**Module 5 Assignment**

**Tasks:**

1. Complete the **Module 5.1a** tutorial and fill in **Table 1**.
2. Complete the **Module 5.1b** tutorial and fill in **Table 2**.
3. Upload your completed Worksheet to the Module 5.1 Assignment element in DigitalChalk, and write a brief explanation detailing which grid(s) you would take forward for prospective structure-based virtual screening making sure to reference the results of your validation studies.

In module 3, we prepared and validated our targets, two PLK1 structures 2OWB and 2RKU. Although, 2RKU appears to be “best” based on protein reliability report, sitemap score, resolution; in this module we are going to test both structures to screen ligands library. In module 4, we prepared our ligand libraries. Some of which contain compounds that are known PLK1 binders. The first ligand library we worked on PLK1 DUD-E library. This library contains a mixture of known actives and decoys. In module 5, they have been used as part of our validation study to assess whether our target can identify known binders effectively. The initial profiled DUD-E library contains 50% druglike, 30% near-druglike and 2% Leadlike ligands and next filter them to remove the non-druglike and fragments ligand and this filtered library contains 5782 compounds. And in section 4.1b these 5782 compounds were prepared with LigPrep tool. In section 4.1C we prepared the cognate ligands of 2OWB and 2RKU with LigPrep tool. It is a good practice to perform docking validation using known binders prior running a structure-based virtual screen. Before running the docking we generate multiple receptor grids for each structure of PLK1 protein in section 4.2.

**Table 1**: In section 5.1a, we used Glide to dock the co-crystalized ligands to the multiple receptor grids of 2OWB and 2RKU. To assess how well Glide can reproduce the known cognate ligand binding poses, we analyzed docking score and RMSD, shown in table below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **PDB** | **Grid** | **2OWB\_ligand** | | **2RKU\_ligand** | |
| **Docking Score** | **RMSD** | **Docking Score** | **RMSD** |
| **2OWB** | **Dry** | -10.931 | 0.7372 | -8.042 | 3.5712 |
| **1 Water** | -12.032 | 0.3104 | -9.010 | 4.1141 |
| **4 Waters** | -13.765 | 0.3883 | -4.940 | 1.3501 |
| **2RKU** | **Dry** | -9.979 | 3.8628 | -7.918 | 1.1345 |
| **1 Water** | -10.524 | 3.1185 | -5.929 | 0.00 |
| **3 Waters** | -11.601 | 5.0303 | -6.327 | 0.00 |

From the table, it appears that, 2OWB receptor grid without water, with 1 water molecule and 4 water molecules show better docking poses in comparison with other arrangements.

**Table 2**: In section 5.1b, we used Glide to dock the filtered PLK1 DUD-E ligand library, both with and without grid-based constraints, for each of the 2OWB and 2RKU receptor grids. Then we used the studies to assess if our structures and docking methodology can demonstrate enrichment of known active compounds over decoys. Following table demonstrate the results of our analysis:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **PDB** | **Constraints** | **Grid** | **ROC** | **RIE** | **% Actives in the top N% of results** | | | | |
| **1%** | **2%** | **5%** | **10%** | **20%** |
| **2OWB** | **No constraints** | **Dry** | 0.91 | 10.95 | 31.6 | 48.1 | 64.6 | 78.5 | 87.3 |
| **1 Water** | 0.94 | 10.30 | 26.6 | 36.7 | 68.4 | 78.5 | 93.7 |
| **4 Waters** | 0.75 | 3.52 | 6.3 | 12.7 | 17.7 | 32.9 | 48.1 |
| **1 or more H-bond constraint** | **Dry** | 0.88 | 6.72 | 11.4 | 20.3 | 41.8 | 58.2 | 82.3 |
| **1 Water** | 0.87 | 6.99 | 13.9 | 20.3 | 44.3 | 65.8 | 79.7 |
| **4 Waters** | 0.71 | 3.50 | 8.9 | 13.9 | 20.3 | 24.1 | 41.8 |
| **2RKU** | **No constraints** | **Dry** | 0.92 | 10.70 | 36.7 | 45.6 | 62.0 | 79.7 | 87.3 |
| **1 Water** | 0.89 | 8.23 | 20.3 | 30.4 | 50.6 | 67.1 | 78.5 |
| **3 Waters** | 0.89 | 9.53 | 25.3 | 38.0 | 57.0 | 67.1 | 79.7 |
| **1 or more H-bond constraint** | **Dry** | 0.91 | 9.75 | 29.1 | 39.2 | 54.4 | 77.2 | 87.3 |
| **1 Water** | 0.88 | 8.86 | 21.5 | 31.6 | 54.4 | 70.9 | 82.3 |
| **3 Waters** | 0.90 | 9.16 | 24.1 | 35.4 | 57.0 | 68.4 | 84.8 |

Both structures without water grids and no constraints docking show good enrichment of known actives over the decoys. 2OWB receptor grid with 1 water molecule (no constraints) shows better enrichment of known actives over decoys in comparison with other arrangement.

our target can identify known binders effectively in the following manner

**2OWB\_1water\_no\_constraints > 2OWB\_dry\_no\_constraints> 2RKU\_dry\_no\_constraints > 2RKU\_dry\_ constraints**

It is normal for enrichment to be worse when constraints and structural waters are used, so unless the cognate ligand redocking results are poor, using the grid that demonstrates the best enrichment for SBVS is generally advised. Often docking scores/RMSD values for cognate ligands can show improvement when constraints or crystallographic waters are included, but water molecules tend to limit the number and diversity of compounds that can be successfully docked into a pocket since waters are treated as rigid during docking.

Constraints can also limit the types of molecules able to dock; those with no/few HBA or HBA groups will be unable to meet the required H-bond criteria. Sometimes this is desirable, for example, if your protein of interest has known structure-activity relationships and the literature demonstrates that a particular interaction must be present, then using constraints during a screen can be a useful filter. For our PLK1 structures, we already observed that some of the H-bonds are ligand-specific when setting up the constraints, so a more in-depth analysis of the literature would be recommended before using a constraint set such as this for a prospective screen, particularly if looking for diverse novel compounds.

The aim of this assignment was to demonstrate the pros and cons of using various grids and constraints. Your data nicely indicate that using constraints and crystallographic waters can be useful for replicating known poses, but are likely to be limiting for virtual screens.