



OPEN ACCESS

EDITED BY

Urban Bren,
University of Maribor, Slovenia

REVIEWED BY

Samo Lešnik,
University of Maribor, Slovenia
Janez Konc,
National Institute of Chemistry, Slovenia
Chris Oostenbrink,
University of Natural Resources and Life
Sciences Vienna, Austria

*CORRESPONDENCE

Olaitan I. Awe,
✉ laitanawe@gmail.com

RECEIVED 21 July 2023

ACCEPTED 13 November 2023

PUBLISHED 30 November 2023

CITATION

Ogbodo UC, Enejoh OA, Okonkwo CH,
Gnanasekar P, Gachanja PW, Osata S,
Atanda HC, Iwuchukwu EA, Achilonu I
and Awe OI (2023), Computational
identification of potential inhibitors
targeting *cdk1* in colorectal cancer.
Front. Chem. 11:1264808.
doi: 10.3389/fchem.2023.1264808

COPYRIGHT

© 2023 Ogbodo, Enejoh, Okonkwo,
Gnanasekar, Gachanja, Osata, Atanda,
Iwuchukwu, Achilonu and Awe. 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.

Computational identification of potential inhibitors targeting *cdk1* in colorectal cancer

Uchechukwu C. Ogbodo ¹, Ojochenemi A. Enejoh ²,
Chinelo H. Okonkwo³, Pranavathiyani Gnanasekar⁴,
Pauline W. Gachanja⁵, Shamim Osata ⁶, Halimat C. Atanda ⁷,
Emmanuel A. Iwuchukwu ⁸, Ikechukwu Achilonu⁸ and
Olaitan I. Awe ^{9,10*}

¹Department of Applied Biochemistry, Nnamdi Azikiwe University, Awka, Nigeria, ²Genomics and Bioinformatics Department, National Biotechnology Development Agency, Abuja, Nigeria, ³Department of Pharmacology and Toxicology, University of Nigeria, Nsukka, Nigeria, ⁴Department of Bioinformatics, Pondicherry University, Puducherry, India, ⁵Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya, ⁶Department of Biochemistry, University of Nairobi, Nairobi, Kenya, ⁷Biotechnology Department, Federal University of Technology, Akure, Nigeria, ⁸Protein Structure-Function Research Unit, School of Molecular and Cell Biology, Faculty of Sciences, University of Witwatersrand, Johannesburg, South Africa, ⁹Department of Computer Science, University of Ibadan, Ibadan, Nigeria, ¹⁰African Society for Bioinformatics and Computational Biology, Cape Town, South Africa

Introduction: Despite improved treatment options, colorectal cancer (CRC) remains a huge public health concern with a significant impact on affected individuals. Cell cycle dysregulation and overexpression of certain regulators and checkpoint activators are important recurring events in the progression of cancer. Cyclin-dependent kinase 1 (CDK1), a key regulator of the cell cycle component central to the uncontrolled proliferation of malignant cells, has been reportedly implicated in CRC. This study aimed to identify CDK1 inhibitors with potential for clinical drug research in CRC.

Methods: Ten thousand (10,000) naturally occurring compounds were evaluated for their inhibitory efficacies against CDK1 through molecular docking studies. The stability of the lead compounds in complex with CDK1 was evaluated using molecular dynamics simulation for one thousand (1,000) nanoseconds. The top-scoring candidates' ADME characteristics and drug-likeness were profiled using SwissADME.

Results: Four hit compounds, namely, spiraeoside, robinetin, 6-hydroxyluteolin, and quercetagenin were identified from molecular docking analysis to possess the least binding scores. Molecular dynamics simulation revealed that robinetin and 6-hydroxyluteolin complexes were stable within the binding pocket of the CDK1 protein.

Discussion: The findings from this study provide insight into novel candidates with specific inhibitory CDK1 activities that can be further investigated through animal testing, clinical trials, and drug development research for CRC treatment.

KEYWORDS

colorectal cancer, cyclin-dependent kinase 1, inhibitors, natural compounds, molecular dynamics simulation

1 Introduction

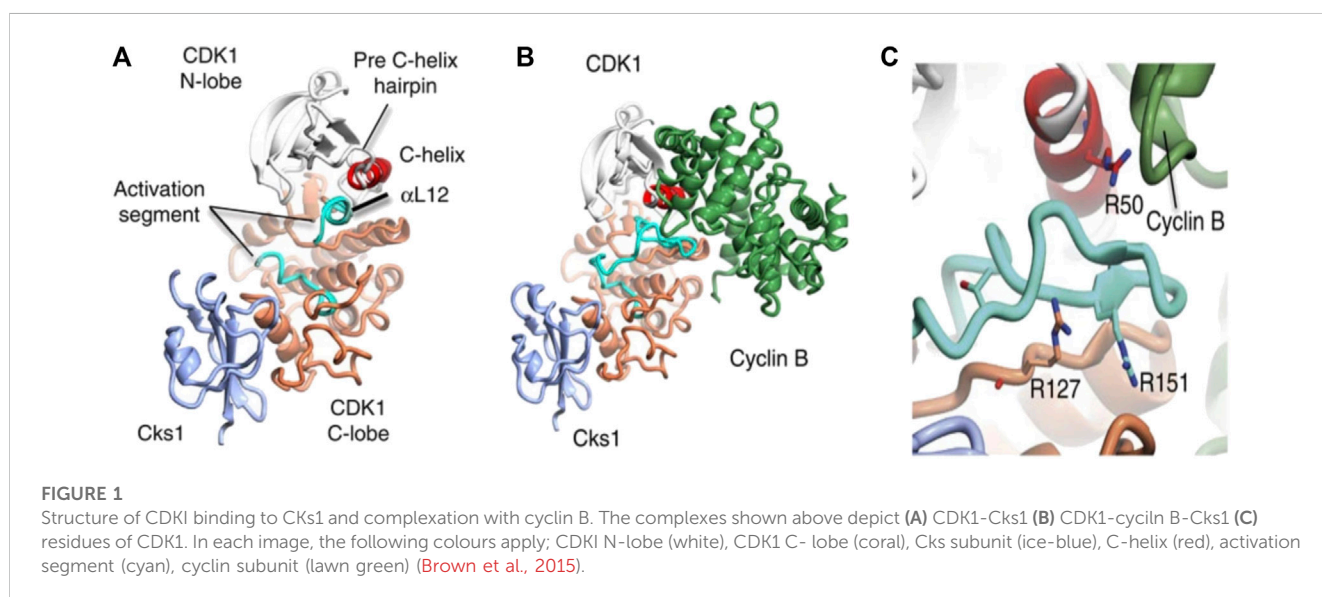
Colorectal cancer (CRC) is the third most common cancer next to lung and breast cancers with an estimated diagnosis of 2 million individuals and a mortality of approximately 900,000 annually (Bray et al., 2020; Holowatyj et al., 2020). It is thought to be a disease in high-income countries; however, studies have shown the incidence of CRC to increase at a projected rate of 70% by 2030 for low and middle-income countries (Ferlay et al., 2015) with a 5% lifetime average-risk incidence globally (Weitz et al., 2005). These gory statistics have significant implications for the public health systems both in the country and globally hence the need for a more dynamic scientific approach to the CRC burden.

The onset of CRC follows a pattern of abnormal growth of the epithelial cells of the colon and/or rectum leading to a tumor mass that fills the intestinal lumen, metastasizing into nearby organs of the abdominal area, lymph nodes, and other parts of the body (Adeoti et al., 2016). It usually begins as a small, adenomatous mass of cells known as polyps that form on the inside of the colon, which may become malignant or remain benign (Ikwu et al., 2020). As with many cancers, the cause of CRC is not known (Simon et al., 2017), however, risk factors such as age, lifestyle habits (alcohol and tobacco use), high-fat diet, predisposing medical conditions such as type 2 diabetes, and a family history of CRC have been associated with the disease course. Although traditional chemotherapy options have yielded remarkable results in CRC treatment with possibilities of relapses, limited molecular understanding of the oncogenic initiation and interplay of biochemical mechanisms/factors have impaired the development of novel pharmacologically therapeutic anti-cancer drug agents.

Cyclin-dependent kinase 1 (CDK1) enzyme has been implicated in the oncogenesis of CRC (Sofi et al., 2022a). It is a serine/threonine kinase known to be the catalytic subunit of the M-phase promoting factor (MPF) complex (Kalous et al., 2020) and belongs to a family of cyclic cell regulatory proteins involved in maintaining cell cycle efficiency (Sofi et al., 2022a; Sofi et al., 2022b). Though involved in

many cellular events that ensure cell survival, CDK1 specifically acts by regulating the G₂/M and G₁/S phases oscillating in activity between the M transition stages to bring about replication, activating checkpoints, and more recently has been essential to replication fork maintenance (Adhikari et al., 2012; Liao et al., 2017; Matthews et al., 2022). Prior to initiating cell replication, CDK1 is activated by binding to cyclin B₁ to form CDK1-cyclin B complex (Figure 1) (Brown et al., 2015). However, the over expression of CDK1 has been linked to unrestricted cell proliferation. Its over expression has been linked to unrestricted cell proliferation, tumor development, and dysregulation associated with the initiation of many cancers including colorectal cancer (Li et al., 2020; Sofi et al., 2022a; Sofi et al., 2022b). Gene expression has been associated with biomarkers in complex diseases like prostate cancer (Chikwambi et al., 2023), epilepsy (El Abed et al., 2023), malaria/COVID-19 (Nzungize et al., 2022) and SARS-CoV-2 infection (Nyamari et al., 2023). Several studies suggest that overexpression of CDK1, which is the hallmark event in many malignancies is associated with 5-fluorouracil resistance through altered activation of Wnt/ β -catenin signaling, mTOR, and p53 pathways leading to poor prognosis in CRC (Zhu et al., 2019; Zhu et al., 2020a; Zhu et al., 2020b; Sofi et al., 2022a).

Natural products have drawn a lot of attention recently as potential sources of treatments for a variety of illnesses. In recent years, bioactive compounds from plants have been proven to have opportunities in the pharmaceutical industry for therapeutic purposes (Sasidharan et al., 2011). Resveratrol, a polyphenol has been documented for its ability to reduce or inhibit cell proliferation, reduce apoptosis and suppress angiogenesis and metastasis in cancer cells (Greenlee et al., 2022; Dar et al., 2023). Also, flavopiridol, a flavonoid based substance has undergone clinical trials for its anti-cancer benefits and as a potent Cyclin-Dependent Kinase (CDK) inhibitor (Deep et al., 2018). By specifically targeting CDK1, cell cycle arrest is initiated, slowing down the growth of tumors. Compared to synthetic alternatives, natural products have fewer adverse effects.



Despite recent advances in diagnostics and treatments for CRC as well as the failure of conventional cancer treatments including chemotherapy and radiotherapy to sustain lower relapse rates, there remains a gap in treatment and increasing incidence of CRC cases. With the vast plethora of computer-aided drug design and combinatorial chemistry technologies available, a potentially wide array of therapeutic options from small molecule compounds that can target significant aspects of CDK1 activity exists. Our study investigated potential inhibitors of CDK1 *in silico* and explored their pharmacokinetic properties using a number of computational techniques.

2 Materials and methods

2.1 Ligand preparation

A library of ten thousand bioactive plant compounds was compiled and downloaded from the PubChem database (<https://pubchem.ncbi.nlm.nih.gov>). Pubchem is a publicly available small molecule database (Li et al., 2010; Gozalbes and Pineda-Lucena, 2011; Cheng et al., 2014) used in drug discovery and other applications. It contains information on over 250 million unique chemical molecules along with their descriptors. Before carrying out docking, the two-dimensional (2D) structure data file (SDF) format of the compounds downloaded from the pubchem repository were imported into Maestro (Schrodinger suite) for transformation in the LigPrep module of the software (Epik, 2017a). Ionization states were generated at a pH range of 7 ± 2.0 and stereoisomers were calculated at default (Sastry et al., 2013).

2.2 Protein preparation

The three-dimensional (3D) human crystal X-ray diffraction structure of Cyclin-dependent kinase 1 (<http://www.rcsb.org/pdb/home>; PDB ID: 4Y72) (Brown et al., 2015), was downloaded as a pdb file and imported into Maestro suite. For appropriate structural preparation of the virtual protein using the Protein Preparation wizard of the Schrodinger software, hydrogen, bond orders, and missing side chains were added appropriately while water molecules beyond 5 Å from het groups were deleted before optimization at pH $7.0 + 2.0$. The protein structure was subsequently minimized using the optimized potentials for liquid simulations (OPLS) 2005 force field to a root-mean-square-deviation (RMSD) of 0.30 Å (Epik, 2017b).

2.3 Molecular docking

The compounds were subjected to molecular docking within the active site of CDK1. Using the receptor grid generation feature in the Schrodinger suite, a grid box where each ligand will be docked was generated. The binding pocket where the co-crystallized ligand is located was used to construct the defining box as well as the residue locations within 4 Å radius of the ligand. Molecular docking was performed via the standard precision (SP) and extra precision (XP) of the Glide tool in the software (Friesner et al., 2006; Tripathi et al.,

2013). Following docking, the top compounds with optimal binding scores were selected for further studies.

2.4 Drug likeness and *in silico* ADME predictions

The drug-likeness of the lead compounds was evaluated by subjecting them to computational pharmacological and pharmacokinetic analyses on the SwissADME web server (<http://www.swissadme.ch>) (Daina et al., 2017; Guan et al., 2019), using the generated Canonical SMILES (simplified molecular-input entry system) transformations of the structures (Cheng et al., 2022). Lipinski's rule of five was used to assess the compounds based on the following parameters of molecular weight, number of hydrogen bond donors and acceptors and octanol-water partition (Lipinski et al., 1997; Veber et al., 2002; Yang et al., 2019). Pro-Tox II online server (https://tox-new.charite.de/protox_II) was used to predict the toxicity of the compounds (Banerjee et al., 2018).

2.5 Induced fit docking

The top four ligands with the least binding scores were selected and subjected to the mixed molecular docking and dynamic protocol used by the induced-fit docking (IFD) module of the Maestro algorithm. An implicit solvent model with the OPLS 2005 force field was used to apply the standard IFD protocol to the chosen (centroid) amino acid side chains (8–10, 18–20, 31–33, 64, 80–89, 135, 145–146). The IFD procedure included a ring conformational sampling with a 2.5 kcal/mol energy barrier and a non-planar conformation penalty on amide bonds. Each ligand was given a maximum of 20 permitted postures, with the scaling for both the receptor and the ligand fixed at 0.5. The Prime Refinement technique was used to further refine residues that were within 5 Å of the docked ligand. The protein-ligand complexes with the improved structure were ranked using prime energy. For a final round of Glide docking and scoring, the receptor structures that were less than 30 kcal/mol from the minimal energy structure were submitted. In the next second docking stage, all ligands were re-docked into all revised low-energy receptor structures using Glide XP's default settings.

2.6 Molecular dynamics simulation

The docked complexes were further subjected to molecular dynamics (MD) simulations using the Desmond module of the Schrodinger package on a GPU-enabled Linux computer. To simulate physiological conditions, the model was created in a TIP3P (transferable intermolecular potential-three point), orthorhombic water buffered solvation box of 10 Å dimensions and 0.15 M salt. The simulation model was then minimized and neutralized by adding sodium (Na^+) and chloride (Cl^-) ions, placed at least 20 Å from the ligands. The protein-ligand complexes were simulated for one thousand nanoseconds in an OPLS-2005 force

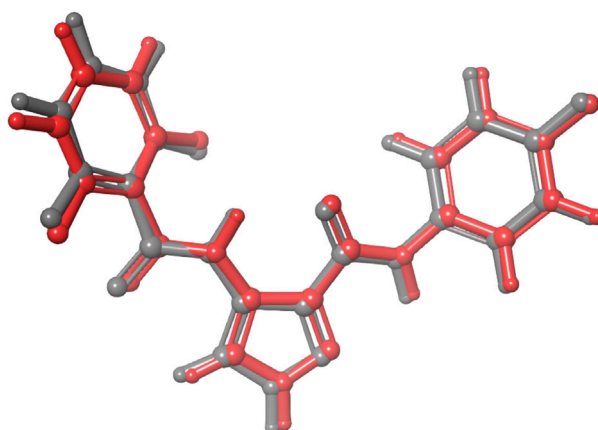


FIGURE 2
Validation of docking showing the co-crystallized ligand (in grey) and the re-docked pose (in red) with an RMSD score of 0.28 Å. The low RMSD value obtained demonstrates the efficacy of the docking protocol used.

field. At the end of the simulation period, the root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (R_g), Ligand RMSD, solvent accessible surface area (SASA), and intramolecular hydrogen bonds (intraHB) were computed and subsequently analyzed (Chow et al., 2008; Bergdorf et al., 2014; Release, 2014).

2.7 Molecular mechanics/generalized born and surface area solvation (MMGBSA) calculations

Molecular mechanics/generalized Born solvent area (MM/GBSA) calculations (Kollman et al., 2000; Massova and Kollman, 2000; Tsui and Case, 2000) were performed to estimate the relative binding energies of the selected ligands to CDK1. This method estimates the amount of free energy involved in interactions of the protein-ligand complex. The binding energy for 2000 independent snapshots of each ligand-protein complex was generated and average binding energies calculated. ΔG of each ligand was calculated following the given equation:

$$\Delta G_{\text{bind}} = \Delta G_{\text{RL}} - (\Delta G_{\text{R}} + \Delta G_{\text{L}}) \quad (1)$$

Where ΔG_{RL} represents the free energy of protein-ligand complex; ΔG_{R} is free energy of receptor, and ΔG_{L} the free energy of ligand.

The free binding energy establishes the binding affinity of ligands to a given protein. A more negative ΔG value indicates better binding of the ligand to the protein (Gilson et al., 1997). The free energy of each state was calculated, using FF14SB force field, by the given equations. . .

$$\Delta G_{\text{bind}} = \Delta E_{\text{gas}} + \Delta G_{\text{sol}} - T\Delta S \quad (2)$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdW}} + E_{\text{ele}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (4)$$

$$G_{\text{SA}} = \gamma \text{SASA} \quad (5)$$

Contribution of each residue to the total binding free energy was also obtained using MM/GBSA method in Amber 18.

3 Results

3.1 Analysis of the binding modes of co-lig and hit compounds in CDK1 active site

We performed molecular docking study of the phyto-compounds against CDK1 target protein. We selected the top four hits with the most favourable binding scores for subsequent studies. To validate the docking procedure, the co-crystallized ligand was re-docked into the binding pocket of CDK1. The RMSD between docked and native pose (Figure 2) is given as 0.28 Å. Figure 3 depicts the result from molecular docking studies of the four top-scoring compounds: Spiraeoside (−12.47 kcal/mol), Robinetin (−12.22 kcal/mol), 6-hydroxy luteolin (−12.07 kcal/mol), Quercetagenin (−12.06 kcal/mol), and the co-crystallized ligand, LZ9 (−13.99 kcal/mol). The 2D protein-ligand interaction diagrams (Figure 4) show the hydrogen and hydrophobic bond network and interactions shared between the compounds and CDK1. In this study, it was observed that all 4 hit compounds formed hydrogen bonds and hydrophobic interactions with the Ile10 residue of CDK1. Hydrogen bond interactions also existed between the compounds and CDK1 via Leu83, Ser84, and/or Gln 132 residues. Hydrophobic interactions were seen between all the compounds and Ile10, Val18, Ala31, Val64, Phe80, Phe82, Leu83, Met85, Leu135 and Ala145 residues of CDK1. LZ9 shared hydrophobic bonding with Tyr15, in addition to the aforementioned residues. Hydrophobic bonds observed from molecular docking were preserved during induced fit docking. Bond interactions for molecular docking, induced fit docking and molecular dynamics simulation are summarized in Table 1. The complexes formed from induced-fit docking are displayed in Figure 5.

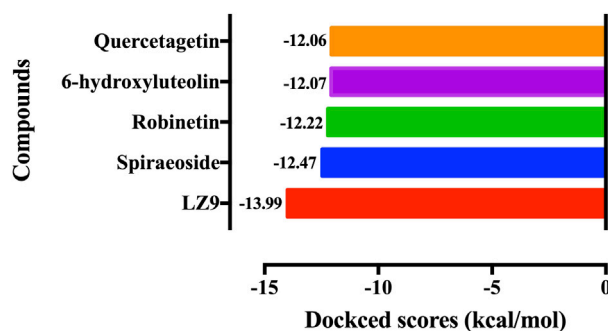


FIGURE 3

Molecular docking results of the top four scoring compounds: Spiraeoside, Robinetin, 6-hydroxyluteolin, Quercetagenin and the co-crystallized ligand (LZ9).

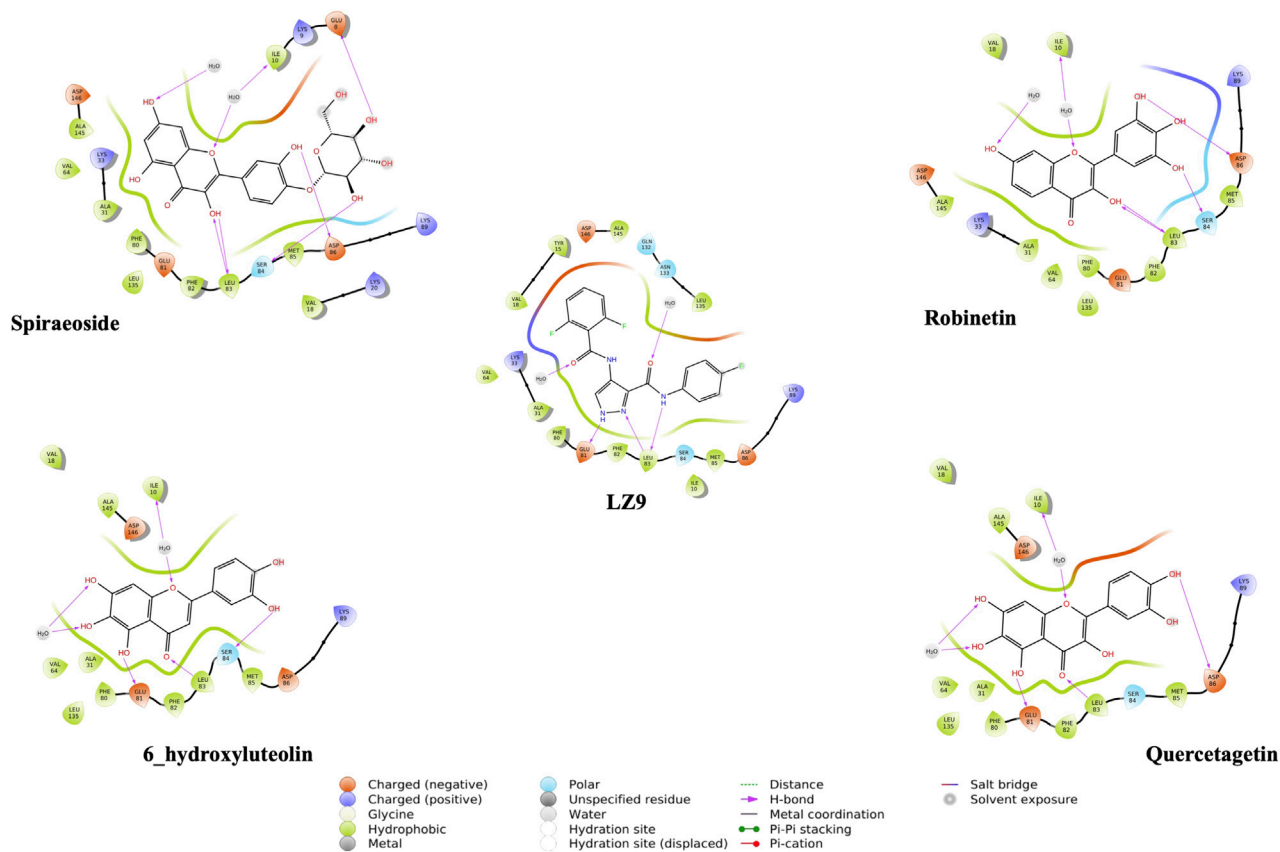


FIGURE 4

2D interaction diagrams of top scoring compounds in complex with CDK1 from molecular docking studies. Maestro 2D interaction diagram (Schrodinger) was used to generate the image.

3.2 *In silico* ADME predictions

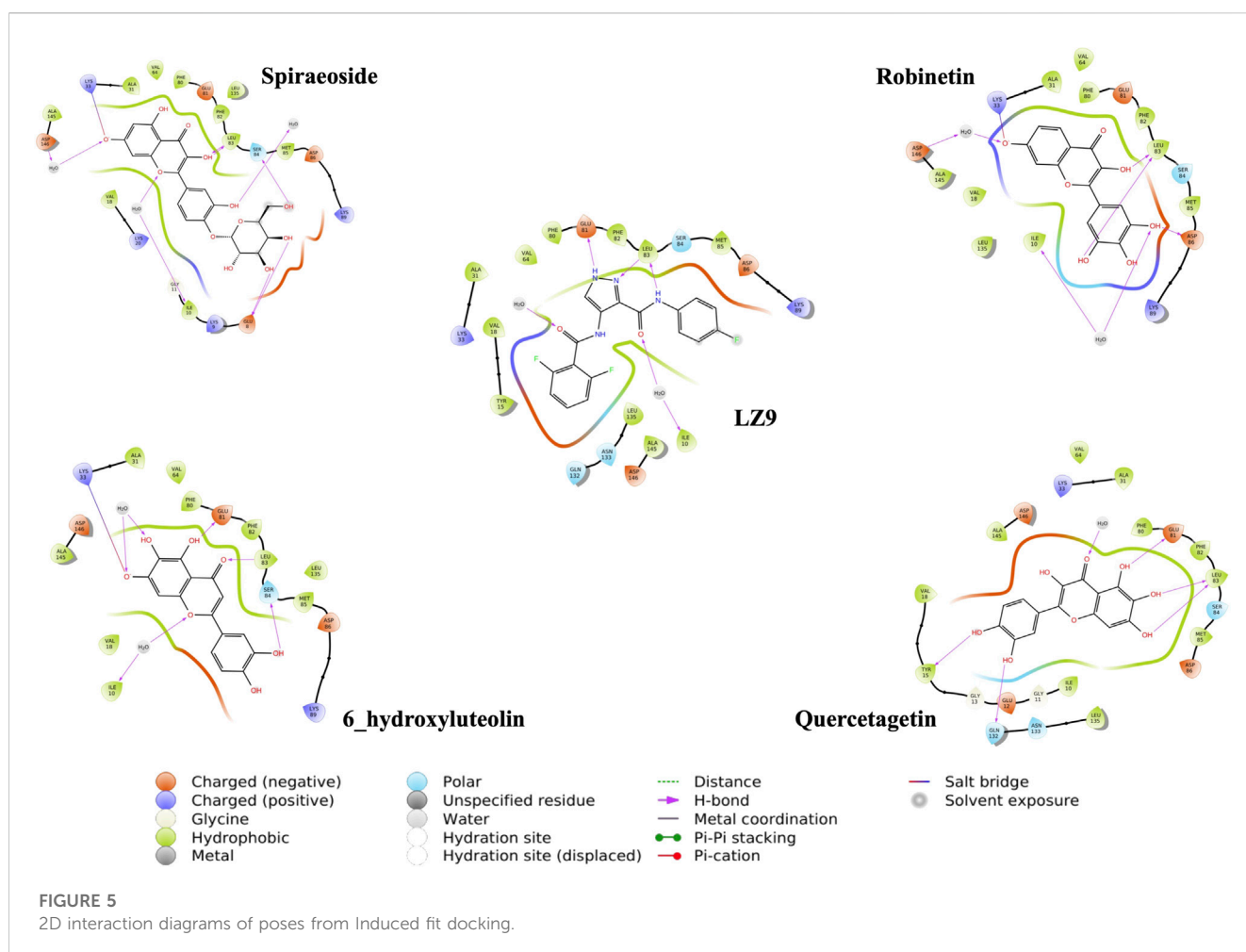
Table 2 gives details of ADME and toxicity predictions. Physicochemical descriptors, ADME (Absorption, Distribution, Metabolism, and Excretion) parameters, and pharmacokinetic properties of the compounds, except spiraeoside, were predicted to be within the permissible parameter ranges.

From toxicity prediction, spiraeoside was predicted to be inactive for all parameters. LZ9 was predicted to be hepatotoxic and carcinogenic, but not mutagenic. Robinetin, 6-hydroxyluteolin, and quercetagenin were predicted to be inactive for hepatotoxicity, while results for carcinogenicity and mutagenicity were predicted to be active. Their probability scores are shown in brackets.

TABLE 1 Bond interactions between ligands and CDK1 during molecular docking, induced fit docking and molecular dynamics simulation.

Compounds	Hydrogen bonds			π - π bonds
	MD	IFD	MDS	MDS
LZ9	Glu81, leu83	Ile10, Glu81, Leu83	Glu81, Leu83, Asp146	Tyr15
Spiraeoside	Glu8, Ile10, Leu83, Ser84, Asp86	Glu8, Ile10, Lys33, Leu83, Ser84, Asp146	—	—
Robinetin	Ile10, Leu83, Ser84, Asp86	Ile10, Lys33, Leu83, Asp86, Asp146	Thr14, Lys33, Val64, Glu81, Leu83, Arg158	Phe80
6_hydroxyluteolin	Ile10, Glu81, Leu83, Ser84	Ile10, Lys33, Glu81, Leu83, Ser84	Thr14, Lys33, Glu51, Leu83, Arg158, Phe147	—
Quercetagetin	Ile10, Glu81, leu83, asp86	Tyr15, Glu81, Leu83, Gln132	Lys33, Glu51, Leu83	—

MD, molecular docking; IFD, induced fit docking; MDS, molecular dynamics simulation.



3.3 RMSD analysis of CDK1 and hit compounds

The protein backbone RMSD ($C\alpha$) & ligand heavy atoms RMSD over the course of simulation time were plotted. The average values of each system were also calculated. The RMSD depicts the dynamic behavior and stability of the complexes over the simulation period and estimates the average change in atom displacement for a given frame relative to a reference frame. All protein frames were first

aligned on the reference frame backbone before determining the RMSD based on the atom selection. RMSD fluctuations give insight to the stability of a ligand within the binding pocket of a protein. A gradual increase in RMSD values was observed in the first 200 ns of the apo protein system, after which it stabilised and had an average value of 2.48 Å (Figure 6A). The CDK1-LZ9 system fluctuated for the first 400 ns, after which it stabilised and had an average of 2.36 Å. CDK1-robinetin system stabilised after 200 ns with an average value of 2.40 Å, while the CDK1-quercetagetin system fluctuated

TABLE 2 Predicted physicochemical descriptors, ADME parameters, pharmacokinetic properties, druglike nature, medicinal chemistry friendliness and toxicity of compounds.

	LZ9	Spiraeoside	Robinetin	6-Hydroxyluteolin	Quercetagenin
Physicochemical properties					
Molecular weight (g/mol)	360.29	464.38	302.24	302.24	318.24
Heavy atoms	26	33	22	22	23
Rotatable bonds	6	4	1	1	1
Hydrogen bond acceptor	6	12	7	7	8
Hydrogen bond donor	3	8	5	5	6
TPSA (Å ²)	86.88	210.51	131.36	131.36	151.59
Lipophilicity					
Log <i>P</i> _{o/w} (XLOGP3)	3.07	1.31	1.61	2.17	1.18
Log <i>P</i> _{o/w} (WLOGP)	4.21	-0.54	1.99	1.99	1.69
Log <i>P</i> _{o/w} (MLOGP)	2.64	-2.59	-0.56	-0.56	-1.08
Consensus Log <i>P</i> _{o/w}	3.03	-0.19	1.12	1.42	0.92
Water solubility					
LogS (ESOL)	-4.10	-3.64	-3.20	-3.55	-3.01
Solubility (mg/mL)	2.89 e-02	1.07e-01	1.91e-01	8.46e-02	3.14e-01
Class	Moderately soluble	Soluble	Soluble	Soluble	Soluble
Pharmacokinetics					
GI absorption	High	Low	High	High	Low
BBB permeant	No	No	No	No	No
P-gp substrate	No	Yes	No	No	No
CYP1A2 inhibitor	Yes	No	Yes	Yes	Yes
CYP2C19 inhibitor	No	No	No	No	No
CYP2C9 inhibitor	No	No	No	No	No
CYP2D6 inhibitor	Yes	No	Yes	Yes	No
CYP3A4 inhibitor	No	No	Yes	Yes	Yes
Druglikeness					
Lipinski	Yes (0 violation)	No (2 violations)	Yes (0 violation)	Yes (0 violation)	Yes (1 violation)
Ghose	Yes	No	Yes	Yes	Yes
Veber	Yes	No	Yes	Yes	No (1 violation)
Bioavailability score	0.55	0.17	0.55	0.55	0.55
Leadlikeness	No	No	Yes	Yes	Yes
Toxicity					
Predicted LD 50 (mg/kg)	1,000	5,000	159	3,919	159
Toxicity Class	4	5	3	5	3
Hepatotoxicity (probability score)	Active (0.65)	Inactive (0.82)	Inactive (0.69)	Inactive (0.69)	Inactive (0.69)
Carcinogenicity (probability score)	Active (0.53)	Inactive (0.85)	Active (0.68)	Active (0.68)	Active (0.68)
Mutagenicity (probability score)	Inactive (0.71)	Inactive (0.76)	Active (0.51)	Active (0.51)	Active (0.51)

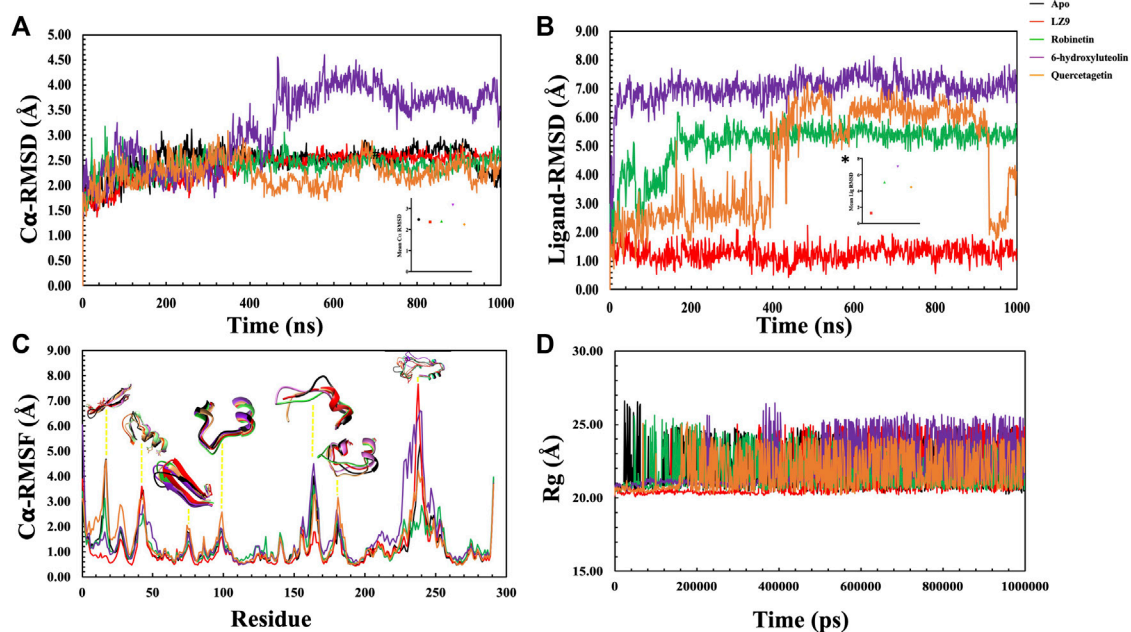


FIGURE 6

Molecular dynamics simulation trajectories of CDK1 in presence of ligands and absence of ligand (Apo) as a function of 1,000 ns simulation time. (A) Root Mean Square Deviation (RMSD) of C α atoms backbone. Insert, # shows the mean C α RMSD values. (B) Root Mean Square Deviation of the ligands with respect to CDK1 protein. The plot displays the RMSD of the ligands after measuring the RMSD of the heavy atoms in the ligand and aligning the protein-ligand complex on the reference protein backbone. If the reported values are noticeably greater than the protein's RMSD, the ligand has probably diffused away from its initial binding site. Insert, * shows the mean RMSD values. (C): Root Mean Square Fluctuation (RMSF) per residue over time. (D): Radius of gyration for the Apo state and CDK1-ligand complexes.

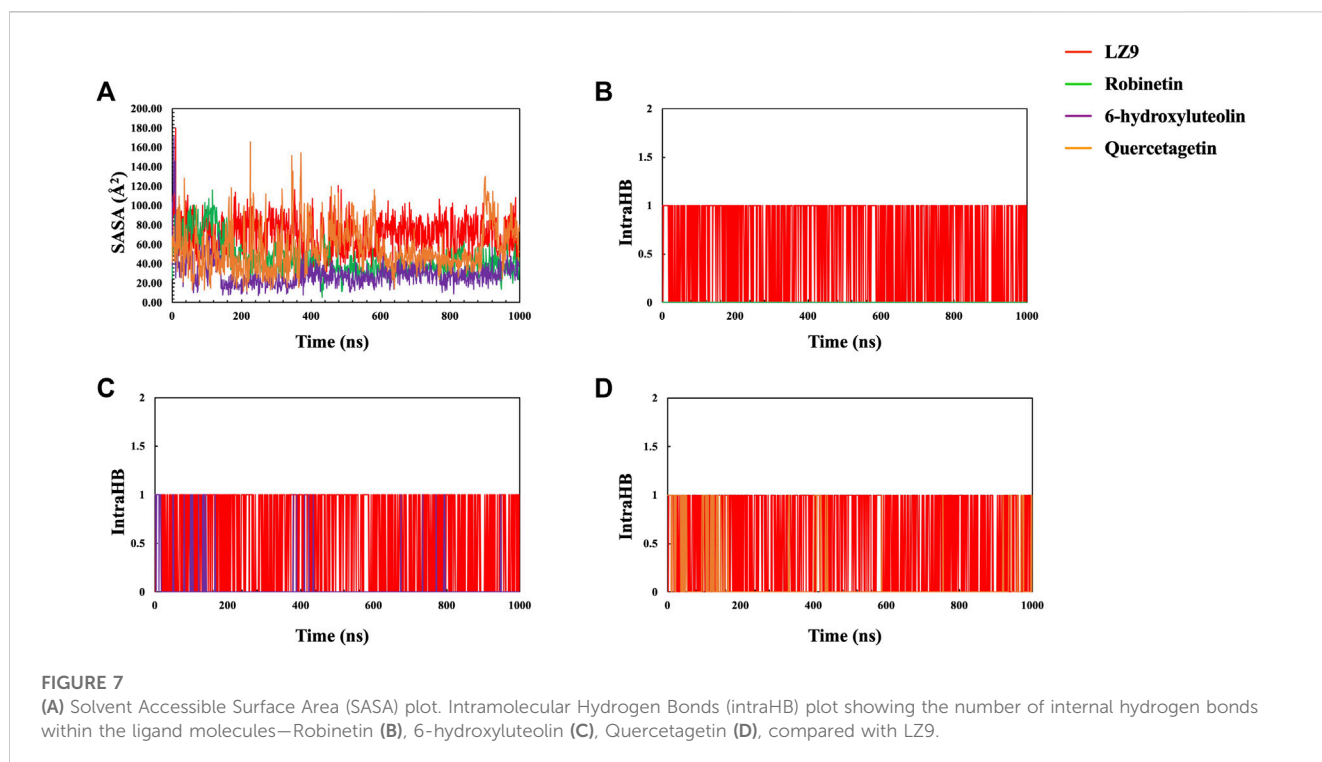
more, however, had an average value of 2.24 Å. The most fluctuations were observed in the CDK1-6-hydroxyluteolin system which also had the highest mean value of 3.17 Å. From the ligand RMSD plot (Figure 6B), the ligands displayed quite stable trends over the simulation period, except for quercetagenin, which fluctuated significantly after 400 ns, even though it had an average RMSD value of 4.48 Å. LZ9 had an average RMSD value of 1.27 Å, robinetin, 5.10 Å and 6-hydroxyluteolin, 7.00 Å.

3.4 RMSF analysis of CDK1 and hit compounds

In general, secondary structure components of a protein like alpha helices and beta strands are more rigid than the unstructured part and fluctuate less than loop regions during molecular dynamics simulation runs. Analysis of the RMSF plots (Figure 6C), show that all the systems displayed similar patterns. Notable distinctions include quercetagenin with the highest peak value at residue 100, robinetin also displayed highest peak at residue 130, whereas LZ9 had highest peak value at residue 240. LZ9 had the least peak at residue 165 while robinetin had the least value between residues 230 and 240 where other system had highest peaks. A comparative 3D representation of the crucial residue regions of the systems which exhibited higher fluctuations has equally been presented.

3.5 Radius of gyration, solvent accessible surface area and intra-hydrogen bond analysis

The radius of gyration for all systems is shown in Figure 6D. When comparing the plots of the apo form to the CDK1 bound states, there is no apparent alteration to the global protein compactness in response to ligand binding. The solvent-accessible surface area describes the surface area of a molecule accessible to water. In addition to examining the impacts on the molecule, it expresses the degree to which a simulated protein interacts with solvent molecules. Evaluating the SASA value is an important way to estimate conformational changes caused by complex interactions. This is essential for characterizing protein-ligand interactions and understanding structural changes during simulations. Solvent Accessible Surface Area (SASA) plots (Figure 7A) show that LZ9 and quercetagenin have high SASA values, while robinetin and 6-hydroxyluteolin have lower SASA values. Intramolecular hydrogen bonds show the number of internal hydrogen bonds (HB) within a ligand molecule and give insight into the behavior of ligands in the active site of a protein. They can also stabilize conformations that are conducive to protein binding, thus increasing the affinity for the receptor. Intra-hydrogen bond (intraHB) plots (Figures 7B–D) reveal that LZ9 had intraHB interactions almost throughout the simulation time, 6-hydroxyluteolin and quercetagenin had more intraHB interactions at the beginning of the simulation, while robinetin maintained a



value of zero, that is zero intraHB interactions throughout the simulation time.

3.6 Protein-ligand contacts

Figure 8 shows a schematic of the detailed ligand atom interactions with the protein residues. The interactions in this trajectory that occur for more than 30% of the simulation period are displayed. LZ9 and robinetin make pi-pi contacts with Tyr15 and Phe80 for more than 50% of the time. All the ligands made contact with Leu83 via hydrogen bonds; LZ9 and robinetin form hydrogen bonds with Glu81, while 6-hydroxy luteolin 6-hydroxy luteolin and quercetagenin form hydrogen bonds with Glu51. Robinetin, 6-hydroxyluteolin, and quercetagenin form salt bridges with Lys33. All the ligands show relatively high solvent exposure. Protein-ligand contacts as illustrated by bar charts (Figure 9) during the simulation, show that hydrogen bonds, hydrophobic interactions, and water bridges were predominant in LZ9. The other ligands had ionic interactions in addition to the aforementioned. All compounds exhibited strong hydrogen bonding on Leu83 and hydrophobic interactions on Leu135. They also showed a water bridge on residue Asp146. Only the lead compounds showed ionic bonds at residue Lys33.

3.7 Potential energy

Figure 10 shows the potential energy plot of the systems in the apo state and in complex with the ligands. When bound to LZ9, robinetin and quercetagenin, we observe energies in the same range

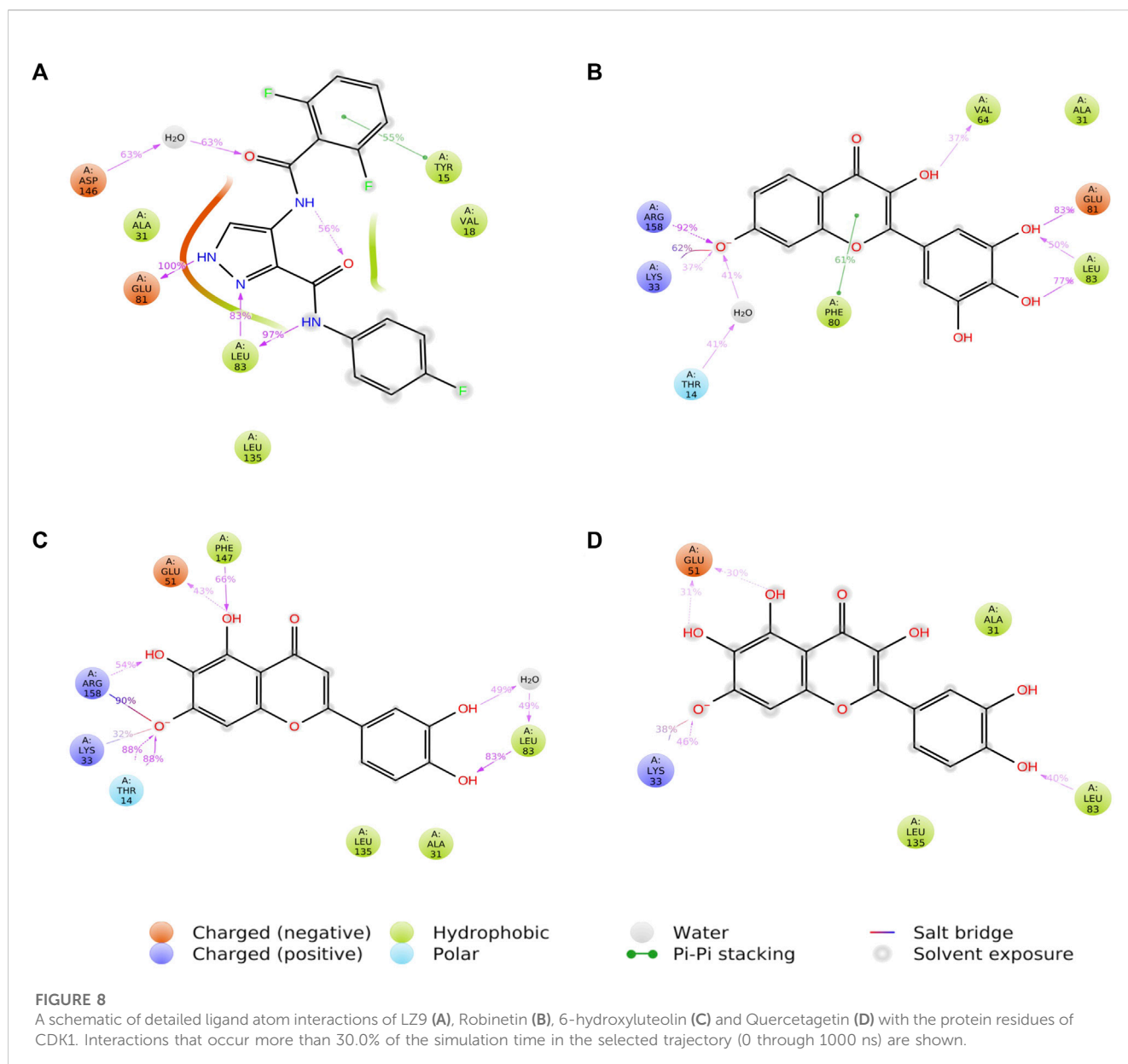
as that of the apo state; $-115,500$ to $-117,000$ kcal/mol. On the other hand, when bound to 6-hydroxyluteolin, a lower energy state of $-144,000$ kcal/mol is observed.

3.8 Free binding energy

Robinetin and 6-hydroxyluteolin exhibited good ΔG values of -30.40 kcal/mol and -30.63 kcal/mol, respectively (Table 3). The co-lig, LZ9, had a value of -34.96 kcal/mol. Quercetagenin had a binding energy of -18.52 kcal/mol. Energy contributions of active site residues responsible for stabilizing the CDK1-ligand complexes are shown in Figure 11. Worthy of note is Glu81 as the residue showing the highest energy contribution in the CDK1-LZ9 complex, while Asp86 is the residue showing highest binding energy contribution to the CDK1-robinetin and CDK1-6-hydroxyluteolin complexes. The CDK1-quercetagenin system has Asp146 as the active site residue contributing the highest energy. Visual inspection of the figure shows that Van der Waals forces contribute more to the interaction between CDK1-LZ9 and CDK1-6-hydroxyluteolin complexes. Robinetin has more electrostatic energy contributing to the stability of the complex it forms with CDK1. Quercetagenin forms weak interaction with CDK1 as a result of the overall low Van der Waals forces and electrostatic energy.

4 Discussion

The prevalence of colorectal cancer is fast becoming a public health concern (Holowatyj et al., 2020). Development of new pharmacologically therapeutic anti-cancer drug agents has been



hampered by the limited molecular understanding of the oncogenic initiation and the interaction of biochemical mechanisms/factors, although traditional chemotherapy options have produced impressive results in the treatment of CRC with the possibility of relapses.

Natural products are a valuable source of novel therapeutic leads given that they contain a wide variety of chemical components (Cragg and Newman, 2013; Newman and Cragg, 2016). They have been essential in the development of new drugs, particularly for the treatment of cancerous disorders (Atanasov et al., 2015; Harvey et al., 2015). The compounds identified from this study were found to be in the class of flavonoids. There is growing evidence that flavonoids may be useful in the prevention and treatment of cancer and they are also reported to have a variety of anti-cancer effects, including inducing apoptosis in cancer cells, inhibiting cell proliferation, and may enhance chemotherapy (Abotaleb et al., 2018). Flavonoids have also

been reported to play a potential role in colorectal cancer therapies (Amado et al., 2014).

Given the importance of cyclin-dependent kinases (CDK) in cancer, targeting CDKs is a viable strategy for combating the cancer menace (Sofi et al., 2022a). Because of the significant role it plays in one of the phases of cell cycle control, which is a crucial one for the growth of tumor cells, specifically targeting CDK1 became a strategic focus for optimized treatment. The drug discovery pipeline greatly benefits from the use of computational screening techniques (Choudhury et al., 2022). In recent years, computer-aided drug development methods have been widely used to find lead compounds with therapeutic value. These technologies include molecular docking, *in silico* pharmacokinetic profiling, and molecular dynamics simulations (MDS) (Wan, 2013; Salmaso and Moro, 2018). MDS are anticipated to become more significant as new pharmacological therapies are developed (Durrant and

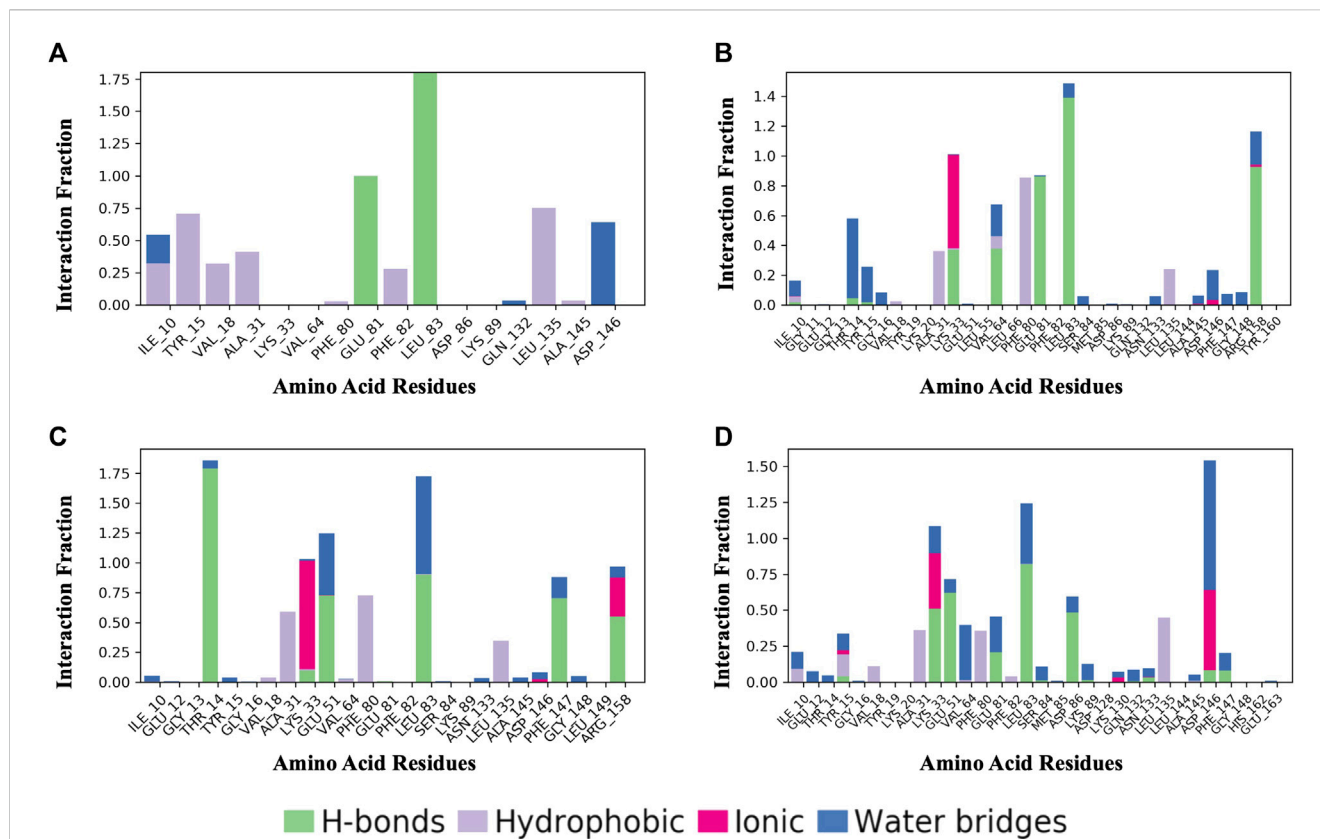


FIGURE 9

Protein-Ligand Contacts. (A) LZ9. (B) Robinetin. (C) 6-hydroxyluteolin. (D) Quercetagenin. Contact interactions can be categorized by type and summarized, as shown in the plot. Protein-ligand contacts are categorized into four types: Hydrogen bonds, Hydrophobic, Ionic and Water Bridges. The stacked bar charts are normalized over the course of the trajectory: for example, a value of 0.7 suggests that 70% of the simulation time the specific interaction is maintained. Values over 1.0 are possible as some protein residue may make multiple contacts of same subtype with the ligand.

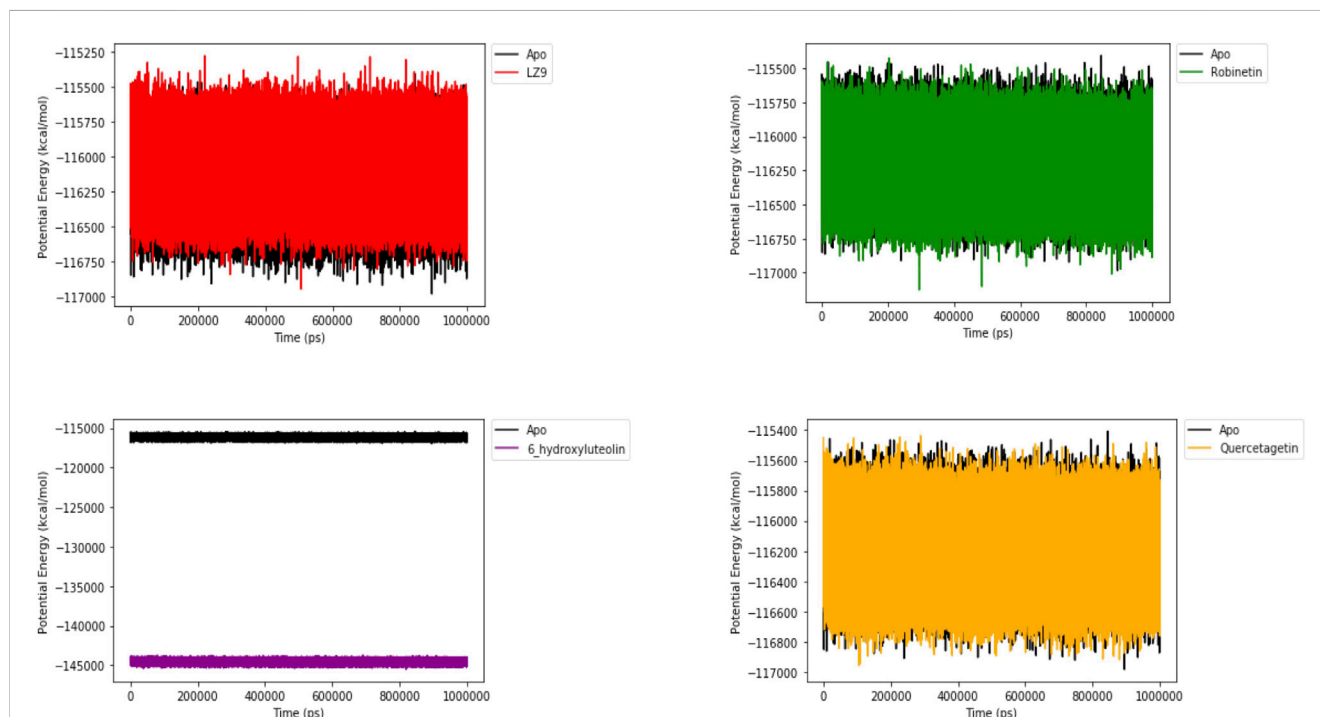


FIGURE 10

Potential energy plot of 1,000 ns simulation time. The plots show the energies of the systems in apo and bound states.

TABLE 3 Energy component calculations obtained by MM/GBSA method.

Compounds	Energy components (kcal/mol)				
	ΔE_{vdw}	ΔE_{ele}	ΔG_{gas}	ΔG_{sol}	ΔG_{bind}
LZ9	-41.88 ± 2.97	-22.02 ± 3.62	-63.90 ± 4.39	28.95 ± 2.98	-34.96 ± 2.90
Robinetin	-28.32 ± 3.22	-36.89 ± 7.32	-65.21 ± 6.41	34.81 ± 5.61	-30.40 ± 2.66
6-hydroxyluteolin	-37.20 ± 2.32	-28.94 ± 7.81	-66.14 ± 7.60	35.51 ± 5.17	-30.63 ± 3.78
Quercetagetin	-29.00 ± 3.043	-15.95 ± 10.06	-44.95 ± 10.04	26.43 ± 6.44	-18.52 ± 5.17

ΔE_{vdw} , Van der Waals energy; ΔE_{ele} , Electrostatic energy; ΔG_{gas} , Gas phase energy; ΔG_{sol} , Solvation energy; ΔG_{bind} , Total binding energy.

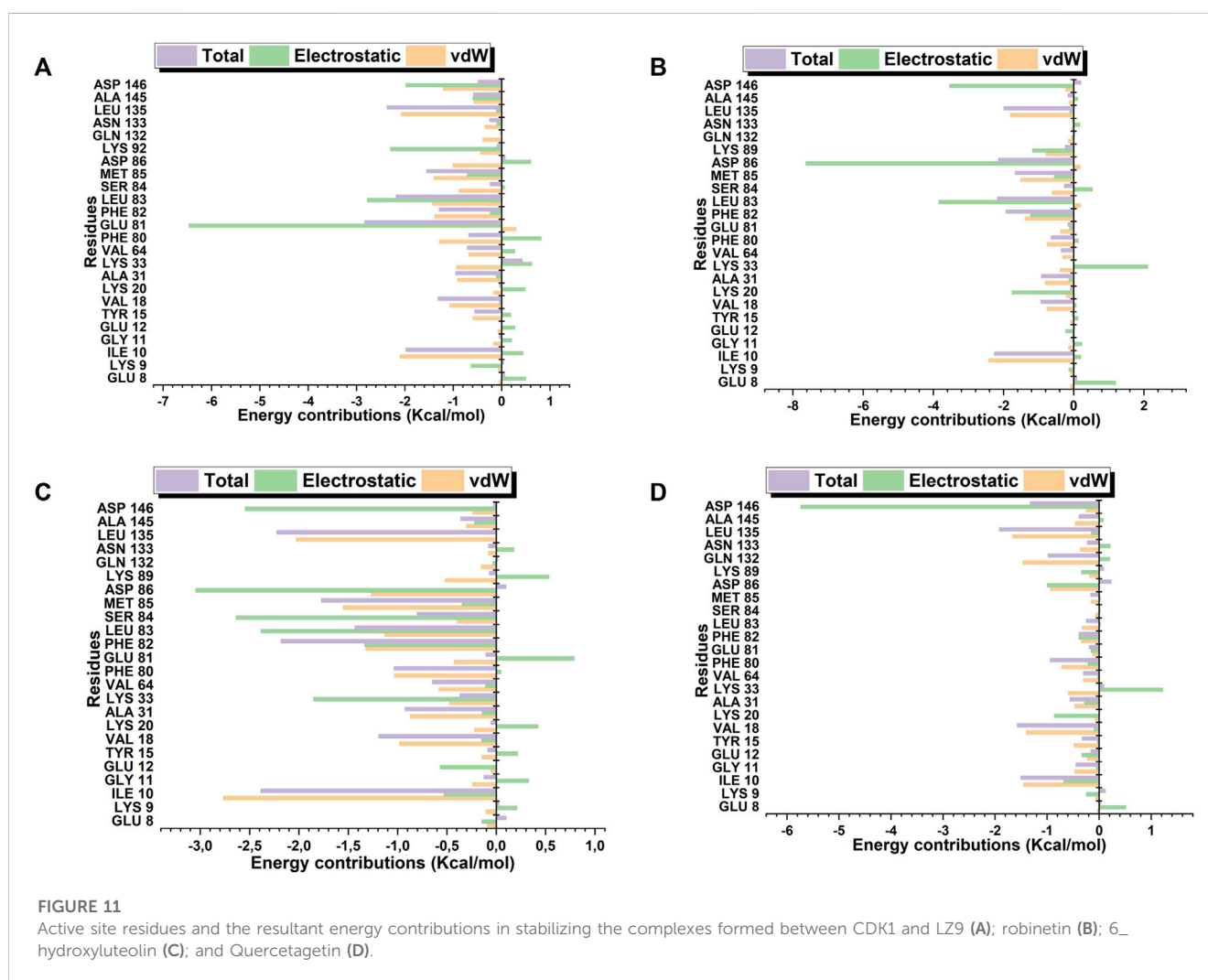


FIGURE 11

Active site residues and the resultant energy contributions in stabilizing the complexes formed between CDK1 and LZ9 (A); robinetin (B); 6-hydroxyluteolin (C); and Quercetagetin (D).

McCammon, 2011). Several studies now use these computational methods to quickly identify potential inhibitors of key proteins in disease progression as treatment options (Choudhary et al., 2020; Sofi et al., 2022a; Rathod et al., 2022; Castrosanto et al., 2023) others have identified experimentally validated leads (Kalani et al., 2013; Yilmaz Goler et al., 2022). Here, we used these computational techniques to screen phyto-compounds and identify possible inhibitors of CDK1.

Molecular docking is a computational method used to predict the preferred binding orientation of a ligand to a receptor. The

software algorithm used predicts possible binding orientations and provides scores ranked according to the predicted binding affinity between the receptor-ligand complex (Huang and Zou, 2010; Salmaso and Moro, 2018). To better understand the binding dynamics of the four compounds selected in this study to CDK1, they were subjected to molecular dynamics simulations (MDS). Prior to running simulation studies, we carried out induced fit docking (IFD) (Achilonu et al., 2020). Induced fit docking is a more realistic model of how proteins and ligands interact in nature as it considers the flexibility of the protein and ligand, allowing them to

move and rotate during the docking process, resulting in more accurate predictions of the binding of ligands. For traditional rigid docking, the proteins and ligands are treated as rigid bodies, and the docking algorithm attempts to find the best orientation of the ligand in the protein's binding site. However, in reality, proteins are flexible molecules, which can undergo conformational changes upon ligand binding. Both minor backbone relaxations and substantial side-chain conformational changes in the receptor structure are accounted for by the induced-fit docking (IFD) technique (Sherman et al., 2006; Xu and Lill, 2013).

All systems ran through the complete one thousand nanoseconds duration during MDS, except spiraeoside, which ran for only 700 ns? It could be that due to conformational changes in the protein, spiraeoside left the binding site of the protein during the simulation run. For this reason, the post-simulation analysis could not be completed, consequently, results for spiraeoside were excluded.

It is important to pay attention to the protein's RMSD value to properly infer structural changes occurring in a protein which could affect ligand binding. C α -RMSD demonstrated a stable structural conformation across the course of the simulation, showing that it was balanced. The low RMSD values of LZ9, robinetin and quercetagenin indicate low fluctuations and suggests that the system maintained stability during simulation and the protein did not undergo significant structural changes as a result of binding of the ligands. The higher RMSD value of 6-hydroxyluteolin system might suggest that the binding of this ligand might have perturbed the structure of the CDK1 protein.

Ligand RMSD provides insight into how stably the bound ligand sits within the protein binding pocket. The Ligand RMSD results show that LZ9, robinetin, and 6-hydroxyluteolin did not move much, therefore forming a stable complex with CDK1, unlike quercetagenin which had a lot of fluctuations. The fluctuations observed with quercetagenin after 400 ns tell us that it is not tightly bound to the protein and has possibly diffused away from the binding pocket, so it forms an unstable protein-ligand complex and is less likely to perform its intended function.

RMSF measures the structural flexibility of a protein and characterizes local perturbations along a protein chain (Cob-Calan et al., 2019). It also provides information on how each amino acid residue is changing relative to its reference position during simulation (Sarkar and Sen, 2022). Here, we analysed the RMSF of the CDK1 protein in the apo and bound states. Peaks represent regions of the protein where the simulation-induced fluctuations were greatest (Tripathi et al., 2022). Illustrative 3D images of the peaks have been presented to clearly elucidate the silent fluctuation dynamics depicted by the represented peaks. The 3D images reveal a comparative conformational variation exhibited by the various system due to the perturbations created by the MD simulation. The 3D images have been color-coded to tally with their respective corresponding RMSF trajectories. The spikes observed from RMSF results may have been caused by increased flexibility within the protein's loop regions. It may be inferred from the RMSF result that CDK1 has a great deal of flexibility to accommodate the ligand at the binding pocket.

The radius of gyration measures how compact a protein structure is and is equivalent to its principal moment of inertia, describing the common separation of each scattered element from

the molecule's mass center (Lobanov et al., 2008). During simulation, a more rigid structure is represented by lower Rg values (Jiang et al., 2019). Since not much difference was observed between the Apo and bound states, we can say that the binding of the ligands did not interfere with the compact structure of CDK1. In this study, the Rg, RMSD, and RMSF were correlated, and this relationship shed light on how complex compressibility affected the variations in complex RMSD and the residual fluctuations seen in RMSF. No noticeable changes are observed in the bound and unbound states of CDK1, suggesting that the protein structure retained its compact conformation. Our results indicate that the surface area of LZ9 and quercetagenin have higher exposure to water than robinetin and 6-hydroxyluteolin. This result is in tandem with that observed from the ligand-atom interaction diagram.

Hydrogen and hydrophobic bonds play key roles in protein-ligand interactions. Crucial hydrophobic and/or hydrogen bonds, including Ile10, Phe80, Leu83, and Ser84 between CDK1 and the lead compounds that were first seen after molecular docking were maintained during MDS. Phe80 has previously been described as a hydrophobic checkpoint enabling water molecules to gain access to the catalytic portion of the ATP catalytic domain for an additional hydrogen bond (McGrath et al., 2005). Robinetin is seen to form a pi-pi bond with Phe80 during simulation run, which may further account for its strong inhibitory properties. 6-hydroxyluteolin also forms a bond with Thr14 for 88% of the simulation time. Inhibition of Thr14 by small molecule compounds has been proposed as a promising option for the regulation of the G2/M transition of the cell cycle (Schmidt et al., 2017). Inhibition of CDK1 due to the phosphorylation of the kinase on Tyr15 residue by flavonoids (Casagrande and Darbon, 2001) as well as induction of G₂/M phase arrest and apoptosis by the flavonoid tamarixetin on human leukemia cells (Nicolini et al., 2014) have previously been reported.

The hydrophobic and hydrogen bonds formed between the compounds and CDK1 may play a vital role in stabilizing the complexes for the compounds to exert their inhibitory efficacies (Zhang et al., 2011). Ionic bonds observed on only the lead compounds might also play a significant role in their inhibition of CDK1. Interestingly, the lead compounds are all of the flavonoid class; another study by (Navarro-Retamal and Caballero, 2016)) also gave insight into flavonoids as CDK1 inhibitors. Natural flavonoids such as Luteolin and Daidzein have also been found to inhibit CDK1 (Ravishankar et al., 2013).

Higher free energy values in fact correlate with more preferable binding affinities of a ligand to a protein. Robinetin, which had good binding energy to CDK1, as revealed by the ΔG values, might have been a result of higher electrostatic and van der Waals energy contributions. Interactions between robinetin and Asp86 residue which was observed from molecular docking and induced fit docking results was maintained after the MM/GBSA run. We can say this interaction might be a key factor to stabilizing the CDK1-robinetin complex. Energetic analysis reveals that van der Waals interaction and non-polar contributions are favorable in the formation of complexes and amine group of the ligand, which plays a crucial role for binding (Tripathi et al., 2012; He et al., 2023). Potential energy (PE) and RMSD are both vital indicators of protein stability. In previous work done by Yuan et al. (2012) to ascertain if the structures were stabilized, MD simulation was

examined using PE and RMSD from the original model structure (Coutsias et al., 2004) The potential energy plot for 6-hydroxyluteolin is a reflection of the trajectory pattern exhibited by the Ca-RMSD plot of 6-hydroxyluteolin system between 230 ns and 420 ns, where a perturbation on the 6-hydroxyluteolin system seems to have created conformational alteration within the system, making it assume a new/different trajectory. The new structural conformational deviation of 6-hydroxyluteolin relative to other systems could presumably have produced the obtained PE.

The physicochemical properties of a compound are important parameters in the pipeline of drug development as they directly affect the absorption, distribution, metabolism, excretion, and toxicity profile (ADMET) of a drug candidate (Manallack et al., 2013). The drug-like and ADMET predictions are used in identifying properties of a ligand suitable for formulation into drugs. Even though all the compounds are adequately hydrophilic and lipophilic, robinetin and 6-hydroxyluteolin showed high intestinal absorption. This implies that the compounds are capable of achieving pharmaceutical and therapeutic bioavailability at the intestinal linings and target cells in the colon. This is further corroborated by their bioavailability scores. A study on natural compounds as inhibitors of CDK1 by Saikat (Saikat, 2021) also reported similar ADME properties.

Lipinski's Rule of five is used to assess and predict the drug-likeness of compounds based on oral absorption, membrane permeability and bioavailability (Lipinski et al., 1997). Although, parenteral administration may be applicable in most cases of cancer treatment, oral administration is a valuable treatment option for cancer patients as it is more convenient and reduces the frequency of their clinic visits or venipuncture. In recent times, oral anti-cancer drugs such as capecitabine, topotecan and vinorelbine are successfully used in the treatment of a variety of tumours like colon, breast, ovarian or lung cancer (Colomer et al., 2010). With the advent of targeted deliveries, most of the drug molecules approved as oral medications for cancer treatment including hormone therapies such as, dasatinib, sorafenib, imatinib, lapatinib and abiraterone have resulted in good clinical outcomes (Gralow et al., 2008; Rehman and Rosenberg, 2012).

Despite the fact that natural products frequently violate these rules due to their complexity, they still serve as reliable sources for novel drug discovery and development (Doak et al., 2014). Some anti-cancer agents such as vincristine, irinotecan, paclitaxel and etoposide, widely used for decades in cancer treatments originate from natural compounds (Huang et al., 2021).

The ProTox-II server is a web-based tool that predicts the potential toxicity of a given small molecule compound on various organs and systems in the body. The toxicity is predicted using a machine learning model that has been trained on a sizable dataset of toxic and non-toxic molecules. The toxicity prediction and probability score are considered important parameters for predicting a compound's toxicity (Han et al., 2019). The probability score can be used to evaluate the prediction's level of confidence. A forecast with a high probability score is more certain to occur, whereas one with a low probability value may not. Additionally, if the organ toxicity and/or toxicity end points of a compound are classified as active, it suggests that the compound may have harmful effects if it is ingested, inhaled, or otherwise exposed to the body. Conversely, "inactive" could mean that the

compound has not been predicted to have a toxic effect. Early identification of toxicity in the drug development steps is critical in reducing late-stage attrition rates (Roberts et al., 2014; Segall and Barber, 2014). The predicted toxicity class exhibited by the compounds indicates that they can be administered within the accepted safety doses. Although the test ligands are predicted to be non-hepatotoxic, lead optimization may be required to modify the other toxicity endpoints.

5 Conclusion and next steps

In this study, computational techniques including molecular docking, induced fit docking, molecular dynamics simulation and MM/GBSA were used to study the potential inhibitory activities of natural compounds against CDK1. The findings provided insight into molecular interactions of significant residues involved in inhibitory CDK1 activities through molecular docking and molecular dynamics simulation results. The current investigation of small molecule inhibitors in the activity of the CDK1 explored viable options for the roles of Robinetin and 6-hydroxyluteolin in the pathogenesis of CRC. This research, however, provides a basis for further examination of potential phytochemicals for medical intervention of several cancers. This study aimed to identify potential small molecule inhibitors of CDK1 hence the use of molecular modeling approaches was deployed to highlight compounds of strong binding energies, good spatial amino acid residue arrangement, and better pharmacokinetic properties among others. The MD simulations showed that the hit compounds stably interacted with the CDK1 enzyme and maintained firm positions within the binding pocket of the enzyme over the simulation timeframe. The findings of this study should undergo experimental validation to confirm the interaction between the top ligands identified in our research and the target protein. Additionally, the ADME/Tox properties that were predicted using computational models need to be verified through experimental testing.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: NCBI PDB.

Author contributions

UO: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Writing—original draft, Writing—review and editing. OE: Data curation, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Writing—original draft, Writing—review and editing. CO: Formal Analysis, Methodology, Writing—review and editing. PrG: Data curation, Formal Analysis, Methodology, Software, Writing—review and editing. PaG: Formal Analysis, Methodology, Writing—review and editing. SO: Formal Analysis, Methodology, Writing—review and editing. HA: Formal Analysis, Methodology, Writing—review and editing. EI: Data

curation, Formal Analysis, Methodology, Software, writing–review and editing. IA: Data curation, Formal Analysis, Methodology, Project administration, Software, Writing–review and editing. OA: Conceptualization, Data curation, Formal Analysis, Methodology, Project administration, Software, Supervision, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors thank the National Institutes of Health (NIH) Office of Data Science Strategy (ODSS), the National Center for Biotechnology Information (NCBI) and the Genetics Society of America for their immense support before and during the

References

- Abotaleb, M., Samuel, S. M., Varghese, E., Varghese, S., Kubatka, P., Liskova, A., et al. (2018). Flavonoids in cancer and apoptosis. *Cancers* 11 (1), 28. doi:10.3390/cancers11010028
- Achilonu, I., Iwuchukwu, E. A., Achilonu, O. J., Fernandes, M. A., and Sayed, Y. (2020). Targeting the SARS-CoV-2 main protease using FDA-approved Isavuconazonium, a P2–P3 α -ketoamide derivative and Pentagastrin: an in-silico drug discovery approach. *J. Mol. Graph. Model.* 101, 107730. doi:10.1016/j.jmgm.2020.107730
- Adeoti, M. L., Oguntola, S. A., Olugbenga-Bello, A. I., Oladimeji, O. J., and Jegede, S. O. (2016). Colorectal cancer: knowledge and risk factors among adults in a sub urban Nigeria community. *J. Med. Sci. Clin. Res.* 4 (9), 1247. doi:10.18535/jmscr/v4i9.28
- Adhikari, D., Zheng, W., Shen, Y., Gorre, N., Ning, Y., Halet, G., et al. (2012). Cdk1, but not Cdk2, is the sole Cdk that is essential and sufficient to drive resumption of meiosis in mouse oocytes. *Hum. Mol. Genet.* 21 (11), 2476–2484. doi:10.1093/hmg/ddso61
- Amado, N. G., Predes, D., Moreno, M. M., Carvalho, I. O., Mendes, F. A., and Abreu, J. G. (2014). Flavonoids and wnt/ β -catenin signaling: potential role in colorectal cancer therapies. *Int. J. Mol. Sci.* 15 (7), 12094–12106. doi:10.3390/ijms150712094
- Atanasov, A. G., Waltenberger, B., Pferschy-Wenzig, E. M., Linder, T., Wawrosch, C., Uhrin, P., et al. (2015). Discovery and resupply of pharmacologically active plant-derived natural products: a review. *Biotechnol. Adv.* 33 (8), 1582–1614. doi:10.1016/j.biotechadv.2015.08.001
- Banerjee, P., Eckert, A. O., Schrey, A. K., and Preissner, R. (2018). ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.* 46 (W1), W257–W263. doi:10.1093/nar/gky318
- Bergdorf, M., Kim, E. T., Rendleman, C. A., and Shaw, D. E. (2014). *Desmond/GPU performance as of november 2014*.
- Bray, F. F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. J. (2020). Erratum: global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca Cancer J. Clin.* 70 (4), 313. doi:10.3322/caac.21609
- Brown, N. R., Korolchuk, S., Martin, M. P., Stanley, W. A., Moukhametdzianov, R., Noble, M. E., et al. (2015). CDK1 structures reveal conserved and unique features of the essential cell cycle CDK. *Nat. Commun.* 6 (1), 6769. doi:10.1038/ncomms7769
- Casagrande, F., and Darbon, J. M. (2001). Effects of structurally related flavonoids on cell cycle progression of human melanoma cells: regulation of cyclin-dependent kinases CDK2 and CDK11. Abbreviations: CDK, cyclin-dependent kinase; CKI, CDK inhibitor; PI 3-kinase, phosphatidylinositol 3-kinase; PKC, protein kinase C; DTT, dithiothreitol; RIPA, radioimmunoprecipitation assay buffer. *Biochem. Pharmacol.* 61 (10), 1205–1215. doi:10.1016/s0006-2952(01)00583-4
- Castrosanto, M. A., Abrera, A. T., Manalo, M. N., and Ghosh, A. (2023). *In silico* evaluation of binding of phytochemicals from bayati (*Anamirta cocculus* Linn) to the glutathione-s-transferase of Asian Corn Borer (*Ostrinia furnacalis* Guenée). *J. Biomol. Struct. Dyn.* 41 (7), 2660–2666. doi:10.1080/07391102.2022.2036240
- October 2022 Omics codeathon organised in collaboration with the African Society for Bioinformatics and Computational Biology (ASBCB).

Conflict of interest

The 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.

Cheng, F., Li, W., Zhou, Y., Shen, J., Wu, Z., Liu, G., et al. (2022). *admetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties*. doi:10.1021/ci300367a

Cheng, T., Pan, Y., Hao, M., Wang, Y., and Bryant, S. H. (2014). PubChem applications in drug discovery: a bibliometric analysis. *Drug Discov. today* 19 (11), 1751–1756. doi:10.1016/j.drudis.2014.08.008

Chikwambi, Z., Hidjo, M., Chikondowa, P., Afolabi, L., Aketch, V., Jayeoba, G., et al. (2023). *BioRxiv*. Multi-omics data integration approach identifies potential biomarkers for Prostate cancer

Choudhary, M. I., Shaikh, M., tul-Wahab, A., and ur-Rahman, A. (2020). *In silico* identification of potential inhibitors of key SARS-CoV-2 3CL hydrolase (Mpro) via molecular docking, MMGBSA predictive binding energy calculations, and molecular dynamics simulation. *Plos one* 15 (7), e0235030. doi:10.1371/journal.pone.0235030

Choudhury, C., Murugan, N. A., and Priyakumar, U. D. (2022). Structure-based drug repurposing: traditional and advanced AI/ML-aided methods. *Drug Discov. Today* 27, 1847–1861. doi:10.1016/j.drudis.2022.03.006

Chow, E., Rendleman, C. A., Bowers, K. J., Dror, R. O., Hughes, D. H., Gullingsrud, J., et al. (2008). *Desmond performance on a cluster of multicore processors*. China, DE Shaw Research Technical Report DESRES/TR–2008-01.

Cob-Calan, N. N., Chi-Uluac, L. A., Ortiz-Chi, F., Cerqueda-García, D., Navarrete-Vázquez, G., Ruiz-Sánchez, E., et al. (2019). Molecular docking and dynamics simulation of protein β -tubulin and antifungal cyclic lipopeptides. *Molecules* 24 (18), 3387. doi:10.3390/molecules24183387

Colomer, R., Alba, E., González-Martín, A., Paz-Ares, L., Martín, M., Llombart, A., et al. (2010). Treatment of cancer with oral drugs: a position statement by the Spanish Society of Medical Oncology (SEOM). *Ann. Oncol.* 21 (2), 195–198. doi:10.1093/annonc/mdp595

Coutsias, E. A., Seok, C., and Dill, K. A. (2004). Using quaternions to calculate RMSD. *J. Comput. Chem.* 25 (15), 1849–1857. doi:10.1002/jcc.20110

Cragg, G. M., and Newman, D. J. (2013). Natural products: a continuing source of novel drug leads. *Biochimica Biophysica Acta (BBA)-General Subj.* 1830 (6), 3670–3695. doi:10.1016/j.bbagen.2013.02.008

Daina, A., Michielin, O., and Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep.* 7 (1), 42717. doi:10.1038/srep42717

Dar, R. A., Shahnawaz, M., Ahanger, M. A., and ul Majid, I. (2023). Exploring the diverse bioactive compounds from medicinal plants: a review. *A Rev.* 12, 189–195. doi:10.31254/phyto.2023.12307

Deep, A., Marwaha, R. K., Marwaha, M. G., Nandal, R., and Sharma, A. K. (2018). Flavopiridol as cyclin dependent kinase (CDK) inhibitor: a review. *New J. Chem.* 42 (23), 18500–18507. doi:10.1039/C8NJ04306j

Doak, B. C., Over, B., Giordanetto, F., and Kihlberg, J. (2014). Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. *Chem. Biol.* 21, 1115–1142. doi:10.1016/j.chembiol.2014.08.013

- Durrant, J. D., and McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biol.* 9 (1), 71–79. doi:10.1186/1741-7007-9-71
- El Abed, F., Baraket, G., Nyamari, M. N., Naitore, C., and Awe, O. I. (2023). Differential expression analysis of miRNAs and mRNAs in epilepsy uncovers potential biomarkers. *bioRxiv*. doi:10.1101/2023.09.11.557132
- Epik, S. (2017a). *Scrodinger release 2017*. New York, NY: LigPrep.
- Epik, S. (2017b). *Impact, schrodinger*. New York, NY: LLC.
- Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136 (5), E359–E386. doi:10.1002/ijc.29210
- 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 (21), 6177–6196. doi:10.1021/jm051256o
- Gilson, M. K., Given, J. A., Bush, B. L., and McCammon, J. A. (1997). The statistical-thermodynamic basis for computation of binding affinities: a critical review. *Biophys. J.* 72 (3), 1047–1069. doi:10.1016/S0006-3495(97)78756-3
- Gozalbes, R., and Pineda-Lucena, A. (2011). Small molecule databases and chemical descriptors useful in chemoinformatics: an overview. *Comb. Chem. high throughput Screen.* 14 (6), 548–558. doi:10.2174/138620711795767857
- Gralow, J., Ozols, R. F., Bajorin, D. F., Cheson, B. D., Sandler, H. M., Winer, E. P., et al. (2008). Clinical cancer advances 2007: major research advances in cancer treatment, prevention, and screening—a report from the American Society of Clinical Oncology. *J. Clin. Oncol.* 26 (2), 313–325. doi:10.1200/JCO.2007.15.4088
- Greenlee, J. D., Liu, K., Lopez-Cavestany, M., and King, M. R. (2022). Piezol mechano-activation is augmented by resveratrol and differs between colorectal cancer cells of primary and metastatic origin. *Molecules* 27 (17), 5430. doi:10.3390/molecules27175430
- Guan, L., Yang, H., Cai, Y., Sun, L., Di, P., Li, W., et al. (2019). ADMET-score—a comprehensive scoring function for evaluation of chemical drug-likeness. *Medchemcomm* 10 (1), 148–157. doi:10.1039/c8md00472b
- Han, Y., Zhang, J., Hu, C. Q., Zhang, X., Ma, B., and Zhang, P. (2019). *In silico* ADME and toxicity prediction of ceftazidime and its impurities. *Front. Pharmacol.* 10, 434. doi:10.3389/fphar.2019.00434
- Harvey, A. L., Edrada-Ebel, R., and Quinn, R. J. (2015). The re-emergence of natural products for drug discovery in the genomics era. *Nat. Rev. Drug Discov.* 14 (2), 111–129. doi:10.1038/nrd4510
- He, F., Wang, X., Wu, Q., Liu, S., Cao, Y., Guo, X., et al. (2023). Identification of potential ATP-competitive cyclin-dependent kinase 1 inhibitors: *de novo* drug generation, molecular docking, and molecular dynamics simulation. *Comput. Biol. Med.* 155, 106645. doi:10.1016/j.compbiomed.2023.106645
- Holowatyj, A. N., Maude, A. S., Musa, H. S., Adamu, A., Ibrahim, S., Abdullahi, A., et al. (2020). Patterns of early-onset colorectal cancer among Nigerians and African Americans. *JCO Glob. Oncol.* 6, 1647–1655. doi:10.1200/GO.20.00272
- Huang, M., Lu, J. J., and Ding, J. (2021). Natural products in cancer therapy: past, present and future. *Nat. Prod. Bioprospect.* 11, 5–13. doi:10.1007/s13659-020-00293-7
- Huang, S. Y., and Zou, X. (2010). Advances and challenges in protein-ligand docking. *Int. J. Mol. Sci.* 11 (8), 3016–3034. doi:10.3390/ijms11083016
- Ikwu, F. A., Isyaku, Y., Obadawo, B. S., Lawal, H. A., and Ajibowu, S. A. (2020). *In silico* design and molecular docking study of CDK2 inhibitors with potent cytotoxic activity against HCT116 colorectal cancer cell line. *J. Genet. Eng. Biotechnol.* 18 (1), 51–52. doi:10.1186/s43141-020-00066-2
- Jiang, Z., You, L., Dou, W., Sun, T., and Xu, P. (2019). Effects of an electric field on the conformational transition of the protein: a molecular dynamics simulation study. *Polymers* 11 (2), 282. doi:10.3390/polym11020282
- Kalani, K., Agarwal, J., Alam, S., Khan, F., Pal, A., and Srivastava, S. K. (2013). *In silico* and *in vivo* anti-malarial studies of 18 β glycyrrhetic acid from *Glycyrrhiza glabra*. *PLoS one* 8 (9), e74761. doi:10.1371/journal.pone.0074761
- Kalous, J., Jansová, D., and Šušor, A. (2020). Role of cyclin-dependent kinase 1 in translational regulation in the M-phase. *Cells* 9 (7), 1568. doi:10.3390/cells9071568
- Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., et al. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts Chem. Res.* 33 (12), 889–897. doi:10.1021/ar000033j
- Li, J., Wang, Y., Wang, X., and Yang, Q. (2020). CDK1 and CDC20 overexpression in patients with colorectal cancer are associated with poor prognosis: evidence from integrated bioinformatics analysis. *World J. Surg. Oncol.* 18 (1), 50–51. doi:10.1186/s12957-020-01817-8
- Li, Q., Cheng, T., Wang, Y., and Bryant, S. H. (2010). PubChem as a public resource for drug discovery. *Drug Discov. today* 15 (23–24), 1052–1057. doi:10.1016/j.drudis.2010.10.003
- Liao, H., Ji, F., and Ying, S. (2017). CDK1: beyond cell cycle regulation. *Aging (Albany NY)* 9 (12), 2465–2466. doi:10.18632/aging.101348
- Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23 (1–3), 3–25. doi:10.1016/S0169-409X(96)00423-1
- Lobanov, M. Y., Bogatyreva, N. S., and Galzitskaya, O. V. (2008). Radius of gyration as an indicator of protein structure compactness. *Mol. Biol.* 42, 623–628. doi:10.1134/s0026893308040195
- Manallack, D. T., Prankerd, R. J., Yuriev, E., Oprea, T. I., and Chalmers, D. K. (2013). The significance of acid/base properties in drug discovery. *Chem. Soc. Rev.* 42 (2), 485–496. doi:10.1039/c2cs35348b
- Massova, I., and Kollman, P. A. (2000). Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. *Perspect. Drug Discov. Des.* 18, 113–135. doi:10.1023/a:1008763014207
- Matthews, H. K., Bertoli, C., and de Bruin, R. A. (2022). Cell cycle control in cancer. *Nat. Rev. Mol. Cell Biol.* 23 (1), 74–88. doi:10.1038/s41580-021-00404-3
- McGrath, C. F., Pattabiraman, N., Kellogg, G. E., Lemcke, T., Kunick, C., Sausville, E. A., et al. (2005). Homology model of the CDK1/cyclin B complex. *J. Biomol. Struct. Dyn.* 22 (5), 493–502. doi:10.1080/07391102.2005.10531227
- Navarro-Retamal, C., and Caballero, J. (2016). Flavonoids as CDK1 inhibitors: insights in their binding orientations and structure-activity relationship. *PLoS One* 11 (8), e0161111. doi:10.1371/journal.pone.0161111
- Newman, D. J., and Cragg, G. M. (2016). Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* 79 (3), 629–661. doi:10.1021/acs.jnatprod.5b01055
- Nicolini, F., Burmistrova, O., Marrero, M. T., Torres, F., Hernández, C., Quintana, J., et al. (2014). Induction of G2/M phase arrest and apoptosis by the flavonoid tamarixetin on human leukemia cells. *Mol. Carcinog.* 53 (12), 939–950. doi:10.1002/mc.22055
- Nyamari, M. N., Omar, K. M., Fayehun, A. F., Dachi, O., Bwana, B. K., and Awe, O. I. (2023). Expression level analysis of ACE2 receptor gene in african-American and non-african-American COVID-19 patients. *bioRxiv*. doi:10.1101/2023.09.11.557129
- Nzungize, L., Kengne-Ouafu, J. A., Wesonga, M. R., Umuhoza, D., Murithi, K., Kimani, P., et al. (2022). Transcriptional profiles analysis of COVID-19 and malaria patients reveals potential biomarkers in children. *bioRxiv*, 2022.06.30.498338. doi:10.1101/2022.06.30.498338
- Rathod, S., Shinde, K., Porlekar, J., Choudhari, P., Dhavale, R., Mahuli, D., et al. (2022). Computational exploration of anti-cancer potential of flavonoids against cyclin-dependent kinase 8: an *in silico* molecular docking and dynamic approach. *ACS omega* 8, 391–409. doi:10.1021/acsomega.2c04837
- Ravishanker, D., Rajora, A. K., Greco, F., and Osborn, H. M. (2013). Flavonoids as prospective compounds for anti-cancer therapy. *Int. J. Biochem. Cell Biol.* 45 (12), 2821–2831. doi:10.1016/j.biocel.2013.10.004
- Rehman, Y., and Rosenberg, J. E. (2012). Abiraterone acetate: oral androgen biosynthesis inhibitor for treatment of castration-resistant prostate cancer. *Drug Des. Dev. Ther.* 6, 13–18. doi:10.2147/DDDT.S15850
- Release, S. I.: Desmond molecular dynamics system. DE Shaw Research, New York, NY, *Maestro-desmond interoperability tools*, version. 2014;3.
- Roberts, R. A., Kavanagh, S. L., Mellor, H. R., Pollard, C. E., Robinson, S., and Platz, S. J. (2014). Reducing attrition in drug development: smart loading preclinical safety assessment. *Drug Discov. today* 19 (3), 341–347. doi:10.1016/j.drudis.2013.11.014
- Saikat, A. S. (2021). An *in silico* approach for potential natural compounds as inhibitors of protein CDK1/cks2. *Chem. Proc.* 8 (1), 5. doi:10.3390/ecsoc-25-11721
- Salmaso, V., and Moro, S. (2018). Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: an overview. *Front. Pharmacol.* 9, 923. doi:10.3389/fphar.2018.00923
- Sarkar, I., and Sen, A. (2022). *In silico* screening predicts common cold drug Dextromethorphan along with Prednisolone and Dexamethasone can be effective against novel Coronavirus disease (COVID-19). *J. Biomol. Struct. Dyn.* 40 (8), 3706–3710. doi:10.1080/07391102.2020.1850528
- Sasidharan, S., Chen, Y., Saravanan, D., Sundram, K. M., and Latha, L. Y. (2011). Extraction, isolation and characterization of bioactive compounds from plants' extracts. *Afr. J. traditional, complementary Altern. Med.* 8 (1), 1–10. doi:10.4314/ajcm.v8i1.60483
- Sastry, M. G., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. computer-aided Mol. Des.* 27, 221–234. doi:10.1007/s10822-013-9644-8
- Schmidt, M., Rohe, A., Platzer, C., Najjar, A., Erdmann, F., and Sippl, W. (2017). Regulation of G2/M transition by inhibition of WEE1 and PKMYT1 kinases. *Molecules* 22 (12), 2045. doi:10.3390/molecules22122045
- Segall, M. D., and Barber, C. (2014). Addressing toxicity risk when designing and selecting compounds in early drug discovery. *Drug Discov. today* 19 (5), 688–693. doi:10.1016/j.drudis.2014.01.006
- Sherman, W., Beard, H. S., and Farid, R. (2006). Use of an induced fit receptor structure in virtual screening. *Chem. Biol. Drug Des.* 67 (1), 83–84. doi:10.1111/j.1747-0285.2005.00327.x

- Simon, L., Imane, A., Srinivasan, K. K., Pathak, L., and Daoud, I. (2017). *In silico* drug-designing studies on flavanoids as anticancer agents: pharmacophore mapping, molecular docking, and Monte Carlo method-based QSAR modeling. *Interdiscip. Sci. Comput. Life Sci.* 9, 445–458. doi:10.1007/s12539-016-0169-4
- Sofi, S., Mehraj, U., Qayoom, H., Aisha, S., Almilaibary, A., Alkhanani, M., et al. (2022a). Targeting cyclin-dependent kinase 1 (CDK1) in cancer: molecular docking and dynamic simulations of potential CDK1 inhibitors. *Med. Oncol.* 39 (9), 133. doi:10.1007/s12032-022-01748-2
- Sofi, S., Mehraj, U., Qayoom, H., Aisha, S., Asdaq, S. M., Almilaibary, A., et al. (2022b). Cyclin-dependent kinases in breast cancer: expression pattern and therapeutic implications. *Med. Oncol.* 39 (6), 106. doi:10.1007/s12032-022-01731-x
- Tripathi, N., Goel, B., Bhardwaj, N., Sahu, B., Kumar, H., and Jain, S. K. (2022). Virtual screening and molecular simulation study of natural products database for lead identification of novel coronavirus main protease inhibitors. *J. Biomol. Struct. Dyn.* 40 (8), 3655–3667. doi:10.1080/07391102.2020.1848630
- Tripathi, S. K., Muttineni, R., and Singh, S. K. (2013). Extra precision docking, free energy calculation and molecular dynamics simulation studies of CDK2 inhibitors. *J. Theor. Biol.* 334, 87–100. doi:10.1016/j.jtbi.2013.05.014
- Tripathi, S. K., Singh, S. K., Singh, P., Chellaperumal, P., Reddy, K. K., and Selvaraj, C. (2012). Exploring the selectivity of a ligand complex with CDK2/CDK1: a molecular dynamics simulation approach. *J. Mol. Recognit.* 25 (10), 504–512. doi:10.1002/jmr.2216
- Tsui, V., and Case, D. A. (2000). Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers Orig. Res. Biomol.* 56 (4), 275–291. doi:10.1002/1097-0282(2000)56:4<275::aid-bip10024>3.0.co;2-e
- Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., and Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 45 (12), 2615–2623. doi:10.1021/jm020017n
- Wan, H. (2013). What ADME tests should be conducted for preclinical studies? *ADMET DMPK* 1 (3), 19–28. doi:10.5599/admet.1.3.9
- Weitz, J., Koch, M., Debus, J., Höhler, T., Galle, P. R., and Büchler, M. W. (2005). Colorectal cancer. *Lancet* 365 (9454), 153–165. doi:10.1016/s0140-6736(05)17706-x
- Xu, M., and Lill, M. A. (2013). Induced fit docking, and the use of QM/MM methods in docking. *Drug Discov. Today Technol.* 10 (3), e411–e418. doi:10.1016/j.ddtec.2013.02.003
- Yang, H., Lou, C., Sun, L., Li, J., Cai, Y., Wang, Z., et al. (2019). admetSAR 2.0: web-service for prediction and optimization of chemical ADMET properties. *Bioinformatics* 35 (6), 1067–1069. doi:10.1093/bioinformatics/bty707
- Yilmaz Goler, A. M., Jannuzzi, A. T., Bayrak, N., Yıldız, M., Yıldırım, H., Otsuka, M., et al. (2022). *In vitro* and *in silico* study to assess toxic mechanisms of hybrid molecules of quinone-benzocaine as plastoquinone analogues in breast cancer cells. *ACS omega* 7 (34), 30250–30264. doi:10.1021/acsomega.2c03428
- Yuan, X. H., Wang, Y. C., Jin, W. J., Zhao, B. B., Chen, C. F., and Yang, J. (2012). Structure-based high-throughput epitope analysis of hexon proteins in B and C species human adenoviruses (HAdVs). *PLoS One* 7 (3), e32938. doi:10.1371/journal.pone.0032938
- Zhang, L., Zhu, H., Wang, Q., Fang, H., Xu, W., and Li, M. (2011). Homology modeling, molecular dynamic simulation and docking studies of cyclin dependent kinase 1. *J. Mol. Model.* 17, 219–226. doi:10.1007/s00894-010-0710-z
- Zhu, Y., Bian, Y., Zhang, Q., Hu, J., Li, L., Yang, M., et al. (2019). Construction and analysis of dysregulated lncRNA-associated ceRNA network in colorectal cancer. *J. Cell. Biochem.* 120 (6), 9250–9263. doi:10.1002/jcb.28201
- Zhu, Y., Bian, Y., Zhang, Q., Hu, J., Li, L., Yang, M., et al. (2020b). LINC00365 promotes colorectal cancer cell progression through the Wnt/ β -catenin signaling pathway. *J. Cell. Biochem.* 121 (2), 1260–1272. doi:10.1002/jcb.29359
- Zhu, Y., Li, K., Zhang, J., Wang, L., Sheng, L., and Yan, L. (2020a). Inhibition of CDK1 reverses the resistance of 5-Fu in colorectal cancer. *Cancer Manag. Res.* 12, 11271–11283. doi:10.2147/cmar.s255895