BIOINFORMATION open access

Supplementary material:

Scripts:

Configuration File to run PyPLIF (config.txt)

Lines **Script**

```
1   protein_reference ER_site.mol2
2   ligand_reference OHT.mol2
3
4   protein_ligand_folder results
5   residue_of_choice LEU327 TYR328 SER329 GLU330 SER341 MET342 MET343 GLY344 LEU345 LEU346 THR347
    ASN348 LEU349 ALA350 ASP351 ARG352 GLU353 LEU354 VAL355 MET357 LEU379 GLU380 CYS381 ALA382 TRP383
    LEU384 GLU385 ILE386 LEU387 MET388 ILE389 GLY390 LEU391 VAL392 ARG394 SER395 LEU402 LEU403 PHE404
    ALA405 LEU408 LEU410 GLY415 VAL418 GLU419 GLY420 MET421 VAL422 GLU423 ILE424 PHE425 LEU428 ILE514
    HIS516 MET517 SER518 ASN519 LYS520 GLY521 MET522 GLU523 HIS524 LEU525 TYR526 SER527 MET528 LYS529
    CYS530 LEU536 LEU539
6   output_file ../../pyplif_result/ligandER_tc.csv
```

Shell Script to perform hydrogen bond to the residue ASP351 filtering (tc_cum_sorted_filter_ASP351.sh):

Lines **Script**

```
1   #!/bin/sh
2
3   # a script to choose the conformation with best Tc-IFP and score from each ligand
4   # then put them into one file and sort them (all the best ligand conf) by Tc-IFP and score.

5   rm tc_all.csv
6   for i in $(cat ligand.lst)
7   do
8      best=`awk '{if (substr($4,103,1)==1) print $0}' pyplif_result/${i}_tc.csv | sort -n -k3rn -k2n |
9   head -n1`
      # to check which csv is missing
10    if [ -z "$best" ]; then
11       echo $i
12    fi
13    echo $best >> tc_all.csv
14  done
15
16  sort -n -k3rn -k2n tc_all.csv > tc_all_sorted.csv
```