CoLibri

CoLibri is a toolkit for chemical space exploration comprised of multiple components, the reaction synthesizer, the fragment space merger and Ftrees-FS

How does CoLibri work?

Synthetic access is the key to success in practically all pharmaceutical research projects. What can be made - easily enough - decides in the end what is actually made. The collective know how of chemical reactions in a company is therefore the most valuable intellectual property pool to draw from. Why not capture this know-how and make it accessible throughout the organization? CoLibri can do exactly that.
Starting from what is known in the literature, it lets you transform your collective know how into a chemical space, constituting huge numbers of virtual but synthetically accessible compounds.

- **reaction synthesizer**

reaction synthesizer is a command line tool that takes one or more reaction definitions as an input and generates an individual fragment space for every one of them. In addition the reaction synthesizer can enumerate product molecules for a reaction.

- **fragment space merger**

Another command line tool implemented in CoLibri is the fragment space merger. It takes the output of the reaction synthesizer (multiple individual fragment spaces) and merges them to a single concise fragment space.

- **FTrees-FS**

The extension module to the similarity search method FTrees, which works with a fragment space as input to search through.

Advantages

- Generate new intellectual property
- Expand your literature compound toolkit
- Find purchasable analogs
- Save time and money

de novo design & synthetic access

De novo design was first invented to think outside the box, to generate new ideas, and to come up with entirely novel solutions. However, it's major drawback so far was the lack of synthetic access, hence medicinal chemists often objected to these designs. CoLibri overcomes this issue.
On the one hand side FTrees-FS provides an extremely fast method to generate new chemical entities (NCEs) based on one or many query compounds. On the other hand side synthetic access of the virtual hits is "built-in" since the search space is based on real chemical reactions.

Can I create my own chemical space?

Many companies have a unique collection of in-house chemistries and building blocks, and therefore want to set up their own space, based on those in-house chemistries and building blocks. This will also create the highest level of IP-value. **Pfizer** did that and built PGVL. They turned 358 combinatorial libraries based in validated reaction protocols into a virtual chemical space, resulting in 3 trillion chemically accessible compounds. Read more in the original publication:

Similarity Searching and Scaffold Hopping in Synthetically Accessible Combinatorial Chemistry Spaces
M. Boehm, T.-Y. Wu, H. Claussen, C. Lemmen
2008 51 (8), pp 2468-2480
http://dx.doi.org/10.1021/jm0707727

Find hidden gems

Scientists at **AstraZeneca** did a systematic study about the orthogonality of different similarity search methods. Specifically it was determined across a number of data sets for a variety of targets, which method uniquely found active molecules, namely active molecules that the other methods missed. Later this study was repeated using FTrees confirming that FTrees performed best according to this benchmark.

Multifingerprint Based Similarity Searches for Targeted Class Compound Selection
M. Boehm, T.-Y. Wu, H. Claussen, C. Lemmen
2006 46 (3), pp 1201-1213
http://dx.doi.org/10.1021/ci0504723

No active site – where do I start?

If you are working on a GPCR project (or any other project lacking a protein structure) and cannot follow the traditional structure-based design approach, mining a virtual chemical space will be a useful approach to generate inspiring new chemistries as starting points.
Boehringer-Ingelheim used this approach to propel a GPR119 project. They started out with 4 structurally different literature compounds as queries to search their in-house virtual, BI-CLAIM to generate 10,000 FTrees hits. These were post-processed by a 3D shape filter based on one of the more rigid query compounds. Specific cores of the molecules were selected by the chemistry team as "activity anchors" and focused libraries around these designed and synthesized.
This approach lead to two new GPR119-agonist hit series, one of which subsequently resulted in a new lead class. Read more details in the original publication:

Identification of New Potent GPR119 Agonists by Combining Virtual Screening and Combinatorial Chemistry
M. Rarey, B. Kramer, T. Lengauer, G. Klebe
2012 55 (24), pp 11031-11041
http://dx.doi.org/10.1021/jm301549a

Discover distant neighbors

Using the FTrees similarity search method, **Gedeon Richter** found novel Histamine H_4 receptor and Serotonin transporter ligands. This study also nicely confirms that FTrees is well able to jump chemical classes and find remotely similar molecules – active against the same target – with distinctly different scaffolds.

Discovery of Novel Histamine H4 and Serotonin Transporter Ligands Using the Topological Feature Tree Descriptor
R. Kiss, M. Sándor, A. Gere, É. Schmidt, G.T. Balogh, B. Kiss, L. Molnár, C. Lemmen, G.M. Keseru
2012 52 (1), pp 233-242
http://dx.doi.org/10.1021/ci0504723