

New Virtual Hits of Pyruvate Kinase of *Leishmania* spp. Identified by Integrated Structure and Ligand-Based Virtual Screening

Cleber C. Melo-Filho¹ (PG), Elias M. Laurentino¹ (IC), Rodolpho C. Braga¹ (PQ), Eugene N. Muratov² (PQ), Carolina H. Andrade^{1*} (PQ)

*carolina@ufg.br

¹LabMol - Laboratory for Molecular Modeling and Drug Design, Universidade Federal de Goiás, Rua 240, Qd.87, Goiânia, Goiás, Brazil.

²Laboratory for Molecular Modeling, Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599, USA.

Keywords: *Leishmaniasis*, LBVS, SBVS, QSAR, PAINS.

Abstract

We report the identification of new potential inhibitors of PK of *Leishmania* spp. by integrated virtual screening approach.

Introduction

Leishmaniasis is a group of diseases caused by protozoan parasites of genus *Leishmania*. According to the World Health Organization, 1.3 million new cases and around 20,000 deaths occur annually.¹ Pyruvate kinase (PK) of *Leishmania* catalyzes the final reaction of glycolysis for production of ATP, playing an important role in energy metabolism.² In this study, validated ligand-based virtual screening (LBVS) and structure-based virtual screening (SBVS) workflows were applied for screening of new potential PK inhibitors from ChemBridge database. Furthermore, a filter was used to exclude Pan-Assay Interference Compounds (PAINS). Finally, QSAR models using different descriptors and machine learning classifiers were combined in a consensus approach and used as final filters to help the selection of compounds. The virtual hits identified in this study will be experimentally evaluated against *Leishmania* spp. amastigotes.

Results and Discussion

The ChemBridge database containing more than 1 million compounds was used for VS. The LBVS model based on molecular shape was generated on ROCS software v.3.2.1.4.³ The SBVS model based on molecular docking was developed using FRED module of OEDocking package v.3.0.1^{4,5}, with the structure of PK co-crystallized with inhibitor suramin (PDB ID: 3PP7). LBVS, SBVS and QSAR models were validated using a dataset of 291,193 compounds tested against PK from *L. mexicana* (PubChem Bioassay AID 1721). Firstly, the dataset was curated according to the protocol reported by Fourches et al.⁶ and a threshold of 10 μ M was set to divide the dataset (307 actives and 288,777 inactives). Then, dataset balancing was performed by linear under-sampling, using MACCS fingerprints and *k*-Nearest Neighbors (*k*-NN) method. The SBVS

and LBVS were validated using a dataset balancing of 1:36 (actives/inactives). QSAR models were generated using balanced (1:1) and unbalanced (1:2, 1:3 and 1:4) datasets. Five different molecular fingerprints (MACCS, Morgan, FeatMorgan, AtomPair and Avalon) and two machine learning classifiers (SVM and Random Forest) were used to generate QSAR models. Consensus QSAR models were generated by different combinations of five individual models. In the first step of VS (1,063,926 compounds), Lipinski's rule of five and Veber's filters were applied and undesirable compounds removed (metals, salts, insoluble, aggregators). Then, the LBVS model was applied and compounds with TanimotoCombo > 1.0 (22,091 compounds) were selected. After the SBVS, the 4,000 top ranked compounds by ChemGauss4 score were selected. An additional filter was used to exclude PAINS. Finally, the best QSAR consensus model was applied as a final filter. All steps of QSAR modeling were performed in our in house KSAR 1.6.2 workflow implemented on KNIME (<http://labmol.farmacia.ufg.br/ksar>).

Conclusions

The validated LBVS, SBVS and QSAR models were able to identify new potential inhibitors of PK of *Leishmania*. The virtual hits identified in this study will be experimentally evaluated against *Leishmania* spp. amastigotes. Then, the most promising hits will be tested in enzymatic assays using PK enzyme.

Acknowledgments

CNPq, CAPES, FAPEG, ChemAxon and OpenEye Scientific Software.

¹WHO. Leishmaniasis Fact sheet N°375 <http://www.who.int/mediacentre/factsheets/fs375/en/>.

²Morgan et al., *J. Biol. Chem.* **2011**, 286, 31232–31240.

³ROCS 3.2.1.4: OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>.

⁴OEDocking version 3.0.1. OpenEye Scientific Software, Santa Fe, NM, USA, <http://www.eyesopen.com>.

⁵McGann M. FRED pose prediction and virtual screening accuracy. *J Chem Inf. Model.* **2011**, 51, 578–596.

⁶Fourches, D.; Muratov, E.; Tropsha, A. *J. Chem. Inf. Model.* **2010**, 50, 1189–1204.