A STUDY OF HIV-1 PROTEASE-INHIBITOR INTERMOLECULAR INTERACTIONS BY USING QUANTITATIVE MOLECULAR GRAPHICS AND *A PRIORI* QSAR

Rudolf Kiralj and Márcia M. C. Ferreira, Laboratório de Quimiometrica Teórica e Aplicada, Instituto de Química, Universidade Estadual de Campinas (UNICAMP), Campinas, SP, 13083-970, Brazil.

E-mails: *marcia@iqm.unicamp.br*, *rudolf@iqm.unicamp.br*, ***URL:*** *http://lqta.iqm.unicamp.br*

1

ABSTRACT

Although peptidic HIV-1 protease inhibitors were studied by sophisticated[1] and simple[2] QSARs, and crystal structures of some protease-inhibitor complexes are known,[3] some aspects of protease-inhibitor intermolecular interactions were not considered. Molecular graphics is usually utilized as a visualization aid. In this work we employ molecular graphics on inhibitor L-700,417 as a quantitative tool. Molecular dimensions (like geometrical parameters), parameters of electron density and frontier orbital isosurfaces were measured/calculated from high quality pictures by an empirical method. The relationships between enzyme-inhibitor geometry and inhibitor electron density distribution, beyond van der Waals radii, were established. Another empirical method and Partial Least Squares (PLS) with *a priori* molecular descriptors[2] were used to predict activities of L-700,417 derivatives. These activities were well related to enzyme-inhibitor interaction energy and inhibitor molecular properties. The group additivity of activities was confirmed. *A priori* QSAR included modeling of L-700,417 derivatives at the active site of frozen enzyme. The number and size of substituents, especially number of hydroxyl groups, and the presence of various hydrophobic hooks attached to a substituent significantly changes the inhibitor hydrophobicity, its biological activity and the enzyme-inhibitor interaction energy.

1 Pérez, C.; Pastor, M.; Ortiz, A. R.; Gago, F. *J. Med. Chem.* **41** (1998) 836.

2 Kiralj, R.; Ferreira, M. M. C. *J. Mol. Graph. Mod.*, submitted for publication.

3 *Database for Anti-HIV Compounds*. National Institute for Allergy and Infectious Diseases, Bethesda, Maryland, US. http://www.niaid.nih.gov/daids/dtpdb/

2

Quantitative molecular graphics – what is it?

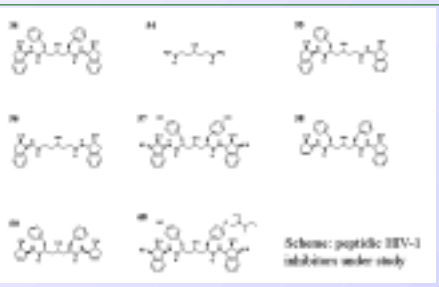
-use of molecular graphics at quantitative level employing high-quality pictures of molecules in various views (orthographic projections primarily) as a source of quantitative information
-examples of measurements on molecular portrayals: 1D- molecular dimensions (bond lengths, maximum and minimum intramolecular distances etc.), 2D-molecular dimensions (projected surface areas, surface areas of planar fragments etc.), 3D-molecular dimensions (van der Waals volume, surface accessible volume etc.), electron density relationships etc.
-methods of measurements: empirical (hand-made) or using specialized software
-advantages: easy way of obtaining required information instead of using expensive, sophisticated software; use of pictures that cannot be easily reproduced

A priori QSAR – why called *a priori*?

-it includes *a priori* molecular descriptors: those "known before" any sophisticated, computer-assisted calculation
-generated by hand- or pocket calculator count/calculation using 1D or 2D chemical formula, chemical knowledge, and well-known facts and a few literature numerical data (not databases); 3D-coordinates are omitted

Peptidic HIV-1 protease inhibitors (Scheme) were studied by using quantitative molecular graphics (**34**) and *a priori* QSAR (**54-60**).[2]

3



4

METHODOLOGY

A) Empirical measurements of molecular dimensions of inhibitor 34 (Figure 1)

-linear dimensions: $L_i = N_i f_i$ where: L_i – the true length, N_i – number of length units (millimeters), f_i – scale factor based on an internal standard (molecular fragment of known dimensions)
-projected surface areas S of volumes V: van der Waals volume, Conolly surface accessible volume, volumes enclosed into electron density and HOMO-LUMO isosurfaces: $S_i = M_i f_i^2$ where: M_i – number of square units (square millimeters)
-estimation of V (for electron density, HOMO-LUMO) from $V\text{-}S^{1/2}$ relationship (based on known relationships for vdW and Conolly volumes, in projections where this relationship is the most accurate)

B) Empirical study of inhibitor 34 – environment (HIV-1 protease&water) interactions

-projected surface areas S, height H and length L of electron density isosurfaces ρ = 0.001, 0.01, 0.05, 0.1, 0.2, 0.5 Å^{-3} was measured in bottom projection; cut-off distance D was calculated from H and L (Figure 2)
-logp-D linear relationship was established and V (vdW) was predicted
-average environment electron density $\tau = N_{ve} / (V\text{-}553.7\ Å^3)$ was calculated, where: Nve, V – environment valence electrons and volume inside cut-off D
-the inhibitor-environment overlap function $F = (\rho\tau)^{1/2}$ (V-553.7 Å^3) was calculated

C) Empirical estimation and PLS prediction of biological activities for 54-60

-empirical estimation of activities for **54-60** was performed on the basis of group additivity properties of activities for **1-48**[2], that is an intrinsic property of both the HIV-1 protease active site and of its inhibitors
-PLS prediction of biological activities for **54-60**: *a priori* molecular descriptors X_1-X_{14} were generated (Table 1) for **54-60** and the PLS *a priori* model I (from the training set **1-32**, previous work[2]) was applied

5

-both sets of biological activities (the empirical and PLS) for **34**, **54-60** were related to molecular vdw volume V and surface area S, number of potential hydrogen bonds N_{HB}, number of environment atoms $N_{5.5}$ and inhibitor-environment interaction energies at 5.5 and 10.0 Å cut-off distance; the energy was based on empirical atom-atom potential of Gavezzotti (*Acc. Chem. Res.* **27** (1994) 309).

D) A priori molecular modeling of HIV-1 protease - inhibitor 60 complex

-a derivative of **57** with *ortho*-positioned $-O\text{-}(CH_2)_2\text{-}N(CH_2CH_3)\text{-}CH(CH_3)CH_2CH_3$ flexible hook on a phenyl group, **60**, was modeled at the active site hole of the frozen protease: among its molecular properties, special attention was paid to number of OH/hydrophobic groups, its size and shape, conformational flexibility, hydrophobicity and interaction mode with the protease; the complex was also optimized using MMFF94 force field.

6

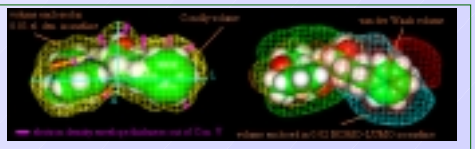


Figure 1. Types of volumes and molecular dimensions measured for **34** by means of quantitative molecular graphics.

Figure 2. Electron density chicken wire cages for 0.5, 0.2, 0.1, 0.05, 0.01, 0.001 isosurface of **34** (left to right).

7

RESULTS

A) Empirical measurements of molecular dimensions of 34 (Figure 1)

-Thickness of 0.01 Å^{-3} electron density envelope out of Conolly volume distinguishes different types of hydrogens (Fig. 1): in polar groups (1.16 Å), in hydrophobic groups (1.02 Å); and also anisotropic distribution around benzene ring: perpendicular to the ring (0.66 Å), in between CH bonds (0.97 Å), inclined 45 deg. to the ring (0.87 Å)
-Volumes are: enclosed into 0.06/0.01 Å^{-3} electron density isosurfaces – 565.6/854.2 Å^3, enclosed into 0.02 Å^{-3} HOMO-LUMO – 847.3 Å^3, van der Waals for inhibitor of maximum size (**1-32** overlapped, see ref. 2) – 725.9 Å^3

B) Empirical study of inhibitor 34 – environment (HIV-1 protease&water) interactions

-F-D curves (Figure 3) reveal: a) 5.5 Å cut-off distance is a reasonable molecular boundary; b) intermolecular interactions occur up to 7.0 Å cut-off, what allows mutual penetration of the components (Figure 4); d) various functional groups of **34** exhibit different behavior in intermolecular interactions: range, geometry, intensity and frequency of the interactions are parameters characteristic for each type of the functional groups

C) Empirical estimation and PLS prediction of biological activities for 54-60

All the data are in Tables 1-3.
-Empirical biological activity y_{emp} is highly correlated to all the variables in Table 2 (corr. coef.>0.7), what explains why the PLS model based on *a priori* variables[2] works well: the bigger the substituent in the inhibitor (or: the bigger the inhibitor in question, meaning increase of V, S), the more intermolecular interactions and environment atoms ($N_{5.5}$, N_{HB} increase) are around the inhibitor, what obviously increases inhibitor-environment intermolecular interaction energy at 5.5 Å, 10.0 Å and at any higher cut-off distance D. This way, size of substituents on the basic peptidic inhibitor structure, expressed by means of effective number of substitutents X_9 or some other extensive descriptor, is directly related to biological activity.

8

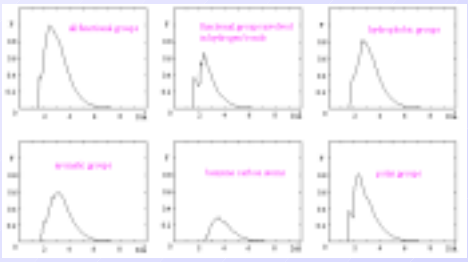


Figure 3. F-D curves for various functional groups of **34**.

9

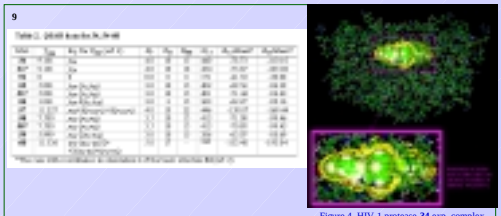


Figure 4. HIV-1 protease-**34** exp. complex.

D) *A priori* molecular modeling of HIV-1 protease - inhibitor 60 complex

Hook inhibitor **60**, a derivative of L-700,417 (**34**), is a highly hydrophobic molecule (ClogP=4.92), with an amino-alkyl side chain and more –OH groups than the parent compound; this compound could be an effective peptidic HIV-1 protease inhibitor. *A priori* modeling shows **60** to be in a position not much different from that one from MMFF94 optimization (Figure 5).

10

11

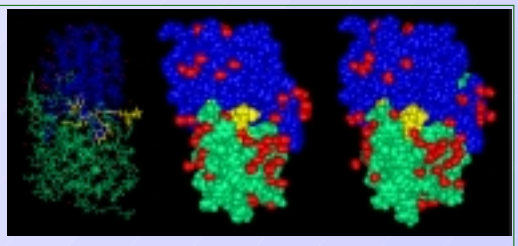


Figure 5. Inhibitor **60** (yellow) inside HIV-1 protease consisting of A (blue) and B (green) chains. Waters are in red. The hook is visible outside the protein. Unoptimized (left and middle, protein from experimental structure) complex is compared to MMFF94 optimized structure (right) in which the hook seems to fit fairly well to the enzyme.

12

CONCLUSION

Quantitative molecular graphics analysis:

-showed to be an effective method for estimation of molecular dimensions and quantification of spatial electron density distribution and enzyme-substrate intermolecular interactions
-revealed that in 1.6-5.5 Å cut-off distance (beyond van der Waals radii) occur the most important inhibitor-environment intermolecular interactions, while interactions exist up to 7.0 Å; this way, 5.5 Å can be defined as a physical boundary of molecules; various functional groups of inhibitor participate differently to intermolecular interactions

***A priori*&PLS model, empirical estimation of activities and *a priori* molecular modeling:**

-introducing both polar (primarily –OH) and hydrophobic groups (rings or chains) into a peptidic inhibitor increases biological activity (the inhibitor should have at least four substituents, by other words it should occupy four protease pockets)
-attaching flexible hook(s) on inhibitor substituent(s) would increase the activity; the hook(s) must be hydrophobic, and so would stick on the hydrophobic HIV-1 protease surface at that part where is the entrance to the active site hole

ACKNOWLEDGEMENT

The authors acknowledge FAPESP for the financial support