EXAMPLE OF USE: LigRMSDLink reference:
**https://ligrmsd.appsbio.utalca.cl/results/20191223220522/**

LigRMSD is a free web server for automatic matching and RMSD calculations among identical or similar chemical compounds. For calculations, LigRMSD requires a set of molecules in MOL format (at least two). Once the user uploads the molecules, one of them (at least) must be selected as reference. Then, all molecules will be compared with the chosen references. It is essential to notice that the user must assure that MOL files submitted were correctly generated and the structures standardized. Special attention for the bond types and atom valences of molecules extracted from PDB structures; the PDB format does not include this type of information.

As it is mentioned in the manuscript, there are two assays available in LigRMSD: (i) matching and RMSD calculation between two compounds; and (ii) matching an RMSD calculation of one (or more) query against several reference compounds (data set). For any of these options, two different matching methods are configured. First, strict matching identifies equivalent atoms by considering identical atom types and bonds. Meanwhile, flexible matching contains the strict match and also may include different atom and bond types, only where their inclusion does not affect the chemical graph of the molecules.

Matching and RMSD calculation between two or more compounds:

In this first example, we are interested in comparing the reproduced protein-ligand docking of the structures of three thrombin-ligands complexes (81A, I50 and NLI) of the PDBids 1T4V, 2R2M and 3LDX (Figure 1).
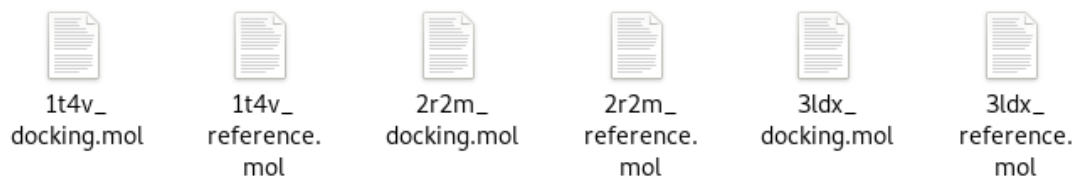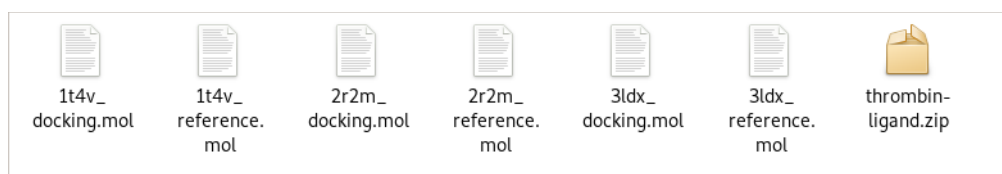

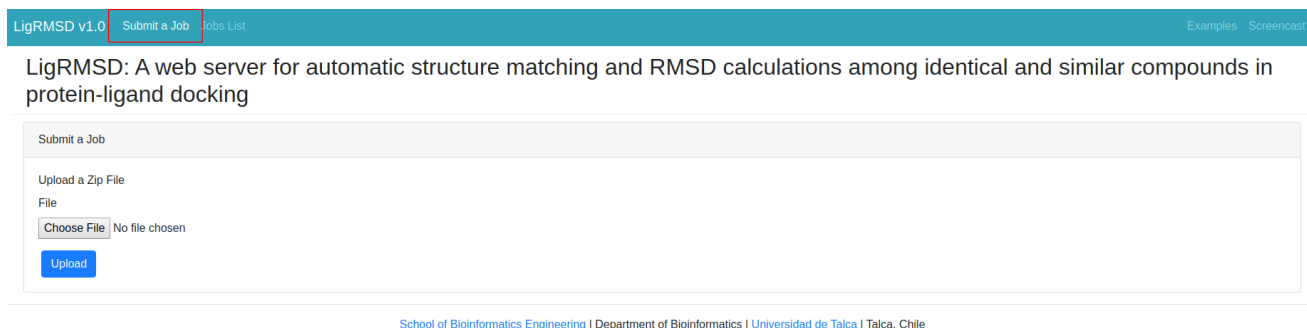**Figure 1:** Set of molecules (thrombin-ligands docked and references) in MOL format

It is essential to mention that LigRMSD will only process molecules submitted in MOL format. Also, the server requires that all MOL files must be compressed as a ZIP file (Figure 2 and Figure 3).
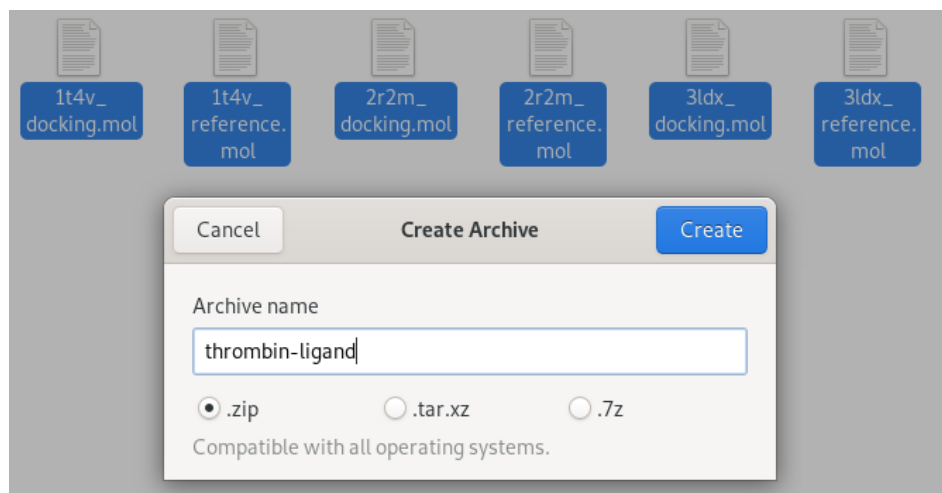
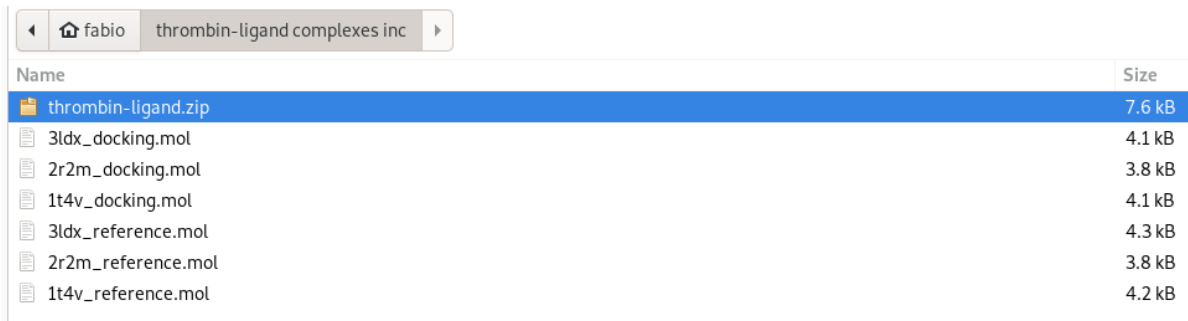**Figure 2:** Creating the ZIP file

Figure 3: The ZIP file created
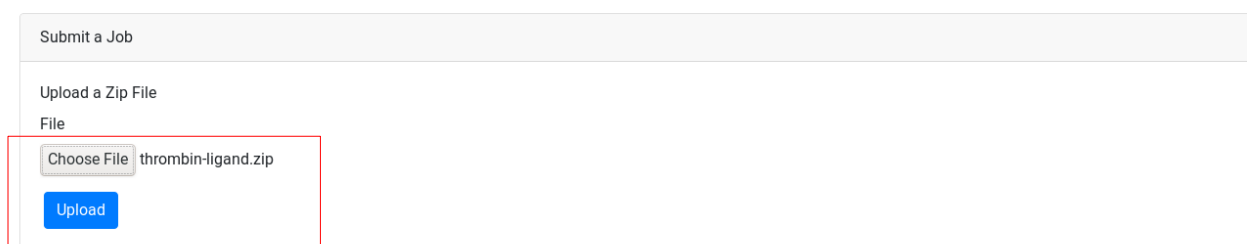

Figure 4: Submitting a new job

The next step is to upload the ZIP file to the server. Figure 4 shows the link for the submission of a new job.

Now the user can press the button "Choose File" and select the previously created ZIP file


(Figure 5 and Figure 6).

Figure 5: Selecting the ZIP file

Figure 6: Uploading the ZIP file

Once the user presses the upload button, the ZIP file is submitted into a temporary directory and then is relocated into the server. The next step to submit a job in LigRMSD is the selection of reference (s). For this, the web-server shows a 2D representation of each molecule included in the ZIP file submitted (Figure 7). Here, the user can select one or more
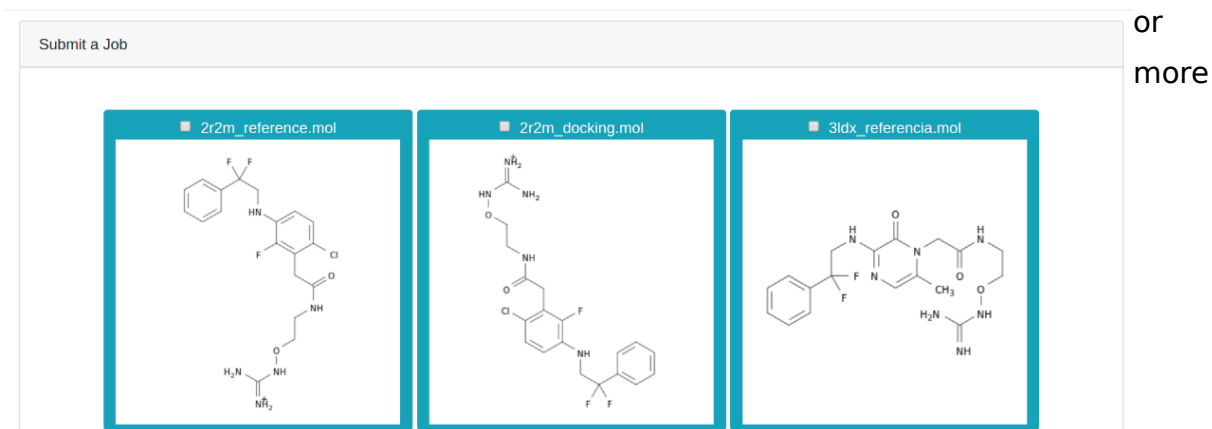
references for the comparisons (Figure 8).

Figure 7: 2D representation of the MOL files submitted
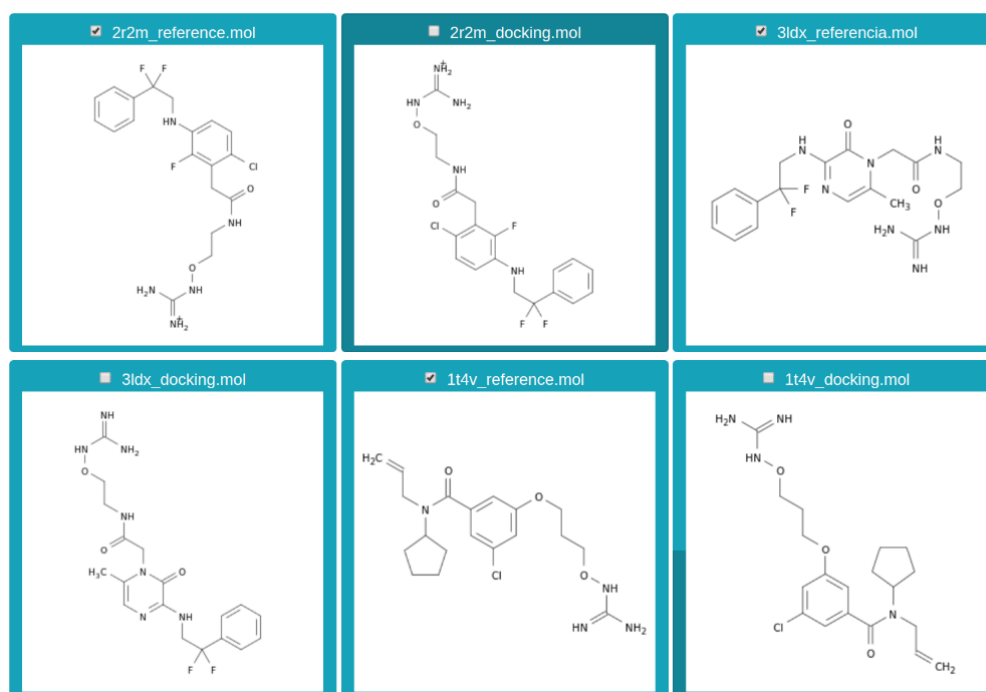
Figure 8: Selecting the references

Below the selection box of the references, the user can complete the form identification for the job (Figure 9). Even when these parameters are not mandatory, it is commendable to include representative job names. At the end of the measures, this information will be sent to the user e-mail. Also, the user must include a "Limit" value. This parameter represents the minimum structural match that must have a pair of molecules for calculation of the RMSD.

Figure 9: Job identification

Once completed the form and pressed the "execute" button thus, LigRMSD redirects the browser to the "wait page". This will refresh every 10 seconds the page until the job is finished. After that, the website will be redirected to the results page.


Figure 10: "wait page"

On the wait page, we recommended bookmark the link denoted by a red box in Figure 10. Is very important save the URL because is unique and self-generated for LigRMSD. Thus with this URL, the user can visualize the results for the job. By the way, if you put an email on the submit job form (figure 9) you should receive the link in that email box.

Visualizing the results

When the calculation was done, the user can visualize the results accessing to URL self-generated for the job. The page of results has four sections. The first section (denoted by a red box in Figure 11) shows the reference molecules selected by the user. If an error is detected on any reference, this molecule would be mentioned in this section as a list next to "References with an error". For exploration of a particular reference result, the user must click on the name of the reference. The second section, named "Matching Method", is denoted by a green box in Figure 11. Here the user can visualize all the pair of molecules evaluated using strict or flexible methods. The third section is the Multi-Reference (center side of the green box in Figure 11). This section is used for comparing an extensive database of compounds against one or several query molecules (Figure 12). In this section, the user can view all the comparisons simultaneously.

The last section is "Summary" (figure 13). This section has a tableau with all results for RMSD, percent of reference match, percent of molecule match, separated by the method match. The tableau can export the results as PDF format pressing the PDF button.
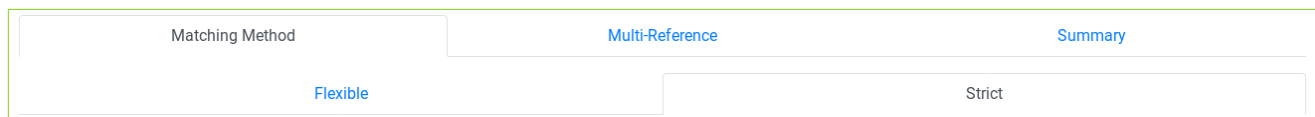
Figure 11: The four sections of results are denoted into the red and green boxes

Figure 12: Multi-Reference results

REF	Molecule	Type	RMSD	Percent of ref match	Percent of molecule match
2r2m_docking.mol	3ldx_docking.mol	flexible	0.80	100.00	100.00
		strict	0.59	40.00	40.00
	1t4v_reference.mol	flexible	1.56	60.00	66.67
		strict	1.12	23.33	25.93
	1t4v_docking.mol	flexible	1.52	60.00	66.67
		strict	2.74	23.33	25.93
2r2m_reference.mol	3ldx_docking.mol	flexible	0.61	100.00	100.00

Figure 13: Summary results
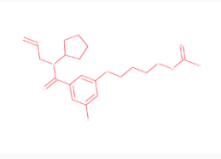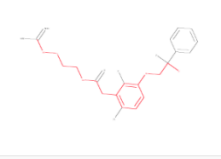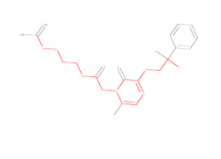
If the user press on one of the references located in the first section of results (red box in Figure 11), then this reference will be shown as a 2D representation and in a 3D viewer (Figure 14).
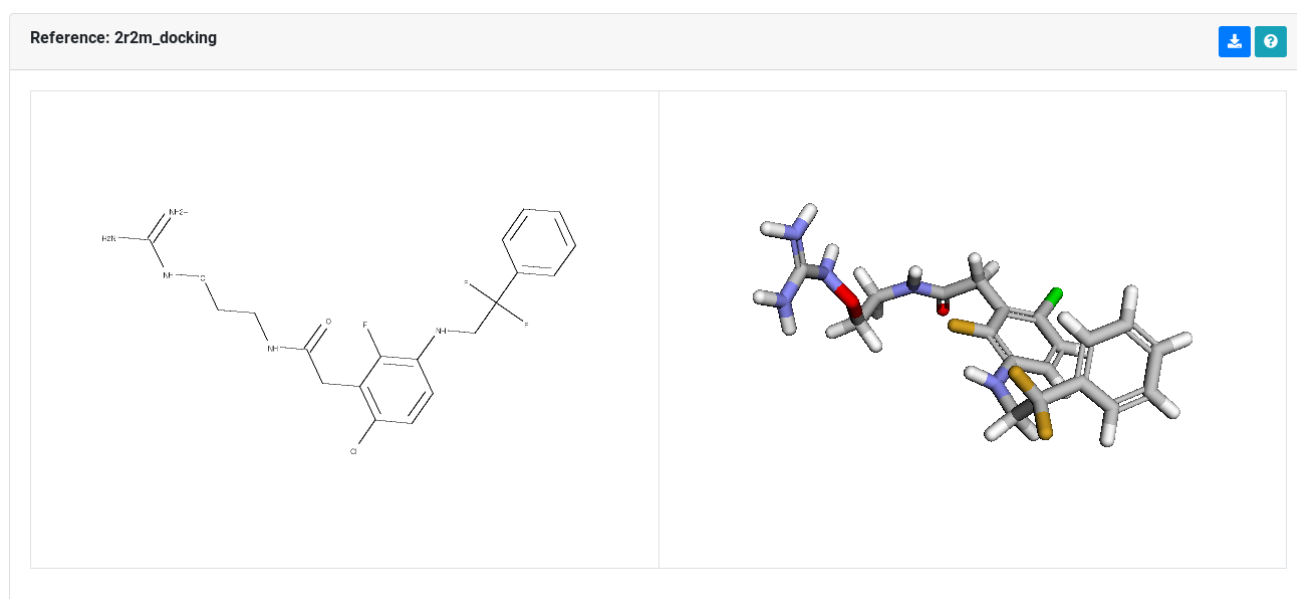
Figure 14: 2D representation and 3D viewer of a reference selected

Once selected one reference, the results can be filtered for both, the strict and the flexible method. As is shown in Figure 15 (flexible) and Figure 16 (strict), all the molecules are listed and represented as 2D structures. Those zones that were included in the

maximum common substructure (MCS) between the molecule and the selected reference are remarked in red. The information about the percentage of the match and the RMSD values are delivered in the same list. 'Percent of Ref Match' refers to the percent of common graphs between the docked and reference compounds concerning the total number of atoms of the reference compound. 'Percent of Molecule Match' refers to the percent of common graphs between the docked and reference compounds regarding the total number of atoms of the docked compound.

Molecule	Image	RMSD	Percent of ref match	Percent of molecule match	Details	Download
1t4v_docking		1.33	100.00	100.00	View	Download
2r2m_docking		1.56	66.67	60.00	View	Download
3ldx_docking		1.45	66.67	60.00	View	Download

Figure 15: The matching results for several molecules against reference (flexible method)

As we mentioned before, a "Limit" value of matching must be defined on each query. In these examples, we used an amount of 30%, denoting the minimum percent of structural matching allowed to realise the calculation of the RMSD. Because this value is related to the MCS, different results might be obtained between the flexible and the strict method. Thus and for comparisons purposes, LigRMSD will show the results for the two methods if at least in one of them, the "Limit" threshold is overcome. In this case, the molecules 2r2m_docking and 3ldx_docking, shows a 66,67 % of 'Percent of Ref Match' considering the flexible method (Figure 15). This value decreases to 23,33 % when the strict method is selected (Figure 16). It should be noted that low values of 'Percent of Ref Match' might bring unsatisfactory results.

Molecule	Image	RMSD	Percent of ref match	Percent of molecule match	Details	Download
1t4v_docking		1.33	100.00	100.00	View	Download
2r2m_docking		1.12	25.93	23.33	View	Download
3ldx_docking		5.28	25.93	23.33	View	Download

Figure 16: The matching results for several molecules against reference (strict method).

By clicking on the View link depicted on the Figure 16, the user can view a detailed comparison between a pair of compounds (Figure 17) or download the structures analyzed.
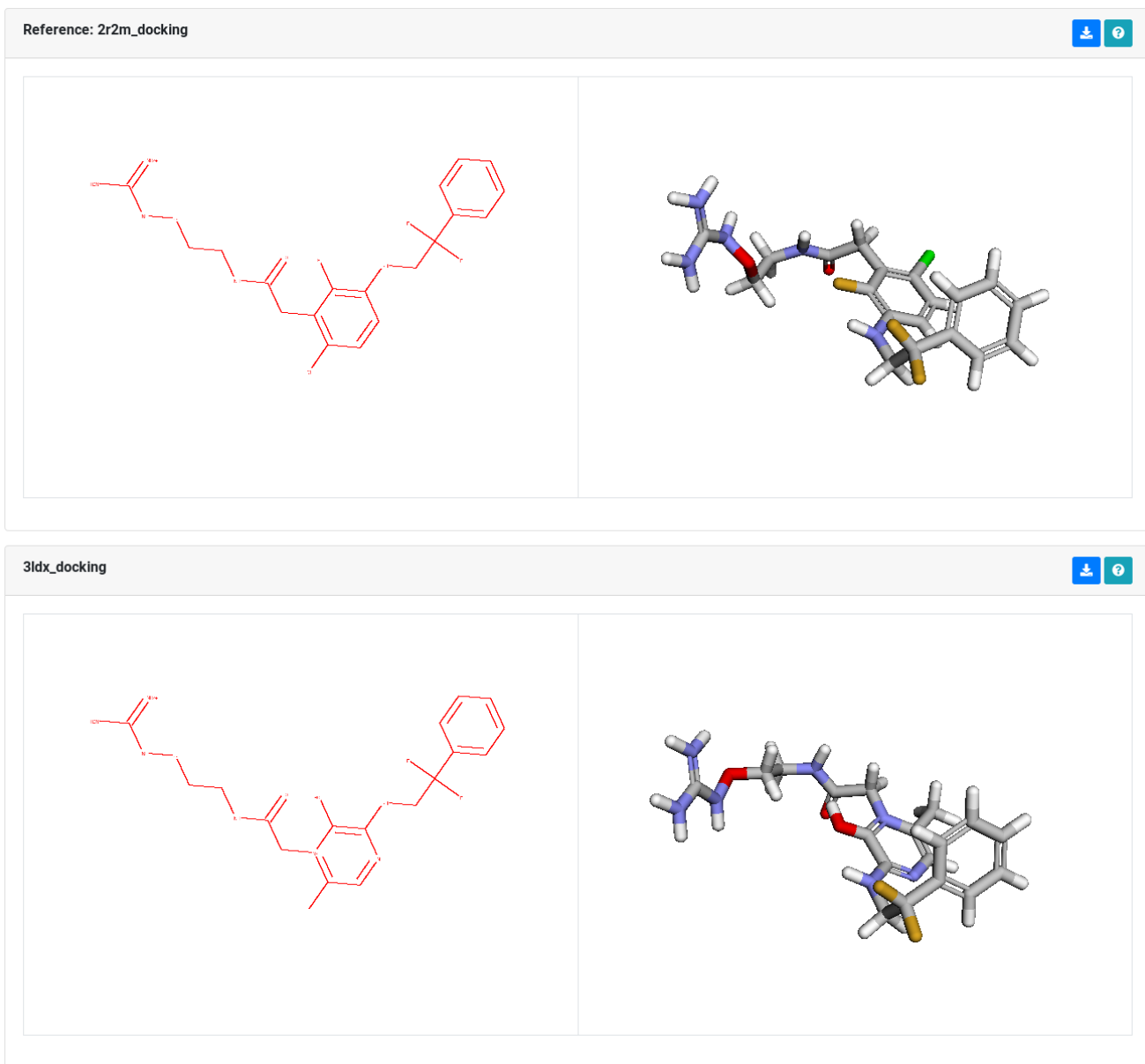
Figure 17: Detailed evaluation between two compounds.

On the same page below the box evaluation between two compounds, LigRMSD offers a 3D visualization (figure 18) with the same two compounds in a superposition view.

Red: 2r2m_docking.mol | Blue: 3ldx_docking.mol

Figure 18: 3D Viewer superposition view

Next to is the Results box (figure 19) with RMSD value and the percentage of reference and molecule match.

Results

RMSD	Percent of ref match	Percent of molecule match
0.80	100.00	100.00

Click here for details for the matched results

BACK

Figure 19: The results match

Finally, if click for more details can found on each comparison and the user can visualize the detailed atoms matched in the MSC (Figure 20)

Visualization

Reference		Molecule	

Line Number	Reference 2r2m_docking Elements	X	Y	Z	Molecule 3ldx_docking Elements	X	Y	Z
1	C	19.235	-15.692	22.6289	N	19.0695	-15.9117	22.5679
2	C	19.7634	-14.8997	21.597	C	19.7184	-15.2767	21.6872

Figure 20: Detailed matched atoms between two compounds.