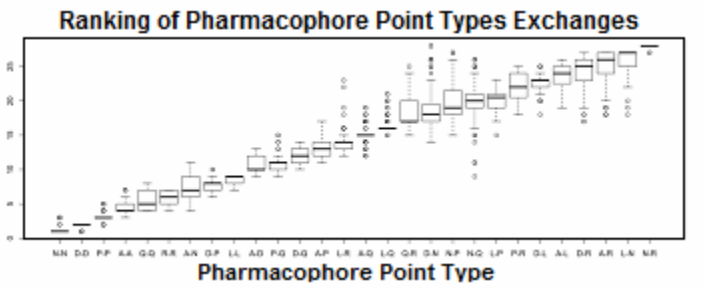
Conservation and Relevance of Pharmacophore Point Types

Tobias Fehlmann, Michael C. Hutter*,

** corresponding author: michael.hutter@bioinformatik.uni-saarland.de*

Center for Bioinformatics, Saarland University, Campus E2.1, 66123 Saarbrücken

Pharmacophore models apply a variety of features to express chemical characteristics, for example hydrogen-bond donors, hydrogen-bond acceptors, positively charged, negatively charged, as well as aromatic and hydrophobic moieties. [1-3] It is generally assumed that features have to match to identical types. [4,5] To clarify if this stringent one-to-one assignment is justified, we investigated a set of 581 unique ligands from the BindingDB with know orientation inside the respective binding pockets and conducted a statistical analysis of the likelihood of observed exchanges in between the pharmacophore features, respectively their degree of conservation. To find out if certain features are obsolete, we furthermore derived a ranking to determine the most relevant ones. We found that the degree of conservation decreases in the following order: negative ionizable (N) > hydrogen-bond donor (D) > positive ionizable (P) > hydrogen-bond acceptor (A) > aromatic (R) > non-aromatic pi-systems (Q) and other hydrophobic moieties (L) The most likely exchanges were found between carboxylate groups and hydrogen-bond acceptors (A-N), and likewise between basic nitrogens and hydrogen-bond donors (D-P), which reflects the characteristics of Lewis acids and bases. Whereas the kind of target (soluble proteins, metal-containing ones, and GPCRs) did not show substantial influence on the degree of conservation, the relevance of the respective pharmacophore feature was found to be strongly dependent on the applied ranking scheme. Overall, lipophilic and aromatic features turned out to be highly important, whereas the positive ionizable feature is less relevant.



Boxplot showing the degree of conservation, respectively the likelihood of exchanges between pharmacophoric point types. Leftmost combinations are highly conserved, respectively most likely. Those in the middle are randomly occurring, whereas those on the right hand side are disfavored.

[1] A. Leach, V. Gillet, R. Lewis, R. Taylor, *J. Med. Chem.*, **2010**, *53*, 539-558.
[2] O. Güner, J. Bowen, *J. Chem. Inf. Model.*, **2014**, *54*, 1269-1283.
[3] M. Löwer, E. Proschak, *Mol. Inf.*, **2011**, *30*, 398-404.
[4] G. Wobler, T. Langer, *J. Chem. Inf. Model.*, **2006**, *45*, 160-169.
[5] V. Hänke, B. Hofmann, T. Grgat, E. Proschak, D. Steinhilber, G. Schneider, *J. Comput. Chem.*, **2009**, *30*, 761-771.