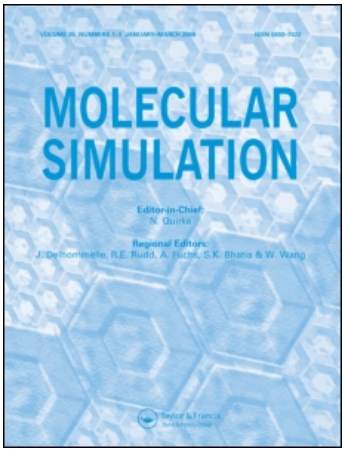

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

QSAR Modeling ANTI-HIV-1 Activities by Optimization of Correlation Weights of Local Graph Invariants

E. A. Castro[a]; F. Torrens[b]; A. A. Toropov[c]; I. V. Nesterov[c]; O. M. Nabiev[c]
[a] INIFTA, Departamento de Química, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, La Plata, Argentina [b] Facultad de Química, Departament de Química Física, Institut Universitari de Ciencia Molecular, Universitat de Valencia, Burjassot, Valencia, Spain [c] Scientifical Research Institute "Algorithm-Engineering", Tashkent, Uzbekistan

To cite this Article Castro, E. A. , Torrens, F. , Toropov, A. A. , Nesterov, I. V. and Nabiev, O. M.(2004) 'QSAR Modeling ANTI-HIV-1 Activities by Optimization of Correlation Weights of Local Graph Invariants', Molecular Simulation, 30: 10, 691 — 696
To link to this Article: DOI: 10.1080/08927020412331272449
URL: http://dx.doi.org/10.1080/08927020412331272449

Molecular Simulation, VOL. 30 (10), 30 August 2004, pp. 691–696

Taylor & Francis
Taylor & Francis Group

QSAR Modeling ANTI-HIV-1 Activities by Optimization of Correlation Weights of Local Graph Invariants

E.A. CASTRO[a,*], F. TORRENS[b], A.A. TOROPOV[c], I.V. NESTEROV[c] and O.M. NABIEV[c]

[a]*INIFTA, Departamento de Química, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Suc.4, C.C. 16, La Plata 1900, Argentina;* [b]*Facultad de Química, Departament de Química Física, Institut Universitari de Ciencia Molecular, Universitat de Valencia, Dr Moliner 50, E-46100 Burjassot, Valencia, Spain;* [c]*Scientifical Research Institute "Algorithm-Engineering", F. Khodjaev Street 25, 700125, Tashkent, Uzbekistan*

(Received March 2004; In final form June 2004)

Results of using descriptors calculated with the correlation weights (CWs) of local graph invariants for modeling of anti-HIV-1 potencies of two groups of reverse transcriptase (RT) inhibitors are reported. Presence of different chemical elements in molecular structure of the inhibitors and the presence of Morgan extended connectivity values of zeroth-, first- and second order have been examined as local graph invariants in the labeled hydrogen-filled graphs. By Monte Carlo method optimization procedure, values of the CWs which produce as large values as possible of correlation coefficient between the numerical data on the anti-HIV-1 potencies and values of the descriptors on the training set have been computed. The model of the anti-HIV-1 activity obtained with compounds of training set by means of optimization of correlation weights of presence of chemical elements together with the presence of Morgan extended connectivity of first order is reasonable well model for the prediction of endpoints under consideration for compounds of the test set.

Keywords: QSAR Modeling; Anti-HIV-1 activity; Correlation weight of local graph invariants; Flexible topological descriptors

INTRODUCTION

Acquired immunodeficiency syndrome (AIDS) is the most devastating pandemic in recent history of the mankind. It portends an increasing toll in human suffering and is a major hurdle in the economic progress of many countries. Current approaches for the treatment of AIDS using single agents are plagued by the development of resistance. Combination therapies employing multiple components, each aimed against different viral enzymes, may potentially provide an effective means of countering such resistance [1]. Since human type 1 virus (HIV-1) is the causative agent of AIDS, extensive works are currently going on to block its replication. In fact, the HIV-1 reverse transcriptase (RT) has become an attractive active target for several antiviral therapeutic agents used in the treatment of AIDS. The main biological function of HIV-1 RT is an essential enzyme involved in the life cycle of the HIV responsible for virus replication from single-stranded RNA viral genome into a double-stranded proviral DNZ, which is then integrated into the host chromosome. During the past decade, several compounds with a wide variety of structures have been identified to counteract the activity of HIV-1 RT [2–4].

Since 1991, compounds possessing anti-HIV activity have been the subject of numerous quantitative structure–activity relationship (QSAR) studies. A search in the chemical abstract database [5] for the "QSAR and HIV" query led to about 200 references describing structure–property relationships established using partial least-squares (PLS), artificial neural network (ANN) and multiple linear regression (MLR) methods involving 1D and/or 2D descriptors [6–11] or 3D descriptors [12,13], the 4D-QSAR technique [14], comparative molecular field analysis (CoMFA) [15–19] and electrostatic potential distribution [20,21].

Quantitative structure–property (activity) relationships (QSPR/QSAR) based on descriptors calculated with chemical graphs have shown to be quite useful tools to predict different physicochemical properties and biological activities of interest in different fields [22–29]. Recently, the optimization of correlation

*Corresponding author. E-mail: castro@quimica.unlp.edu.ar; jubert@arnet.com.ar

ISSN 0892-7022 print/ISSN 1029-0435 online © 2004 Taylor & Francis Ltd
DOI: 10.1080/08927020412331272449

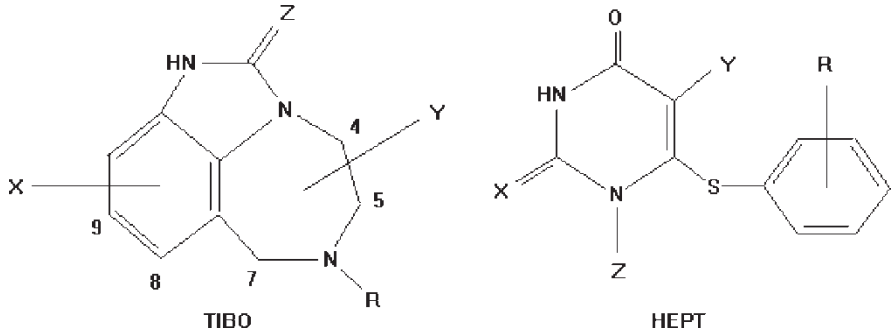

FIGURE 1 The TIBO and the HEPT derivatives, differing in *X,Y,Z* and *R* substituents.

weights of local graph invariants (OCWLGI) has been suggested as a suitable approach of the QSPR/QSAR analyses [30–40]. This particular set of molecular descriptors belongs to the more encompassing category of flexible descriptors. The concept of flexible topological descriptors, originally introduced by Randic [41–43], is a major advance with regard to the possibility to extract a maximum amount of chemical information and, at the same time, the descriptors used in a multiple regression equation are not intercorrelated among themselves. The difficulties of multiple regression are not present in such an approach, which is based on regression with a single descriptor. Unlike usual fixed topological descriptors, flexible topological descriptors do not have a definite predetermined value, which can be applied to any sets of compounds for the modeling of physicochemical property or/and biological activity. In fact, the formalism of such descriptors is defined on the base of an optimization procedure to get the best optimal relation for a particular data set. Thus, the definition of the descriptors will vary from data set to the other set, and the ultimate purpose of the iterative optimization procedure is to obtain the best predictive model.

The aim of the present study is to apply the optimization of correlation weights scheme for modeling of the anti-HIV-1 activities of inhibitors from Ref. [22] to the show usefulness of the scheme. Two sets of data have been employed for the present analysis: TIBO derivatives and HEPT derivatives (see Fig. 1). In a series of studies, Ho *et al.* demonstrated that certain compounds of type 1 show anti-HIV-1 activity, functioning as nonnucleoside RT inhibitors [44]. They systematically synthesized and tested numerous members of this family, differing in *X*, *Y*, *Z* and *R*. Here, *Z* was restricted to oxygen and or sulfur. Potency was found to be enhanced by substitution on the 8-position, by letting *R* be the 3,3-dimethylallyl group, and when $Z = S$. Another class of HIV-1 RT inhibitors, HEPT derivatives, are of type 2 [6,45]. The *X* can be oxygen or sulfur, but now it is the former that tends to produce a somewhat higher level of activity.

Effectiveness in inhibiting HIV-1 was measured by the concentration of the compound, C_{50}, required to achieve 50% protection of MT-4 cells against the virus [6,44,45]. We have considered these experimental results for 38 TIBO and 19 HEPT derivatives as representative databases for developing regression equations for the anti-HIV-1 potencies of these classes of compounds.

METHOD

Models of anti-HIV-1 activities which have been examined in the present study are based on labeled hydrogen filled graph (LHFG) in the following manner:

$$\mathrm{DCW}(a_k, {}^{x}\mathrm{EC}_k) = \sum_{k=1}^{n} [\mathrm{CW}(a_k) + \mathrm{CW}({}^{x}\mathrm{EC}_k)] \qquad (1)$$

In the above equation, the DCW term represents the molecular descriptor, CW terms represent the correlation weights, a_k is the chemical element which is image of the kth vertex of the LHFG and ${}^{x}\mathrm{EC}_k$ is the Morgan extended connectivity [13]. As local invariants, we have used the Morgan extended connectivity of zero, first and second order (denoted by ${}^{0}\mathrm{EC}_k$, ${}^{1}\mathrm{EC}_k$ and ${}^{2}\mathrm{EC}_k$, respectively).

Zero-order Morgan connectivity of an atom k is the adjacency count of that atom. Again, the first-order Morgan connectivity value of atom k is the sum of the zero-order Morgan connectivity values of the atoms that are connected to atom k. Similarly, the second-order Morgan connectivity value of atom k is the sum of the first-order Morgan connectivity values of the atoms that are connected to atom k. An example of calculations of Morgan connectivity indices for ethanol is given in Fig. 2.

The starting value of each correlation weight was 1 and resorting to the Monte Carlo interative optimization procedure [29,46,47] the best values of correlation weights (i.e. $\mathrm{CW}(a_k)$ and $\mathrm{CW}({}^{x}\mathrm{EC}_k)$ were found out. The "best values of CW's" means those which give largest possible correlation coefficient

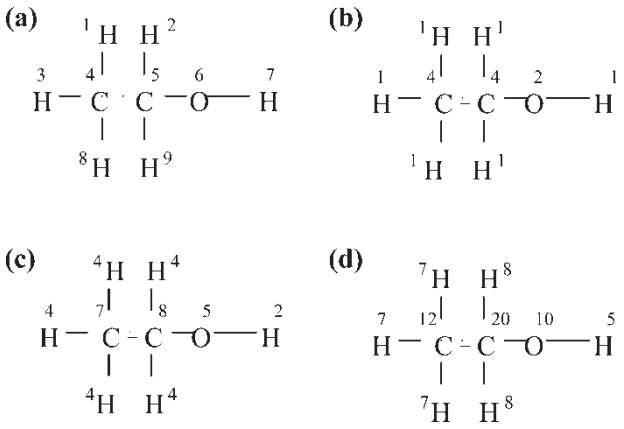

FIGURE 2 Local invariants of the LHFG of ethanol: (a) arbitrary numbering of the vertices, (b) ^{0}EC, (c) ^{1}EC, (d) ^{2}EC.

between the $\log(10^6/C_{50})$ values of the molecular set and the molecular descriptor DCW. Finally, the molecular descriptor is defined on the basis of the optimized correlation weights and it was then used to derive all the relations with $\log(10^6/C_{50})$ values, employing the least-squares method of regression

$$\log\left(10^6/C_{50}\right) = C_0 + C_1 \cdot \mathrm{DCW}(a_k, {}^x\mathrm{EC}) \qquad (2)$$

The data set was divided into a training and a test set, as listed in Table III. The optimization of correlation weights were done resorting to a program developed by one of the authors (AAT). Least-squares linear regression analyses were done employing a standard computer program and statistical quality of equations were judged by examining the parameters r (correlation coefficient), F (variance ratio), s (standard error of estimate) and AVRES (average of absolute values of residuals).

RESULTS AND DISCUSSION

Table I lists results of OCWLGI of three probes based on optimization of correlation weights of different chemical elements present, CW(a_k), together with the CW(^{x}EC$_k$). The correlation weights of first OCWLGI probe are presented in Table II. The results of the optimized correlation weights vary for each try due to the statistical nature of the Monte Carlo method. However, differences are not significant for different probes corresponding to a given descriptor, as shown in Table I for the three cases reported for each OCWLGI based on the DCW(a_k,^{x}EC$_k$, $x = 0, 1, 2$).

From Table I one can see that best QSAR model of the anti-HIV-1 activity takes place in case of the [DCW(a_k,^{1}EC$_k$), Table II lists relevant CWs values for calculating of the descriptor.

Among the molecular descriptors (DCW) defined in different ways [DCW(a_k,^{x}EC$_k$), $x = 0, 1, 2$), DCW(a_k,^{1}EC$_k$) gave the best correlation between the descriptor and the property [$\log(10^6/C_{50})$].

TABLE II Correlation weights of first OCWLGI based on the DCW(a_k,^{1}EC$_k$)

LHFG local invariants	*Correlation weights*
Chemical elements, a_k, CW(a_k)	
C	−0.575
N	−3.178
H	0.150
Br	1.298
S	2.333
F	−0.266
Cl	0.695
I	1.482
O	0.953
Morgan extended connectivity of first order, ^{1}EC$_k$, CW(^{1}EC$_k$)	
0002	3.390
0003	1.359
0004	0.275
0005	−3.775
0006	0.071
0007	−0.509
0008	1.408
0009	0.775
0010	0.625
0011	2.066
0012	1.906
0013	2.933

TABLE I Results of OCWLGI based on the DCW(a_k,^{0}EC$_k$), DCW(a_k,^{1}EC$_k$) and DCW(a_k,^{2}EC$_k$)

	Training set n = 37			*Test set n = 20*			*All compounds N = 57*		
Probe	*R*	*S*	*F*	*R*	*S*	*F*	*R*	*S*	*F*
OCWLGI based on the DCW(a_k,^{0}EC$_k$)									
1	0.8712	0.754	110	0.9121	0.625	89	0.8834	0.706	195
2	0.8660	0.768	105	0.9075	0.639	84	0.8784	0.720	186
3	0.8664	0.767	105	0.9012	0.649	78	0.8773	0.722	184
OCWLGI based on the DCW(a_k,^{1}EC$_k$)									
1	0.9320	0.557	231	0.9343	0.590	124	0.9295	0.564	349
2	0.9317	0.558	230	0.9357	0.576	127	0.9303	0.559	354
3	0.9321	0.557	232	0.9359	0.588	127	0.9300	0.562	352
OCWLGI based on the DCW(a_k,^{2}EC$_k$)									
1	0.9499	0.480	324	0.8990	0.730	76	0.9281	0.574	342
2	0.9498	0.481	322	0.9067	0.705	83	0.9308	0.563	356
3	0.9499	0.480	323	0.9010	0.705	78	0.9300	0.563	352

TABLE III Observed [22] and calculated (Eq. (2) with $x = 1$) $\log(10^6/C_{50})$ values of TIBO and HEPT structural types of RT inhibitors

No	*X*	*Y*	*Z*	*R*	*Type*	*Exp.*	*Calc.*	*Exp.-Calc.*
*Training set**								
1	8-Br	5-CH_3	S	DMA	TIBO	8.52	8.28	0.25
2	8-F	5-CH_3	S	DMA	TIBO	8.24	6.94	1.30
3	8-Cl	7-CH_3	S	DMA	TIBO	7.92	7.76	0.16
4	8-CH_3	5-CH_3	S	DMA	TIBO	7.87	7.63	0.24
5	9-Cl	5-CH_3	S	DMA	TIBO	7.47	7.76	−0.29
6	8-Cl	H	S	DMA	TIBO	7.34	7.42	−0.08
7	8-I	5-CH_3	S	DMA	TIBO	7.32	7.55	−0.23
8	8-CN	5-CH_3	S	DMA	TIBO	7.25	6.84	0.41
9	8-I	5-CH_3	O	DMA	TIBO	7.06	7.25	−0.19
10	H	5-CH_3	O	DMA	TIBO	7.01	6.11	0.90
11	9-Cl	7-CH_3	O	DMA	TIBO	6.80	6.58	0.22
12	9-CF_3	5-CH_3	S	DMA	TIBO	6.31	6.56	−0.25
13	8-CH_3	5-CH_3	O	DMA	TIBO	6.00	6.45	−0.45
14	8-CN	5-CH_3	O	DMA	TIBO	5.94	6.39	−0.45
15	9-NO_2	5-CH_3	S	CPM	TIBO	5.61	5.60	0.01
16	10-OCH_3	5-CH_3	S	DMA	TIBO	5.33	5.39	−0.06
17	9-CF_3	5-CH_3	O	DMA	TIBO	5.23	5.38	−0.15
18	H	7-CH_3	O	DMA	TIBO	4.92	6.11	−1.19
19	H	5-CH_3	O	$CH_2C(C_2H_5)=CH_2$	TIBO	4.43	4.43	0.01
20	H	5-CH_3	O	$CH_2CH_2CH=CH_2$	TIBO	4.30	4.45	−0.15
21	H	5-CH_3	O	C_3H_7	TIBO	4.22	3.62	0.60
22	H	5-CH_3	O	$CH_2CH=CH_2$	TIBO	4.15	4.09	0.06
23	H	4-$CH(CH_3)_2$	O	C_3H_7	TIBO	4.13	4.79	−0.66
24	8-NH_2	5-CH_3	O	CPM	TIBO	3.07	3.05	0.03
25	O	C_2H_5	C_6H_5	3,5-$(CH_3)_2$	HEPT	8.62	8.83	−0.21
26	O	$CH(CH_3)_2$	CH_2OH	3,5-$(CH_3)_2$	HEPT	8.48	7.90	0.58
27	O	C_2H_5	C_6H_5	H	HEPT	8.31	8.16	0.16
28	S	C_2H_5	CH_3	3,5-$(CH_3)_2$	HEPT	8.25	8.56	−0.31
29	O	C_2H_5	CH_3	3,5-$(CH_3)_2$	HEPT	8.21	7.38	0.84
30	O	$CH(CH_3)_2$	CH_3	H	HEPT	8.09	7.47	0.63
31	S	$CH(CH_3)_2$	CH_3	H	HEPT	7.92	8.65	−0.73
32	O	C_2H_5	CH_3	H	HEPT	7.66	6.70	0.96
33	S	C_2H_5	CH_3	H	HEPT	7.59	7.88	−0.29
34	O	C_2H_5	CH_2OH	H	HEPT	6.92	6.45	0.47
35	O	CH_3	CH_2OH	3,5-$(CH_3)_2$	HEPT	6.59	6.73	−0.14
36	O	CH_3	CH_2OH	3-CH_3	HEPT	5.59	6.39	−0.80
37	O	CH_3	C_2H_5	H	HEPT	5.52	6.70	−1.18
Test set†								
1	8-Cl	5-CH_3	S	DMA	TIBO	8.37	7.76	0.61
2	9-F	5-CH_3	S	DMA	TIBO	7.60	6.94	0.66
3	H	5,7-CH_3	S	DMA	TIBO	7.38	7.63	−0.25
4	H	5-CH_3	S	DMA	TIBO	7.36	7.29	0.07
5	8-Br	5-CH_3	O	DMA	TIBO	7.33	7.10	0.24
6	H	7-CH_3	S	DMA	TIBO	7.11	7.29	−0.18
7	8-Cl	7-CH_3	O	DMA	TIBO	6.84	6.58	0.26
8	9-Cl	H	S	DMA	TIBO	6.80	7.42	−0.62
9	H	7-CH_3	S	C_3H_7	TIBO	5.61	4.80	0.81
10	H	5-CH_3	O	DMA	TIBO	5.48	6.11	−0.63
11	10-OCH_3	5-CH_3	O	DMA	TIBO	5.18	4.21	0.97
12	9-NO_2	5-CH_3	O	CPM	TIBO	4.48	4.42	0.06
13	H	5-CH_3	O	$CH_2(CH_2)_2CH_3$	TIBO	4.00	4.39	−0.39
14	H	5-CH_3	O	$CH_2CO(O)CH_3$	TIBO	3.07	2.03	1.04
15	O	$CH(CH_3)_2$	C_6H_5	H	HEPT	8.47	8.92	−0.45
16	O	C_2H_5	CH_2OH	3,5-$(CH_3)_2$	HEPT	7.80	7.13	0.67
17	O	$CH(CH_3)_2$	CH_2OH	H	HEPT	7.14	7.22	−0.08
18	O	CH_3	C_6H_5	H	HEPT	7.03	7.75	−0.72
19	O	CH_3	CH_3	H	HEPT	6.48	6.30	0.18
20	O	CH_3	CH_2OH	H	HEPT	5.19	6.05	−0.86

* AVRES = 0.42. † AVRES = 0.49.

The analysis of data displayed in Table III show the existence of a quite sensible agreement among experimental and theoretical data. The AVRES for training and test sets, are similar, although, as expected, it is somewhat smaller for the training set. Since results for test set are true predictions, this similarity is encouraging. Among the 20 molecules comprising the test set there is only one of them having a relatively large deviation (molecule No. 14).

When comparing our theoretical results with respect to those previously published [22] for identical molecular set [22], we verify that present

ones are clearly superior. In fact, although statistical parameters associated to previous regression equations (see Eqs. (8)–(11) in Ref. [22]) and ours are nearly the same, numerical data corresponding to our test set are true predictions, while those reported by Politzer *et al.* corresponds to two separate sets (TIBO and HEPT derivatives) considered both as training sets. Besides, Politzer *et al.*'s equations employ three and four independent variables (i.e. molecular descriptors for TIBO and HEPT derivatives, respectively, while our regression equation depends upon just one variable (i.e. DCW(a_k, ^{x}EC)).

CONCLUSIONS

The employment of optimization of correlation weights of local graph invariants makes up a reasonable good approximation to predict the anti-HIV-activities for two representative molecular sets (TIBO and HEPT derivatives) resorting to Morgan extended connectivity index of first order. The power of flexible indices based on the optimization of correlation weights of local graph invariants is shown when comparing with other approach based on computed molecular surface electrostatic potentials [22] since this last method resorts to the use of several variables to attain similar results as present model grounded in one-variable regression equation. Moreover, the present scheme does not require complex calculation of diverse descriptors and statistical analysis for proper selection of descriptors and intercorrelation among them. Thus, the model merits additional assessment on exploring quantitative structure–property (activity) of different physicochemical properties and biological data using several different local invariants to justify its suitability in modeling studies.

Acknowledgements

Authors wish to express their gratefulness to scientific international fund of the Third Word Academy of Sciences (TWAS) for supporting this study.

References

[1] Makhija, M.T. and Kulkarni, V.M. (2001) "Eigen value analysis of HIV-1 integrase inhibitors", *J. Chem. Inf. Comput. Sci.* **41**, 1569–1577.
[2] Arnold, E., Das, K., Yadav, P.N.S., Hsiou, Y., Boyer, P.L. and Hughes, S.H. (1996) "Targeting HIV recerse transcriptase for anti-AIDS drug design", *Drug Des. Discov.* **13**, 29–47.
[3] De Clercq, E. (1998) "The role of nonnucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection", *Antiviral Res.* **38**, 153–179.
[4] Jonckheere, H., Anne, J. and De Clercq, E. (2000) "The HIV-1 reverse transcription (RT) process as target for RT inhibitors", *Med. Res. Rev.* **20**, 129–154.
[5] (2000) *SciFinder Scholar, version 2000.1*, American Chemical Society: Washington, DC, http://www.cas.org/SCIFINDER/SCHOLAR/index.html.
[6] Garg, R., Gupta, S.P., Gao, H., Babu, M.S., Debnath, A.K. and Hansch, C. (1999) "Comparative quantitative structure–activity relationship studies on anti-HIV drugs", *Chem. Rev.* **99**, 3525–3602.
[7] Jalali-Heravi, M. and Parastar, F. (2000) "Use of artificial neural networks in a qsar study of anti-HIV activity for a large group of HEPT derivatives", *J. Chem. Inf. Comput. Sci.* **40**, 147–154.
[8] Knaggs, M.H., McGuigan, C., Harris, S.A., Heshmati, P., Cahard, D., Gilbert, I.H. and Balzarini, J. (2000) "A QSAR study investigating the effect of L-Alanine ester variation on the anti-HIV activity of some phosphoramidate derivatives of d4T", *Bioorg. Med. Chem. Lett.* **10**, 2075–2078.
[9] Tronchet, J.M.J., Grigorov, M., Dolatshahi, N., Moriaud, F. and Weber, J.A. (1997) "A QSAR study confirming the heterogeneity of the HEPT derivative series regarding their interaction with HIV reverse transcriptase", *Eur. J. Med. Chem.* **32**, 279–299.
[10] Houskonen, J. (2001) "Modeling QSAR with the electrotopological state: TIBO derivatives", *J. Chem. Inf. Comput. Sci.* **41**, 425–429.
[11] Maw, H.H. and Hall, L.H. (2002) "E-State modeling of HIV-1 protease inhibitor binding independent of 3D information", *J. Chem. Inf. Comput. Sci.* **42**, 190–298.
[12] Gancia, E., Bravi, G., Mascagni, P. and Zaliani, A. (2000) "Global 3D-QSAR methods: MS-WHIM and autocorrelation", *J. Comput.-Aided Mol. Des.* **14**, 293–306.
[13] Klein, C.T., Lawtrakul, L., Hannongbua, S. and Wolschann, P. (2000) "Accessible charges in structure–activity relationships a study on HEPT-based HIV-1 RT inhibitors", *Sci. Pharm.* **68**, 25–40.
[14] Santos-Filho, O.A. and Hopfinger, A.J. (2002) "The 4D-QSAR paradigm: application to a novel set of Nonpetidic HIV protease inhibitors", *Quant. Struct.-Act. Relat.* **21**, 369–381.
[15] Kireev, D.B., Chretien, J.R. and Raevsky, O.A. (1995) "Molecular modeling and quantitative structure–activity studies of anti-HIV-1,2-heteroarylquinoline-4-amines", *Eur. J. Med. Chem.* **30**, 395–402.
[16] Hennongbua, S., Nivesanond, K., Lawtrakul, L., Pungpo, P. and Wolschann, P. (2001) "3D-Quantitative structure–activity relationships of HEPT derivatives as HIV-1 reverse transcriptase inhibitors, based on *Ab Initio* calculations", *Chem. Inf. Comput. Sci.* **41**, 848–855.
[17] Jayatilleke, P.R.N., Nair, A.C., Zauhar, R. and Welsh, W.J. (2000) "Computational studies on HIV-1 protease inhibitors: influence of calculated inhibitor-enzyme binding affinities on the statistical quality of 3D-QSAR CoMFA models", *J. Med. Chem.* **43**, 4446–4451.
[18] Debnath, A.K. (1999) "Three-dimensional quantitative structure–activity relationhsips study on cyclic urea derivatives as HIV-1 protease inhibitors: application to comparative molecular field analysis", *J. Med. Chem.* **42**, 249–259.
[19] Buolamwini, J.K. and Assefa, H. (2002) "CoMFA and CoMSIA 3D QSAR and docking studies on conformationally-restrained cinnamoyl HIV-1 integrase inhibitars: exploration of a binding mode at the active site", *J. Med. Chem.* **45**, 841–852.
[20] Mickle, T. and Nair, V. (2000) "Anti-human immunodeficiency virus activities of nucleosides and nucleotides: correlation with molecular electrostatic potential data", *Antimicrob. Agents Chemother.* **44**, 2939–2947.
[21] Mickel, T. and Nair, V. (2000) "Predictive QSAR analysis of anti-HIV agents", *Drugs Future* **25**, 393–400.
[22] Gonzáles, O.G., Murray, J.S., Peralta-Inga, Z. and Politzer, P. (2001) "Computed molecular surface electrostatic potentials of two groups of riverse transcriptase inhibitors: relationships to anti- HIV-1 activities", *Int. J. Quantum Chem.* **83**(3–4), 115–124.
[23] Firpo, M., Gavernet, L., Castro, E.A. and Toropov, A.A. (2000) "Maximum topological distances based indices as molecular descriptors for QSPR. Part 1. Application to alkyl benzenes boiling points", *J. Mol. Struct. (Theochem)* **501–502**, 419–425.
[24] Castro, E.A., Tueros, M. and Toropov, A.A. (2000) "Maximum topological distances based indices as molecular descriptor

for QSPR. 2-Application to Aromatic Hydrocarbons", *Comput. Chem.* **24**, 571–576.
[25] Mercader, A., Castro, E.A. and Toropov, A.A. (2001) "Maximum topological distances based indices as molecular descriptors for QSPR. 4-Modeling the enthlpy of formation of hydrocarbons from elements", *Int. J. Mol. Sci.* **2**, 121–134.
[26] Toropov, A., Toropova, A., Ismailov, T. and Bonchev, D. (1998) "3D weighting of molecular descriptors for QSPR/QSAR by the method of ideal symmetry (MIS). 1. Application to boiling points of alkanes", *J. Mol. Struct. (Theochem)* **424**, 237–247.
[27] Krenkel, G., Castro, E.A. and Toropov, A.A. (2002) "3D and 4D Molecular models derived from the ideal symmetry method: prediction of alkanes normal boiling points", *Chem. Phys. Lett.* **355**, 517–528.
[28] Toropov, A.A. and Toropova, A.P. (1997) "Method of ideal symmetry in four-dimensional space: implementation in the QSPR studies on the thermochemistry of complex compounds", *Russ. J. Coord. Chem.* **23**, 741–747.
[29] Toropov, A.A. and Toropova, A.P. (1998) "Optimization of correlation weights of the local graph invariants: use of the enthalpies of formation of complex compounds for the QSPR modeling", *Russ. J. Coord. Chem.* **24**, 81–85.
[30] Toropov, A.A. and Toropova, A.P. (2001) "Modeling QSPR of stability of complexes of adenosine phosphate derivatives with metals absent from the complexes of the teaching access", *Russ. J. Coord. Chem.* **27**, 574–578.
[31] Toropov, A.A. and Toropova, A.P. (2002) "Modeling of acyclic compounds normal boiling points by correlation weighting of nearest neighboring codes", *J. Mol. Struct. (Theochem)* **581**, 11–15.
[32] Toropov, A.A. and Toropova, A.P. (2001) "Prediction of heteroaromatic amine mutagenicity by means of correlation weighting of atomic orbital graph of local invariants", *J. Mol. Struct. (Theochem)* **538**, 287–293.
[33] Toropov, A.A. and Toropova, A.P. (2002) "QSAR modeling of mutagenicity based on graph of atomic orbitals, internet electron", *J. Mol. Des.* **1**, 109–113, http://www.biochempress.com..
[34] Toropov, A.A. and Toropova, A.P. (2002) "Modeling QSAR of toxicity on optimization of correlation weights of morgan extended connectivity", *J. Mol. Struct. (Theochem)* **578**, 129–134.
[35] Toropova, A.P., Toropov, A.A., Ishankhodzhaeva, M.M. and Parpiev, N.A. (2000) "Modeling QSPR of stability constants of coordination compounds by optimization of correlation weights of local graph invariants", *Russ. J. Inorg. Chem.* **45**, 1057–1059.
[36] Krenkel, G., Castro, E.A. and Toropov, A.A. (2001) "Improved molecular descriptors to calculate boiling points based on the optimization of correlation weights of local graph invarinats", *J. Mol. Struct. (Theochem)* **542**, 107–113.
[37] Krenkel, G., Castro, E.A. and Toropov, A.A. (2001) "Improved molecular descriptors based on the optimization of correlation weights of local graph invariants", *Int. J. Mol. Sci.* **2**, 57–65.
[38] Mercader, A., Castro, E.A. and Toropov, A.A. (2000) "Modeling QSPR of the enthalpy of formation from elements by means of correlation weighting of local invariants of atomic orbital molecular graphs", *Chem. Phys. Lett.* **330**, 612–623.
[39] Duchowicz, P.R., Castro, E.A. and Toropov, A.A. (2002) "Improved QSPR analysis of standard entropy of acyclic and aromatic compounds using optimized correlation weights of linear graph invariants", *Comput. Chem.* **26**, 327–332.
[40] Peruzzo, P.J., Marino, D.J.G., Castro, E.A. and Toropov, A.A. (2001) "Calculation of pK values of flavylium salts from the optimization of correlation weights of local graph invariants", *J.Mol. Struct. (Theochem)* **572**, 53–60.
[41] Randic, M. (1991) "On computation of optimal parameters for multivariate analysis of structure–property relationship", *J. Comput. Chem.* **12**, 970–980.
[42] Randic, M. (1991) "Novel graph theoretic approach to heteroatoms in quantitative structure–activity relationships", *Chemom. Intell. Lab. Syst.* **10**, 213–227.
[43] Randic, M. (1991) "Resolution of ambiguities in structure–property studies by use of orthogonal descriptors", *J. Chem. Inf. Comput. Sci.* **31**, 311–320.
[44] Ho, W., Kukla, M.J., Breslin, H.J., Ludovici, D.W., Grous, P.P., Diamond, C.J., Miranda, M., Rodgers, J.D., Ho, C.Y., De Clercq, E., Pauwels, R., Andries, K., Janssen, M.A.C. and Janssen, P.A.J. (1995) *J. Med. Chem.* **38**, 794.
[45] Baba, M., Shigeta, S., Tanaka, H., Miyasaka, T., Ubasawa, M., Umezu, K., Walker, R.T., Pauwles, R. and De Clerq, E. (1992) *Antiviral Res.* **17**, 245.
[46] Toropov, A.A., Toropova, A.P., Voropaeva, N.L., Ruban, I.N. and Rashidova, S.Sh. (1988) "Generalized zero-order molecular connectivity indez: enthalpies of crystalline aquo and ammino complexes in QSPR modeling", *Russ. J. Coord. Chem.* **24**, 525–529.
[47] Toropov, A.A., Voropaeva, N.L., Ruban, I.N. and Rashidova, S.Sh. (1999) "Quantitative structure–property relationships for binary polymer solvent systems: correlation weighting of the local invariants of molecular graphs", *Polym. Sci. Ser. A* **41**, 975–985.