Application of Structural Interaction Fingerprints (SIFt) in Identification and Analysis of GPCR Binding Sites

Stefan Mordalski, Tomasz Kościółek and Andrzej J. Bojarski
Department of Medicinal Chemistry, Institute of Pharmacology, Polish Academy of Sciences Smętna 12, 31-343 Kraków, Poland
e-mail: tomek.kosciolek@gmail.com

Introduction:

One of the most troublesome stages of Computer Aided Drug Design (CADD) process is analyzing huge amount of data provided by docking studies. Simple scoring functions alone can provide only shallow information about ligand-receptor interactions, since they do not distinguish neither residues nor single atoms. Very often a visual inspection is the only way to determine a binding mode. Here, we introduce an implementation of interaction profiles (Chuaqui, 2005) based on Structural Interaction Fingerprints (SIFt) (Deng, 2004), which allow precise and rapid binding site description.

Fingerprints preparation:

SIFt is a bit string representing interactions between ligand and receptor (Deng, 2004; Chuaqui, 2005). It can be divided into chunks expressing contacts for individual amino acids (Fig. 1). In this research nine bits were used to describe those associations: any contact, backbone, side chain, polar, hydrophobic, hydrogen bond donor/acceptor, aromatic and charged. Basic algorithm is to find amino acids around bound ligand, and from distance and residue types, to determine class of interaction. This approach was successfully applied in interaction fingerprints module of Schrodinger suite.

In our approach Schrodinger python libraries are used to handle 3D structures of ligand-receptor complexes and to determine possible interactions. The algorithm itself was tweaked to enhance computations speed, and to simplify batch SIFt generation and analysis. To reduce computation time each atom of receptor structure was indexed using red-black tree hash table. Distance from (0,0,0) was a search key. Analogous orbit, calculated for complexed ligand atoms, was then queried, and for each fragment returned as a result, atom-atom distance was measured to reject false positives. For every approved amino acid fragment, interaction type was determined and appropriate fingerprint bits were switched on.
So generated SIFts were then analyzed to determine amino acids contributing to ligand binding (Fig. 2).

SIFt analysis:

Every ligand-receptor pair has its individual interaction fingerprint. On this basis an average SIFt (Fig. 2) may by generated for the population of ligands and/or receptors (e.g. alternative conformational states). Implemented workflow is focused on finding consensus interactions for the set of ligands docked to a particular model (receptor-based approach).

For each receptor conformation (model) a list of interactions is constructed. The list is then sequentially recalculated for every amino acid in the population of ligands docked into each receptor, comparing between alternative complexes. At this stage, only ‘any contact‘ bit is taken into account. The most frequent matches (present in at least 50% of entries) are then put into a separate list for the construction of averaged/consensus fingerprint. When all the alternative models are analyzed this way, the list is recalculated, and again, a consensus fingerprint for the whole population is generated.

Output .sift file contains a matrix of real values representing relative presence of particular interaction in analyzed population of ligands/receptors.
Analysis thereof allows i.e. rapid determination, whether amino acids preferred by docking experiments stay in agreement with experimental data (Table 1), and therefore determine credibility of obtained results.
Visualization of acquired output is also possible (Fig. 4). Our in-house sctipt utilizes PyMol to display chosen interactions during visual inspection of ligand-receptor complexes.

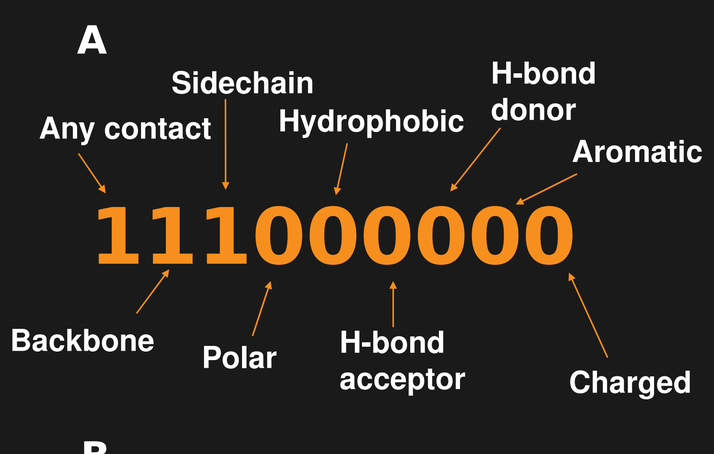
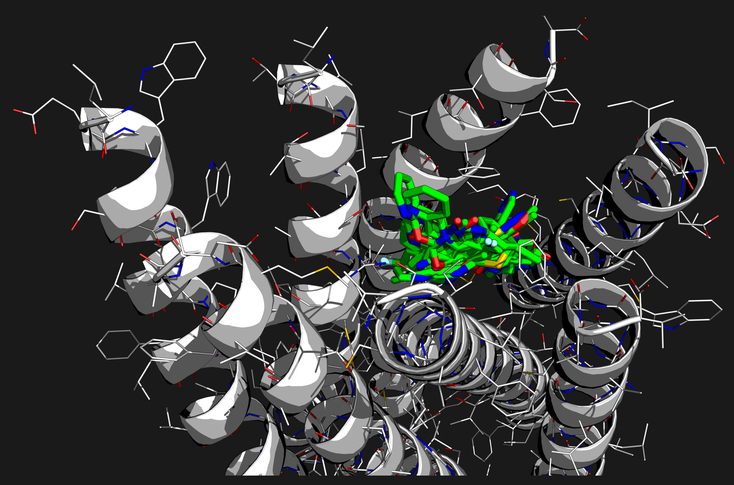
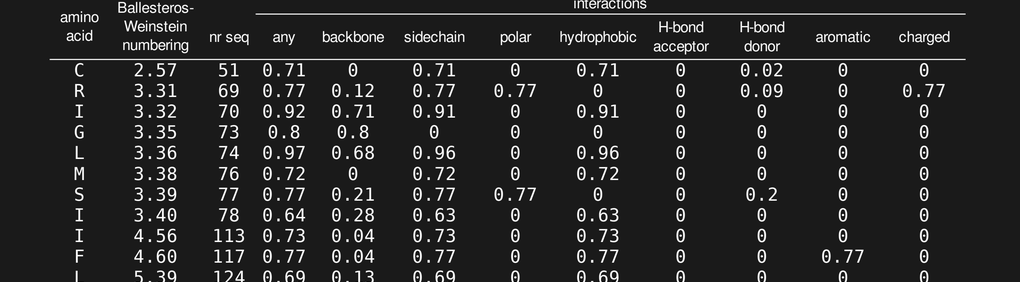
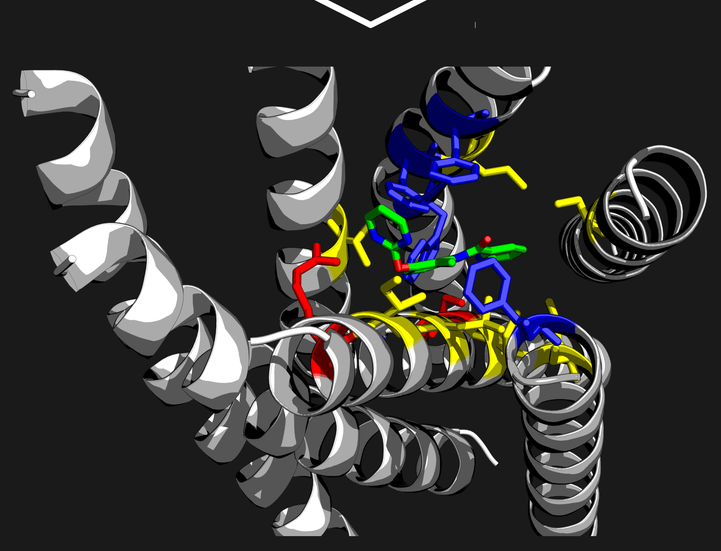


Fig. 1. (A) Ligand-receptor contacts classified by interaction type as used in the current implementation of SIFt.
(B) All created fingerprints are analyzed to find common interactions for all reference ligands.



amino acid	Ballesteros-Weinstein numbering	nr seq	interactions								
			any	backbone	sidechain	polar	hydrophobic	H-bond acceptor	H-bond donor	aromatic	charged
C	2.57	51	0.71	0	0.71	0	0.71	0	0.02	0	0
R	3.31	69	0.77	0.12	0.77	0.77	0	0	0.09	0	0.77
I	3.32	70	0.92	0.71	0.91	0	0.91	0	0	0	0
G	3.35	73	0.8	0.8	0	0	0	0	0	0	0
L	3.36	74	0.97	0.68	0.96	0	0.96	0	0	0	0
M	3.38	76	0.72	0	0.72	0	0.72	0	0	0	0
S	3.39	77	0.77	0.21	0.77	0.77	0	0	0.2	0	0
I	3.40	78	0.64	0.28	0.63	0	0.63	0	0	0	0
I	4.56	113	0.73	0.04	0.73	0	0.73	0	0	0	0
F	4.60	117	0.77	0.04	0.77	0	0.77	0	0	0.77	0
L	5.39	124	0.69	0.13	0.69	0	0.69	0	0	0	0
I	5.42	127	0.65	0.16	0.63	0	0.63	0	0	0	0
C	5.43	128	0.79	0.72	0.75	0	0.75	0	0	0	0
G	5.46	131	0.63	0.63	0	0	0	0	0	0	0
Y	5.47	132	0.67	0.56	0.67	0	0.67	0	0	0.67	0
L	5.50	135	0.62	0	0.62	0	0.62	0	0	0	0
W	6.48	165	0.83	0.05	0.83	0	0.83	0	0	0.83	0
F	6.51	168	0.99	0.06	0.99	0	0.99	0	0	0.99	0
I	6.52	169	0.79	0.27	0.78	0	0.78	0	0	0	0
I	6.54	171	0.65	0.37	0.64	0	0.64	0	0	0	0
F	6.55	172	0.88	0.22	0.88	0	0.88	0	0	0.88	0
F	6.56	173	0.73	0.64	0.73	0	0.73	0	0	0.73	0
T	6.58	175	0.67	0.14	0.67	0.67	0	0	0.05	0	0
S	6.59	176	0.65	0.28	0.64	0.64	0	0	0.04	0	0
S	7.35	181	0.63	0.43	0.63	0.63	0	0	0.01	0	0
S	7.39	185	0.77	0.49	0.76	0.76	0	0.01	0.03	0	0
V	7.42	188	0.82	0.08	0.82	0	0.82	0	0	0	0
S	7.43	189	0.98	0.58	0.98	0.98	0	0.01	0	0	0





* ≥ 75% interaction cutoff

Fig. 2. Schematic flowchart of active-site description.
(top) a population of ligands docked to a receptor.
(center) An averaged structural interaction fingerprint (SIFt) presented for all identified interacting residues. Representation of SIFt on schematic helices aligned according to relative amino acids positions. Only three types of interactions are presented for the sake of clarity of the figure.
(bottom) A representative ligand docked to the receptor (the same as on the top). Residues identified by SIFt are colored according to the scheme presented above.

Metabotropic glutamate receptors:

The mGluR family consists of eight proteins divided into three groups corresponding to sequence similarities, pharmacology and physiological role. These groups are: I (mGluR1, -5), II (mGluR2, -3) and III (mGluR4, -6, -7, -8). All mGluR receptors consist of 2 topological domains (Fig. 2): an extracellular Venus Flytrap – binding glutamate or other orthosteric ligands, and transmembrane 7TM – bearing allosteric site.
mGluR allosteric potentiators lie in field of our interest due to their potential as therapeutic target for antidepressant and anxiolytic drugs.

This research was performed on population of 100 mGluR4 models created on rhodopsin crystal structure template. Building that many virtual receptors provided us with semi-conformational search on residues assembling incriminated receptor.
The Library of 53 known allosteric modulators of group III mGluR was used for docking studies and thus forging the binding mode.

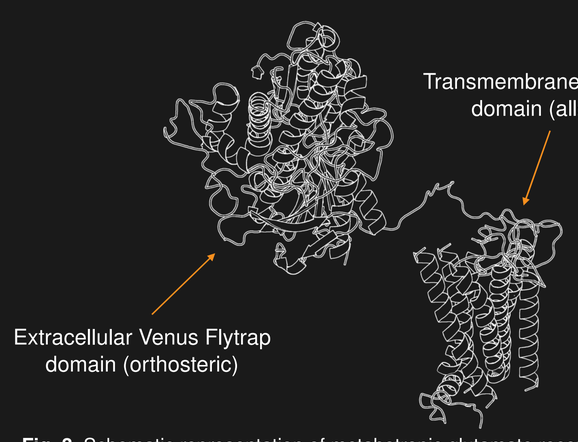


Fig. 3. Schematic representation of metabotropic glutamate receptors' domain organization.

Table 1. The comparison between amino acid composition of an active site determined by SIFt and emerging from mutagenetic experiments (Malherbe, 2003, 2006). Unique mGluR mutations are grouped according to their Ballesteros-Weinstein numbering scheme.

receptor	mutation	Ballesteros-Weinstein numbering	SIFt in mGluR4?
mGluR5 (rat)	P654S	**3.36**	+
	Y658V	**3.40**	(weak)
	W784A	**6.48**	+
	F787A	**6.51**	+
	A809V	**7.47**	-
mGluR1 (rat)	Y805A	**6.55**	+
	T815M	**7.39**	+

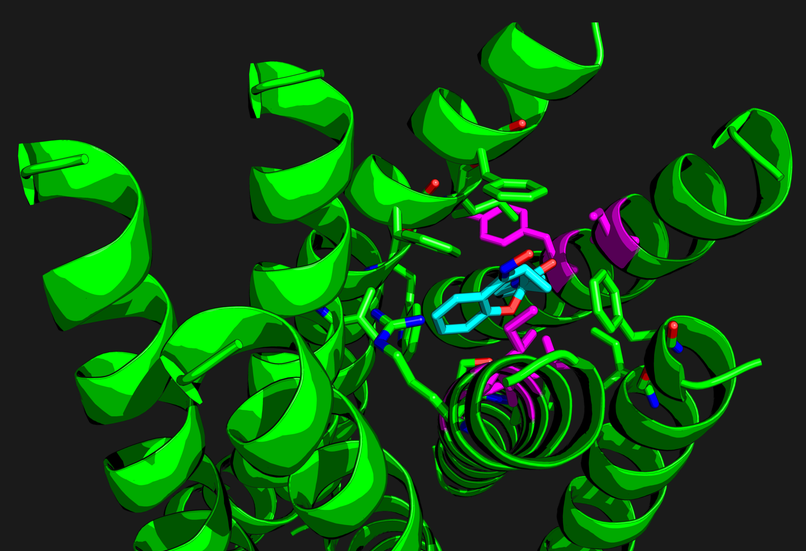


Fig. 4. Our scripts for generation and analysis of SIFts (available for download via our website; see footnotes) allow easy imaging of identified interactions. This enables quick visualization of statistical results embedded in an average fingerprint over all ligands and chosen population of receptors.
Backbone interactions shown in **magenta**.

Scripts for generating SIFts, performing their analysis and allowing visualisation are available via http://www.if-pan.krakow.pl/ifpan_ww/index.php/medicinal-chemistry.html

This study was partly supported by a grant PNRF-103-AI-1/07 from Norway through the Norwegian Financial Mechanism.

http://www.cns-platform.eu

Polish-Norwegian Research Fund
norway grants

REFERENCES:

Ballesteros, JA. and Weinstein, H. (1995) Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors, *Methods in Neurosci*, 25, 366–428.

Chuaqui, C., Deng, Z. and Singh, J. (2005) Interaction profiles of protein kinase-inhibitor complexes and their application to virtual screening., *J Med Chem*, 48, 121–133.

Deng, Z., Chuaqui, C. and Singh, J. (2004) Structural interaction fingerprint (SIFt): a novel method for analyzing three-dimensional protein-ligand binding interactions, *J Med Chem*, 47, 337–344.

Malherbe P., Kratochwil N., Knoflach F., et al. (2003) Mutational analysis and molecular modeling of the allosteric binding site of a novel, selective, noncompetitive antagonist of the metabotropic glutamate 1 receptor.
J Biol Chem 278, 8340-8347.

Malherbe P., Kratochwil N., et al. (2006) Comparison of the binding pockets of two chemically unrelated allosteric antagonists of the mGlu5 receptor and identification of crucial residues involved in the inverse agonism of MPEP.
J Neurochem 98, 601-615.

Maestro, version 9.1, Schrödinger, LLC, New York, NY, 2010; http://www.schrodinger.com

The PyMOL Molecular Graphics System, Version 1.2r, Schrödinger, LLC; http://www.pymol.org