QSAR studies of combretastatin-like chalcones using PLS and ANN methods

Célio F. Lipinski (PG), Aline A. Oliveira (PG), Albérico B. F. da Silva (PQ)

Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, SP, Brazil.
alineoliveira@iqsc.usp.br

Keywords: chalcone, QSAR, PLS, MLP-ANN

INTRODUCTION

Several compounds with antimitotic properties have been studied for the cancer treatment and most of them interact with components of the microtubules[1]. Although many drugs are available for the cancer treatment, the high toxicity and resistance limit their use. Therefore, it is necessary to develop new therapeutic alternatives. For this purpose, chalcones have been extensively studied since they inhibit the polymerization of microtubules in cancer cells by interacting with β-tubulin[2]. In order to obtain predictive mathematical models for the biological activity of 87 combretastatin-like chalcones[3] and to identify structural features associated to their biological activity we have developed QSAR studies using Partial Least Square (PLS) and Artificial Neural Network (ANN) methods.

METHODS

Gaussian 09 software was used to perform optimization and frequency calculations on the chalcones structures using DFT/B3LYP with 6-31G++(d,p) basis set. Molecular descriptors were obtained by Dragon software and electronic descriptors were obtained by the quantum calculations. Correlation coefficients between descriptors and biological activity and genetic algorithm analysis (BuildQSAR software) were used to select the most relevant descriptors. Piroutte and Matlab softwares were used to perform PLS and Multilayer Perceptron ANN (MLP-ANN) analysis, respectively. Additional validation tests were performed on QSARModeling software.

RESULTS AND DISCUSSION

Correlation coefficients and genetic algorithm analysis allowed us to select the following descriptors: RDF045e, RTv, RDF155u, RDF035m, SP02, UNIP, PI and E_{HOMO-3}. We have obtained an acceptable PLS model using the selected descriptors (r^2 = 0.756, q^2 = 0.681 and r^2_{test} = 0.843; Figure 1a). The validation tests showed that the PLS model is robust and it was not obtained by chance correlation. An ANN model was also obtained by training a MLP-ANN with 8-12-1 architecture (r^2 = 0.803, q^2 = 0.761 and r^2_{test} = 0.663; Figure1b).

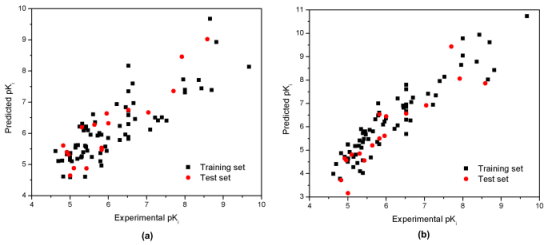

Figure 1. Biological activity prediction by (a) PLS and (b) ANN models.

CONCLUSIONS

The PLS and ANN models were able to predict the biological activity of the studied combretastatin-like chalcones and the PLS model showed better results indicating a linear behavior for the selected descriptors. Additionally, the selected descriptors have shown important features of the combretastatin-like chalcones structures that can be used for the rational drug design of new analogues.

ACKNOWLEDGMENTS

CNPq, CAPES and FAPESP.

[1]Wilson, L. and Jordan, M. A. Nature 4 (2004) 252-265.
[2]Bergstralh, D. T. and Ting, J. P. Cancer Treat. Rev. 32 (2006) 166-179.
[3]Ducki, S. *et al*. Bioorg. Med. Chem. 17 (2009) 7698-7710.