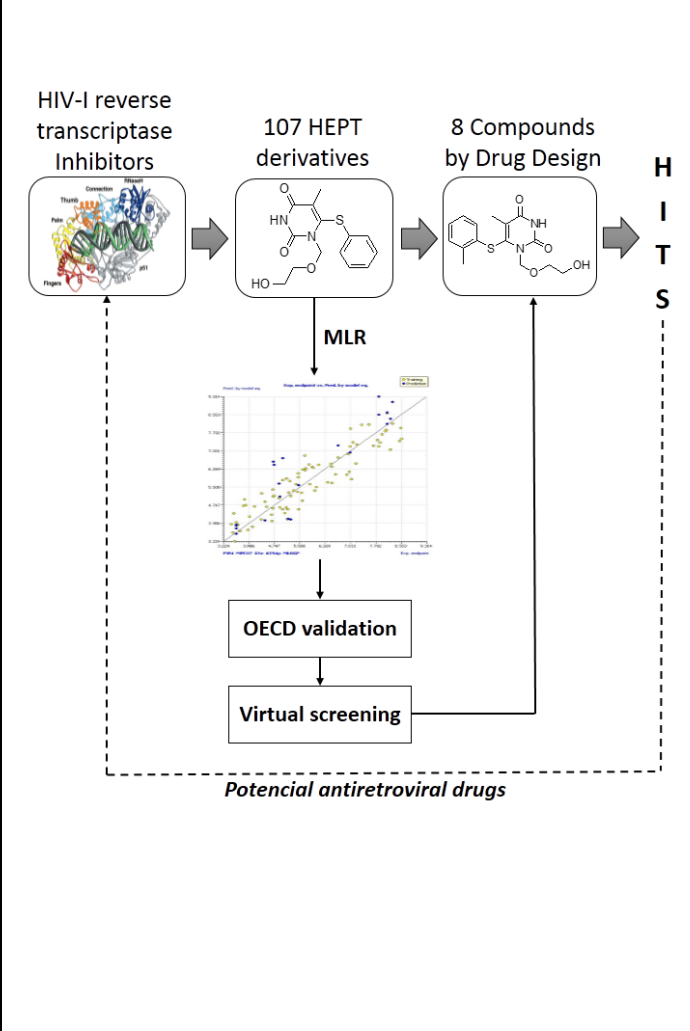


Chemometric study of antiretroviral activities of HEPT derivatives using Multiple Linear Regression

Yudith Cañizares-Carmenate^{a*}, Juan Alberto Castillo-Garit^b, Yoanny Hernández Machado^a

^a Unit of Computer-Aided Molecular “Biosilico” Discovery and Bioinformatic Research (CAMD-BIR Unit), Facultad de Química-Farmacia, Universidad Central “Marta Abreu” de Las Villas, Santa Clara 54830, Villa Clara, Cuba yudithc@uclv.edu.cu

^b Unidad de Toxicología Experimental, Universidad de Ciencias Médicas de Villa Clara, Santa Clara 50200, Villa Clara, Cuba

Graphical Abstract	Abstract
 <p>The graphical abstract illustrates a workflow for identifying potential antiretroviral drugs. It starts with HIV-I reverse transcriptase inhibitors, leading to 107 HEPT derivatives. These derivatives are analyzed using Multiple Linear Regression (MLR), which is validated using OECD validation. The results are then used for virtual screening to identify 8 compounds by drug design. The final output is labeled as 'Potencial antiretroviral drugs'.</p>	<p>Abstract</p> <p>Reverse transcriptase is an enzyme whose function is to synthesize double-stranded DNA using the single-stranded RNA of HIV as the template. Antiretroviral drugs are effective against HIV infection because they reduce virus replication and slow the progression of the disease, even though it does not mean a definitive cure. In addition, existing drugs present a large number of adverse reactions, so the search for new antiretroviral drugs continues to be a major challenge today. In this study, a series of 107 experimentally tested 1-[(2-hydroxyethoxy)methyl]-6-(phenylthio)thymine (HEPT) derivatives were used against HIV-I reverse transcriptase to model antiretroviral activity using molecular descriptors of DRAGON. Subsequently, a mathematical model of Multiple Linear Regression, implemented in the QSARINS software, was developed by common least squares and genetic algorithm for the variables</p>

selection, with a good performance of its parameters. The model is extensively validated in accordance with OECD regulations so that its robustness ($R^2=0.84$), stability ($Q^2_{LOO}=0.81$; $Q^2_{LMO}=0.80$), low correlation of the descriptors ($K_{XX}=0.58$) and good predictive power ($R^2_{EXT}=0.85$). In addition, it was found that the fit of the model is not a product of a random correlation ($R^2_{Y-scr}=0.06$; $Q^2_{Y-scr}=-0.09$), so it was used to predict the antiretroviral activity of eight compounds obtained by rational drug design.

References

1. Weiss RA. How does HIV cause AIDS? *Science*. 1993;260(5112):1273-9.
2. Cunningham AL, Donaghy H, Harman AN, Kim M, Turville SG. Manipulation of dendritic cell function by viruses. *Current opinion in microbiology*. 2010;13(4):524-9.
3. Tanaka H, Baba M, Hayakawa H, Sakamaki T, Miyasaka T, Ubasawa M, et al. A new class of HIV-1 specific 6-substituted acyclouridine derivatives: synthesis and anti-HIV activity of 5- or 6-substituted analogues of 1-[(2-hydroxyethoxy)-methyl]-6-(phenylthio)thymine (HEPT). *J Med Chem*. 1991; 34:349-57.
4. Toropova AP, Toropov AA, Veselinovi JB, Miljkovi FN, Veselinovi AM. QSAR models for HEPT derivatives as NNRTI inhibitors based on Monte Carlo method. *European Journal of Medicinal Chemistry*. 2014;77:298-305.
5. Montessori V, Press N, Harris M, Akagi L, Montaner JSG. Adverse effects of antiretroviral therapy for HIV infection. *CMAJ*. 2004;170(2):229-38.
6. Todeschini R, Consonni V. *Handbook of Molecular Descriptors*. Mannhold R KH, Timmerman H editor. Federal Republic of Germany 2000.
7. Estrada E. *Chem Inf Comput Sci*. 1996;36:844.
8. Roy K. On some aspects of validation of predictive quantitative structure-activity relationship models. *Expert Opin Drug Discov*. 2007;2 (12):1567-77.
9. Switzer WM, Garcia AD, Yang C, Wright A, Kalish ML, Folks TM. Coinfection with HIV-1 and simian foamy virus in West Central Africans. *J Infect Dis*. 2008;197:1389-93.
10. Plantier JC, Leoz M, Dickerson JE, De Oliveira F, Cordonnier F, Leme V. A new human immunodeficiency virus derived from gorillas. *Nat Med*. 2009;15:871-2.
11. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a Covalently Trapped Catalytic Complex of HIV-1 Reverse Transcriptase: Implications for Drug Resistance. *Science* 1998;282:1669-75.
12. Dalglish AG, Beverley PCL, Clapham PR, Crawford DH, Greaves MF, Weiss RA. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature*. 1984;312:763-8.
13. Carter CC, Onafuwa-Nuga A, McNamara LA, Riddell J, Bixby D, Savona MR. HIV-1 infects multipotent progenitor cells causing cell death and establishing latent cellular reservoirs. *Nat Med*. 2010;16:446-51.
14. Kieffer TL, Finucane MM, Nettles RE, Quinn TC, Broman KW, Ray SC. Genotypic analysis of HIV-1 drug resistance at the limit of detection: virus production without evolution in treated adults with undetectable HIV loads. *J Infect Dis*. 2004;189:1452-65.
15. Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, et al. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. *Proc Natl Acad Sci USA*. 1995;92(4):1222-6.

16. Sarafianos SG, Marchand B, Das K, Himmel DM, Parniak MA, Hughes SH, et al. *Structure and Function of HIV-1 Reverse Transcriptase: Molecular Mechanisms of Polymerization and Inhibition. J Mol Biol* 2009; 385:693-713.
17. Huang H, Chopra R, Verdine GL, Harrison SC. *Structure of a Covalently Trapped Catalytic Complex of HIV-1 Reverse Transcriptase: Implications for Drug Resistance. Science.* 1998;282:1669-75.
18. Duda-Seiman C, Duda-Seimana D, Putz MV, Ciubotariu D. *QSAR Modelling of Anti-HIV Activity with HEPT Derivatives. Digest Journal of Nanomaterials and Biostructures.* 2007;2(2):207 - 19.
19. Todeschini R, Consonni V. *Molecular Descriptors for Chemoinformatics: WILEY-VCH; 2009.*
20. Gramatica P, Chirico N, Papa E, Cassani S, Kovarich S, Sangion A. *QSAR Research Unit in Environmental Chemistry and Ecotoxicology. Includes QSARINS-Chem. 2015.*