High level summary

1. Task description

a. Compounds were downloaded from Enamine Store that were similar to the co-ligand from PDB: 1H1Q. The dataset was analysed & top 300 **diverse** compounds were selected, and prepared for docking with AutoDock Vina (notebook 1_PurchasableCompounds.ipynb).
b. Protein files were prepared (2_PrepareProtein.ipynb) and docking was finally carried out (3_Docking_with_Vina.ipynb).
c. A workflow for co-folding (4_Boltz_predictions.ipynb) using Boltz2 was also prepared and executed.

2. Results obtained

a. Top 16 compounds are identified and ranked for synthesis. These were identified after thoroughly analysing their docking scores, ligand efficiencies (LE), lipophilicity (clogP) and other descriptors (5_Analyse_Vina_Output.ipynb).
b. Boltz2 predictions failed as my machine lacks the memory needed.

3. Challenges

a. I have concerns about the quality of ligand alignment, pose generation and accuracy of the Vina scoring function.

b. Running Boltz2 is quite memory-heavy.

4. Next steps

a. This has been described thoroughly in the notebook 5_Analyse_Vina_Output.ipynb under the sections “What could be improved further?”, “Post-Docking methods for more accurate binding free energy scores” & “Next experimental steps”