Characteritzation of Protein-Protein Interfaces and Identification of Transient Cavities for its Modulation.

Mireia Rosell and Juan Fernandez-Recio
Barcelona Supercomputing Center, Barcelona, Spain
mireia.rosell@bsc.es

Abstract- *Protein-protein interactions (PPIs) play an essential role in many biological processes, including disease conditions. Strategies to modulate PPIs with small molecules have therefore attracted increasing interest over the last few years, where successful PPI inhibitors have been reported into transient cavities from previously flat PPIfs.*

Recent studies emphasize on hot-spots (those residues contribute for most of the energy of binding) as promising targets for the modulation of PPI. PyDock is the only computational method that uses docking to predict PPIfs and hot-spots (HS) residues. Using Normalized Interface Propensity (NIP) values derived from rigid-body protein docking simulation, we are able to predict the PPIfs and HS residues without any prior structural knowledge of the complex.

We benchmarked the protocol in a small set of protein-protein complexes for which both structural data and PPI inhibitors are known. We present an approach aimed at identifying HS and transient pockets from predicted PPIfs in order to find potential small molecules capable of modulating PPIs. The method uses pyDock to identify PPIfs and HS and molecular dynamics (MD) techniques to describe the possible fluctuations of the interacting proteins in order to suggest transient pockets. Afterwards, we evaluated the validity of predicted HS and pockets for in silico drug design by using ligand docking.

We present a strategy based on MD and NIP which allows to identify cavities as potentially good targets to bind inhibitors when there is no information at all about the protein-protein complex structure.

I. Introduction

Protein-protein interactions (PPI) play an essential role in regulating biological processes, involved in the majority of diseases, highlighting the interest in protein-protein interfaces (PPIfs) as an attractive target for therapeutic intervention. A detailed structural knowledge of PPIs is needed to understand disease at molecular level, to identify new targets for therapeutic intervention and also to find small molecules capable of inhibiting PPIs [1].

It has been reported that only a few amino acids (so-called "hot-spot" residues) usually contribute to the majority of the free energy of binding. Experimental approaches typically define hot-spots (HS) as those residues that decrease binding energy in more than 1 or 2 kcal/mol upon mutation to alanine [2]. These HS residues are important in the context of drug discovery targeting PPIs because blocking them seems the only way for a small-molecule to compete with a protein-protein interaction. The reason is that PPIfs are usually large and involve higher number of atomic interactions, and hence have higher affinity as compared to protein-ligand interfaces. Other difficulties are that PPIfs do not have clear binding pockets for drug binding, and that very often, both the location of the interface and the binding mode of the PPI are not known. Successful PPI inhibitors have been reported into transient cavities from previously flat PPIfs [3].

Computational approaches such as protein-protein docking and molecular dynamics (MD) are becoming increasingly important tools in drug discovery in order to help solving the difficulties mentioned above. PyDock algorithm (a tool developed in our lab to perform protein-protein docking) is the only computational method that uses docking to predict PPIfs and HS residues when there is no structural information available of the protein-protein complex [4,5,6,7]. The method applies the fast Fourier transform algorithm to the unbound proteins of the complexes, followed by the energy-based scoring from pyDock to calculate the Normalized Interface Propensity (NIP). Using pyDock and MD techniques to suggest putative transient cavities, we present an approach addressed to targeting PPIs and to find potential small molecules capable of modulating PPIs [8].

II. Methods

A. Generation of small benchmark.

From the 2P2I database, we benchmarked the protocol in a small set of protein-protein complexes for which both structural data and PPI inhibitors are known.

B. Hot-spot prediction from protein-protein docking.

We used ZDOCK 2.1 [9] to generate 2000 rigid-body docking poses. We used the top 100 lowest-energy solutions proposed by pyDock algorithm to

calculate the Average Buried Surface (ABS) and Normalized Interface Propensity (NIP) [7]. We applied a cutoff of NIP $\geq$ 0.2 to predict HS residues.

C. *Molecular dynamics and transient cavities detection.*

In the correct predictions of HS in PPIfs, we used AMBER10 to detect transient pockets on the unbound proteins, which were selected based on the predicted HS. For each case, using Fpocket [10], we analyzed 1000 out of 10000 snapshots resulting from 10ns of simulation.

D. Ligand docking.

In those selected snapshots with a putative transient cavity, we used MAESTRO to prepare the structures for docking, as well as the inhibitors. We generated 1000 docking poses from RDOCK (flexible ligand docking).

Results

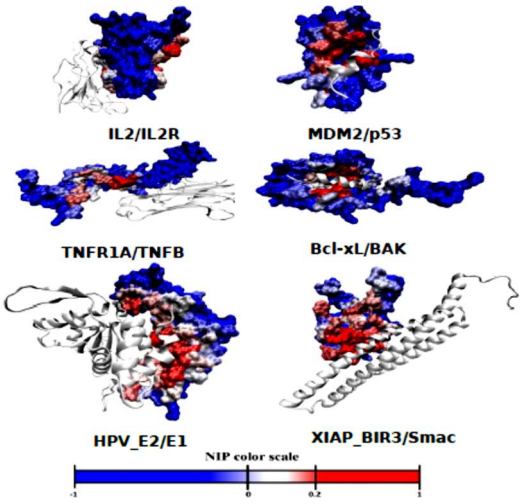
Assuming the knowledge of the PPI, 6 out of 10 cases are successful. We have focused on these cases to continue the analysis to identify transient cavities. HS and PPIfs predictions are shown in Table I (Figure 1).

[a] Number of predicted hot-spots (NIP$\geq$ 0.2). [b] Number of predicted hot-spots that are located at the PPIfs. [c] Number of predicted hot-spots that are located at the protein-inhibitor interface (PII).

* Correct predictions (these were selected for a more thorough analysis).

TABLE I

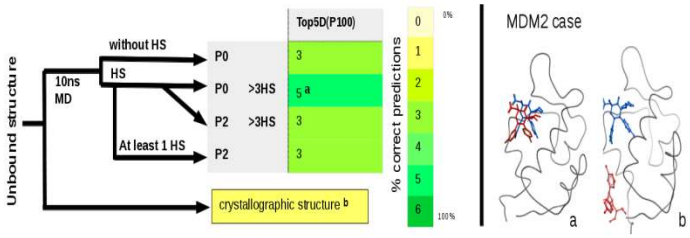
Complex	***HSpred*** [a]	***HSpred at PPIfs*** [b]	***HSpred at PII*** [c]	
Bcl-X_L/BAK	9	2	2	
Xiap_BIR3/Caspase	12	0	2	
HPV_E2/E1	21	12	7	
IL2/IL2R	4	4	4	
Nos/iNos	0	0	0	
Integrase/LEDGF	16	0	1	
MDM2/p53	7	4	4	
Xiap_BIR3/Smac	19	6	7	
TNFR1A/TNFB	14	1	0	
ZipA/FtsZ	0	0	0	



We proposed two strategies to analyze the pockets and identify the putative transient cavity from MD. One is using the top ranked pockets predicted by Fpocket and the second is using the pockets that have at least 2 most frequent residues from all those pockets located at a concrete place (defined using HS) during the simulation (Figure 2.).

In order to evaluate the weight of HS in the role of selection of possible candidates with an interesting transient cavity, we also evaluated both strategies without HS. We compared both strategies with unbound cases. From all strategies, we selected the snapshots with a putative transient cavity in different ways: the top scored by fpocket, the top druggable defined by fpocket and the top druggable from the top 100 scored by fpocket (topD(P100)). We propose the top 5 candidates (from topD(P100)) for further analysis using ligand docking. If we compare unbound structures with respect to the selected cases, we obtain better results in these selections (Figure 3).

Fig. 3. At left, different strategies of selection of candidates with the best transient cavity. In unbound structures selected according to the results obtained from PPIfs selection, we analyzed the pocket and the transient pockets resulting from the simulation. P0 means the top ranked pockets strategy and P2 means the strategy applying most frequent residues. Transient pockets were analyzed: Using HS (at least 3HS) and without HS. Correct predictions were considered with a PPV&COV$\geq$40%. At right, results of ligand docking in MDM2 applying P0 with at least 3HS strategy (a) and directly with the unbound structure (b) of selection.



III. CONCLUSIONS

The characterization of druggable cavities in PPIfs is still unknown where predicting PPIfs from a three dimensional structure is a key task for the modulation of PPIs. The use of the NIP-based HS prediction method improves the identification of transient cavities from MD simulation when compared to known binding cavities. We propose a new tool to predict and characterize PPIs, PPIfs and HS residues. We present a strategy based on MD and NIP which allows to identify cavities as potentially good targets to bind inhibitors. This approach can be extremely useful in a realistic scenario of drug discovery targeting PPIfs, when there is no information at all about the protein-protein complex structure.

REFERENCES

1. Wells J.A., McClendon CL., (2007). Nature, 450,1001-1009.
2. Bogan A.A. and Thorn K.S., (1998). J. Mol. Biol, 280, 1-9.
3. Basse MJ., Betzi S., Bourgeas R., et al. (2013). Nucleic Acid Research, 41, 824-827
4. Teng S., Madej T., Panchenko A., Alexov E., (2009). Biophys J. 96, 2178–2188.
5. Cheng T.M-K.,Blundell T.L.,and Fernandez-Recio J., (2007).Proteins, 68, 503_515.
6. Fernandez-Recio J., Totrov M., and Abagyan R.,(2004). J Mol Biol, 335, 843-865..
7. Grosdidier S., and Fernandez-Recio J., (2008). BMC Bioinformatics, 9:447.
8. Hubert Li, Vinod Kasam,. (2014) J Chem. Inf. Model 54, 1319-1400.
9. Chen R., Weng Z. (2003). Proteins 51, 397-408.
10. Le Guilloux V., Schmidtke P. and Tuffery P., (2009). BMC Bioinformatics, 10-168