A QSAR Study of Human Thymidine Phosphorylase Inhibitors with SMILES-Based Descriptors

Adriana Santos Costa, Theoretical, Medicinal and Environmental Chemistry Laboratory (LQMAT), Department of Pharmacy, Western Paraná State University (UNIOESTE), Cascavel, Brazil
Eduardo Borges de Melo, Theoretical, Medicinal and Environmental Chemistry Laboratory (LQMAT), Department of Pharmacy, Western Paraná State University (UNIOESTE), Cascavel, Brazil

ABSTRACT
Thymidine phosphorylase (TP) is a multifunctional protein frequently overexpressed in many types of cancer. Considering the interest in new anticancer compounds, a QSAR study was carried out to investigate a set of uracil derivatives described as TP inhibitors. The only molecular descriptors used were derived from SMILES notation. Ordered Predictors Selection (OPS) was used for variable selection and the models were built using the PLS method. The authors validated the internal and external prediction capabilities of the obtained model. The model was also tested using a set of 12 molecules obtained by similarity search in the ZINC database. The results showed that it is possible to describe the variation of biological activity of the selected dataset using only SMILES-derived molecular descriptors, and the obtained model shows potential for use as an aid in the design of new TP inhibitors.

KEYWORDS
Applicability Domain, Cancer, Mechanistic Interpretation, Model Validation, OPS, PLS, Structure Activity Relationships, Thymidine Phosphorylase

INTRODUCTION
Currently, cancer treatment remains a challenge owing to the wide variety of cancers and the similarity between normal and tumor cells. Thus, novel chemical entities and biological targets for the treatment are constantly being investigated (Ganesan, 2007).

Among these targets is human thymidine phosphorylase (TP), a multifunctional protein that has an important role in pyrimidine nucleoside metabolism and is associated with the maintenance of healthy mitochondria and the recovery of cells from pathological stress. However, TP is overexpressed in a wide range of human cancers (bladder, colorectal, esophageal, pancreatic, and others) and has been associated with the promotion of tumorigenesis, angiogenesis, metastasis, and the inhibition of apoptosis. TP is thus useful for metabolic activation in loco of capecitabine (Xeloda®) and doxifluridine, two 5-fluouracil prodrugs (Figure 1) (Jain, Rasheed, & Kalman, 2010; Toi, Rahman, Bando, & Chow, 2005; Norman et al., 2004; Ye & Zang, 2013).

As various types of anticancer treatments (e.g. chemotherapy and radiotherapy) stimulate TP expression, probably owing to their pro-angiogenic and antiapoptotic properties, the suppression of this protein could hypothetically slow down or suppress the recovery of cancer cells, thus providing beneficial supplementary effects (Yano et al., 2004; Toi, Rahman, Bando, & Chow, 2005). Lonsurf®
(TAS-102, Taiho Pharmaceuticals Inc.) is approved in Japan for the treatment of unresectable advanced or recurrent colorectal cancer. One of the active principles is 5-chloro-6-[(2-iminopyrrolidin-1-yl)methyl]uracil (TPI or tipracil) (Figure 2), which is formulated in association with 5-trifluoromethyl-2’-deoxyuridine (Jain, Rasheed, & Kalman, 2010). However, despite its great therapeutic value, tipracil is still the only TP inhibitor clinically available.

In the field of computer-aided molecular design (CADD), the quantitative structure-activity relationship (QSAR) method is useful in lead optimization, since it increases the probability of finding novel drug candidates by avoiding the synthesis and biological screening of few potential molecules and thus saving time and money (Roy, Kar, & Das, 2015). This method considers that a specific biological activity can vary as a function of molecular descriptors encoding information about the chemical structure of a set of molecules. Thus, a model containing the calculated descriptors could be used to predict the responses of new compounds (Ribeiro & Ferreira, 2003; Molfetta, Bruni, Rosseli, & Silva, 2007).

Considering the need for new anticancer agents as well as the clinical validation of TP as a biological target, a QSAR study was carried out in the present study with the aim of obtaining models that could aid in the design of new inhibitors of this protein. For this, a dataset of 37 uracil derivatives previously described by Yano et al. (2004a; 2004b) (Table 1) which proved able to inhibit human
Related Content

Multilayered Variable Selection in QSPR: A Case Study of Modeling Melting Point of Bromide Ionic Liquids
www.igi-global.com/article/multilayered-variable-selection-in-qspr/171149?camid=4v1a

Mining a Nanoparticle Dataset, Compiled Within the MODENA-COST Action
www.igi-global.com/article/mining-a-nanoparticle-dataset-compiled-within-the-modena-cost-action/216895?camid=4v1a
Evolution of Optimal Descriptors: Solved, Unsolved, and Unsoluble Tasks
www.igi-global.com/article/evolution-of-optimal-descriptors/146170?camid=4v1a

The History and Development of Quantitative Structure-Activity Relationships (QSARs): Addendum
www.igi-global.com/article/the-history-and-development-of-quantitative-structure-activity-relationships_qsars/181618?camid=4v1a