The compressed directory we provided contains 3 directories:
- files
- pdb
- bin

files directory contains:
• 1KRL_D_surf.csv
• 1KRL_C_surf.csv
The computed surfaces for the sample pdb.

• Zernike_inv_protein_1KRL_D_patch_0.txt
• Zernike_inv_protein_1KRL_C_patch_0.txt
Zernike descriptors for a sample patch, which is centered in the first point of the surface.

• All_points_zernike_invariant_1KRL_D_verso_down.dat
• All_points_zernike_invariant_1KRL_C_verso_up.dat
Zernike descriptors for all the point of the surfaces.

• surf_newrot_1KRL_C.txt
• surf_newrot_1KRL_D.txt
Surfaces with the binding propensity.

pdb directory contains the sample pdb files.

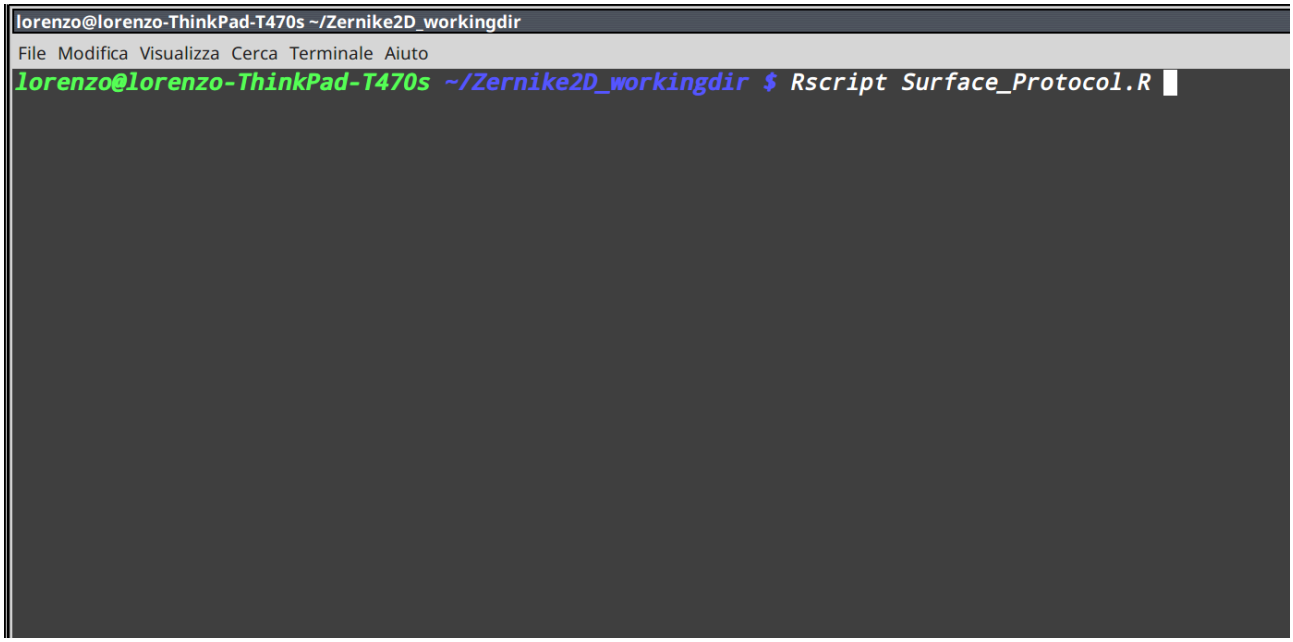
bin directory contains all functions to calculate Zernike 2D descriptors.

dms.zip is a compressed package to install the *dms* software.

##

This version of the Zernike 2D software takes as an input a pdb file.

Running the following R script, the molecular surface for all the pdb files contain in the ***pdb*** directory are evaluated.

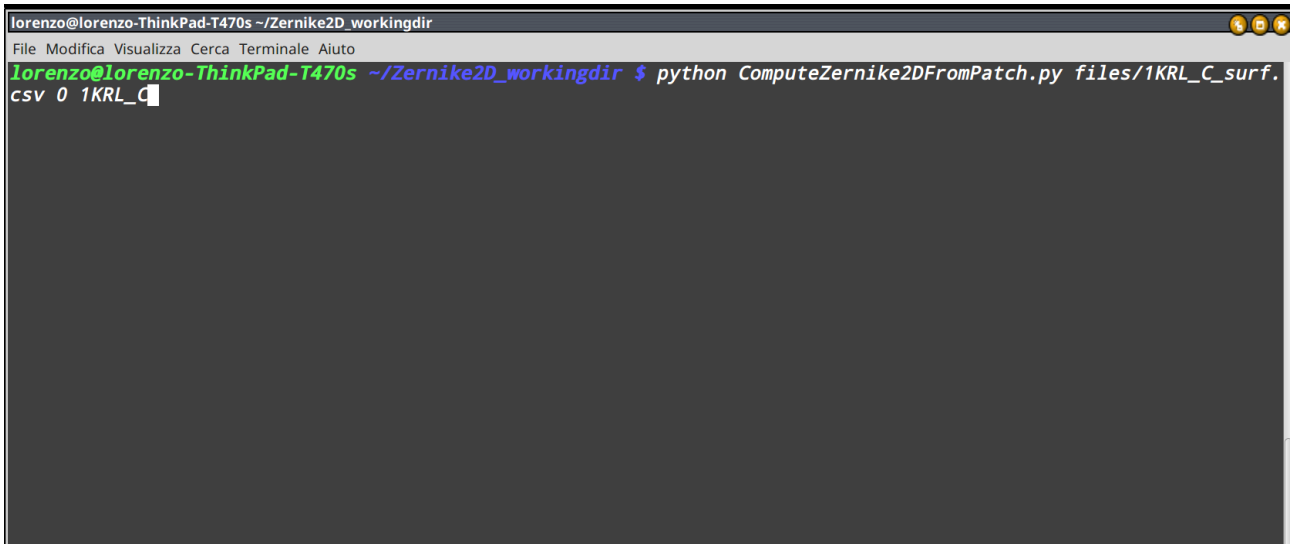


The sample molecular surface are in the ***files*** directory.

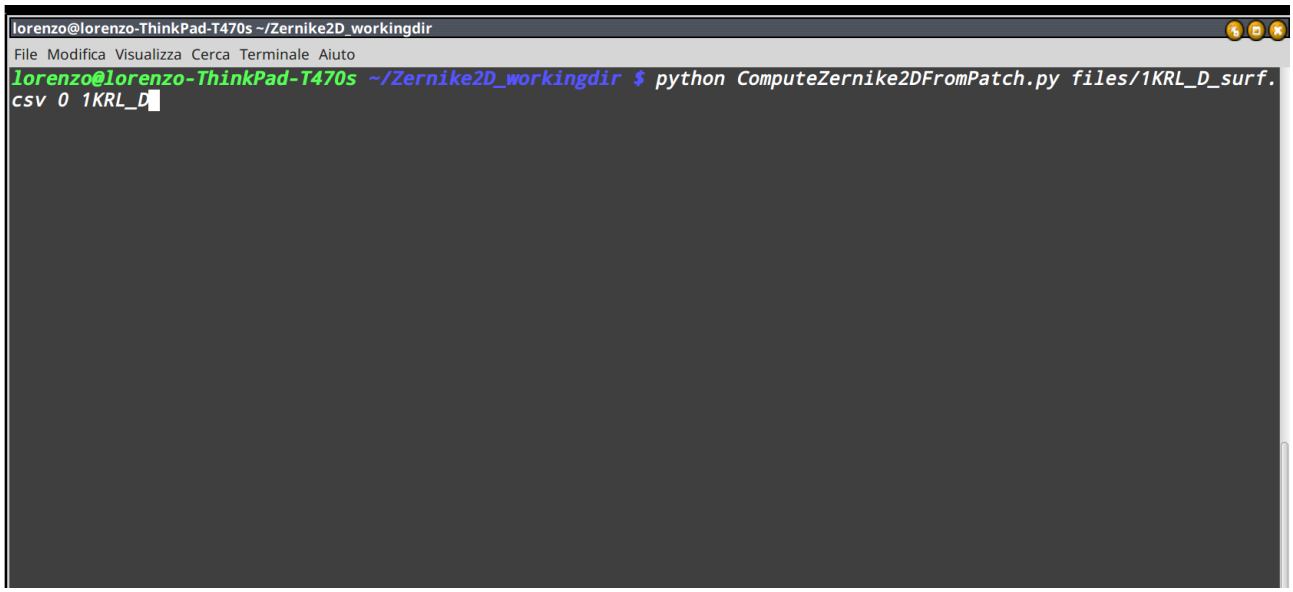
The python script “ComputeZernike2DFromPatch.py” evaluates the Zernike descriptors of a patch of radius Rs = 6 A, centered around a given point of the surface.

In this case:
files/1KRL_C_surf.csv —> the surface.
0 —> point of the surface around which patch is centered.
1KRL_C —> label for the output.

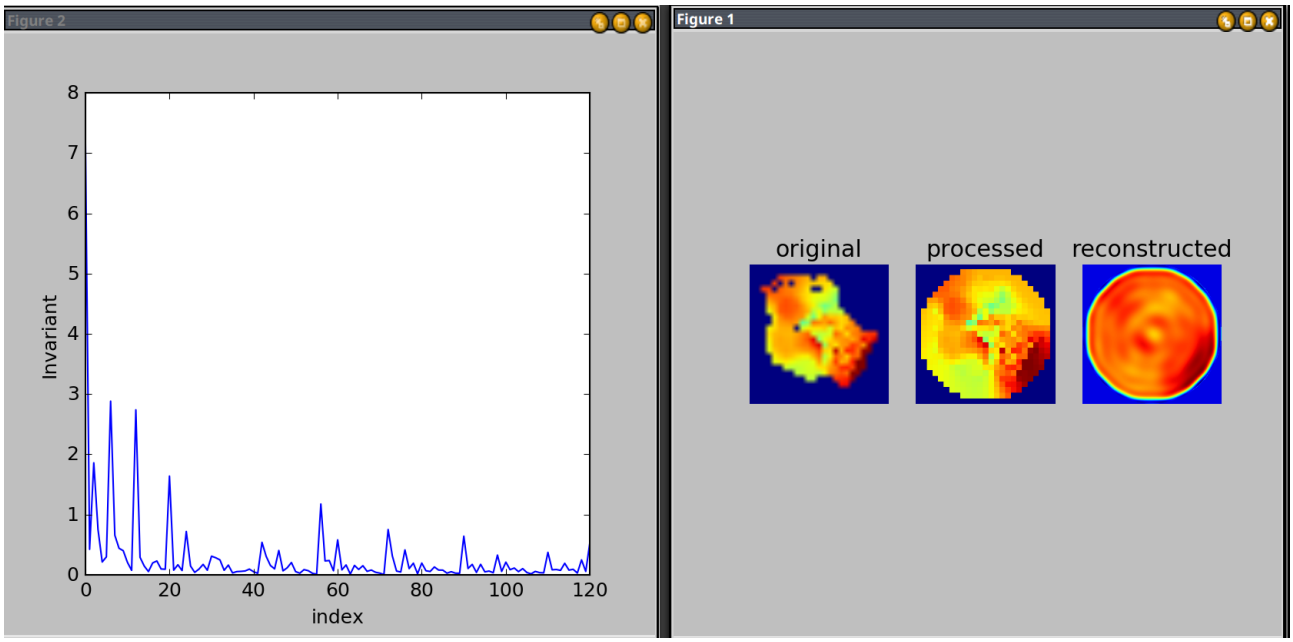
Processing 1KRL_C_surf.csv file.



Processing 1KRL_D_surf.csv file.



This is the output of the previous command:



The invariant Zernike descriptor (plot on the left) are saved in the files:

- Zernike_inv_protein_1KRL_D_patch_0.txt
- Zernike_inv_protein_1KRL_C_patch_0.txt

Running the code for alla the point of the surface, we obtain the following files:

- All_points_zernike_invariant_1KRL_D_verso_down.dat
- All_points_zernike_invariant_1KRL_C_verso_up.dat

Running the following script:

we obtain the “binding propensity” information for each point of the surface, which is defined as the minimal distance between Zernike descriptors associated to a point of surface “C” and all points belonging to the surface “D” (and viceversa).

The output of this script are:

- surf_newrot_1KRL_C.txt
- surf_newrot_1KRL_D.txt

Running the following command line:

the script performs the smoothing procedure and evaluates the AUC of the ROC Curve comparing the real and the predicted “binding site propensity” groups.

The output for this example is the following:

```
"**********"
"The process is running..."
"**********"
"The complex analyzed is: 1KRL"
"The AUC of the first protein is: 0.75"
"The AUC of the second protein is: 0.65"
"**********"
```