Table of Periodic Properties of Human Immunodeficiency Virus Inhibitors

Torrens F[1], Castellano G[2]

Motivation

Classification algorithms are proposed based on information entropy. The feasibility of mixing a given human immunodeficiency virus (HIV) inhibitor with dissimilar ones is studied. The 31 inhibitors are classified by their structural chemical properties.

Methods

Many classification algorithms are based on information entropy. An excessive number of results appear compatible with the data and suffer combinatorial explosion. However after the equipartition conjecture one has a selection criterion. According to this conjecture, the best configuration is that in which entropy production is most uniformly distributed. The structural elements of an inhibitor can be ranked according to their inhibitory activity.

Results

In didanosine (ddI) the base is a guanine derivative and the furan contains only one O heteroatom; ddI is selected as reference. In most inhibitors the furan contains one O heteroatom. The analysis is in agreement with principal component analysis and compares well with other classification taken as good based on docking, etc. Provisional conclusions follow. (1) Several criteria, selected to reduce the analysis to a manageable quantity of structures from the large set of HIV-1 inhibitors, refer to the structural parameters related with the base derivative, furan ring, number of N atoms, etc. Many algorithms for classification are based on information entropy. For sets of moderate size an excessive number of results appear compatible with data, and the number suffers a combinatorial explosion. However, after the equipartition conjecture, one has a selection criterion between different variants resulting from classification between hierarchical trees. According to the conjecture, the best configuration of a flowsheet is the one in which the entropy production is most uniformly distributed. The method avoids the problem of other methods of continuum variables because, for the four compounds with constant <11111> vector, the null standard deviation always causes a Pearson correlation coefficient of r = 1. The lower-level classification processes show lower entropy and may be more parsimonious. (2) Program MolClas is a simple, reliable, efficient

[1] Institut Universitari de Ciencia Molecular, Universitat de Valencia, Edifici dInstituts de Paterna, P. O. Box 22085, 46071 Valencia, Spain [2] Departamento de Ciencias Experimentales y Matematicas, Universidad Catolica de Valencia San Vicente Martir, 46003 Valencia, Spain

and fast procedure for molecular classification, based on the equipartition conjecture of entropy production. It has been written not only to analyze the equipartition conjecture of entropy production, but also to explore the world of molecular classification. (3) The good comparison of our classification results, with other taken as good, confirm the adequacy of the property vector selected for the molecular structures of the HIV-1 inhibitors. Information entropy and principal component analyses permit classifying the inhibitors and agree. The inhibitors are grouped into different classes. In general the classical classes are recognized, viz. non-nucleoside reverse transcriptase inhibitors, nucleoside reverse transcriptase inhibitors, nucleotide reverse transcriptase inhibitors, and protease inhibitors. The final classification is shown more precise and with lower bias. The classification model calculates in each case the contribution of signal and noise. (4) Classification algorithms are proposed based on information entropy. The feasibility of mixing a given HIV-1 inhibitor with dissimilar ones in a complex drug is studied. The 31 inhibitors (including eight nucleoside/nucleotide reverse transcriptase inhibitors) are classified, by structural chemical properties. The analysis includes inhibitors fitting the general scheme: (base derivative)-(furan ring). The base portion is often a guanine or cytosine derivative, and the furan normally contains one O heteroatom. The structural elements of an inhibitor can be ranked according to their inhibitory activity in the order: base > furan. In ddI, the base is a guanine derivative and the furan contains only one O heteroatom (N4O3S0P0X0, X = F, Cl, Br); its associated vector is <11111>. The ddI was selected as a reference. In most inhibitors the furan contains only one O heteroatom (ddI, ddC, d4T, novel proposed ligand, N3-4O3S0P0X0), while in 3TC the furan includes one O and one S heteroatoms (N3O3S1P0X0). The analysis is in agreement with principal component analysis. It compares well with other classification taken as good based on docking, density functional, molecular dynamics, the Rule of Five, and absorption, distribution, metabolism, excretion and toxicity. (5) The periodic law has not the rank of the laws of physics: (a) the properties of the human immunodeficiency virus type 1 inhibitors are not repeated; perhaps their chemical character; (b) the order relationships are repeated with exceptions. The analysis forces the statement: The relationships that any inhibitor p has with its neighbour p + 1 are approximately repeated for each period. Periodicity is not general; however, if a natural order of the inhibitors is accepted the law must be phenomenological.

Contact e-mail
Francisco.Torrens@uv.es