

QSAR MODELING OF ANTI-HIV-1 ACTIVITY OF HEPT DERIVATIVES. OPTIMIZATION OF CORRELATION WEIGHTS OF MORGAN EXTENDED CONNECTIVITY IN GRAPH OF ATOMIC ORBITALS

Castro, E.A.¹; Toropov, A.A.²; Toropova, A.P.²; Mukhamedzhanove, D.V.²

¹INIFTA, Theoretical Chemistry Division, Chemistry Department, Faculty of Exact Sciences, La Plata National University, Suc.4, C.C. 16, La Plata, B1906ZAA, Argentina.
Fax: +54-221-425-4642, Email: jubert@arnet.com.ar

²Uzbekistan Academy of Sciences, Geology & Geophysics Institute, 700041, Khodzhibaev, 49 Street, Tashkent, Uzbekistan

Received August 12th, 2005. In final form August 29th, 2005

Dedicated to the memory of the late Prof. Hans J. Schumacher on the occasion of his 100th birthday

Abstract

Quantitative Structure – Activity Relationships (QSAR) for Anti-HIV-1 activity of 1-[2-hydroxyethoxy-methyl]-6-(phenylthio) thymine] (HEPT) derivatives have been obtained by optimization of correlation weights of local invariants in Labelled Hydrogen – Filled Graphs (LHFGs) and Graph of Atomic Orbitals (GAOs). The best model is based on correlation weights of Morgan extended connectivity of first order in the GAO. Statistical characteristics of this model are the following: $n = 59$, $R^2 = 0.91$, $s = 0.41$, $F = 577$ (Training Set); $n = 20$, $R^2 = 0.91$, $s = 0.42$, $F = 183$ (Test Set).

Resumen

Se han determinado las Relaciones Cuantitativas Estructura-Actividad (QSAR) para la actividad Anti-HIV-1 de los derivados del 1-[2-hidrometoxi-metil]-6-(feniltio)timina (HEPT) por medio de la optimización de los pesos de correlación de los invariantes locales de los Grafos Completos de Hidrógeno Señalados (LHFGs) y de los Grafos de los Orbitales Atómicos (GAOs). El mejor de los modelos obtenidos basado en los pesos de correlación de la conectividad extendida de Morgan del primer orden es el GAO. Las características estadísticas de este modelo son las siguientes: $n = 59$, $R^2 = 0,91$, $s = 0,41$, $F = 577$ (Conjunto de Entrenamiento); $n = 20$, $R^2 = 0,91$, $s = 0,42$, $F = 183$ (Conjunto de Prueba).

Introduction

The Acquired Immuno Deficiency Syndrome (AIDS) epidemic has taken more than 3 million lives in 2004, and an estimated 5 million people have acquired the Human Immunodeficiency Virus (HIV) in 2004, bringing to 42 million people the number of people globally living with the

virus. HIV/AIDS has become the most devastating pandemic in the recorded contemporary history. It has killed 40 million people in the last 20 years and the World Health Organization has estimated that there would be up to 100 million new infections in the next 10 years [1]. Because human type immunodeficiency virus (HIV-1) is the causative agent of AIDS, extensive research clinical works are currently going on to try to block its replication [2]. Quantitative structure–property/activity relationships (QSPR/QSAR) may be used as a tool for prediction anti-HIV-1 activity. From a literature survey it was found that faced with the need to develop effective anti-HIV drugs, QSAR studies have been used on potential anti-HIV drug candidates acting on different targets [3-13].

There are number of ways to represent molecular structures in order to employ the QSPR/QSAR analyses, such as the hydrogen – suppressed graph (i.e., hydrogen atoms are ignored), hydrogen-filled graph, 3D molecular model, and quantum chemical models. In the present study two versions of elucidation of molecular structure have been used. These are the Labelled Hydrogen–Filled Graph (LHFG) and Graph of Atomic Orbitals (GAO). GAO is an attempt to take into account the structure of atoms in QSPR/QSAR analyses. Recently, we have obtained satisfactory predictions of physical chemistry properties and biological activities within the realm of QSPR/QSAR theory via molecular graphs representation by means of optimization of correlation weights of local graph invariants (OCWLI) [14-23].

The present study is aimed to estimate the capability of the OCWLI to predict numerical values of the Anti-HIV-1 activity, in $\log(1/LC_{50})$ taken from Ref. [2] (see Fig. 1).

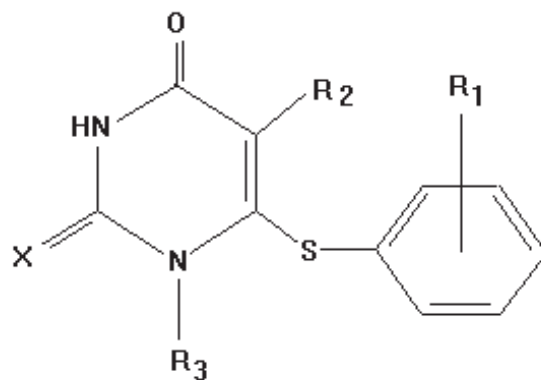


Figure 1. General structure of HEPT derivatives studied.

Method

The descriptors used in the present study are defined as

$${}^0X_{CW}(G^xEC) = \sum_{\text{all vertices}} CW(V_k) \cdot CW({}^xEC_k) \quad (1)$$

where G is the LHFG or GAO, V_k stands for a chemical element in case of the LHFG or Atomic Orbital in case of the GAO. The xEC_k is the Morgan extended connectivity of x -th order calculated with the recursion formula [14-23]

$${}^x\text{EC}_k = \sum_{(k,l)} ({}^{x-1}\text{EC}_l) \quad (2)$$

The correlation coefficient between the ${}^0\text{X}_{\text{CW}}(\text{G}, {}^x\text{EC})$ and $\log(1/\text{LC}_{50})$ is a function of the correlation weights (CWs). Under such circumstances the following steps can be taken [14-23]:

1. Using the Monte Carlo method, the values of the Cws giving rise to the largest possible values of the correlation coefficient can be calculated;
2. Using the least squares method, the following equation may be obtained with compounds of the training set

$$\log(1/\text{LC}_{50}) = C_0 + C_1 {}^0\text{X}_{\text{CW}}(\text{G}, {}^x\text{EC}) \quad (3)$$

3. Predictive capability of Eq.(3) may be validated for compounds belonging to a test set.

The conversion of the LHFG into the corresponding GAO may be carried out by the following scheme:

- i) Each vertex of the LHFG is replaced by a group of Atomic Orbitals (AOs). Such groups of AOs on all atoms under consideration are listed in Table 1.
- ii) The element a_{ij} of adjacency matrix of the GAO is defined as 1 if the i-th and j-th vertices of the GAO fall in groups of different atoms from the LHFG and these atoms have joint edge in the LHFG; otherwise, the $a_{ij} = 0$.

The LHFG and GAO on compounds No 49 (training set) are demonstrated in Fig 2 and Fig.3, respectively.

Table 1. AO groups on atoms under consideration.

Chemical element	AO group of the chemical element
H	1s ¹
C	1s ² , 2s ² , 2p ²
N	1s ² , 2s ² , 2p ³
O	1s ² , 2s ² , 2p ⁴
F	1s ² , 2s ² , 2p ⁵
S	1s ² , 2s ² , 2p ⁶ , 3s ² , 3p ⁴
Cl	1s ² , 2s ² , 2p ⁶ , 3s ² , 3p ⁵
Br	1s ² , 2s ² , 2p ⁶ , 3s ² , 3p ⁶ , 3d ¹⁰ , 4s ² , 4p ⁵
I	1s ² , 2s ² , 2p ⁶ , 3s ² , 3p ⁶ , 3d ¹⁰ , 4s ² , 4p ⁶ , 4d ¹⁰ , 5s ² , 5p ⁵

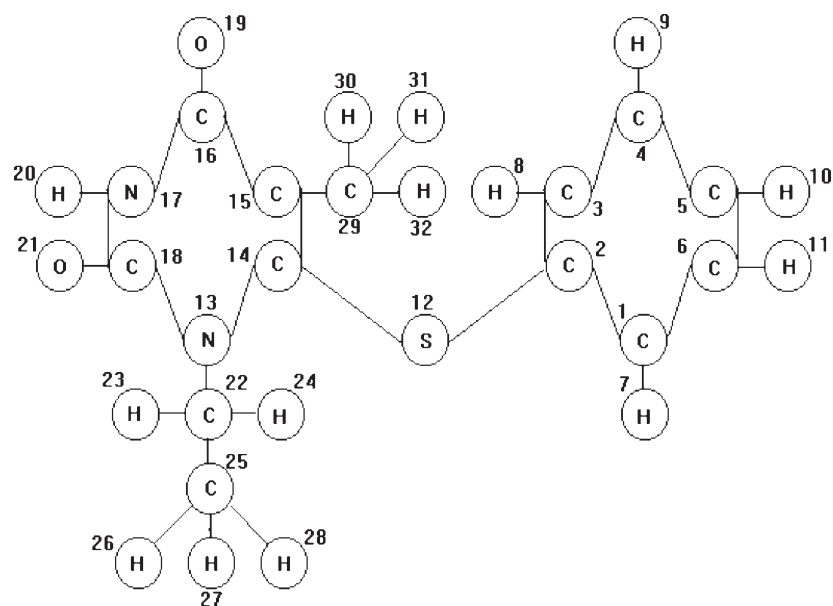


Figure 2. The LHFG of compound No 49 (training set).

Results and Discussion

Table 2 lists statistical characteristics of the OCWLI models of Anti-HIV-1 activity. One can see from the Table 2 that best OCWLI model of the activity under consideration takes place for the ${}^0\mathbf{X}(\mathbf{GAO}, {}^1\mathbf{EC})$. Table 3 lists results of three probes OCWLI based on the ${}^0\mathbf{X}(\mathbf{GAO}, {}^1\mathbf{EC})$. Numerical values of the CWs are displayed in Table 4. Calculation of the ${}^0\mathbf{X}(\mathbf{GAO}, {}^1\mathbf{EC})$ for the 49th compound from the training set is illustrated in Table 5. The results of this calculation are based on CWs obtained for the first OCWLI probe from the Table 4.

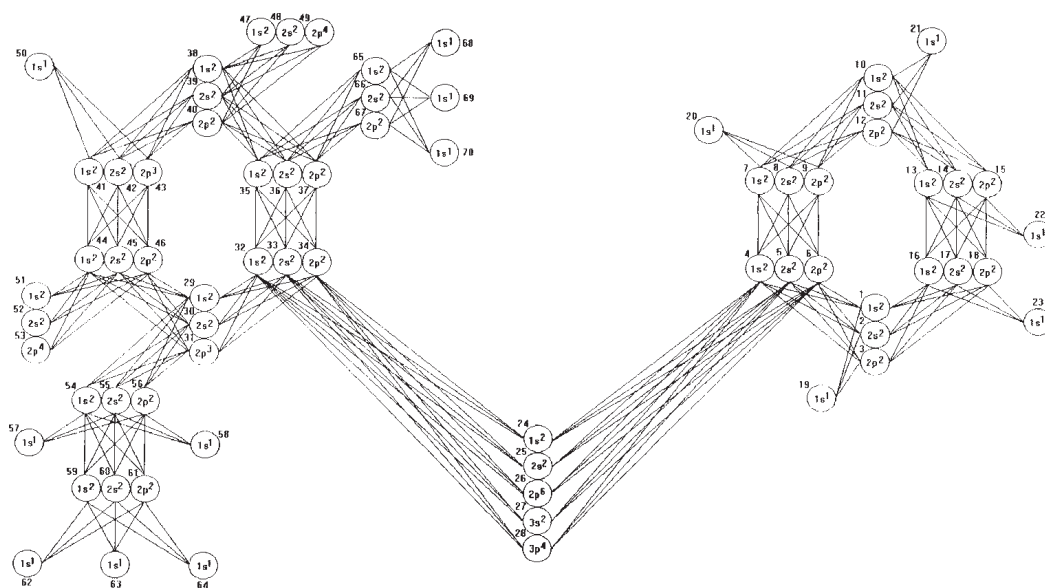


Figure 3. The GAO for Compound No 49 (training set)

Table 2. Number of OCWLI parameters and statistical characteristics of QSAR on different versions of flexible descriptor calculated by Eq.(1).

Descriptor	Number of local invariants for the OCWLI	R ² for the training set n = 59	R ² for the test set n = 20
⁰ X(LHFG, ⁰ EC)	13	0.55	0.37
⁰ X(LHFG, ¹ EC)	22	0.77	0.53
⁰ X(LHFG, ² EC)	40	0.91	0.68
⁰ X(GAO, ⁰ EC)	31	0.68	0.60
⁰ X(GAO, ¹ EC)	48	0.91	0.91

Observed values of the Anti-HIV-1 activity from Ref. [2], numerical values of the ⁰X(GAO,¹EC), and structure compounds of general formula shown in Fig.1 are listed in Table 6. OCWLI model of log(1/LC₅₀) calculated as

$$\log(1/LC_{50}) = 4.977 + 0.06745 \text{ } ^0X_{cw}(G^sEC) \quad (4)$$

are also shown in Table 6. Statistical features of Eq.(4) can be seen from the Table 3 (first OCWLI probe).

According to Ref. [2], the 79 HEPT derivatives under consideration have been studied by means of neural networks technique and via a multiple linear regression. Standard errors in this previous research examined in Ref. [2] fall within an interval going from 0.607 till 0.185. Thus statistics data associated with Eq.(4), i.e., 0.41 – 0.42 may be considered as reasonably satisfactory.

Table 3. Statistical characteristics of three OCWLI probes to model of Anti-HIV-1 activity based on ⁰X(GAO,¹EC)

	Training set n = 59					Test set n = 20			Complete set n = 79		
	C ₁	C ₀	R ²	s	F	R ²	s	F	R ²	s	F
1	0.067	4.977	0.9107	0.406	581	0.9150	0.419	194	0.9103	0.407	781
2	0.059	4.666	0.9089	0.410	569	0.9125	0.418	188	0.9090	0.409	769
3	0.072	4.862	0.9104	0.407	579	0.9140	0.418	191	0.9102	0.407	780

Table 4. Correlation weights for calculating of the ${}^0X(\text{GAO}, {}^1\text{EC})$ obtained in three OCWLI probe

GAO invariant	CWs of OCWLI probe 1	CWs of OCWLI probe 2	CWs of OCWLI probe 3
AO			
1s ¹	0.027	0.042	0.028
1s ²	-1.635	-2.094	-2.054
2s ²	-1.502	-1.015	-1.008
2p ²	8.973	9.669	9.252
2p ³	-0.599	-0.069	-0.568
2p ⁴	-2.768	-3.793	-3.159
2p ⁵	1.495	1.085	1.411
2p ⁶	0.048	-1.399	0.122
3s ²	0.171	-0.257	-0.486
3p ⁴	6.597	5.999	6.492
3p ⁵	3.890	6.087	4.569
3p ⁶	-3.351	-2.813	-5.563
3d ¹⁰	-2.014	-0.890	-0.659
4s ²	-1.866	-1.696	-0.901
4p ⁵	-2.123	-3.935	-2.537
4p ⁶	42.870	34.752	18.159
4d ¹⁰	35.721	40.509	49.326
5s ²	32.882	50.686	26.336
5p ⁵	30.384	25.369	36.256
¹EC			
0012	685.447	469.583	623.920
0015	335.458	238.560	307.686
0018	-0.957	-1.407	-0.974
0021	9.852	7.898	11.384
0024	-9.411	-8.305	-8.873
0027	0.834	0.815	0.742
0030	5.227	4.891	4.621
0033	2.874	2.845	2.602
0036	1.491	1.482	1.358
0039	-0.246	-0.242	-0.212
0042	-2.394	-2.276	-2.128
0045	0.714	0.670	0.622

0048	-1.900	-1.822	-1.646
0051	-0.125	-0.118	-0.100
0054	-0.525	-0.475	-0.453
0057	0.639	0.611	0.533
0060	-1.562	-1.579	-1.384
0063	2.854	2.837	2.470
0066	-0.845	-0.779	-0.761
0069	1.174	1.072	1.005
0072	-3.805	-3.521	-3.288
0075	3.632	3.493	3.208
0078	-0.491	-0.526	-0.456
0081	-2.427	-2.287	-2.114
0084	2.181	2.227	1.919
0087	-2.177	-2.251	-2.513
0090	7.938	7.930	7.095
0093	6.301	5.407	4.439
0108	4.027	4.516	4.087

Table 5. Calculation of ${}^0X_{cw}(GAO,{}^1EC)$ for compound No 49 (training set) with correlation weights of first OCWLI probe. The ${}^0X_{cw}(GAO,{}^1EC)$ is equal to 17.285

Atom	No(Atom)	1EC in LHFG	AO	No (AO)	1EC in GAO	CW(AO)	CW(1EC)
C	1	7	1s ²	1	57	-1.635	0.639
			2s ²	2	57	-1.502	0.639
			2p ²	3	57	8.973	0.639
C	2	8	1s ²	4	72	-1.635	-3.805
			2s ²	5	72	-1.502	-3.805
			2p ²	6	72	8.973	-3.805
C	3	7	1s ²	7	57	-1.635	0.639
			2s ²	8	57	-1.502	0.639
			2p ²	9	57	8.973	0.639
C	4	7	1s ²	10	45	-1.635	0.714
			2s ²	11	45	-1.502	0.714
			2p ²	12	45	8.973	0.714
C	5	7	1s ²	13	45	-1.635	0.714
			2s ²	14	45	-1.502	0.714
			2p ²	15	45	8.973	0.714
C	6	7	1s ²	16	45	-1.635	0.714

			$2s^2$	17	45	-1.502	0.714
			$2p^2$	18	45	8.973	0.714
H	7	3	$1s^1$	19	21	0.027	9.852
H	8	3	$1s^1$	20	21	0.027	9.852
H	9	3	$1s^1$	21	21	0.027	9.852
H	10	3	$1s^1$	22	21	0.027	9.852
H	11	3	$1s^1$	23	21	0.027	9.852
S	12	6	$1s^2$	24	66	-1.635	-0.845
			$2s^2$	25	66	-1.502	-0.845
			$2p^6$	26	66	0.048	-0.845
			$3s^2$	27	66	0.171	-0.845
			$3p^4$	28	66	6.597	-0.845
N	13	10	$1s^2$	29	84	-1.635	2.181
			$2s^2$	30	84	-1.502	2.181
			$2p^3$	31	84	-0.599	2.181
C	14	8	$1s^2$	32	84	-1.635	2.181
			$2s^2$	33	84	-1.502	2.181
			$2p^2$	34	84	8.973	2.181
C	15	10	$1s^2$	35	78	-1.635	-0.491
			$2s^2$	36	78	-1.502	-0.491
			$2p^2$	37	78	8.973	-0.491
C	16	7	$1s^2$	38	57	-1.635	0.639
			$2s^2$	39	57	-1.502	0.639
			$2p^2$	40	57	8.973	0.639
N	17	7	$1s^2$	41	57	-1.635	0.639
			$2s^2$	42	57	-1.502	0.639
			$2p^3$	43	57	-0.599	0.639
C	18	7	$1s^2$	44	57	-1.635	0.639
			$2s^2$	45	57	-1.502	0.639
			$2p^2$	46	57	8.973	0.639
O	19	3	$1s^2$	47	27	-1.635	0.834
			$2s^2$	48	27	-1.502	0.834
			$2p^4$	49	27	-2.768	0.834
H	20	3	$1s^1$	50	21	0.027	9.852
O	21	3	$1s^2$	51	27	-1.635	0.834
			$2s^2$	52	27	-1.502	0.834
			$2p^4$	53	27	-2.768	0.834

C	22	9	1s ²	54	51	-1.635	-0.125
			2s ²	55	51	-1.502	-0.125
			2p ²	56	51	8.973	-0.125
H	23	4	1s ¹	57	24	0.027	-9.411
H	24	4	1s ¹	58	24	0.027	-9.411
C	25	7	1s ²	59	33	-1.635	2.874
			2s ²	60	33	-1.502	2.874
			2p ²	61	33	8.973	2.874
H	26	4	1s ¹	62	18	0.027	-0.957
H	27	4	1s ¹	63	18	0.027	-0.957
H	28	4	1s ¹	64	18	0.027	-0.957
C	29	6	1s ²	65	36	-1.635	1.491
			2s ²	66	36	-1.502	1.491
			2p ²	67	36	8.973	1.491
H	30	4	1s ¹	68	18	0.027	-0.957
H	31	4	1s ¹	69	18	0.027	-0.957
H	32	4	1s ¹	70	18	0.027	-0.957

Table 6. Chemical Structures of the Compounds Studied and Their Anti-HIV-1 Activity ($\log (1/EC50)$)(see Fig. 1)

Training set

ID	R ₁	R ₂	R ₃	X	DCW	exp.	calc.	exp. -calc.
1	2-Me	Me	CH ₂ OCH ₂ CH ₂ OH	O	-3.618	4.150	4.733	-0.583
2	2-NO ₂	Me	CH ₂ OCH ₂ CH ₂ OH	O	-16.717	3.850	3.849	0.001
3	2-OMe	Me	CH ₂ OCH ₂ CH ₂ OH	O	-11.675	4.720	4.190	0.530
4	3-Et	Me	CH ₂ OCH ₂ CH ₂ OH	O	13.626	5.570	5.896	-0.326
5	3-t-Bu	Me	CH ₂ OCH ₂ CH ₂ OH	O	-1.645	4.920	4.866	0.054
6	3-CF ₃	Me	CH ₂ OCH ₂ CH ₂ OH	O	-10.168	4.350	4.291	0.059
7	3-F	Me	CH ₂ OCH ₂ CH ₂ OH	O	1.012	5.480	5.045	0.435
8	3-Cl	Me	CH ₂ OCH ₂ CH ₂ OH	O	4.294	4.890	5.267	-0.377
9	3-Br	Me	CH ₂ OCH ₂ CH ₂ OH	O	3.831	5.240	5.235	0.005
10	3-I	Me	CH ₂ OCH ₂ CH ₂ OH	O	0.387	5.000	5.003	-0.003
11	3-NO ₂	Me	CH ₂ OCH ₂ CH ₂ OH	O	-2.553	4.470	4.805	-0.335
12	3-OH	Me	CH ₂ OCH ₂ CH ₂ OH	O	-12.098	4.090	4.161	-0.071
13	3-Ome	Me	CH ₂ OCH ₂ CH ₂ OH	O	-5.435	4.660	4.610	0.050
14	3,5-Cl ₂	Me	CH ₂ OCH ₂ CH ₂ OH	O	12.621	5.890	5.828	0.062

15	3-COOMe	Me	CH ₂ OCH ₂ CH ₂ OH	O	2.540	5.100	5.148	-0.048
16	3-COMe	Me	CH ₂ OCH ₂ CH ₂ OH	O	10.597	5.140	5.692	-0.552
17	3-CN	Me	CH ₂ OCH ₂ CH ₂ OH	O	6.276	5.000	5.400	-0.400
18	H	CH ₂ CH =CH ₂	CH ₂ OCH ₂ CH ₂ OH	O	12.401	5.600	5.813	-0.213
19	H	Pr	CH ₂ OCH ₂ CH ₂ OH	S	13.445	5.000	5.884	-0.884
20	3,5-Me ₂	Et	CH ₂ OCH ₂ CH ₂ OH	S	36.143	8.110	7.415	0.695
21	3,5-Cl ₂	Et	CH ₂ OCH ₂ CH ₂ OH	S	33.682	7.370	7.249	0.121
22	H	Et	CH ₂ OCH ₂ CH ₂ OH	O	21.943	6.920	6.457	0.463
23	H	i-Pr	CH ₂ OCH ₂ CH ₂ OH	O	32.270	7.200	7.154	0.046
24	3,5-Me ₂	Et	CH ₂ OCH ₂ CH ₂ OH	O	37.501	7.890	7.506	0.384
25	3,5-Me ₂	i-Pr	CH ₂ OCH ₂ CH ₂ OH	O	47.828	8.570	8.203	0.367
26	3,5-Cl ₂	Et	CH ₂ OCH ₂ CH ₂ OH	O	35.040	7.850	7.340	0.510
27	4-Me	Me	CH ₂ OCH ₂ CH ₂ OH	O	-15.203	3.660	3.952	-0.292
28	H	Me	CH ₂ OCH ₂ CH ₂ OH	O	-0.476	5.150	4.945	0.205
29	H	Me	CH ₂ OCH ₂ CH ₂ OH	S	-1.834	6.010	4.853	1.157
30	H	I	CH ₂ OCH ₂ CH ₂ OH	O	6.931	5.440	5.444	-0.004
31	H	CH=C H ₂	CH ₂ OCH ₂ CH ₂ OH	O	7.228	5.690	5.465	0.225
32	H	CH=C HPh	CH ₂ OCH ₂ CH ₂ OH	O	7.730	5.220	5.498	-0.278
33	H	CH ₂ Ph	CH ₂ OCH ₂ CH ₂ OH	O	4.328	4.370	5.269	-0.899
34	H	Me	CH ₂ OCH ₂ CH ₂ OMe	O	7.531	5.060	5.485	-0.425
35	H	Me	CH ₂ OCH ₂ CH ₂ OAc	O	-2.781	5.170	4.789	0.381
36	H	Me	CH ₂ OCH ₂ CH ₂ OCH OPh	O	-2.275	5.120	4.824	0.296
37	H	Me	CH ₂ OCH ₂ CH ₂ Cl	O	17.774	5.820	6.176	-0.356
38	H	Me	CH ₂ OCH ₂ CH ₂ N ₃	O	-1.987	5.240	4.843	0.397
39	H	Me	CH ₂ OCH ₂ CH ₂ F	O	13.716	5.960	5.902	0.058
40	H	Me	CH ₂ OCH ₂ Ph	O	18.028	7.060	6.193	0.867
41	H	Et	CH ₂ OCH ₂ Me	O	36.440	7.720	7.435	0.285
42	3,5-Me ₂	Et	CH ₂ OCH ₂ Me	O	51.998	8.240	8.484	-0.244
43	H	Et	CH ₂ OCH ₂ Ph	O	40.447	8.230	7.705	0.525
44	3,5-Me ₂	Et	CH ₂ OCH ₂ Ph	S	54.647	8.140	8.663	-0.523
45	H	i-Pr	CH ₂ OCH ₂ Ph	O	50.774	8.510	8.402	0.108
46	H	i-Pr	CH ₂ OCH ₂ Me	S	45.410	7.890	8.040	-0.150
47	H	i-Pr	CH ₂ OCH ₂ Ph	S	49.417	8.140	8.310	-0.170

48	H	Me	CH ₂ Ome	O	19.515	5.680	6.293	-0.613
49	H	Me	Et	O	17.285	5.660	6.143	-0.483
50	H	Me	Bu	O	6.565	5.920	5.420	0.500
51	3,5-Cl ₂	Et	CH ₂ OCH ₂ Me	S	48.179	7.890	8.227	-0.337
52	H	Et	CH ₂ O-i-Pr	S	25.605	6.660	6.704	-0.044
53	H	Et	CH ₂ OCH ₂ -c-Hex	S	20.639	6.450	6.369	0.081
54	H	Et	CH ₂ OCH ₂ C ₆ H ₄ (4-Cl)	S	40.302	7.920	7.695	0.225
55	H	Et	CH ₂ O-i-Pr	O	26.963	6.470	6.796	-0.326
56	H	Et	CH ₂ O-c-Hex	O	3.114	5.400	5.187	0.213
57	H	Et	CH ₂ OCH ₂ -c-Hex	O	21.996	6.350	6.461	-0.111
58	H	c-Pr	CH ₂ OCH ₂ Me	S	31.250	7.020	7.085	-0.065
59	H	c-Pr	CH ₂ OCH ₂ Me	O	32.607	7.000	7.176	-0.176

Test Set

1	3-Me	Me	CH ₂ OCH ₂ CH ₂ OH	O	2.622	5.590	5.154	0.436
2	3,5-Me ₂	Me	CH ₂ OCH ₂ CH ₂ OH	O	15.082	6.590	5.994	0.596
3	3,5-Me ₂	Me	CH ₂ OCH ₂ CH ₂ OH	S	13.724	6.660	5.903	0.757
4	H	Et	CH ₂ OCH ₂ CH ₂ OH	S	20.585	6.960	6.365	0.595
5	H	i-Pr	CH ₂ OCH ₂ CH ₂ OH	S	30.913	7.230	7.062	0.168
6	3,5-Me ₂	i-Pr	CH ₂ OCH ₂ CH ₂ OH	S	46.471	8.300	8.111	0.189
7	H	Pr	CH ₂ OCH ₂ CH ₂ OH	O	14.802	5.470	5.975	-0.505
8	H	Me	CH ₂ OCH ₂ Me	O	14.021	6.480	5.923	0.557
9	H	Me	CH ₂ OCH ₂ CH ₂ Me	O	5.301	5.480	5.335	0.145
10	H	Et	CH ₂ OCH ₂ Me	S	35.083	7.580	7.343	0.237
11	3,5-Me ₂	Et	CH ₂ OCH ₂ Me	S	50.640	8.300	8.393	-0.093
12	3,5-Me ₂	Et	CH ₂ OCH ₂ Ph	O	56.004	8.550	8.754	-0.204
13	H	Et	CH ₂ OCH ₂ Ph	S	39.089	8.090	7.614	0.476
14	H	i-Pr	CH ₂ OCH ₂ Me	O	46.768	7.990	8.131	-0.141
15	H	Me	CH ₂ OBu	O	1.721	5.330	5.093	0.237
16	H	Et	CH ₂ O-C-Hex	S	1.757	5.790	5.095	0.695
17	H	Et	CH ₂ OCH ₂ C ₆ H ₄ (4-Me)	S	24.363	7.110	6.620	0.490
18	H	Et	CH ₂ OCH ₂ CH ₂ Ph	S	31.948	7.040	7.132	-0.092
19	3,5-Cl ₂	Et	CH ₂ OCH ₂ Me	O	49.537	8.130	8.318	-0.188
20	H	Et	CH ₂ OCH ₂ CH ₂ Ph	O	33.306	7.020	7.223	-0.203

Conclusions

The optimization of correlation weights of local invariants in graph of atomic orbitals may be used as a useful tool for predicting numerical values of Anti-HIV-1 activity. This sort of

molecular descriptor belongs to the category of flexible molecular descriptors and they have shown to be quite suitable for QSAR/QSPR studies to predict several kinds of physicochemical properties and biological activities [24-29]. The advantages and usefulness of this method can be summarized as follows:

- The method is relatively simple to be implemented in QSAR/QSPR studies.
- Numerical accuracy is satisfactory and it gives a solid basis for interpretative purposes in order to assess the main factors governing biological activities and physicochemical properties.
- The comparison between rigid and flexible molecular descriptors shows that the last category of molecular descriptor yields better predictive results.
- Present results on Anti-HIV-1 activity makes a significant contribution with regard to several previous studies on this issue.
- Main conclusions derived from this research allow us to extend present analysis in order to study other molecules having potential Anti-HIV-1 activity. At present some complementary calculations are under development and results will be presented elsewhere in the forthcoming future.

References

- [1] Mwau, M.; McMichel, A.J., *J. Gene Med.*, **2003**, *5*, 3-10.
- [2] Douali, L.; Villemin, D.; Cherqaoui, D., *J. Chem. Inf. Comput. Sci.*, **2003**, *43*, 1200-1207.
- [3] Castro, E.A.; Torrens, F.; Toropov, A.A.; Nesterov, I.V.; Nabiev, O.M., *Molec. Simul.*, **2004**, *30* (10), 691-696.
- [4] Garg, R.; Gupta, S.P.; Gao, H.; Babu, M.S.; Debnath, A.K.; Hansch, C., *Chem. Rev.*, **1999**, *99*, 3525-3601.
- [5] H. Bazoui, M. Zahouily, S. Boulajaaj, S. Sebti, D. Zakarya, *SAR-QSAR Environ. Res.*, **13** (2002), 567-577.
- [6] Luco, J.M.; Ferretti, F.H., *J. Chem. Inf. Comput. Sci.*, **1997**, *37*, 392-401.
- [7] Katritzky, A.R.; Oliferenko, A.; Lomaka, A.; Karelson, M., *Bioorg. Med. Chem. Lett.*, **2002**, *12*, 3453-3457.
- [8] Gayathri, P.; Pande, V.; Sivakumar, R.; Gupta, S.P., *Bioorg. Med. Chem.*, **2001**, *9*, 3059-3063.
- [9] Pungpo, P.; Hannongbua, S., *J. Mol. Graph. Model.*, **2000**, *18*, 581-590.
- [10] Siddiqui, A.Q.; McGuigan, C.; Ballatore, C.; Zuccotto, F.; Gilbert, I.H.; De Clercq, E.; Balzarini, J., *J. Med. Chem.*, **1999**, *42*, 4122-4128.
- [11] Huang, X.; Xu, L.; Luio, X.; Fan, K.; Ji, R.; Pei, G.; Chen, K.; Jiang, H., *J. Med. Chem.*, **2002**, *45*, 333-343.
- [12] Chan, J.H.; Hong, J.S.; Hunter III, R.N.; Orr, G.F.; Cowan, J.L.; Sherman, D.L.; Sparks, S.M.; Reitter, B.E.; Andrews III, C.W.; Hazen, R.J.; St. Clair, M.; Boone, L.R.; Ferris, R.G.; Creech, K.L.; Roberts, G.B.; Short, S.A.; Weaver, K.; Ott, R.J.; Ren, J.; Hopkins, A.; Stuart, D.I.; Stammers, D.K., *J. Med. Chem.*, **2001**, *44*, 1866-1882.

- [13] Leonard, J.T.; Roy, K., *QSAR Comb. Sci.*, **2004**, *23*, 23-35.
- [14] Toropov, A.A.; Toropova, A.P., *J. Mol. Struct. (THEOCHEM)*, **2001**, *538*, 197-199.
- [15] Toropov, A.A.; Toropova, A.P., *J. Mol. Struct. (THEOCHEM)*, **2001**, *538*, 287-293.
- [16] Toropov, A.A.; Toropova, A.P., *J. Mol. Struct. (THEOCHEM)*, **2002**, *581*, 11-15.
- [17] Toropov, A.A.; Toropova, A.P., *J. Mol. Struct. (THEOCHEM)*, **2003**, *637*, 1-10
- [18] Toropov, A.A.; Toropova, A.P., *Internet Electron. J. Mol. Des.*, **2002**, *1*, 109-113, <http://www.biochempress.com>.
- [19] Peruzzo, P.J.; Marino, D.J.G.; Castro, E.A.; Toropov, A.A., *Internet Electron. J. Mol. Des.*, **2003**, *2*, 334-347, <http://www.biochempress.com>.
- [20] Toropov, A.A.; Toropova, A.P.; Nesterov, I.V.; Nabiev, O.M., *J. Mol. Struct. (THEOCHEM)* **2003**, *640*, 175-181.
- [21] Toropov, A.A.; Toropova, A.P.; Nesterova, A.I.; Nabiev, O.M., *Chem. Phys. Letters*, **2004**, *384*, 357-363.
- [22] Toropov, A.A., Roy, K., *J. Chem. Inf. Comput. Sci.*, **2004**, *44*, 179-186.
- [23] Toropov, A.A.; Benfenati, E., *J. Mol. Struct. (THEOCHEM)*, **2004**, *676*, 165-169.
- [24] Toropov, A.A.; Duchowicz, P.R.; Castro, E.A., *Molecules*, **2004**, *9*, 1019-1033.
- [25] Castro, E.A.; Toropova, A.P.; Toropov, A.A.; Mekhamedjanova, D.V., *Struct. Chem.*, **2005**, *16*(3), 311-330.
- [26] Castro, E.A.; Toropov, A.A.; Nesterova, A.I.; Nabiev, O.M., *Cent. Eur. J. Chem.*, **2004**, *2*(3), 1-24.
- [27] Duchowicz, P.R.; Castro, E.A.; Toropov, A.A.; Nesterova, I.V.; Nabiev, O.M. *Molec. Diversity* **2004**, *8*(4), 325-330.
- [28] Duchowicz, P.R.; Castro, E.A.; Toropov, A.A.; Nesterova, A.I.; Nabiev, O.M., *J. Argent. Chem. Soc.*, **2004**, *92*(1-3), 29-42.
- [29] Castro, E.A.; Duchowicz, P.R.; Toropov, A.A.; Nesterova, A.I.; Nazarov, A.U., *J. Theor. Comput. Chem.*, **2004**, *3*(1), 31-41.