Cryptic Site Identification and Evaluation of Their Role for Allostery

Denis Schmidt[1], Christopher Pfleger[1], Susanne Hermans[1], Markus Boehm[2], Holger Gohlke[1]

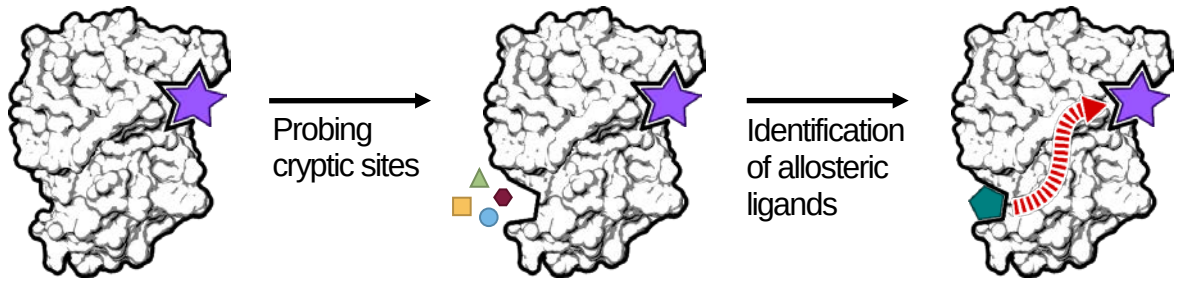
[1]Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

[2]Pfizer Worldwide Research & Development, Cambridge, Massachusetts 02139, USA

It is widely accepted that allosteric modulators have a huge potential to overcome typical hurdles in drug design. Our group recently showed that the "Constraint Network Analysis" (CNA) approach is capable of identifying allosteric pathways by means of rigidity analysis [1]. CNA hence has the potential to predict the allosteric effect of new ligands.

Even in the absence of a defined binding site, allosteric regulation might be possible via so-called "cryptic sites" ("transient pockets"). Such cryptic sites are closed in the *apo* state due to their lipophilicity. Consequently, they are hard to identify, although they are predicted to exist in many proteins [2]. Molecular dynamics simulations have been proposed as promising approach to sample cryptic sites in their open states [3,4,5,6].

Here, we systematically investigated the influence of different organic solvents on the opening of cryptic sites during molecular dynamics simulations, starting from the *apo* state. The addition of 10% phenol to the solvent significantly increased the number of frames where the cryptic site was open compared to water alone. By averaging pockets formed during these simulations, we identified and ranked spatial regions where pockets are more probable to form in the presence of organic solvent ("pocket cores"). For most test cases, i) the cryptic site opened, ii) the high-ranked pocket cores matched the known cryptic sites, and iii) the crystal ligand could be re-docked into MD-generated structures. Notably, the docking setup was based exclusively on the identified cores, thus excluding ligand bias. Unlike perturbation-based approaches [5,6], equilibrium simulations do not require *a priori* knowledge of the location of the binding site. Rather, pocket cores have the potential to locate yet unknown cryptic sites.



After their identification, we employ molecule surrogates ("fuzzy ligands") [7] to fill the cryptic sites. This way, CNA is able to quantify the allosteric potential of a ligand in this binding site. The generated fuzzy ligands subsequently serve as seeds for virtual screening.

The cryptic site identification in combination with the "fuzzy ligand" approach and CNA yields a powerful workflow to prospectively identify new sites, evaluate their allosteric potential, and ultimately predict new allosteric compounds.

[1] C. Pfleger, A. Minges, M. Boehm, C.L. McClendon, R. Torella & H. Gohlke, *J. Chem. Theory Comput.*, **2017**, *13*, 6343–6357
[2] P. Cimermancic, *et al.*, *J. Mol. Biol.*, **2016**, *428*, 709–719.
[3] S. Eyrisch & V. Helms, *J. Med. Chem.*, **2007**, *50*, 3457–3464.
[4] S.R. Kimura, *et al.*, *J. Chem. Inf. Model.*, **2017**, *57*, 1388–1401.
[5] D.B. Kokh, *et al.*, *J. Chem. Theory Comput.*, **2016**, *12*, 4100–4113.
[6] V. Oleinikovas, *et al.*, *J. Am. Chem. Soc.*, **2016**, *138*, 14257–14263.
[7] S.M.A. Hermans, *Master thesis*, Heinrich Heine University Düsseldorf, **2015**.