

Interaction Finger Prints for Docking Analysis of Scaffolds

*D. Rognan, E. Kellenberger, G. Marcou**

Abstract

The authors have developed a C++ library to design intermolecular interaction finger prints between a ligand and its receptor. Those finger prints take the form of bit vectors. They subsequently show how to retrieve relevant poses in docking output, using a simple Tanimoto metric to compare bit vector to given references. They show that using this approach, one can retrieve relevant scaffolds from a scaffold library for a particular target. The results seem promising to help in rational drug design.

Keywords: interaction finger prints, docking

1 Introduction

Docked poses are typically scored to define a ranking then filtered on geometric criteria, consensus with different scoring functions, visual analysis and a bit of black magic. Upon filtering, one usually relies on the similarities of a docked pose to one that is known or believed to be of good quality. This is usually done by visual inspection of 3D or 2D representations of the docked poses. Another option is to check if similar molecular features are in the same environment. This is the goal of the intermolecular interaction finger prints library presented here.

The notion of finger print to describe chemical features is relatively old. In fact binary representation of data is as old as computers and allows for numerous data treatment more or less sophisticated. The main advantage is to introduce related metrics that can be used for various kernel methods or statistical analysis. Nonetheless the application of the concept to the analysis of docking output was for the first time published in a truly mature form, only recently.

This intermolecular interaction finger print library was applied to a challenging, yet promising approach in computer aided drug design. Modern technics has been developed to experimentally screen scaffolds on target bio-polymers. The goal being to identify a high affinity scaffold that can further be decorated generating small and cheap compound libraries. It is expected that it will then be easier to find in these libraries more potent, specific and drugable molecules for the target protein.

But, because scaffolds are by definition small and rigid, they proved to be difficult to screen by computational means. Scoring functions get easily lost in multiple binding modes with similar potential intermolecular interactions. A scaffold's top scored pose does not always makes sense for a particular target. The bio-informatician has then to retrieve in multiple poses proposed by his automated docking tools, the scaffolds that were successfully docked, according to some hypothesis. The intermolecular interaction finger prints are designed to help in this task. The original contribution of this paper is the application of finger prints to help docking analysis of scaffolds.

2 Results

The intermolecular interaction finger prints library is used to generate several applications. One of them, LBMFP, takes in entry a cavity, a set of docked ligand poses and a set of reference poses to be compared to. It generates finger prints for each input pose and evaluates a Tanimoto score¹ between each reference and each docked ligand pose.

The program was applied to two test sets. The first one is a test set of scaffold-like molecules and the second one was a test set of scaffolds. At last the program was applied to retrieve active scaffolds using virtual screening output.

2.1 Scaffold-like molecules test set

The program was applied to a first test set of scaffold like molecules. It consists of a set of 60 ligand-protein complexes extracted from the scPDB. They were chosen so that the ligand were small, rigid but not too simple (see 4.2).

Each ligand was docked into the binding site of the corresponding protein. From each docking experiment, at most 30 poses were conserved. The validity

¹ Tanimoto score is defined as the number of 1 common to two bit vectors divided by the number of one that appear indifferently in the first or second bit vector.

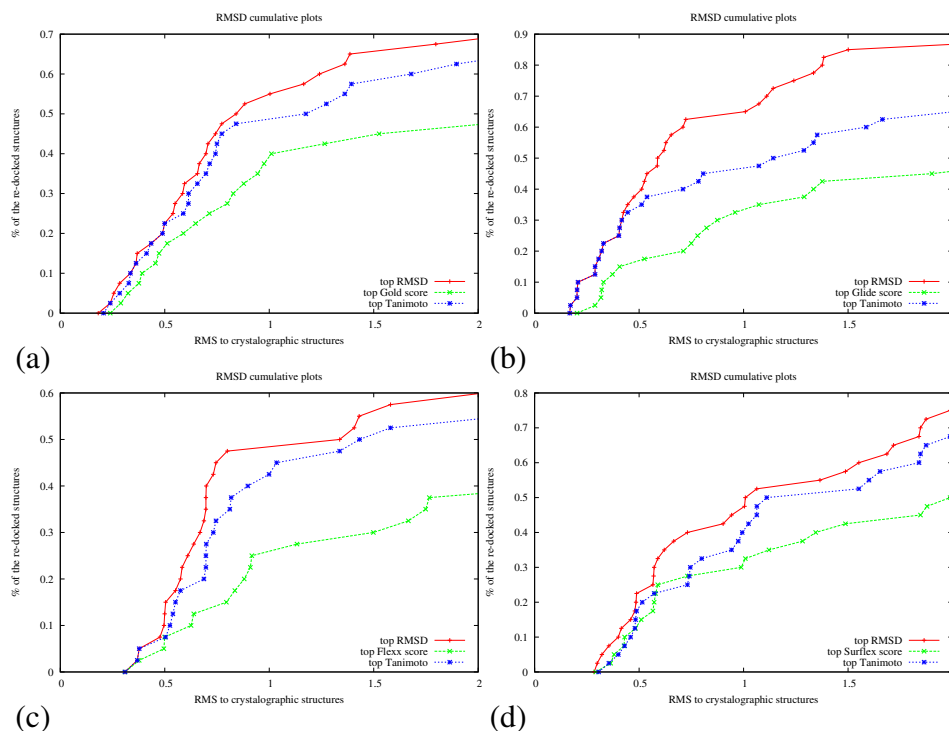


Fig. 1: Cumulative rmsd plot for (a) Gold, (b) Glide, (c) FlexX, (d) Surflex. At each value of rmsd, the curve gives the percentage of the prioritized library selected that is closer than this value from the reference crystallographic position. Ligands poses were prioritized by rmsd to their reference crystallographic position (red curves), to the docking score (green curve), to the Tanimoto score (blue curve).

of each pose was assessed by evaluating the rmsd between the docked and the crystallographic poses. The crystallographic position of the ligand was also used as reference to get a Tanimoto score for each of the docking pose.

For each ligand three poses were considered particularly: the closest one to the crystallographic position according to rmsd, the best docking scored one and the best Tanimoto scored one. The docked ligands were sorted -prioritized- according to each of these criteria.

For each docking tool, it was verified that no significant correlation existed between the docking tool's score and the Tanimoto score. Some weak correlation existed between the Tanimoto score and the rmsd, though some outliers were iden-

tified that had both significant Tanimoto similarity and high rmsd values. Those outliers were expected since they are alternative positions of partially symmetric ligands, able to fulfill most of the characteristic interactions of their references. Docking scores had weaker correlation with the rmsd, and outliers were often found having a bad scores and low rmsd values.

As a consequence, it is observed that the Tanimoto score is, in general, a better choice than the docking score to select the best pose according to the crystallographic pose (figure 1). No docking tool was able to retrieve more than 50% of the ligands closer than 2Å rmsd from the crystallographic position references. It is particularly spectacular for Glide, for which in 90% of the cases a pose was generated that was closer than 2Å rmsd from the crystallographic pose but it selected it in less than 30% of the cases.

2.2 Scaffolds test set

The program was then used to retrieve the docking pose most relevant to user's hypothesis. For a set each protein of set of 10, four scaffolds were proposed extracted from ligands that were co-crystallized with them. The best docking pose possible was considered to be the closest one, according to rmsd, to the crystallographic position.

Four docking hypothesis were possible to analyze the docking poses of the scaffolds:

1. for each scaffold docked, the position of the same scaffold in the crystal structure.
2. for each scaffold docked, the position of the ligand it was extracted from, in the crystal structure.
3. for each scaffold docked, the position of the three other scaffolds docked to the same protein, in the crystal structure.
4. for each scaffold docked, the position of the three other ligands docked to the same protein, in the crystal structure.

As for the preceding test, for each scaffold, only one pose was kept. The scaffolds were then sorted according to rmsd to their crystallographic position.

When a docked scaffold pose had to be compared to the position of the other three scaffolds or ligands crystallographic positions in the same protein, some

basic data fusion was used. That is the docked pose with the highest Tanimoto score over the other three references was kept.

The best docking score pose and the smallest rmsd to crystallographic position pose for each ligand were extracted and sorted the same way.

The cumulative plots shows similar features as for the first test set and get some expected behavior (figure 2). It is observed that the Tanimoto score, in general is a better choice than the docking score. Also, using the same scaffold's crystallographic position as reference is the best choice possible. And using the crystallographic position of the ligand from which was extracted the scaffold as reference is a second choice.

The rmsd from crystallographic position of the selected docked poses averaged over the four scaffolds per protein were calculated (figure 3). Mean values over four points do not bear a lot of sense and there is even less for variance estimators. So no error bars were proposed for the plot since the plot would get more confuse without much information added. Yet it must be noted that there was no case in which a docking tool was not able to find for each protein, at least for one of the four scaffolds, a pose closer than 2Å of the crystallographic position.

2.3 Scaffold library virtual screening

See Didier

3 Discussion

The fast virtual screening docking methods are based on one fundamental hypothesis: the affinity of a ligand for a protein can be explained, essentially by one conformer of the ligand docked in one conformer of the protein. This hypothesis is also behind experimental scaffold screening. It is expected that good scaffolds would, essentially, bind in a specific manner to a given conformer of the protein. This is essentially the interpretation sought in NMR scaffold screening.

But virtual screening of scaffold is at the limits of the capacities of docking tools. The fragments are rigid and small so that their conformational space is fast explored. But they make a low number of interactions that make different binding modes difficult to distinguish on the sole basis of the docking score. Some supplementary information has to be given. This is the goal of the intermolecular interaction finger prints.

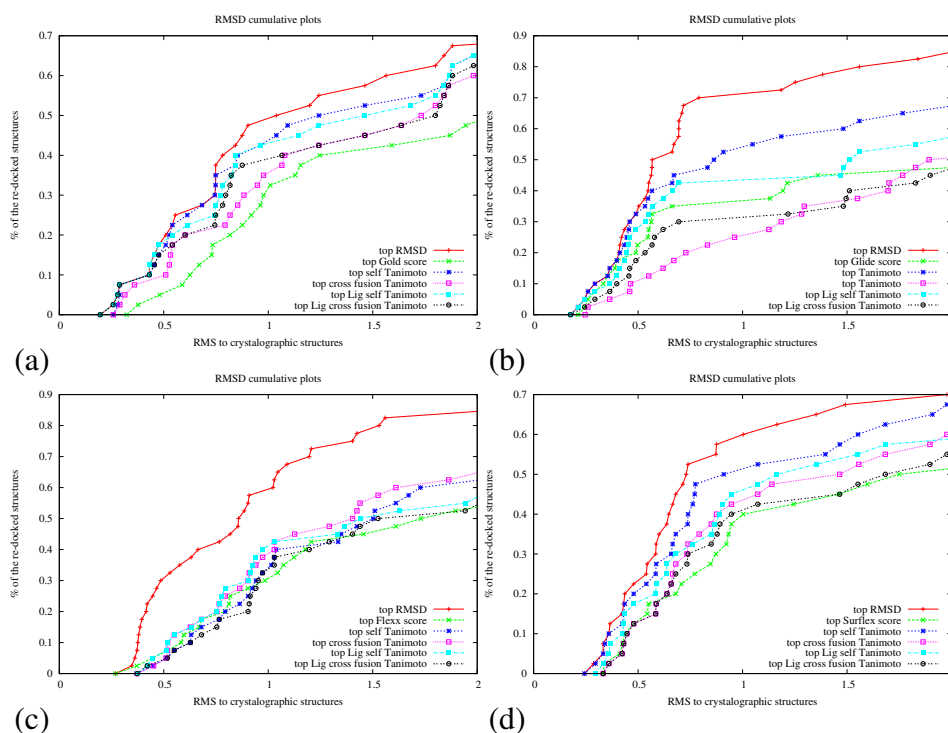


Fig. 2: Cumulative rmsd plot for (a) Gold, (b) Glide, (c) FlexX, (d) Surflex. At each value of rmsd, the curve gives the percentage of the prioritized library selected that is closer than this value from the reference crystallographic position. Scaffolds poses were prioritized by rmsd to their reference crystallographic position (red), to the docking score (green), to the Tanimoto score using the same scaffold's crystallographic position (blue) or the crystallographic position of the ligand from which it was extracted (magenta) or one of the three other scaffold of the same protein (cyan) or one of the three other ligands of the same protein (black).

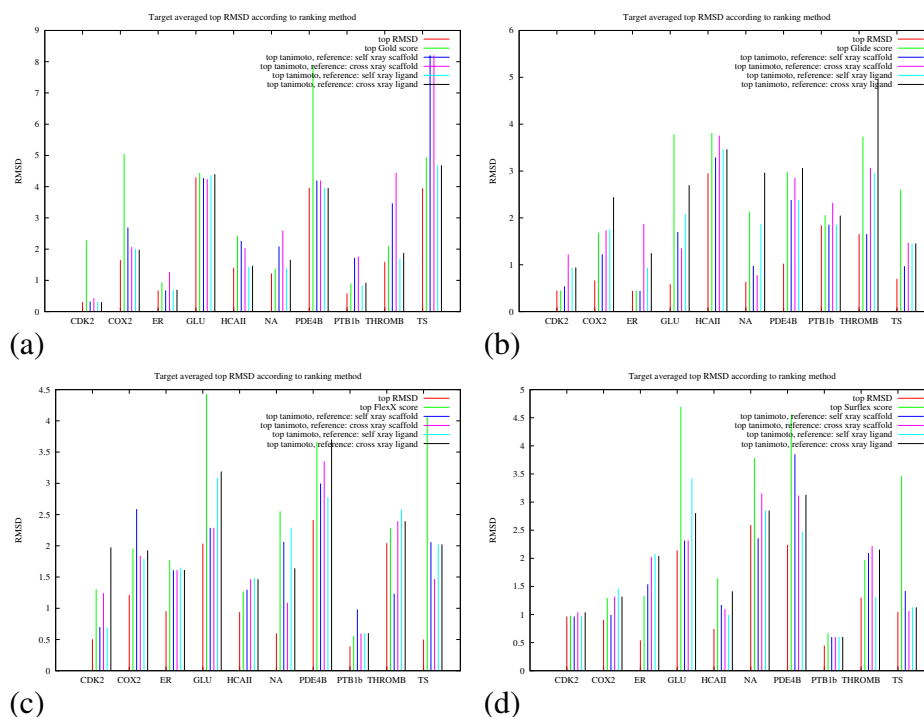


Fig. 3: Cumulative rmsd plot for (a) Gold, (b) Glide, (c) FlexX, (d) Surflex. At each value of rmsd, the curve gives the percentage of the prioritized library selected that is closer than this value from the reference crystallographic position. Ligands poses were prioritized by rmsd to their reference crystallographic position (red curves), to the docking score (green curve), to the Tanimoto score (blue curve).

It shall be noted that in what follows we are not comparing our docking tools relative performances but their relative behavior. In all the test sets, the preparation of the protein and the ligand was fully controlled so that each docking program had exactly the same input. This way of preparing the input might not be fully optimized and some docking tools might give better results using their standard preparation protocols.

In the first test set, the finger print library was shown to be able to retrieve docking poses of small scaffold-like ligands closer to the crystallographic position than the docking score. This is already a desirable feature since the docked position can help combinatorial chemistry optimization.

The most spectacular result belongs to Glide (figure 1 (b)). The systematic conformer exploration of the algorithm is very efficient on such small sized, rigid ligands. While being very successful in generating at least one good pose, the score is misleading. The use of G-Score for pose selection did not change anything.

As FlexX's algorithm is based on recognition on fragment positioning we expected it to perform better than any other tool. So the average performance of FlexX was disappointing (figure 1 (b)). FlexX also is very sensible to input preparation. Since this part has been considerably simplified in the newer versions of FlexX, results shall improve.

Overall, Surflex and Gold are performing quite well. They differ in their ability to retrieve poses close to crystallographic position. As the rate they generate good poses is more similar.

Overall the intermolecular interaction finger prints based curves on all four docking output are quite comparable. This is expected since it shows some independence of the analysis over the docking tool.

Two concepts are underlying behind the second test set. The first one was to provide a test closer to scaffold screening. The second was to allow for some kind of cross docking analysis.

Glide remained the best tool to generate docking poses but was still easily confused (figure 2 (b)). Some expected trends also appear clearer with those curves. First the crystal position of the scaffold represents a perfect binding mode hypothesis. Around this hypothesis, taking the crystal position of the complete ligand, doing data fusion using the other scaffolds or complete ligands for the same protein would represent some ways to add different kind of noise to the binding mode analysis. The performances of the finger print analysis can be deeply affected by these.

The behavior of Gold and Surflex remained also similar to what they were for

the previous test set (figure 2 (a) and (d)).

The FlexX docking tool behavior changed in a more expected manner (figure 2 (c)). The scaffold being more similar to the fragments used by the docking algorithm, FlexX proved to be very good both for scaffold docking and pose scoring. This account for greater sensibility of the tool compared to the others to the ligand size.

The histograms of rmsd average over the proteins (figure 3) shows a particular behavior of Glide. In fact Gold, Surflex and FlexX tend to fail on the same proteins on which Glide seems to outperform them all. This provide some self-consistency to the test set since it tends to show that no major bias in protein or ligand preparation was included.

Gold scoring performances seems closer to those of the optimum pose selection than any of the other docking tools. It is also the docking tool that can be most easily wrong in pose generation. This is particularly true for the tryptophane synthase. Among the tryptophane synthase scaffolds three out of four has a sulfur. As they are protonated these sulfur might be considered as hydrogen bond donors even though they are not in the ligand the scaffolds are extracted from.

The histograms confirm that it is usually a better idea to do finger print analysis rather to rely on the docking score. Indeed, when the average best rmsd is greater than 3 Å, so the finger prints are losing sense too.

There is a few other notable points. The phosphodiesterase seems to be the more difficult target of the test set. This not so much a surprise since we performed unconstrained docking and docking programs tried to fill some of the coordination site of the two metals. Though, we dock IBMX-related fragments that should be found in hydrophobic clamp of the PDE.

Another difficult target is the ionotropic glutamate receptor. Here also, this entry was particularly sensible because of the presence of two water molecules that were oriented manually.

Thrombine seems a mild difficult target for all the docking tools. None is performing particularly well in generating poses.

Discussion Didier...

4 Materials and Methods

We used a C++ library for generating the intermolecular interactions finger prints. The library, programs generated from this library and the present test sets are freely available on the web site: <http://www.bioinfo.u-strasbg.fr>.

4.1 The intermolecular interaction finger prints library

The library makes use of three other libraries. The first one, the Boost library is used for user interface and regular expressions processing. The second one, the C cluster library, is used for Tanimoto scores based clustering of high throughput docking outputs. Both of these libraries are released in some kind of public licensing and are free. The last library used is OpenEye's oechem. This last library is free for academics but can be substituted with little efforts with OpenBabel. The library makes use of the object paradigm of C++ in an effort to be easily expandable and adaptable to particular problems. There exists some documentation for use of the library and of the derived tools as well as some examples of how to manage and expand the basic finger prints proposed.

Our finger prints describe the existence or not of hydrogen bonds -one bit for donor-acceptor, one bit for acceptor donor-, hydrophobic contacts -one bit-, aromatic systems face to face -one bit- or edge to face -one bit- and electrostatic match -one bit for plus-minus, one bit for minus-plus. They were 7 bits per residue long. Atoms were labeled according to their potential to fulfill some potential interactions:

- Hydrogens linked to carbon atoms and carbon atoms were labeled hydrophobic.
- Oxygen, nitrogen and sulfur atoms were labeled hydrogen bond acceptors.
- Hydrogens bonded to oxygen, nitrogen and sulfur atoms were labeled hydrogen bond donors.
- Aromatic atoms were considered aromatic.
- Atoms bearing formal positive partial charge were considered cations.
- Atoms bearing formal negative partial charge were considered anions.

Interactions were detected if atom labels were compatible and their position fulfilled the simple geometric requirements:

- Hydrophobic interactions occurred between hydrophobic atoms closer than 3.5Å
- Hydrogen bond interactions occurred between acceptor and donor atoms closer than 3.5Å and if the angle of the bond was $\in [-\frac{\pi}{6}, \frac{\pi}{6}]$.

- Aromatic edge to face interactions occurred between aromatic atoms closer than 3.5Å and if the angle between the aromatic planes was $\in [\frac{\pi}{6}, \frac{5\pi}{6}]$.
- Aromatic face to face interactions occurred between aromatic atoms closer than 4.5Å and if the angle between the aromatic planes was $\in [-\frac{\pi}{6}, \frac{\pi}{6}]$.
- electrostatic match occurred between cations and anions closer than 4.0Å

4.2 Description of the scaffold-like molecules test set

This test set is composed of 60 ligand-protein complexes extracted from the scPDB. The ligand were selected so that:

- Molecular weight ≤ 250
- Number of cycle $\in [1, 3]$
- Number of Hydrogen bonds donors $\in [0, 2]$
- Number of Hydrogen bonds acceptors $\in [0, 5]$
- Number of rotamer degree of freedom $\in [0, 2]$
- Polar Surface Area (PSA) $\in [20, 120]$

Protonation of the ligand was done using Filter. Hydrogens were added to the protein using Tripos's biopolymer module. Protonation state and tautomeric state of each residue near the binding site were checked and corrected when necessary. Rotamers of polar hydrogen and water molecule orientations were checked and set manually when necessary.

Docking of each ligand in its native pocket was done using: Gold, Glide, Surflex, FlexX.

4.3 Description of the scaffolds test set

This test set is composed of 40 ligand-proteins crystallographic structures from the PDB. Those entries are set so that they represent 10 proteins, 4 PDB entries each.

For each ligand in these complexes, a small scaffold was extracted. The extracted scaffolds have the following properties:

CDK2	COX2	ER	GLU	HCA2
1DM2	1CX2	1XP1	1P1O	1BNU
1FVT	1PXX	1XP1	1WJV	1EOU
1JVP	3PGH	1MQJ	1INV	1G45
1PXP	4COX	1MY2	1IVD	1TTM
NA	PDE4B	PTB	THROMB	TS
1INV	1XOR	1C84	1D4P	1C29
1IVD	1TB7	1PTY	1D6W	1CW2
1L7G	1Y2E	1XBO	1D9I	1CX9
2QWG	1ZKN	2BGD	1JWT	1K7E

Tab. 1: PDB entries list at the basis of the scaffold test set. The proteins described are the CDK2, the COX2, the estrogen receptor (ER), the ionotropic glutamate receptor type 2 (GLU), the neuraminidase (NA), the phosphodiesterase 4B (PDE4B), the protein tyrosine phosphatase (PTB1b), the thrombin (THROMB) and the triptophane synthase (TS).

- Molecular weight ≤ 150
- 1 or 2 cycles
- Number of Hydrogen bonds donors $\in [0, 2]$
- Number of Hydrogen bonds acceptors $\in [0, 2]$
- Number of rotamer degree of freedom $\in [0, 2]$
- Polar Surface Area (PSA) $\in [20, 120]$

Protonation of the ligands was decided the help of Filter. Protonation and tautomerization of the protein was done using Tripos's Biopolymer module, then checked and corrected manually if necessary. Rotamer state of all polar hydrogen in the binding site was checked and corrected manually when necessary.

4.4 Description of the scaffolds virtual screening experiment

See Didier

5 Conclusion

Intermolecular interaction finger prints are a promising method to help analyze docking poses. As it introduces some obvious metrics, it opens the field to numerous kernel and learning machine methods; the bit vector representation allows for immediate application of a lot of statistical methods.

In this paper we proposed a way to do scaffold docking. In this context, intermolecular interaction finger prints prove to be an efficient tool. First because it helps in retrieving the closes docking pose to a set of binding mode hypothesis. Second because we showed that it can produce enrichment in prioritized library where none could be obtained by the sole use of the docking scored.

Doing so we could assess some of the limit of the approach. Because of the coarser grain of the Tanimoto score compared to usual scoring function², several poses can sometime get the same Tanimoto score.

6 Acknowledgments

² The Tanimoto scores are taken from a smaller subset of the rationals