



OPEN ACCESS

EDITED BY
Chandrabose Selvaraj,
Alagappa University, India

REVIEWED BY
Young Ho Seo,
Keimyung College, South Korea
Arul Murugan Natarajan,
Indraprastha Institute of Information
Technology Delhi, India

*CORRESPONDENCE
Kwan-Young Jung,
krjeong@kriect.re.kr
Jae-Sung Kim,
jaesung@kirams.re.kr
Sangwook Wu,
s.wu@pharmacadd.com

SPECIALTY SECTION
This article was submitted to Biophysics,
a section of the journal
Frontiers in Molecular Biosciences

RECEIVED 17 October 2022
ACCEPTED 16 November 2022
PUBLISHED 25 November 2022

CITATION
Jang SH, Sivakumar D, Mudedla SK,
Choi J, Lee S, Jeon M, Bvs SK, Hwang J,
Kang M, Shin EG, Lee KM, Jung K-Y,
Kim J-S and Wu S (2022), PCW-A1001,
AI-assisted *de novo* design approach to
design a selective inhibitor for FLT-
3(D835Y) in acute myeloid leukemia.
Front. Mol. Biosci. 9:1072028.
doi: 10.3389/fmolb.2022.1072028

COPYRIGHT
© 2022 Jang, Sivakumar, Mudedla, Choi,
Lee, Jeon, Bvs, Hwang, Kang, Shin, Lee,
Jung, Kim and Wu. This is an open-
access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

PCW-A1001, AI-assisted *de novo* design approach to design a selective inhibitor for FLT-3(D835Y) in acute myeloid leukemia

Seong Hun Jang¹, Dakshinamurthy Sivakumar¹,
Sathish Kumar Mudedla¹, Jaehan Choi¹, Sungmin Lee¹,
Minjun Jeon¹, Suneel Kumar Bvs¹, Jinha Hwang¹,
Minsung Kang², Eun Gyeong Shin^{3,4}, Kyu Myung Lee³,
Kwan-Young Jung^{3,4*}, Jae-Sung Kim^{2*} and Sangwook Wu^{1,5*}

¹R&D Center, PharmCADD, Busan, South Korea, ²Division of Radiation Biomedical Research, Korea Institute of Radiological and Medical Sciences, Seoul, South Korea, ³Therapeutics & Biotechnology Division, Korea Research Institute of Chemical Technology, Daejeon, South Korea, ⁴Department of Medicinal Chemistry and Pharmacology, University of Science & Technology, Daejeon, South Korea, ⁵Department of Physics, Pukyong National University, Busan, South Korea

Treating acute myeloid leukemia (AML) by targeting FMS-like tyrosine kinase 3 (FLT-3) is considered an effective treatment strategy. By using AI-assisted hit optimization, we discovered a novel and highly selective compound with desired drug-like properties with which to target the FLT-3 (D835Y) mutant. In the current study, we applied an AI-assisted *de novo* design approach to identify a novel inhibitor of FLT-3 (D835Y). A recurrent neural network containing long short-term memory cells (LSTM) was implemented to generate potential candidates related to our in-house hit compound (PCW-1001). Approximately 10,416 hits were generated from 20 epochs, and the generated hits were further filtered using various toxicity and synthetic feasibility filters. Based on the docking and free energy ranking, the top compound was selected for synthesis and screening. Of these three compounds, PCW-A1001 proved to be highly selective for the FLT-3 (D835Y) mutant, with an IC₅₀ of 764 nM, whereas the IC₅₀ of FLT-3 WT was 2.54 μM.

KEYWORDS

FLT-3, AML, kinase inhibitors, *de novo* drug design, LSTM, AI

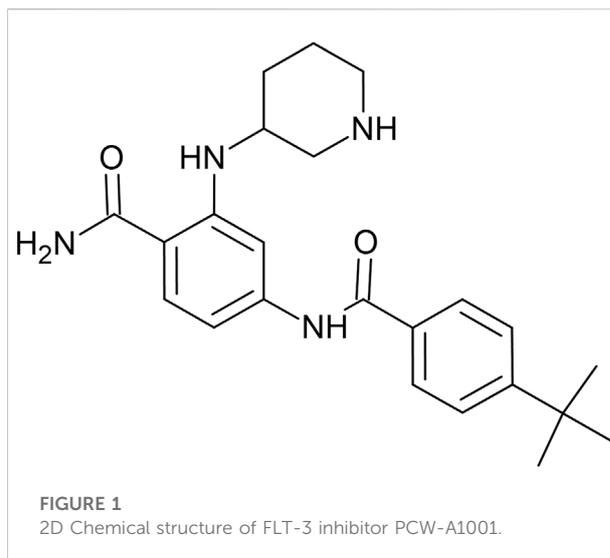
Introduction

Overexpression or mutation of some signaling proteins leads to cancer development (Kazi and Rönnstrand, 2019). Among the most mutated extracellular signaling mediators in cancer are the receptor tyrosine kinases (RTKs) (McDonnell et al., 2015). Among the five known types of RTKs, FMS-like tyrosine kinase (Kindler et al., 2010) (FLT-3) belongs to type III. It plays an essential role in regulating early hematopoiesis because it is selectively expressed on CD34+ hematopoietic stem cells and immature hematopoietic progenitors (Rosnet et al., 1996; Kindler et al., 2010). It is also expressed in the liver, spleen, lymph nodes, thymus, placenta, gonads, and brain (Del Zotto et al., 2001; Stirewalt and Radich, 2003; Brown and Small, 2004). Our work is focused on FLT-3, a gene that is highly mutated in acute myeloid leukemia (AML) (Grafone et al., 2012).

Interestingly, thousands of mutations (mostly insertions) have been reported for FLT-3. Many FLT-3 point mutations are commonly found in AML, and the activation loop residue D835, which stabilizes the inactive conformation is the predominant site of mutations (Yamamoto et al., 2001; Liang et al., 2003; Smith et al., 2012). Overexpression and frequent FLT-3 mutations are associated with poor prognoses and AML pathogenicity and activate downstream signaling molecules, which leads to stimulation and survival of cancerous cells (Zhang and Broxmeyer, 1999; Hayakawa et al., 2000; Lin et al., 2012). Treating AML patients by targeting FLT-3 and its mutants with small molecules is considered a promising strategy (Assouline et al., 2012; Leung et al., 2013; Gebru and Wang, 2020; Ambinder and Levis, 2021).

Since approval by the Food and Drug Administration (FDA) of the first tyrosine kinase inhibitor imatinib (Savage and Antman, 2002) two decades ago, several drugs targeting FLT-3 have entered clinical trials. Nevertheless, only midostaurin and gilteritinib have been approved by the FDA (Scholl et al., 2020). FLT-3 inhibitors are classified as type I or type II based on their binding with the protein. Type I inhibitors such as sunitinib (Schittenhelm et al., 2006), midostaurin (Stone et al., 2004), lestaurtinib (Smith et al., 2004), crenolanib (Heinrich et al., 2012), and gilteritinib (Grunwald and Levis, 2013) bind with the active state (DFG-in) of FLT-3, whereas type II inhibitors such as sorafenib (Auclair et al., 2007), ponatinib (O'Hare et al., 2009), and quizartinib (Zarrinkar et al., 2009) bind only with the inactive (DFG-out) FLT-3 conformation (Scholl et al., 2020). Studies have shown that type I inhibitors are more promising for use in AML treatment, as they target the predominant mutated kinase (Wodicka et al., 2010; Smith et al., 2012). There has been tremendous interest in developing FLT-3 inhibitors using classic computer-aided drug design approaches (Chang Hsu et al., 2014; Ke et al., 2015). In this study, we focus on developing a more rational approach for preparation of FLT-3 type-I inhibitors.

Recent breakthroughs show the significance of artificial intelligence (AI) in drug discovery, and AI reduces costs and



increases the speed of the drug discovery pipeline (Mak and Pichika, 2019). One of the main bottlenecks of traditional *de novo* drug design methods is the complicated synthetic routes; reported AI methods suggest synthetically feasible molecules or synthetic pathways that can help chemists (Corey and Wipke, 1969; Hessler and Baringhaus, 2018). Using AI, identification of a DDR1 kinase inhibitor was completed in just 60 days, including synthesis and experimental validation (Zavoronkov et al., 2019). Excientia prepared the first AI-designed drug (DSP-1181) to treat obsessive-compulsive disorder (OCD), which subsequently entered clinical trials (Luo et al., 2022). They also discovered the AI-designed molecule EXS-21546 for immuno-oncology, which entered clinical trials in 8 months. Insilico medicine (www.insilico.com) used its AI program to develop a novel inhibitor (ISM001-055) for antifibrotic targets, and it reached clinical trials in 9 months. Recently, they have announced a preclinical candidate for the main protease of SARS-CoV-2, which was discovered with their novel AI platform, Chemistry42.

Network-based approaches are widely used to infer relationships between diseases and drugs (Guney et al., 2016) and are more focused on predicting novel protein targets and new uses of known drugs (Berger and Iyengar, 2009; Wu et al., 2013). In the current study, we used our reverse network theory approach developed in-house to identify a potential therapeutic target for PCW-1001. Based on the network theory and docking results, FLT3 was considered a potential target. Further biological screening studies showed that PCW-1001 exhibited an inhibitory IC_{50} of 13.6 μ M against FLT-3 WT and 1.83 μ M against the FLT-3 (D835Y) mutant (Kang et al., 2022). An AI-assisted *de novo* design approach was applied to identify a potent and selective inhibitor for the FLT3/FLT-3 (D835Y) mutant. This parent compound (PCW-1001) was considered for further optimization, and more than 10,416 analogs were generated using the LSTM approach. These hits were further evaluated for synthetic

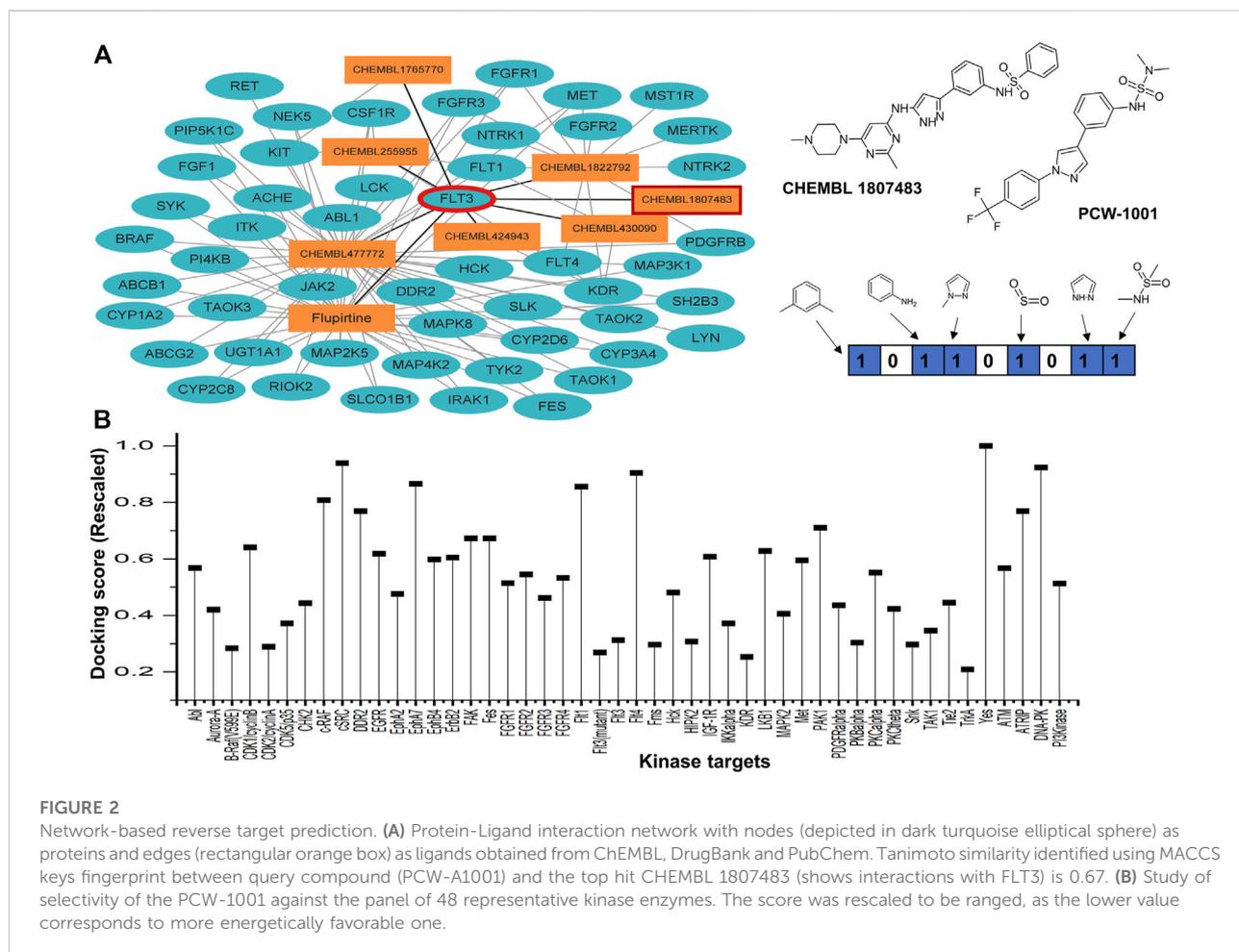


FIGURE 2

Network-based reverse target prediction. (A) Protein-Ligand interaction network with nodes (depicted in dark turquoise elliptical sphere) as proteins and edges (rectangular orange box) as ligands obtained from ChEMBL, DrugBank and PubChem. Tanimoto similarity identified using MACCS keys fingerprint between query compound (PCW-A1001) and the top hit CHEMBL 1807483 (shows interactions with FLT3) is 0.67. (B) Study of selectivity of the PCW-1001 against the panel of 48 representative kinase enzymes. The score was rescaled to be ranged, as the lower value corresponds to more energetically favorable one.

feasibility by in-house machine learning models and assessed for potential structural alerts. The resulting hits were subjected to docking studies, binding mode reviews, and free energy calculations for prioritization. Based on the binding mode review and free energy calculation data, the top compound was prioritized for synthesis and further screening. Screening data showed that PCW-A1001 (Figure 1) proved to be a potential and selective inhibitor against the FLT3 (D835Y) mutant.

Results and discussion

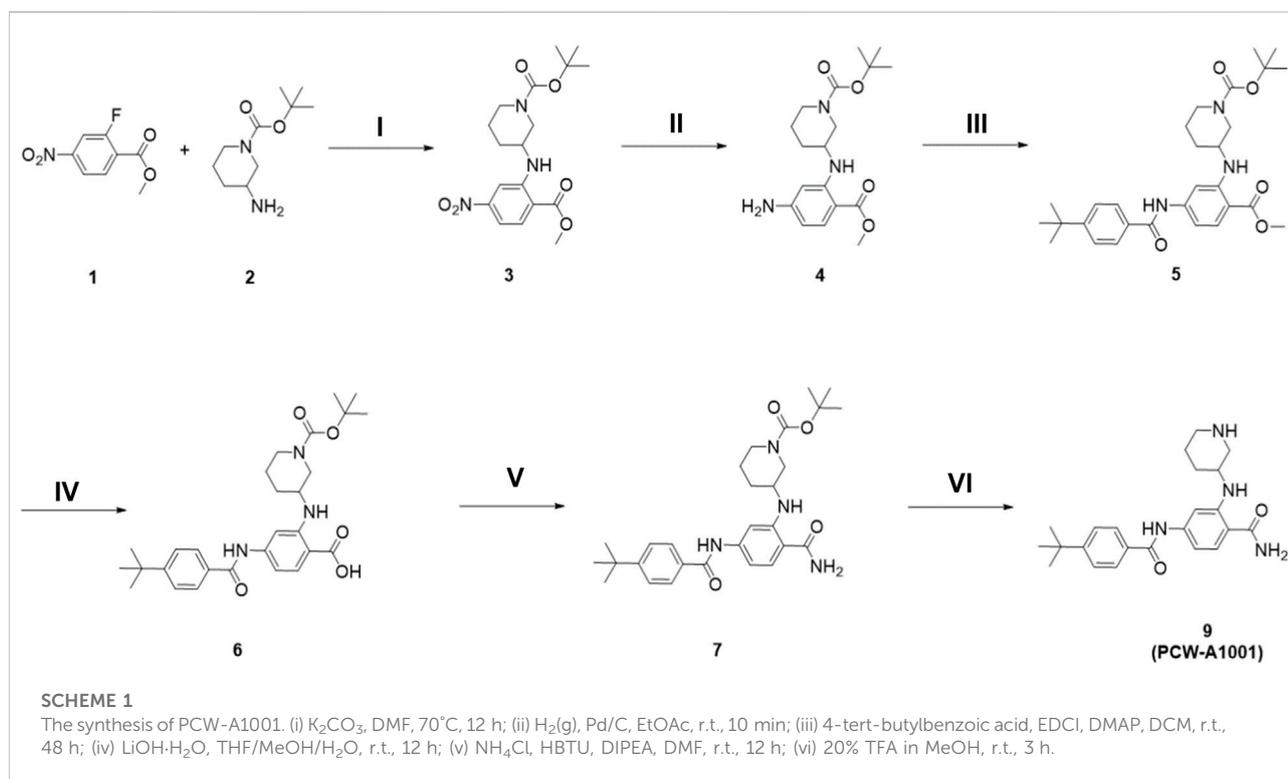
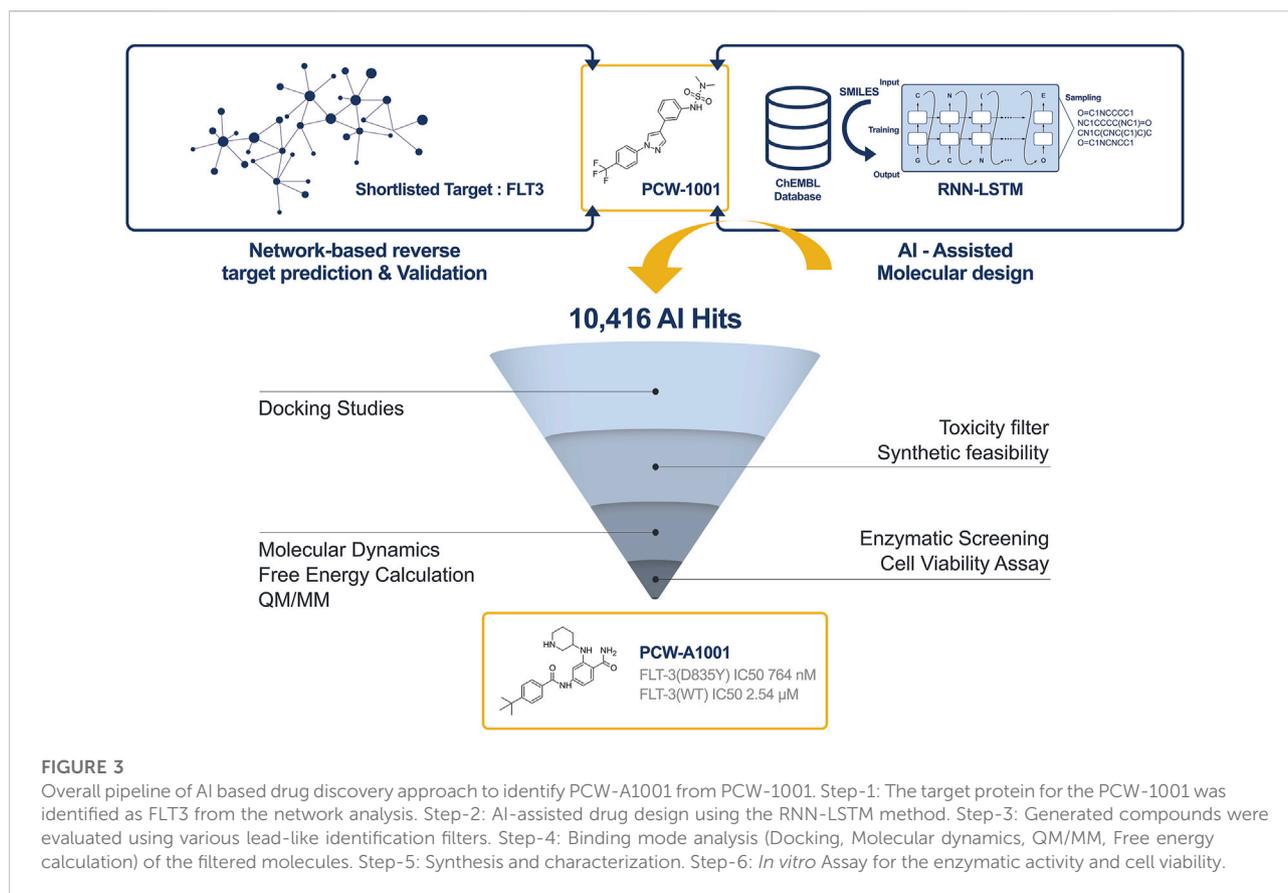
Network-based reverse target prediction

We implemented an in-house network-based reverse target prediction module to identify a protein target for PCW-1001 (Figure 2). Our parent compound, PCW-1001, exhibited a significant inhibitory profile against various breast cancer cell lines, but a substantial protein target was unknown (Kang et al., 2022). PCW-1001 compound structural similarity (atom pair descriptors) was computed against the ChEMBL chemical database. The

generated similarity matrix of the ChEMBL database with PCW-1001 and its corresponding protein target information was considered for further analysis. Ensemble docking studies were carried out for PCW-1001 against all 2,000 unique targets [with a known crystal structure database (www.rcsb.org)]. Of the top 10 scored (docking score) targets, five kinases (FLT3, JAK2, NTRK, MKNK2, and TGFBR1) were observed to be potential targets for PCW-1001. All five kinase targets are known to play a critical role in treating various cancers; among the five targets, we selected FLT-3 based on the score. Furthermore, FLT-3 point mutations are frequently found in AML, where the mutations occur in the activation loop residue D835 and stabilize the active conformation.

Artificial intelligence-assisted *de novo* design of a novel FMS-like tyrosine kinase 3 inhibitor using the long short-term memory approach

In the current study, we applied a deep recurrent neural network (RNN) with long short-term memory (LSTM) cells



for *de novo* drug design (Merk et al., 2018) to generate potential hit candidates around PCW-1001 (Figure 3). We fine-tuned the model by using the transfer learning approach to optimize *de novo* generation of FLT-3 active compounds (Kang et al., 2022). We sampled 10,416 SMILES (Simplified Molecular Input Line Entry System) strings from 20 epochs from the resulting fine-tuned model. AI-generated hits were further evaluated using MOSES (Polykovskiy et al., 2020) for novelty, validity, diversity, scaffold similarity, and uniqueness. Benchmarking analyses indicated that 98.8% of the hits were valid, 85% of the hits were unique, and 90% of the hits were novel. The chemical space of AI-generated hits falls within the range of FLT3 known actives and PCW-1001 (Supplementary Figure S1). Furthermore, violin plot analysis also suggested that the distribution of molecular weights and LogP of AI-generated hits were within the range of known FLT3 actives (Supplementary Figure S2). Overall, AI-generated hits fell within the chemical space of known actives, and MOSES analysis suggested that AI-generated hits were diverse and novel compared with known FLT3 actives and PCW-1001.

Based on the binding analysis and interaction pattern studies, we identified 1750 compounds out of 10,416 hits as suitable for further studies. Following this preliminary evaluation, we assessed the resulting structures for toxicity endpoints *via* our Pharmulator™. We generated 9 toxicity models (different endpoints) using available literature data, and validated models were deployed in Pharmulator™ to assess the hit moieties quickly. Of 1750 hits, only 190 compounds passed synthetic feasibility, novelty, drug-like, all toxicity, and PAINS filters and were further subjected to a binding pose analysis and free energy calculations. We selected the top compound for synthesis and *in vitro* screening.

Synthesis and structural characterization of the *de novo* compound PCW-A1001

The synthetic route to PCW-A1001 is summarized in Scheme 1. Methyl 2-fluoro-4-nitrobenzoate **1** and tert-butyl 3-aminopiperidine-1-carboxylate **2** were reacted in the presence of K₂CO₃ to obtain **3** through nucleophilic aromatic substitution. The nitro group of resulting compound **3** was converted to an amino functional group *via* hydrogenation. Amine Compound **4** was coupled with 4-(tert-butyl)benzoic acid in the presence of EDCI and a catalytic amount of DMAP and then hydrolyzed using LiOH·H₂O to produce intermediate **6**. The acid functional group was efficiently converted to an amide with the HBTU coupling reagent. The Boc protecting group of the secondary amine in the piperidine ring was removed to obtain the desired compound PCW-A1001. The step-by-step synthesis and structural characterizations are shown in the Supplementary Material.

Structural interaction and stability analysis of PCW-A1001 with wild-type FLT-3 and the FLT-3 (D835Y) mutant

The top predicted binding mode of PCW-A1001 with FLT-3 WT showed two key hydrogen bonding interactions with Cys694 and Cys695 and a π - π interaction with Phe830. The docking complex of PCW-A1001 with FLT-3 WT was considered for molecular dynamics (MD) simulation for 100 ns. The MD simulation results also showed that the compound binding interactions observed in the initial docked complex were retained in PCW-A1001. The compound bound perfectly in the ATP binding site by forming hydrogen bonds with the two cysteine residues (Cys694 and Cys695) located in the hinge region. As observed in several inhibitor-kinase complexes, hydrogen bonding interactions with the inhibitor are essential for kinase inhibitory activity (Banks et al., 1979; Ke et al., 2015). The carbonyl moiety of the benzamide group formed a hydrogen bond with the NH group of the Cys694 residue (Figure 4).

The NH moiety of the next benzamide group in the compound formed a hydrogen bond with the backbone carbonyl group of the Cys695 residue. Phe830 in the DFG loop in wild-type (WT) FLT-3 formed π - π interactions with compound PCW-A1001. The binding free energy of PCW-A1001 was -13.4 kcal/mol with FLT-3 WT but -14.8 kcal/mol with the FLT-3 (D835Y) mutant, whereas those of the precursor compound PCW-1001 were -7.2 kcal/mol and -8.07 kcal/mol for FLT-3(WT) and FLT-3 (D835Y), respectively (Table 1).

In the FLT-3 (D835Y) mutant, the NH group of Cys694 formed a hydrogen bond with the CO moiety of the benzamide group in PCW-A1001, as seen with WT FLT-3. Cys695 also maintained its hydrogen bonding interactions, as in the WT; additionally, Lys614 and Tyr693 interacted with the protein. The C α -RMSD of the WT FLT-3 and the mutant complex showed that the complex was stable throughout the simulation (Figure 4).

Based on the binding study of PCW-A1001 and its precursor compound PCW-1001 against the panel of kinase enzymes, selectivity was achieved by PCW-A1001 for FLT-3 (D835Y). The selectivity scores of PCW-A1001 and PCW-1001 were calculated from the dock score (rescaled) of the selected kinase panel of enzymes as 0.33 and 0.46, respectively, for the Flt-3 (D835Y) mutant (Supplementary Figure S3).

QM/MM analysis of PCW-A1001

QM/MM optimization was used to validate the interactions between PCW-A1001 and the FLT-3 (D835Y) mutant in the MD-determined complex to study the electronic and structural properties of the ligand and selected atoms of the protein (Figure 5). The electrostatic and van der Waals interactions

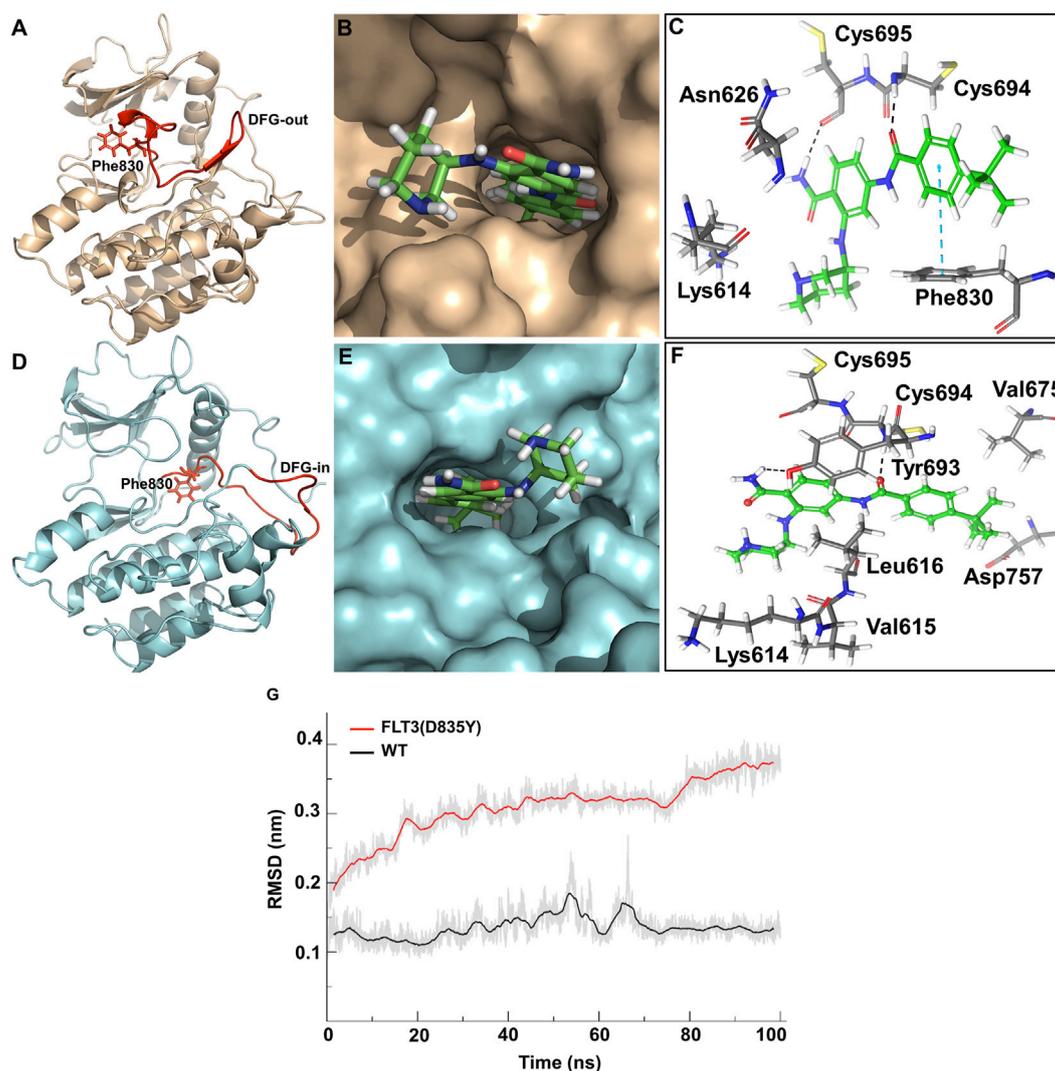


FIGURE 4

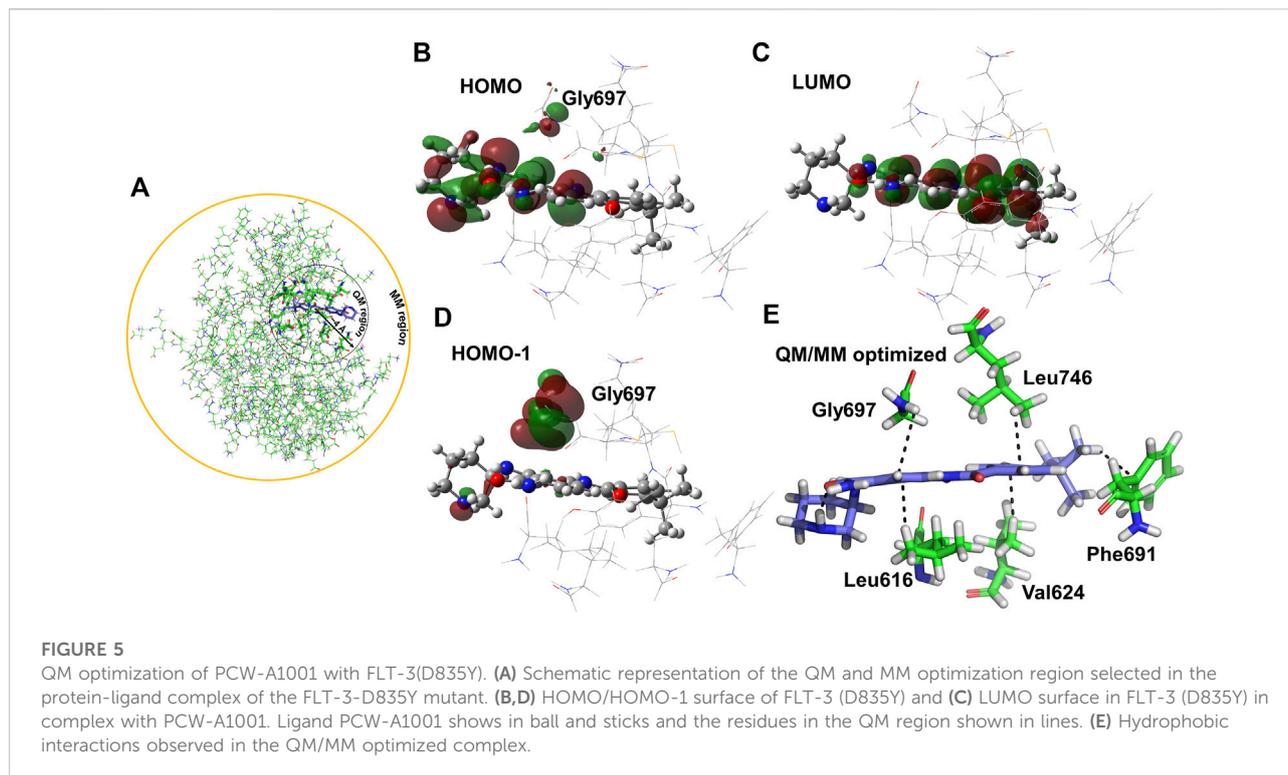
Binding analysis of PCW-A1001. (A) DFG-out conformation of FLT-3 wild type (B) Binding mode of PCW-A1001 with FLT-3 wild type (C) Crucial interactions observed in the FLT-3 wild-type (sticks). (D) DFG-in conformation of FLT-3 (D835Y) mutant, (E) Binding mode of PCW-A1001 with FLT-3 (D835Y), (F) Crucial interactions observed in the FLT-3 (D835Y) (sticks). Hydrogen bond interactions are represented in black dashed lines and pi-pi interactions in the blue dashed lines. (G) RMSD (Ca) plot of FLT-3 wild type and D835Y mutant.

TABLE 1 Binding free energy calculation for PCW-A1001 with WT and mutant FLT-3.

Compound	FLT-3 (WT)	FLT-3 (D835Y) mutant
PCW-A1001	-13.4 kcal/mol	-14.8 kcal/mol
PCW-1001	-7.2 kcal/mol	-8.07 kcal/mol

were intact in the protein-ligand complex, as with the MD structure. The ligand was stabilized at the binding site through hydrogen bonding and -CH- π and -NH- π interactions. The backbone -NH group from

Gly697 interacted with the phenyl ring of the ligand. In addition to hydrogen bonding interactions, -CH- π interactions were dominant in the complex formed between the ligand and protein. The Leu746, Phe691, Val624, and Leu616 residues were involved in the -CH- π interaction, as shown in Figure 5. The same pattern was also observed in the case of WT protein binding with the ligand, except for the -CH- π interactions with Ala642. The calculated interaction energies for the ligand and protein were -46.91 and -34.04 kcal/mol for the mutant and WT, respectively. These binding affinities were in good agreement with the free energy calculations for an explicit water environment.



The electron distribution was determined from the calculated orbital energies. The highest occupied molecular orbital (HOMO), HOMO-1, and lowest unoccupied molecular orbital (LUMO) were computed and are displayed in Figure 6. The HOMO surface was predominantly localized in the hydrogen bonding interaction regions of the ligand, whereas the LUMO surface was distributed evenly across the ligand (Figure 6). Gly697, which was involved in the -NH- π interaction with the ligand, contributed less to the HOMO, whereas HOMO-1 was highly localized on Gly697. The piperidine ring in the ligand formed an intramolecular hydrogen bond and stabilized the ligand orbitals. Thus, the HOMO was localized on and near the piperidine ring. Furthermore, atomic charges were calculated with natural population analysis. The sum of the atomic charges on the ligand was found to be 0.02 au. No significant charge transfer from the ligand to the protein was observed.

We also analyzed the Frontier orbitals of PCW-A1001, as shown in Figure 6. Frontier orbitals direct the mode of interaction between drugs and proteins. The HOMO and LUMO contribute to the chemical stability of the molecule. If the energy gap is zero or negligible, the molecule is highly reactive. PCW-A1001 was stable and showed an energy gap of 3.7 eV. The HOMO was localized on the phenyl ring, and the LUMO was distributed across two phenyl rings of PCW-A1001. This indicated that intramolecular charge transfer might enhance the stability of PCW-A1001. The molecular electrostatic potential illustrates the charge distribution of a molecule. This explains how one molecule can interact with another. The electrostatic potential helps determine the electrophilic

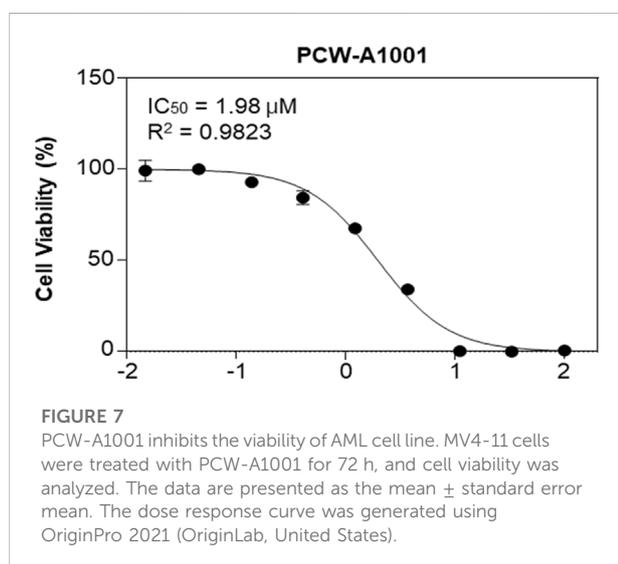
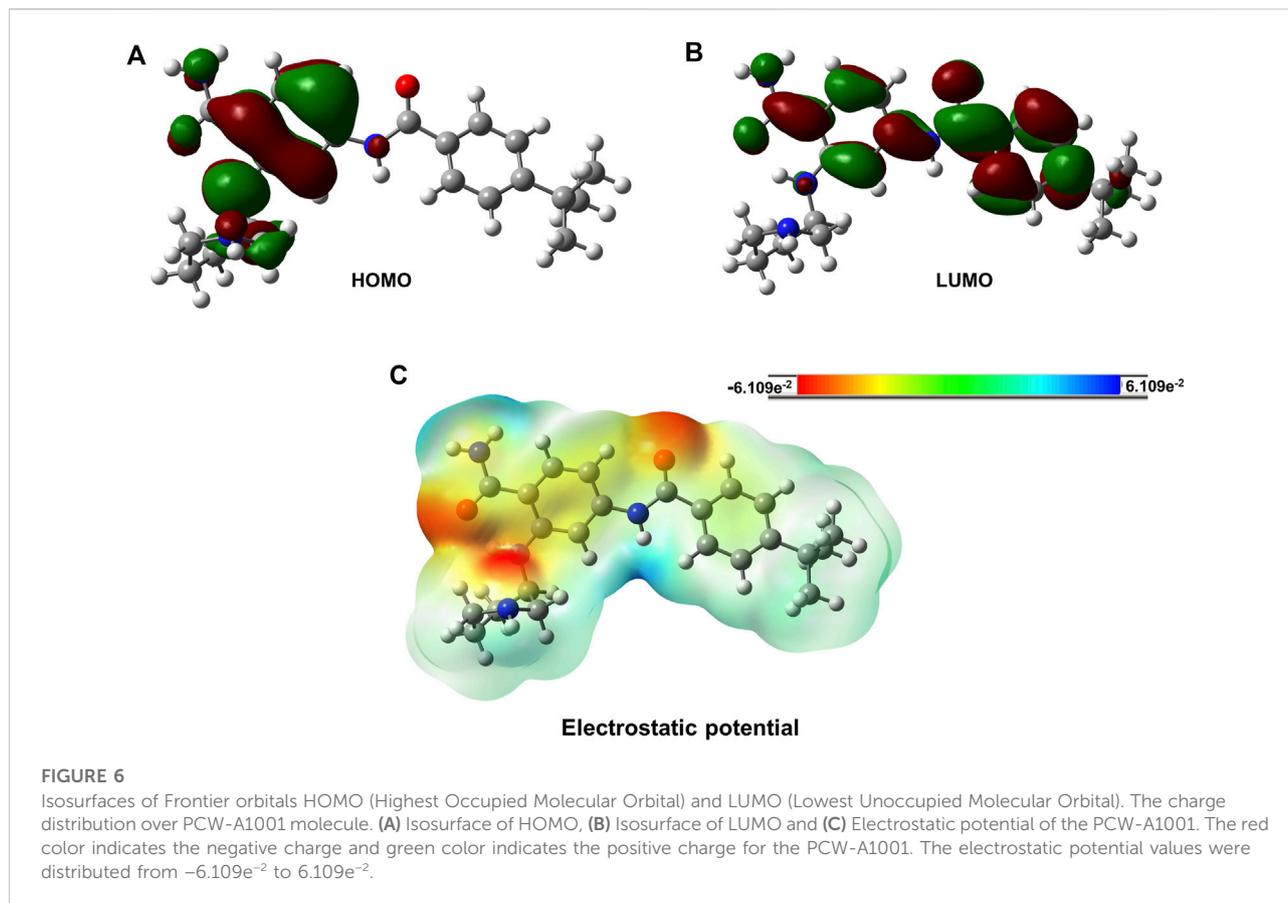
and nucleophilic sites involved in hydrogen bond formation. The calculated electrostatic potential surface is shown in Figure 6. The positive and negative potentials are indicated by blue and red colors, respectively. Atoms in the positive potential region act as electron acceptors, whereas atoms with a negative potential behave as electron donors during hydrogen bond formation with FLT-3. The aromatic phenyl rings involved in -CH- π interactions were found between the positive and negative potentials. The results show that the charge distribution over PCW-A1001 was favorable for interacting with the binding pocket of FLT-3.

Inhibition of MV4-11 and acute myeloid leukemia cell lines by PCW-A1001

MV4-11 cells and FLT-3-mutated AML cells (Quentmeier et al., 2003) were used to examine the anticancer activity of PCW-A1001. It inhibited the proliferation of MV4-11 cells, with an IC_{50} of 1.98 μ M, showing that PCW-A1001 has potent anticancer activity in AML cells (Figure 7).

In vitro kinase activity of PCW-A1001

Next, we performed an *in vitro* kinase assay to evaluate the inhibitory activity of PCW-A1001 for WT FLT-3 and the D835Y-mutant kinase. Interestingly, our data indicated that PCW-



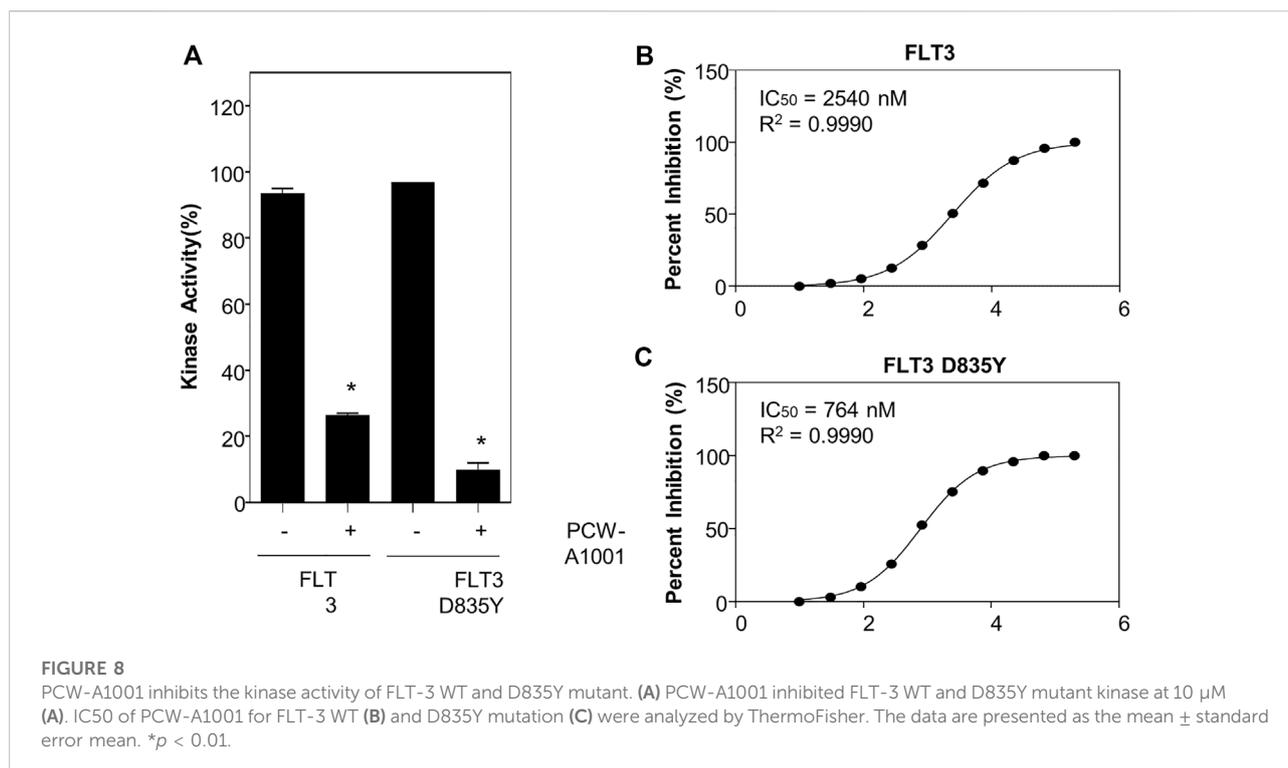
A1001 inhibited the mutant kinase more effectively than the WT (Figure 8A). The IC₅₀ determined from the kinase assay was 764 nM for the FLT-3 D835Y mutant, which was only one-third the IC₅₀ for WT (2.54 μ M) (Figures 8B,C). Thus, our data

provide proof-of-concept evidence for the AI-assisted *de novo* drug design approach.

Materials and methods

Reverse target prediction—a network-based approach

The input structures were used as SMILES structures and converted into fingerprints. We used six kinds of fingerprints: RDKit, MACCSkeys, AtomPair, Torsion, Morgan, and Morgan with Features (Landrum, 2016). A total of six fingerprints and six similarities were calculated and compared with the precalculated scores for the ligands in our protein–ligand interaction network database. A total of 10,647 compounds were used for the analysis among 262,327 compounds in the ChEMBL database. The list of compounds used with SMILES is included in the (Supplementary Table S2). In detail, the similarities between the input compound and all drugs in the network were calculated. Here, we used six similarity measures: Tanimoto, Dice, Sokal, Cosine, Kulczynsk, and McConnaughey. We selected drugs with high similarity scores (sum of similarity values). Using protein–drug relationships in the



network, the similarity score of a drug was assigned to all proteins connected to the drug. Then, we obtained a list of proteins with similarity scores for each fingerprint. Based on the similarity score, the list was sorted in descending order, and a rank value, which is a descending value from the maximum rank, was assigned. The rank score was calculated using the following formula:

$$\text{rank score} = \frac{\text{rank value}}{\sum \text{rank values}} \quad (1)$$

The proteins obtained a total rank score R , which was the sum of the rank score of each fingerprint. The rank score was further modified as \tilde{R} by applying the following formula and assigned to each protein. The modified total rank score showed the potential of the target protein. The network model was validated using eight known Bruton's tyrosine kinase (BTK) inhibitors, the model prediction rank was given (Supplementary Figure S4), and the details of the targets along with their inhibitors are given (Supplementary Table S1).

$$\tilde{R} = \frac{1}{R_{\max}} \sum_{\text{Fingerprint}} \text{rank score} \quad (2)$$

AI-assisted *de novo* design

A recurrent neural network (RNN) is a type of neural network that is widely used for natural language processing

(NLP) tasks from simple language processing to complex cheminformatics problems. RNNs have been successfully applied for protein structure and function predictions from sequences (Liu, 2017; Zhang et al., 2018), property predictions, fragment-based hit generation (Awale et al., 2019), and hit identification (Segler et al., 2018; Erikawa et al., 2021).

For *de novo* drug design, we successfully applied generative recurrent neural networks (RNN) containing long short-term memory (LSTM) cells (Merk et al., 2018). The model considers the SMILES strings for molecular representation and learns the patterns and their probabilities from pretraining for use in generation of the SMILES structures. We fine-tuned the generated structures (SMILES) for specific molecular targets or chemical series by employing transfer learning. The generative LSTM approach has proven helpful in low-data drug discovery, hit expansion, molecular design (fragment-based), and lead optimization (Gupta et al., 2018; Segler et al., 2018; Erikawa et al., 2021).

All deep learning models were applied using TensorFlow (v2.1, www.tensorflow.org) in Python (v3.7, www.python.org). We have used RDKit (www.rdkit.org) for most cheminformatics activities (property calculations, SMILES string validity calculations, molecular fingerprint calculations, and molecular clustering calculations). A detailed analysis of the generated SMILES strings was performed using the Jupyter notebook (www.anaconda.org).

RNNs were used as autoencoders, and the deep learning model employed for this study was initially pretrained to capture the structures of 438,552 bioactive small molecules (after carefully excluding the FLT-3 actives) retrieved from ChEMBL25 (KD, Ki, EC50, IC50 < 1 μ M) and represent them as simplified molecular-input line-entry system (SMILES) strings (Weininger, 1988). Using this pretrained model, we fine-tuned the model by transfer learning to bias *de novo* molecule generation toward the desired bioactivities of the templates (Merk et al., 2018). This fine-tuning step was employed to train the model for designing functional mimetics. Generated hits were further evaluated and benchmarked using Molecular sets (MOSES) (Polykovskiy et al., 2020).

Synthetic feasibility

The synthetic feasibility of each compound generated by AI was obtained with the retrosynthesis-associated fragment-based synthetic feasibility (RAFSF) score module. The fundamental idea of the module was that after cleaving synthetically meaningful bonds of the given compound, the bonds and the resulting fragments were searched from a bond/fragment space extracted from the ChEMBL (Mendez et al., 2019) or USPTO (Lowe, 2017) grants database in the same way. In the ChEMBL small molecule database, 1,917,863 molecules with molecular weights of less than 1,000 were used, as were 1,808,937 reactions from the USPTO grants database. To break bonds, we used the modified BRICS (Degen et al., 2008) module included in the rdkit (Landrum, 2016). If the bond/fragment from the given compound was not contained or rarely appeared in the bond/fragment space, a RAFSF score with a high value was assigned, meaning it was synthetically unfeasible. The RAFSF score is a value ranging from 1 (highly feasible synthesis) to 10 (highly unfeasible).

Protein preparation and modeling

The DFG-out WT FLT-3 protein structure was downloaded from the protein databank (www.rcsb.org) with PDB ID: 1RJB in the DFG-out conformation. Protein structures were prepared by correcting the bond orders, adding missing hydrogens, optimizing H-bonding with the protonation states of residues at pH 7.0, and restraining minimization for added hydrogens using the OPLS2005 forcefield of Protein Preparation Wizard (Sastry et al., 2013). The DFG-in conformation of the FLT-3 (D835Y) mutant was modeled using Modeler 9.25 (Sali and Blundell, 1993) with two templates, as reported previously (Ke et al., 2015). The first template was the DFG-out conformation of FLT-3 (PDB: 1RJB) with its DFG motif removed, and the other template was the DFG-in conformation of the colony-stimulating factor-1 receptor (CSF-1) crystal structure (PDB

id: 3LCD). These templates shared 93% and 63% sequence identity with the target protein, respectively. The model was subjected to loop refinement and minimization, followed by validation using a standard protocol discussed elsewhere (Sivakumar et al., 2013; Ke et al., 2015).

Ligand preparation and molecular docking

Hit compounds were initially optimized using the DFT method in Gaussian 16 with B3LYP functionals and the 6-31G** basis set (Tirado-Rives and Jorgensen, 2008). The antechamber obtained GAFF atom types with RESP charges from the Gaussian output file. The atom types and all needed parameters for the ligand were obtained from the above process along with parm and leap (Shirts et al., 2017). Molecular docking was carried out using Glide XP (Friesner et al., 2006) with default parameters; initially, the receptor was prepared with a grid box set covering the centroid of the active site, followed by flexible ligand sampling of the ligand docking.

Molecular dynamics simulations

The stabilities of the complexes were studied by MD simulations using Gromacs 2019 (Abraham et al., 2015). The Charmm36 force field (Huang and MacKerell, 2013) was used for the protein parameters. The protein–ligand complexes were solvated explicitly using the TIP3P water model inside the cubic box, and their sizes extended 0.1 nm away from the protein on the edges of the box in each direction. The overall charge of the system was neutralized by adding a 0.15 M salt (Na^+Cl^-). All simulations were carried out on GPU-enabled Linux clusters. The entire system was minimized with a maximum step size of 50,000 until the maximum force was less than 10 kJ/mol. The system was then equilibrated for 5 ns under NVT conditions with temperature coupling for two separate groups, protein–ligand and water-ions, at 300 K. The Lincs algorithm was used to constrain the bonds of the hydrogen atoms (Hess et al., 1997). A Berendsen thermostat and V-rescale were used to keep the temperature and pressure constant, respectively (Lemak and Balabaev, 1994; Bussi et al., 2007). The cutoff distances for Coulomb and van der Waals interactions were set as 1.2 nm. The particle mesh Ewald method (PME) was used to calculate the long-range electrostatic interactions (Darden et al., 1993). The final production run was carried out for 100 ns at a temperature of 300 K and a pressure of 1 bar.

Free energy calculation

The binding free energies for protein and ligand complexes were calculated in an explicit water environment by employing the alchemical method (Supplementary Figure S5). The final

snapshot from the MD simulations (100 ns) was used as the starting point for the free energy simulations. The alchemical method involves two steps: 1) decoupling of the ligand from the protein–ligand complex in an explicit water environment and 2) decoupling of the ligand from the water environment. The decoupling process includes turning off the van der Waals and electrostatic interactions responsible for complex formation (protein–ligand or water–ligand) with the help of the coupling parameter (λ). First, electrostatic interactions were turned off slowly, while the van der Waals interactions were still present. Then, the van der Waals interactions between the protein and ligand (water and ligand) were turned off using the coupling parameter (λ). The electrostatic interactions were turned off by changing λ (0 0.25 0.5 0.75 1.0) from 0 to 1 with a step size of $\Delta\lambda = 0.25$, and the van der Waals interactions were turned off with nonuniformly distributed values of λ (0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1.0). The same procedure was applied to decouple the ligands from the protein–ligand complex and ligands from the water environment. Therefore, 21 windows, each of 1 ns, were employed to decouple the ligand from the protein–ligand and water–ligand complexes. The free energy difference between the two end states was calculated using the Bennett acceptance ratio (BAR) method (Bennett, 1976). The BAR method is used to estimate the free energy difference between two states with the following equation:

$$\left\langle \frac{1}{1 + \exp\{\beta(\Delta U_{ij} - \Delta G)\}} \right\rangle_i = \left\langle \frac{1}{1 + \exp\{\beta(-\Delta U_{ij} + \Delta G)\}} \right\rangle_j \quad (3)$$

where β is the reciprocal of the thermodynamic temperature, ΔG is the free energy difference between states i and j , and $\Delta U_{ij} = U_j - U_i$ is the potential energy difference.

At each λ -point, the structures were subjected to energy minimization using the steepest descent method. Using Langevin dynamics, the resulting structures were equilibrated in an isothermal-isobaric (NPT) ensemble at 300 K and at a pressure of 1 bar. The rest of the simulation protocol was similar to the protocol followed in the classic MD section.

QM/MM approach

The final snapshots of protein and ligand complexes determined from MD simulations were optimized in the gas phase at the (B3LYP-D3/6-31G*)/Universal force field level of theory with the help of the Gaussian16 package. It has been found in earlier studies that density functionals such as M06-2X, B3LYP-D, and ω B97XD are suitable for investigating noncovalent interactions. Hence, in all calculations, the QM region was optimized with dispersion-corrected B3LYP with the Grimme empirical dispersion functional (B3LYP-D3)

using the 6-31G* basis set. The ligand and surrounding region within 4 Å were treated as the QM region, and the remaining parts were considered the MM region. We extracted only the QM region from the optimized geometries and added terminal hydrogens to calculate binding affinities. The resulting structures were used to calculate the interaction energies with the supermolecule approach at the B3LYP-D3/6-31G* level of theory.

$$IE = EC - (EM1 + EM2) \quad (4)$$

where IE is the interaction energy of the complex, EC is the energy of the complex, EM1 is the energy of the protein part of the complex, and EM2 is the energy of the ligand in the complex. All IEs were corrected for basis set superposition error (BSSE) using the counterpoise method suggested by Boys and Bernadi (Gutowski et al., 1993), as implemented in the Gaussian16 package (Frisch et al., 2016).

Selectivity

In total, 49 kinases were evaluated *via* selectivity score calculation, and those 49 kinases were previously used for the actual kinase panel assay. Representative PDB structures for the 49 kinases were extracted from the RCSB Protein Data Bank (<https://www.rcsb.org/>). The ligands were docked to the binding pocket of each PDB using AutoDock-Vina (Trott and Olson, 2010). The resulting docking score was rescaled to observe and compare compound trends. Quantitatively, the selectivity score was calculated to measure the overall selectivity across different kinase families. A lower selectivity score indicates better selectivity for the tested compound.

Selectivity score (S)

$$= \frac{\text{number of kinases with rescaled docking score less than 0.5}}{\text{total number of kinases tested}} \quad (5)$$

Cell culture and cell viability assay

MV4-11 cells were purchased from American Type Culture Collection (ATCC, VA, United States). The cells were passaged for less than 1 month, and mycoplasma infection was checked by PCR once a week. The growth medium was Iscove's Modified Dulbecco's Medium (IMDM; ThermoFisher, United States) supplemented with 10% fetal bovine serum (FBS; Corning, United States) and 1% penicillin/streptomycin (GenDEPOT, United States). The cells were maintained in a humidified atmosphere with 5% CO₂ at 37°C. Cell viability was determined using the WST-8 assay (Cyto XTM cell viability

assay kit; LPS solution, Daejeon, South Korea) in accordance with the manufacturer's protocol.

Synthesis of PCW-A1001

Unless otherwise stated, all reactions were performed under an inert (N₂) atmosphere. Reagents and solvents were reagent grade and purchased from Sigma-Aldrich, Alfa Aesar, and Combi-Blocks. Anhydrous solvents were purchased from Sigma-Aldrich and used as provided. Reactions were monitored by TLC and visualized with a UV lamp and/or KMnO₄ staining. Silica gel 60 (230–400 mesh, Merck) was used for flash column chromatography. ¹H and ¹³C NMR spectra were recorded on BRUKER Ultrashield 300 and 400 MHz NMR spectrometers at 25°C. Chemical shifts are reported in parts per million (ppm). Data for ¹H NMR are reported as follows: chemical shift (δ ppm) [multiplicity, coupling constant (Hz), integration]. Multiplicities are reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, dd = doublet of doublets, m = multiplet. Data from ¹³C spectra are reported as chemical shifts (δ ppm). The residual solvent peak was used as an internal reference. Mass spectra were obtained on Acquity™ Waters A06UPD9BM and Agilent Technologies SG12109048 systems. Prior to biological testing, the final compound was confirmed to be > 98% pure by UPLC chromatography using a Waters ACQUITY H-class system fitted with a C₁₈ reverse-phase column (ACQUITY UPLC BEH C₁₈: 2.1 mm × 50 mm, Part No. 186002350) according to the following eluent conditions: (A) H₂O + 0.1% formic acid, (B) CH₃CN + 0.1% formic acid, (C) MeOH + 0.1% formic acid; (I) a gradient of 95% A to 95% B over 5 min; and (II) a gradient of 95% A to 95% C over 5 min.

Conclusion

In this work, we used an AI-assisted *de novo* drug design (LSTM) approach to identify a novel FLT-3 inhibitor that selectively targets the FLT-3 (D835Y) mutant. The deep learning model was pretrained on a known bioactive chemical space (ChEMBL22), and the generated hits were fine-tuned using our in-house FLT-3 inhibitors. The generated hits were further evaluated and filtered using various parameters focusing on their novelty, similarities, diversities, etc. We further evaluated the toxicities of the *de novo* molecules with our in-house program Pharmulator™. Among the screened hits, only 146 compounds passed the toxicity filters. The binding affinities, conformations and interaction patterns of these screened compounds were studied with WT FLT-3 and its mutant (D835Y). Since the FLT-3 (D835Y) mutant structure in the DFG-in conformation was unavailable, we modeled the protein to validate the compounds in terms of the binding interactions. The

stabilities of complexes were further validated qualitatively with MD simulations and quantitatively with free energy calculations.

The top compound, named PCW-A1001, was considered for synthesis and screening studies. The anticancer activity was tested against MV4-11 cells to verify the effectiveness of these compounds in AML treatment. PCW-A1001 was found to be a promising inhibitor of FLT3, and it showed an IC₅₀ of 764 nM against the FLT-3 (D835Y) mutant and 2.54 μM against WT FLT-3. PCW-A1001 also showed an IC₅₀ of 1.98 μM against MV4-11-cell line screening. We successfully implemented reverse network theory and AI-based *de novo* design strategies and identified a potential inhibitor of the FLT3/FLT3 (D835Y) mutant, PCW-A1001. AI generated a hit, PCW-A1001 exhibited better activity than the parent compound, PCW-1001. Further fine-tuning of PCW-A1001 is in progress to optimize the selectivity and activity and will be reported in due course.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

SJ, DS, and SM—performed and analyzed the computational part, JC and SL—designed and analyzed the network part, MJ and SB—performed the AI part, ES, KL, and K-YJ—performed synthesis and characterization of the compound, MK and J-SK—performed the Cell assay, JH—analysis of experimental data, SW—conceived and designed the experiments, DS, SM, SB, K-YJ, and J-SK—writing-original draft, and SW—conceptualization and draft review.

Conflict of interest

Authors SJ, DS, SM, JC, SL, MJ, SB, JH, and SW were employed by PharmCADD.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their

affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al. (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1 (2), 19–25. doi:10.1016/j.softx.2015.06.001
- Ambinder, A. J., and Levis, M. (2021). Potential targeting of FLT3 acute myeloid leukemia. *Haematologica* 106, 671–681. doi:10.3324/haematol.2019.240754
- Assouline, S., Cocolakis, E., and Borden, K. L. B. (2012). The development of novel therapies for the treatment of acute myeloid leukemia (AML). *Cancers* 4, 1161–1179. doi:10.3390/cancers4041161
- Auclair, D., Miller, D., Yatsula, V., Pickett, W., Carter, C., Chang, Y., et al. (2007). Antitumor activity of sorafenib in FLT3-driven leukemic cells. *Leukemia* 21, 439–445. doi:10.1038/sj.leu.2404508
- Awale, M., Sirockin, F., Stiefl, N., and Reymond, J.-L. (2019). Drug analogs from fragment-based long short-term memory generative neural networks. *J. Chem. Inf. Model.* 59, 1347–1356. doi:10.1021/acs.jcim.8b00902
- Banks, R. D., Blake, C. C. F., Evans, P. R., Haser, R., Rice, D. W., Hardy, G. W., et al. (1979). Sequence, structure and activity of phosphoglycerate kinase: A possible hinge-bending enzyme. *Nature* 279, 773–777. doi:10.1038/279773a0
- Bennett, C. H. (1976). Efficient estimation of free energy differences from Monte Carlo data. *J. Comput. Phys.* 22, 245–268. doi:10.1016/0021-9991(76)90078-4
- Berger, S. I., and Iyengar, R. (2009). Network analyses in systems pharmacology. *Bioinformatics* 25, 2466–2472. doi:10.1093/bioinformatics/btp465
- Brown, P., and Small, D. (2004). FLT3 inhibitors: A paradigm for the development of targeted therapeutics for paediatric cancer. *Eur. J. Cancer* 40, 707–721. doi:10.1016/j.ejca.2003.08.030
- Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126, 014101. doi:10.1063/1.2408420
- Chang Hsu, Y., Ke, Y.-Y., Shiao, H.-Y., Lee, C.-C., Lin, W.-H., Chen, C.-H., et al. (2014). Facile identification of dual FLT3–aurora A inhibitors: A computer-guided drug design approach. *ChemMedChem* 9, 953–961. doi:10.1002/cmdc.201300571
- Corey, E. J., and Wipke, W. T. (1969). Computer-assisted design of complex organic syntheses. *Science* 166, 178–192. doi:10.1126/science.166.3902.178
- Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98, 10089–10092. doi:10.1063/1.464397
- Degen, J., Wegscheid-Gerlach, C., Zaliani, A., and Rarey, M. (2008). On the art of compiling and using “drug-like” chemical fragment spaces. *ChemMedChem* 3, 1503–1507. doi:10.1002/cmdc.200800178
- Del Zotto, G., Luchetti, F., and Zamai, L. (2001). CD135. *J. Biol. Regul. Homeost. Agents* 15, 103–106.
- Erikawa, D., Yasuo, N., and Sekijima, M. (2021). Mermaid: An open source automated hit-to-lead method based on deep reinforcement learning. *J. Cheminform.* 13, 94. doi:10.1186/s13321-021-00572-6
- Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., et al. (2006). Extra precision Glide: Docking and scoring incorporating a model of hydrophobic enclosure for Protein–Ligand complexes. *J. Med. Chem.* 49, 6177–6196. doi:10.1021/jm051256o
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian 16 revision C.01*. Willingford, CT.
- Gebru, M. T., and Wang, H.-G. (2020). Therapeutic targeting of FLT3 and associated drug resistance in acute myeloid leukemia. *J. Hematol. Oncol.* 13, 155. doi:10.1186/s13045-020-00992-1
- Grafone, T., Palmisano, M., Nicci, C., and Storti, S. (2012). An overview on the role of FLT3-tyrosine kinase receptor in acute myeloid leukemia: Biology and treatment. *Oncol. Rev.* 6, e8. doi:10.4081/oncol.2012.e8
- Grunwald, M. R., and Levis, M. J. (2013). FLT3 inhibitors for acute myeloid leukemia: A review of their efficacy and mechanisms of resistance. *Int. J. Hematol.* 97, 683–694. doi:10.1007/s12185-013-1334-8
- Guney, E., Menche, J., Vidal, M., and Barabasi, A.-L. (2016). Network-based *in silico* drug efficacy screening. *Nat. Commun.* 7, 10331. doi:10.1038/ncomms10331
- Gupta, A., Müller, A. T., Huisman, B. J. H., Fuchs, J. A., Schneider, P., and Schneider, G. (2018). Generative recurrent networks for de novo drug design. *Mol. Inf.* 37, 1700111. doi:10.1002/minf.201700111
- Gutowski, M., van Duijneveldt-van de Rijdt, J. G. C. M., van Lenthe, J. H., and van Duijneveldt, F. B. (1993). Accuracy of the Boys and Bernardi function counterpoise method. *J. Chem. Phys.* 98, 4728–4737. doi:10.1063/1.465106
- Hayakawa, F., Towatari, M., Kiyoi, H., Tanimoto, M., Kitamura, T., Saito, H., et al. (2000). Tandem-duplicated Flt3 constitutively activates STAT5 and MAP kinase and introduces autonomous cell growth in IL-3-dependent cell lines. *Oncogene* 19, 624–631. doi:10.1038/sj.onc.1203354
- Heinrich, M. C., Griffith, D., McKinley, A., Patterson, J., Presnell, A., Ramachandran, A., et al. (2012). Crenolanib inhibits the drug-resistant PDGFRA D842V mutation associated with imatinib-resistant gastrointestinal stromal tumors. *Clin. Cancer Res.* 18, 4375–4384. doi:10.1158/1078-0432.CCR-12-0625
- Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. G. E. M. (1997). Lincs: A linear constraint solver for molecular simulations. *J. Comput. Chem.* 18, 1463–1472. doi:10.1002/(sici)1096-987x(199709)18:12<1463::aid-jcc4>3.0.co;2-h
- Hessler, G., and Baringhaus, K.-H. (2018). Artificial intelligence in drug design. *Molecules* 23, E2520. doi:10.3390/molecules23102520
- Huang, J., and MacKerell, A. D. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* 34, 2135–2145. doi:10.1002/jcc.23354
- Kang, M., Pandit, N., Kim, A.-Y., Cho, S. J., Kwon, Y.-J., Ahn, J., et al. (2022). PCW-1001, a novel pyrazole derivative, exerts antitumor and radio-sensitizing activities in breast cancer. *Front. Oncol.* 12, 835833. doi:10.3389/fonc.2022.835833
- Kazi, J. U., and Rönstrand, L. (2019). FMS-Like tyrosine kinase 3/FLT3: From basic science to clinical implications. *Physiol. Rev.* 99, 1433–1466. doi:10.1152/physrev.00029.2018
- Ke, Y.-Y., Singh, V. K., Coumar, M. S., Hsu, Y. C., Wang, W.-C., Song, J.-S., et al. (2015). Homology modeling of DFG-in FMS-like tyrosine kinase 3 (FLT3) and structure-based virtual screening for inhibitor identification. *Sci. Rep.* 5, 11702. doi:10.1038/srep11702
- Kindler, T., Lipka, D. B., and Fischer, T. (2010). FLT3 as a therapeutic target in AML: Still challenging after all these years. *Blood* 116, 5089–5102. doi:10.1182/blood-2010-04-261867
- Landrum, G. (2016). RDKit: Open-Source cheminformatics software. Available at: https://github.com/rdkit/rdkit/releases/tag/Release_2016_09_4.
- Lemak, A. S., and Balabaev, N. K. (1994). On the berendsen thermostat. *Mol. Simul.* 13, 177–187. doi:10.1080/08927029408021981
- Leung, A. Y. H., Man, C.-H., and Kwong, Y.-L. (2013). FLT3 inhibition: A moving and evolving target in acute myeloid leukaemia. *Leukemia* 27, 260–268. doi:10.1038/leu.2012.195
- Liang, D.-C., Shih, L.-Y., Hung, I.-J., Yang, C.-P., Chen, S.-H., Jaing, T.-H., et al. (2003). FLT3-TKD mutation in childhood acute myeloid leukemia. *Leukemia* 17, 883–886. doi:10.1038/sj.leu.2402928
- Lin, W.-H., Jiaang, W.-T., Chen, C.-W., Yen, K.-J., Hsieh, S.-Y., Yen, S.-C., et al. (2012). BPR1J-097, a novel FLT3 kinase inhibitor, exerts potent inhibitory activity against AML. *Br. J. Cancer* 106, 475–481. doi:10.1038/bjc.2011.564
- Liu, X. (2017). Deep recurrent neural network for protein function prediction from sequence. *arXiv* 2017, 08318. doi:10.48550/arXiv.1701.08318
- Lowe, D. (2017). *Chemical reactions from US patents (1976-Sep2016)*. doi:10.6084/m9.figshare.5104873.v1
- Luo, Y., Peng, J., and Ma, J. (2022). Next decade’s AI-based drug development Features tight integration of data and computation. *Health Data Sci.* 2022, 1–3. doi:10.34133/2022/9816939

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.1072028/full#supplementary-material>

- Mak, K.-K., and Pichika, M. R. (2019). Artificial intelligence in drug development: Present status and future prospects. *Drug Discov. Today* 24, 773–780. doi:10.1016/j.drudis.2018.11.014
- McDonnell, L. M., Kernohan, K. D., Boycott, K. M., and Sawyer, S. L. (2015). Receptor tyrosine kinase mutations in developmental syndromes and cancer: Two sides of the same coin. *Hum. Mol. Genet.* 24, R60–R66. doi:10.1093/hmg/ddv254
- Mendez, D., Gaulton, A., Bento, A. P., Chambers, J., De Veij, M., Félix, E., et al. (2019). ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Res.* 47, D930–D940. doi:10.1093/nar/gky1075
- Merk, D., Friedrich, L., Grisoni, F., and Schneider, G. (2018). De novo design of bioactive small molecules by artificial intelligence. *Mol. Inf.* 37, 1700153. doi:10.1002/minf.201700153
- O'Hare, T., Shakespeare, W. C., Zhu, X., Eide, C. A., Rivera, V. M., Wang, F., et al. (2009). AP24534, a pan-BCR-ABL inhibitor for chronic myeloid leukemia, potently inhibits the T315I mutant and overcomes mutation-based resistance. *Cancer Cell* 16, 401–412. doi:10.1016/j.ccr.2009.09.028
- Polykovskiy, D., Zhebrak, A., Sanchez-Lengeling, B., Golovanov, S., Tatanov, O., Belyaev, S., et al. (2020). Molecular sets (MOSES): A benchmarking platform for molecular generation models. *Front. Pharmacol.* 11, 565644. doi:10.3389/fphar.2020.565644
- Quentmeier, H., Reinhardt, J., Zaborski, M., and Drexler, H. G. (2003). FLT3 mutations in acute myeloid leukemia cell lines. *Leukemia* 17, 120–124. doi:10.1038/sj.leu.2402740
- Rosnet, O., Bühring, H. J., Marchetto, S., Rappold, I., Lavagna, C., Sainy, D., et al. (1996). Human FLT3/FLK2 receptor tyrosine kinase is expressed at the surface of normal and malignant hematopoietic cells. *Leukemia* 10, 238–248.
- Sali, A., and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815. doi:10.1006/jmbi.1993.1626
- Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013). Protein and ligand preparation: Parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided. Mol. Des.* 27, 221–234. doi:10.1007/s10822-013-9644-8
- Savage, D. G., and Antman, K. H. (2002). Imatinib mesylate — a new oral targeted therapy. *N. Engl. J. Med.* 346, 683–693. doi:10.1056/NEJMra013339
- Schittenhelm, M. M., Yee, K. W. H., Tyner, J. W., McGreevey, L., Haley, A. D., Town, A., et al. (2006). FLT3 K663Q is a novel AML-associated oncogenic kinase: Determination of biochemical properties and sensitivity to Sunitinib (SU11248). *Leukemia* 20, 2008–2014. doi:10.1038/sj.leu.2404374
- Scholl, S., Fleischmann, M., Schnetzke, U., and Heidel, F. H. (2020). Molecular mechanisms of resistance to FLT3 inhibitors in acute myeloid leukemia: Ongoing challenges and future treatments. *Cells* 9, 112493. doi:10.3390/cells9112493
- Segler, M. H. S., Kogej, T., Tyrchan, C., and Waller, M. P. (2018). Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Cent. Sci.* 4, 120–131. doi:10.1021/acscentsci.7b00512
- Shirts, M. R., Klein, C., Swails, J. M., Yin, J., Gilson, M. K., Mobley, D. L., et al. (2017). Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset. *J. Comput. Aided. Mol. Des.* 31, 147–161. doi:10.1007/s10822-016-9977-1
- Sivakumar, D., Gorai, B., and Sivaraman, T. (2013). Screening efficient BH3-mimetics to hBcl-B by means of peptidodynamic method. *Mol. Biosyst.* 9, 700–712. doi:10.1039/C2MB25195G
- Smith, B. D., Levis, M., Beran, M., Giles, F., Kantarjian, H., Berg, K., et al. (2004). Single-agent CEP-701, a novel FLT3 inhibitor, shows biologic and clinical activity in patients with relapsed or refractory acute myeloid leukemia. *Blood* 103, 3669–3676. doi:10.1182/blood-2003-11-3775
- Smith, C. C., Wang, Q., Chin, C.-S., Salerno, S., Damon, L. E., Levis, M. J., et al. (2012). Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature* 485, 260–263. doi:10.1038/nature11016
- Stirewalt, D. L., and Radich, J. P. (2003). The role of FLT3 in hematopoietic malignancies. *Nat. Rev. Cancer* 3, 650–665. doi:10.1038/nrc1169
- Stone, R. M., De Angelo, J., Galinsky, I., Estey, E., Klimek, V., Grandin, W., et al. (2004). PKC 412 FLT3 inhibitor therapy in AML: Results of a phase II trial. *Ann. Hematol.* 83 (1), S61–S64. doi:10.1007/s00277-004-0850-2
- Tirado-Rives, J., and Jorgensen, W. L. (2008). Performance of B3LYP density functional methods for a large set of organic molecules. *J. Chem. Theory Comput.* 4, 297–306. doi:10.1021/ct700248k
- Trott, O., and Olson, A. J. (2010). AutoDock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem.* 31, 455–461. doi:10.1002/jcc.21334
- Weininger, D. (1988). SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Model.* 28, 31–36. doi:10.1021/ci00057a005
- Wodicka, L. M., Ciceri, P., Davis, M. I., Hunt, J. P., Floyd, M., Salerno, S., et al. (2010). Activation state-dependent binding of small molecule kinase inhibitors: Structural insights from biochemistry. *Chem. Biol.* 17, 1241–1249. doi:10.1016/j.chembiol.2010.09.010
- Wu, Z., Wang, Y., and Chen, L. (2013). Network-based drug repositioning. *Mol. Biosyst.* 9, 1268–1281. doi:10.1039/C3MB25382A
- Yamamoto, Y., Kiyoi, H., Nakano, Y., Suzuki, R., Kodera, Y., Miyawaki, S., et al. (2001). Activating mutation of D835 within the activation loop of FLT3 in human hematologic malignancies. *Blood* 97, 2434–2439. doi:10.1182/blood.v97.8.2434
- Zarrinkar, P. P., Gunawardane, R. N., Cramer, M. D., Gardner, M. F., Brigham, D., Belli, B., et al. (2009). AC220 is a uniquely potent and selective inhibitor of FLT3 for the treatment of acute myeloid leukemia (AML). *Blood* 114, 2984–2992. doi:10.1182/blood-2009-05-222034
- Zhang, B., Li, J., and Lü, Q. (2018). Prediction of 8-state protein secondary structures by a novel deep learning architecture. *BMC Bioinforma.* 19, 293. doi:10.1186/s12859-018-2280-5
- Zhang, S., and Broxmeyer, H. E. (1999). p85 subunit of PI3 kinase does not bind to human Flt3 receptor, but associates with SHP2, SHIP, and a tyrosine-phosphorylated 100-kDa protein in Flt3 ligand-stimulated hematopoietic cells. *Biochem. Biophys. Res. Commun.* 254, 440–445. doi:10.1006/bbrc.1998.9959
- Zhavoronkov, A., Ivanenkov, Y. A., Aliper, A., Veselov, M. S., Aladinskiy, V. A., Aladinskaya, A. V., et al. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat. Biotechnol.* 37, 1038–1040. doi:10.1038/s41587-019-0224-x