**Targeting Cyclin-Dependent Kinase 2:**

**Molecular Docking Studies of Roscovitine**

**1-** Cyclin-dependent kinases (CDKs) are pivotal enzymes in regulating the cell cycle, orchestrating the phosphorylation of specific substrates upon activation by cyclin binding. These interactions ensure proper cell division and proliferation. However, dysregulation of CDKs has been closely linked to the development of cancers and other proliferative disorders, highlighting their significance as critical therapeutic targets.

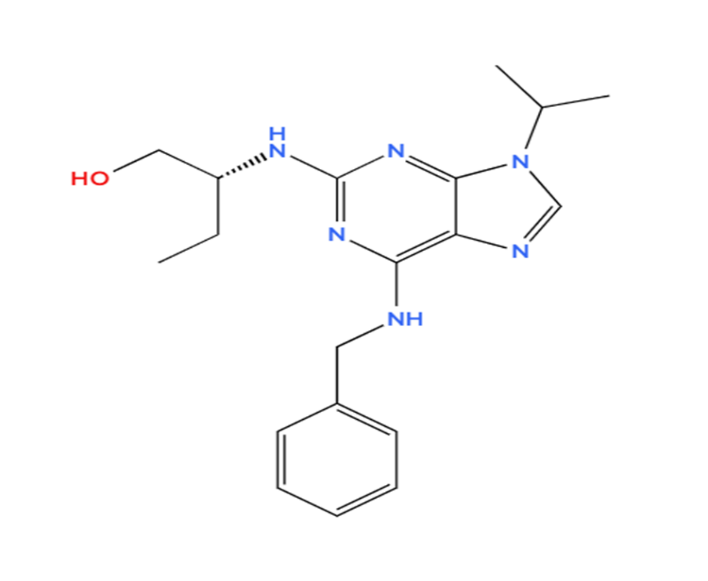
In this study, the selected complex features CDK2 bound to the small-molecule inhibitor Roscovitine. This ligand binds competitively to the ATP-binding pocket of CDK2, effectively inhibiting its kinase activity. By halting the enzyme's functionality, the cell cycle is arrested, preventing uncontrolled cellular proliferation—a hallmark of cancer progression.

This mechanism underscores the therapeutic potential of CDK inhibitors like Roscovitine in targeting specific pathways associated with cancer and other cell cycle-related diseases. The structural insights into such interactions with CDK2 offer valuable information for designing highly selective and potent anticancer agents.

The complex has the PDB ID: 2A4L.

To isolate the ligand from the macromolecule, we utilized Sublime Text and employed regular expressions to streamline the process. Specifically, we identified and extracted the lines beginning with "ATOM," which correspond to the protein, and those starting with "HETATM," which represent the ligand.

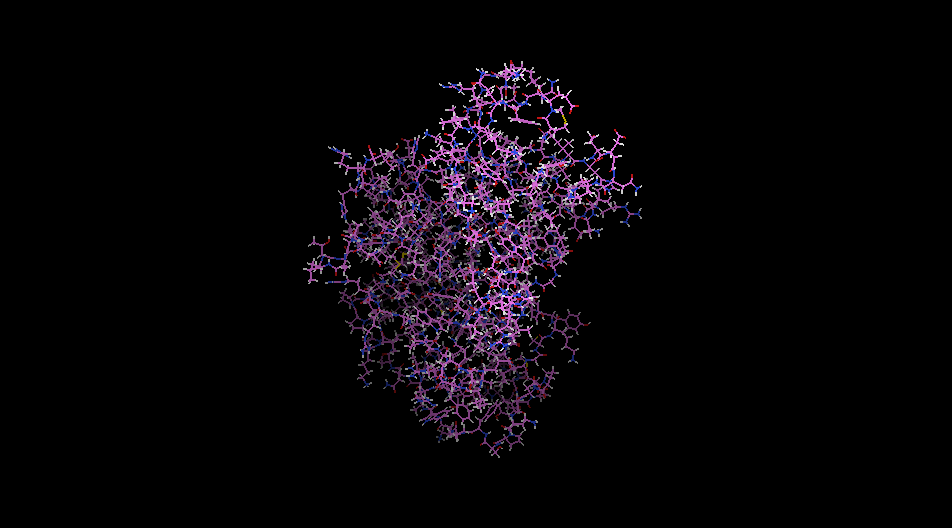
Additionally, we refined the ligand file by removing lines associated with water molecules (HOH), ensuring a cleaner and more precise dataset for analysis. This step was essential for subsequent structural and functional studies.



Document 1: The secondary structure of Roscovitine using *MolView*

**2-** Using AutoDockTools, we prepared the receptor and ligand files for docking analysis.

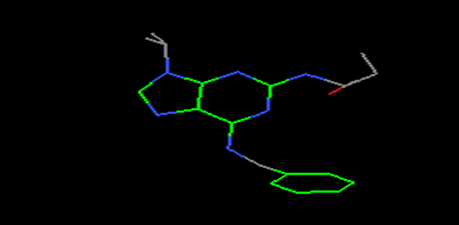
* For the receptor file, we began by cleaning the structure, which involved removing crystallographic waters as well as the water molecules previously identified and eliminated using regular expressions. Next, we added the missing hydrogen atoms, specifically focusing on polar hydrogens as they are the most relevant for docking studies. Finally, we computed the Gasteiger partial atomic charges to assign accurate charge distributions to the receptor (total charges= 6.97).



Document 2: The receptor CDK2

* For the ligand file, we similarly removed any remaining water molecules and calculated the Gasteiger charges. Additionally, we identified the root of the torsion tree, represented by a green sphere on the ligand. This step is critical for understanding the ligand's flexibility during docking simulations.

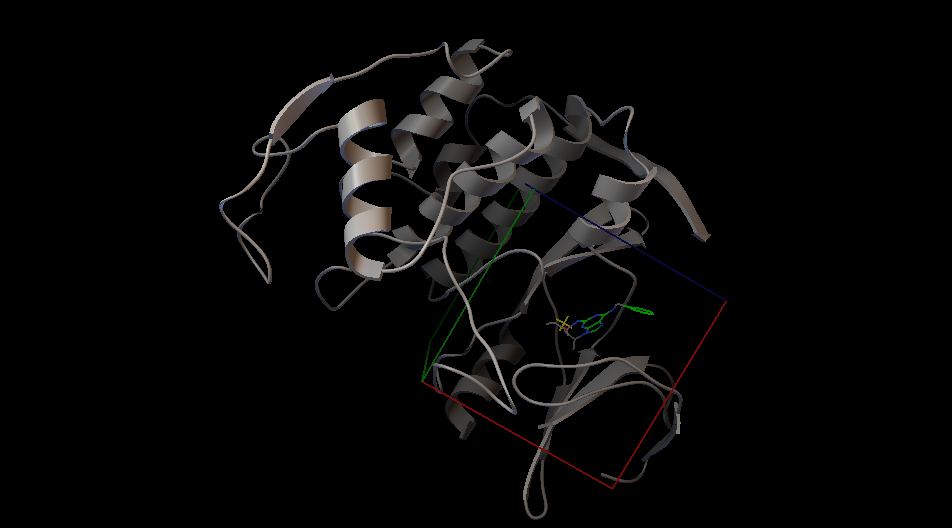
The ligand was found to have eight rotatable bonds (torsions), a number well within the acceptable range for efficient computation, ensuring that the docking process would not be computationally intensive.

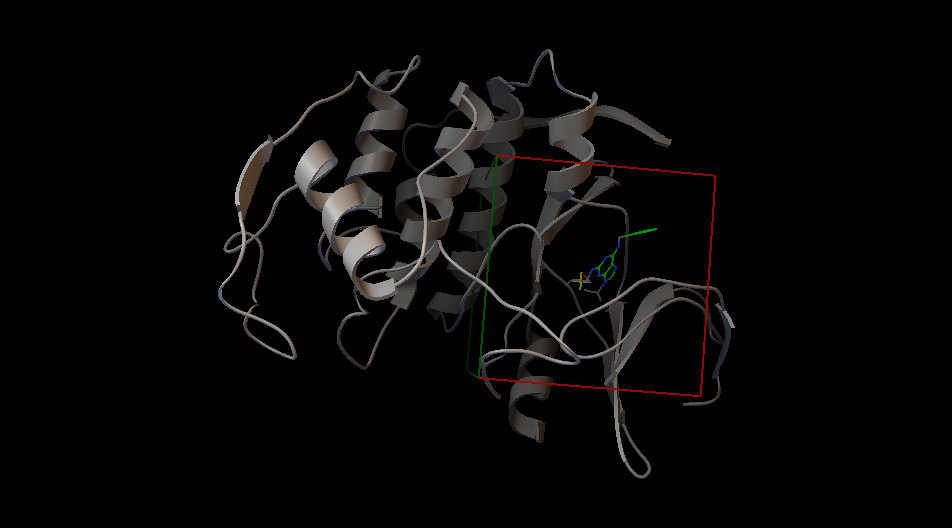


Document 3: The ligand Roscovitine

**3-** To compute the interaction energy maps for the binding site of CDK2, we began by placing a grid box centered on the ligand to ensure accurate coverage of the interaction site. The dimensions of the grid box were adjusted to accommodate the largest extent of the ligand, with the dimensions in each direction (x, y, and z) set to 60. Each axis is visually represented by a distinct color: red for x, green for y, and blue for z. The spacing between grid points was kept at the default value of 0.375 Å, balancing precision with computational efficiency.

The alignment of the grid box was crucial to the accuracy of the docking analysis. We centered the grid box on the ligand by matching its center coordinates to those of the ligand, resulting in x = 102.909, y = 99.042, and z = 81.973. This ensured that the grid accurately encompassed the ligand and its potential binding interactions.





Documents 4-5: The grid box around the ligand and the binding site

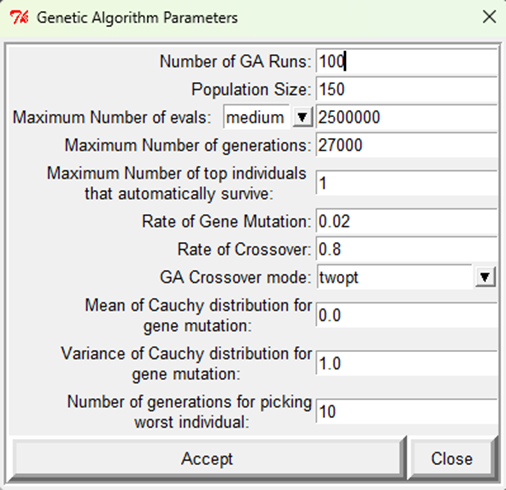
After configuring the grid box, we saved the grid parameter file (GPF), an essential input for the docking process. This file defines the grid box dimensions and the parameters needed for generating the grid maps. Using the AutoGrid4 tool and the GPF, we generated the grid log files (GLG), which provide detailed information about the execution process and the resulting interaction energy maps.

By analyzing these maps, we gain valuable insights into the binding affinities and potential interaction sites, forming a foundation for accurate and reliable docking simulations.

**4-** To perform the docking analysis using the Genetic Algorithm (GA), we carefully selected and configured the parameters to ensure both efficiency and accuracy. Specifically, we opted for 100 GA runs with a population size of 150 individuals, striking a balance between computational demand and thorough exploration of the ligand's conformational space. The medium-range docking was set with a maximum of 2,500,000 energy evaluations, providing adequate sampling for reliable results.

Key genetic algorithm parameters included a gene mutation rate of 0.02 and a crossover rate of 0.8, ensuring an appropriate level of diversity within the population while preserving high-quality solutions. Additionally, the mean of the Cauchy distribution for gene mutation was set to 1, which is particularly effective in balancing local and global search capabilities within the conformational space.

With these settings in place and leveraging pre-calculated interaction energy maps, we employed the Lamarckian Genetic Algorithm for docking. This hybrid approach combines the global optimization capabilities of GA with local refinement, improving the precision of the final docked conformations. After that, we ran the docking with autodock4 and retrieved the “dlg” file. The entire process took approximately 40 minutes, culminating in detailed insights into the ligand's conformational behavior and interactions within the receptor’s binding site.



Document 6: Running the GA 100 times

**5- a)** The 5 identified atom types are: **A, NA, C, OA, N.** These represent the atom types present in the ligand based on the docking setup. The total charge of the ligand is +1.0004e.

**b)**  The ligand has **8 torsional degrees of freedom**, all active, as indicated in the docking setup. These torsions are active between the following pairs of atoms:

1. CAQ\_2311 and CAR\_2312
2. CAR\_2312 and NAS\_2315
3. CAR\_2312 and CAK\_2313
4. NAS\_2315 and CAT\_2316
5. NAW\_2319 and CAZ\_2320
6. CAM\_2327 and NAJ\_2328
7. NAJ\_2328 and CAD\_2329
8. CAD\_2329 and CAE\_2330

The torsional free energy contribution is +2.3864 kcal/mol, as per the docking results.

**c)** The file lists several docking results. Among them, one of the best binding energy values between the ligand and the receptor is **-7.76 kcal/mol**, observed for the run 99 in the Lamarckian Genetic Algorithm (LGA) run.

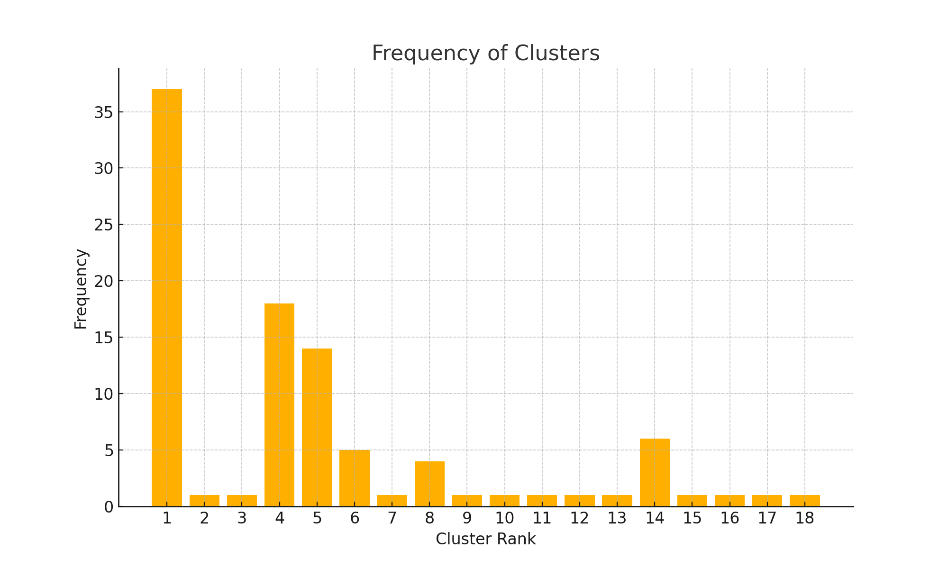
**d)** The docking results were analyzed, and the ligand poses were grouped into **distinct conformational clusters** based on structural similarity, using an RMSD tolerance of **2.0 Å**. This clustering provides insights into the stability and diversity of the ligand's binding conformations.

From the analysis, **18 distinct conformational clusters** were identified from **100 docking runs**. Each cluster is characterized by its **lowest binding energy**, the **run** that produced it, the **mean binding energy** of poses in the cluster, and the **number of docking poses assigned to the cluster**. These clusters vary in both their stability and frequency, indicating which conformations are most likely favored during binding.

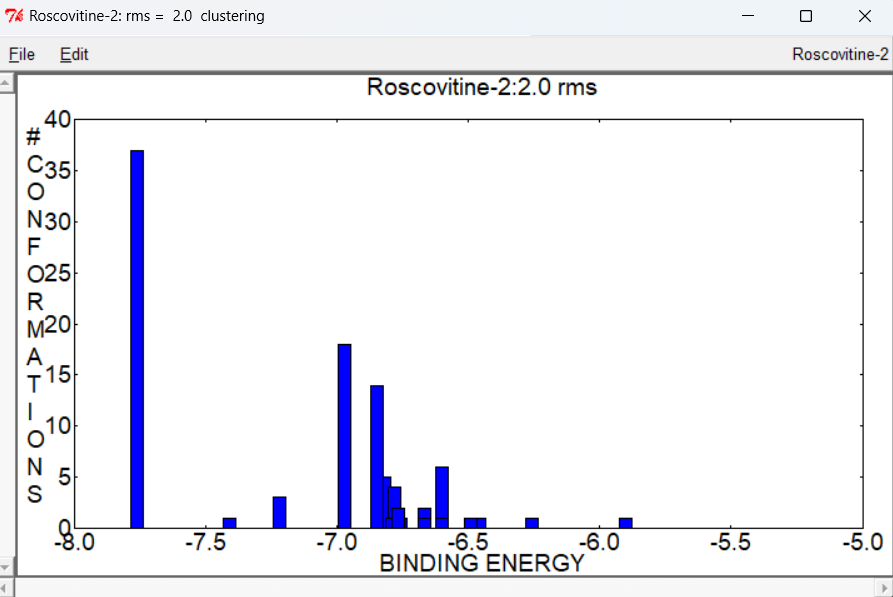
There are **9 multi-member clusters**, indicating significant conformational diversity among the docking results. These clusters represent conformations that occurred more frequently across the docking runs, implying they are likely energetically or structurally favorable.

The **most stable cluster**, Cluster 1, has a lowest binding energy of **-7.76 kcal/mol** (from Run 99) and contains 37 members. This cluster represents the dominant binding conformation of the ligand, appearing in **37%** of the docking runs.  
Other clusters, such as Cluster 4 and Cluster 5, occur less frequently but indicate alternative binding poses, with 18 and 14 members, respectively.  
Clusters 2, 3, 7–13, and 15–18 are rare, each appearing in only 1–4 docking runs, suggesting they represent less favorable binding conformations.

|  |  |
| --- | --- |
| **Cluster Rank** | **Frequency (%)** |
| 1 | 37 |
| 2 | 1 |
| 3 | 1 |
| 4 | 18 |
| 5 | 14 |
| 6 | 5 |
| 7 | 1 |
| 8 | 4 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |
| 13 | 1 |
| 14 | 6 |
| 15 | 1 |
| 16 | 1 |
| 17 | 1 |
| 18 | 1 |

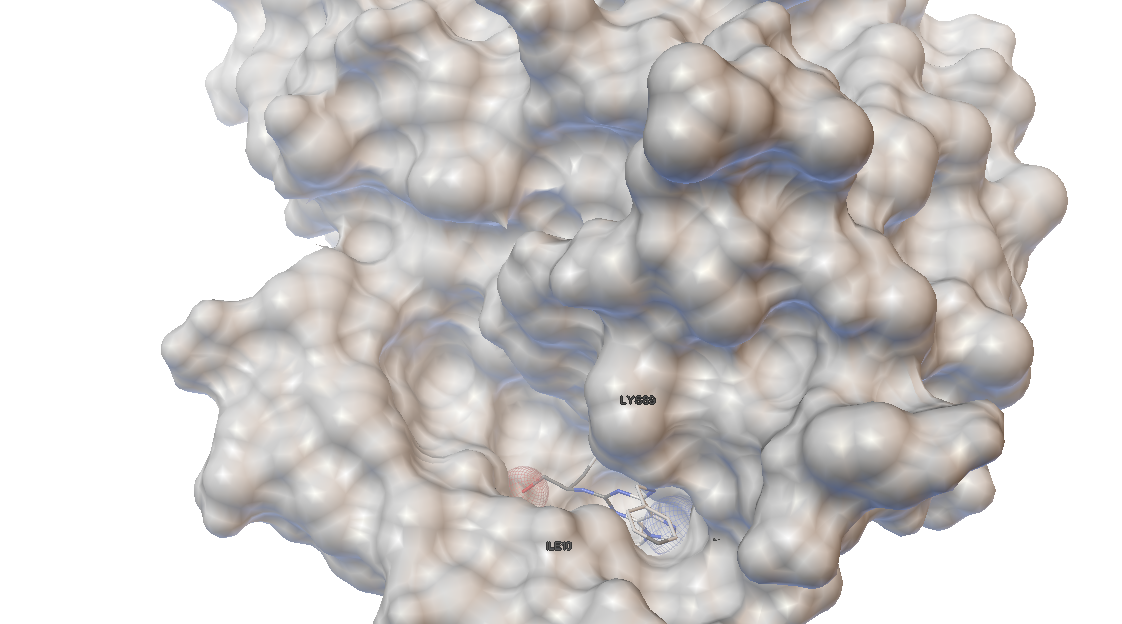
Document 8: Graph of frequency of each cluster

Document 7: Table of frequency of each cluster



Document 9: Graph showing the binding energy of each cluster with rmsd= 2.0

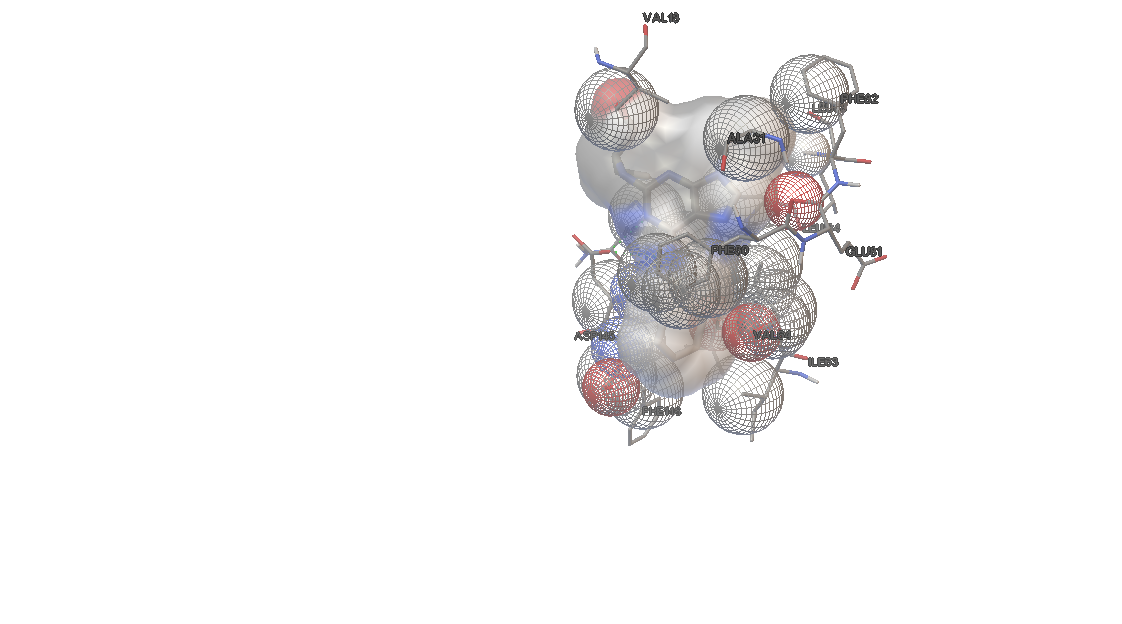
e**)** Our receptor, CDK2, features a well-defined binding pocket that accommodates the ligand Roscovitine with precision. The analysis reveals that Roscovitine fits seamlessly into this pocket, aligning perfectly with the receptor's structural features. Among the different clusters examined, variations in ligand orientations and poses were observed; however, all conformations consistently reside within the binding pocket without any steric clashes between the residues of CDK2 and those of Roscovitine.



Document 10: Docking of the first cluster

In Document 10, the first pose within the first cluster is identified as having the lowest energy, making it the most stable conformation.

For the ligand to bind to the receptor, several interactions occur. In the documents below, the spheres highlight the residues in close contact that participate in these interactions.



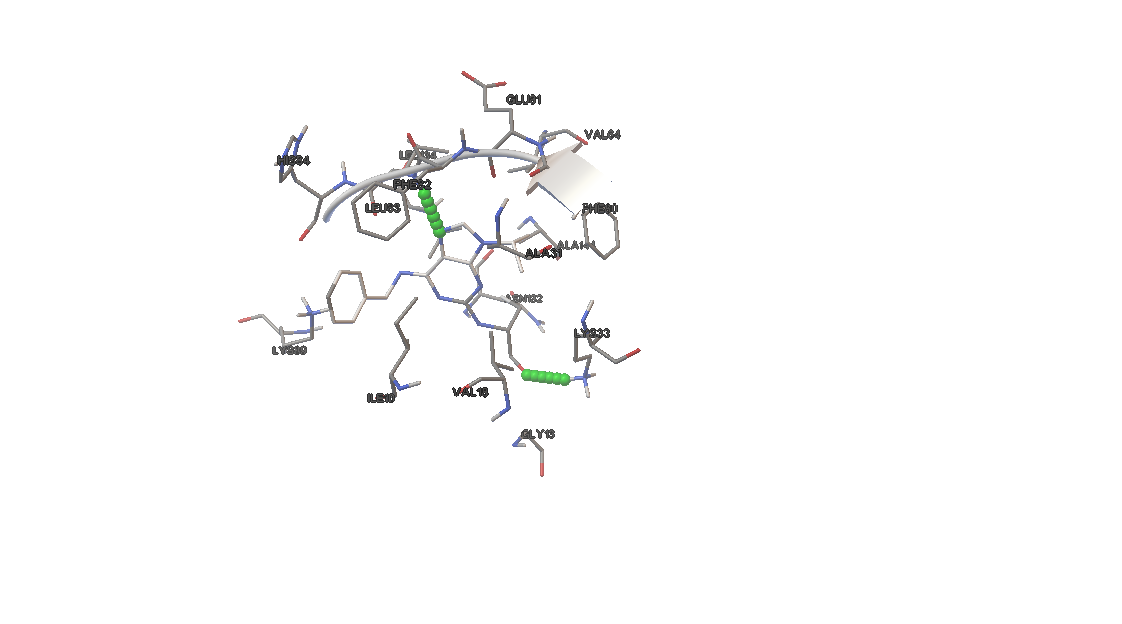
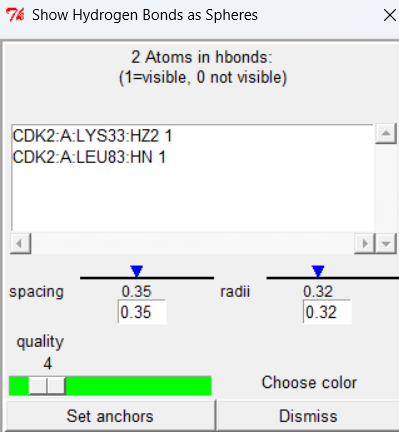
Document 11: Interactions between the ligand and the receptor

As we go through the different clusters and conformations, the interactions vary.

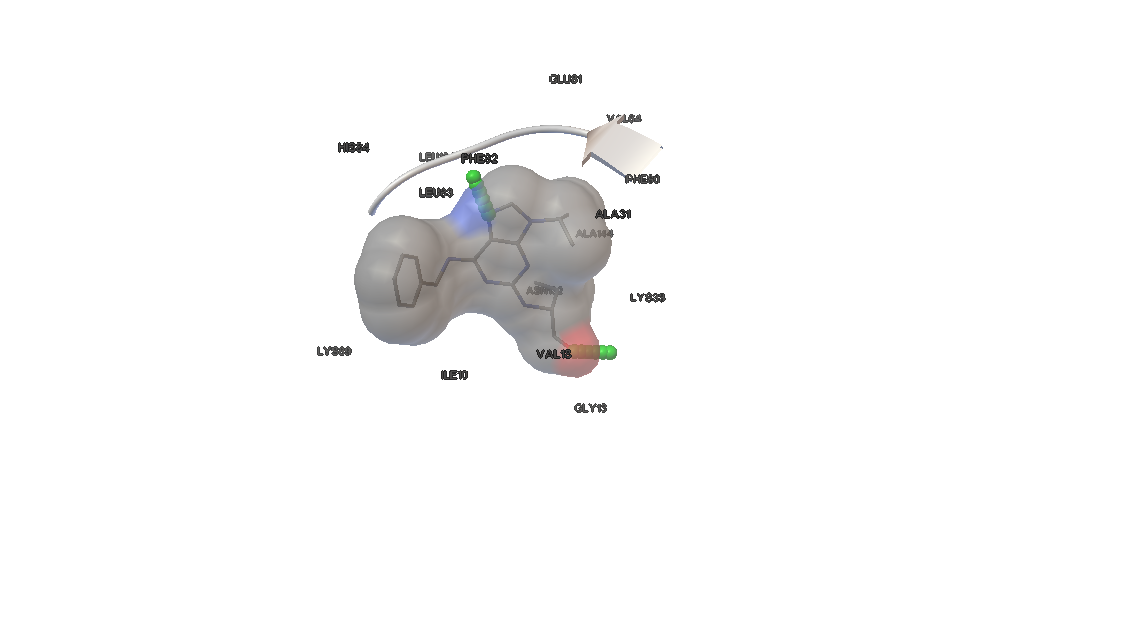
For instance, the first cluster forms two hydrogen bonds, the fifth cluster forms one, and the ninth cluster forms three.

These hydrogen bonds are visually represented by small green spheres.

**For Cluster 1:**

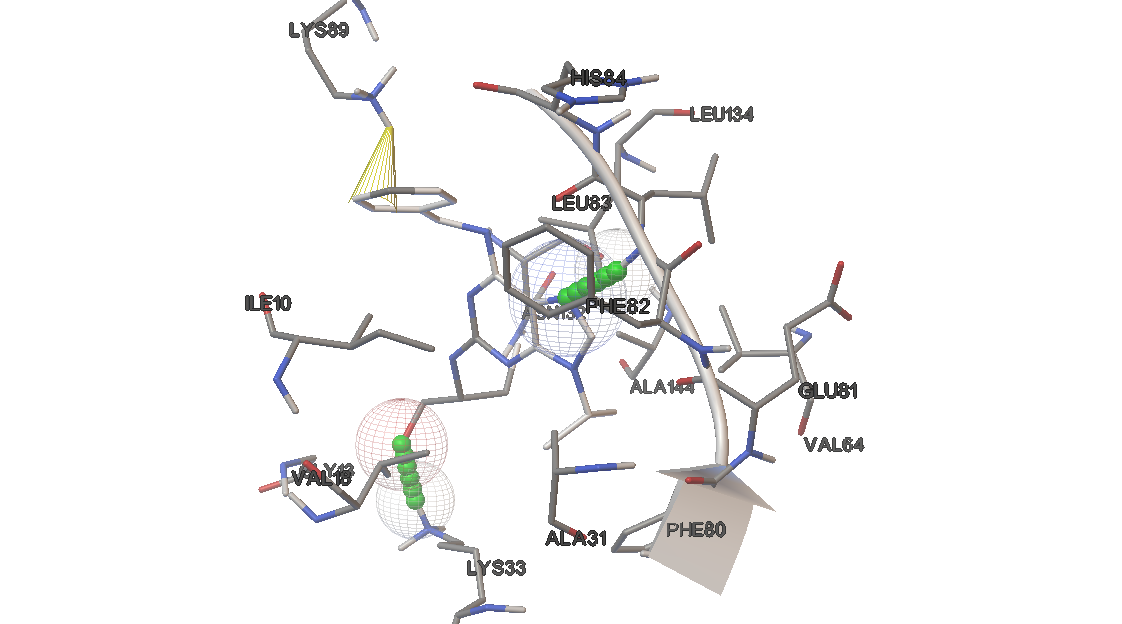
 

LYS33 and LEU83 from the receptor form a hydrogen bond with the ligand.

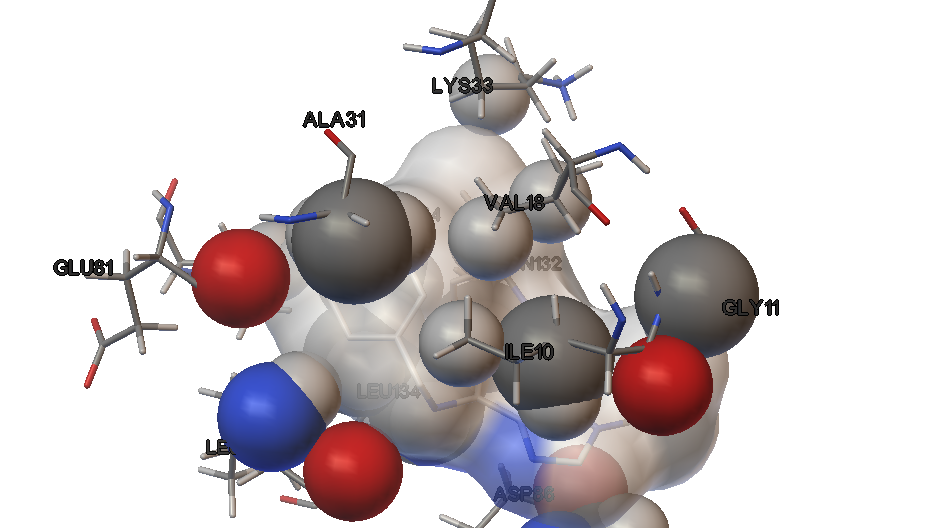


Document 12: This image provides a clearer view for identifying the ligand.

In the first cluster (most stable conformation), no pi-pi interactions are observed. However, there is one pi-cation interaction, van der Waals interactions, and hydrophobic interactions.



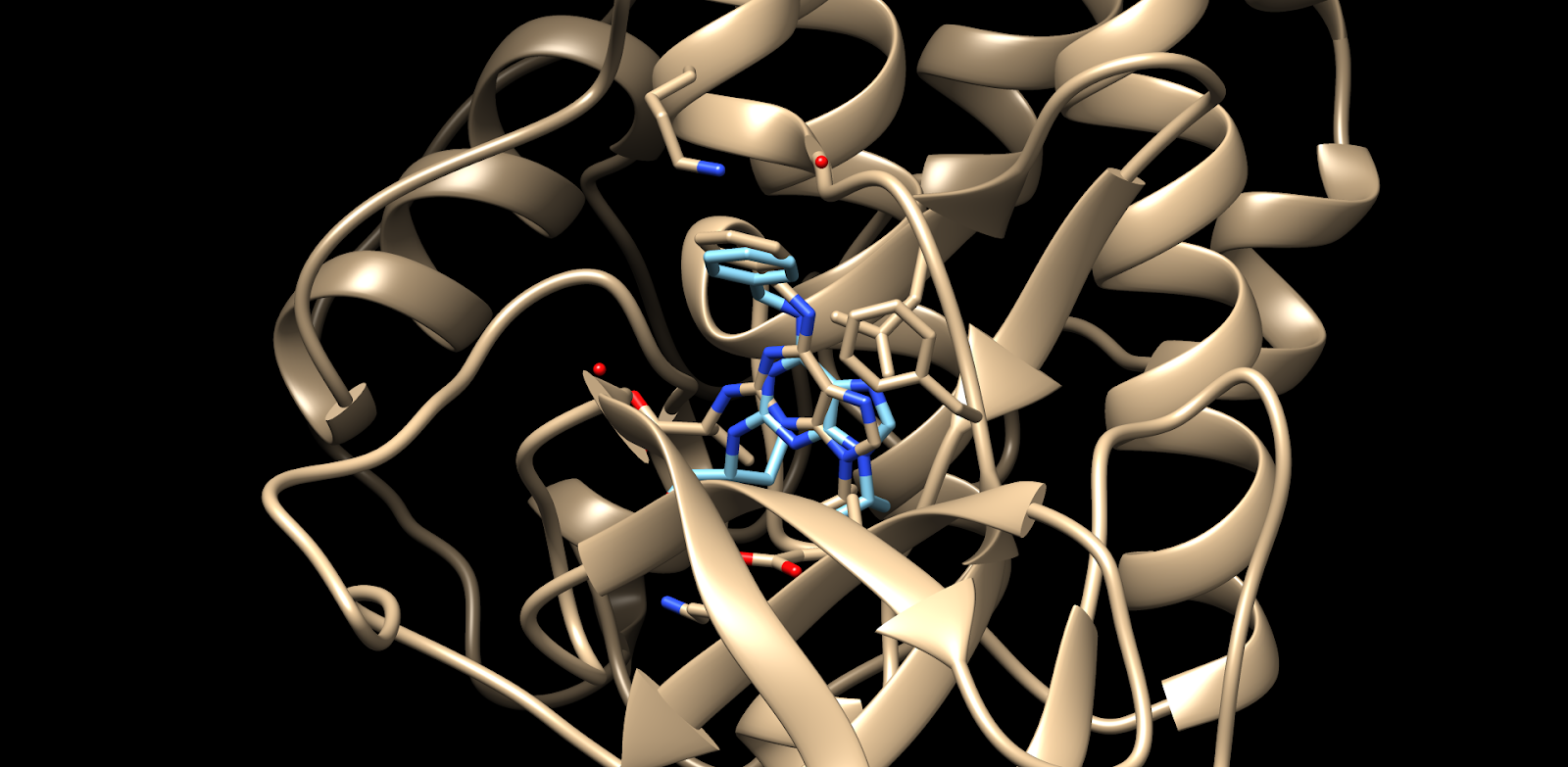
Document 13: pi-cation interaction observed in **yellow**.



Document 14:

**Val18** (hydrophobic) is located in the ATP-binding pocket of CDK2 and is a common participant in **hydrophobic** **interactions** with ATP-competitive inhibitors like Roscovitine.

f**)** A comparison between the position of the ligand in the experimental structure and the best-docked position reveals, upon visual inspection, that they are virtually identical.



Document 15: Superposition of the ligand from the experimental structure the one from the best-docked cluster on Chimera

When comparing the RMSD of the best-docked position obtained from run 99 to the reference structure, we find an RMSD of **1.205 Å**. This low RMSD value indicates a high degree of similarity between the docked pose and the experimental structure, demonstrating that the docking procedure effectively reproduced the experimental conformation.

All docked poses maintain an RMSD below 4.13 Å, though the orientation of the ligand varies in some cases. For instance, in certain poses, the branched ring points upward, while in others, it shifts downward. The position of this benzyl ring explains the specificity of Roscovitine in inhibiting CDK2.

Importantly, these variations do not cause any clashes.

Furthermore, a low entropy value of **0.45**, indicates that the docking poses within the clusters are fairly similar and tightly grouped. This suggests that the docking algorithm consistently places Roscovitine in similar poses within CDK2's binding pocket and the poses do not vary significantly from each other, which is a sign of convergence and consistency in docking results.

**CDK2 vs CDK5 docking with Roscovitine:**

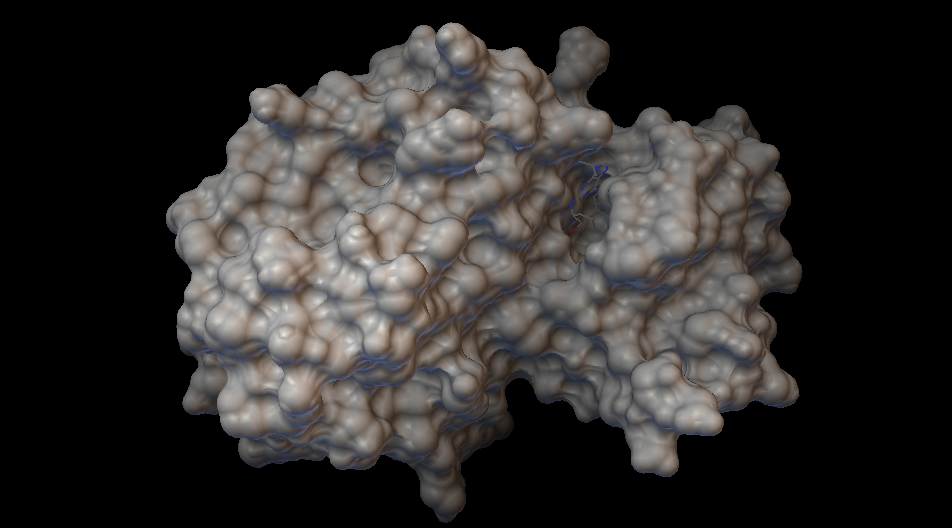
First, for the docking part:

We cleaned the pdb file of CDK5 in order to only keep the atoms of chain A which is the one that binds to the ligand.

The docking analysis of Roscovitine with CDK2 and CDK5 revealed key similarities and differences in their binding profiles, considering they are both part of the same family.

Both kinases exhibited stable docking poses, with binding energies ranging from -7.16 to -5.76 kcal/mol for CDK2 and from -7.68 to -5.67 kcal/mol for CDK5. The ligand predominantly occupied similar spatial regions within each binding pocket (both having their entropy equal to 0.45), although CDK5 displayed slightly more variability in ligand positioning, reflecting a potentially more adaptable binding site. Quaternion data supported these observations, with CDK2 showing consistent orientations across docking runs, while CDK5 presented greater angular diversity.

Intermolecular interactions were dominated by hydrogen bonding and van der Waals forces in both cases, with conserved hydrogen-bonding patterns observed in CDK2 and additional unique interactions in CDK5 that might contribute to its slightly more favorable binding energies.

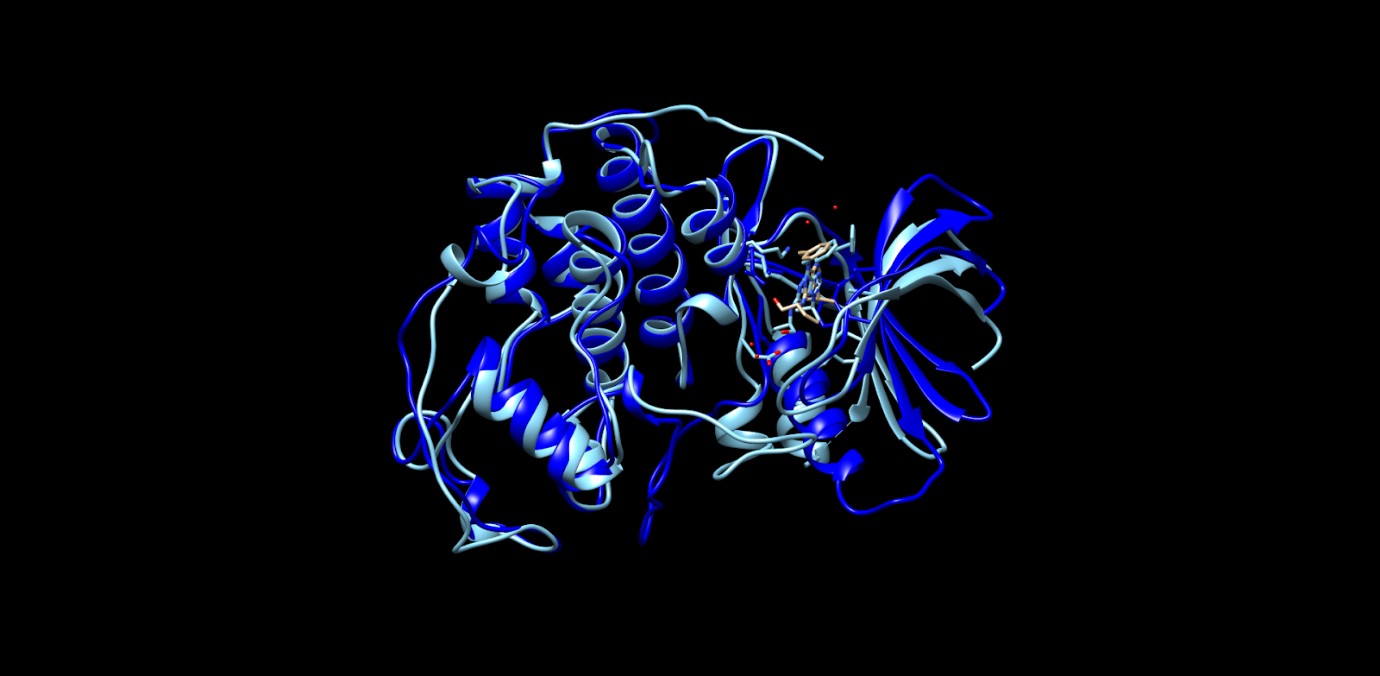
Overall, while Roscovitine binds effectively to both kinases, CDK5 demonstrated a marginally stronger and more flexible binding profile, likely due to subtle differences in the architecture of its binding site. 

Document 16: Best-docked position of Roscovitine in CDK5

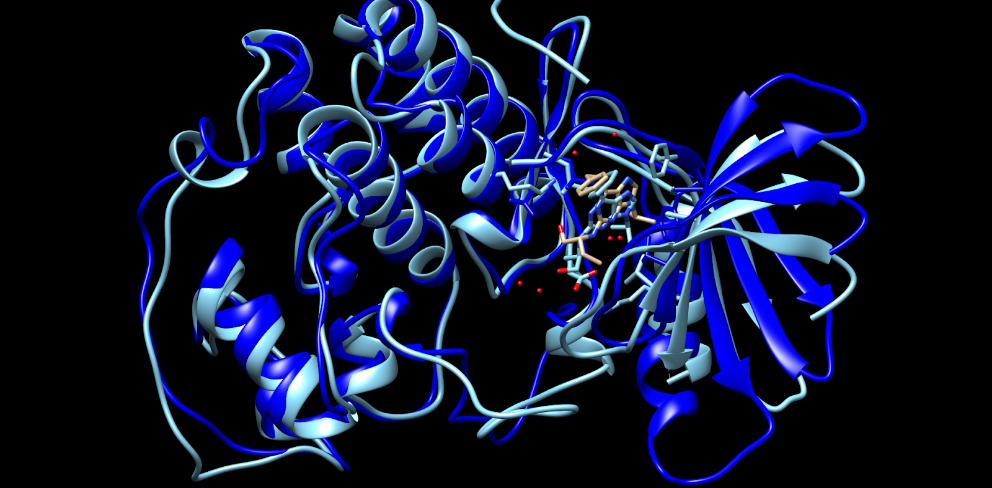
Now for the visualization part:

CDK5 consists of four chains, but for our analysis, we focused exclusively on Chain A, as it contains the ligand of interest. Using Chimera, we isolated Chain A (represented in dark blue) by hiding the remaining chains (B, C, and D). Subsequently, we superposed the structure of CDK5 Chain A with that of CDK2, which consists of a single chain (depicted in light blue).

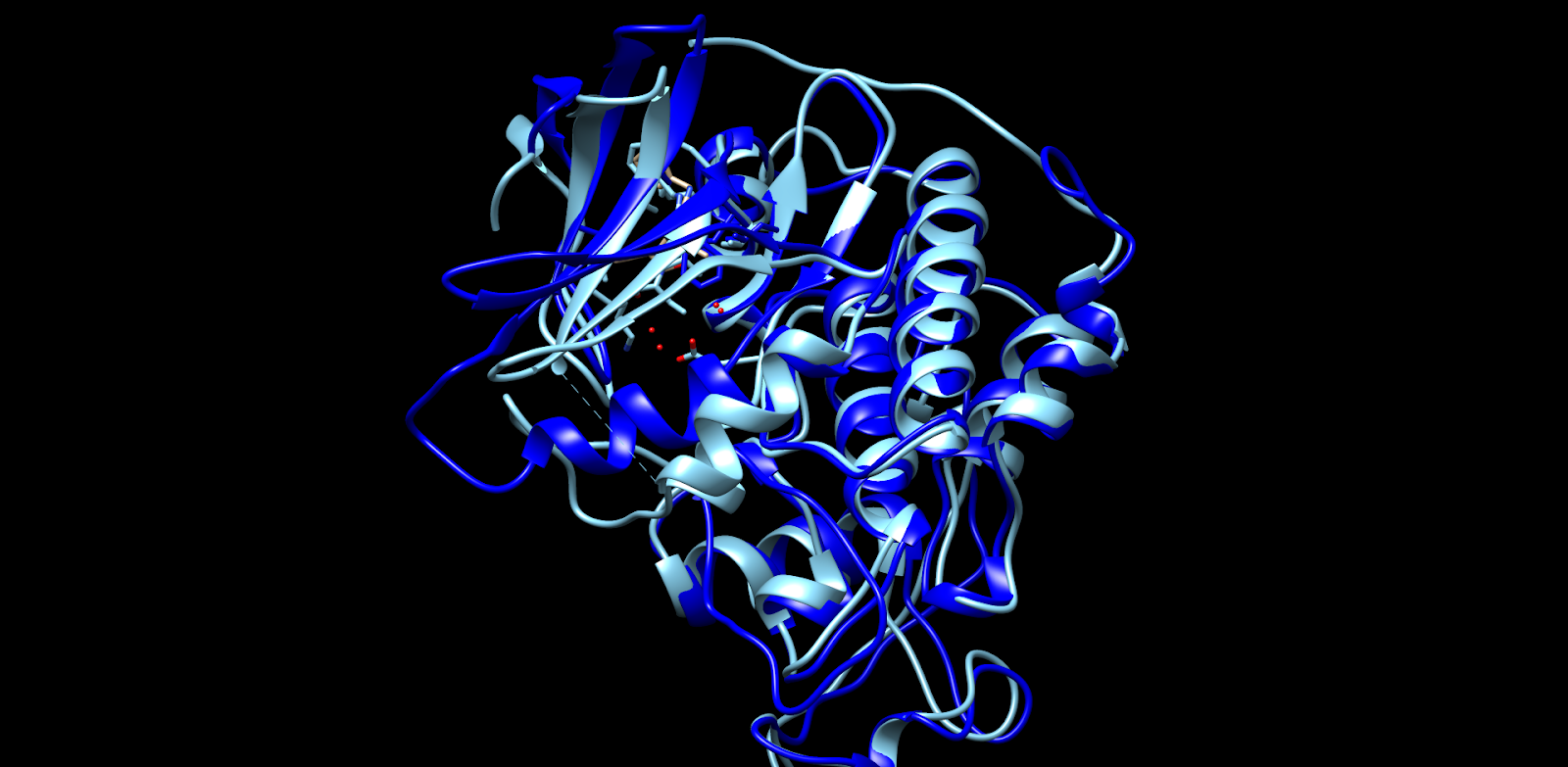
The comparison revealed a striking similarity between the two structures. Both proteins exhibit highly similar overall architectures, including the loops and the spatial location of their binding sites. Additionally, the ligands in both receptors are aligned, further highlighting their structural resemblance.



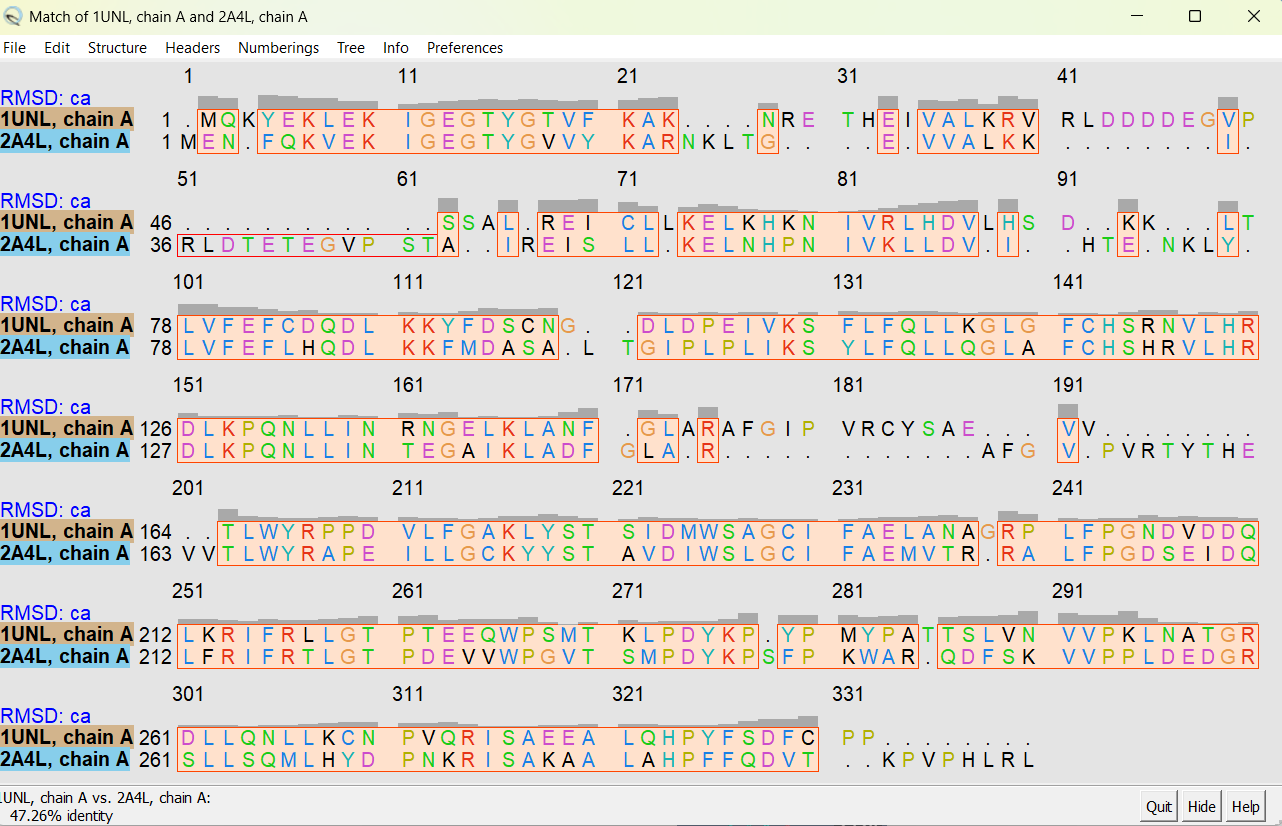
Document 17: Superposition of CDK5 and CDK2 on Chimera



Document 18: Superposition of Roscovitine Ligands in CDK5 and CDK2 with a slight difference in position

However, CDK2 has a chain break that is not present in CDK5. The chain break is represented by a dashed line (in light blue): 

CDK5 and CDK2 share significant sequence similarity, with 47.26% identity and numerous conserved residues along their amino acid sequences.



Document 19: Sequence similarity between CDK2 and CDK5

These findings highlight important pharmacological insights into Roscovitine's interaction with CDK2 and CDK5, which are critical enzymes in regulating cell growth and division.

The slightly stronger and more flexible binding observed with CDK5 suggests that Roscovitine might inhibit this kinase more effectively, potentially making it a better target for diseases where CDK5 plays a significant role, such as neurodegenerative disorders. On the other hand, the stable binding with CDK2 reinforces its potential use in cancers, where CDK2 activity drives uncontrolled cell proliferation. The differences in binding patterns, including the distinct interactions with residues unique to CDK5, could be leveraged to design drugs that are more selective, reducing side effects caused by non-specific binding to other kinases.

In conclusion, these findings offer valuable information for optimizing Roscovitine or developing new inhibitors tailored to treat specific diseases involving CDK2 or CDK5 dysregulation.