

# DESIGN OF TARGET-FOCUSED LIBRARIES FOR VIRTUAL SCREENING

Elena Fioravanzo (a), Maria Alessandra Alisi (b), Arianna Bassan (a), Nicola Cazzolla (b), Rosella Ombrato (b), Gaia Corso (b).

(a) S-IN Soluzioni Informatiche Via Salvemini 9, 36100, Vicenza, Italy ([www.s-in.it](http://www.s-in.it))  
(b) Angelini Research Center, ACRAF, P.le Stazione I-00040, S. Palomba, Pomezia (Rome), Italy

## INTRODUCTION

Structure-based virtual screening (VS) has evolved over the past few decades as a powerful tool for identification of novel and diverse lead structures[1]. The problem of false positives and negatives in virtual screening is particularly apparent when docking large databases, where the number of compounds satisfying a "good" docking score can be much larger than assay throughput capabilities. Assaying only the top-scoring compounds is not a reasonable strategy because a high percentage of promising compounds might be disregarded.

A common strategy includes pre-docking filters on calculated physical or ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) properties in order to reduce the number of compounds to dock. However imposing hard-cut ADME or property-based filters may prevent finding new scaffolds.

Here we present a flexible and intelligent strategy for designing and filtering large libraries using pre-docking and post-docking filters based both on physical-chemical and ADMET ligands properties and structural information of a binding site.

## METHODS

**Library generation.** Optimized 3D structures are generated by means of LigPrep[2], including Epik[2,3] enantiomers, tautomers and protonation states at pH= 7±1 enumeration.

**ADMET properties.** A wide variety of pharmaceutically relevant properties are predicted by means of QikProp[2,4] and employed to optimize ADMET library profile.

**Diversity analysis.** A cheminformatics approach is employed to taper down the large compound library without losing its information of original molecular properties as well as structural diversity. A diverse subset is selected by means of a cluster analysis based on fingerprints carried out with Canvas[2] in order to filter out redundant or similar compounds and design optimal diverse subsets to enhance the structural diversity of the large initial library.

**Docking.** A workflow for virtual screening is applied. Flexible and ensemble docking to account for both the ligands and the receptor flexibility is adopted and performed with Glide[2,6]. A progression of precision docking is performed: 1) High-throughput virtual screening (HTVS) docking, intended for the rapid screening of very large numbers of ligands; 2) SP (standard precision) docking; 3) XP (Extra-precision) docking and scoring, a more powerful and discriminating procedure.

**Pre-docking analysis.** If a collection of known active molecules are available, abstract 3D pharmacophore models are generated with Phase[2,7] from these compounds by extracting the common spatial arrangement of pharmacophoric features. These models are applied to filter the library and identify compounds that also satisfy the pharmacophore.

**Post-docking analysis.** Various post-docking filters were applied to eliminate false positive: 1) Structural Interaction Fingerprints[8] 1D binary bit-strings that represent important target-ligand binding interactions useful to reproduce the true binding modes of the compounds and thereby to obtain improved library enrichments from virtual screening; 2) geometrical filters; 3) visual inspection.

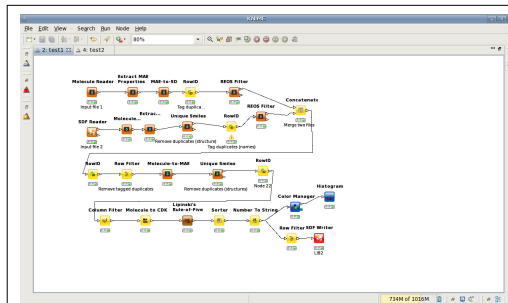
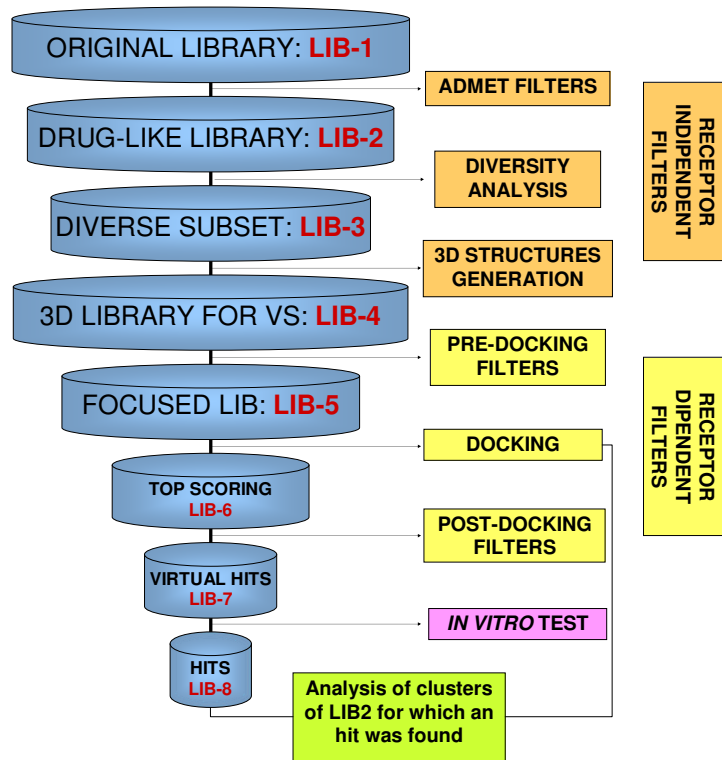
**Integration tool.** The entire procedure is implemented in workflows built employing Knime[9] and the Schrödinger Knime Extensions[2]. Knime is a popular modular data exploration platform that enables the user to visually create data flows (often referred to as pipelines), selectively execute some or all analysis steps, and later investigate the results through interactive views on data and models. The Knime Extensions are the Schrödinger nodes built upon the existing Knime infrastructure, providing access to a wealth of ligand- and structure-based tools.

## CONCLUSION

The development of a flexible and intelligent strategy for focusing both large and small libraries for a successful virtual screening is described. The main two challenges were: 1) to achieve hit lists in which the likelihood that the selected compounds display the desired affinity is increased; 2) and to integrate different techniques and tools by means of a workflow technology in order to reduce time and cost-intensive steps in the drug development process. Therefore, a combination of virtual techniques was employed to reduce and to focus the starting library with the aim of increasing the efficiency in identifying novel active structures applying a rational approach. The developed workflows allow the user to easily:

- 1) set up filters cut-off values;
- 2) set up docking parameters;
- 3) prepare all-atom 3D library structures starting with 2D or 3D structures in SD format;
- 4) prepare receptors models;
- 5) run pre-docking filtering, docking and post-docking filtering;
- 6) perform a diversity selection;
- 7) explore the results.

## RESULTS



Various workflows were developed in Knime employing both native nodes and the Schrödinger Knime Extensions in order to integrate different tools into a single infrastructure. The displayed workflow generates the drug-like library (LIB-2) and includes the following steps: 1) library pre-processing (merging of different files and duplicate eliminations); 2) ADMET filters.

## REFERENCES

- [1] *Virtual Screening in Drug Discovery*, edited by Juan Alvarez and Brian Shoichet. Taylor & Francis Group, LLC, CRC Press, 2005. ISBN 0824754794.
- [2] Schrödinger, L.L.C., New York, [www.schrodinger.com](http://www.schrodinger.com).
- [3] *J. Comput. Aided Mol. Des.*, 2007, 21, 681–691.
- [4] *J. Chem. Inf. Model.*, 2007, 47, 2416–2428.
- [5] *J. Comput. Aided Mol. Des.*, 2009, 23, 527–539.
- [6] *Chem. Biol. Drug Des.*, 2006, 67, 370–372.
- [7] *J. Med. Chem.* 2006, 49, 490–500
- [8] Prof. Dr. Michael Berthold, University of Konstanz, Germany, [www.knime.org/index.html](http://www.knime.org/index.html).