Identifying the Novel Pyrimidine-Based CDK2 Inhibitors as Anticancer Agents Using Text-Mining and Combined Molecular Modeling Approaches

Kader ŞAHİN[1]* ✉ iD, Serdar DURDAĞI[1] ✉ iD

[1]Computational Biology and Molecular Simulations Laboratory, Department of Biophysics, School of Medicine, Bahcesehir University, Istanbul, Turkey

Abstract: The cycline-dependent kinase (CDK) protein is a vital target used in anti-cancer drug designing studies. Many FDA-approved drugs, including anti-cancer drugs, use pyrimidine as a crucial fragment. In the current study, a small molecule database (Specs SC) that text mining studies include more than 210.000 compounds, and we filtered 6668 molecules that carry "pyrimidine" fragments. We then screened these compounds at the binding pocket of CDK-2 target using molecular docking and molecular dynamics (MD) simulations approaches. Besides, we compared the binding free energies of screened compounds with pyrimidine-based FDA-approved anti-cancer drug Abemaciclib which targets CDK. Based on the comparison of docking scores of screened compounds, we used top-7 hits in 100 ns MD simulations. We also applied the same MD simulations protocol (100 ns) to the Abemaciclib-bound CDK-2 complex structure. We then calculated the average Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) energies. Our MM/GBSA results showed that 6 out of 7 compounds have better MM/GBSA scores than FDA-approved compound Abemaciclib. Thus, together with a combination of text mining and integrated molecular modeling approaches, we identified novel pyrimidine-based hits against CDK.

*Corresponding author. E-mail: sahkader@hotmail.com.

INTRODUCTION

Cycline-dependent kinases (CDKs) are the catalytic subunits of a large family of serine/threonine protein kinases, having crucial roles in the regulation of cell cycle progression. There are two main groups in the family: Cell cycle CDKs and transcriptional CDKs (1). We often observe the over-expression of CDKs in many cancer types; thus, their inhibition may lead to cell cycle arrest or apoptosis (2). Although the human genome encodes 21 CDKs, researchers have shown that only seven (CDK1-

4, 6, 10, 11) of them indeed play a direct role in the progression of the cell cycle. Thus, targeting the CDK enzyme function is a rational approach to prevent unregulated cancer cell proliferation, and the authors considered these targets in many anti-cancer drug design studies (3, 4). Pyrimidine derivatives have attracted considerable interest in recent years in diverse biological problems (5, 6). Experts use pyrimidine derivatives in many different biological applications, including anti-tumor (7), antiviral (7, 8), antioxidant (9), antifungal (10), and hepatoprotective (11) roles. FDA-approved drug Brigatinib (marketed as Alunbrig) is a pyrimidine-based small-molecule used in targeted cancer therapy (12). Another FDA-approved drug Abemaciclib (commercial terms Verzenio and Verzenios) specifically targets CDK. Physicians mainly use them in advanced or metastatic breast cancer treatment. However, it has some adverse side effects such as diarrhea, nausea, vomiting, leukopenia (i.e., low white blood cell count) including neutropenia, anemia, thrombocytopenia (i.e., low platelet count), stomach pain, fatigue, decreased appetite, and headache (13, 14). Therefore, in the current study, pyrimidine-based new lead compounds targeting CDK were investigated using integrated text-mining and molecular modeling approaches. For this aim, we used MarvinSketch software (15) in 210,596 molecules obtained from Specs-SC drug-like small molecule database in an IUPAC text file format. We then screened this text file by a Python-based text mining in-house script and identified 6668 molecules as compounds that include the "pyrimidine" phrase in their text files. Moreover, we used these filtered compounds at the active site of the CDK-2 enzyme, and screened compounds were ranked based on their interaction energy scores. We used top-7 compounds that show better docking scores than FDA-approved drug Abemaciclib in molecular dynamics (MD) simulations. Trajectory frames collected throughout the MD simulations were used in Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) calculations and based on average MM/GBSA scores, we identified novel 6 pyrimidine-based small molecules that have better MM/GBSA scores than Abemaciclib (Table 1).

MATERIALS AND METHODS

Data Collection

Text mining helps find significant molecules by quickly screening them from large databases using keywords (16). In this study, 210,596 molecules obtained from Specs-SC were prepared in the IUPAC text file format using MarvinSketch software. We then screened this text file to find pyrimidine-based molecules using a Python-based text mining and obtained 6,668 molecules.

Ligand Preparation

Ligand preparation consists of the analysis of a 3D structure of a ligand, generation, and optimization steps. We have prepared 6,668 compounds using the OPLS-2005 forcefield (17) LigPrep module (18) of the Maestro Molecular Modeling Suit. A problem to be addressed is the ionization of the ligand in physiological conditions. We have used the Epik module (18) at the physiological pH of 7.4 for potential ionization states. We also produced all possible stereoisomers and tautomers and considered 32 structures per ligand.

Protein Preparation

We obtained the CDK-2 target from the protein data bank (PDB), and a 2XMY PDB-coded protein structure was used (19). We fixed the missing side chains, backbone atoms, and loops with Prime (20). For protonation states, structural optimization and minimization, the PROPKA and OPLS-2005 force field were used, respectively (21). We considered water molecules around 5.0 Å of the co-crystallized ligand in the docking simulations.

Molecular Docking Simulations

Molecular docking studies investigate interactions occurring in protein-protein or ligand-protein complexes and rank candidate poses according to predicted affinity scoring functions (22). Docking processes predict ligands to bind with the most appropriate conformation in the binding pocket of the target protein using different algorithms. We used Glide to carry out the molecular docking between the investigated ligands and CDK-2 protein to obtain the binding poses as well as docking scores of the studied ligands (23). The conformations obtained during the docking were ranked using the Glide score function (24). Residues set the binding pocket within 10 Å vicinity of a co-crystallized ligand. All ligands were initially docked into the binding pocket of CDK using a grid-based docking program Glide standard precision (SP) of Maestro Molecular Modeling pocket (20, 23) and requested 10 docking poses for each ligand (25, 26).

Molecular Dynamics (MD) Simulations

There is evidence that some systems require MD simulations to discover proper binding fit (27). Long MD simulations can find more and energetically favorable configurations. MD simulations were implemented up to 100 ns using Desmond V.4.9 to investigate the conformational stability of the complexes of selected hit molecules with CDK-2 (28). We solvated the complex structures in the orthorhombic simple point charge (SPC) water model (29).

Moreover, we neutralized the systems with counter ions (0.15 M NaCl solution). We set the system as Lennard-Jones interactions cutoff of 10 Å on periodic boundary conditions (23). We used a 2.0 fs time step in the integration steps.

We have used Nose-Hoover thermostat (30) and Martyna-Tobias Klein protocols (31) to control the temperature and pressure of the systems at 310 K and 1 bar, respectively.

Molecular Mechanics/Generalized Born Surface Area (MM/GBSA)

Protein-ligand complexes were also analyzed by MM/GBSA to estimate the free binding energies of studied ligands. The MM/GBSA calculations were applied to complex structures using Schrödinger's Prime module (20). The software extracted the frames of ligand-protein complexes from the MD trajectory of each complex at every 10 ps. For protein flexibility, we used the VSGB solvation model (32), which is a realistic parameterization of the solvation and OPLS-2005 force field.

Therapeutic Activity Prediction

Binary QSAR models in MetaCore/MetaDrug underwent therapeutic activity prediction, pharmacokinetic predictions, and ADME/toxicity profile for the compounds under investigation. It is an advanced data analysis program based on protein-protein interactions, protein-DNA interactions, disease, and toxicity. In Metacore/Metadrug, the software normalized the predicted therapeutic activity values to 0 and 1 (i.e., 0 meaning inactive, 1 meaning active molecules). Therapeutic activity results of selected 7 hit compounds, and the FDA-approved two drugs (Abemaciclib and Ribociclib) are in Table 1. Also, we carried out the therapeutic activity calculations for specific CDK2 inhibitors (A-674563 and MK-8776). Activity values of A-674563 and MK-8776 are 0.44 and 0.79, respectively.

RESULTS AND DISCUSSION

The current study aims to identify novel pyrimidine-based small compounds against CDK-2 from a small molecule library using combined text mining and molecular modeling approaches. For this aim, we used the Specs SC database, which includes 210,596 drug-like compounds. Figure 1 summarizes the flowchart of all applied procedures in the current study. All 210,596 compounds imported from the Specs-SC database in .sdf file format were converted to .name IUPAC text file format with MarvinSketch program (MarvinSketch, 2018) and identified 6,668 compounds that contain the phrase "pyrimidine." We then converted these compounds to .sdf file format and prepared 2D molecular structures of selected pyrimidine-derivatives to energetically optimized 3D structures using the LigPrep module of Maestro molecular modeling Suite. Afterward, we used these prepared compounds in molecular docking simulations. These 6,668 pyrimidine derivative compounds were screened at the active site of the CDK-2 (PDB ID: 2XMY) using a grid-based docking program Glide/SP (standard precision). Docking scores of these compounds were measured, and selected top-7 compounds for MD simulations to better understand their structural and dynamical profiles at the active site of the enzyme. Trajectory frames of each complex system were collected and analyzed throughout the MD simulations. Figure 2 shows MM/GBSA free energy plots of each selected hits, as well as FDA-approved drugs Abemaciclib and Ribociclib. Table 1 summarizes the top-docking scores and average MM/GBSA binding free energies of selected hit compounds as well as FDA-approved compounds. Docking scores of Glide/SP for the selected hit compounds are similar. However, a few selected compounds (i.e., compounds 2523, 2657, and 6447) have in a different range of MM/GBSA scores. There are a few reasons. (i) Some of the hit compounds (i.e., compound 2523) do not maintain their initial docking poses (Figure S13); (ii) Used docking scoring functions and MM/GBSA scoring functions are different; thus they may not necessarily produce the same ranking. For this reason, we also performed docking simulations using another docking program GOLD. As well as, instead of a prolonged (100-ns) MD simulations, we also repeated simulations with three times with short (10-ns) MD simulations for selected hit compounds and FDA-approved drugs (Abemaciclib, Ribociclib) and specific CDK-2 inhibitors (A-674563 and MK-8776), (Table S1). GOLD docking scores showed that selected hit compounds 6447 and 2657 also have high docking scores. Besides, Figures S1, S2, and S3 give the protein RMSD plot of compounds 2657, 5925, and 6447 during 10 ns simulations,

respectively. The results show that the average ΔG values obtained as a result of 10 ns simulations are very close to each other.

We provided, in Figure 3, the protein (Cα) RMSD plots of selected 7 hit compounds and FDA-approved drug Abemaciclib. RMSD plots of all studied systems show that all the systems reach the equilibrium after the 20-ns time scale. The deviations from the initial conformations are not significant (<4.0 Å). Within the studied systems, the Mol-236 lead compound had higher structural fluctuations compared to other systems.

Together with protein-RMSDs for the studied systems, we further measured the ligand-RMSDs. We plotted two different ligand RMSD graphs: LigFitLig and LigFitProt. While we considered the rotational motion in LigFitLig RMSDs, throughout the MD simulations, we adopted the translational motion of the ligand in LigFitProt RMSDs. Thus, in LigFitLig, heavy atoms of the ligand align themselves, and we recorded the deviations from the initial conformation throughout the simulation. However, in LigFitProt RMSDs, heavy atoms of ligand aligned to the backbone atoms of protein, and we measured the deviations from the initial positions as RMSDs. Ligand RMSD shows how stable the ligand is regarding the protein and its binding pocket. We showed the LigFitProt's RMSD of a ligand when we first matched the protein-ligand complex with the reference protein backbone, and then the RMSD of the ligand heavy atoms is determined. If the measured values are significantly larger than the protein's RMSD, then the ligand is likely to have diffused away from its original binding position. LigFitLig displays a ligand's RMSD that is aligned and measured just according to its reference conformation. This value for RMSD measures the internal fluctuations in the ligand atoms.

Figure 4 shows the LigFitProt RMSD plots of studied ligands. As it can be seen clearly from the Figure, most of the studied ligands do not have significant translational changes throughout the MD simulations. We observed the maximum deviations for FDA-approved drug Abemaciclib, Mol-2655 and Mol-2657, which have around 5.0 Å throughout the simulations. Results show that studied systems tightly bound to the binding pocket of the target structure; thus, they do not diffuse from their initial positions.

Figure 5 shows the LigFitLig RMSD plots of studied ligands. As we can see from the RMSD plots, all the ligands have less than 3.0 Å average RMSD values that represent the structural stability of studied compounds at the active site of the target protein. We also investigated the root-mean-square fluctuation (RMSF) values of studied compounds to check the effect of each ligand to the target protein structure. (Figure 6) Residue numbers around

40, 70, and 150 have higher fluctuations compared to other regions when the studied compounds bind.

We showed the 2D ligand interaction diagrams of hit compound Mol-2657 in Figure 7. The compound constructs stable non-bonded chemical interactions with mainly following residues throughout the MD simulations: Lys20, Lys129, Asn132, and Asp145.

Figure 8 displays a time-line representation of the interactions and contacts (H-bonds, Hydrophobic, Ionic, Water Bridges). The top panel shows the total number of specific contacts produced throughout the MD simulation period by the protein with the ligand. The bottom panel displays which residues interact with the ligand in each trajectory frame. Many residues allow more than one direct interaction with the ligand, defined by darker orange color, depending on the scale to the right of the map. Figure 8 displays that Mol-2657 constructs critical chemical interactions with residues Lys129, Asn132, and Asp 145, which are stable throughout the simulation time.

We screened the therapeutic activity values of selected 7 hit compounds against cancer using binary quantitative structure-activity relationship (QSAR) models in Metacore/Metadrug. Therapeutic activity values of all selected compounds greater than 0.5 (Table 1).
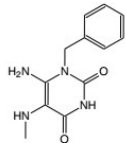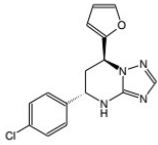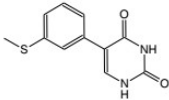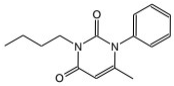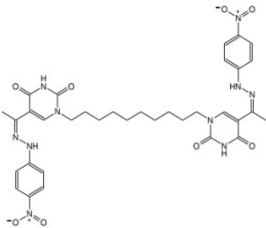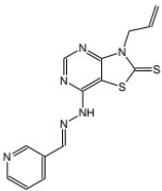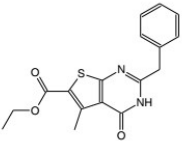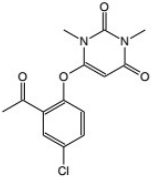
CONCLUSION

In this study, we carried out advanced integrated text mining, virtual screening, and hybrid molecular modeling strategies to identify the pyrimidine-based inhibitors targeted to the binding site of CDK-2. Thus, we have selected 6,668 pyrimidine derivatives from the small molecule database of Specs-SC for this aim. We screened these identified compounds against the CDK-2 target using molecular docking and used the top-docking poses of selected 7 hits in long MD simulations (100 ns). MM/GBSA binding free energy calculations have been conducted for all ligands. We also performed the same protocol (100 ns molecular dynamics simulation) for the FDA-approved drug Abemaciclib. The therapeutic activities of the selected 8 compounds were screened in Metacore/Metadrug using binary QSAR models. We evaluated the docking scores and MM/GBSA scores in Table 1. We observed that the compound 2523 has a high docking score, however its corresponding MM/GBSA score was low compared to MM/GBSA scores of other compounds. When we visualize the trajectory
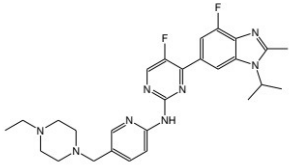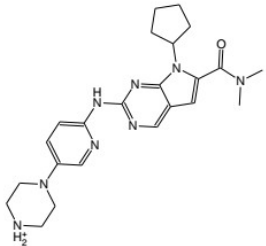
frames throughout the simulations, we observed that this compound does not stay at the binding pocket throughout the simulations and loses its initial docking pose. (see Figure S11 Thus, we have not evaluated the binding energy of compound 2523. Here, we showed that integrated text-mining and molecular modeling approaches leading to 7 novel pyrimidine-based CDK-2 inhibitors. Thus, such new compounds may open new paths for developing small inhibitors against CDK-2.

Table 1. 2D structures, top-docking scores in Glide/SP and Gold, and average MM/GBSA scores of selected 7 hits and FDA-approved 2 CDK inhibitors.

Compounds	2D Structure	Glide Docking Score (kcal/mol)	Gold Docking Score (kcal/mol)	MM/GBSA (kcal/mol)	Predicted Therapeutic activity
Mol-30		-9.359	-4.956	-48.91±3.29	0.53
Mol-236		-9.385	-6.453	-46.52±3.95	0.75
Mol-2623		-9.382	-6.882	-52.85±4.64	0.52
Mol-2655		-9.298	-6.963	-47.32±3.58	0.53

Compounds	2D Structure	Glide Docking Score (kcal/mol)	Gold Docking Score (kcal/mol)	MM/GBSA (kcal/mol)	Predicted Therapeutic activity
Mol-2657		-9.573	-7.533	-74.38±7.91	0.69
Mol-5925		-9.746	-7.00	-51.90±3.02	0.74
Mol-6447		-9.317	-7.675	-62.14±3.20	0.56
Mol-2523		-9.373	-5.969	NA	0.30

Compounds	2D Structure	Glide Docking Score (kcal/mol)	Gold Docking Score (kcal/mol)	MM/GBSA (kcal/mol)	Predicted Therapeutic activity
Abemaciclib		-6.818	-9.170	-45.49±5.90	0.78
Ribociclib		-5.187	-7.568	-47.40±3.16	0.68
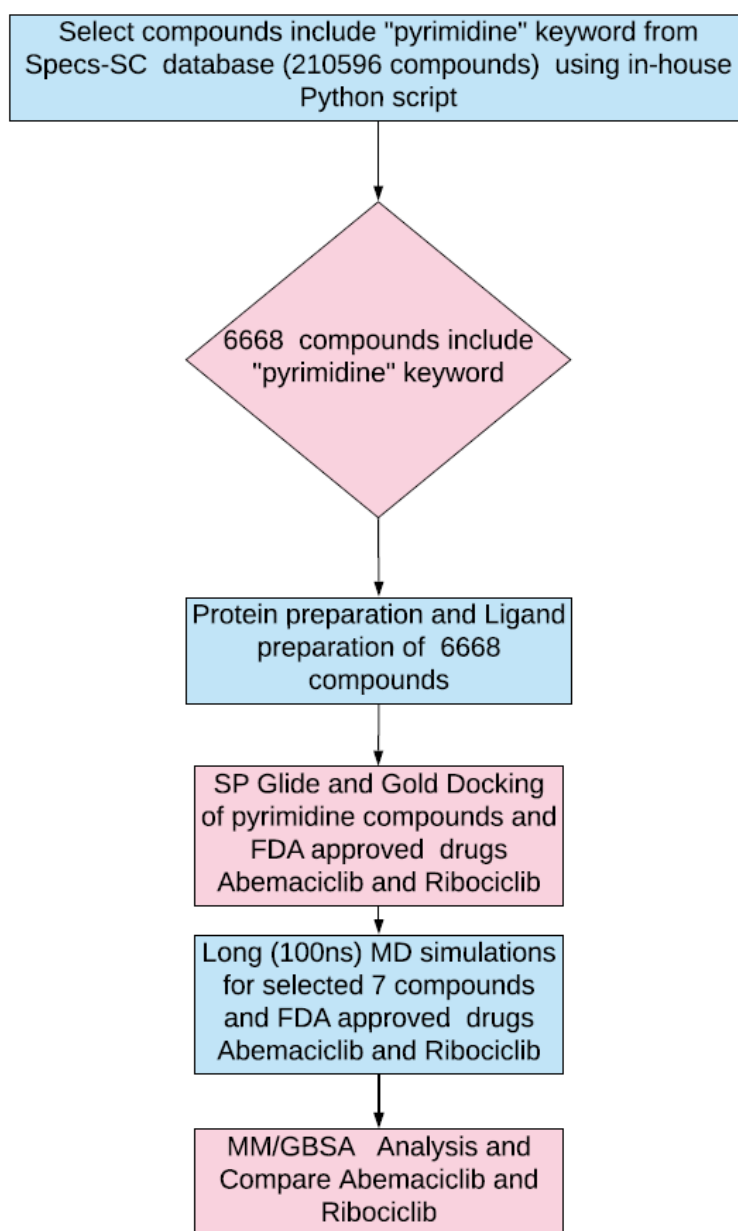
Figure 1: Applied virtual screening workflow at the current study.
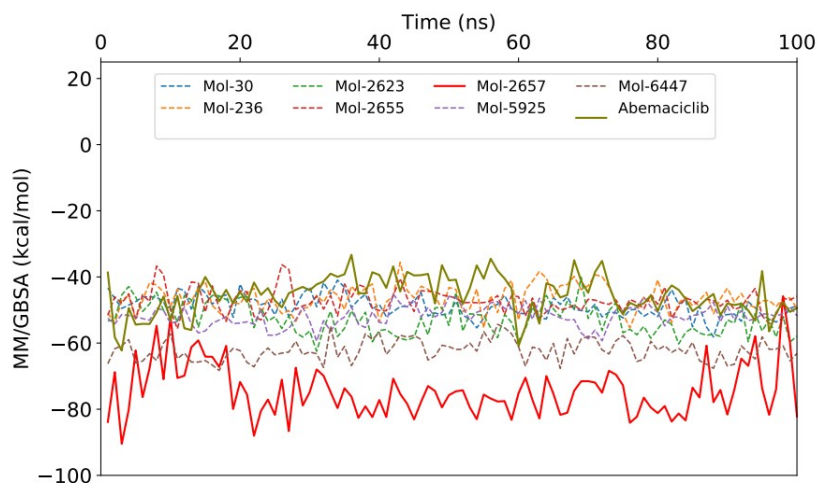
Figure 2: MM/GBSA plots of selected 7 compounds and FDA-approved drug Abemaciclib during the MD simulations.
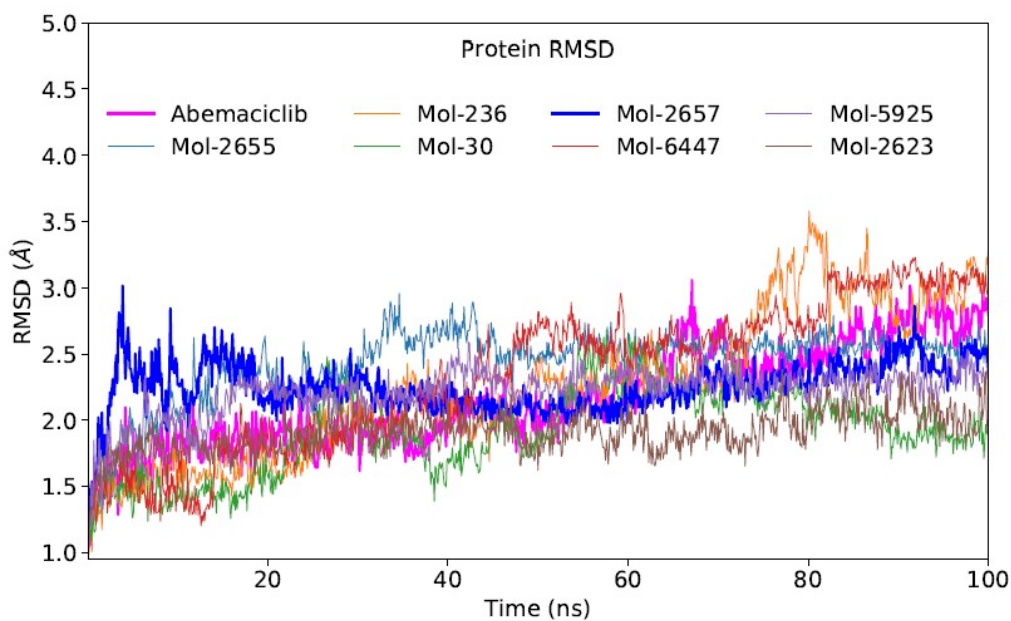
Figure 3: Protein-RMSD plots of selected 7 compounds and FDA-approved drug Abemaciclib throughout the MD simulations.
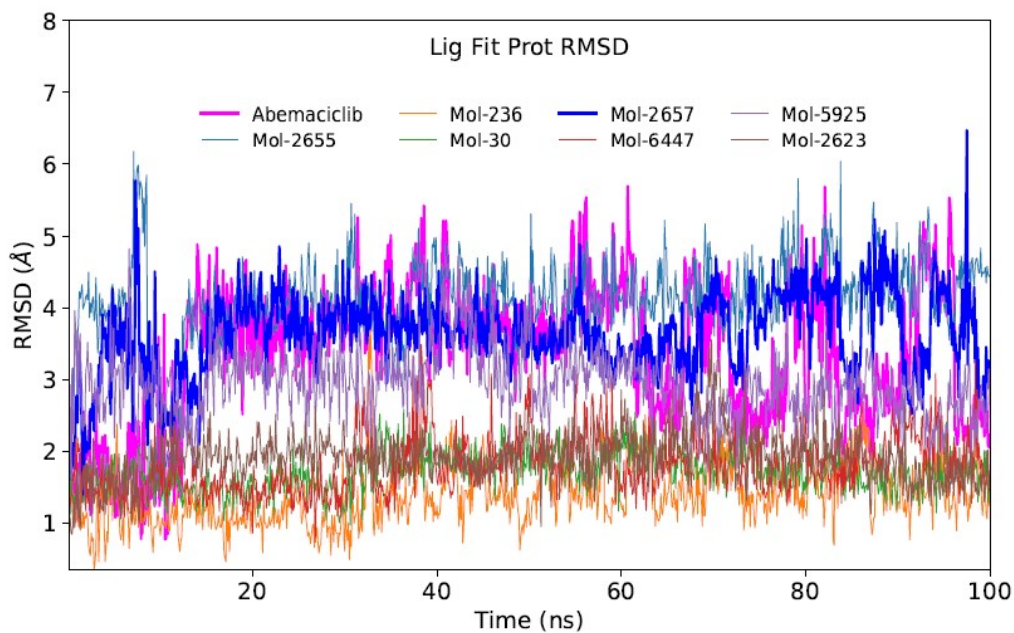
Figure 4: LigFitProt RMSD graph for selected 7 compounds and FDA-approved drug Abemaciclib.
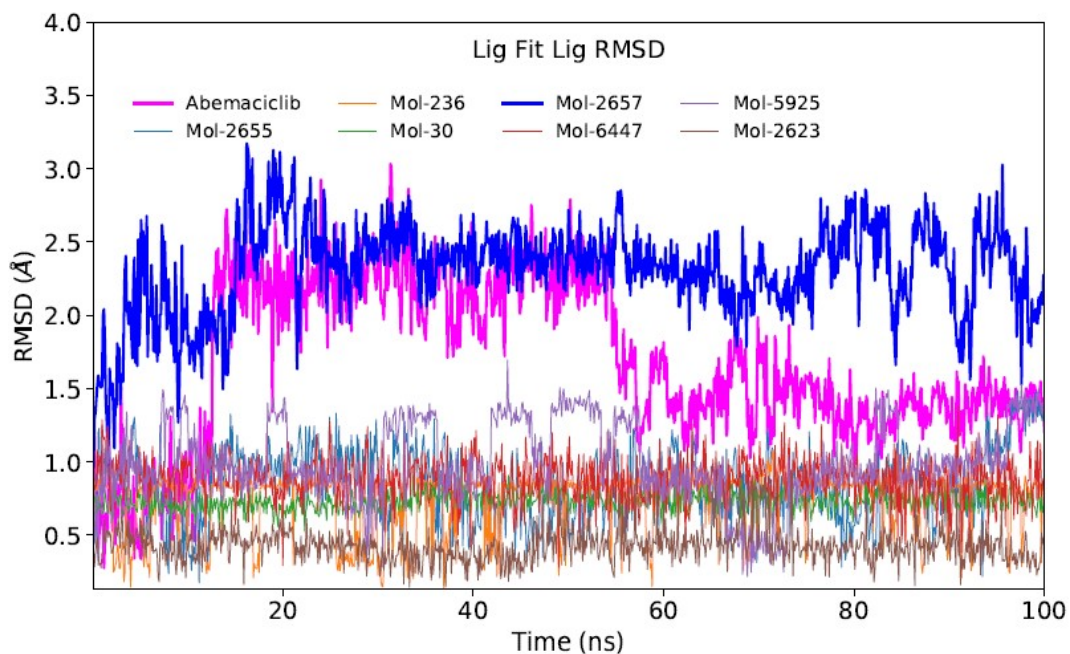

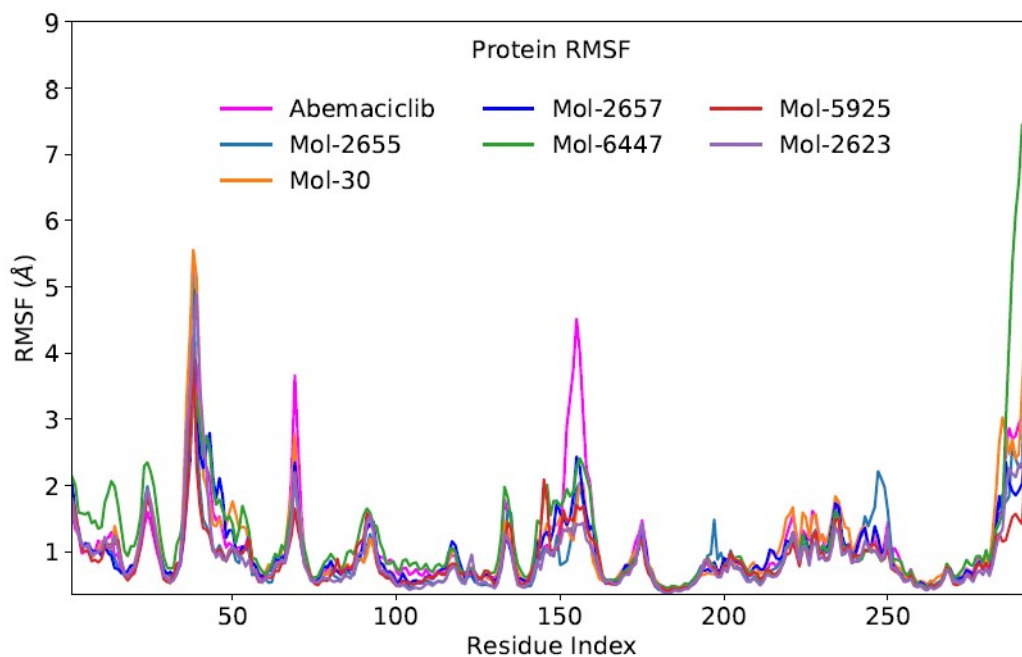**Figure 5:** LigFitLig RMSD graph for selected 7 compounds and FDA-approved drug Abemaciclib.

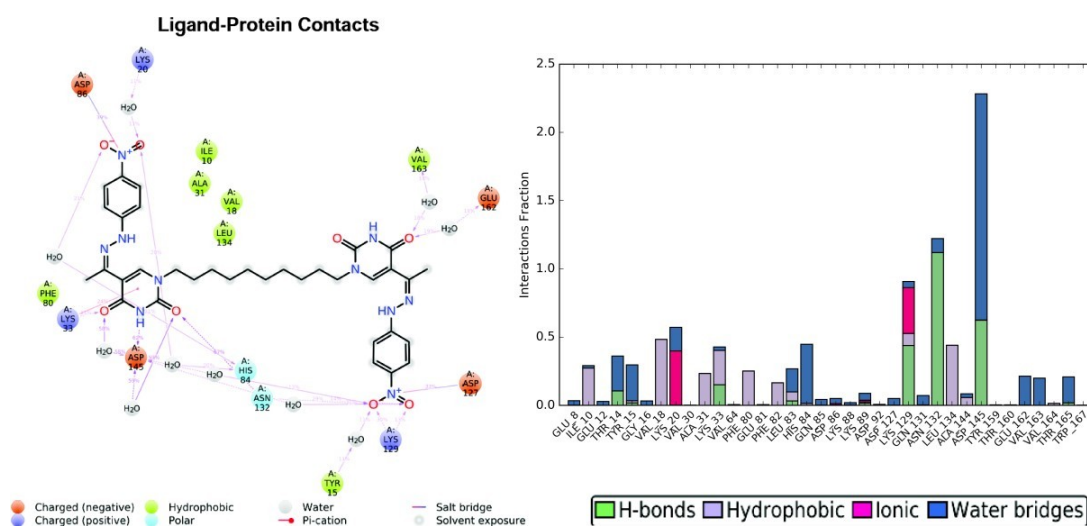**Figure 6:** RMSF plots for selected 7 hit compounds and FDA-approved drug Abemaciclib.

Figure 7: Mol-2657's 2D ligand interaction graph. The figure also shows interaction fractions throughout the simulations.
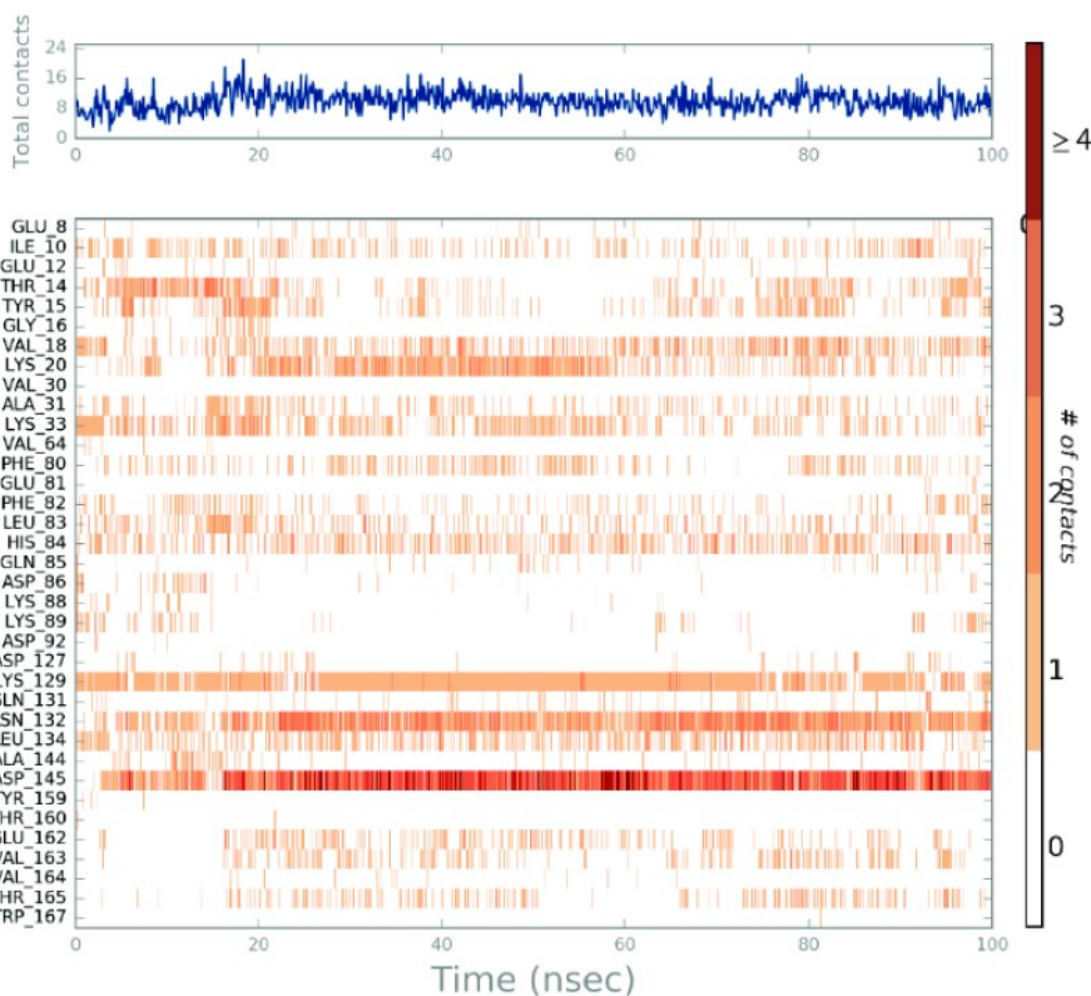
Figure 8: Protein-ligand interactions throughout the simulations. While top-panel shows total contacts with the ligand, the bottom-panel represents the formed and broken interactions between the protein and ligand during the simulation.

REFERENCES

1. Hirose Y, Ohkuma Y. Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. Journal of Biochemistry. 2007;141(5):601–8.

2. Shapiro GI, Harper JW. Anticancer drug targets: cell cycle and checkpoint control. Journal of Clinical Investigation. 1999 ;104(12):1645–53.

3. Barriere C, Santamaria D, Cerqueira A, Galan J, Martin A et al. Mice thrive without Cdk4 and Cdk2. Molecular Oncology. 2007;1(1):72–83.

4. Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. Nature Reviews Cancer. 2009;9(3):153–66.

5. Clark A. Natural products as a resource for new drugs. Pharmaceutical Research. 1996;13:1133–44.

6. Melik-Ogandzhanyan RG, Khachatryan VE, Gapoyan AS. Furo-, Thieno-, and Pyrrolo-[2,3-d]pyrimidines. Russian Chemical Reviews. 1985;54(3):262–76.

7. Sanghvi YS, Larson SB, Matsumoto SS, Nord LD, Smee DF et al. Antitumor and antiviral activity of synthetic alpha and beta-ribonucleosides of certain substituted pyrimido[5,4-d]pyrimidines: a new synthetic strategy for exocyclic aminonucleosides. Journal of Medicinal Chemistry. 1989;32(3):629–37.

8. Tenser RB, Gaydos A, Hay KA. Inhibition of herpes simplex virus reactivation by dipyridamole. Antimicrobial Agents and Chemotherapy. 2001;45 (12):3657–9.

9. de-la-Cruz JP, Carrasco T, Ortega G, Sanchez de-la-Cuesta F. Inhibition of ferrous induced lipid peroxidation by pyrimido-pyrimidine derivatives in human liver membranes. Lipids. 1992; 27(3):192–4.

10. Sharma P, Rane N, Gurram VK. Synthesis and QSAR studies of pyrimido[4,5-d]pyrimidine-2,5-dione derivatives as potential antimicrobial agents. Bioorganic & Medicinal Chemistry Letter. 2004;14(16):4185–90.

11. Ram VJ, Goel A, Sarkhel S, Maulik PR. A convenient synthesis and hepato protectiveactivity of imidazo[1,2-]pyrimido[5,4-]pyrimidine, tetraazaacenaphthene and tetraazaphenalene from cyclic ketene aminals through tandem addition-cyclization reactions. Bioorganic & Medicinal Chemistry. 2002;10(5):1275–80.

12. Huang WS, Liu S, Zou D, Thomas M, Wang Y. et al. Discovery of Brigatinib (AP26113), a Phosphine Oxide-Containing, Potent, Orally Active Inhibitor of Anaplastic Lymphoma Kinase. Journal of Medicinal Chemistry. 2016;59(10):4948–64.

13. Corona SP, Generali D. Abemaciclib: a CDK4/6 inhibitor for the treatment of HR+/HER2- advanced breast cancer. Drug Design, Development and Therapy. 2018;12:321–30.

14. Krallinger M, Valencia A. Text-mining and information-retrieval services for molecular biology. Genome Biology. 2005;6(7):224.

15. MarvinSketch 18.30.0, Chemaxon, 2018.

16. Banks JL, Beard HS, Cao Y, Cho AE, Damm W et al. Integrated Modeling Program, Applied Chemical Theory (IMPACT). Journal of Computational Chemistry. 2005;26(16):1752–80.

17. Kakarala KK, Jamil K , Devaraji V. Structure and putative signaling mechanism of Protease activated receptor 2 (PAR2) - a promising target for breast cancer. Journal of Molecular Graphics and Modelling. 2014;53:179–99.

18. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR et al. Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules. Journal of Computer-Aided Molecular Design. 2007;21(12):681–91.

19. Wang S, Griffiths G, Midgley CA, Barnett AL, Cooper M et al. Discovery and characterization of 2-anilino-4- (thiazol-5-yl)pyrimidine transcriptional CDK inhibitors as anticancer agents. Cell Chemical Biology. 2010;17(10):1111–21.

20. Jacobson MP, Pincus DL, Rapp CS, Day TJ, Honig B et al. A hierarchical approach to all-atom protein loop prediction. Proteins. 2004;55(2):351–67.

21. Protein preparation, version 2.5, Schrödinger, LLC, New York, 2011.

22. Sledz P, Caflisch A. Protein structure-based drug design: from docking to molecular dynamics. Current Opinion in Structural Biology. 2018;48:93–102.

23. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of Medicinal Chemistry. 2004;47(7):1739–49.

24. Raha K, Merz KM. A quantum mechanics-based scoring function: Study of zinc ion-mediated ligand binding. Journal of the American Chemical Society. 2004;126(4):1020–1.

25. Subhani S, Jayaraman A, Jamil K. Homology modelling and molecular docking of MDR1 with chemotherapeutic agents in non-small cell lung cancer. Biomedicine & Pharmacotherapy. 2015;71:37–45.

26. Roy S, Kumar A, Baig MH, Masajik M, Provaznik I. Virtual screening, ADMET profiling, molecular docking and dynamics approaches to search for potent selective natural molecules based inhibitors against metallothionein-III to study Alzheimer's disease. Methods. 2015; 83:105–10.

27. Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. Journal of Chemical Information and Modeling. 2011;51(1):69-82.

28. Desmond, version 4.9, Schrödinger, LLC, New York, 2011.

29. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction Models for Water in Relation to Protein Hydration pages 331–342. Springer Netherlands, Dordrecht. 1981.

30. Hoover WG. Canonical dynamics: Equilibrium phase-space distributions. Physical Review A. 1985;31(3):1695–7.

31. Ma Z, Tuckerman M. Constant pressure ab initio molecular dynamics with discrete variable representation basis sets. The Journal of Chemical Physics. 2010;133(18):184110.

32. Li J, Abel R, Zhu K , Cao Y, Zhao S, Friesner RA. The VSGB 2.0 model: a next generation energy model for high resolution protein structure modeling. Proteins. 2011;79 (10):2794–812.

SUPPLEMENTARY INFORMATION

Table S1. Molecular dynamics simulations results of selected 7 hit compound, FDA-approved two drugs (Abemaciclib and Ribociclib) and two selective CDK2 inhibitors (A-674563 and MK-8776)

Compounds	Gold Docking Scores (kcal/mol)	MD_1 (10 ns) (ΔG/kcal/mol)	MD_2 (10 ns) (ΔG/kcal/mol)	MD_3 (10 ns) (ΔG/kcal/mol)	Average (ΔG/kcal/mol)
30	-4.956	-46.13	-46.15	-47.59	-46.62
236	-6.453	-47.37	-46.99	-42.95	-45.77
2657	-7.533	-75.19	-82.62	-75.62	-77.81
5925	-7.000	-51.46	-53.85	-55.04	-53.45
6447	-7.675	-60.63	-56.98	-60.32	-59.31
2523	-5.969	-43.47	-40.95	-39.53	-41.31
2655	-6.963	-45.13	-45.08	-45.99	-45.40
Abemaciclib	-9.170	-41.75	-39.04	-40.05	-40.28
Ribociclib	-7.568	-49.40	-52.64	-50.99	-51.01
A-674563	-9.685	-46.98	-47.75	-47.24	-47.32
MK-8776	-7.529	-42.12	-44.35	-44.90	-43.79
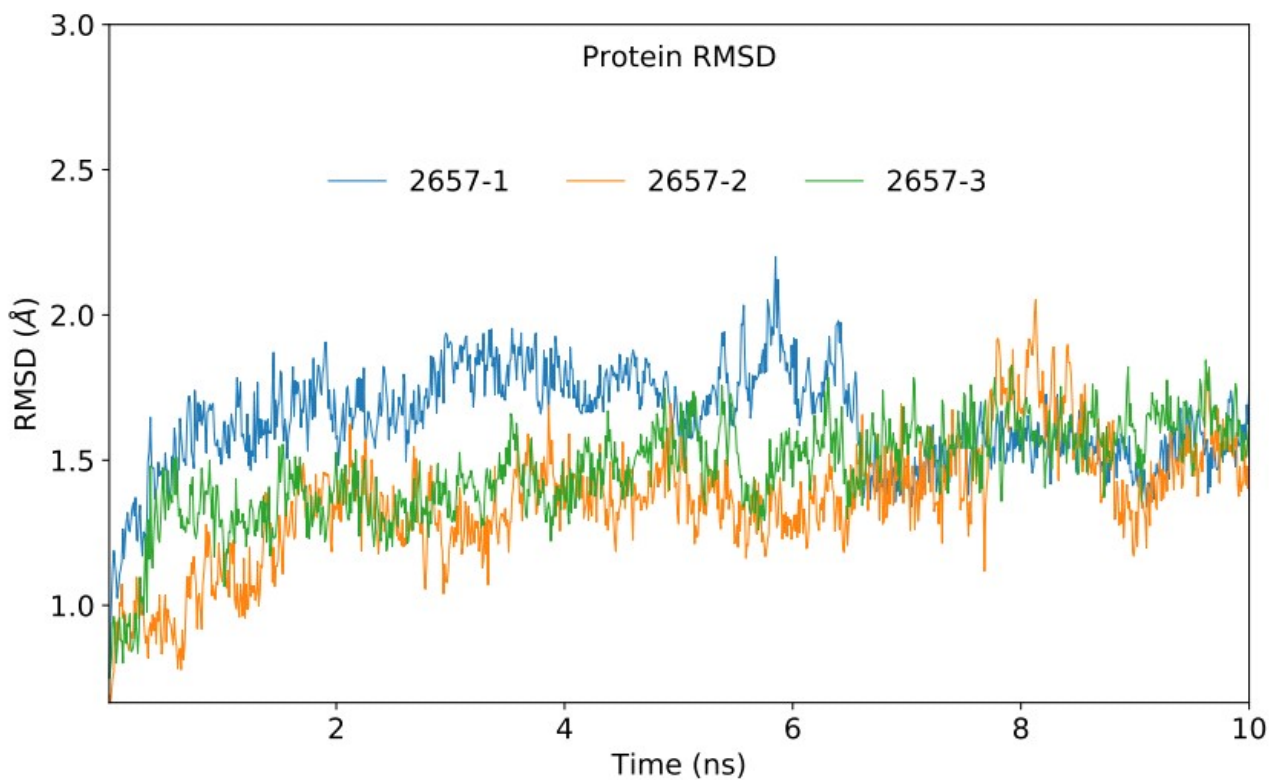
Figure S1: Protein RMSD plot of compound 2657 as a result of 3 separate 10 ns simulations.
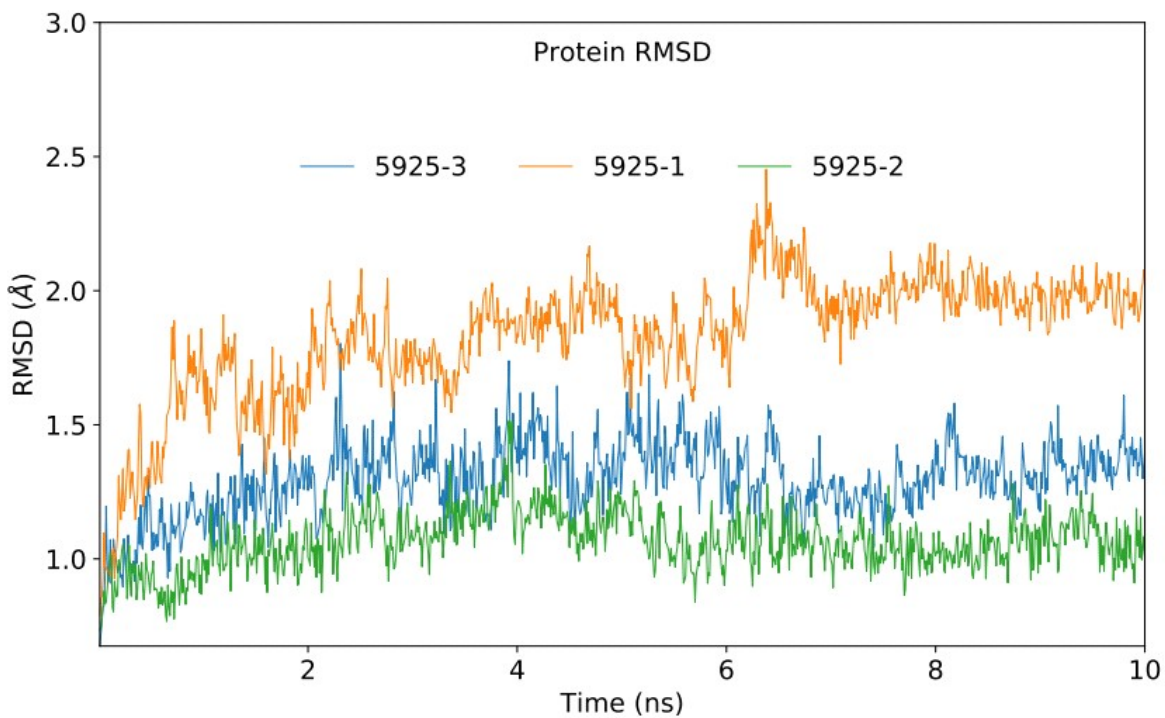

Figure S2: Protein RMSD plot of compound 5925 as a result of 3 separate 10 ns simulations.
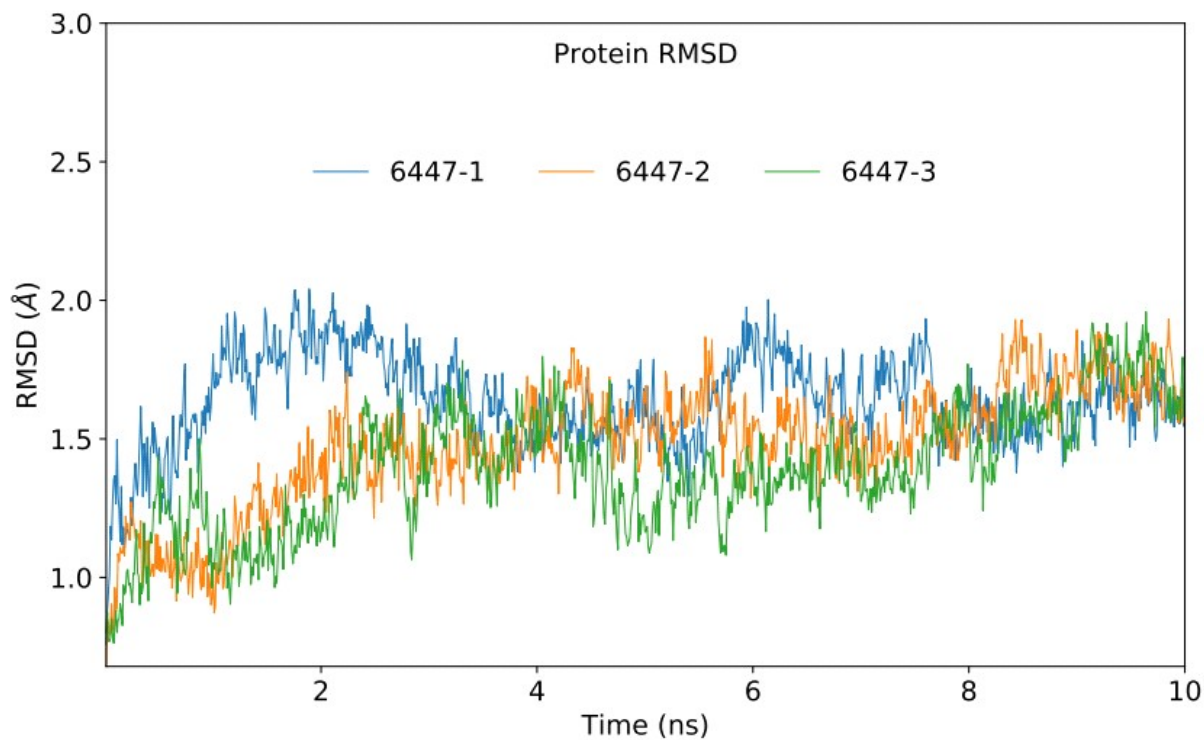
Figure S3: Protein RMSD plot of compound 6447 as a result of 3 separate 10 ns simulations.
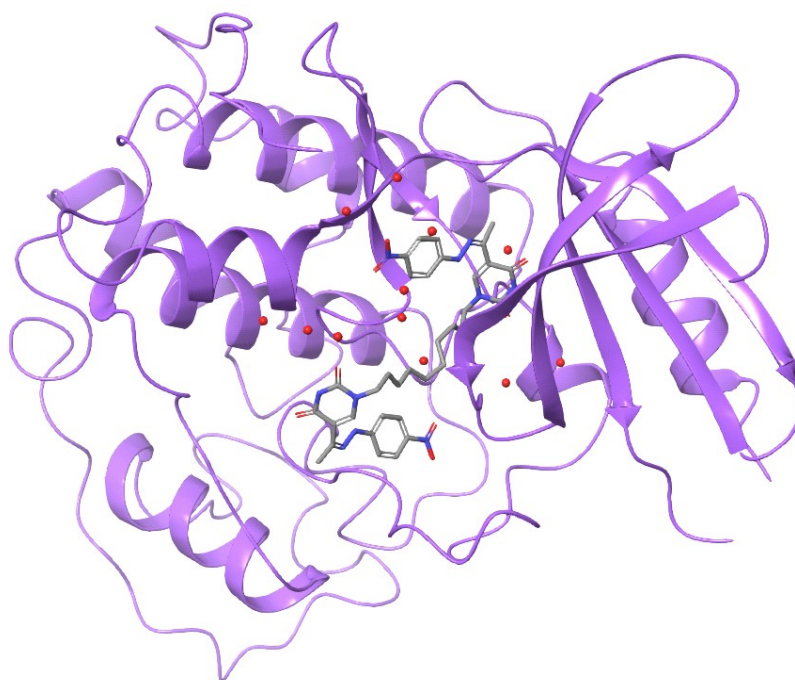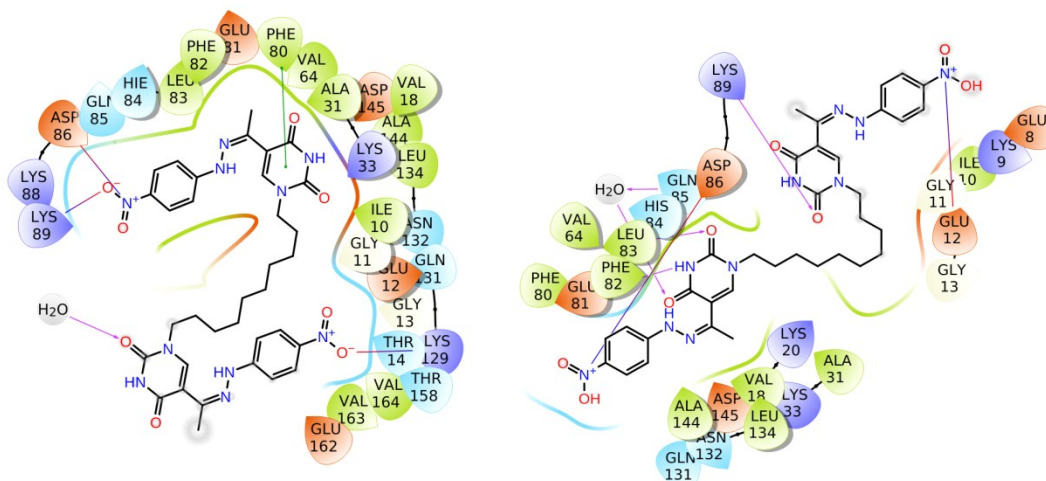

Figure S4: 2XMY-Compound 2657 complex structure.

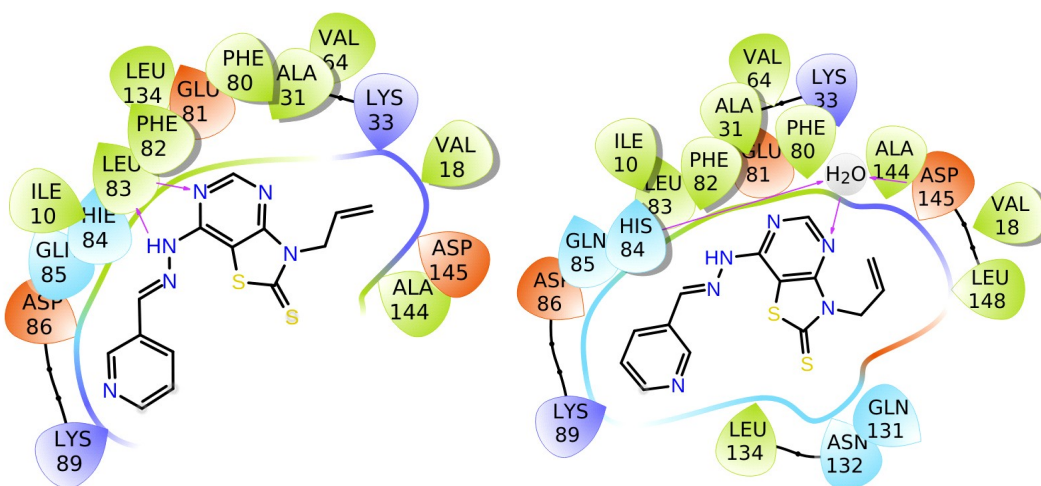**Figure S5:** Ligand interactions diagram of compound 2657 (left Glide docking and right Gold docking)

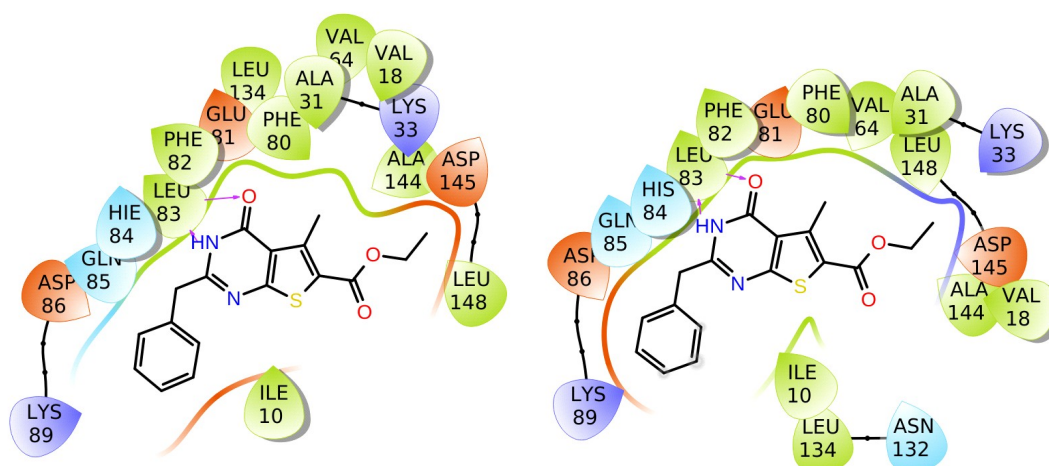**Figure S6:** Ligand interactions diagram of compound 5925 (left Glide docking and right Gold docking)

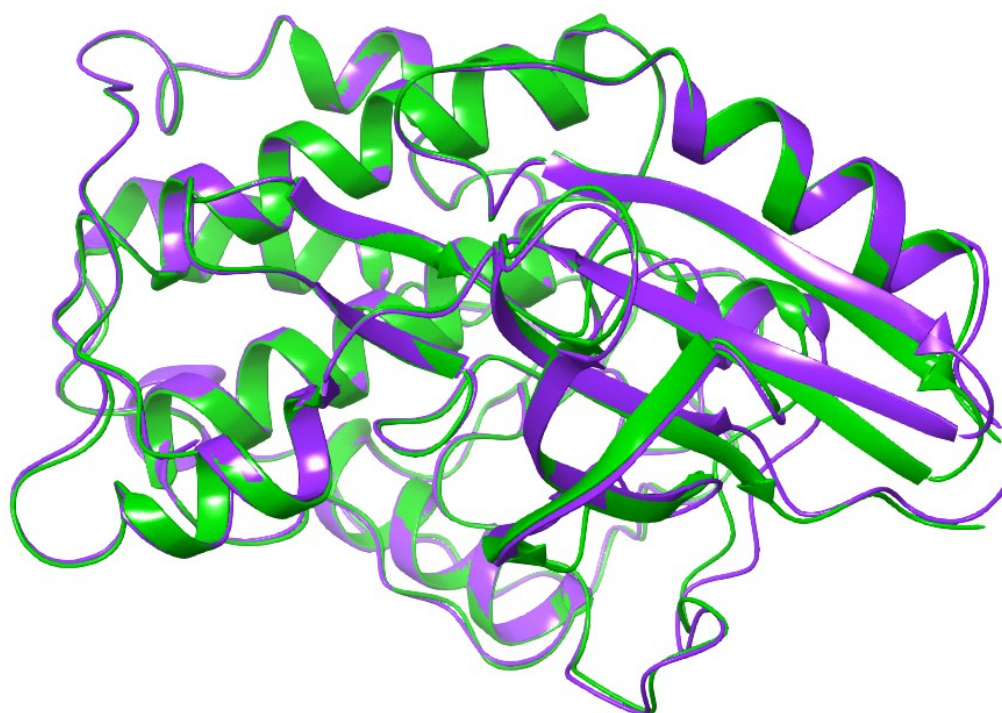**Figure S7:** Ligand interactions diagram of compound 6447 (left Glide docking and right Gold docking)

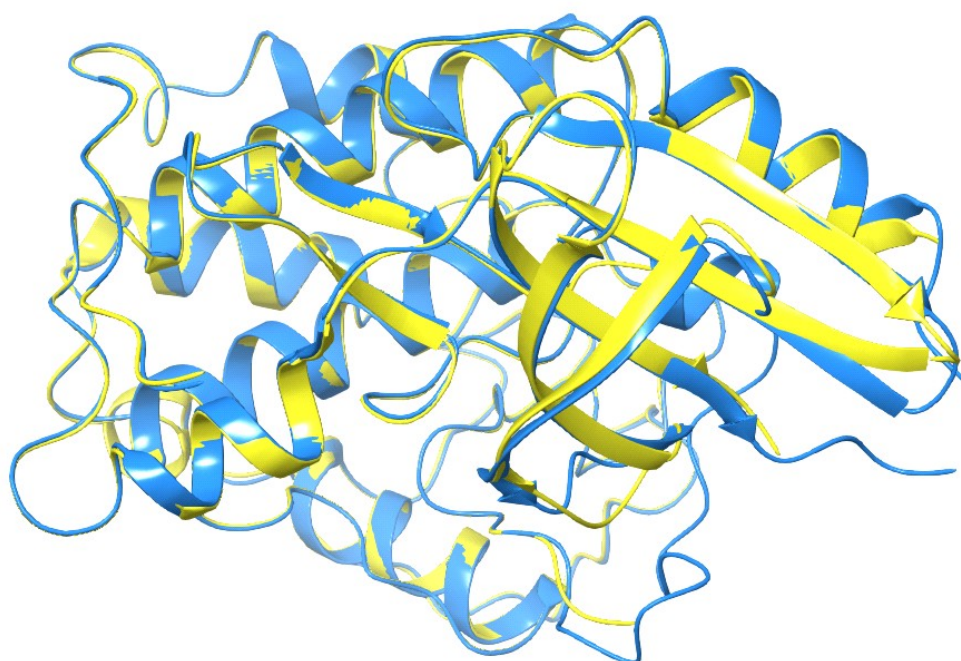**Figure S8:** Alignment of 2XMY and 2XNB (Alignment Score: 0.014 ; RMSD: 0.583 A)

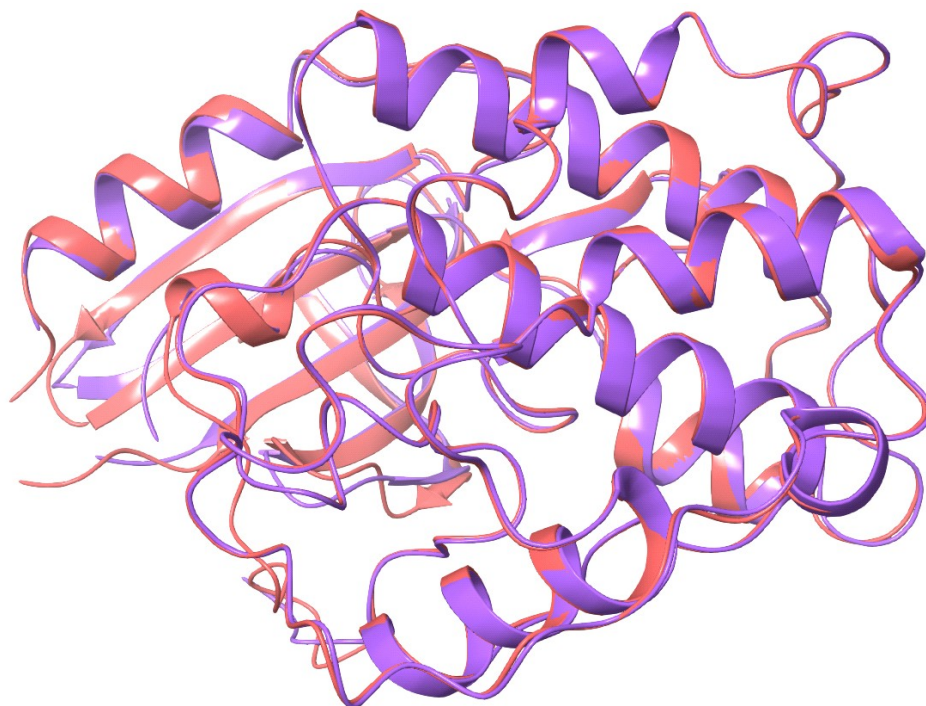**Figure S9:** Alignment of 2XMY and 2W05 (Alignment Score: 0.016; RMSD: 0.629 A)

Figure S10: Alignment of 2XMY and 1Y91 (Alignment Score: 0.018; RMSD: 0.670 A)

Table S2: A method development-validation1.

Title	Stars	t Q	Job Name	docking score
▼ glide-dock_SP_validation_se…				
2XMY - minimized	☆☆☆		glide-dock_SP_validat…	
2m	☆☆☆	0	glide-dock_SP_validat…	-8.629
2c	☆☆☆	0	glide-dock_SP_validat…	-7.510
2c	☆☆☆	0	glide-dock_SP_validat…	-7.510
2d	☆☆☆	0	glide-dock_SP_validat…	-7.428
2n	☆☆☆	0	glide-dock_SP_validat…	-7.077
2h	☆☆☆	0	glide-dock_SP_validat…	-6.729
▼ glide-dock_SP_validation_se…				
2XMY - minimized	☆☆☆		glide-dock_SP_validat…	
2i	☆☆☆	0	glide-dock_SP_validat…	-7.499
2l	☆☆☆	-1	glide-dock_SP_validat…	-7.490
2j	☆☆☆	0	glide-dock_SP_validat…	-6.880
2k	☆☆☆	-1	glide-dock_SP_validat…	-4.472

Table S3. A method development-validation2.

Title	...omer probability	docking score
▼ glide-dock_SP_validation_set3_inactive_…		
2XMY - minimized		
12	1.000	-8.839
17	1.000	-8.174
16	1.000	-6.361
▼ glide-dock_SP_validation_set3_active_pv…		
2XMY - minimized		
11e	1.000	-9.554
11a	1.000	-9.273
11h	1.000	-7.286
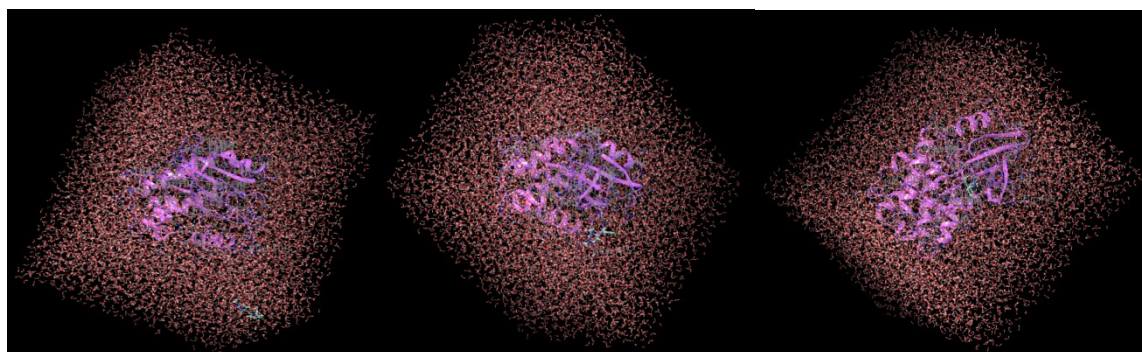
FigureS11: Initial (right), after 60 ns (center) and last (left) poses of compound 2523 throughout 100 ns molecular dynamics simülations.