12th International Electronic Conference on Synthetic Organic Chemistry (ECSOC-12) 1-30 November 2008
http://www.usc.es/congresos/ecsoc/12/ECSOC12.htm & http://www.mdpi.org/ecsoc-12

[G0006]

Theoretical Prediction of Antiproliferative Activity against Murine Leukemia Tumor Cell Line (L1210). 3D-Morse Descriptors and its Application in Computational Chemistry.

Liane Saíz-Urra [a,c] ,Yunierkis Pérez-Castillo [a,c], Maykel Pérez González[c], Reinaldo Molina Ruiz [a,b], M. Natália D. S. Cordeiro [a] , J. Enrique Rodríguez-Borges [b,*] , Xerardo García-Mera[d]

[a]*REQUIMTE,* [b]*CIQ,* Chemistry Department, Faculty of Sciences, University of Porto, 4169-007 Porto Portugal. [c]Molecular Simulation and Drug Design Group, Chemical Bioactive Center, Central University of Las Villas, Santa Clara, Villa Clara, C.P. 54830, Cuba. [d]Departamento de Química Orgánica, Facultade de Farmacia, Universidade de Santiago de Compostela, E-15706 Santiago de Compostela, Spain.

ABSTRACT: Cancer is among the top ten causes of death in the world but in spite of the efforts of the pharmaceutical companies and many governmental organizations, new and more effective drugs are urgently needed. Computer assisted studies have been widely used to predict anticancer activity taking into account different molecular descriptors, statistical techniques, cell lines and data sets of congeneric and non-congeneric compounds. This paper describes a QSAR study and the successful application of 3D-MoRSE descriptors for developing Linear Discriminant Analysis (LDA) to predict the anticancer potential of a diverse set of indolocarbazoles derivatives. Despite the structural complexity of this sort of compounds the used descriptors are able to identify the most remarkable features like the incidence of polarizability of the substituents and the interatomic distance in the 7-azaindole moiety in the antiproliferative activity. A comparison with other approaches such as the Getaway, Randić molecular profile, Geometrical, RDF descriptors, was carried out showing the model with 3D-MoRSE descriptors resulted in the best accuracy and predictive capability. An LDA based desirability analysis was conducted to select the levels of the predictor variables which should generate more desirable drugs, i.e. with higher posterior probability to be classified cytotoxic.

Keywords: QSAR; Anticancer activity; Indolocarbazoles derivatives; 3D-MoRSE.

*Corresponding author. Fax: +351 226082959. e-mail: jrborges@fc.up.pt

Introduction

One of the most important issues in Medicinal Chemistry is cancer, which encompasses a group of diseases characterized by the excessive and uncontrolled growth of cells invading and impairing tissues and organs and can, eventually, result in the death. In 2005, 7.6 million of the 58 million deaths registered in the world were caused by cancer. Over 70% of these deaths were in countries with low or average incomes where the resources for the diagnosis and the treatment of the disease are limited or even nonexistent [1].

The search for new anticancer drugs plays a central role in the research programs of pharmaceutical companies but also those of many governmental organizations [2]. However, it is estimated that the rate of incidence of cancer far from decreasing will rise to about 9 million in 2015 and 11.4 million in 2030. Hence, new and effective drugs are increasingly and urgently needed. A large number of anticancer agents have been discovered that act at different levels [3] and have higher efficacy and lower toxicity than existing treatments. These databases can be exploited with the help of automated and multivariate data analysis methods [4-6]. The latter relates the molecular structures with their biological properties by establishing computational models able to assign activity values to new untested compounds [7, 8].

QSAR techniques in anticancer activity studies have previously reported the use of different molecular descriptors, statistical techniques, cell lines and data sets of congeneric and non-congeneric compounds as well as the respective toxicological assays of these compounds [4, 9-17].

An interesting group of compounds is indolocarbazole derivatives whose properties as protein kinase C and topoisomerase I inhibitors have been widely studied [18, 19]. Rebeccamycin, a microbial metabolite isolated from cultures of Saccharothrix aerocolonigenes, which belong to this group, is an antitumor antibiotic that inhibits topoisomerase I by stabilizing the topoisomerase I-DNA cleavable complex [20, 21]. Also, it has been shown that although topoisomerase I is a target for most of the rebeccamycin derivatives, the inhibition of other enzymes may also be a contributing factor to their cytotoxicity. However, its toxicity prohibits its use in cancer chemotherapy. Structure–activity relationship studies have been carried out with the purpose of improving the pharmacological profile of rebeccamycin,[19, 22] and have led to the development of a schematic representation of a drug-topoisomerase I-DNA ternary complex.

In spite of its promise and the diversity and quantity of derivatives developed, no anticancer QSAR studies taking into account these types of compounds have been reported. In this paper we report a QSAR model for the rational selection of anticancer compounds which involves a diverse data set of indolocarbazoles derivatives. The use of the 3D descriptors is reported as well, specifically the 3D-MoRSE, owing to the flexibility of these descriptors. They afford the possibility for choosing an appropriate atomic property and in this way we could adapt them to the specific problem under study. Besides, these descriptors present an advantage as they code with fixed-length representation of 3D molecular structure, allowing us to compare the datasets comprising of molecules of different sizes, and number of atoms [23, 24].

Materials and methods

Data sets:

In the present study we used a data set of 125 compounds whose anticancer activity against murine leukemia tumor cell line (L1210) has been previously reported. Eligible compounds were determined by reviewing the literature [22, 25-37].

The data encompasses rebeccamycin analogues from indolo[2,3-*c*]carbazole, indolocarbazoles bearing amino acid residues, sugar units linked to both indole nitrogens, 7-azaindole moieties or different substituents on the indolocarbazole framework. Another group consists of dipyrrolo[3,4-a:3,4-c]carbazole-1,3,4,6-tetraones, substituted with various saturated and unsaturated side chains, indolylpyrazolones and indolylpyridazinedione. Finally, we studied isogranulatimide and bis-imide granulatimide analogues modified on the indole moiety and on the imide heterocycles. Cytotoxicity was measured by the microculture tetrazolium assay as described by Leonce, S. *et. al.* [38]. Results are expressed as IC_{50}, the concentration at which the optical density of treated cells with respect to untreated controls is reduced by 50%.

Resulting from the need for more potent and less toxic new anticancer drugs, we established the threshold value of activity IC_{50} equal to 10 µM, thereby only the compounds with an activity value lower than the aforementioned were considered as active.

In order to obtain validated QSAR models the dataset was divided into training and test sets. Ideally, this division should be performed such that points representing both training and test sets are distributed within the whole descriptor space occupied by the entire dataset, and each point of the test set is close to at least one point of the training