ProtInter: Protein Interaction Calculator

Maxime Borry[a]

[a] *Paris Diderot University,*

September 21, 2017

ProtInter is a tool developed to predict non-covalent protein interactions such as hydrophobic interactions, disulphide bridges, hydrogen bonds, ionic interactions, aromatic interactions, and cation-pi interactions.

keywords: Protein, Interaction, Python, non-covalent.

1 Introduction

Non-covalent interactions in protein play a major role in the establishment of three-dimensional protein structure. After the determination of the 3D structure of a protein, either by modeling, NMR, or crystallography, the coordinates of each atom of a protein are saved in a Protein Data Bank (pdb) file. However, the encoding of a protein in the pdb format saves the atoms composing the protein, and their coordinates, but do not save non-covalent interactions. The accounting of these non-covalent interaction plays a crucial role in further structural biology analysis, such as remote homology studies, protein fold recognition, protein modeling... For example, in remote homology, the detection of homologs can be performed using the conservation of side chains interactions Bhaduri et al. (2004). Here I present tool to compute several of these non-covalent interactions: hydrophobic interactions, disulphide bridges, hydrogen bonds, ionic interactions, aromatic-aromatic interactions, aromatic-sulphur interactions, and cation-pi interactions. This tool, which I name ProtInter is replicating in part the one published by Tina et al. (2007).

2 Material and Methods

2.1 Structural criteria

These criteria were used to infer the following non-covalent interactions as in Tina et al. (2007):

1. **Hydrophobic interactions:** The residues `ALA, VAL, LEU, ILE, MET, PHE, TRP, PRO, TYR` participate in the interaction if one their atom in their side-chain is situated less than 5 Å from another.

2. **Disulphide bridges:** A pair of `CYS` are considered part of a disulphide bridge if their sulphur atoms are located less than 2.2 Å one from another.

3. **Ionic interactions:** A pair of `ARG, LYS, HIS, ASP, GLU` participates in the interaction if at least one of their atom `O` or `N` is located less than 6 Å from another.

4. **Aromatic-aromatic interactions:** A pair of `PHE, TRP, TYR` participates in the interaction if the center of mass of their phenyl ring is located between 4.5 and 7 Å from the other.

5. **Aromatic-Sulphur interactions:** A `CYS` sulphur atom can interact with an aromatic amino-acid `PHE, TYR, TRP` if it is located less than 5.3 Å from the center of mass of the phenyl rings.

6. **Cation-pi interactions:** A cationic side chain (`LYS, ARG`) can participate in the interaction with an aromatic side chain (`PHE, TYR, TRP`) if they are less than 6 Å from each other.

7. **Hydrogen bonds:** Two cutoffs are used depending on the donor-acceptor pair. In the case of the Nitrogen-Oxygen Pair, the cutoff is 3.5 Å while in the case of of Sulphur, it is 4 Å .

2.2 Programming

All the programming was realized using Python 3 and the library Biopython (Cock et al. (2009)) and its pdb and structure module for parsing pdb files.
ProtInter, its documentation and source code, is available at the following address: `https://github.com/maxibor/protinter`

2.3 Comparison with PIC

The results of ProtInter were compared with the ones of PIC (Tina et al. (2007)) using the 1BTC file, the structure of the soybean beta-amylase determined by crystallography (Mikami et al. (1993)). Here I'll investigate a bit deeper the differences between ProtInter and PIC, of the Cation-Pi and Disulphide bridges, as other interactions calculated by ProtInter (hydrophobic, aromatic-aromatic, aromatic-sulphur, and ionic) give a same or very similar results with PIC. Hydrogen Bonds (hbond) inference also give different results, however as inferring hbond can differ very much between different methods, results won't be compared here.

- Disulphide bridges: PIC inferred 6 disulphide bridges (table 1) whereas ProtInter finds none. At first this discrepancy can be confusing, however, looking at the

Table 1: Disulphide bridges inferred by PIC

Pos1	Res1	Chain	Pos2	Res2	Chain	Distance(A)
288	CYS	A	343	CYS	A	0.00
288	CYS	A	448	CYS	A	0.00
343	CYS	A	448	CYS	A	0.00
95	CYS	A	288	CYS	A	0.00
95	CYS	A	343	CYS	A	0.00
95	CYS	A	448	CYS	A	0.00

different residues and their positions, one can find that Disulphide bridges inferred by PIC probably do not exist, because of a greater distance exceding greatly the cutoff value of 2.2 Å (fig 1).

- Cation-pi: ProtInter infers all the interaction found PIC, and some more. To verify the additional interactions inferred only by ProtInter, distance measurement were taken using PyMol (fig 2), and indeed, those Cation-pi interactions are valid as distance were lower than the cutoff of 6 Å .

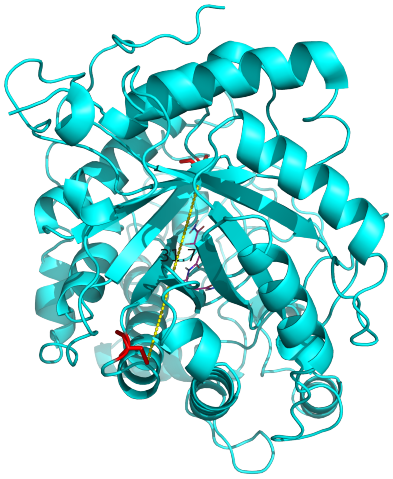

Figure 1: 1BTC Protein. Amino acids highlighted in red are the two CYS inferred by PIC as participating in a disulphide bridge (CYS 288 and CYS 343). The distance between them is 31.7 Å

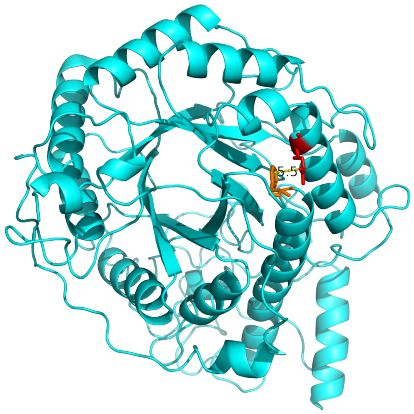

Figure 2: 1BTC Protein. Amino acids highlighted in orange and red are respectively PHE 285 and ARG 162 inferred by ProtInter as participating in a cation-pi interaction. The distance between them is 5.5 Å

3 Conclusion

Here, I have demonstrated that ProtInter performs similarly or better than PIC on the 1BTC protein, on the inference of hydrophobic, aromatic-aromatic, aromatic-sulphur, ionic, cation-pi, and disulphide bridges interactions. I also addressed some of the shortcomings of PIC, and propose ProtInter, an open-source command line interface tool, readily available.

References

Bhaduri, A., Ravishankar, R., and Sowdhamini, R. (2004). Conserved spatially interacting motifs of protein superfamilies: Application to fold recognition and function annotation of genome data. *Proteins: Structure, Function, and Bioinformatics*, 54(4):657–670.

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., Hoon, D., and L, M. J. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423.

Mikami, B., Hehre, E. J., Sato, M., Katsube, Y., Hirose, M., Morita, Y., and Sacchettini, J. C. (1993). The 2.0-A resolution structure of soybean beta-amylase complexed with alpha-cyclodextrin. *Biochemistry*, 32(27):6836–6845.

Tina, K. G., Bhadra, R., and Srinivasan, N. (2007). PIC: Protein Interactions Calculator. *Nucleic Acids Research*, 35(Web Server issue):W473–W476.