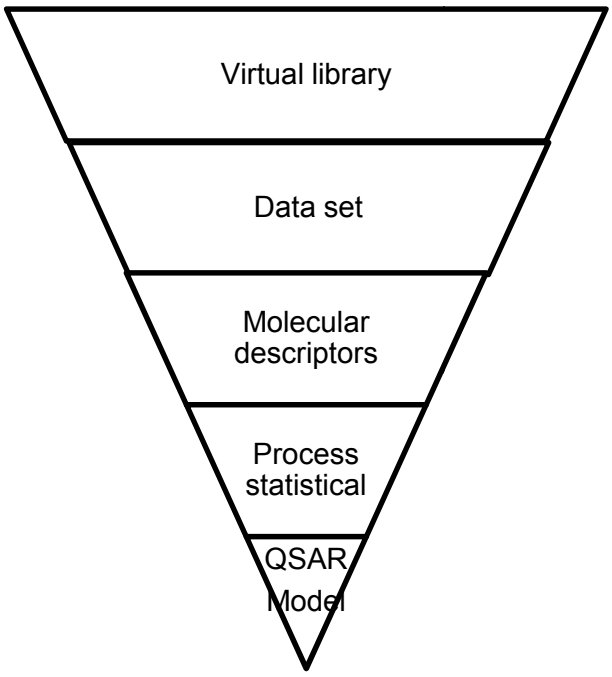
MOL2NET, International Conference Series on Multidisciplinary Sciences
Insert the title of the workshop here

Obtaining a computer-assisted QSAR model for the prediction of anti-inflammatory activity

Luis A Torres Gomez [a], *Dayrenis Garcia* [b], *Juan Carlos Polo* [a], *Laura Machin* [a]
[a] *Institute of Pharmacy and Foods. University of the Havana*
[b] *Center of Investigation and Development of Drugs. Cuba*

.
.
.

Graphical Abstract



Abstract.

The main objective of this study was to develop quantitative structure-activity relationships (QSAR) for the classification and prediction of anti-inflammatory activity. To this end, the ToSS-MoDE approximation was applied for the calculation of the spectral moments of the adjacency matrix between edges of the molecular graph with suppressed hydrogens, weighted on the main diagonal with moments of bond dipoles, bond distance, Van der Waals radius, polarizability and hydrophobicity to 509 active and inactive compounds. The calculated descriptors were used in the design of a training series and a prediction series. With the training series, a discriminant function was developed for the anti-inflammatory activity and another function to characterize the potential of these drugs using the Multivariate Linear Discriminant analysis, obtaining a good total classification of 96.07%. The model was validated by using the external prediction series, obtaining a good classification of 92.59%.

Introduction

.

..Researchers around the world are engaged in the search for new leading compounds that exhibit anti-inflammatory activity and lack the undesirable adverse effects characteristic of these drugs.
Traditionally, the method used to search for new active ingredients is based on the “trial and error”

system through massive tests of a large number of chemical substances, it is increasingly inefficient, since it is done It is necessary to test more than 10,000 compounds, of which 10 pass all the tests and only 1 can become a prescription medicine.

These unfavorable characteristics (ineffectiveness, high cost and high time consumption) make the "traditional" method of random evaluation irrelevant for developing countries and even for large pharmaceutical transnationals.

If we take into account that the limiting step in the discovery and development of new drugs continues to be the identification and optimization of leading compounds in an effective way (in the shortest possible time and at a reasonable cost), the drug design / discovery approach Computer-assisted offers an alternative to the real world of synthesis and evaluation. (1-3). This procedure encompasses all computer-assisted techniques used in the design / discovery and optimization of compounds with desired properties and has played a fundamental role in the development of a number of drugs that are now on the market. (4-6)

This type of 'in silica' procedure avoids costly tasks for current syntheses and bioassays, which are done only after the exploration of the initial concepts with Quantitative Structure-Activity / Property Relationships [internationally known by its acronym in English QSAR / QSPR (Quantitative Structure-Activity / Property Relationships). (7,8).

Materials and Methods

.Using the ACDLABS_v.10_0 software, the molecular structures of the compounds included in the series were represented, which were saved as SMILES (Simplified Molecular Input Line Entry Specification), abbreviated representation codes that can be imported by other softwares. The ACDLABS_v.10_0 was selected given the possibility it offers to use a large database available that contains more than one hundred thousand organic molecules, which minimizes the possibility of making mistakes by manual representation of the molecules under study. In this software you can find the names, synonyms and structural formulas for each of the compounds.

. Molecular descriptors

Once the molecular representations (SMILES) were imported using the MODESLAB software, the atomic or binding parameters that were used for the calculation of the molecular descriptors (spectral moments) of each compound were selected. Of the total of 15 weighting parameters for the calculation of the molecular descriptors that Modeslab allows, the five that are considered most related to biological activity are selected

.

.

Results and Discussion

.

.The applicability and representativeness of the present method critically depend on the selection of the compounds that are used as a training series to construct the classification model. The most critical aspect in the construction of the training series is to ensure the great structural diversity of the data that guarantees a greater domain of application of the model. In order to achieve this diversity, a data composed of a large number of molecular entities was selected, some reported as anti-inflammatory (are considered active) and the rest with a series of compounds from other pharmacological activities

(inactive). The data of both active and inactive compounds was selected considering a broad representativeness of the different structural nuclei.
The latter drugs include antibiotics, antifungals, antivirals, antibacterials, antihypertensives, vasodilators, antineoplastic agents, cardiotonics, antihistamines, sedatives, antidepressants, diuretics, etc. (approximately 15 compounds per pharmacological group)
In the process of reducing variables using the general discriminant analysis, their significance was considered; As a result of this analysis, 36 variables were selected, from the initial 76 (molecular descriptors previously calculated for the training series) with which the linear discriminant analysis (ADL) was carried out in order to identify the set of descriptors with high capacity to classify the compounds under study, according to their activity. As there are two groups there is only one discriminant function.

Biological activity **(FD)** =- 6,4579 μ_0– 8,6160 STD_1+1,6640 STD_3 -0,0021 STD_8 +5,0368 DIP_1 - 2,3309 DIP_2 + 0,5767 DIP_4– 0,1708 DIP_5 + 0,0145 DIP_6 - 4,2269 HYD_1+ 1,5187 HYD_3 -0,1424 HYD_5 + 3,2138 POL_1+ 0,9574 POL_4 – 0,6255 POL_5+ 0,1225 POL_6-0,0077 Pol_7 + 2,1927 VAN_3-2,7119

Lambda de Wilks	Chi squared	**Sig.**	**F de Fisher**	Canonical correlation	**D^2**
0,4396	362,4858	0,000	36,472	0,748620	7,14

Classification matrix for the training series

Biological activity	Well classified	Badly classified
actives	**81**	**16**
inactives	**408**	**4**
total	**485**	**20**

References(

.

1. .Ali, H., et al., Mechanisms of inflammation and leukocyte activation. Adv Reumatol, 1997: 1-28.

2. Brune K, Rainsford KD, Schweitzer A. Biodistributión of mild analgesics Br J Clin Pharmacol 1980:279-84.

3. Carda M. Máster en Química Aplicada y Farmacología. Universidad Jaume I. Inflamación: síntesis de antiinflamatorios. 2011.

4. Cramer III, R. D., Patterson, D. E. and Bunce, J. D. J. Am. Chem. Soc., (1988)

5. *Crofford LJ. COX-1 and COX-2 tissue expression: implications and predictions. J Rheumatol 1997: 15-9.*

6. *David E. Golan y col. Principles of Pharmacology: The Pathophysiologic Basic of Drug Therapy, 2nd Edition. Chapter 40: Principles of Inflammation and the Immune System, p 736-74. Chapter 41: Pharmacology of Eicosanoids. p. 748-762,2007.*
7. *Gerlag, D.M., et al., The effect of a T cell-specific NF-κB inhibitor on in vitro cytokine production and collagen-induced arthritis. J Immunol, 2000: 1652-1658.*

8. *Golbraikh A 2002, Tropsha A. Predictive QSAR modeling based on diversitysampling of experimental datasets for the training and test set selection.Molecular Diversity. 2002: 231-243*

.