Subpocket-based fingerprint for structural kinase comparison

Dominique Sydow[1], Eva Aßmann[1], Albert Kooistra[2], Friedrich Rippmann[3], Andrea Volkamer[1]

[1] *In silico* Toxicology, Institute for Physiology, Universitätsmedizin Berlin, Virchowweg 6, 10117 Berlin, Germany

[2] Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Jagtvej 162, DK-2100 Copenhagen, Denmark

[3] Computational Chemistry & Biology, Merck Healthcare KGaA, Frankfurter Str. 250, 64293 Darmstadt, Germany

Introduction

Kinases are important and well studied drug targets for cancer and inflammatory diseases. Due to the highly conserved structure of kinases, especially at the ATP binding site, the main challenge when developing kinase inhibitors is achieving selectivity, which requires a comprehensive understanding of kinase similarity. [1] Here, we present our work on a novel fingerprinting strategy designed specifically for kinase pockets, allowing for similarity studies across the structurally covered kinome.

Methods

The kinase fingerprint is based on the KLIFS [2] pocket alignment, which defines 85 aligned pocket residues for all kinase structures. This enables a residue-by-residue comparison across the kinome without a computationally expensive alignment step.

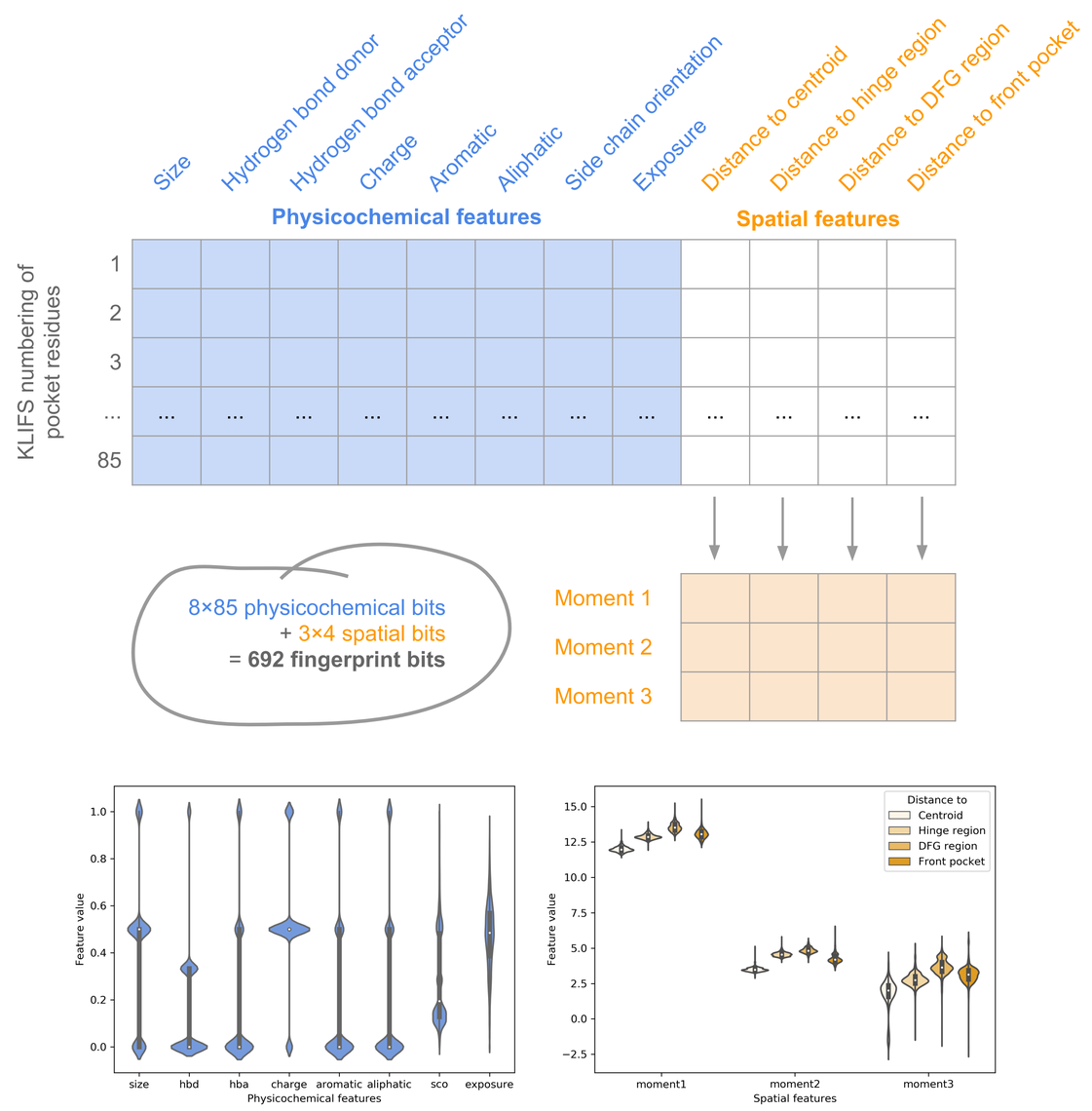
Data preparation. KLIFS dataset was filtered for structures (i) describing human kinases, (ii) in DFG-in conformation, (iii) with the best quality score per kinase-structure pair, and (iv) with a resolution ≤ 4 and a quality score ≥ 4. The resulting kinase dataset consists of 3,875 structures.

Kinase fingerprint. The *pocket fingerprint* consists of 85 concatenated *residue fingerprints*, each encoding a residue's spatial and physicochemical properties. Inspired by the ligand-based USR approach [3], the *spatial properties* describe the residue's position in relation to the kinase pocket centroid and important kinase subpockets, i.e. the hinge region, the DFG region, and the front pocket. The resulting distance distributions per subpocket are reduced in complexity to the first three moments, i.e. the mean, standard deviation and skewness. The *physicochemical properties* encompass for each residue its size, side chain orientation and pharmacophoric features as described by SiteAlign [4] in addition to its solvent exposure as implemented in Biopython's module `Bio.PDB.HSExposure` [5, 6].

Kinase comparison & scoring. Kinases are compared pairwise using the inverse, translated and scaled Manhattan distance as described for the USR method [3]. Per kinase, the best score is used for further analysis, resulting in a 253x253 similarity matrix.

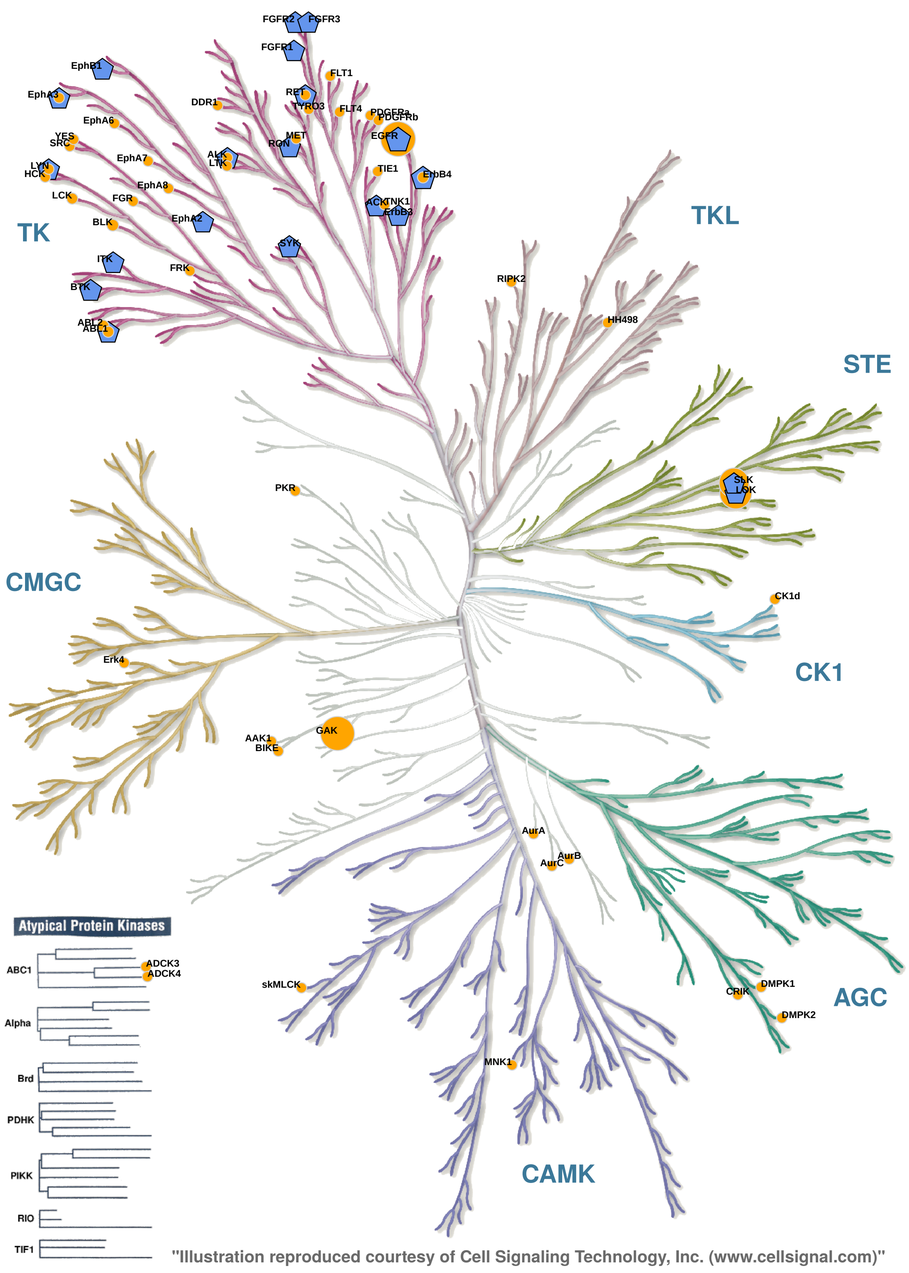
$$d(A,B) = \frac{1}{1 + \frac{1}{N}\sum_{i=1}^{N}|A_i - B_i|} \in]0,1]$$

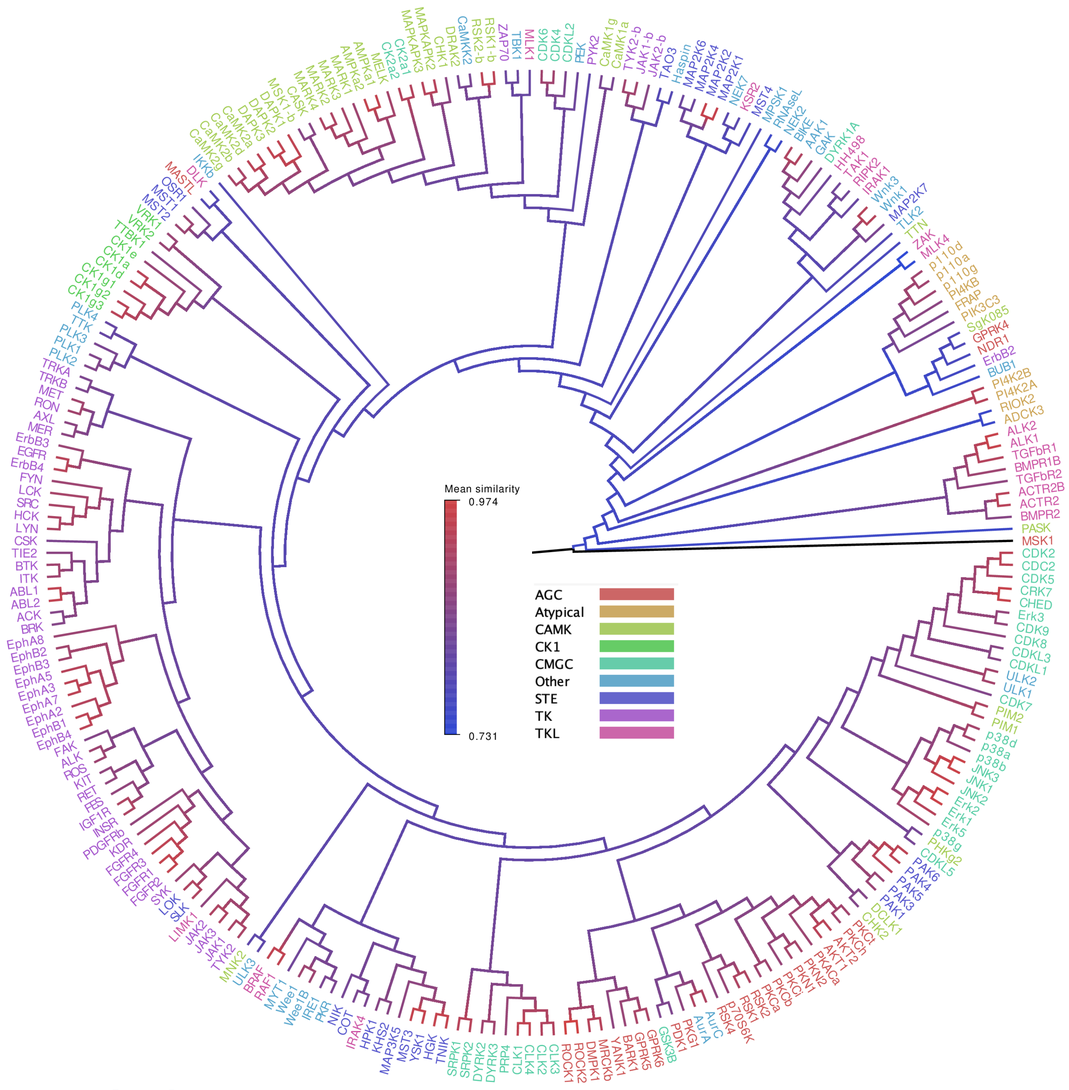
(inverse; translated; scaled; Manhattan distance)



Results

The potential of our subpocket-based kinase comparison is demonstrated by comparing our structure-based clustering results to the sequence-based Manning kinome tree [5], uncovering retrospectively ligand-based on- and off-targets, as well as assessing structure-based conservation of residue positions.





Conclusion

We believe that our analysis of the structurally covered kinome can help researchers (i) to detect potential promiscuities and off-targets at an early stage of inhibitor design and (ii) to conduct structure-informed polypharmacology studies.

References

[1] Kooistra and Volkamer. Ann Rep Med Chem. 2017, 50, 263-299. [2] van Linden et al. J Med Chem. 2014, 57, 249-77. [3] Ballester and Richards. J Comp Chem. 2007, 28, 1711-23. [4] Schalon et al. Proteins. 2008, 71, 1755-78. [5] Cock et al. Bioinformatics 2009, 25, 1422-3. [6] Hamelryck Proteins 2005, 59, 38-48. [7] Manning et al. Science. 2002, 298, 1912-34.