UENF-CBB-LQFPP 2020

HTP SurflexDock Tutorial.

Dr. João Luiz de Almeida Filho

&

Prof. Dr. Jorge Hernandez Fernandez

December / 2019

1. Introduction

In this tutorial, we will use the HTP SurflexDock web service to do a structure-based virtual screening (SBVS). The goal of SBVS is to filter out possible inhibitors of a large amount of compounds for a particular protein through the docking methodology. However, docking does not consider the effects of receptor flexibility on ligand interaction and therefore generates false negatives in SVBS experiments with other receptor ligands that have distinct chemical properties from the original ligand problem (the receptor memory effect).

HTP SurflexDock provides a solution to this problem by using the ensemble docking technique. In this technique an ensemble of receptor conformations is generated by molecular simulation, afterwards, the docking of the compounds is performed with each of these conformations. In HTP SurflexDock, each ensemble created has 4 conformations: the original receptor conformation (submitted by the user) and 3 other conformations representing more than 50% of the structural space obtained in 5 ns of molecular simulation.

For this tutorial, we will use the C-terminal domain of human Angiotensin I Converter Enzyme (ACE) as a receptor. ACE is a zinc dependent metalloprotease involved in vascular problems such as hypertension, myocardial infarction and diabetic nephropathy (Bateman et al., 2017; Evans et al., 2016) . In the human body, ACE converts angiotensin I to angiotensin II, the latter an important vasoconstrictor (MASUYER, et al., 2012). The three-dimensional structure of this enzyme was determined by X-ray crystallography in 2003 (NATESH, et al., 2003). A zinc (II) ion is identified at its catalytic site, coordinated by 2 His and 1 Glu residues (GUAN et al., 2016; JALKUTE et al., 2013; ZHANG et al., 2013). Due to their therapeutic applicability, the development of ACE inhibitors is quite common and today we have a variety of these drugs on the market such as Captopril, Lisinopril or Enalaprilat.

In this context, we will use human ACE to do an inhibitor enrichment experiment, that is, let's see if HTP SurflexDock is able to filter real inhibitors from a library containing 100 compounds between inhibitors and decoy molecules.

Our experiment with HTP SurflexDock will consist of the following steps:

(I) File preparation,

(II) Submitting the project to HTP SurflexDock,

(III) Result analysis and

(IV) Result refinement.

2. File Preparation

To submit a job on HTP SurflexDock, we need to have the structure of the receptor in PDB format and a compound library. Thus, to organize our experiment, we will create the folder VS_HTP on the desktop of the computer, where we will save all the files generated on this tutorial.

2.1 Receptor File preparation.

In this step, we will get and prepare the receptor file. Download the C-terminal domain of human ACE receptor here and save it as hACEcOk.pdb in the VS_HTP folder. The preparation consists of opening the file VS_HTP/hACEcOk.pdb with any text editor (wordpad or gedit for example) and checking that the file has no missing residues or atoms. For this, we will observe if the second and fifth columns are ascending order without skips. As our file looks ok, so let's just close the file.

2.2 Compound Library Preparation

HTP SurflexDock only supports compounds in PDBQT format. However, since the virtual screening experiment can use hundreds to thousands of compounds, you should compress these compounds into a single .zip or tar.gz file with a file compression software such as 7-zip before submitting to HTP SurflexDock.

For this SBVS project, download the compound library (download) and save it with the name "sample_15_lig_85_dec_hace_c.tar.gz" in the VS_HTP folder.

The VS_HTP / sample_15_lig_85_dec_hace_c.tar.gz library contains:

- 15 known ligands for hACE. That is, compounds that have high affinity for the enzyme. These compounds have the prefixes "*new_ligand"*, "*BPP"* and "*LPR"*.
- 85 molecules supposedly without activity with ACE. These molecules have the prefix "*decoys"*.

3. Submitting the project to HTP SurflexDock

For this step, we will access HTP SurflexDock at http://biocomp.uenf.br:81 using the web browser (Google Chrome, Firefox or Edge). The HTP SurflexDock homepage has the layout as shown in Figure 1 below and has 3 main sections to fill out.

Thus, we will fill in the information needed to perform the virtual screening for ACE. Let's start with the "**Experiment Identification**" section, giving our project a name so we can get a direct link to the experiment results.

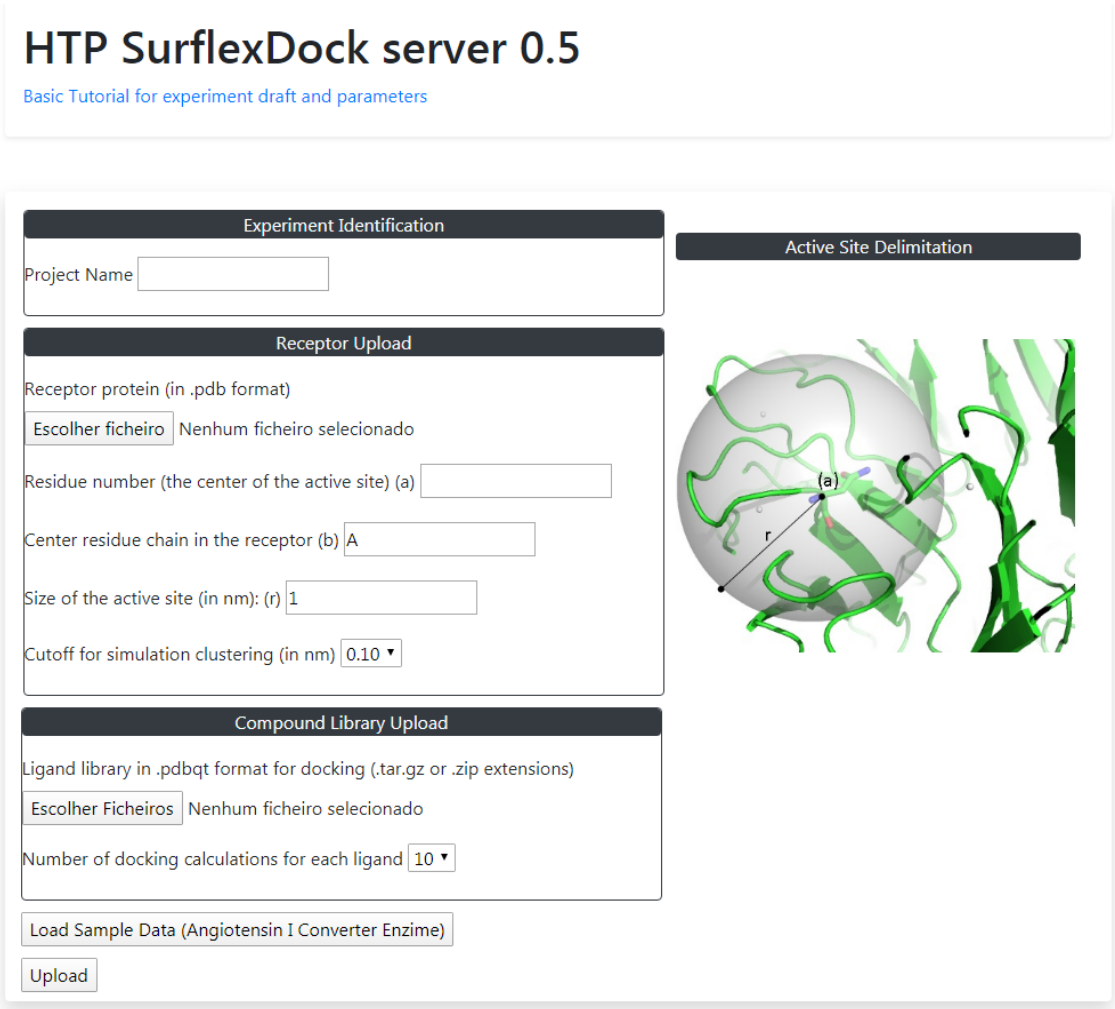

Figure 1: HTP SurflexDock homepage. You can click the "**Load Sample Data (Angiotensin I Converter Enzyme)**" button to automatically load the experiment data from this tutorial.

In the "**Receptor Upload**" section we will enter the receptor information and the parameters defining the active site of hACE. In HTP SurflexDock, the active site is defined by a sphere centred on the alpha carbon of the residue specified by the "**Residue number"** (a) and with radius (r) in nanometers (Figure 2).

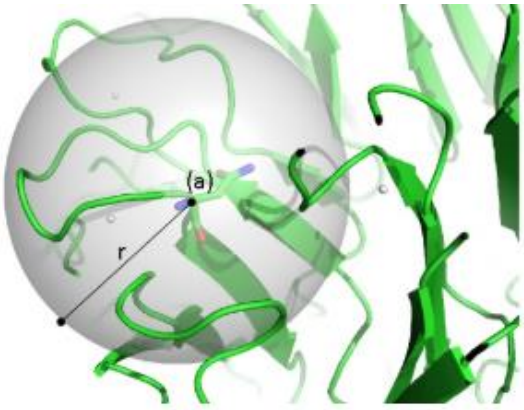

Figure 2: Active site definition for HTP SurflexDock. The SurflexDock HTP to define the receptor's active site utilizes a sphere of radius 'r' (nanometers) and centered on the alpha carbon of a residue (a).

Thus, in this section we must:

- Upload the receptor file (VS_HTP / hACEcOk.pdb) into the "**Receptor protein (in .pdb format)**" field.
- Let's use Glu372 alpha carbon near the zinc ion present in the enzyme as the centre of the active site. So enter the number 372 in the **"residue number**" field.
- The Glu372 is in the "A" chain of our receptor. Therefore, keep "A" in the "**Center residue chain in the receptor"** field.
- Let's define the 1.2 nm radius for the sphere that defines the active site in the "**Size of the active site (in nm)"** field.
- For this experiment, we will keep the simulation-clustering cutoff at 0.10 nm in the "**Cutoff for simulation clustering (in nm)**" field. This parameter is used by HTP SurflexDock to create clusters that differ by a maximum R.M.S.D. of 0.10nm of the defined active site over the molecular simulation trajectory. HTP SurflexDock will compute a representative structure for the 3 largest clusters to create the ensemble.

In the "**Compound Library Upload**" section, we should enter information about the compound library:

- Load the compound library (file VS_HTP / sample_15_lig_85_dec_hace_c.tar.gz) into the "**Ligand library in .pdbqt format for docking**" field and;
- Select 10 in the "**Number of docking calculations for each ligand**" field to have HTP SurflexDock do 10 docking evaluations for each submitted compound.

We completed the project submission by pressing the "**Upload**" button. HTP SurflexDock will upload the data to the server and start the experiment. At the beginning of the experiment, you will receive a complex URL containing a link to the results page. This experiment takes normally 11 to 14 hours.

4. Result analysis

HTP SurflexDock presents the results in a layout that can be divided into three sections (Figure 3): The **Experiment Detail** section contains information about the submitted SBVS project. The **Clusters of Simulation** Section graphically presents the matrix of the clusterization experiment over the obtained trajectory in 5ns of molecular simulation. HTP SurflexDock calculates a representative conformation of the 3 largest clusters to compose the receptor ensemble, in addition to the structure submitted by the user (control). Finally, The **Compound Ranking** section contains a table that has the results of calculated docking of the compounds in each receptor conformation.

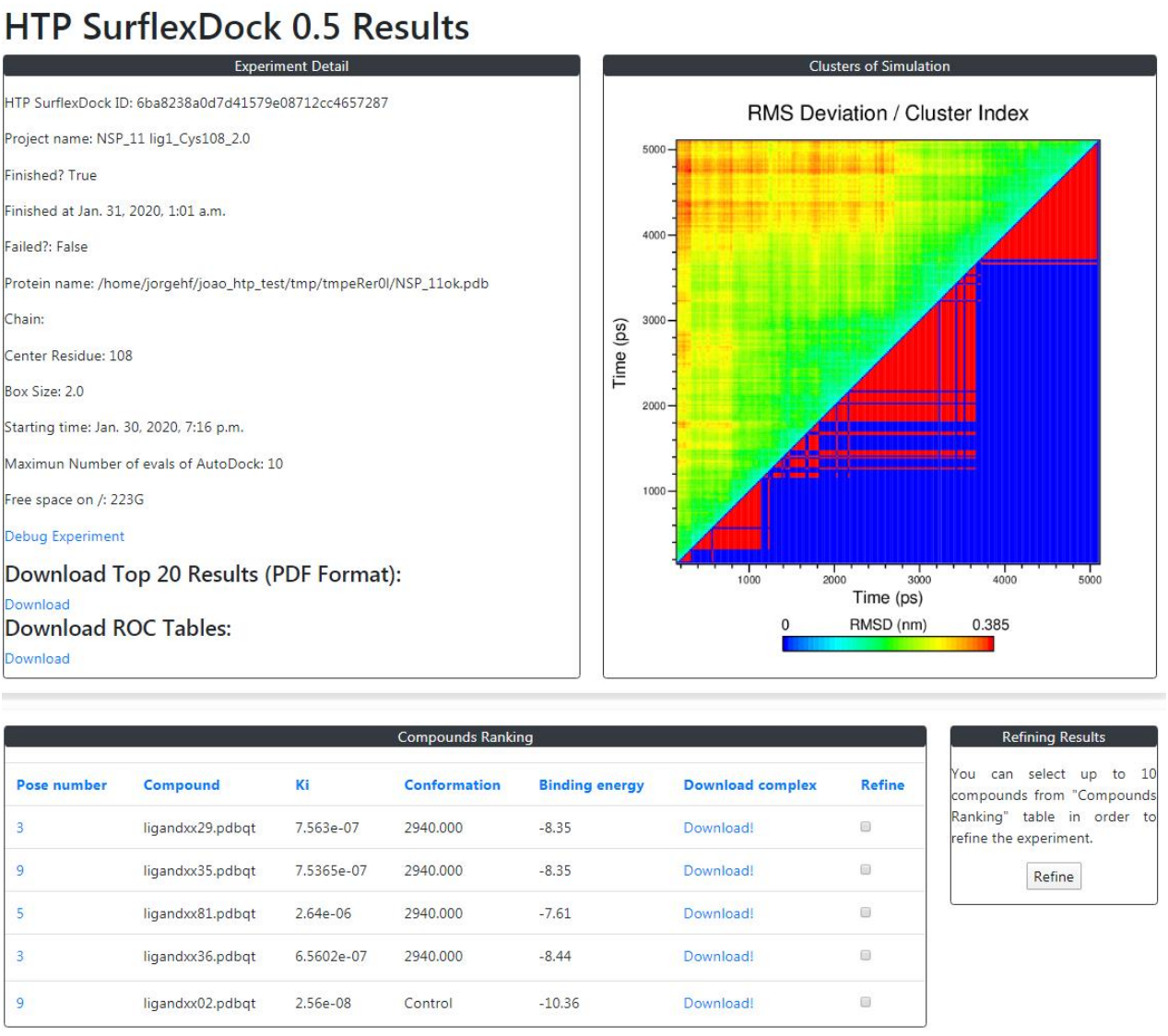

Figure 3: HTP SurflexDock results page.

When our SBVS experiment is completed, the table in **Compound Ranking** section will show a total of 400 calculated dockings. That is, the best docking result for each submitted compound multiplied by the number of four ensemble conformations. In an SBVS, we typically are interested in the compounds that obtained the lowest binding

energy. So let's sort the table in ascending order of the binding energy by clicking on the "**binding energy**" heading on result table. Note that HTP SurflexDock has been able to enrich a large number of ligands among the “best” compounds in this table.

4.1 Viewing complexes.

A low binding energy is a strong indication that the evaluated compound is a promising ligand for the receptor protein in the SBVS experiment. However, in a real experiment other parameters should also be considered. Among them, the most important are the structural shape and the binding pocket of the ligand in the receptor. Thus, we will visualize the interaction of the first twenty best ACE inhibitors obtained in our experiment. For this, we will click on the pose number of the respective compound and we will have access to a three-dimensional structure viewer (Figure 4), where the compound will appear in the pocket of the receptor. In the viewer, you can zoom in on the complex using the mouse scroll and rotate using the mouse drag feature.

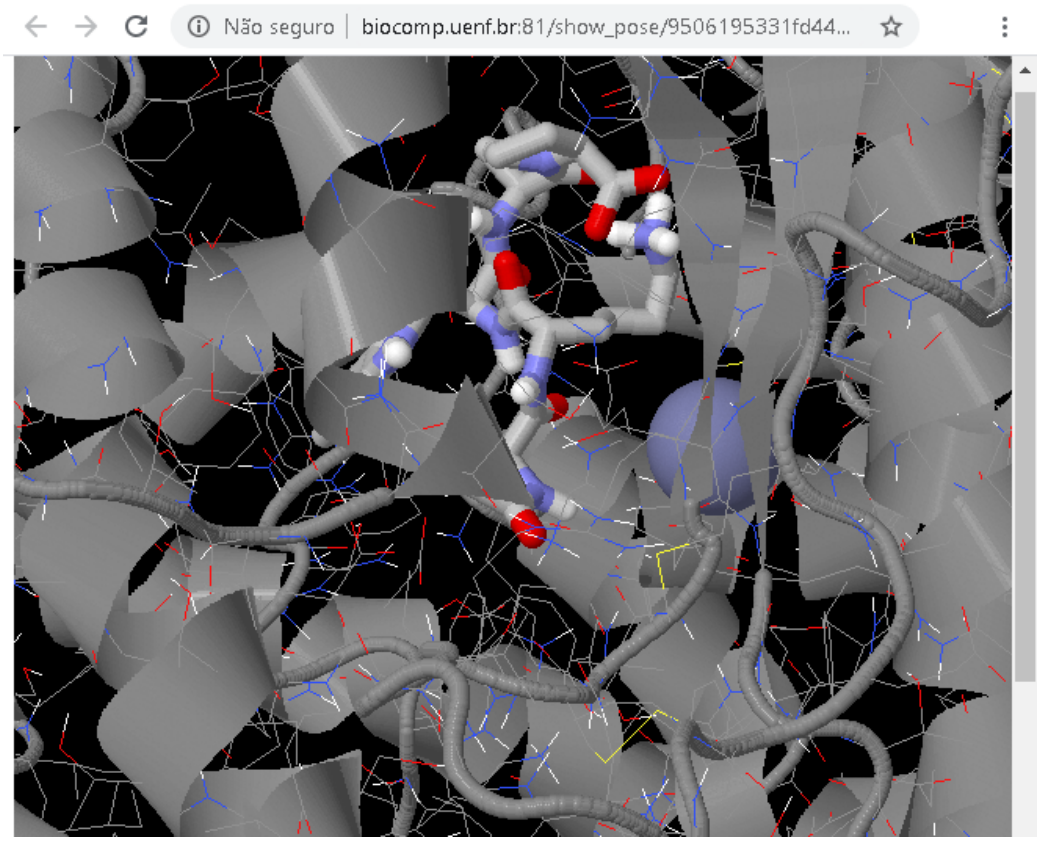

Figure 4: Viewing a complex. In HTP SurflexDock, ligands are represented by thicker sticks style and the receptor is represented by cartoon style.

An important quality factor in ACE inhibitor studies is the ability of these compounds to interact with the zinc ion present at the centre of the active site. Because, this ion represent more of a 50% of the interaction force of ACE with its substrate (Jiang

et al., 2019). So let's look at this interaction of the top 20 compounds with this. However, ions are not represented in the default viewer setting. To enable their visualization, we will select all ions Zn present in the complex by right clicking anywhere in the viewer, selecting the menus "**Select**" and "**Element"**. Finally, click on **Zn**. After that, we will change the representation of the selected Zn ion: Right click, select the "**Style**" and "**Scheme**" menus and click on the item "**CPK Spacefill**". Notice how these inhibitors interact with the zinc ion present in ACE.

4.2 Compare performance of receptor ensemble conformations

We can compare the performance of the ligands between the 4 different receptor conformations present in the ensemble through a table generated by HTP SurflexDock. To access it, let's click on the **Download** link in the item "**Download Top 20 Results**". The downloaded table contains the performance of the top 20 compounds for each ensemble conformation (Figure 5).

HTP SurflexDock results
Top 20 Results Generated for Each Protein Conformation of hACEcOk.pdb

Conformation Control		Conformation 220.000		Conformation 1720.000		Conformation 3640.000	
Compound	ΔG (kcal/mol)	Compound	ΔG (kcal/mol)	Compound	ΔG (kcal/mol)	Compound	ΔG (kcal/mol)
BPP3	-10.49	BPP5a	-12.64	BPP5_3	-10.23	BPP5_3	-10.61
new_ligand_xx15	-10.41	BPP5_3	-11.95	new_ligand_xx15	-9.6	BPP5a	-9.98
new_ligand_xx37	-10.25	new_ligand_xx15	-9.78	BPP5a	-9.45	new_ligand_xx15	-9.68
new_ligand_xx24	-10.0	BPP3	-9.51	new_ligand_xx24	-8.82	BPP3	-9.46
BPP5_3	-9.8	new_ligand_xx24	-8.76	new_ligand_xx33	-8.62	new_ligand_xx24	-8.35
decoys_xx529	-9.79	new_ligand_xx37	-8.71	new_ligand_xx37	-8.31	decoys_xx475	-8.12
BPP5a	-9.77	decoys_xx1617	-8.31	BPP5_4	-8.11	new_ligand_xx23	-8.06
new_ligand_xx48	-9.34	new_ligand_xx23	-8.04	BPP3	-8.07	decoys_xx1507	-8.04
new_ligand_xx25	-9.29	decoys_xx1262	-7.94	new_ligand_xx34	-8.04	decoys_xx109	-7.83
new_ligand_xx33	-9.28	decoys_xx529	-7.92	new_ligand_xx30	-7.56	new_ligand_xx37	-7.58
decoys_xx1273	-9.26	decoys_xx1596	-7.85	decoys_xx979	-7.55	decoys_xx1273	-7.45
new_ligand_xx34	-9.08	decoys_xx979	-7.84	decoys_xx370	-7.41	decoys_xx979	-7.45
decoys_xx986	-8.94	decoys_xx977	-7.82	decoys_xx1596	-7.39	decoys_xx1408	-7.43
decoys_xx974	-8.77	decoys_xx998	-7.7	decoys_xx529	-7.36	new_ligand_xx22	-7.41
decoys_xx229	-8.65	decoys_xx01	-7.68	decoys_xx1273	-7.35	new_ligand_xx34	-7.39
decoys_xx1218	-8.61	decoys_xx398	-7.62	decoys_xx524	-7.35	decoys_xx1142	-7.34
decoys_xx1659	-8.54	decoys_xx110	-7.62	LPR	-7.32	new_ligand_xx48	-7.33
decoys_xx1262	-8.53	decoys_xx76	-7.6	decoys_xx109	-7.31	decoys_xx110	-7.26
decoys_xx1596	-8.48	new_ligand_xx33	-7.57	decoys_xx475	-7.29	decoys_xx748	-7.21
BPP5_4	-8.47	new_ligand_xx22	-7.55	decoys_xx23	-7.29	new_ligand_xx33	-7.19

Figure 5: A table representing the "Top 20" Compounds in each conformation of ensemble.

5. Result refinement

HTP SurflexDock calculates 10 to 30 dockings of each compound for each ensemble conformation and presents only the lowest energy solution of each compound in the table in **Compound Ranking** section. However, calculated dockings are mathematical solutions that do not always faithfully represent a biological solution. So we can explore other docking solutions, HTP SurflexDock allows us to select some

complexes and recalculate docking for those complexes. In this case, at the end of the calculation you will be redirected to a new results page, which will present all the solutions found for the selected compounds. We will have access to the HTP SurflexDock filtering page (Figure 6), on this page we can select the number of docking experiments for each compound previously selected. Let's change the sliding menu to make 30 evaluations and press "**Submit"**.

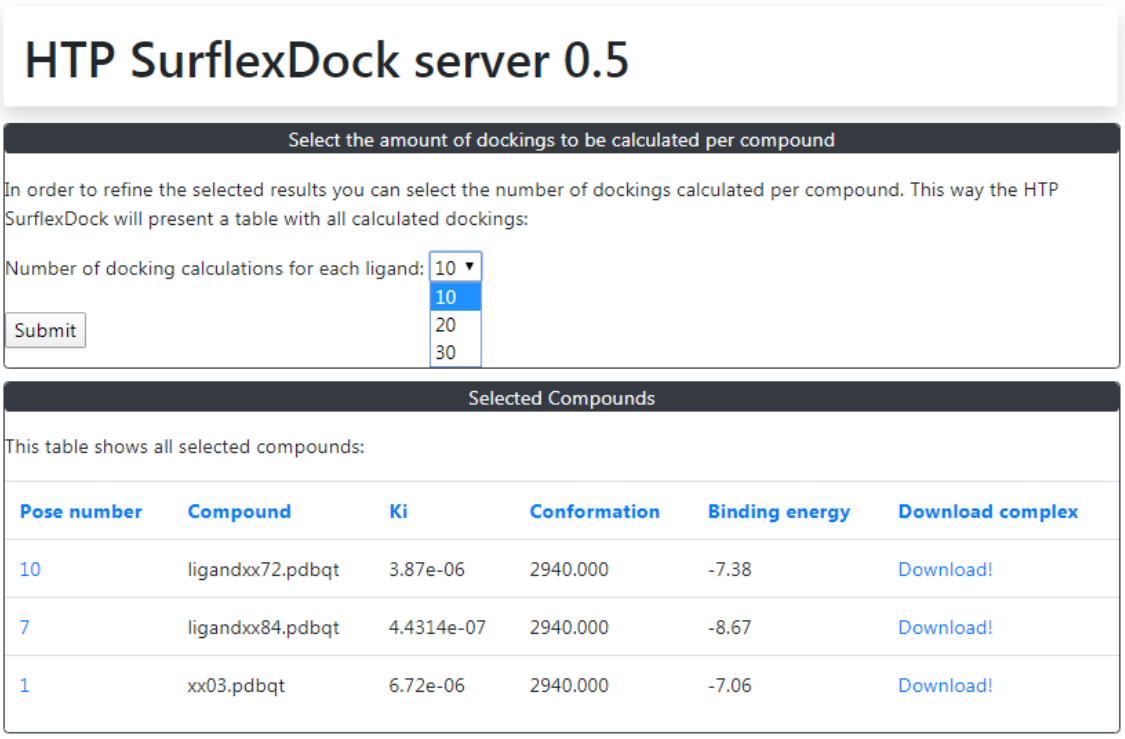

Figure 6: Selecting the number of dockings calculated per compound in the refinement process.

With the submission of this new experiment, You will access the results page. This new results page has a layout very similar to the previous result layout but with some differences: the first one is in **Ki of the compounds in each cluster** section which has boxplot graphics to present the top 10 dockings performance of each compound in the interaction with every conformation of the ensemble and **Compound Ranking** section has a same table form but containing all calculated dockings for every selected compound. This new results page has a layout very similar to the previous result page but with some differences: the first one is in **Ki of the compounds in each cluster** section which has boxplot graphics to present the top 10 dockings performance of each compound in the interaction with the conformations of the ensemble and **Compound Ranking** section has a table containing all calculated dockings for all selected compounds. As in previous results page, **Experiment Detail** section presents the information of project submission.

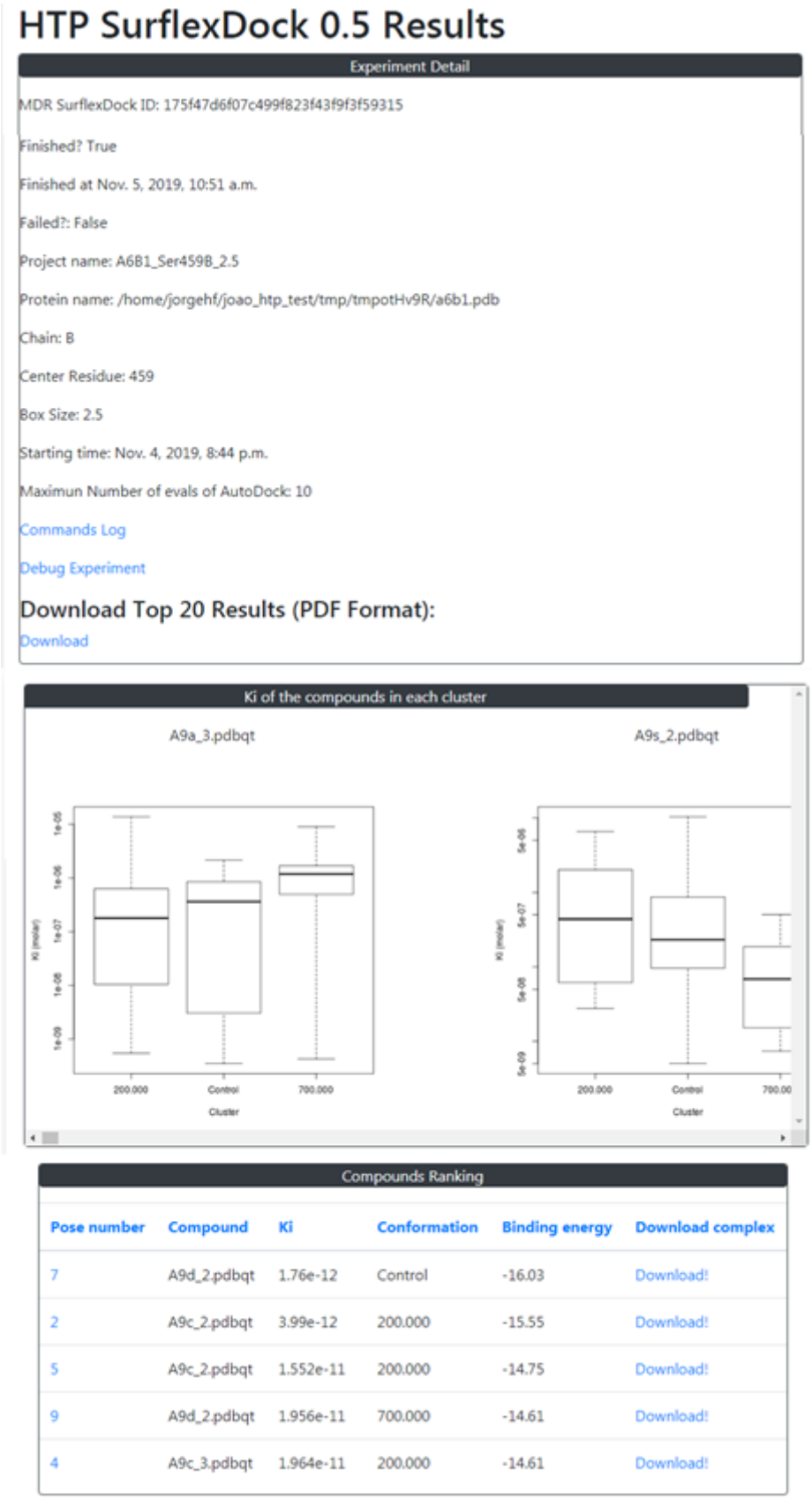

Figure 7: Results page of refinement procedure.

Finally, let's go and visualize the calculated complexes by clicking on the pose number and download the results in PDB format by clicking on the link in the folder

column. You can save these results for viewing in other tools like Pymol or AutoDockTools.

6. References

Bateman, B. T., Patorno, E., Desai, R. J., Seely, E. W., Mogun, H., Dejene, S. Z., … Huybrechts, K. F. (2017). Angiotensin-converting enzyme inhibitors and the risk of congenital malformations. *Obstetrics and Gynecology*, *129*(1), 174.

Evans, M., Carrero, J.-J., Szummer, K., Åkerblom, A., Edfors, R., Spaak, J., … Jernberg, T. (2016). Angiotensin-converting enzyme inhibitors and angiotensin receptor blockers in myocardial infarction patients with renal dysfunction. *Journal of the American College of Cardiology*, *67*(14), 1687–1697.

Guan, S., Han, W., Zhang, H., Wang, S., & Shan, Y. (2016). Insight into the interactive residues between two domains of human somatic Angiotensin-converting enzyme and Angiotensin II by MM-PBSA calculation and steered molecular dynamics simulation. *Journal of Biomolecular Structure and Dynamics*, *34*(1), 15–28.

Jalkute, C. B., Barage, S. H., Dhanavade, M. J., & Sonawane, K. D. (2013). Molecular dynamics simulation and molecular docking studies of angiotensin converting enzyme with inhibitor lisinopril and amyloid beta peptide. *Protein Journal*, *32*(5), 356–364. https://doi.org/10.1007/s10930-013-9492-3

Jiang, Z., Zhang, H., Bian, X., Li, J., Li, J., & Zhang, H. (2019). Insight into the binding of ACE-inhibitory peptides to angiotensin-converting enzyme: a molecular simulation. *Molecular Simulation*, *45*(3), 215–222. https://doi.org/10.1080/08927022.2018.1557327

Masuyer, G., Schwager, S. L. U., Sturrock, E. D., Isaac, R. E., & Acharya, K. R. (2012). Molecular recognition and regulation of human angiotensin-I converting enzyme (ACE) activity by natural inhibitory peptides. *Scientific Reports*, *2*, 1–10. https://doi.org/10.1038/srep00717

Natesh, R., Schwager, S. L. U., Sturrock, E. D., & Acharya, K. R. (2003). Crystal structure of the human enzyme – lisinopril complex. *Nature*, *421*(1995), 1427–1429. https://doi.org/10.1038/nature01370

Zhang, C., Wu, S., & Xu, D. (2013). *Catalytic Mechanism of Angiotensin-Converting Enzyme and E ff ects of the Chloride Ion*. (Md).