

We report on a module compatibility between FlexX^C and FlexX-Pharm. Three sample scenarios growing in size and complexity (combinatorial libraries of 1 million, 1.3 million, and 22.4 million compounds) have been created to illustrate a method inspired by A. R. Leach et. al. [1] (see Figure 1), which is essentially a combination of multimillion compound combinatorial docking and receptor-based pharmacophore constraints. Docking performed with this method was accomplished within minutes and several hours, depending on the outset. [4]

Learning Phase

Following Leach's approach, we conducted three preparing steps:

1. The first combinatorial library constructed of MTX analogues consisted of 1,095,510 compounds, and we retrieved the closest possible analogue to MTX on rank 1 (Figure 2), using a three-point pharmacophore of the DHFR active site
2. a) We took a subset (22,500) of a 1.356,250 library from the ACD spiked with 16 actives and retrieved up to 4 actives in the first 6 compounds (1st hit among first 3!) using a three-point pharmacophore of the DHFR active site
2. b) We screened the entire 1.3 Mio library to find the first hit on rank 14, 4 actives among the first 50 and 15 actives overall after having screened 30,000 compounds. This amounts in an approximate enrichment factor of 7000! (see Figure 3)

Application Phase

In this real life experiment we show how we found the original compound Gleevec, an anti-cancer drug marketed by Novartis, in a library consisting of 22,400,000 compounds. Gleevec was decomposed and a combinatorial library re-assembled from fragment analogues ($22^4 \times 7 \times 5 \times 4 = 22,400,000$ compounds, see Figure 5). After docking this combinatorial library under pharmacophore constraints we found the closest analogue to Gleevec on rank 141 (0.0005% of the library), with 70898 compounds fulfilling the pharmacophore. Running time for this experiment was slightly over 3hrs.

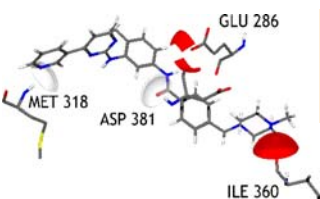


Figure 4: Pharmacophore definition for Gleevec. (PDB code 1IEP). The protein ligand hydrogen bonds most conserved have been defined as pharmacophore constraints.

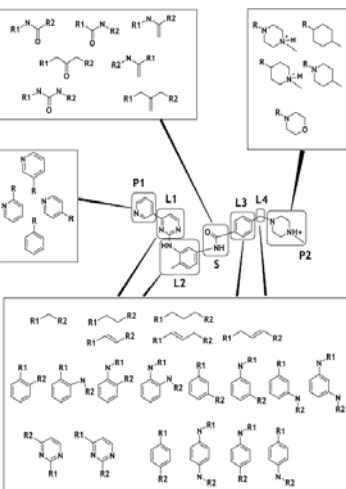


Figure 5: Combinatorial library generated from Gleevec with notation of the R-groups. Central amide scaffold has been denoted S. Other R-groups are denoted P1, L1, L2, L3, L4, and P2 (terminal, on the right). The methyl group in L2 has been left out in the building blocks of the combinatorial library we generated.

References:
 [1] Leach et. al., J. Mol. Graph. Model. 18 (4-5), 358-367 (2000)
 [2] Rarey et. al. J. Comp. Aid. Mol. Des., 12, 471-490, (1998)
 [3] Schindler et. al., Science, 289, 1938-1942 (2000)
 [4] Gastreich et. al. J. Comp. Aid. Mol. Des., submitted

Summary:

- *de novo* design protocol for ultra fast combinatorial compound docking under pharmacophore constraints
 - significant enhancement of enrichment and reduction of false positives
 - speed-up of combinatorial approach (order of 10) could be multiplied by 5 by using pharmacophore definitions
- Applications: *de novo* design, gap bridging between pockets, scaffold hopping, scanning combinatorial libraries

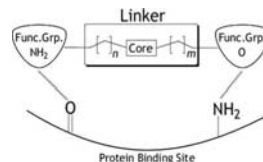


Figure 1: Leach's approach to novel ligand inhibitor compounds [1]: The distance between so-called templates (in our case pharmacophore definitions) is varied by linkers consisting of the core plus side chains.



Figure 2: Result of learning phase, step 1: Methotrexate (MTX) crystal structure and rank 1 docking solution (left). Amino acids of DHFR involved in the pharmacophore definition are labeled. On the right are 2D depictions of the same molecules. Rank 1 solution is the closest possible analogue to MTX (MTX not present in library).

Figure 3: Result of learning phase, step 2b): Logarithmic plot of enrichment for the full 1.35 million library docked with pharmacophore constraints. Enrichment factor ~7000 at 0.0025% of the library screened.

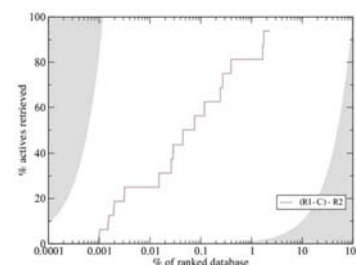


Figure 6: Single point (connected crosses, green) and average (thick blue line) FTrees [2] similarities for the first 1000 virtual compounds of the 22,400,000 library, according to FlexX score. Closest analogue ① to Gleevec found by FlexX out of 22,400,000 compounds was found on rank 141 (0.0005% of the library)!

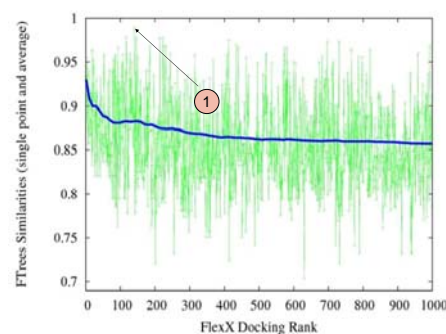


Figure 7: Docking pose of rank 1 solution and the Gleevec crystal structure. [3] The Gleevec structure is depicted in thinner stick representation.

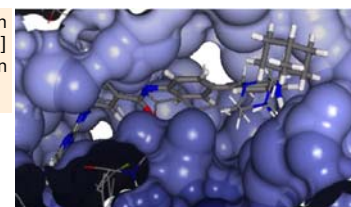


Figure 8: Docking pose of the closest analogue to Gleevec within the 22.4 Mio member combinatorial library (rank 141). The crystal structure of Gleevec is depicted in green.