

# FAST DATABASE SCREENING BY GENERATION OF VIRTUAL PHARMACOLOGICAL PROFILES

Massimo Mabilia, Elena Fioravanzo, Marco Daniele Parenti

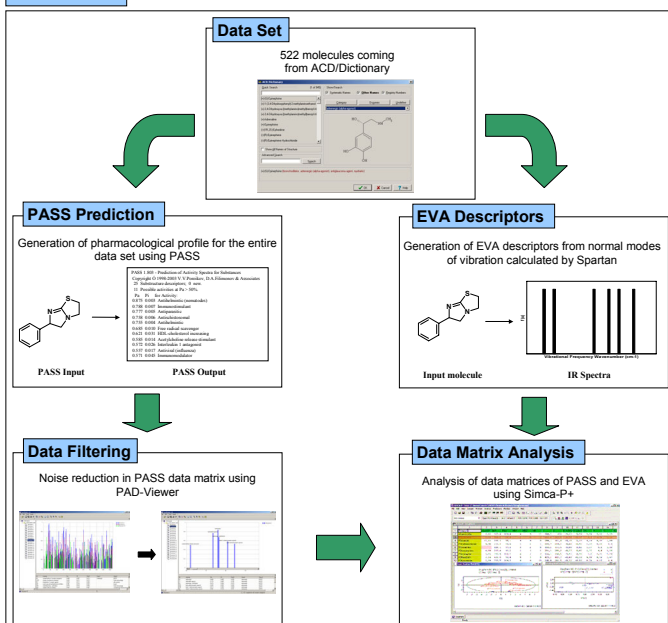
S.IN - Soluzioni Informatiche S.a.s., via Salvemini 9, I-36100 Vicenza, Italy

## INTRODUCTION

The computational analysis of structure-activity relationship (SAR) is widely applied in pharmaceutical research, especially in lead finding and optimization steps; nevertheless, the majority of the available SAR/QSAR methods are applicable only to a small number of biological activities and within the same chemical series. For these reasons, those techniques are often inappropriate to describe a complete pharmacological profile of potential drugs that have, in most cases, a wide range of biological effects (e.g. therapeutic effects, secondary effects, side effects, etc).

The software program PASS (Prediction of Activity Spectra for Substances)[1] predicts simultaneously several hundreds biological activity probability values, such as pharmacological effects, mechanisms of action, toxicity and metabolism reactions. In this communication, we demonstrate the capability of this software to rapidly screen a molecular database, and easily point out all molecules that possess a specific pharmacological profile at a high precision level.

## METHODS



## DATASET

The starting database is constituted by 522 known drugs, extracted among more than 20,000 molecules contained in ACD/Dictionary (ACD/Labs)[3], classified as active against five common receptorial subtypes:

- 178 adrenergic
- 111 cholinergic
- 69 dopaminergic
- 112 histaminergic
- 52 serotonergic

## DESCRIPTORS

Two different types of descriptors are used to classify the molecules:

(1) **PASS prediction:** is based on the analysis of structure-activity relationship of a training set including a great number of non-congeneric compounds with different biological activities (knowledge base), using a substructure descriptor called "Multilevel Neighborhoods of Atoms" (MNA)[2]. The output of PASS is a "descriptor", expressed as probability of presence (Pa) or absence (Pi) for each activity type in the knowledge base.

(2) **EVA descriptors:** a descriptor of molecular structure based on theoretically derived normal coordinate EigenValue (EVA)[4], related to IR spectra and then strictly dependent on nature and size of neighboring atoms. The normal modes of vibration are calculated using the software Spartan[5].

## GENERATION OF PHARMACOLOGICAL PROFILES AND DATA FILTERING

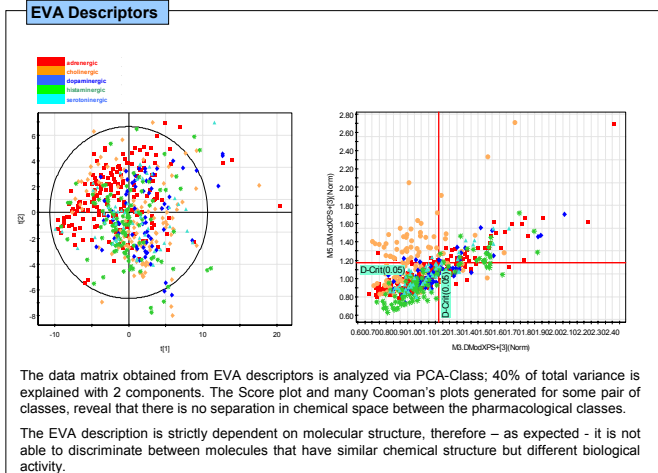
Each molecule in the dataset is submitted to a pharmacological profile generation with PASS, including in the output file all the biological activities predicted with  $Pa > Pi$ ; the PASS output file is then filtered using PAD Viewer[6], and only the data related to the five considered activity classes, with Pa values greater than 0.5 (or 0.7 in a second test), are exported to the final data matrix.

## DATA MATRIX ANALYSIS

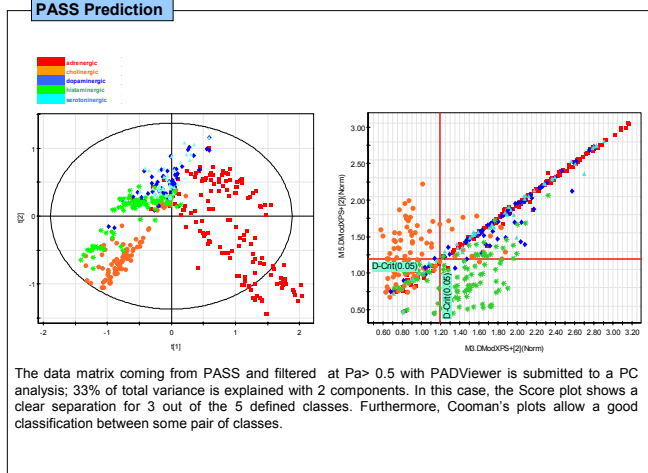
The two data matrices are analyzed via Principal Component Analysis (PCA) and PCA-Class using SIMCA-P+ software [7], in order to assess the presence of some separation in space between defined classes. Furthermore, the DModX (Class Distance) for each classes are plotted against each other in a scatter plot (Cooman's Plot).

## RESULTS

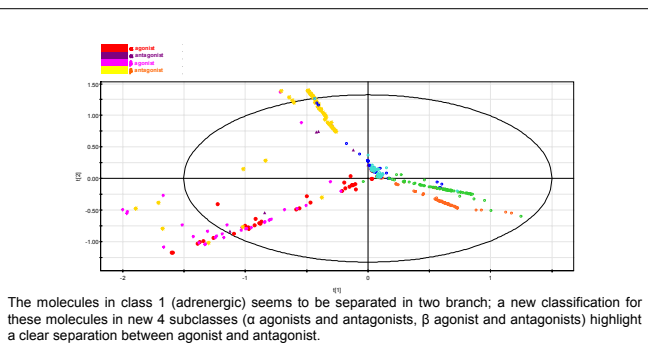
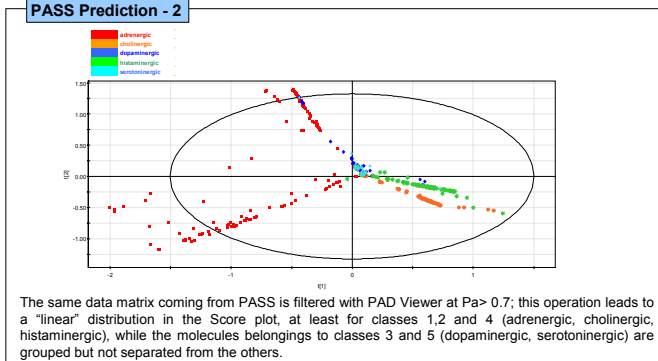
### EVA Descriptors



### PASS Prediction



### PASS Prediction - 2



## CONCLUSIONS

Even if the EVA descriptors are often successfully used to develop highly predictive and robust QSAR model, in this particular case the performance in classification was not pretty good, probably due to the highly dependence of this type of descriptors from the molecular structure. On the contrary, the pharmacological profile predicted by PASS has demonstrated its capability to discriminate between different classes of activity, even for molecules that possess very similar structure but different biological activity. These results are particularly interesting and encouraging, even more so if one considers that some compounds in the database are characterized by mixed activity profiles.

As revealed by the results of data filtering at  $Pa > 0.7$ , the PAD Viewer software plays an important role in these procedure, not only in analysis and visualization of the pharmacological profile, but also in reduction of noise and subsequently improvement of molecules classification.

## REFERENCE

- [1] V. Poroikov, D. Filimonov & Associates; [www.ibmh.msk.su/PASS/](http://www.ibmh.msk.su/PASS/)
- [2] Filimonov D.A. et al. J. Chem. Inf. Comput. Sci., 1999, 39, 666-670.
- [3] Advanced Chemistry Development, Inc.; [www.acdlabs.com](http://www.acdlabs.com)
- [4] A.M. Ferguson, T. Heritage, P. Jonathon, S.E. Pack, J. CAMD 11, 143-152, 1997.
- [5] "Pharmacological Activity Data Viewer"; developed by S.IN – Soluzioni Informatiche; [www.s-in.it](http://www.s-in.it)
- [6] Wavefunction Inc.; [www.wavefunction.com](http://www.wavefunction.com)
- [7] Umetrics AB, Sweden; [www.umetrics.com](http://www.umetrics.com)



S-IN Soluzioni Informatiche

di Ciampinelli Lucina & C. s.a.s.  
via G. Salvemini 9  
I-36100 Vicenza – Italy

tel. ++39 0444 240341  
fax ++39 0444 533954  
C. Fisc. e P.IVA 02397280245

info@s-in.it  
[www.s-in.it](http://www.s-in.it)