QSAR-based Virtual Screening to Identify New Potential Anticancer Agents: Targeting Bcl-2 pro-survival proteins

Arthur C. Silva[1]* (PG), Ney R. Toledo[1] (IC), Driele C. P. Almeida[1] (IC), Rodolpho C. Braga[1] (PQ), Eugene Muratov[1,2] (PQ), Carolina H. Andrade[1] (PQ)**

***arthurcs.farma@gmail.com** ****carolina@ufg.br**

[1]LabMol – Laboratory for Molecular Modeling and Drug Design, Faculty of Pharmacy, Federal University of Goiás, Goiania – GO, Brazil.
[2]Laboratory for Molecular Modeling, Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA.

Keywords: *Apoptosis, Bcl-2, QSAR, drug design, virtual screening, cancer chemotherapy*

Abstract

We report a QSAR-based virtual screening to select new promising anticancer agents acting as Bcl-xL protein inhibitors.

Introduction

Pro-survival proteins are Bcl-2 family members, related to the intrinsic pathway of apoptosis[1]. These proteins are overexpressed in many malignancies and correlated to therapeutic resistance and poor prognostics, once they turn the cancer cells able to evade apoptosis[2]. Bcl-xL is an anti-apoptotic protein, member of Bcl-2 family and is a validated target in cancer drug discovery pipeline[3,4]. The aim of this work was divided in three steps: (i) to compile, integrate and curate the largest publicly available data set of compounds with reported activity against Bcl-xL; (ii) to generate and rigorously validate QSAR models in compliance with OECD principles; (iii) to perform a virtual screening of a protein-protein interaction focused database in an attempt to identify new Bcl-xL inhibitors as potential new anticancer agents candidates.

Results and Discussion

The largest publicly available data sets containing biological annotation related to biological assays with Bcl-xL pro-survival protein were found in ChEMBL and PubChem Bioassay. Initially, the total number of entries was around 214,000 inactive compounds, e.g. activity > 10μM and 840 active compounds, e.g. activity ≤ 10μM. After rigorous curation process, 212,717 inactive and 428 active compounds remained and, to balance the final dataset (1:1), a hierarchical cluster analysis was conducted. QSAR models were built through the combination of three machine learning methods (Support Vector Machine, Random Forest, Gradient Boosting Machine) and five molecular descriptors (MACCS, DRAGON, Morgan, CDK, Atom Pairs). Consensus models were built by the combination of the best individual models. The consensus model 2 was used in the virtual screening to search for potential active compounds inside a protein-protein focused ASINEX library of 11,387 compounds. The softwares FILTER[5], OMEGA[5] and QUACPAC[5] were used in the pre-processing of ASINEX database to filter, generate conformers and calculate partial charges, respectivelly.

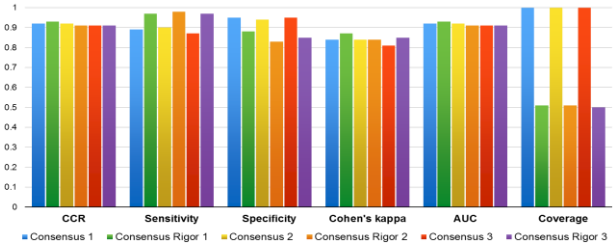

Figure 1. Results for the best QSAR consensus models.

The top ten virtual hits were selected for experimental validation as Bcl-xL inhibitor candidates and, hence, as new potential anticancer agents.

Conclusions

The largest publicly available data sets for Bcl-xL were compiled, integrated and curated and served as source to generate robust and highly predictive QSAR models accordingly to OECD guidelines. The best QSAR consensus model was used in the virtual screening to prioritize compounds in a protein-protein interaction focused database and the top 10 hits were selected for acquisition and experimental validation.

Acknowledgements

CNPq, CAPES, FAPEG, OpenEye Scientific Software and Inte:Ligand.

[1] Cerella, C.; Teiten, M.-H.; Radogna, F.; Dicato, M.; Diederich, M. *Biotechnol. Adv.* **2014,** 32**,** 1111–1122.
[2] Goard, C. A e Schimmer, A. D. *Core Evid.* **2013,** 8**,** 15–26.
[3] Akl, H. *et al. Biochim. Biophys. Acta*. **2014,** 1843**,** 2240–52.
[4] Leverson, J. D. *et al*. *Sci. Transl. Med.* **2015, 7,** 279.
[5] OpenEye Scientific Software, Inc., Santa Fe, NM, USA, **2011**.