The information system in a small biotech company: overview of the cheminformatics developments at Aptanomics

Dr. Jérôme Pansanel*, Dr. Matthieu Schapira
Aptanomics S.A.
181-203, avenue Jean Jaurès
69007 Lyon, France
jerome.pansanel@aptanomics.com

Aptanomics discovers and develops new drugs in major markets of unmet medical need. The Company focuses on well-validated but hard-to-crack therapeutic targets that are involved in signalling pathways that are deregulated in a broad range of cancers. Our technology platform, based on peptide aptamers, enables us to identify new druggable sites on these targets. Small molecule hits to the newly identified sites are discovered by virtual screening or by AptaScreen™.

To support our lead discovery and development activity, we have developed an intranet chemical/biological information management system based on Open Source software. It can be accessed through a web portal for strategic planning or day-to-day operational support. As our business is to develop new drugs, the main component of this system is dedicated to chemistry. Cheminformatics developments cover the entire life cycle of molecules in our pipeline:

1-Chemical library management. An interface with both our internal collection of over 70,000 compounds and a virtual library of several million molecules is in place. Molecular descriptors and fingerprints are calculated and stored within a relational database environment, for rapid querying and analysis. Since part of these libraries feed the screening platform and the hit2lead team, the project includes also a stock management system. The structural data, available to all, is used for virtual screening and SAR.

2-HTS data management and analysis. Tools are implemented for the management of data coming from our high throughput screening platform, AptaScreen™. HTS data is linked to molecular properties and can be analyzed with varying level of complexity, from visual inspection of HTS signals in plate format, to scattered plot analysis that allow for any combination of successive graphical selections through heterogeneous data types in order to zoom-in onto most relevant hits.

3-Lead development and pipeline snapshot. We have developed a system to support our lead optimization activity. Different data types can be attached to molecules, such as drug-likeness, *in vitro* and *in vivo* biological activity, and traced within graphical representations of chemical series, for a synthetic depiction of the advancement of our pipeline.