

Palbociclib and Ribociclib analogs targeting CDK4 and CDK6 for breast cancer treatment: A molecular modeling study

Dhoha Triki¹, Sravya Kuchibhotla¹ and Denis Fourches*¹.

¹ Department of Chemistry, Bioinformatics Research Center, North Carolina State University,
Raleigh, North Carolina, USA

** To whom email should be sent: dfourch@ncsu.edu*

Abstract

Palbociclib and Ribociclib are FDA approved drugs that target cyclin dependent kinases CDK4 and CDK6 to treat breast cancer. Herein, we conducted a cheminformatics analysis of a large set of their analogs. The study highlights *(i)* several clusters of similar compounds with excellent inhibitory profiles, *(ii)* their shared CDK-ligand intermolecular interactions as predicted by 3D docking, and *(iii)* key dynamic interactions shared by highly active CDK4/6 inhibitors.

1. Introduction

Under normal circumstances, the cell cycle is well controlled resulting in the perfect balance between cell turnover and cell death. However, in cells where normal growth mechanisms are disrupted, uncontrolled cell growth can give rise to clones that can divide and grow in a disorderly way, which can lead to cancer. Several proteins are involved in the regulation of the cell cycle, especially the Cyclin Dependent Kinases (CDKs) family, which consists of serine/threonine kinases. CDK4 and CDK6 are amongst the key regulators; they bind to the cyclin D to be active. For the negative regulation, the retinoblastoma protein (Rb) acts as a tumor suppressor, and its dysfunction causes cancer. The transition from G1 to S phase (DNA synthesis) is regulated by the cyclin D-CDK4/6-Rb protein pathway where cyclin D1-CDK4/6 complex phosphorylates Rb proteins and dissociates them from E2F transcription factor (Pandey et al., 2019). However, diverse mechanisms such as the overexpression of cyclin D, mutation or amplification of CDK4/6 and loss of Rb, can induce the activation of cyclin D-CDK4/6 and subsequent hyperphosphorylation of Rb, leading to uncontrolled cell proliferation (Hamilton & Infante, 2016; Pandey et al., 2019).

Breast cancer is a malignant tumor that can develop in mammary gland, particularly in women. The use of CDK4 and CDK6 inhibitors has been proposed as an efficient strategy to treat breast cancer (Chen et al., 2019). These inhibitors separate the CDK4/6-cyclin D1 complex and block Rb protein phosphorylation, which results in the arrest of the cell cycle from progressing to S phase and preventing subsequent cancer proliferation (Liu et al., 2018; Vanarsdale et al., 2015). Three drugs, which target the ATP binding site of CDK4, have currently been approved by the FDA (Food and Drug Administration): Palbociclib, Ribociclib and Abemaciclib (Hamilton & Infante, 2016; Pernas et al., 2018; Sherr et al., 2016). Studies have shown that they are well

tolerated, although several side effects were reported such as neutropenia (few neutrophils, a type of white blood cells), along with diarrhea and nausea (Chen et al., 2019; Iwata, 2018). Palbociclib was developed by Pfizer and approved in 2015; Ribociclib was developed by Novartis and Astex Pharmaceuticals and approved later in 2017; and Abemaciclib was developed by Eli Lilly and approved in 2017 (Sobhani et al., 2019). These three drugs have similar efficiency, however, the two analogs Palbociclib and Ribociclib drugs were associated with a significantly reduced risk of diarrhea and anemia, relative to Abemaciclib (Petrelli et al., 2019). Moreover, Palbociclib and Ribociclib have higher affinity for CDK4 ($pIC_{50} = 7.95$ and $pIC_{50} = 8$ respectively) and CDK6 ($pIC_{50} = 7.89$ and $pIC_{50} = 7.41$ respectively) comparing to other cell cycle CDKs. On the contrary, Abemaciclib is less selective (Portman et al., 2013), although it has very good activity ($pIC_{50} = 8.70$ for CDK4 and 8.30 for CDK6) (Poratti & Marzaro, 2019). For these reasons, this study focuses on analogs of Palbociclib and Ribociclib that inhibit both CDK4 and CDK6, and we studied their properties comparing to the two approved drugs.

Some previous works have been done to study analogs of Palbociclib. Wang et al. have conducted an *in vitro* study to synthesize compounds by adding carbamates or amides to the tail piperazine ring of Palbociclib, with measuring their inhibitory activity (Wang et al., 2016). They selected one analog that has a higher activity than Palbociclib and showed that it has a similar binding mode to Palbociclib based on molecular docking. An *in vivo* study suggests that this compound is a promising tumor inhibitor, which provided additional information for the design of new drugs. In another study, Rondla et al applied *in silico* analysis of six compounds using molecular docking and 3D-QSAR pharmacophore mapping to identify new leads that target only CDK4. They selected six compounds that are considered as leads from a set of 287 hits (Rondla et al., 2017).

Surprisingly, there is no integrative analysis of all published analogs in the current literature. Herein, we aimed at characterizing the physical-chemical properties of a series of Palbociclib and Ribociclib analogs extracted from the ChEMBL database in regard of their CDK4/6 inhibitory activity. To do so, we performed a 2D analysis by clustering 234 selected analogs (Davies et al., 2015; Gaulton et al., 2017) according to their structural properties. Then, we pursued 3D molecular docking calculations towards CDK4 and 6 binding sites for these analogs. For the compounds that achieved the top docking scores (some achieving even better scores than both Palbociclib and Ribociclib), molecular dynamic simulations were performed. Some of the findings could help in designing new CDK inhibitors.

2. Material and methods

2.1. Dataset collection and curation

All chemical and biological data was extracted from the ChEMBL database (Davies et al., 2015; Gaulton et al., 2017). We searched for analogs of the two FDA-approved drugs Palbociclib and Ribociclib that had a similarity greater than 70% and obtained 5,803 Palbociclib analogs and 6,198 Ribociclib analogs. Then, we merged the results from the two queries and removed structural duplicates. We selected all compounds that both inhibit CDK4 and CDK6 and have an experimental biological activity expressed as IC₅₀ (concentration at which a drug inhibits a biological process by 50% expressed in nM). In this study, we converted IC₅₀ to pIC₅₀ in order to scale all values ($\text{pIC}_{50} = -\log_{10}(\text{IC}_{50} \times 10^{-9})$). All analogs were curated following the protocol developed earlier (Fourches et al., 2010, 2016). Briefly:

- We kept compounds for which we had both an experimental pIC₅₀ value for CDK4 and one for for CDK6. These values were extracted from the same source (*e.g.*, CHEMBL14762) where authors did the same assay for both CDK4 and CDK6;

- For compounds with more than one pIC₅₀ value corresponding for CDK4 and/or CDK6, we calculated the average. This is the case for 20 compounds where authors did assays for multiple stereoisomers (*e.g.*, CHEMBL3641995). It is also the case for one compound where pIC₅₀ values were extracted from different reference sources (*e.g.*, CHEMBL189963).

After curation, we obtained a set of 701 analogs that target CDK4 and 235 analogs that target CDK6. Thus, 467 analogs that target CDK4 don't have a corresponding assay value for CDK6, and one compound has an assay for CDK6 but not for CDK4. These 468 compounds were removed. Overall, we obtained a final dataset of 234 unique analogs with both experimental CDK4

and CDK6 activities. Amongst these compounds, 20 correspond to a particular stereoisomer. Both configurations were considered for the 3D analysis. Thus, we accounted for 254 structures in the molecular docking simulation.

SDF files of 2D structures of the selected analogs were generated from SMILES codes using OpenBabel (O'Boyle et al., 2011) in KNIME software (Berthold et al., 2008). The stereochemistry of the 20 structures was checked with Maestro from Schrödinger Suite (Schrödinger, LLC, New York, 2019) according to the document reference structure. Analogs compounds were prepared using Ligprep from Schrödinger, where protonation states were generated at $\text{pH } 7 \pm 2$ and an energy minimization using the OPLS3 force field was performed (Harder et al., 2016). The curated compound structures are available in the Supplementary Material.

2.2. Distribution and clustering

We studied the distribution of pIC_{50} in CDK4 and CDK6, and molecular properties of the 234 analogs using barplots in R software (R Development Core Team). We then generated 2D descriptors of 234 analogs with RDKit tool from KNIME software (Berthold et al., 2008). Descriptors were cleaned by removing null variance and redundant descriptors with a correlation coefficient greater than 0.9. Then we performed a hierarchical clustering of the 234 analogs based on 2D generated descriptors and we presented them in a circular dendrogram. We used Euclidean distance and ward linkage and we colored the compounds according to pIC_{50} values of CDK4, the more is green, the more is active.

2.3. Protein structure preparation

The crystal structure of human CDK6/Vcyclin complexed with Palbociclib inhibitor was selected from the Protein Data Bank (PDB code: 2EUF) (Berman et al., 2000). Only the chain B corresponding to CDK6 was kept for the 3D study.

For CDK4, no wild type structure complexed to Palbociclib or Ribociclib exists in the PDB. In fact, Day et al. crystallized four structures of CDK4 in complex with cyclin D1, and all CDK4 are mutants (PDB codes: 2W96, 2W99, 2W9F, and 2W9Z) (Day et al., 2009). Moreover, Takaki et al. crystallized wild-type CDK4 complexed to cyclin D3 (PDB code: 3G33), but it is in the apo form (Takaki et al., 2009). CDK4 and CDK6 have a sequence identity of 67%. Therefore, we built an homology model based on the wild type CDK4 sequence, the 3G33 chain A structure, as well as taking the structure of CDK6 as a additional template (PDB code: 2EUF chain B). Prime's energy-based method from Schrödinger suite was used (Jacobson et al., 2002, 2004) to build and refine the homology model.

The receptors were prepared using Protein Preparation Wizard from Schrödinger Suite, where missing side chains were generated with Prime (Jacobson et al., 2002, 2004), hydrogen atoms were added, H-bonds assignment were performed at pH=7 with PROPKA, and an energy minimization was performed with the OPLS3 force field (Harder et al., 2016).

2.4. Molecular Docking

Molecular docking simulation of 254 compounds was done with the CDK4 homology model and the CDK6 structures using Glide from Schrödinger Suite 2019-1 in standard precision (SP) and extra precision (XP) modes, with a flexible ligand sampling (Friesner et al., 2004, 2006;

Halgren et al., 2004). For the 20 compounds that have two stereoisomers, average docking scores were calculated when they were represented in the circular dendrogram.

A grid box of 35 x 35 x 35 Å was generated for CDK6 based on the centroid of the Palbociclib ligand using Receptor Grid Generation from Schrödinger. The outer box defines the volume in which the grid potentials are computed. The grid center has as coordinates x=30.3, y=21.97 and z=60.27. All ligand atoms of a valid pose must be located within this outer box. The inner box defines the volume that the ligand center explores during the exhaustive site-point search and has as dimensions 12 x 12 x 12 Å. For the CDK4 model, the grid with the same dimension was generated based on the selected residues of the binding pocket defined in the literature (Rondla et al., 2017).

2.5. *Molecular Dynamics Simulations*

The top-scored docking poses and the FDA-approved drugs were taken as a template for solvent-explicit and all-atom molecular dynamics simulations that were conducted using the GPU-accelerated Desmond from the Schrödinger Suite (Desmond Molecular Dynamics System, 2018; Bowers et al., 2006). Complexes were neutralized by adding counter-ions. The system was immersed in a $10 \times 10 \times 10$ Å buffered orthorhombic box with a TIP3P water model. We have worked in the isothermal-isobar NPT ensemble class, in which N (number of atoms), P (pressure = 1.01325 bar) and T (temperature = 300K) remain constant. OPLS3e force field (Harder et al., 2016; Jacobson et al., 2004; Jorgensen et al., 1996; Tirado-rives & Jorgensen, 1988) is used. The simulation time for each run is 50 ns, with a recording interval of 50.0 ps for trajectory calculation. Prior to each simulation, Desmond's default relaxation protocol was performed to equilibrate the system of interest. Protein-Ligand interactions plots were then generated and analyzed.

3. Results and discussion

3.1. *Distribution of analogs' activity and properties*

From the ChEMBL database, we collected a series of Palbociclib and Ribociclib analogs with a similarity greater than 70%. After curation (see Material and Methods), we obtained a set of 234 compounds with their corresponding experimental pIC₅₀ values for both CDK4 and CDK6 kinases. The values of pIC₅₀ are ranging from 4.87 to 9.89 for CDK4 and from 4.63 to 9.44 for CDK6 (**Table S1**). As expected, pIC₅₀ values of CDK4 and CDK6 are correlated ($R^2 = 0.85$, see **Figure S1**). Moreover, we calculated the difference between pIC₅₀ values of CDK4 and CDK6 for each compound. We showed that all compounds have values in the same range, with an absolute difference less or equal to 1.2 log units (except for two compounds CHEMBL3702048 and CHEMBL3972254, see **Figure S2**). For CHEMBL3702048, pIC₅₀ is equal to 7.03 for CDK4 and as low as 5.0 for CDK6. This compound is reported in Patent No. WO 2009/085185 A1 (2009, example 140). For CHEMBL3972254, pIC₅₀ is equal to 5.0 for CDK4 and 6.96 for CDK6. This compound is reported in the same Patent No. WO 2009/085185 A1 (2009, example 162).

Interestingly, 75% of the total number of compounds that inhibit CDK4 and 71% of the total number of compounds that inhibit CDK6 in our dataset have a pIC₅₀ value comprised between 7.5 and 9. Thus, the vast majority of those compounds are very active. Again, Palbociclib has a pIC₅₀ of 7.95 for CDK4 and 7.89 for CDK6, whereas Ribociclib has a pIC₅₀ of 8.0 for CDK4 and 7.41 for CDK6 (**Figure 1**).

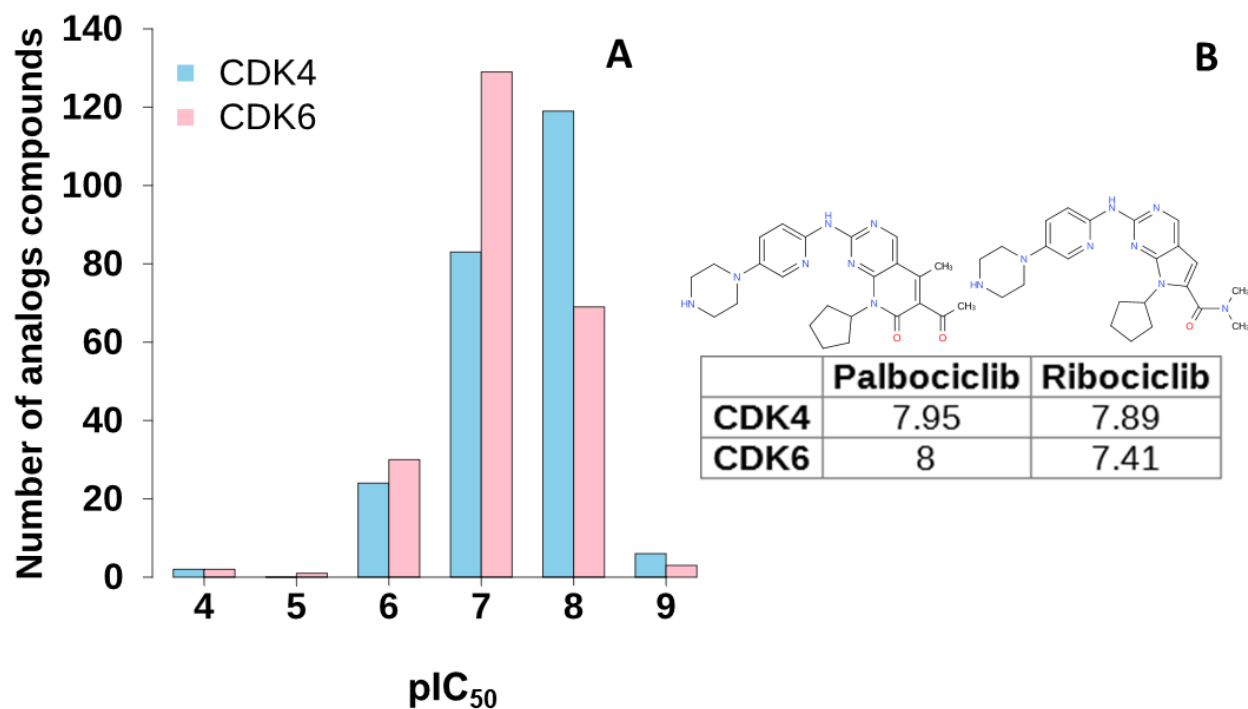


Figure 1. A. Distribution of pIC_{50} values of Palbociclib and Ribociclib analogs for CDK4 and CDK6. **B.** Chemical structures and pIC_{50} values of Palbociclib and Ribociclib drugs towards CDK4 and CDK6.

Three key physical chemical properties (e.g., Molecular Weight (MW), lipophilicity (AlogP) and Polar Surface Area (PSA)) were computed for these compounds (**Figure 2**). First, the mean of MW is $445 \pm 33.6 \text{ g.mol}^{-1}$ and 91% of analogs have a MW less than 500 g.mol^{-1} . As references, Palbociclib has a MW of $447.54 \text{ g.mol}^{-1}$ and Ribociclib has a MW of $434.55 \text{ g.mol}^{-1}$ (**Table S2**). Second, the mean of AlogP is 3.8 ± 0.8 and 95% have an AlogP less than 5. AlogP of Palbociclib and Ribociclib are equal to 2.97 and 2.80 respectively (**Table S2**). Thus, the vast majority of analogs respect those two Lipinski's Rule of Five for druglikeness (Lipinski et al., 2001). Concerning PSA, the dataset average is $94.4 \text{ \AA}^2 \pm 15.1$. For Palbociclib, it is equal to 105.04 \AA^2 and to 91.21 \AA^2 for Ribociclib. When PSA is greater than 140 \AA^2 , a compound tends to have a poor permeability (Veber et al., 2002). All analogs have a PSA between 58 and 151, thus, likely have a good oral bioavailability, with only two analogs not respecting the threshold of 140 \AA^2 .

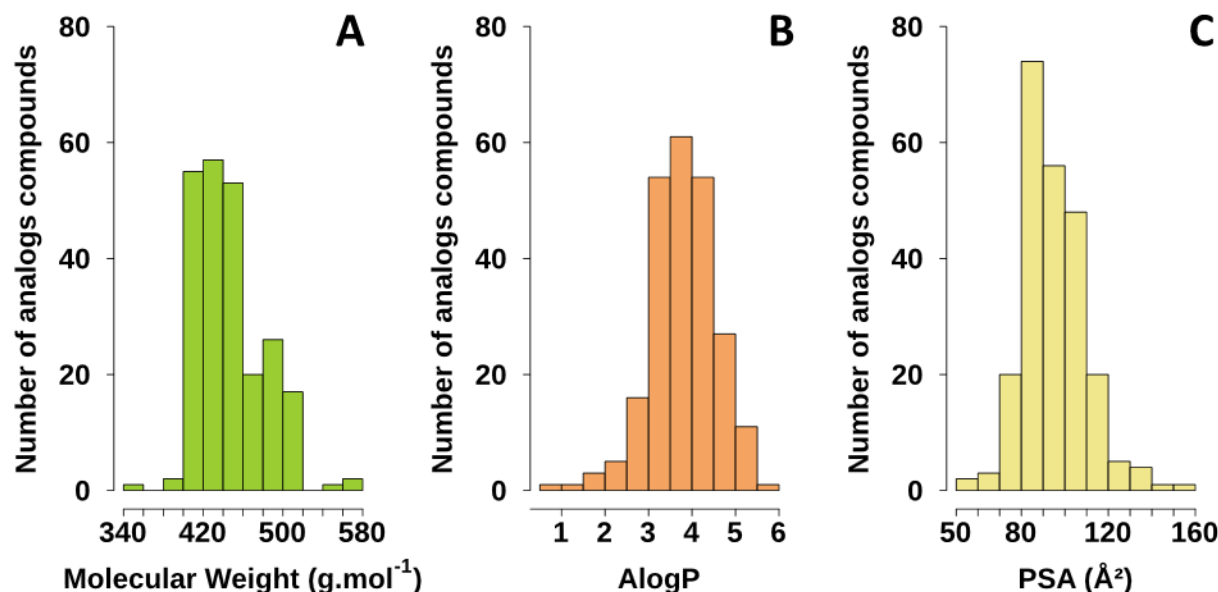


Figure 2: Distribution of molecular properties of Palbociclib and Ribociclib analogs: **A.** Molecular Weight, **B.** Hydrophobicity – AlogP, and **C.** Polar Surface Area – PSA.

3.2. CDK4 homology model and CDK6 Structures

Sequences of CDK4 (PDB code: 3G33 chain A) and CDK6 (PDB code: 2EUF chain B) were aligned using Prime and we found 67% sequence identity. Then we built a CDK4 homology model based on the CDK6 crystal structure (See Material and methods). The RMSD (Root Mean Square Deviation) between CDK4 model and the template CDK6 was equal to 0.14 Å. RMSD of CDK4 model on CDK4 crystal structure (PDB code: 3G33 chain A) is equal to 1.94 Å.

Binding site residues were identified according to the literature and represented in **Figure 3**. In total, 16 residues are the same for both CDK4 and CDK6, and three residues are different, i.e. Val19, Arg106 and Glu149 of CDK4 substituted respectively by Gly21, Thr109 and Gln152 in CDK6.

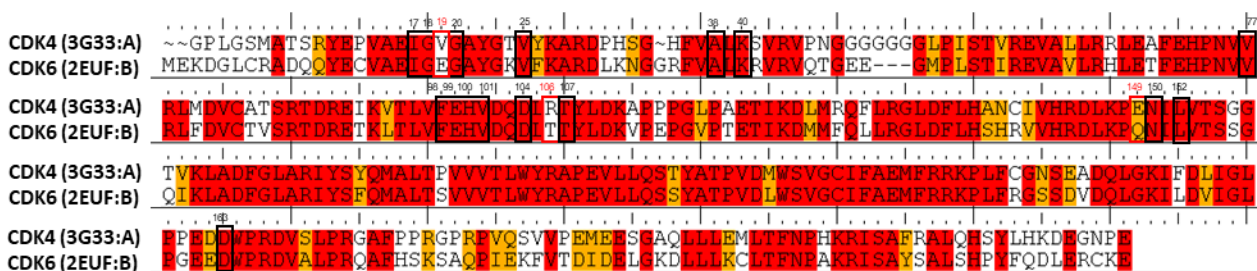


Figure 3: Alignment of CDK4 (PDB code: 3G33 chain A) and CDK6 (PDB code: 2EUF chain B) colored according to Residue Homology. The residue pairs are white when there is no homology, red when there is an exact match and yellow for similar residues (BLOSUM62 (Henikoff & Henikoff, 1992)). Binding sites residues are framed in black and numbered according to CDK4 numbering in PDB code: 3G33 chain A. Residues that are mutated between CDK4 and CDK6 are framed in red.

3.3. Molecular Docking

Molecular docking of all analogs was performed towards CDK4 and CDK6 structures (**Figure 4**). Importantly, we also performed molecular docking taking as receptor mutant and apo crystal structure of CDK4 (PDB codes: 2W96 and 3G33 respectively), and it led to fairly higher docking scores. This confirms the results previously reported by (Rondla et al., 2017), i.e., -5.8 kcal/mol for Palbociclib and -6.0 kcal/mol for Ribociclib. This could be due to the fact that apo binding pocket residues are not favorably oriented (**Figure S3**) to allow for ligand docking (*data not shown*).

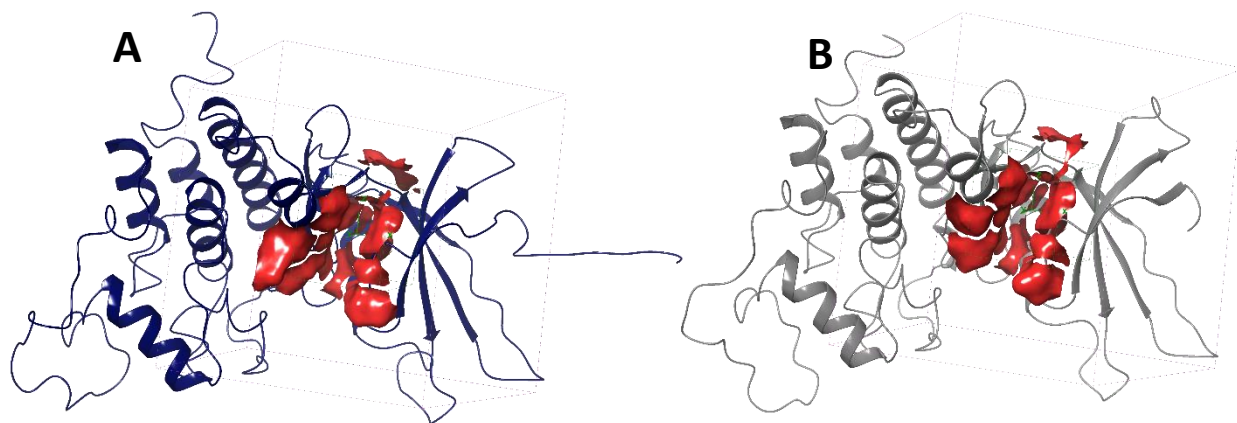


Figure 4: 3D structures of CDK4 homology model and CDK6. **A.** The homology model of CDK4 is based on PDB code: 3G33 chain A sequence, taking CDK6 (PDB code: 2EUF) as template structure. It is represented in dark blue cartoon. **B.** CDK6 structure is represented in grey cartoon. Binding site residues are represented in red surface. Residues were defined from Rondla et al. 2017. The docking grid box was generated based on these binding residues.

in a scatter plot colored according to pIC_{50} values (**Figure 6**). Docking results showed that there is an acceptable separation between very active and less active analogs. Indeed, **Figure 6** illustrates the most active analogs in light green which have the lowest DS and the lowest Emodel. In contrast, less active compounds (in orange and red) have high DS and high Emodel. According to **Figure 6**, SP DS match better with pIC_{50} values than XP DS. CHEMBL3702087 (**Figure S4**) is the top1 analog for CDK4 and CDK6 with a XP DS equal to -11.86 kcal/mol for CDK4 and -12.22 kcal/mol for CDK6, but its pIC_{50} values are respectively 7.80 and 6.95 (**Table S2**).

We particularly focused on complexes that afforded the best docking scores (DS) with SP method (**Figure 5**) to be selected for MD simulations. These compounds also achieved better experimental pIC_{50} than that of both Palbociclib and Ribociclib. The top-ranked compound for CDK4 (top1-CDK4) is CHEMBL3698681 (**Figure 9**) (DS=-11.13 kcal/mol, pIC_{50} for CDK4=8.70, pIC_{50} for CDK6=9.12) and the top ranked compound for CDK6 (top1-CDK6) is CHEMBL3702089 (**Figure 9**) (DS=-11.16 kcal/mol, pIC_{50} for CDK4=8.07, pIC_{50} for CDK6=7.51). We have also been interested in the fifth top-ranked compound with SP docking in

CDK6 (top5-CDK6), which is CHEMBL3698657 (**Figure 9**), because it affords a very high pIC₅₀ of 9 for CDK4 and 9.39 for CDK6.

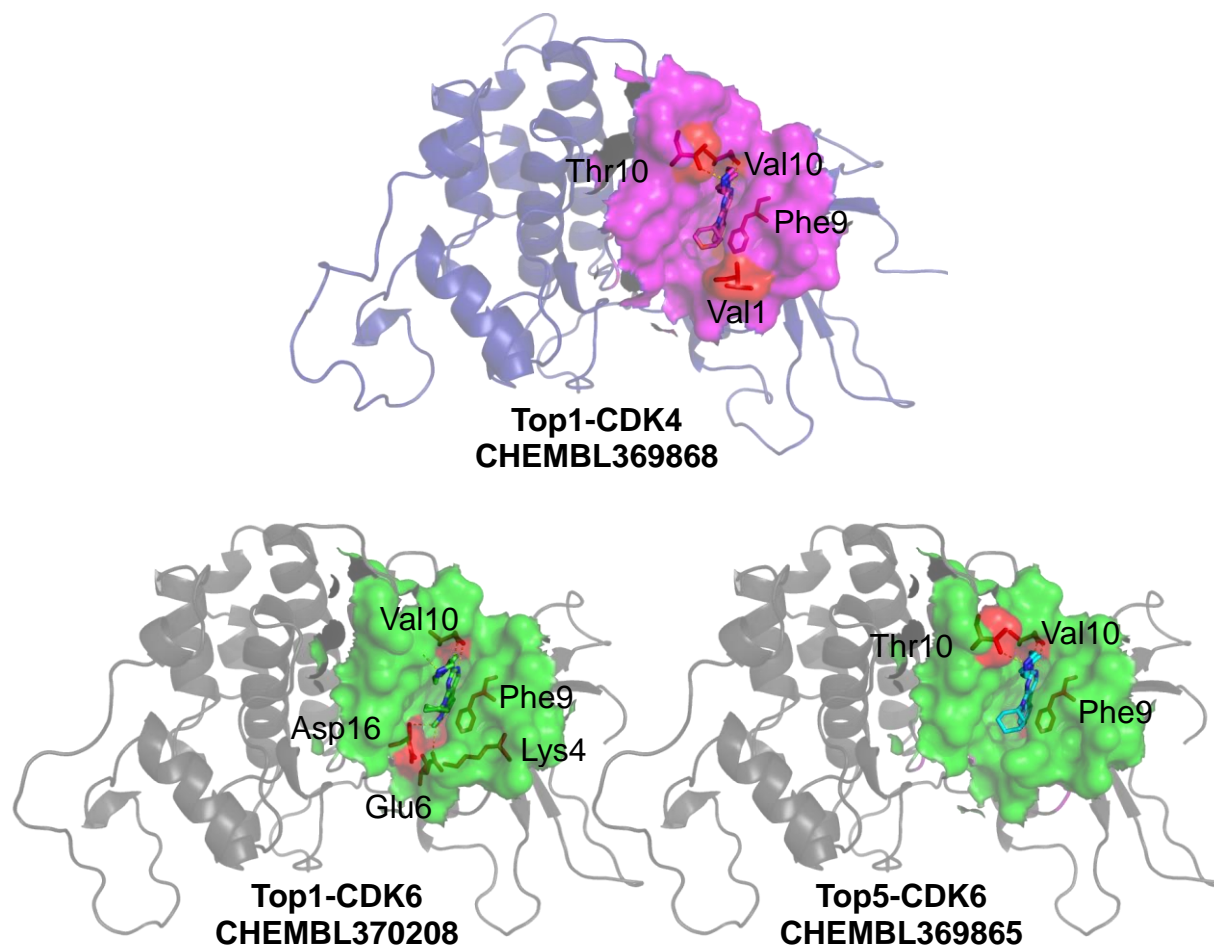


Figure 5. Docking poses of three interesting compounds. **Top1-CDK4:** the protein is represented in dark blue cartoon, ligand in magenta stick, and pocket residues are represented in magenta surface. **Top1-CDK6:** the protein is represented in grey cartoon, ligand in green stick, and pocket residues are represented in green surface. **Top5-CDK6:** the protein is represented in grey cartoon, ligand in cyan stick, and pocket residues are represented in green surface. Residues involved in interactions are represented in red stick.

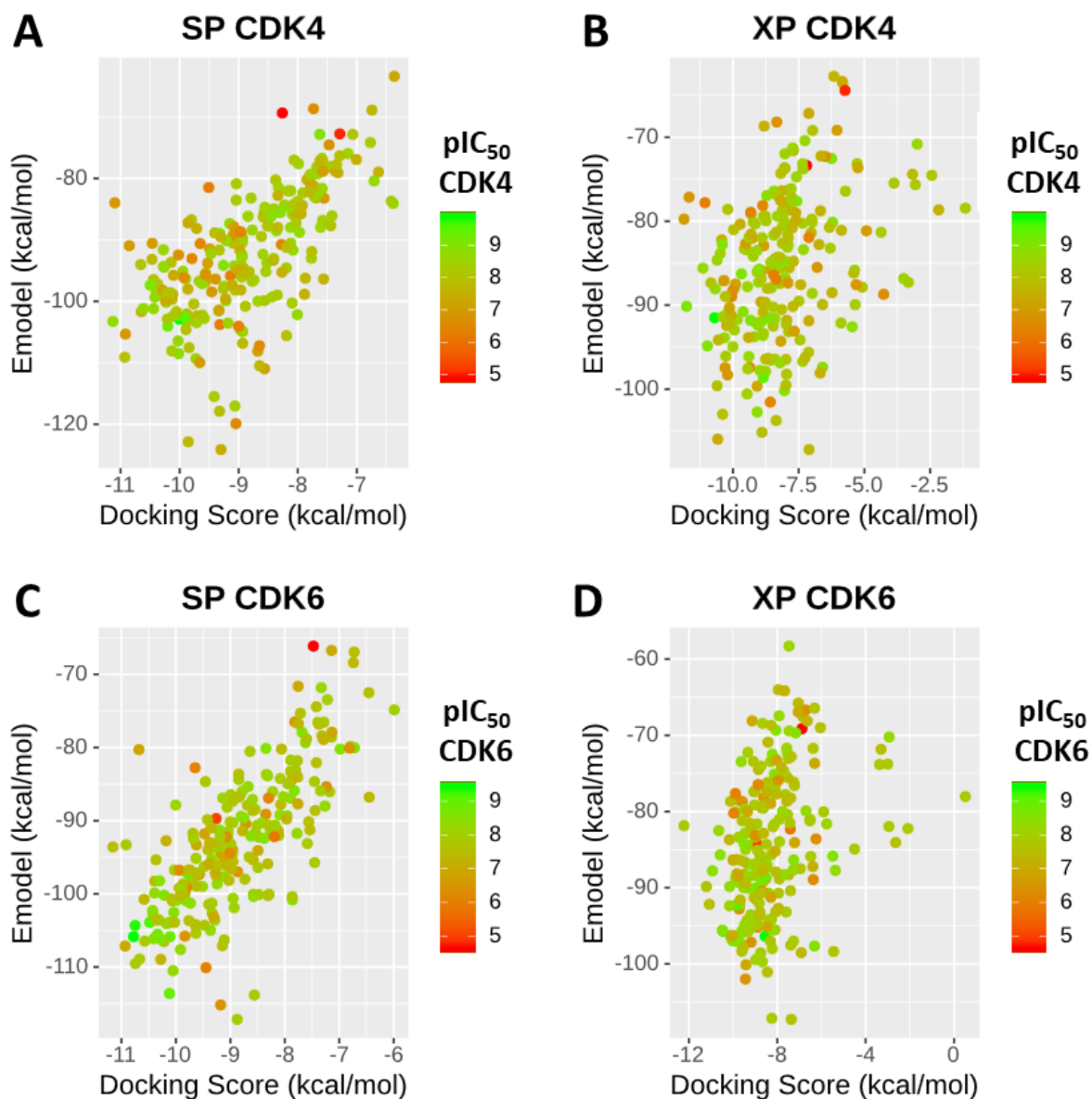


Figure 6: Scatter plots of Docking Score versus Emodel colored according to pIC_{50} values. **A.** SP Docking Score for CDK4 homology model. **B.** XP Docking Score for CDK4 homology model. **C.** SP Docking Score for CDK6. **D.** XP Docking Score for CDK6.

3.4. RDKit descriptors and hierarchical clustering analysis

Physicochemical 2D descriptors were calculated with RDKit for the 234 analogs. Then, a hierarchical clustering was performed based on these descriptors, a Euclidean-based distance, and a Ward-linkage. Results were represented in a circular dendrogram and colored according to pIC_{50} values for CDK4 and CDK6 (**Figure 7**). In **Figure 9**, we added layers for most relevant properties. In **Figure 8**, we added layers that represent DS with SP and XP models to visualize DS with pIC_{50} for each compound. We identified four main clusters that we analyzed cluster by cluster.

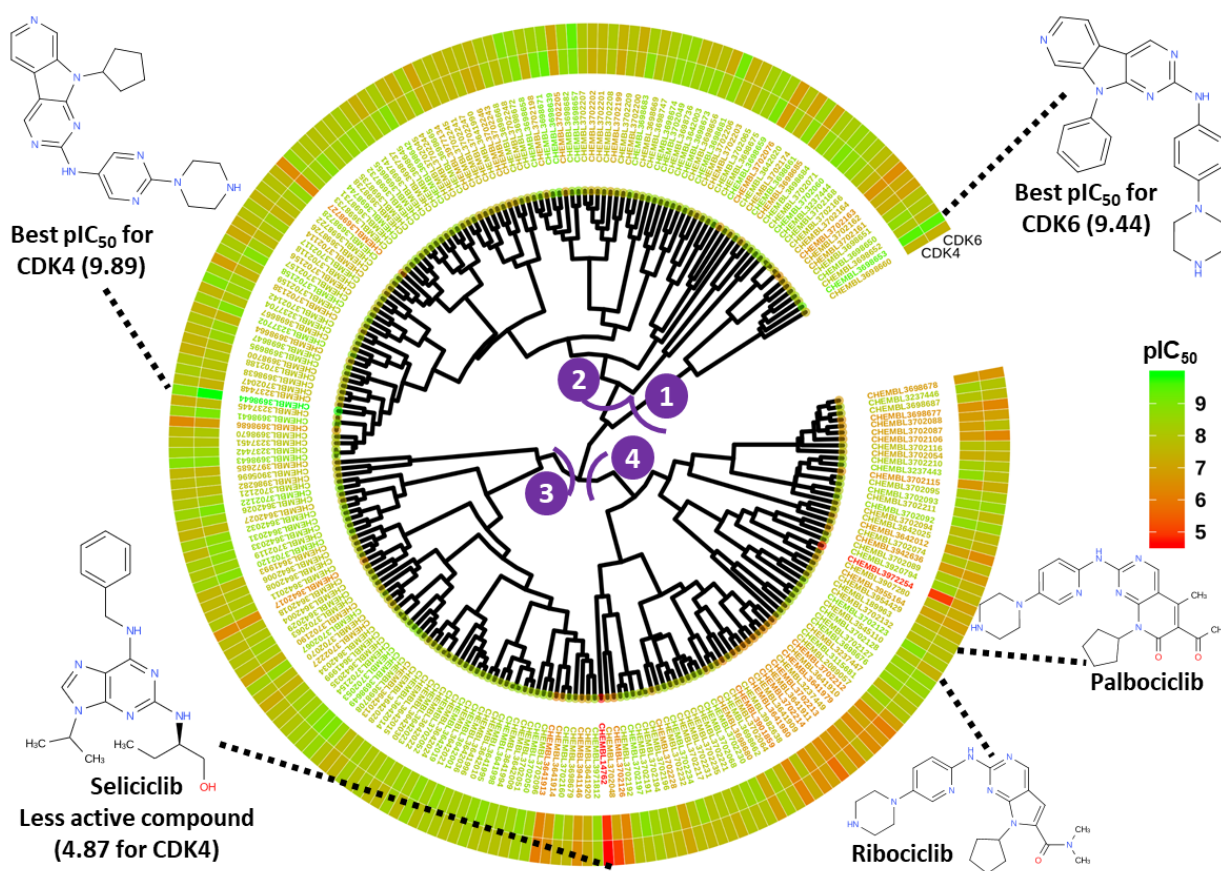


Figure 7: Hierarchical clustering of Palbociclib and Ribociclib analogs based on 2D descriptors generated with RDKit. Euclidean distance and ward linkage were used. Ligands are colored by pIC_{50} of CDK4 (high: green, low: red). Cells around the circular dendrogram corresponding to each ligand are colored by pIC_{50} of CDK4 and pIC_{50} of CDK6. The number of the studied clusters are represented in purple circles.

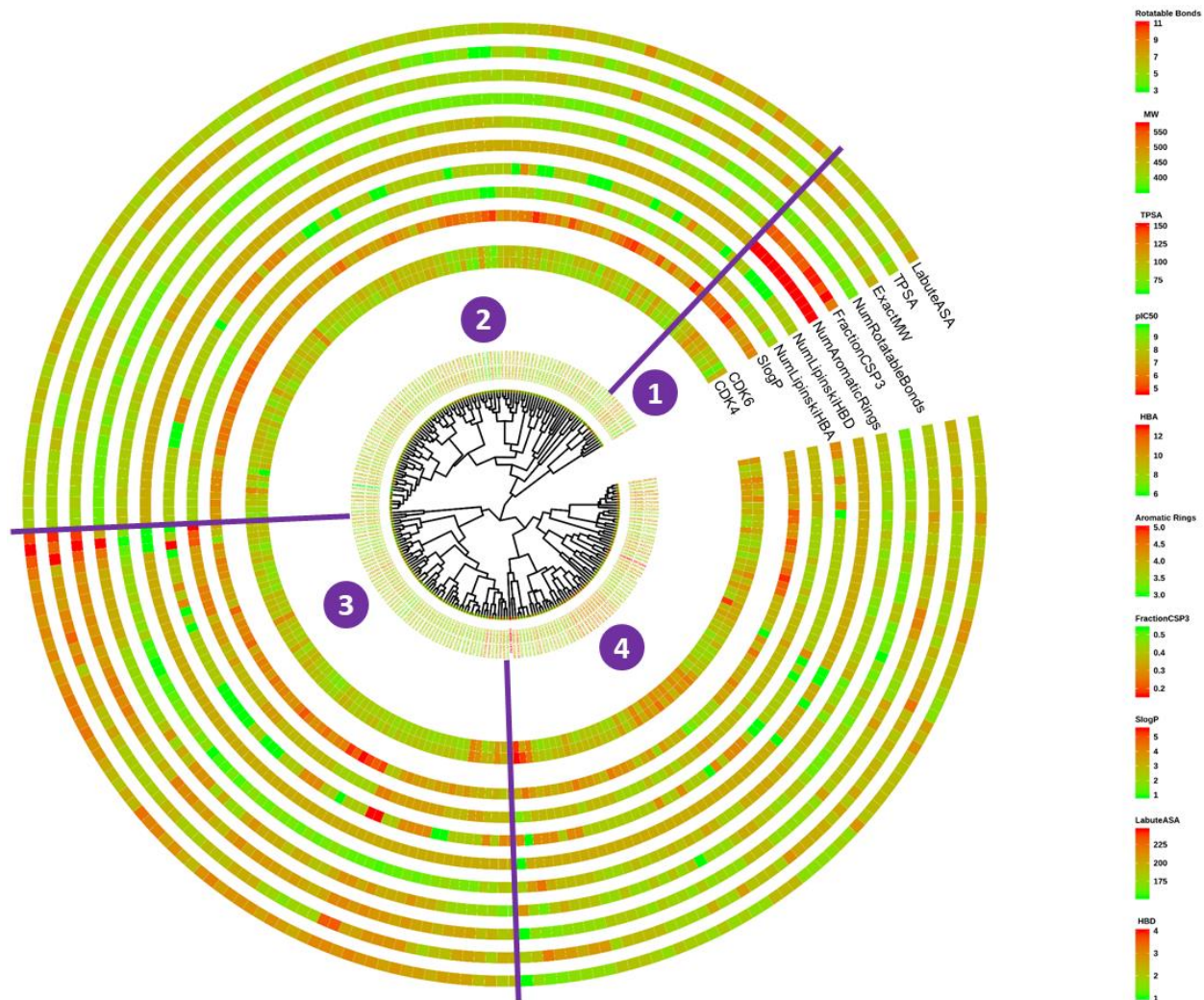


Figure 8: Hierarchical clustering of Palbociclib and Ribociclib analogs based on 2D descriptors generated with RDKit. Ligands are colored by pIC₅₀ of CDK4 (high: green, low: red). Cells around the circular dendrogram corresponding to each ligand are colored by pIC₅₀ of CDK4 and pIC₅₀ of CDK6. Cells around the pIC₅₀ values correspond to selected physicochemical properties.

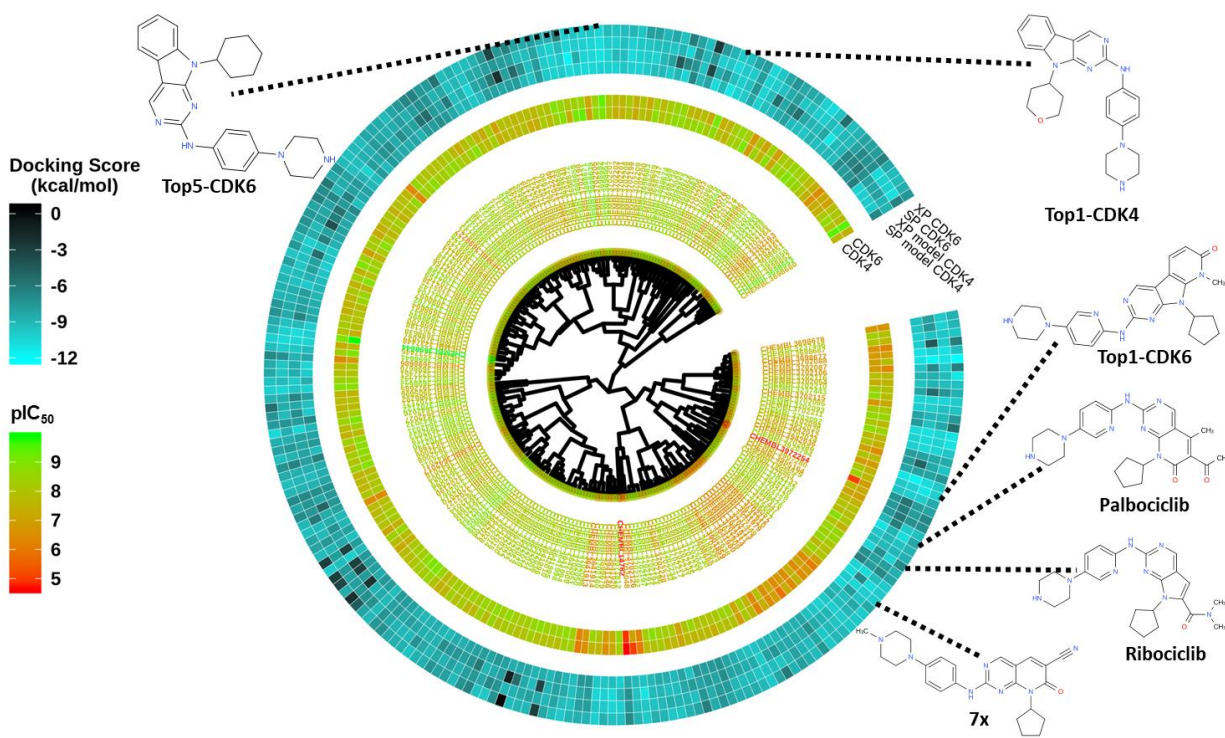


Figure 9: Hierarchical clustering of Palbociclib and Ribociclib analogs based on 2D descriptors generated with RDKit. Ligands are colored by pIC_{50} of CDK4. Cells around the circular dendrogram corresponding to each ligand are colored by pIC_{50} of CDK4 and pIC_{50} of CDK6. Cells around the pIC_{50} values correspond to Docking scores (SP and XP) in CDK4 and CDK6 (high: black, low: cyan).

Clusters 1 and 2 contain the most active compounds with highest pIC_{50} .

Cluster 1 is characterized by including CHEMBL3698644 with the best pIC_{50} for CDK6 (9.44), which is also the second best for CDK4 (9.49). Moreover, all compounds of this cluster are characterized by having the lowest fraction of sp^3 carbons but also the highest number of aromatic rings (5), which are inversely related (Ritchie & Macdonald, 2014).

Cluster 2 contains the most favorable analogs regarding their physicochemical properties, i.e. HBA (Hydrogen Bond Acceptors) less than 10, HBD (Hydrogen Bond Donors) less than 5, MW less than 500 g.mol^{-1} (Lipinski et al., 2001) and TPSA less than 100 \AA^2 . All compounds of this cluster have 4 aromatic rings. It contains the most active compound in CDK4 (CHEMBL3698653) with a pIC_{50} of 9.88. It also contains the promising compound top5-CDK6

which has a better docking score than Palbociclib and Ribociclib drugs (**Table S2**). This cluster contains also top1-CDK4 (CHEMBL3698681).

Cluster 3 is characterized by having compounds with the highest MW, and high LabuteASA and also high PSA. LabuteASA is the approximative van Der Waals Surface Area of an atom in a compound defined by Labute as the amount of surface area of that atom not contained in any other atom of the molecule (Labute, 2004). Analogs of this cluster are also characterized by a high number of rotatable bonds.

Cluster 4 contains both Palbociclib (CHEMBL189963) and Ribociclib (CHEMBL3545110). This cluster also contains the top1-CDK6 (CHEMBL3702089). This compound affords better pIC₅₀ than the FDA approved drugs, and it has similar values of physicochemical properties (**Table S1**).

Two compounds present a little effect on CDK4/6 with a pIC₅₀ less than 5, i.e. CHEMBL14762 and CHEMBL3972254. Amongst them there is Seliciclib (CHEMBL14762), a drug candidate in phase 2. This shows that even if Seliciclib is a structural analog of Palbociclib and Ribociclib, it is not targeting CDK4/6. Indeed, Seliciclib inhibits CDK2/E, CDK2/A, CDK7 and CDK9 with much higher pIC₅₀ than for CDK4/6, e.g. the pIC₅₀ is equal to 7 for CDK2/E and equal to 4.87 for CDK4 (Heathcote et al., 2010).

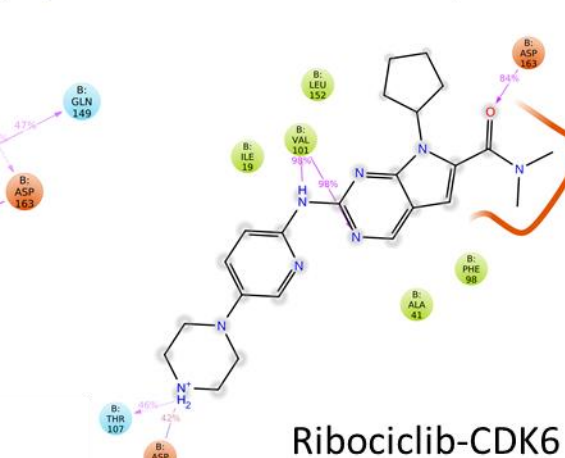
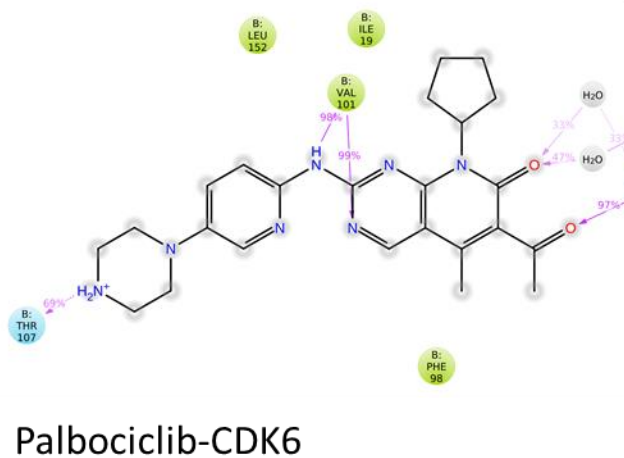
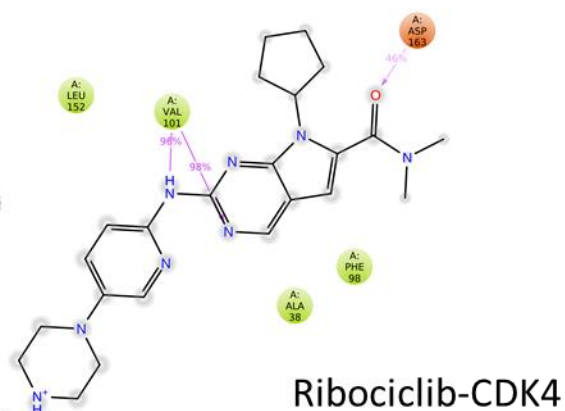
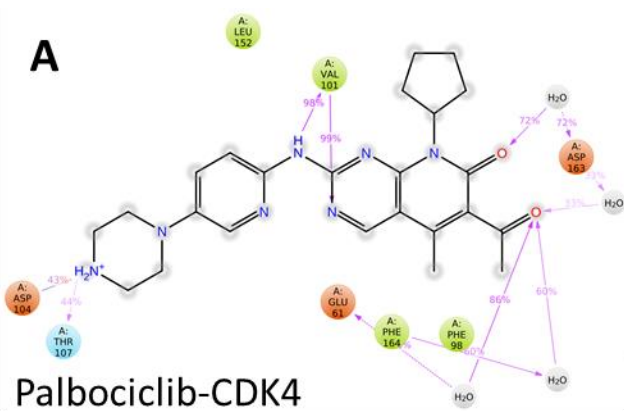
Cluster 4 contains top5-CDK6. It has better pIC₅₀ for both CDK4 and CDK6 than Palbociclib and Ribociclib. It also has better physicochemical properties for TPSA, MW and number of HBA. Top1-CDK4 is in this cluster and it has better pIC₅₀ in both CDK4 and CDK6 than Palbociclib and Ribociclib, it has better physicochemical properties for TPSA, MW and number of HBA (**Table S1**).

We were interested in the CHEMBL3115681 compound, described by Reddy et al. as a potent inhibitor of CDK4/6 (Reddy et al., 2014). This compound belongs to cluster 4 (**Figure 9**) and achieves a better pIC_{50} than the FDA approved drugs with favorable physicochemical properties (**Table S1**). However, this analog is not well ranked compared to Palbociclib and Ribociclib despite the good docking scores (**Table S2**).

3.5. Molecular Dynamics Simulations

MD simulations were performed for top-ranked complexes with SP docking and for Palbociclib and Ribociclib complexed with CDK4 and CDK6. We analyzed their protein-ligand interactions illustrated in **Figure 10** and reported in **Table 1**. For CDK4 complex, residues that interact directly with Palbociclib are Val101, Asp104, and Thr107. We note that Glu61, Asp163, and Phe164 interact via a water bridge. With Ribociclib, residues are Val101 and Asp163. Residues that interact with Palbociclib in CDK6 are Thr107, Val101, Gln149, and Asp163. For Ribociclib in CDK6, residues involved in interactions are Thr107, Val101, Asp163, and Asp104. Top1-CDK6 has key interactions with Val101, Asp163, Glu61 and Phe98, whereas Top5-CDK6 has interactions with Thr107, Val101, and Phe98. Residues involved in interactions with top1-CDK4 are Thr107, Val101, Phe98, and Val19.

A



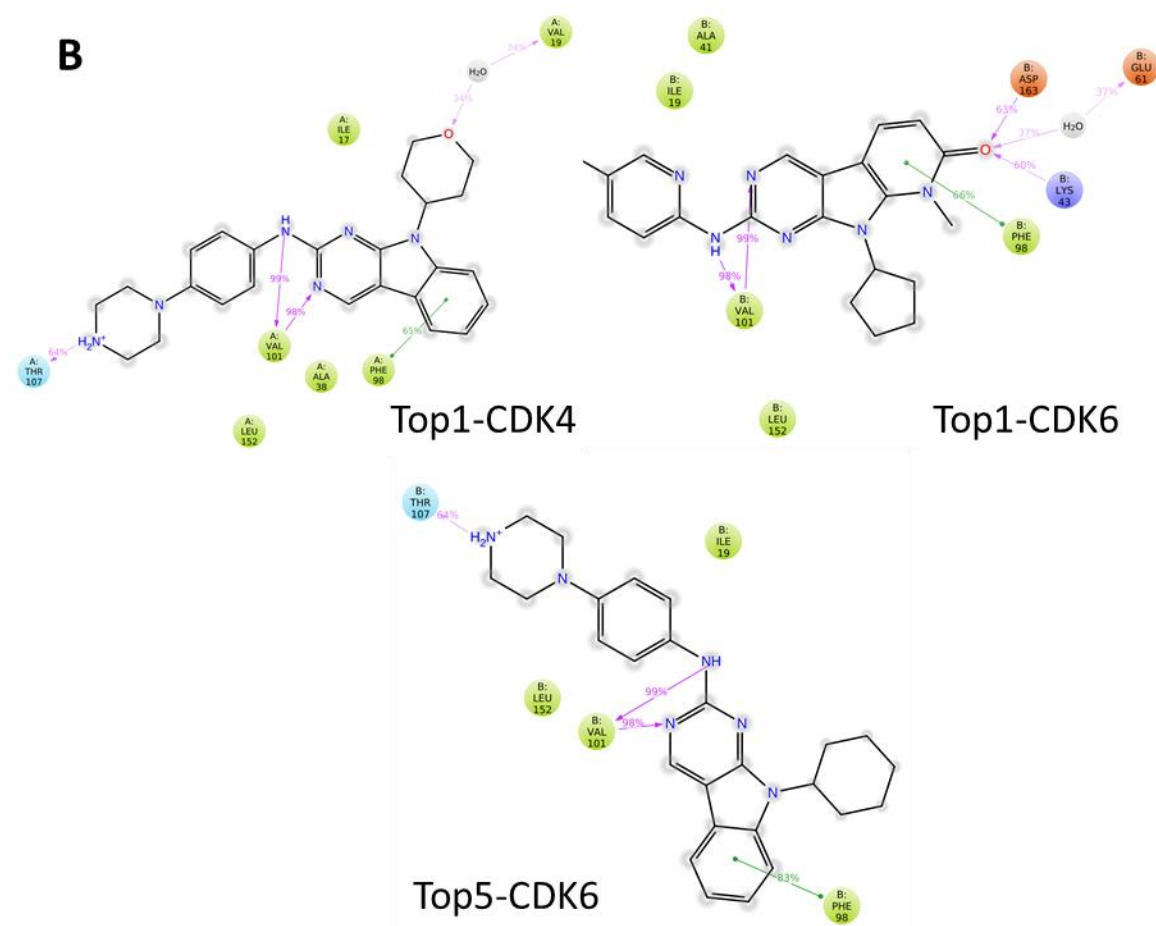


Figure 10: Protein-ligand interactions of selected compounds generated with Desmond. **A.** FDA approved drugs. **B.** Interesting compounds resulted from docking. Interactions that occur more than 30.0% of the simulation time in the selected trajectory (0 through 50 ns), are shown.

Residues	Palbociclib- CDK4	Ribociclib- CDK4	Palbociclib- CDK6	Ribociclib- CDK6	Top1-CDK4 CHEMBL3698681	Top1-CDK6 CHEMBL3702089	Top5-CDK6 CHEMBL3698657
Val19							
Lys43							
Glu61							
Phe98							
Val101							
Asp104							
Thr107							
Gln149							
Asp163							
Phe164							

Hbond
Hbond via H ₂ O
π interaction
No interaction

Table 1. Recapitulative summary of the residues involved in interactions for Palbociclib, Ribociclib and top ranked analogs complexed with CDK4 and CDK6.

These results show that selected analogs share similar interactions as the two FDA approved drugs, such as Val101 involved in all complexes, or Thr107 present in all complexes except Ribociclib-CDK4 and top1-CDK6. Interestingly, π -interaction involving Phe98 is absent in Palbociclib and Ribociclib with both CDK4 and CDK6, and this interaction is specific to all other analogs reported in **Table 1**. This dynamic interaction seems to lead to a better pIC₅₀, as these selected analogs all achieve higher inhibition activity than Palbociclib and Ribociclib. Residue Asp163 is involved in dynamic interactions for all complexes with FDA-approved drugs and with top1-CDK6 (CHEMBL3702089) (**Figure 10**); however, we note that for Palbociclib this residue interacts via a water bridge, contrary to Ribociclib and top1-CDK6 with the direct hydrogen bond. This difference has minor effect on the solubility of Palbociclib and Ribociclib (**Table S1**). Our results also showed that there is no specific dynamic interaction that clearly leads to more specificity toward CDK4 or toward CDK6. The design of a new CDK6 inhibitor with high selectivity is thus highly challenging, even though CDK6 has unique functions that are specific to certain cell types and have distinct development mode (Tadesse et al., 2015).

Overall, analogs CHEMBL3702089, CHEMBL3698657 and CHEMBL3698681 could be of high interest for additional experimental investigations, considering their high inhibitory activity, favorable physicochemical properties and their good affinity towards CDK4 and CDK6. This study could also be conducted in the same way to analyze analogs of the FDA approved drug Abemaciclib and to characterize their properties. Its inhibitory activity (pIC_{50}) is of 8.70 for CDK4 and 8.30 for CDK6 (Poratti & Marzaro, 2019). Abemaciclib is not an analog of Palbociclib and Ribociclib at 70% of similarity but could represent an interesting candidate for an extended cheminformatics study (Hamilton & Infante, 2016; Pernas et al., 2018).

4. Conclusions

Our cheminformatics study allowed us to characterize and describe a large set of analogs of Palbociclib and Ribociclib FDA approved drugs. We combined these results with molecular docking and molecular dynamics simulations to identify the key dynamic CDK-inhibitor interactions that could help for future molecular design.

Acknowledgement

DT acknowledges A. Borrel, M. Kuenemann, PP. Kyaw Zin, X. Li, G. Van Den Driessche, J. Ash and K. Song for sharing their knowledge in ggplot2 package from R and Schrödinger suite.

References

- Aldeghi, M., Malhotra, S., David, L., Wing, A., & Chan, E. (2014). Two- and Three-dimensional Rings in Drugs. *Chemical Biology and Drug Design*, 83, 450–461. <https://doi.org/10.1111/cbdd.12260>
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235–242. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=102472&tool=pmcentrez&rendertype=abstract>
- Berthold, M. R., Cebron, N., Dill, F., Gabriel, T. R., Kotter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K., & Wiswedel, B. (2008). KNIME : The Konstanz Information Miner. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data Analysis, Machine Learning and Applications* (Springer, p. 319:326).
- Bowers, K. J., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., & Shaw, D. E. (2006). *Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters*.
- Cadoo, K. A., Gucalp, A., & Traina, T. A. (2014). Palbociclib : an evidence-based review of its potential in the treatment of breast cancer. *Breast Cancer: Targets and Therapy*, 20(4–6), 123–133.
- Chen, X., Xu, D., Li, X., Zhang, J., Xu, W., Hou, J., Zhang, W., & Tang, J. (2019). Latest Overview of the Cyclin-Dependent Kinases 4 / 6 Inhibitors in Breast Cancer : The Past , the Present and the Future. *Journal of Cancer*, 10(26), 6608–6617. <https://doi.org/10.7150/jca.33079>
- Connors, Richard, V., Dai, K., Eksterowicz, J., Fan, P., Fisher, B., Fu, J., Li, K., Li, Z., Mcgee, L., Sharma, R., Wang, X., Mcminn, D., Mihalic, J., & Deignan, J. (2009). *Fused Pyridine, Pyrimidine and Triazine Compounds as cell cycle inhibitors* (Patent No. WO 2009/085185 A1).
- Davies, M., Nowotka, M., Papadatos, G., Dedman, N., Gaulton, A., Atkinson, F., Bellis, L., & Overington, J. P. (2015). ChEMBL web services : streamlining access to drug discovery data and utilities. *Nucleic Acids Research*, 43, W612–W620. <https://doi.org/10.1093/nar/gkv352>

- Day, P. J., Cleasby, A., Tickle, I. J., Reilly, M. O., Coyle, J. E., Holding, F. P., Mcmenamin, R. L., Yon, J., Chopra, R., Lengauer, C., & Jhoti, H. (2009). Crystal structure of human CDK4 in complex with a D-type cyclin. *Proc Natl Acad Sci USA*, 106(11), 4166–4170. <https://doi.org/10.1073/pnas.0809645106>
- Desmond Molecular Dynamics System, D. E. Shaw Research, New York, NY. (2018).
- Egan, W. J., Merz, K. M., & Baldwin, J. J. (2000). *Prediction of Drug Absorption Using Multivariate Statistics*. 3867–3877. <https://doi.org/10.1021/jm000292e>
- Fourches, D., Muratov, E., & Tropsha, A. (2010). Trust , But Verify : On the Importance of Chemical Structure Curation in Cheminformatics and QSAR Modeling Research. *Journal of Chemical Information and Modeling*, 50(7), 1189–1204.
- Fourches, D., Muratov, E., & Tropsha, A. (2016). Trust , but Verify II : A Practical Guide to Chemogenomics Data Curation. *Journal of Chemical Information and Modeling*, 56(7), 1243–1252. <https://doi.org/10.1021/acs.jcim.6b00129>
- Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., & Shenkin, P. S. (2004). Glide : A New Approach for Rapid , Accurate Docking and Scoring . 1 . Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry*, 47, 1739–1749.
- Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C., & Mainz, D. T. (2006). Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein - Ligand Complexes. *Journal of Medicinal Chemistry*, 49(21), 6177–6196.
- Gaulton, A., Hersey, A., Patr, A., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F., Bellis, L. J., Cibri, E., Davies, M., Dedman, N., Karlsson, A., Magari, P., Overington, J. P., Papadatos, G., & Smit, I. (2017). The ChEMBL database in 2017. *Nucleic Acids Research*, 45, D945–D954. <https://doi.org/10.1093/nar/gkw1074>

- Ghose, A. K., Viswanadhan, V. N., & Wendoloski, J. J. (1998). Prediction of Hydrophobic (Lipophilic) Properties of Small Organic Molecules Using Fragmental Methods : An Analysis of ALOGP and CLOGP Methods. *Journal of Physical Chemistry A*, 102(21), 3762–3772. <https://doi.org/10.1021/jp980230o>
- Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide : A New Approach for Rapid , Accurate Docking and Scoring . 2 . Enrichment Factors in Database Screening. *Journal of Medicinal Chemistry*, 47(7), 1750–1759.
- Hamilton, E., & Infante, J. R. (2016). Targeting CDK4 / 6 in patients with cancer. *Cancer Treatment Reviews*, 45, 129–138. <https://doi.org/10.1016/j.ctrv.2016.03.002>
- Harder, E., Damm, W., Maple, J., Wu, C., Reboul, M., Xiang, J. Y., Wang, L., Lupyan, D., Dahlgren, M. K., Knight, J. L., Kaus, J. W., Cerutti, D. S., Krilov, G., Jorgensen, W. L., Abel, R., & Friesner, R. A. (2016). OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation*, 12(1), 281–296. <https://doi.org/10.1021/acs.jctc.5b00864>
- Heathcote, D. A., Patel, H., Kroll, S. H. B., Hazel, P., Periyasamy, M., Alikian, M., Kanneganti, S. K., Jogalekar, A. S., Scheiper, B., Barbazanges, M., Blum, A., Brackow, J., Siwicka, A., Pace, R. D. M., Fuchter, M. J., Snyder, J. P., Liotta, D. C., Freemont, P. S., Aboagye, E. O., ... Ali, S. (2010). A Novel Pyrazolo [1,5-a] pyrimidine Is a Potent Inhibitor of Cyclin-Dependent Protein Kinases 1, 2, and 9, Which Demonstrates Antitumor Effects in Human Tumor Xenografts Following Oral Administration. *Journal of Medicinal Chemistry*, 53(24), 8508–8522. <https://doi.org/10.1021/jm100732t>
- Henikoff, S., & Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of National Academy of Sciences USA*, 89, 10915–10919.
- Iwata, H. (2018). Clinical development of CDK4 / 6 inhibitor for breast cancer. *Breast Cancer*, 25(4), 402–406. <https://doi.org/10.1007/s12282-017-0827-3>
- Jacobson, M. P., Friesner, R. A., Xiang, Z., & Honig, B. (2002). On the role of the crystal environment in determining protein side-chain conformations. *Journal of Molecular Biology*, 320(3), 597–608.

[https://doi.org/10.1016/S0022-2836\(02\)00470-9](https://doi.org/10.1016/S0022-2836(02)00470-9)

Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., Honig, B., Shaw, D. E., & Friesner, R. A. (2004).

A Hierarchical Approach to All-Atom Protein Loop Prediction. *Proteins: Structure, Function and Genetics*, 55(2), 351–367. <https://doi.org/10.1002/prot.10613>

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45), 11225–11236. <https://doi.org/10.1021/ja9621760>

Labute, P. (2004). Derivation and Applications of Molecular Descriptors Based on Approximate Surface Area. *Chemoinformatics*, 275, 261:278.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Advanced Drug Delivery Reviews*, 46(1–3), 3–26. [https://doi.org/10.1016/s0169-409x\(00\)00129-0](https://doi.org/10.1016/s0169-409x(00)00129-0)

Liu, M., Liu, H., & Chen, J. U. N. (2018). Mechanisms of the CDK4 / 6 inhibitor palbociclib (PD 0332991) and its future application in cancer treatment (Review). *Oncology Reports*, 39, 901–911. <https://doi.org/10.3892/or.2018.6221>

O’Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An Open chemical toolbox. *Journal of Cheminformatics*, 3(10), 33. <https://doi.org/10.1186/1758-2946-3-33>

Pandey, K., An, H., Kim, S. K., Lee, S. A., Kim, S., Lim, S. M., & Kim, G. M. (2019). Molecular mechanisms of resistance to CDK4/6 inhibitors in breast cancer: A review. *International Journal of Cancer*, 145(5), 1179–1188. <https://doi.org/10.1002/ijc.32020>

Pernas, S., Tolaney, S. M., Winer, E. P., & Goel, S. (2018). CDK4/6 inhibition in breast cancer : current practice and future directions. *Therapeutic Advances in Medical Oncology*, 10, 1–15. <https://doi.org/10.1177/https>

Petrelli, F., Ghidini, A., Pedersini, R., Cabiddu, M., Borgonovo, K., Parati, M. C., Ghilardi, M., Amoroso, V., Berruti, A., & Barni, S. (2019). Comparative efficacy of palbociclib , ribociclib and abemaciclib

- for ER + metastatic breast cancer : an adjusted indirect analysis of randomized controlled trials. *Breast Cancer Research and Treatment*, 174(3), 597–604. <https://doi.org/10.1007/s10549-019-05133-y>
- Polêto, M. D., Rusu, V. H., Grisci, B. I., Dorn, M., Lins, R. D., & Verli, H. (2018). Aromatic Rings Commonly Used in Medicinal Chemistry : Force Fields Comparison and Interactions With Water Toward the Design of New Chemical Entities. *Frontiers in Pharmacology*, 9(395), 1–20. <https://doi.org/10.3389/fphar.2018.00395>
- Poratti, M., & Marzaro, G. (2019). Third-generation CDK inhibitors : A review on the synthesis and binding modes of Palbociclib, Ribociclib and Abemaciclib. *European Journal of Medicinal Chemistry*, 172, 143–153. <https://doi.org/10.1016/j.ejmech.2019.03.064>
- Portman, N., Alexandrou, S., Carson, E., Wang, S., Lim, E., & Caldon, C. E. (2013). Overcoming CDK4/6 inhibitor resistance in ER-positive breast cancer. *Endocrine-Related Cancer*, 26(1), R15–R30. <https://doi.org/10.1530/ERC-18-0317>
- R Development Core Team, in *R: A language and environment for statistical computing*, Vienna, Austria, URL: [http:// www.R-project.org/](http://www.R-project.org/). (2013).
- Reddy, M. V. R., Akula, B., Cosenza, S. C., Athuluridivakar, S., Mallireddigari, M. R., Pallela, V. R., Billa, V. K., Subbaiah, D. R. C. V., Bharathi, E. V., Carpio, R. V., Padgaonkar, A., Baker, S. J., & Reddy, E. P. (2014). Discovery of 8-Cyclopentyl-2-[4-(4-methyl-piperazin-1-yl)- phenylamino]-7-oxo-7,8-dihydro-pyrido[2,3-d]pyrimidine-6- carbonitrile (7x) as a Potent Inhibitor of Cyclin-Dependent Kinase 4 (CDK4) and AMPK-Related Kinase 5 (ARK5). *Journal of Medicinal Chemistry*, 57, 578–599.
- Ritchie, T. J., & Macdonald, S. J. F. (2014). Physicochemical Descriptors of Aromatic Character and Their Use in Drug Discovery. *Journal of Medical Virology*, 57, 7206–7215. <https://doi.org/10.1021/jm500515d>
- Rondla, R., Souda, L., & Ramatenki, V. (2017). Selective ATP competitive leads of CDK4 : Discovery by 3D-QSAR pharmacophore mapping and molecular docking approach. *Computational Biology and Chemistry*, 71, 224–229. <https://doi.org/10.1016/j.compbiolchem.2017.11.005>

- Roughley, S. D., & Jordan, A. M. (2011). The Medicinal Chemist's Toolbox : An Analysis of Reactions Used in the Pursuit of Drug Candidates. *Journal of Medicinal Chemistry*, 54, 3451–3479. <https://doi.org/10.1021/jm200187y>
- Schrödinger, LLC, New York. (2019).
- Sherr, C. J., Beach, D., & Shapiro, G. I. (2016). Targeting CDK4 and CDK6: From Discovery to Therapy. *Cancer Discovery*, 6(4), 353–368. <https://doi.org/10.1158/2159-8290.CD-15-0894>
- Sobhani, N., D'Angelo, A., Pittacolo, M., Roviello, G., Miccoli, A., Corona, S. P., Bernocchi, O., Generali, D., & Otto, T. (2019). Updates on the CDK4/6 Inhibitory Strategy and Combinations in Breast Cancer. *Cells*, 8(4), E321. <https://doi.org/10.3390/cells8040321>
- Tadesse, S., Yu, M., Kumarasiri, M., Le, B. T., Tadesse, S., Yu, M., Kumarasiri, M., Le, B. T., & Wang, S. (2015). Targeting CDK6 in cancer : State of the art and new insights. *Cell Cycle*, 14(20), 3220–3230. <https://doi.org/10.1080/15384101.2015.1084445>
- Takaki, T., Echaliier, A., Brown, N. R., Hunt, T., Endicott, J. A., & Noble, M. E. M. (2009). The structure of CDK4/cyclin D3 has implications for models of CDK activation. *Proc Natl Acad Sci USA*, 106(11), 4171–4176. <https://doi.org/https://doi.org/10.1073/pnas.0809674106>
- Tirado-rives, J., & Jorgensen, W. L. (1988). The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *Journal of the American Chemical Society*, 110(4), 1510–1657.
- Vanarsdale, T., Boshoff, C., Arndt, K. T., & Abraham, R. T. (2015). Molecular Pathways: Targeting the Cyclin D – CDK4/6 Axis for Cancer Treatment. *Clinical Cancer Research*, 21(13), 2905–2910. <https://doi.org/10.1158/1078-0432.CCR-14-0816>
- Veber, D. F., Johnson, S. R., Cheng, H., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *Journal of Medicinal Chemistry*, 45(12), 2615–2623. <https://doi.org/10.1021/jm020017n>
- Wang, P., Huang, J., Wang, K., & Gu, Y. (2016). New palbociclib analogues modified at the terminal piperazine ring and their anticancer activities. *European Journal of Medicinal Chemistry*, 122, 546–

556. <https://doi.org/10.1016/j.ejmech.2016.07.020>

Ward, S. E., & Beswick, P. (2014). What does the aromatic ring number mean for drug design? *Expert Opin. Drug. Discov.*, 9(9), 995–1003.

Wildman, S. A., & Crippen, G. M. (1999). Prediction of Physicochemical Parameters by Atomic Contributions. *Journal of Chemical Information and Computer Sciences*, 39(5), 868–873.
<https://doi.org/10.1021/ci990307l>