathi Meenal Kathiresan, Accenture Technology, Chennai, India

Abstract:

Health care Analysis spread wider throughout the globe enables proper structured resolutions for preventing the avalanche of health-related issues. The foremost goal lies in ensuring the proper safety and security of the patients. Talking about Patient safety, early diagnosis is the predominant key in finding any deadly diseases at an earlier stage, reducing the treatment costs, and increasing the overall survival rates. Also, an important factor that comes in parallel along with patient safety is 'Drug Discovery'. In current day scenario, Pharma industries are striving a lot to improvise clinical development productivity by examining important potential strategies. The main point of optimization lies in the 'Drug Discovery' process. Drug Discovery plays a crucial role in clinical programs, as it helps towards diagnosis of illnesses. When a drug of satisfactory evidence is administered to a target it should ensure sufficient safety of the target. One of the foremost risks here is the availability of better drugs for a disease, or in other words appropriate potency should be ensured by the drug over a target. Despite investing very huge amount over a drug, 90% of it gets failed in clinical trials. In order to overcome the above-mentioned issue with Drug discovery, many pharma industries are relying on AI technologies to ensure, the drug actively engages the target and produce expected therapeutic effect. Along AI another promising technology which is set to revolutionize a wide range of health industry is Quantum Computing. As pharma industries are in urge to provide quality efficient and easily accessible drugs within reach, Quantum computing with built in Quantum mechanics perfectly fits in to derive proper compositions of drugs by trying out various *drug-protein-gene* interactions and in forecasting drug's development using quantum algorithms and simulation techniques.

Drug Development Schemata:

Housing the huge ecosystem, our globe is balancing various factors in terms of distinct topographies, atmosphere, weather conditions, calamities etc., As for the humans are concerned, they rear many benefits out of them and by some discrepancies left in above cases they become prone to diversified diseases such as genetic, infectious, metabolic, physiological etc., With outbreak of varied diseases, R & D towards developing new molecules, drugs, biosimilars, determined laboratory tests and medical instruments started to grow on day to day basis.

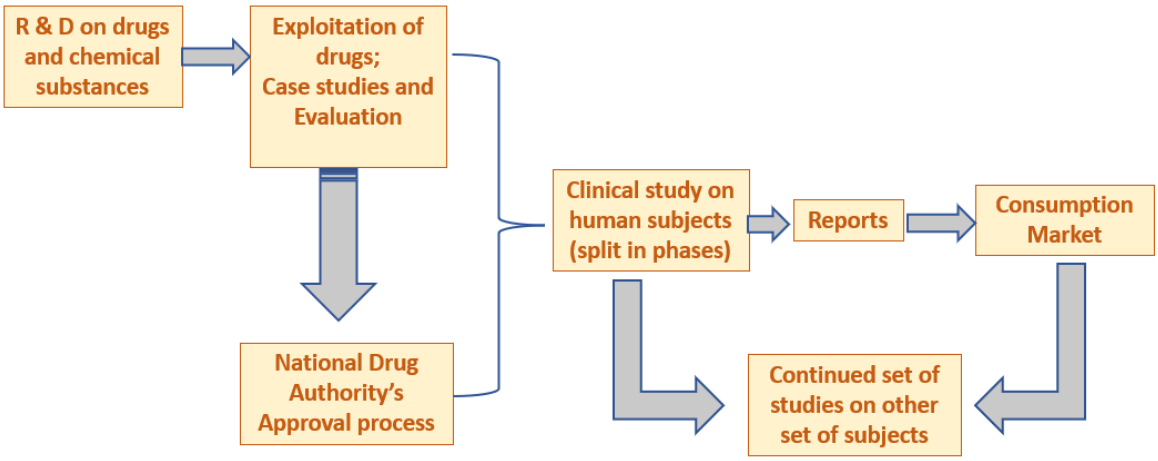

Fig. 1., Basic Schemata of Drug Development Process

Lifecycle of drug's development are never limited as they undergo chain of processes, right from R&D to manufacturing optimal drugs, performing case studies, and evaluating the outcomes. Once the outcomes are proved to be elegant, the same are approved by National Drug Authorities as Central Drugs Standard Control Organization (CDSCO), The U.S. Food and Drug Administration (FDA) etc., Next, they are properly trialed over humans or other subjects under study, and again the reports are validated before deploying them in Market and hence to make easily available for human diagnosis.

NGS:

Next Generation Sequencing (NGS) also known as high-throughput sequencing, takes care of genome sequencing or sequencing millions of DNA in parallel, to yield avalanche amount of data at once. NGS majorly involves three types of sequencing genome sequencing (*that sequence the entire genome*), exome sequencing (*that sequences the coding regions in a genome*) and targeted gene sequencing (*that sequences certain regions of interest in a genome*).

The study majorly assesses the role of every individual gene and their capability to cause diseases, and thus helps in providing distinct therapies to specific set of patients by generating data under multiple experimental factors each resulting in different observatory findings. NGS data, generated through a kind of exploratory analysis helps us to study and understand different patterns of diseases by spotting the main characteristics of genome, thereby we can validate the susceptibility of a genome towards a disease and

diagnose it further. The onset of NGS data helps in the safe and efficient discovery of drugs to treat several diseases, tailoring to the needs of patients.

R & D over drug molecules are considered best important as any adverse effects may lead to the deterioration of human population. Accurate pharmacological activities must be executed to achieve skilled profile, lowered toxicity, and non-allergic moderate drug reactions. Drug discovery process comprises of the following steps as, selecting the required target, Validating the target, compound screening and optimization of lead candidate.

Steps Involved:

1. Data Collection: Initially Genome data from a population, with the subject of interest (like genetic disease, cancer disease, dental etc.,) should be collected. Stratify the collected data with data processing techniques, by associating valid genetic information.
2. DNA data are collected through proper DNA extraction method, to obtain a high quality, robust and sufficient quantity of DNA, following the standards of DNA Extraction or Isolation protocol.
3. Design a species model – To design the species model, as a first step data input has to be collected, that includes the functional details of DNA (See Appendix 1).
4. Creating a genome model is little difficult as it may generate more than exabytes of data [1].
5. Sample data fields that have to be covered in sequence data includes locus (See Appendix 1) name (with organism info, gene product) to group similar sequences, nucleotides and their indicator values, sequence length (number of nucleotide base pairs), exons (protein coding part of DNA) and introns (non-coding portion) details, GenBank division (sequences of a specific organism like plant, rodent, bacterial, synthetic etc.,), molecular type (like RNA, DNA), protein translation identifier number and goes on.
6. Next it undergoes library preparation, where the samples will be purified and fragmented to form ‘Library’, these in turn will be ligated using adapters to form synthetic DNA sequence or DNA fragments of specified length.
7. The DNA sequence will be amplified either using Polymerase chain reaction (PCR) or by using traditional methods of DNA cloning and amplification.
8. Now the sequencing starts by reading or extracting the nucleotides one by one. These on further processing gets converted into significant data.

9. The DNA sequences obtained, helps to develop better drugs and improvised therapies for the targets and to develop tests in preventing adverse drug reactions. The sequences are used to identify the genetic variants that causes diseases in humans.
10. Parting the individual DNA sequences, drugs can be developed for specific DNA signatures found in the *subjective* cells [2] (as cancel cells, or genetically engineered stem cells etc.,), after addressing the expected and unexpected abnormalities.

Drug Discovery:

Drug discovery gets succeeded with optimal selection of chemical compounds or chemical entities in right proportion. The drug molecules should be of greatest potential in reacting towards a target process, as for instance Quercetin (a certain drug molecule) and its related flavonoids are quipped to inhibit the growth of tumor cells in a target [5].

Drug's impacts in modulating the metabolism and dispositioning any substance that are responsible for causing a disease without damaging the cell.

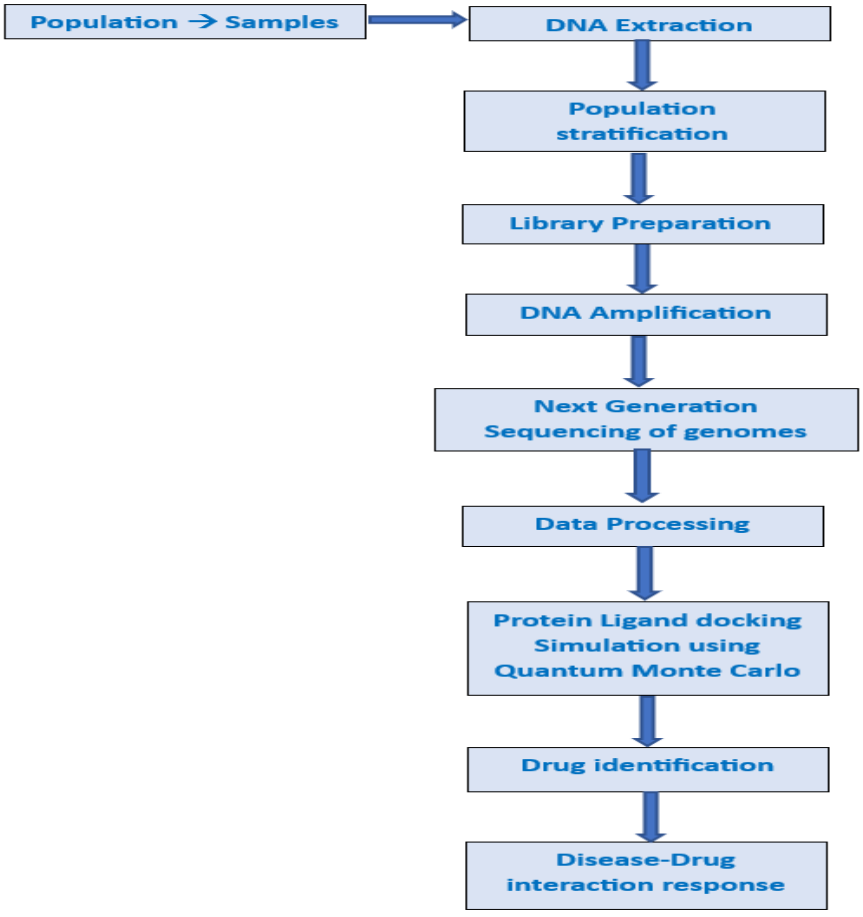

Fig 2. Drug Discovery using Quantum Simulation flow.

Quantum Computing Systems:

Quantum Computing, encapsulating the theories of Quantum mechanics and Quantum Physics, is all set to conquer the whole technical world. To manipulate and manage the quantum computing system, Quantum Computers are taking shape. Bits (that represents either zero or one) in classical computers are replaced with Quantum bits or Qubits (that represents superposition (See Appendix 1) of one and zero simultaneously i.e., it can have a zero, one or both) here. This feature makes Quantum computing the forerunner in performing any operations at an exponential speed. A unique feature with Quantum computing that favors correlation of behavior between two separate things is known as Entanglement, the relationship identified helps us to find solutions to complex algorithms.

Same as CPU, a Quantum chip will be acting as the processing unit of a Quantum computer. The chip equips the Qubits [3], which are made up of 'Quantum dots' in it, they are designed in such a manner to handle trillions of operations within the stipulated time. Using the qubits, Quantum computers run complex multidimensional algorithms. Technically quoting, Qubits also can be represented as cold atoms in optical lattices, atoms in arrays of cavities, ions trapped by electric or magnetic fields in 2D crystals, electrons in arrays of quantum dots, Rydberg atoms, polar molecules, nuclear magnetic resonance (NMR), linear and nonlinear quantum optics, superconducting circuits [4].When two qubits are entangled, changes to one qubit directly impacts the other

The main challenge in building Quantum Computer is that it must integrate thousands of Qubits to control the manipulation and readout large number of Qubits. Hence in order to develop scalable quantum computers with better electronics layout [5], it is closely integrated with classical system (Controlling the spin qubits of Quantum Computer by classical system appears to be one of main roadblock in developing scalable Quantum Computers). To make Quantum Algorithms work efficiently, the effects of noise stored on the Quantum information must be reduced, that is achieved by fault tolerant-quantum error correction (FT – QEC) method.

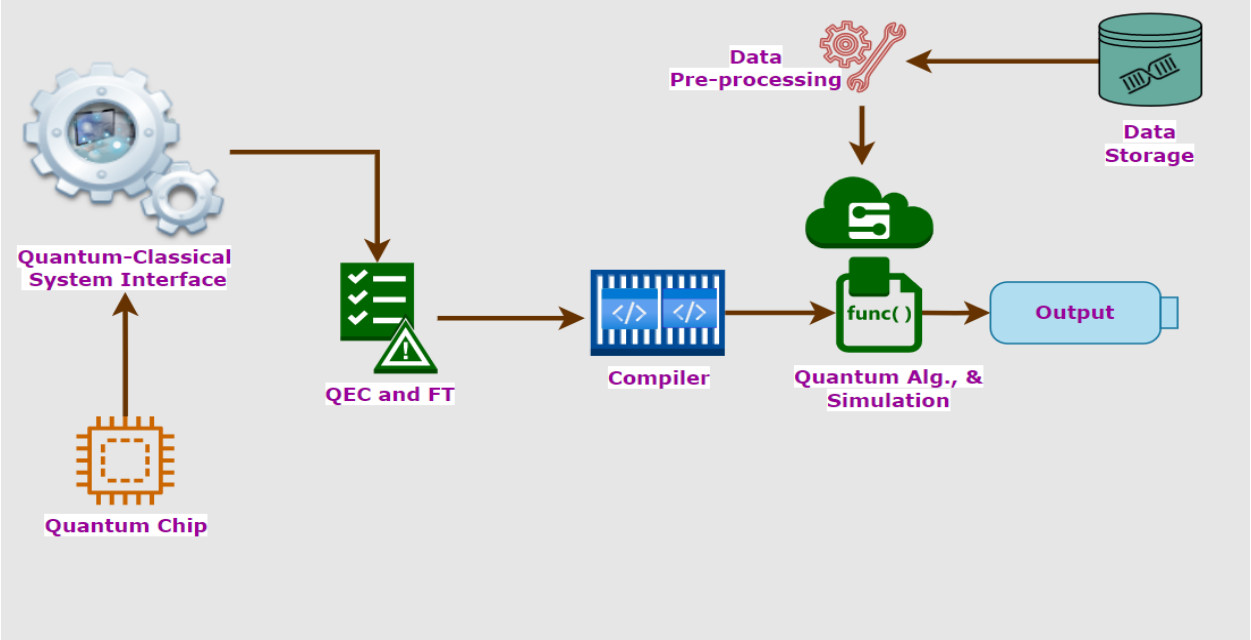

Fig 3. Quantum Computing Systems

Quantum Model and Simulation:

As mentioned above, a unique feature in Quantum computing is *Entanglement*. The phenomenon makes up an entangled pair of objects, where the state of an object seldom acts independently or predominantly cannot be independent at all, hence always the state of an object interacts with the other and this creates a range of correlation, irrespective of the separating distance or any hindrance between them. This concept gets enroute to S*uperposition - a quantum theory, which explains existence of objects in multiple states at the same time.*

Drug designing, a process that involves interactions between ligands and molecules, is modeled using 'molecular modeling techniques' – *a technique to find a region within a molecule that prefers in interacting with specific chemical groups and models those interactions made between partnered molecules*. As it involves molecular designs to a greater extent, we opt for Quantum mechanics.

Quantum Computing and simulation, leveraged with Quantum mechanics and Entanglement, performs simulation of every possible interaction between molecules in a single attempt generating multiple correlated records, overcoming the disadvantage of classical system simulation where the interaction will be staged one by one.

Simulating the interaction is of utmost importance since drugs don't act by themselves, they get coupled with other molecules (either of same or different chemicals) or enzymes and accordingly change their behavior. This kind of interaction is not that simple to be achieved, hence quantum simulation helps to get better prediction and accuracy by trying out '*n*' number of interactions. For E.g., if an operation following the mundane principles considers five different methods to arrive at an optimal solution using classical system, quantum simulation gives a try for over 50 or more methods to arrive a solution with greater accuracy.

A point of consideration in using Quantum Simulation is that the calculations involved are rather computationally expensive, as almost it tries out every possible combination of molecules, thus the quantum states grows exponentially whereas classical does linearly. Of many simulation methods available Monte Carlo Simulation are known best for conducting correlation studies and thus various relationships can be better predicted. In our context of Drug delivery, Monte Carlo Simulation – *a proposed method of simulation procedure* (See Appendix 1) *requires performing Y number of simulations, and each simulation has to calculate X approximations of states, thus making in total of XY state approximations for a quantum system. MSE can be used to overcome the systematic errors.*

Rolling back to our procedure, following NGS, the sequence data extracted has to be processed to remove tainted and contaminated DNA and map the sequence to a reference genome to attain the degree of association between the extracted and the existing. The processed DNA helps in the formulation of drug and thereby pharmacological action (*studies biochemical interaction of drug molecules*) of drugs over receptors or by the inhibition of enzymes can be simulated using Quantum Computing. When the ligands bind to a receptor and produce a response it's called as agonists (See Appendix 1) and when it does not produce any response it is known as antagonists.

- The simulation model consists of an experimental setup, which in our case is Quantum system, NGS modeling and Simulation technique. As a result of simulations, multiple correlations of DNA binding protein interactions are retrieved, that helps in attaining better predicted drug discovery process to treat genetic and other kind of diseases.
- Normally, using quantum mechanics *ligand binding* modeling takes place. As already Quantum computing works by the principle of quantum mechanics, exploring binding mechanism methods can be skipped as it is taken care by entanglement process.
- The protein ligand docking[6] (See Appendix 1), one of molecular modeling technique favored with Quantum mechanics, is simulated (*with techniques similar to the assumption methods of Quantum Monte Carlo Simulation*) to find the optimal binding between the ligand (a small molecule) and the

protein by generating multiple approximated quantum states of data, and the resulting predictions leads to the identification of potential drug candidate; the robustness of the bindings can be evaluated using root mean square deviation or mean square error.

- Finally, the modeling and simulation performed over the population study and the impacts on drug-disease interactions can be recorded to extract the response of the *gene* (towards a disease) against a *drug*.

Target x response	Drug *a*	Drug *b*	Drug *c*
Disease *x*	*Impact*	*No Impact*	
Disease *y*			*Correlated Impact* (See Appendix 1)

Table 1. Drug-Disease Interaction response

- The study can also be further included for *drug – gene* interactions, where the protein-drug interactions, in other words the pharmacological action of a ligand towards to a receptor can be bought by the actions of agonists and antagonists.

Conclusion:

Quantum Computing, expected to reach the market by 2030, currently set onto heap a lot of expectation to achieve fault tolerant computing at an exponential rate. Standing by the context of Drug Discovery, Quantum Simulation are in scales of research by diverse pharma industries to endure better throughput of drugs over targets, particularly in studying the effects of molecular interactions. Focusing the importance of Drug discovery, many tools and platforms are already proposed in the market which relies on simulation techniques, but the success rate is not so high as it is cost effective in incorporating all known proteins and molecule interactions into a classical simulation. Quantum system research, as exhibited by many Organizations appears to be a promise in reforming drug development process. *On a concluding note*, practically when the Quantum computers involving many challenges takes shape, algorithms and simulations can be staged in them much similar to that of classical computers. But the complexity in building out algorithms and computations involved in simulations must be focused from a much different exceptional angle, as the Quantum systems both by physical (temperature) and functional (wave interference techniques) aspects, are susceptible to noises and error at a faster rate, hence the algorithmic approach or simulation approach should be augmented a lot to achieve supreme accuracy, as expected from

Quantum systems. If proper expertise is meant for designing, then there is no way in compromising the technology from stepping into an unexpected era of computing.

References:

[1] https://www.genome.gov/about-genomics/fact-sheets/Genomic-Data-Science

[2] Institute of Medicine (US). Genome-Based Therapeutics: Targeted Drug Discovery and Development: Workshop Summary. Washington (DC): National Academies Press (US); 2012. 4, Emerging Technologies in Drug Development. Available from: https://www.ncbi.nlm.nih.gov/books/NBK116452/

[3] TY - JOUR
AU - Mouny, Pierre-Antoine
AU - Beilliard, Yann
AU - Graveline, Sebastien
AU - Roux, Marc-Antoine
AU - Mesoudy, Abdelouadoud
AU - Dawant, Raphael
AU - Gliech, Pierre
AU - Ecoffey, Serge
AU - Alibart, Fabien
AU - Pioro-Ladriere, Michel
AU - Drouin, Dominique
PY - 2023/04/01
SP - 1
EP - 7
T1 - Memristor-Based Cryogenic Programmable DC Sources for Scalable In Situ Quantum-Dot Control
VL - PP
DO - 10.1109/TED.2023.3244133
JO - IEEE Transactions on Electron Devices
ER -

Available from:
https://www.researchgate.net/publication/368672032_Memristor-Based_Cryogenic_Programmable_DC_Sources_for_Scalable_In_Situ_Quantum-Dot_Control

[4] *ACS Eng. Au* 2022, 2, 3, 151–196
Publication Date: January 25, 2022
https://doi.org/10.1021/acsengineeringau.1c00033
https://pubs.acs.org/doi/10.1021/acsengineeringau.1c00033

[5] A cryogenic CMOS chip for generating control signals for multiple qubits - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-quantum-classical-interface-of-a-quantum-computer-a-The-generic-stack-of-elements_fig4_348758110 [accessed 2 Jul 2023]

[6] Pinzi L, Rastelli G. Molecular Docking: Shifting Paradigms in Drug Discovery. Int J Mol Sci. 2019 Sep 4;20(18):4331. doi: 10.3390/ijms20184331. PMID: 31487867; PMCID: PMC6769923.
Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6769923/

APPENDIX 1: Glossary

DNA: Every Organism's cell consists of nucleus, with strands of nucleotides known as DNA. Collectively a complete set of DNAs in an organism is Genome and Gene is a portion of DNA that helps in encoding a species traits or characteristics. Gene contains the information to make functional molecules known as proteins.

LOCUS: Represents the specific location of a gene in a chromosome. Locus as home to gene, will appear in different forms known as Allele. The complete set of alleles in an individual organism are known as Genotype, whereas the physical or observable or identifiable characteristics in an organism is known as Phenotype. Phenotypes are simply influenced by Genotypes and gets disturbed or impacted by environmental factors such as stress, climate, nutrition etc.,

SUPERPOSITION: Allows the quantum objects to exist in more than one state.

RECEPTORS: Specific type of protein situated in cell membranes. For each type or receptor there is a specific group of drugs or endogenous substance (called ligands) that binds to a receptor and produce a pharmacological effect.

AGONISTS: Ligands/Drug substance gets bind to a receptor and produce an appropriate response; ANTAGONISTS: Ligands/Drug substance gets bind to a primary site or another site of receptor to stop receptor from producing an appropriate response.

CORRELATED IMPACT: Degree of association will be increasingly/decreasingly positive or increasingly/decreasingly negative.

PROTEIN LIGAND DOCKING: Normally protein-ligand docking is carried out using Monte Carlo Simulation, when using Quantum computing it can be replaced with Quantum Monte Carlo Simulation methods.