Personalizing pharmacological treatments for Chronic Myeloid Leukemia (CML) with a hybrid quantum-classical computing stack

Venkata Shashish Vasireddi[1], Collin Szczepanski[2]

[1] Heritage High School, Texas, United States of America
[2] Columbia University, New York, United States of America

* Corresponding author
E-mail: shashish.v.vasireddi@gmail.com (VV)

Word count: 3126

Funding: The author(s) received no specific funding for this work.

Conflicts of interest: All authors declare no conflict of interest.

Personalizing pharmacological treatments for Chronic Myeloid Leukemia (CML) with a hybrid quantum-classical computing stack

Abstract

Chronic myeloid leukemia (CML) is one of the first tumors whose molecular drivers were understood by researchers, revolutionizing treatment by providing one of the first targeted therapies with Tyrosine Kinase Inhibitors (TKIs). Despite CML being a low variability cancer, it contains different protein isomers, signaling the importance of selecting an inhibitor specific to the patient. For that, we propose a hybrid quantum-classical pipeline to efficiently promote a treatment with reduced complications. The pipeline takes the patient's sequenced tumor genome, identifies key mutations, and then retrieves its respective mutated protein from either the Protein Data Bank (PDB) or AlphaFold. Each of the potential candidate ligands are tested on the mutated protein by simulating docking via a Quantum Approximate Optimization Algorithm (QAOA). The pipeline is concluded with classical methods that take the Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) profile before retrieving the best match for the patient. We then tested this method on the kinase domain of the BCR::ABL1 protein, successfully testing and ranking eight CML approved treatments. Bafetinib was identified as the candidate with the highest binding score but also showed significant predicted risks. However, Nilotinib, which holds a binding score just under Bafetinib, is the best overall candidate for the patient. It showed strong potency while having a lower risk of side effects, suggesting that the drug be effective at low dosages. These results show the pipeline's ability to make suggestions by balancing a drug's strength with its safety profile to achieve the best treatment.

Keywords: Chronic myeloid leukemia, Quantum Approximate Optimization Algorithm, hybrid quantum-classical computing, molecular docking, BCR::ABL1

Introduction

Chronic Myeloid Leukemia (CML) is a myeloproliferative neoplasm that stems from the formation of the Philadelphia chromosome and its signature BCR::ABL1 oncogene [1]. This gene derives from the fusion of the Abelson murine leukemia (ABL1) gene on chromosome 9 and the breakpoint cluster region (BCR) gene on chromosome 22. It results in constitutively active tyrosine kinase that supports CML cell growth through the activation of downstream signaling pathways such as RAS, RAF, JUN kinase, MYC, and signal transducer and activator of transcription. This in turn affects leukemogenesis by allowing an uncontrolled cycle without cytokines, preventing the cells from going through apoptosis and sending abnormal signals to prevent them from dying [2].

Over the last 20 to 30 years, treatment of chronic myeloid leukemia has gone from destructive chemotherapy to less dangerous targeted therapy with Tyrosine Kinase Inhibitors (TKIs). Since the introduction of TKI treatment in 2000, they have brought the annual death rate of CML from between 10% to 20% in the first couple years to roughly 1% to 2% [1]. The inhibitors work by impeding the interaction between the adenosine triphosphate (ATP) with the BCR::ABL1 oncoprotein, blocking the rapid cellular division of the cancer [2]. There are three generations of TKI drugs, of which the first two are generally used. Imatinib is referred to as a first-generation TKI while dasatinib, bosutinib, and nilotinib are all a part of the second-generation. The second-generation contains stronger inhibitors of BCR::ABL1 and has been shown to work against certain variants that resist imatinib. Ponatinib, asciminib, and other unfinished therapies third-generation TKIs that maintain potency against the T3151 variant in which first and second generation TKIs are prevented from binding to the ATP binding site [2-5].

With a multitude of treatments available, it becomes increasingly important to take into consideration the adverse effects that the patient is at risk of facing. Some of these effects may be

specific to the TKI or be an issue for all. For instance, all TKIs hold the risk of causing hematologic toxicity, while specifically Imatinib has been seen to commonly cause rashes, edema, and myalgias [1, 4, 6]. These well documented side effects show that the ideal treatment would work with high efficacy even at low dosages to prevent complications.

This is the issue the proposed hybrid quantum-classical stack aims to mitigate. Quantum computing is a novel approach to computation that follows and is based on the principles of quantum mechanics. It sports algorithms and theories unique to quantum computers, which have their own respective traits such as running on quantum bits (qubits) instead of bits, quantum logic gates over logic gates, and more. Quantum computers especially excel when it comes to computations and simulations at small scales as the fundamental nature of qubits works well at representing and simulating quantum systems [7-8]. This makes it useful for small molecular interactions, something classical computers struggle with, and an advantage that the proposed pipeline takes advantage of. Despite these possibilities, quantum computing is currently in the noisy intermediate-scale quantum (NISQ), limiting the number of qubits available for use and preventing us from making complete use of quantum computers [9-10]. While the fundamentals of quantum computing are beyond the scope of this paper, we focus mainly on its application to molecular docking through the Quantum Approximate Optimization Algorithm (QAOA) [10]. Using which, we propose a quantum-classical pipeline that integrates quantum docking via the QAOA with classical ADMET filtering to identify an appropriate inhibitor for CML.

Methods

Protein Target Preparation and Ligand Selection.

Treatment with TKIs specifically target the dysregulated ABL1 found in CML as it is a crucial target which resists apoptosis and causes unregulated cellular proliferation. The kinase domain of the BCR-ABL1 fusion protein was selected as the target for this study, specifically the

ATP binding site within it which is an essential part of the process [11]. The specific amino acid sequence corresponding to the ABL1 kinase domain was obtained from the UniProt database (Accession No. UPI000013D6D4) [12]. Since a crystal structure of the mutated protein was unavailable, its three-dimensional structure was predicted using AlphaFold2. The eight candidate ligands, including Imatinib, Dasatinib, and Nilotinib, were selected as they are all well documented Tyrosine Kinase Inhibitors used in CML treatment [1-3, 5, 13, 14]. Their chemical structures were then retrieved in the SMILES (Simplified Molecular Input Line Entry System) format from the PubChem database [15-22].

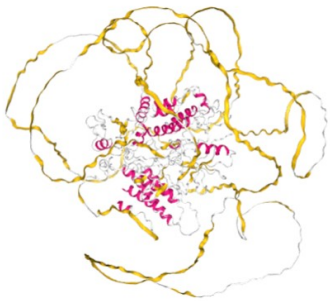


Figure 1. Predicted Structure of the ABL1 Kinase Domain. The three-dimensional model was generated by AlphaFold2 using the ABL1 kinase domain sequence from UniProt [12]. This figure is an original work created for this study.

Graph Representation of Molecular Structures.

To model the molecular interactions, both the protein target and the candidate ligands were converted into graph representations. Before this pharmacophores, structural features critical for a molecule's biological activity, are selected within the radius of the docking site of the protein and off the ligand. Then both the ligand and protein are turned into Labeled Distance

Graphs (LDGs), which work to map the 3D positions of their key chemical features. Within them pharmacophores are represented by vertices, each annotated according to their type (e.g., hydrogen bond acceptors, hydrophobic centers). The edges of the graph are the distances between the vertices, with the weights being the Euclidean distance between them. This allows for the understanding of the geometric relationships between pharmacophores, allowing us to focus on points of interest. Subsequently we create a Binding Interaction Graph (BIG) whose main purpose is to represent all potential intermolecular interactions. Unlike LDGs, in the BIG, vertices represent compatible pharmacophore pairs between the ligand and the protein. If the ligand has n pharmacophores and the protein has m, the BIG will have a total of n x m vertices. In the BIG, edges are only drawn in case of geometric compatibility which signals that the pairs can coexist physically. If a clique, a group of where every point is directly connected to every other point in the group, exists within the BIG then it is a feasible docking position. This is because for this to happen every vertex (ligand-protein pharmacophore pair) within the clique must coexist, suggesting a possible docking posture spatially. From there solving the Maximum Weighted Clique Problem reveals the most probable docking configuration based on the combined weight of each BIG configuration [10].

In the case of this study, compatibility between contact points is based on the Euclidian distance between pharmacophore points. In order to be considered compatible, a pair must at least meet the threshold of the sum of interaction distances $\epsilon 1$, $\epsilon 2$, and the flexibility constant τ [10]. The equation exists as below:

Quantum Approximate Optimization Algorithm for Docking Simulation.

With the creation of the BIG the molecular docking problem has changed into Maximum Weighted Clique (MWC) problem. This graph theory problem involves trying to find the densest clique where each vertex has a weight and the goal is to maximize the total weight. In the context of the initial problem, this retrieves the ideal docking position and strength of it, the latter of which required for the proposed pipeline. To achieve this the problem is taken from a Binary Quadratic Programming (BQP) classical optimization problem to a quantum Hamiltonian for the Quantum Approximate Optimization Algorithm (QAOA) algorithm [10]. The objective function to maximize is seen below:

$$\sum_{i=1}^{N} w_i x_i + \sum_{i=1, j \neq i}^{N} w_{ij} x_i x_j$$

The goal is to maximize ($\sum w_i x_i$) which works as the "reward" for each vertex that is chosen to be included in the graph. This reward is equal to the biochemical strength of the vertex, working as its weight (w_i) mathematically. ($\sum w_{ij} x_i x_j$) is the penalty term that is applied should invalid interactions be included. The penalty is so severe that any docking pose including an incompatible interaction will have a low score, forcing the system to try again.

To solve this with QAOA, the problem must be translated into something in accordance with quantum computing. This is done by converting BQP objective function into a cost Hamiltonian (H_C). Each classical binary variable x_i (0 or 1) is mapped to a quantum Pauli-Z spin operator σ_i^z which acts on one qubit and has possible measurement outcomes of +1 or -1. The relation can be seen below:

$$x_i = (\sigma_i^z - 1)/2.$$

This relationship is then substituted into the classical BQP formula. Giving the cost Hamiltonian which describes the total energy of the system, as seen below:

$$H_C = \sum_{i \in V} w_i \sigma_i^z + \frac{P}{2} \sum_{(i,j) \notin E, i \neq j} \left(\sigma_i^z \sigma_j^z - \sigma_i^z - \sigma_j^z \right) + const.$$

In this equation the constant does not affect the optimization and can be left out during implementation. The Hamiltonian is constructed so that its lowest possible energy (ground state) correlated exactly to the solution of the MWC problem solution with the highest score [10].

QAOA is an iterative algorithm that finds this ground state by continuously applying two different operators onto an initial quantum state, the cost operator and the mixer operator.

The cost operator ($U(H_C, \gamma)$) encodes the problem we are attempting to solve. Every time it is applied it adds a phase, a quantum property that doesn't change the result itself but rather influences how it will superpose and influence observable quantities. It is most often understood as the angle of the complex number within the wave function that quantum states are defined in. This phase is added to every single docking pose at the same time, for each pose it calculates its cost, reflecting how good or bad the pose is. Following this it rotates the phase of that specific state proportional to its cost, a good docking pose with a low cost will have its phase rotated by a small amount, the opposite is also true. Following this step, all the good solutions have a similar phase, with their vector pointing in roughly the same direction, while all the bad solutions have been spun around into random directions.

This sets the states up for the mixer operator ($U(H_M, \beta)$), which makes all the states interfere with each other. When states with similar phases interfere, they add up and reinforce each other, growing their amplitude, increasing their probability of being the final observed answer. On the other hand, states with different random phases will cancel each other out, and

with all the bad solutions being given a variety of phases, their amplitudes will shrink, decreasing their probability of being the final answer.

The QAOA algorithm itself starts with the initial state that is an equal superposition of all possible solutions. For *p* steps, the cost and mixer operators are consecutively applied to the quantum state. This process is controlled by the classical parameters γ and β, which are tuned by a classical optimizer. The goal of the optimization is to get as close as possible to the ground state of the cost Hamiltonian. After the last step, the quantum state is measured, and the resulting bitstring left is the optimal docking configuration found by the algorithm.

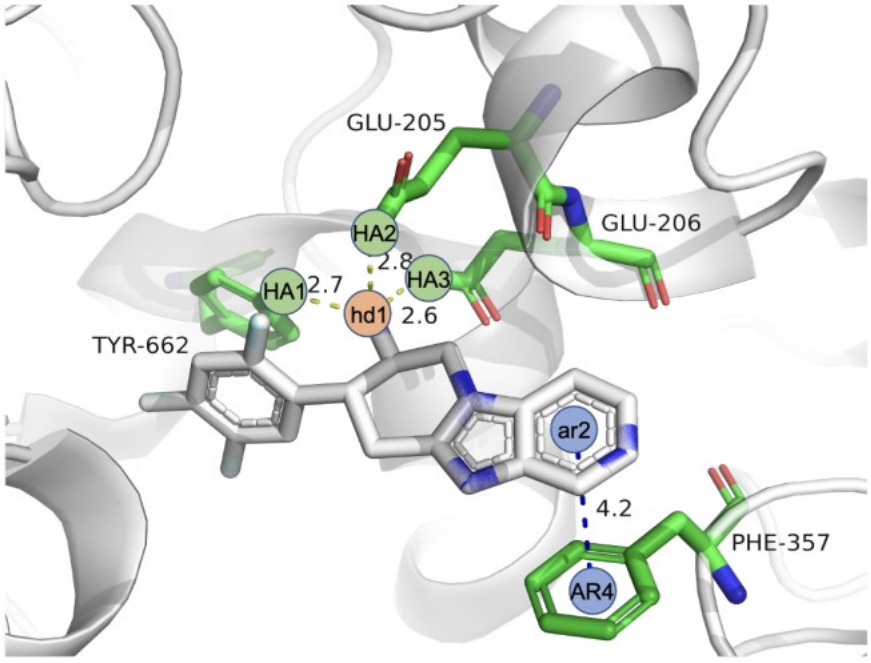


Figure 2. Visualization of a Predicted Protein-Ligand Docking Pose. The figure illustrates an example binding configuration derived from QAOA in the same method of the proposed pipeline. The ligand's carbon atoms shown in white and the interacting protein residues' carbons are shown in green. Key predicted interactions, such as hydrogen bonds (hd1-HA1) and π-π stacking (ar2-AR4), are highlighted with yellow dashed lines to show the primary forces stabilizing the complex [10].

In Silico ADMET Profiling and Candidate Ranking.

Following all docking simulations, all candidate ligands were evaluated for their Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) as well as their results for the Lipinski Rule of 5. We specifically checked human intestinal absorption (HIA), blood-brain barrier (BBB) penetration, inhibition of the cytochrome P450 enzyme CYP3A4, and potential cardiotoxicity via hERG channel inhibition. Finally, the candidate drugs were ranked based off their QAOA-derived binding score, this score was not according to any standardized metric but was rather the negative of the ground state that was calculated. This way the higher the number, the greater the binding strength, giving us a relative method to rank the ligands. We utilize that relative score and the predicted ADMET safety profiles to identify the optimal treatment specific to the patient.

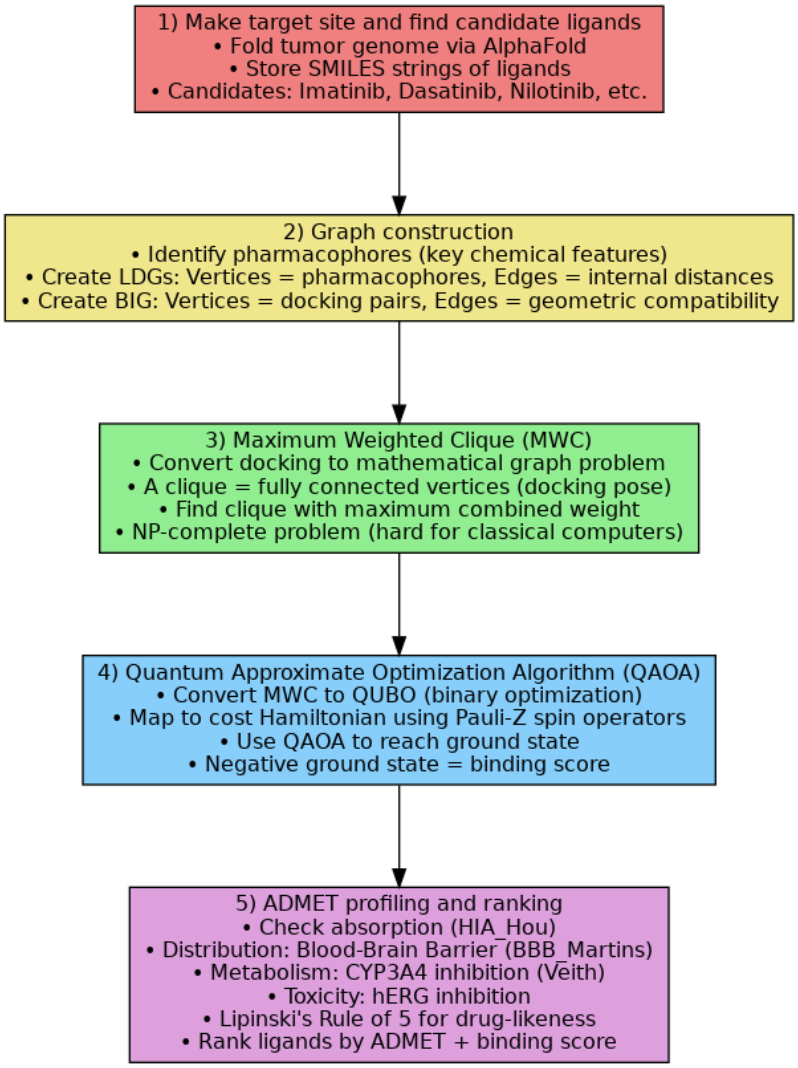


Figure 3. A flowchart detailing the full pipeline from to start to finish, beginning with Protein Target Preparation and Ligand Selection and concluding with In Silico ADMET Profiling and Candidate Ranking.

Results

We ran the outlined pipeline on a folded sequence of ABL1 on the eight candidate drugs we collected based on well-documented reports, Imatinib, Dasatinib, Nilotinib, Bosutinib, Ponatinib, Asciminib, Bafetinib, Axitinib [1-3, 5, 13, 14]. From this process we got the results seen on the below table which pointed to Nilotinib as the ideal treatment.

Table 1. This table summarizes the derived binding scores and key ADMET properties for eight candidate Tyrosine Kinase Inhibitors targeting the ABL1 kinase domain. Binding scores correlate to the predicted potency, while the ADMET profile looks at risks like Blood-Brain Barrier (BBB) crossing, hERG-mediated cardiotoxicity, and potential for drug-drug interactions via CYP3A4 inhibition.

Ligand	Binding Score	HIA_Hou[a]	BBB_Martins[b]	hERG[c]	CYP3A4_Veith[d]	Lipinski pass[e]
Bafetinib	1.103034	0.998798	0.786797	0.980993	0.978987	False
Nilotinib	1.058147	0.998859	0.285234	0.888585	0.780947	True
Asciminib	0.951903	0.992016	0.597829	0.965251	0.857186	True
Dasatinib	0.830372	0.998198	0.892039	0.988828	0.881840	False
Bosutinib	0.779530	0.999184	0.695049	0.801234	0.889655	True
Axitinib	0.767429	0.985370	0.778019	0.984515	0.676099	False
Imatinib	0.491420	0.995878	0.492672	0.933132	0.511831	True
Ponatinib	0.330828	0.998695	0.478613	0.964295	0.958168	False

[a] HIA stands for Human Intestinal Absorption, this test predicts whether a drug, if taken as a pill, can pass through the walls of the gut and enter the bloodstream. A drug passes the test if it scores **> 0.8**

[b] BBB stands for Blood-Brain Barrier, this test predicts whether a drug can bypass that protective barrier around the brain that stops most substances. A drug passes the test if it scores **< 0.5.**

[c] The hERG gene is responsible for a critical potassium channel in the heart that helps regulate the heartbeat. This test predicts the Cardiotoxicity risk by seeing if the drug inhibits that channel. A drug passes the test if it scores **< 0.3.**

[d] CYP3A4 is responsible for breaking down over half of all common medications (like painkillers, statins, etc.). This test predicts the drug interaction risk by seeing whether the drug will inhibit this enzyme. A drug passes this test if it scores **< 0.5.**

[e] The Lipinski Rule of 5 isn't a strict test but a set of guidelines based on the physical properties of a molecule. Molecules that pass this test have a better chance of having good absorption and distribution properties.

In the compiled table we see that all ligands failed the Cardiotoxicity Risk and Drug Interaction tests, this is to be expected as it is a core risk with TKI treatment [23-24]. However, the other columns and tests showed great variety with their values, something visible with the binding scores, the brain permeability test, and the Lipinski test. The range of values, especially the former two, shows that despite doing the same function in principle, each ligand has different issues and benefits depending on the individual. This is seen with Bafetinib which had the highest value on the binding scores, 1.103034, reflecting its high potency on this specific target. However, the drug is also predicted to penetrate the Blood-Brain Barrier, bringing upon harmful effects alongside treatment. Nilotinib on the other hand displays a binding score just short of

Bafetinib at 1.058147, but unlike Bafetinib, it has reduced predicted side effects. The ligand passes the Blood-Brain Barrier test, passes the Lipinski test, and even has reduced risk compared to other ligands regarding Cardiotoxicity Risk and Drug Interaction. These results predict Nilotinib will work at low dosages with high potency while also mitigating harmful side effects. Making it the optimal treatment for this specific genome and cancer.

Discussion

Limitations.

A main limitation of this study was the lack of access to a quantum computer. Without it, the Quantum Approximate Optimization Algorithm (QAOA) was forced to be run on a classical simulator. The resources and time needed for classical simulation restricted the scale of the tests conducted, preventing us from utilizing a larger library of ligands and testing of multiple protein isomers of the target protein. Another consideration that should be made is that the ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) analysis performed was rudimentary compared to what is generally used. While this provided a general overview of each drug's potential safety profile, it does not have the capabilities and insight like a true screening does. The predictions should be considered a preliminary screening at best instead of drug safety evaluation. Lastly, the parameters used to define the pharmacophores were based on default values from other implementations [10] instead of being optimized for the ABL1 kinase. This generalization may affect the precision of the calculated binding scores as the ideal interactions can vary for different protein targets.

Future Work.

The computing stack suggested within this paper works as a proof of concept for the future of cancer treatment, in which there is large room for improvement. And while the current pipeline may be rudimentary it still was able suggest treatment candidates, future iterations would benefit from more sophisticated quantum algorithms, such as the Variational Quantum Eigensolver (VQE), and larger amounts of data to work with. With access to greater computational resources, including quantum hardware, this hybrid framework could be scaled up to screen chemical libraries against a large number of protein variations, creating a more robust and safe drug discovery process. The successful application of this quantum-classical approach to CML shows its potential as a novel approach for personalizing medicine, capable of speeding up the development of targeted therapies for diseases outside of CML.

Conclusion

The pipeline proposed in this paper is a successful, first of its kind, proof of concept for a hybrid quantum-classical approach to personalized drug selection for Chronic Myeloid Leukemia. By integrating quantum docking algorithm (QAOA) with classical ADMET profiling, the computational stack allowed us to optimize between a drug's potency and its potential for adverse effects. The analysis of eight approved Tyrosine Kinase Inhibitors against the target mutated protein correctly identified Nilotinib as the best candidate, not because it had the absolute highest binding score, but because it had both a high predicted efficacy alongside a superior safety profile.

These results prove that a hybrid model like the one seen within this study can provide personalized treatment recommendations that go beyond what is possible with classical computers. Although the scope and results of the paper were constrained by classical simulation of quantum algorithms, it still proved there is a viable path toward the utilization of quantum

computing for real-world clinical and pharmaceutical reasons. As quantum hardware continues to improve, the structure here holds promise in its ability to be scaled into a usable system to accelerate drug development and personalize targeted therapies for a large range of diseases.

Acknowledgements

Special thanks to my mentor Collin Szczepanski of Columbia University for bringing his expertise to my project and helping me out, without him many tools like AlphaFold would be out of my scope and my project would not be as complete as it is now. I also have the Harvard OpenBIO SRI program to thank for this opportunity and the incredible experience I have gained.

References

1. Jabbour E, Kantarjian H. Chronic myeloid leukemia: a review. JAMA. 2025 May 13.
2. Jabbour E, Kantarjian H. Chronic myeloid leukemia: 2025 update on diagnosis, therapy, and monitoring. American journal of hematology. 2024 Nov;99(11):2191-212.
3. Pottier C, Fresnais M, Gilon M, Jérusalem G, Longuespée R, Sounni NE. Tyrosine kinase inhibitors in cancer: breakthrough and challenges of targeted therapy. Cancers. 2020 Mar 20;12(3):731.
4. Fachi MM, Tonin FS, Leonart LP, Rotta I, Fernandez-Llimos F, Pontarolo R. Haematological adverse events associated with tyrosine kinase inhibitors in chronic myeloid leukaemia: A network meta-analysis. British Journal of Clinical Pharmacology. 2019 Oct;85(10):2280-91.
5. Rossari F, Minutolo F, Orciuolo E. Past, present, and future of Bcr-Abl inhibitors: from chemical development to clinical efficacy. Journal of hematology & oncology. 2018 Jun 20;11(1):84.
6. Pretel-Irazabal M, Tuneu-Valls A, Ormaechea-Pérez N. Adverse skin effects of imatinib, a tyrosine kinase inhibitor. Actas Dermo-Sifiliográficas (English Edition). 2014 Sep 1;105(7):655-62.
7. Balamurugan KS, Sivakami A, Mathankumar M. Quantum computing basics, applications and future perspectives. Journal of Molecular Structure. 2024 Jul 15;1308:137917.
8. Cicero A, Maleki MA, Azhar MW, Kockum AF, Trancoso P. Simulation of quantum computers: Review and acceleration opportunities. arXiv preprint arXiv:2410.12660. 2024 Oct 16.
9. Preskill J. Quantum computing in the NISQ era and beyond. Quantum. 2018 Aug 6;2:79.
10. Ding QM, Huang YM, Yuan X. Molecular docking via quantum approximate optimization algorithm. Physical Review Applied. 2024 Mar 1;21(3):034036.\
11. Irgit A, Kamıs R, Sever B, Tuyun AF, Otsuka M, Fujita M, Demirci H, Ciftci H. Structure and dynamics of the ABL1 tyrosine kinase and its important role in chronic myeloid leukemia. Archiv der Pharmazie. 2025 May;358(5):e70005.
12. UniProt Consortium. UniParc entry UPI000013D6D4 [Internet]. The UniProt Consortium; 2025 [cited 2025 Aug 8]. Available from: https://www.uniprot.org/uniparc/UPI000013D6D4/entry
13. Pemovska T, Johnson E, Kontro M, Repasky GA, Chen J, Wells P, Cronin CN, McTigue M, Kallioniemi O, Porkka K, Murray BW. Axitinib effectively inhibits BCR-ABL1 (T315I) with a distinct binding conformation. Nature. 2015 Mar 5;519(7541):102-5.
14. Santos FP, Kantarjian H, Cortes J, Quintas-Cardama A. Bafetinib, a dual Bcr-Abl/Lyn tyrosine kinase inhibitor for the potential treatment of leukemia. Current Opinion in Investigational Drugs (London, England: 2000). 2010 Dec 1;11(12):1450-65.
15. PubChem. Imatinib [Internet]. pubchem.ncbi.nlm.nih.gov. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/5291
16. PubChem. Dasatinib [Internet]. pubchem.ncbi.nlm.nih.gov. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/3062316
17. PubChem. Nilotinib [Internet]. Nih.gov. PubChem; 2024. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/644241
18. PubChem. Bosutinib [Internet]. pubchem.ncbi.nlm.nih.gov. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/5328940

19. PubChem. Ponatinib [Internet]. Nih.gov. PubChem; 2025 [cited 2025 Aug 9]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/24826799
20. PubChem. Asciminib [Internet]. pubchem.ncbi.nlm.nih.gov. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/72165228
21. PubChem. Bafetinib [Internet]. Nih.gov. PubChem; 2025 [cited 2025 Aug 9]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/11387605
22. PubChem. Axitinib [Internet]. pubchem.ncbi.nlm.nih.gov. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/6450551
23. Sayegh N, Yirerong J, Agarwal N, Addison D, Fradley M, Cortes J, Weintraub NL, Sayed N, Raval G, Guha A. Cardiovascular toxicities associated with tyrosine kinase inhibitors. Current cardiology reports. 2023 Apr;25(4):269-80.
24. van Leeuwen RW, van Gelder T, Mathijssen RH, Jansman FG. Drug–drug interactions with tyrosine-kinase inhibitors: a clinical perspective. The Lancet Oncology. 2014 Jul 1;15(8):e315-26.