wjpmr, 2023,9(5), 128-134

SJIF Impact Factor: 5.922

WORLD JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.wjpmr.com

Review Article

ISSN 2455-3301

WJPMR

QUANTUM COMPUTING- AN EMERGING TECHNIQUE FOR SPEEDY DRUG DISCOVERY AND DEVELOPMENT

Adarsh Sahu, Aarti Rai, Aditya Choudhary, Aasim Hussain and Dr. Vivek Jain*

Adina Institute of Pharmaceutical Sciences, NH86A, Lahdara, Sagar, MP, 470001.

***Corresponding Author: Dr. Vivek Jain**

Adina Institute of Pharmaceutical Sciences, NH86A, Lahdara, Sagar, MP, 470001.

Article Received on 08/03/2023 **Article Revised on 28/03/2023** **Article Accepted on 18/04/2023**

ABSTRACT

Quantum computing hardware and software have made enormous strides over the last years. Questions around quantum computing are impact on research and society has changed from "if" to "when/how". Classical computing works by processing bits, or 0s and 1s representing electrical signals of on and off. Quantum computing employs a very different technique for information processing. It uses qubits, which can exist as both a 1 and 0 at the same time,and uses the properties of subatomic particles in quantum physics such as interference, entanglement, and superposition to extendcomputational capabilities to hitherto unprecedented levels. The efficacy of quantum computing for important verticals such as healthcare where quantum computing can enable important breakthroughs in the development of life-saving drugs, performing quick DNA sequencing, detecting diseases in early stages, and performing other compute-intensive healthcare related tasks isnot yet fully explored. Involvement of multiple stakeholders in drug discovery and development, even the simple healthcare problems become complex due to classical approach to treatment. In the Covid-19 era where quick and accurate solutions in healthcare are needed along with quick collaboration of stakeholders such as patients, insurance agents, healthcare providers and medicine supplier etc., a classical computing approach is not enough. A semi-structured interview approach is adopted to gauge the expectations of professionals from drug discovery and development regarding quantum computing. A structured approach of coding, using open, axial and selective approach is adopted to map the themes under quantum computing for drug discovery and development. The findings indicate the potential applications of quantum computing for pharmaceuticalalong with patients to have precise and quick solutions to the problems, where greater accuracy and speed can be achieved. Existing research focuses on the technological background of quantum computing, whereas this study makes an effort to mark the beginning of quantum computing research with respect to organizational management theory.

KEYWORDS: Quantum computing, Healthcare services, Qubits, High performance.

INTRODUCTION TO QUANTUM COMPUTING

Quantum Computing (QC) is underpinned by quantum mechanics, and hence often explained through concepts of superposition, interference, and entanglement. In quantum physics, a single bit can be in more than one state simultaneously (i.e. 1 and 0) at a given time, and a QC system leverages this very behavior and recognizes it as a qubit (Quantum bit). Having roots in quantum physics, QC has the potential of becoming the fabric of tomorrow's highly powerful computing infrastructures, enabling the processing of gigantic amounts of data in real time. Quantum computing has recently seen a surge of interest by researchers who are looking to take computing prowess to the next level as we move past the era of Moore's law, however, there is a need for an in-depth systematic survey to explain possibilities, pitfalls, and challenges.[1]

Quantum computing has rapidly advanced in recent years due to substantial development in both hardware and algorithms. These advances are carrying quantum computers closer to their impending commercial utility. Drug discovery is a promising area of application that will find a number of uses for these new machines.[2] As a prominent example, quantum simulation will enable faster and more accurate characterizations of molecular systems than existing quantum chemistry methods. Furthermore, algorithmic developments in quantum machine learning offer interesting alternatives to classical machine learning techniques, which may also be useful for the biochemical efforts involved in early phases of drug discovery. Meanwhile, quantum hardware is scaling up rapidly into a regime where an exact simulation is difficult even using the world's largest supercomputers. We review how these recent advances can shift the paradigm with which one thinks about drug discovery, focusing on both the promises and caveats

associated with each development. In particular, we highlight how hybrid quantum-classical approaches to quantum simulation and quantum machine learning could yield substantial progress using noisy-intermediate scale quantum devices, whereas fault-tolerant, error-corrected quantum computers are still in their development phase.[3]

Drug discovery is the process of developing a drug from an initial hypothesis to a fully commercialized product. This process can often take more than a decade and billions of dollars in expenditure before a molecule can be recognized as a drug.[1] A significant portion of these resources is invested in the identification of molecules that exhibit significant medicinal activity against a disease, usually referred to as a hit. Most of the research in drug discovery focuses on hits of low molecular weight (< 900 Daltons, with sizes of 1 nm or less[2]), which constitute around 78% of the drug market.1 In this case, the medicinal activity of a particular drug candidate or ligand is associated to its ability to bind to a biological target, usually a protein, whose activity regulates the metabolism of the disease. Typically, the first stage in the discovery process is to generate a library of potential drug candidates that is subsequently screened based on medicinal activity to identify hits.[3] Along with the activity, other factors that determine the efficacy and potency of the hits, such as the absorption, distribution, metabolism, excretion, and toxicity (ADMET) profile, among other pharmacokinetic properties, are optimized to produce a smaller set of better candidates called lead compounds.[4] Further screening and optimization generally delivers a small set of leads that proceed through the stages of drug development and clinical trials before one of them becomes a viable commercial product. Traditionally, the search for hits was accomplished by high-throughput screening (HTS) on large molecular libraries using in vitro activity experiments. These searches generally have low hit rates and require the synthesis of a large number of compounds, which in turn demands a significant investment in resources and time.[4] This approach was completely transformed by the advent of commercial computers in the 1970s and 1980s, which enabled computational chemistry and statistical analysis, among other tools[5], to accelerate HTS, improve the hit rate, and increase the quality of the leads obtained in the process.[4] The increase in computational power and the improvement of computational chemistry techniques fostered the practice of computer-aided drug design (CADD), which constitutes a significant portion of the drug discovery pipeline today.[3] The ultimate goal of CADD is to answer the inverse-design question: What are the best chemical structures associated with a desired therapeutic effect?.[6,7] To answer this question accurately, CADD faces two main challenges: first, the accurate simulation of the interaction of drug candidates with biological targets, and second, accurate statistical modeling of activities and ADMET profiles based on the available simulated and experimental data. The former is largely constrained by the computational cost of simulating the physics of molecular systems for both small molecules and biological targets. The latter is constrained by the effectiveness of existing statistical techniques. Quantum computing could potentially shift the paradigm with which one thinks about quantum chemical simulation. By efficiently preparing highly entangled states that are otherwise intractable to describe on classical computers, quantum computers can perform certain important quantum chemistry and machine learning tasks in ways that are beyond the ability of classical computers. Furthermore, efficient manipulation of quantum states also allows for certain linear algebraic operations to be performed far more efficiently than what is possible with classical devices.[4] With these unique abilities, quantum computing promises to deliver efficient and highly accurate solutions to otherwise intractable problems, for instance finding the ground state energy of a molecular system,[8] As we discuss in detail in Section 3, a common method for treating electronic structure calculation on a quantum computer is via second quantization, where an electronic state over N spin orbitals is represented using N qubits—one qubit for each spin orbital. In the coming years, we are anticipating quantum devices with N > 50 qubits,[9,10], making it possible to map onto a quantum computer problems whose exact solution (say, via exact Hamiltonian diagonalization) is beyond current classical computation. Quantum machine learning is also a rapidly emerging field exploring how quantum computers can perform machine learning tasks with improved performance over classical computers.[11] As we discuss in Section 3.2, there are plausible reasons to believe that quantum computers may enable solutions to machine learning tasks that are beyond classical computation. Pinpointing the precise regimes of quantum advantage is a main mission of the field of quantum machine learning. In this paper, we review developments in quantum computing relevant to drug discovery through quantum chemistry and machine learning, outlining the promises as well as caveats. This paper is organized as follows: First, we describe the general pipeline of CADD and some of the methodologies employed in the industry and their challenges. Second, we outline some of the latest quantum computing algorithms that we consider relevant for CADD, namely quantum simulation and quantum machine learning. Finally, we share our perspective on how these methods could benefit CADD by addressing some of its biggest challenges. The purpose of this perspective is to initiate a mutually beneficial cross-disciplinary discussion and collaboration between the fields of CADD and quantum computing.[5] For the quantum computing community, such dialogue will help to outline the practically useful regimes where quantum computers may have an advantage over classical counterparts. For the drug discovery community, our hope is to bring an alternative perspective on classical computing for solving some of the crucial computational problems, which arise in practice. Our approach is by no means exhaustive, and for more in-depth discussion of

specific technical subjects, the reader is encouraged to refer to the relevant citations.[6]

Pharma's focus on molecular formations makes it well suited for QC

Identifying and developing small molecules and macromolecules that might help cure illnesses and diseases is the core activity of pharmaceutical companies. Given its focus on molecular formations, pharma as an industry is a natural candidate for quantum computing.[7] The molecules (including those that might be used for drugs) are actually quantum systems; that is, systems that are based on quantum physics. QC is expected to be able to predict and simulate the structure, properties, and behavior (or reactivity) of these molecules more effectively than conventional computing can. Exact methods are computationally intractable for standard computers, and approximate methods are often not sufficiently accurate when interactions on the atomic level are critical, as is the case for many compounds. Theoretically, quantum computers have the capacity to efficiently simulate the complete problem, including interactions on the atomic level. As these quantum computers become more powerful, tremendous value will be at stake.[8]

QC could make current CADD tools more effective by helping to predict molecular properties with high accuracy. That can affect the development process in several ways, such as modeling how proteins fold and how drug candidates interact with biologically relevant proteins. Here, QC may allow researchers to screen computational libraries against multiple possible structures of the target in parallel.[9] Current approaches usually restrict the structural flexibility of the target molecule due to a lack of computational power and a limited amount of time. These restrictions may reduce the chances of identifying the best drug candidates.[10] In the longer term, QC may improve generation and validation of hypotheses by using machine-learning (ML) algorithms to uncover new structure-property relationships. Once it has reached sufficient maturity, QC technology may be able to create new types of drug-candidate libraries that are no longer restricted to small molecules but also include peptides and antibodies. It could also enable a more automated approach to drug discovery, in which a large structural library of biologically relevant targets is automatically screened against drug-like molecules via high-throughput approaches.[11]

One could even envision QC triggering a paradigm shift in pharmaceutical R&D, moving beyond today's digitally enabled R&D toward simulation-based or in silico drug discoveries—a trend that has been seen in other industries as well.[12]

The following QC use cases apply to different aspects of drug discovery and will emerge at different points over an extended timeline. All of them, however, may enable more accurate and efficient development of targeted compounds.[13]

Overview of computational methods in drug discovery

At the risk of gross simplification, we summarize the overall drug discovery process. The usual drug discovery pipeline requires the identification and characterization of a suitable biological target, which can be effectively proved to intervene in the mechanism of disease. This step often requires intense experimentation as well as extensive statistical analysis of the collected data. Once a biological target is in place, the next step is the search for hits, which usually involves extensive biological and virtual screening over libraries of molecules, or the generation of completely new compounds (de novo design), which must be synthesized and tested.[14] The group of hits collected on this stage undergoes further optimization of the pharmacokinetics and ADMET properties, involving a combination of biological and in silico tests, to generate the final group of leads. These stages, going from target identification to lead optimization, benefit the most from CADD techniques.[12,13] The subsequent steps in the drug discovery pipeline, which involve clinical studies in animals and humans prior to the Federal Drug Administration review and approval, are less intensive in the use of CADD tools but might require further rounds of lead optimization. The final success of a drug discovery campaign depends, to a great extent, on the quality of the CADD approaches applied in the early stages. CADD approaches employed on the stages of hit search, lead discovery, and lead optimization are generally classified into two main categories.[4]

CADD relies on knowledge of the target protein three-dimensional (3-D) structure to predict the ability of a candidate to bind to the target, whereas ligand-based CADD employs information of known active and inactive molecules to predict the activity of new candidates. Structure-based CADD is preferred over ligand based if the structural information of the biological target is available. This information is usually obtained experimentally using nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography studies on crystallized protein.[4] Predicting the protein structure from the knowledge of the amino acid sequence requires simulating the protein folding process, which is so far out of reach except for small peptides and fast folders.[14]; however, in the absence of experimental structures, it is still possible to approximate the 3-D structure of an unknown target protein by comparing its sequence with related known proteins, a process known as comparative modeling.[15] Along with the structure, it is necessary to characterize the target by identifying the binding (active) sites that are responsible for the biological activity and where the potential drug candidate (ligand) is expected to bind. Assuming that a model for the target structure is available, structured-based CADD approaches attempt to

find suitable drug candidates by analyzing the interaction between the candidate (ligand) and the biological target, generally a protein; therefore, most CADD approaches require the following: first, determining the pose or conformation of the ligand that fits best the binding site of the target, and second, assigning a numerical score that expresses the strength of the interaction of the ligand-target complex.[16] The process of finding the best conformation is generally called docking and the process of computing the affinity is referred as scoring.[4] These procedures are generally intertwined since docking requires a score function that ranks different conformations according to their ability to form bound ligand-protein complexes. Extensive sampling of conformations is often required in structured-based CADD approaches to account for the mobility of protein and ligands in biological conditions (aqueous solutions at room temperature).[17] When no information of the 3-D tertiary structure of the protein can be obtained, ligand-based CADD is the main tool. In this case, the selection of the candidates proceeds by comparison of the structures with a set of known active ligands using molecular similarity indexes and by evaluation of the activity using a quantitative.[15]

Target identification and validation

During target identification, QC can be leveraged to reliably predict the 3-D structures of proteins. Obtaining high-quality structural data is a lengthy process often leading to low-quality results. Despite all efforts, researchers have yet to crystallize many biologically important proteins—be it due to their size, solubility (for example, membrane proteins), or inability to express and purify in sufficient amount. Pharma companies sometimes develop drugs without even knowing the structure of a protein—accepting the risk of a trial-and-error approach in subsequent steps of drug development—because the business case for a given drug is potentially so strong.[16]

AlphaFold, developed by Google's DeepMind, was a breakthrough in AI-driven protein folding but has not resolved all of the challenges of classical computing-based simulation, including, for example, formation of protein complexes, protein-protein interactions, and protein-ligand interactions. It's the interactions that are most difficult to classically solve and, thus, may benefit from QC, which allows for the explicit treatment of electrons. Additionally, QC may allow for strong computational efficiencies here given that Google's AI model—which is trained on around 170,000 different structures of protein data—requires more than 120 high-end computers for several weeks.[17]

Hit generation and validation

QC's ability to parallel process complex phenomena would be particularly valuable during hit generation and validation. With existing computers, pharma companies can only use CADD on small to medium-size drug candidates and largely in a sequential manner. Computing power is the bottleneck. With powerful enough QC, pharma companies would be able to expand all use cases to selected biologics as well, for instance, semi-synthesized biologics or fusion proteins, and perform in silico search and validation experiments in a more high-throughput fashion. This use case would go beyond the identification of the protein and eventually encompass almost the entire known biological world. With a robust enough hit-generation and validation approach, this step would already deliver potential lead molecules that are much easier and quicker to optimize. The process of hit search generally involves HTS of a database of candidate compounds. Traditionally, this process has required the synthesis and experimental determination of the activity of the compounds, which is extremely expensive and slow. Nowadays, the process is accelerated using virtual HTS (vHTS). Different score functions are employed to rank the activity of the candidates depending on whether a structured-based or ligand-based approach is used. Some ligand-based approaches score the candidates based on their similarity with a set of known active compounds. Another option is QSAR, which constructs a statistical model based on experimental information of the activities and chemical information of the ligands. In both approaches, the chemical information is expressed with molecular descriptors that encode physicochemical and structural information of the molecules in a digital format, suitable for comparison. Molecular descriptors can be generated by knowledgebased, graph-theoretical, molecular mechanical, or quantum-mechanical methods . Arguably, the most popular descriptors are molecular fingerprints, which encode various molecular properties as predefined bit settings. Other descriptors are computed solely from the 2-D or 3-D topology of the molecule based on graph-theoretical methods.[18]

Lead optimization

During lead optimization, which is a top-three parameter to improve R&D productivity,[1] QC may allow for enhanced absorption, distribution, metabolism, and excretion (ADME); more accurate activity and toxicity predictions for organ systems; dose and solubility optimization; and other safety issues. Once hit compounds have been identified, they enter an optimization phase to produce a smaller set of better candidates, called leads. The set of leads undergoes further optimization in a process that iterates between CADD development and in vitro and animal experiments.[4] The purpose of this second phase of screening is to optimize the druglike properties of the hit compounds, which includes not only the biological activity, but also the ADMET profile and other pharmacokinetic properties. The general assumption behind this process is that small changes on the chemical structure will produce incremental changes of the druglike properties; therefore, the optimization involves the synthesis of the drug-candidates along with testing of their biological activities accompanied by CADD.[19]

Target identification and characterization

The initial stage of drug discovery concerns collecting evidence of therapeutic effects in activation or inhibition of certain biological pathway associated to a disease. The biological entity responsible for such response is called the target, which is generally a protein. In a broader sense, the term could also refer to the genes or RNA associated to the protein. Ideal targets should be "druggable," meaning that the drug candidate should be able to access the target and affect a biological response that is measurable in vitro and in vivo. The identification of suitable targets and their corresponding validation by studies of the mechanism of action increases the chance of success during the discovery process and allows one to foresee side effects associated to the modulation of the target.[12] Traditional approaches to target identification employ chemical proteomic techniques, such as affinity chromatography, biochemical fractionation, and radioactive ligand binding assays.[21] These methods employ a small molecule with proven activity to isolate the target from a mixture of other proteins. In the case of affinity chromatography, the most widely used approach, the active compound is immobilized in a porous matrix. Subsequently, a solution containing the protein mixture is passed through the matrix, and those proteins that bind to the immobilized active compound are retained. In the final stage, the retained proteins, which correspond to the potential targets, are eluted from the matrix.[22–24]

Data linkage and generation

The metalevel of R&D very much consists of linking appropriate data together—for instance, creating sensible connections between data points through effective (semantic) management. The more complex the biological information that can be processed, the more extensive the graphs that inform the drug discovery research process become. There is currently research on "topological data analysis" under way that aims to identify "holes" and "connections" across large data sets.[2] This may at some point enable R&D specialists to identify concrete cases and "industry verticals" where such algorithms are applicable, for example, in identifying connections across brain cells in response to a drug.[19]

Moreover, QC could be used to "deepfake" missing data points throughout the research process, that is, generate a type of fake data by using ML algorithms. This could be particularly useful wherever there is a scarcity of data, such as in rare diseases that can then be mitigated through artificial data sets. QC will set a new bar here regarding speed in training ML models, amount of initial data needed, and level of accuracy.[20]

Clinical trials

Clinical trials could be optimized through patient identification and stratification and population pharmacogenetic modeling.[3] In trial planning and execution, QC could optimize the selection of the trial sites. QC could also augment causality analyses for side effects to improve active safety surveillance.[21]

Beyond research and development

While the potential value of QC in pharma R&D is immense, it will also likely play a role further down the value chain. In the production of active ingredients, QC may aid in the calculation of reaction rates, optimize catalytic processes, and, ultimately, create significant efficiencies in the development of new product formulations. In the business-related value pools, QC in pharma could include the optimization of logistics (for instance, the optimization of on-site flows of materials, heat, and waste in production facilities) and improvements in the supply chain. Finally, toward market access and commercial, QC may even enable automatic drug recommendations.[22]

Quantum computing

Most digital devices use bits as the building blocks for information processing. Each bit expresses a discrete, "classical" state of 0 or 1. Devices that perform computation by manipulating bits are referred to as classical computers. Quantum computers manipulate quantum states of matter for performing computation. A standard choice for constructing those quantum states is to combine two-level quantum systems called qubits. By manipulating the qubit states and taking advantage of uniquely quantum-mechanical phenomena, such as superposition and entanglement, quantum computers can perform computational tasks in ways that are beyond what is possible on their classical counterparts. A predefined way of manipulating quantum states to solve a computational problem is referred to as a quantum algorithm. In many cases, by analyzing the number of steps that quantum algorithms take, it can be proved that they outperform classical algorithms for specific problems with reduced number of steps required. This capability is known as quantum speedup. Well-known examples of quantum speedup include Shor's algorithm for factorization, Grover search, and simulating quantum systems. On the experimental side, a wide variety of physical systems have been explored as candidates for quantum computers. Some of the hardware platforms, such as ion traps and superconducting qubits, have been scaling up rather rapidly in recent years toward the threshold regime, beyond which it becomes intractable to simulate these physical systems with a classical computer. Recent results in both theory and experiments have pointed to how the so-called noisy intermediate-scale quantum (NISQ) devices[20] with moderate numbers of qubits can in principle produce quantum states whose measurement outcomes follow distributions that are justifiably hard to sample from on a classical computer.[23]

Quantum machine learning

In recent years, there has been a rapidly expanding area of research seeking quantum techniques for enhancing machine learning methods. Although the full extent to

which quantum computers can provide advantages on machine learning is far from known, there are a few heuristic arguments to support the belief that such advantages may exist. (1) Quantum computers can generate quantum states that give rise to probability distributions that are justifiably hard to sample from classically. Because of this ability to generate statistical patterns that are hard to generate classically, one hopes that quantum computers may also be able to recognize patterns in data that are hard to recognize classically.[11] (2) For a physical system of n qubits, the space in which the quantum state of the n-qubit system dwells has dimension 2n. Such exponential size may allow for an exponentially more compact encoding of classical information. For instance, a quantum state of merely 30 qubits can represents a unit vector of length 230 ¼ 1;073;741;824. In some cases, processing the 30-qubit state for machine learning purposes may be more advantageous than treating a vector of more than a billion entries on classical hardwares. (3) Many (classical) machine learning algorithms involve a large amount of linear algebraic operations, whereas quantum computers are known to provide speedups in problems related to some of the most elementary linear algebraic operations, such as Fourier transforms, vector inner products , matrix eigenvalues and eigenvectors, and solving linear systems of equations. The quantum techniques therein can be used as a toolkit for building quantum machine algorithms.[24]

Opportunities for quantum computing in drug discovery

The potential to efficiently deliver quantum chemical calculations with accuracy comparable to FCI methods and find solutions to optimization problems can impact several of the areas of CADD described before. Here, we outline a few potential use cases for quantum computing an important part of the input concerns the structure of the target protein. Some progress has been made in the past decade on quantum techniques for protein folding based on the amino acid sequence. In particular, the quantum computing community has considered two simple models: the hydrophobic-polar model and the Miyazawa–Jernigan model, both of which model the protein as a self-avoided walk on a lattice. Solutions in both quantum annealers and gate-model quantum devices4 have been explored. Current capability of these methods is limited to proof-of-concept examples, such as Chignolin and TrpCage, small peptides with less than 21 amino acid residues. Significant venture capital has been invested in quantum computing for life sciences,5 and future work will further unveil the magnitude of quantum advantage for larger protein folding problems as quantum devices scale up. For molecular docking, one of the prevalent methods is atomistic modeling, which relies on force-field simplifications whose parameters need to match with quantum-mechanical calculations. With the advent of VQEs and the quantum phase estimation algorithm, the size of physical systems that can be treated with accurate ab initio quantum calculations will be greatly expanded as quantum devices scale up. This allows for force-field constructions based on exact quantum calculations for molecular fragments that are larger than what can be handled using existing quantum chemical methods. In de novo design, one of the pressing issues is synthesizability of a drug candidate, which involves simulation of different reaction paths. Quantum computers offer an avenue to potentially tackle electronic structure problems in the strongly correlated regime using, which would allow us to simulate transition states and thermodynamic properties to accuracies comparable to FCI methods. As a result, this can improve the effectiveness of de novo design. One of the bottlenecks for vHTS is the efficiency and accuracy with which one can calculate the scoring function. Ideally, the scoring function should be directly based on binding affinity, which comes from ab initio quantum-mechanical calculations, whereas in practice, empirical approximations are used. Hence, with the quantum subroutine boosted by quantum computers, one may evaluate the scoring functions more efficiently and accurately. This could be achieved using methods where different parts of the system are computed with different levels of approximations, such as QM/MM. The ability of computing binding affinities will also have a major impact on the lead optimization phase of drug discovery and mechanism of action studies, where understanding and quantitatively predicting the interaction of a drug candidate with multiple biological targets provides clues into toxicity, pharmacokinetics, and multitarget action. For ligand-based drug discovery, QSAR models have, in many cases, incorporated quantum-mechanical properties. Generally, the quality and accuracy of these properties significantly affect the quality and predictivity of the model. Most of these approaches use descriptors derived from DFT calculations, and quantum computation could serve as a more efficient and accurate alternative for those calculations. Another major aspect of QSAR is statistical and machine learning models. For example, in virtual screening, a common technique for classification in chemical space is by using kernels that map molecular structures to highdimensional features (see, for example, the “graph kernel” that has been used in cheminformatics literature). Commonly evaluating the kernel function requires handling vectors of extremely high dimensions, making computational efficiency a major issue for deploying kernel-based classification methods. In chemoinformatics, kernels accounting for the similarity between molecules are usually calculated from fingerprints or descriptor vectors using either some standard functions (linear, polynomial, Gaussian) or other popular similarity measures, such as Euclidean distance or Tanimoto coefficient.[25]

CONCLUSION

As quantum hardware becomes more powerful, we expect quantum algorithms for chemistry and machine learning to be progressively integrated into CADD. While FTQC devices are not expected to be available within the next decade, NISQ devices would be more

likely commercialized in the next two to four years. The size of the problems that will be solved on these computers will be linked to their specifications (number of qubits and coherence time). For instance, quantum devices with qubit count N being a few hundreds of qubits and OðNÞ coherence time would be able to perform quantum simulation on molecular systems with the same number of spin orbitals using VQE. Using techniques such as active space approach, we could study molecules of the size of typical drug candidates. This type of calculation could be useful in the parameterization of force fields, in synthesizability and bio-catalysis studies, and in the generation of QSAR descriptors. Calculation of binding energies using QM/MM techniques will likely require in the order of a few thousand qubits and will take advantage of the integration with classical tools. We are entering a new era of quantum computing where quantum hardware currently available already allow for rapid prototyping of quantum algorithms. As a result, the field is open to early explorations of how quantum devices can be used for concrete application settings. Drug discovery is a unique area in the sense that it benefits from advances in both quantum chemistry and machine learning, making it one of the first areas that are likely to adopt quantum computing into its pipelines. This perspective is an invitation to both the quantum computing and the drug discovery communities to bridge the technical gap needed to fully materialize the potential of quantum computing for drug discovery.

REFERENCE

1. J. A. DiMasi, R. W. Hansen, and H. G. Grabowski, "The price of innovation: New estimates of drug development costs," J. Health Econ., 2003; 22(2): 151–185.
2. T. J. Dougherty and M. J. Pucci, Antibiotic Discovery and Development. Berlin, Germany: Springer, 2011.
3. Jain S, Purohit A, Nema P, Vishwakarma H, Jain PK. A brief review on nutraceuticals and its application. Asian Journal of Dental and Health Sciences, Mar 15, 2022; 2(1): 7-13.
4. G. Sliwoski, S. Kothiwale, J. Meiler, et al., "Computational methods in drug discovery," Pharmacological Rev., 2014; 66(1): 334–395.
5. J. H. Van Drie, "Computer-aided drug design: The next 20 years," J. Comput.-Aided Mol. Des., 2007; 21(10–11): 591–601.
6. T. Weymuth and M. Reiher, "Inverse quantum chemistry: Concepts and strategies for rational compound design," Int. J. Quantum Chem., 2014; 114(13): 823–837.
7. A. Aspuru-Guzik, R. Lindh, and M. Reiher, "The matter simulation (r)evolution," ACS Central Sci., 2018; 4(2): 144–152.
8. A. Aspuru-Guzik, A. Dutoi, P. Love, et al., "Simulated quantum computation of molecular energies," Science, 2005; 309(5741): 1704–1707.
9. E. Conover, "Google moves toward quantum supremacy with 72-qubit computer," Sci. News, 2018; 193(6): 13.
10. C. Rigetti, "The Rigetti 128-qubit chip and what it means for quantum," Medium, Aug. 2018. [Online]. Available: https:// medium.com/rigetti/the-rigetti-128-qubit-chip-and-what-itmeans-for-quantum-df757d1b71ea
11. J. Biamonte, P. Wittek, N. Pancotti, et al., "Quantum machine learning," Nature, 2016; 549(7671): 195–202.
12. Purohit A, Rangrej AK, Sahu A, Guru A, Singhai A, Shukla A, Jain S, Jain H. Impact of Covid-19 and Online Education on the Mental Health of Members of Educational Sphere-a Case Study. International Journal of Current Pharmaceutical Research, 2022; 69-76.
13. J. P. Hughes, S. S. Rees, S. B. Kalindjian et al., "Principles of early drug discovery," Brit. J. Pharmacology, 2011; 162(6): 1239–1249.
14. K. Stromgaard, P. Krogsgaard-Larsen, and U. Madsen, Eds. Introduction to Drug Design and Discovery, 5th ed. Boca Raton, FL, USA: Taylor & Francis, 2017.
15. M. Compiani and E. Capriotti, "Computational and theoretical methods for protein folding," Biochemistry (Mosc.), 2013; 52(48): 8601–8624.
16. Jain S, Kirar M, Bindeliya M, Sen L, Soni M, Shan M, Purohit A, Jain PK. Novel drug delivery systems: an overview. Asian Journal of Dental and Health Sciences, Mar 15, 2022; 2(1): 33-9.
17. S. Kalyaanamoorthy and Y.-P. Phoebe Chen, "Structure-based drug design to augment hit discovery," Drug Discovery Today, 2011; 16(17–18): 831–839.
18. A. C. Anderson, "The process of structure-based drug design," Chem. Biol., 2003; 10(9): 787–797.
19. A. Golbraikh, X. S. Wang, H. Zhu et al., "Predictive QSAR modeling: Methods and applications in drug discovery and chemical risk assessment," in Handbook of Computational Chemistry. Berlin, Germany: Springer, 2012; 1309–1342.
20. J. Bajorath, "Computer-aided drug discovery," F1000 Res., 2015; 4: 630.
21. J. Preskill, "Quantum computing in the NISQ era and beyond," Quantum, Aug. 2018; 2: 79.
22. B. Lomenick, R. W. Olsen, and J. Huang, "Identification of direct protein targets of small molecules," ACS Chem. Biol., 2010; 6(1): 34–46.
23. L. Burdine and T. Kodadek, "Target identification in chemical genetics: The (often) missing link," Chem. Biol., 2004; 11(5): 593–597.
24. J. N. Y. Chan, C. Nislow, and A. Emili, "Recent advances and method development for drug target identification," Trends Pharmacological Sci., 2010; 31(2): 82–88.
25. M. Schenone, V. Dan, B. K. Wagner, et al., "Target identification and mechanism of action in chemical biology and drug discovery," Nature Chem. Biol., 2013; 9(4): 232–240.