June 2007 Volume 3 Issue 1

Trade Science Inc.

Materials Science

An Indian Journal

Full Paper

MSAIJ, 3(1), 2007 [55-63]

Quantum Chemistry Based QSAR Study On HIV Drugs Of Protease (PR) Groups And New Drugs Proposed.

Corresponding Author

P.P.Singh
Department of Chemistry, Bareilly College
Bareilly, U.P., (INDIA)
E-Mail: dr_opsingh@sify.com

Co- Author

V.K.Sahu
Department Of Chemistry, Bareilly college Bareilly, U.P., (INDIA)

Received: 16th January, 2007
Accepted: 21st January 2007

ABSTRACT

With the help of PM3 calculations using cache software QSAR study of two set of derivatives of HIV inhibitors have been made. These belong to protease inhibitors group. The parent compounds of protease inhibitors are urea isostere and other isostere derivatives. The correlation coefficient values of QSAR models are above 0.70. The combination of descriptors providing the best correlation coefficient value are heat of formation(ΔH_f), total energy(TE), highest occupied molecular orbital (εHOMO) and electronegativity (χ). The best combinations have been used to predict the activity of fourteen new derivatives of urea isostere. The predicted activities of new derivatives have correlation coefficient above 0.80.

KEYWORDS

HIV inhibitors;
QSAR;
MLR;
Protease inhibitor;
Protease Enzyme.

INTRODUCTION

The HIV inhibitors have been classified in to seven groups: viral binding inhibitors[1-2], virus cell fusion inhibitors[3], virus uncoating inhibitors[4], reverse transcriptase inhibitors[5], integrase inhibitors[6], gene expression inhibitors[7] and protease inhibitors [8-11]. However, we have confined our study to protease inhibitors. Protease is employed at the cleavage events and virion maturation of the viral replication cycle, when new virus particles are being produced within an HIV-infected cell. When the protease enzyme is inhibited, an HIV-infected cell can only produce immature, non-infectious viral progeny. We in this paper present QSAR models of protease inhibitor[12] with the help of quantum chemical descriptors, recently employed by us for QSAR models[13-19].

MATERIAL AND METHOD

The study materials of this paper are presented

in two sets. The first set comprises of derivatives of urea isostere and the second set of derivatives of another isostere. A third set of new derivatives of urea isostere have been suggested as HIV inhibitors, and have been studied.

For QSAR prediction, the 3D modeling and geometry optimization of all the derivatives of protease inhibitors have been done with the help of PCMODEL software using the semiemperical PM3 Hamiltonian. The MOPAC calculations have been performed with Win MOPAC 7.21 software by applying key words charge=0, gnorm-0.1, bonds, geook, vector density. The values of quantum chemical descriptors that have been used for QSAR models have been evaluated using the same software by PM3 methods. The descriptors that have been used are-

(1) Heat of formation (ΔH_f).
(2) Molecular Weight (Mw).
(3) Total Energy (TE).
(4) Eigen value of HOMO (εHOMO).
(5) Eigen value of LUMO (εLUMO).
(6) Absolute Hardness(η).
(7) Electronegativity (χ).

The values of descriptors have been derived by solving the relevant equation given below:-

Parr et al.[20]defined electronegativity as the negative of chemical potential:

$$\chi = -\mu = -(\partial E/\partial N)v(r) \quad (1)$$

The absolute hardness,η,is defined as[21]

$$\eta = 1/2\,(\delta\mu/\delta N)_{v(r)} = 1/2\,(\delta^2 E/\delta N^2)_{v(r)} \quad (2)$$

Where E is the total energy, N the number of electrons of the chemical species, and v(r) the external potential.

The operational definition of absolute hardness and electronegativity[22] is defined as:

$$\eta = 1/2\,(IP-EA) \quad (3)$$

$$\chi = -\mu = 1/2\,(IP+EA) \quad (4)$$

Where IP and EA are the ionization potential and electron affinity respectively, of the chemical species. According to the Koopman's theorem, the IP is simply the eigen value of the HOMO with change of sign[23] and the EA is the eigen value of the LUMO with change of sign hence the equations 5 and 6 can be written as

$$\eta = 1/2\,(\varepsilon LUMO-\varepsilon HOMO) \quad (5)$$

$$\eta = 1/2\,(\varepsilon LUMO+\varepsilon HOMO) \quad (6)$$

The heat of formation is defined as:

$$\Delta H_f = E_{elect.} + E_{nuc.} + E_{isol.} + E_{atom} \quad (7)$$

where $E_{elect.}$ is the electronic energy, $E_{nuc.}$ is the nuclear-nuclear repulsion energy, E_{isol} is the energy required to strip all the valence electrons of all the atoms in the system, and E_{atom} is the total heat of atomization of all the atoms in the system.

Total energy of a molecular system is the sum of the total electronic energy, E_{ee} and the energy of internuclear repulsion, E_{nr}.

The total electronic energy of the system is given by[24]

$$E = 1/2\,P\,(H+F) \quad (8)$$

where P is the density matrix and H is the one-electron matrix

Finally a more general but important property of a molecular system is the molecular weight (Mw) which has been tested as descriptor.

Mode of Action of Protease [PR] Inhibitors: The X-rays study of protease inhibitor complexes has provided deeper insight into the mechanism of the protease inhibition. An extensive network of hydrogen bonds could be illustrated between the enzyme and the polar atom in the inhibitor. These postulated hydrogen bonds are formed primarily with the back bone atoms of the floor and flap regions of HIV protease.

RESULT AND DISCUSSION

The compounds of first set are the derivatives of urea isostere and of second set are the derivatives of a different isostere. The two sets have been

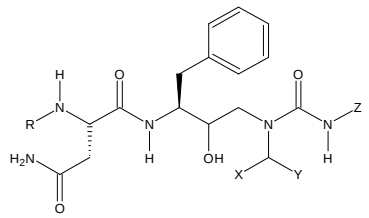

Figure 1 : Urea Isothere

TABLE 1 : First set: Derivatives of urea isostere Inhibitor and their observed activity in terms of inhibitory concentration (IC^{50}).

Comp. No.	Substitutients				IC_{50}
	R	X	Y	Z	
1	Cbz	H	$CHMe_2$	Me	5.82
2	Cbz	H	$CHMe_2$	n-Bu	6.03
3	Qua	H	$CHMe_2$	n-Bu	6.9
4	Cbz	H	$CHMe_2$	n-Pr	6.29
5	Cbz	H	$CHMe_2$	Et	6.48
6	Cbz	H	$CHMe_2$	i-Pr	6.59
7	Cbz	H	$CHMe_2$	t-Bu	7.46
8	Qua	H	$CHMe_2$	t-Bu	8.22
9	Cbz	H	CH_2CHMe_2	t-Bu	7.89
10	Qua	H	CH_2CHMe_2	t-Bu	8.52
11	Cbz	H	C_6H_{11}	t-Bu	7.54
12	Qua	H	C_6H_{11}	t-Bu	8.3
13	Cbz	H	C_6H_5	t-Bu	7.72
14	Qua	H	C_6H_5	t-Bu	8.52
15	Cbz	Me	C_6H_5	t-Bu	5.29
16	Cbz	H	4-Py	t-Bu	6.98
17	Qua	H	4-Py	t-Bu	7.72

Cbz = Carbobenzyloxy,
Qua = Quinolinyl-2-Carboxamide

studied separately.

First Set: The parent compound of this set is urea isostere and its structure is shown in figure 1. The derivatives are listed in TABLE 1. The inhibitors listed in this TABLE are protease (PR) inhibitors which are peptidic in nature. The biological activity of these inhibitors is reported in term of IC_{50}. The values of various descriptors of the derivatives listed in TABLE 1, have been evaluated and are included in TABLE 2. The quantities of descriptors in 37 combinations have been used for MLR analysis and for QSAR models. With the help of values of various descriptors, 37 MLR equations have been developed using different combinations of descriptors. Out of them only 13 QSAR models, presented below, have been found successful. The predicted activities of these models are included in TABLE 3 to 4.

Regression equation

PA1 = -0.00515022 ΔH_f -0.0516896 TE-0.873228 εHOMO-0.139844 χ-19.772
rCV^2 = 0.349993
r^2 = 0.71103

PA2 = -0.00515022 DH_f -0.0516896 TE-0.94315 εHOMO-0.0699221 εLUMO-19.772

TABLE 2 : The values of various descriptors of derivatives of first set.

Comp. No.	Heat of formation (kcal/mol) (Δ_f)	Molecular weight (Mw)	Total energy (TE)	HOMO energy (eV)? εHOMO	LUMO energy (eV)? εLUMO	Absolute hardness (η)	Electro negativity (Y)
01	-183.44	541.646	-300.706	-9.511	0.007	4.759	-4.752
02	-204.404	583.726	-322.205	-9.602	-0.068	4.767	-4.835
03	-173.765	620.747	-340.879	-9.559	-1.088	4.235	-5.324
04	-199.295	569.7	-315.041	-9.564	0.087	4.826	-4.738
05	-193.733	555.673	-307.875	-9.646	0.138	4.892	-4.754
06	-199.16	569.7	-315.039	-9.491	-0.1	4.696	-4.796
07	-203.15	583.726	-322.187	-9.565	0	4.782	-4.782
08	-165.812	620.747	-340.83	-9.248	-1.027	4.111	-5.137
09	-209.067	597.753	-329.349	-9.518	-0.238	4.64	-4.878
10	-206.142	623.791	-342.211	-9.394	-0.253	4.571	-4.823
11	-172.379	634.774	-347.988	-9.172	-1.148	4.012	-5.16
12	-169.077	660.812	-360.879	-9.438	-0.913	4.263	-5.176
13	-214.019	623.791	-342.245	-9.025	-0.137	4.444	-4.581
14	-121.387	654.764	-355.541	-9.406	-0.873	4.267	-5.14
15	-119.476	617.744	-336.538	-7.949	-0.111	3.919	-4.03
16	-152.069	618.731	-339.028	-9.492	-0.107	4.692	-4.8
17	-35.911	669.779	-364.62	-8.479	-1.456	3.512	-4.967

Full Paper

TABLE 3: Predicted biological activity PA1-PA9 of first set.

SN	IC_{50}	PA1	PA2	PA3	PA4	PA5	PA6	PA7	PA8	PA9
01	5.82	5.686	5.686	5.686	5.686	5.686	5.673	5.698	5.698	5.698
02	6.03	6.997	6.997	6.997	6.997	6.997	7.002	6.99	6.99	6.99
03	6.9	7.834	7.834	7.834	7.834	7.834	7.803	7.852	7.852	7.852
04	6.29	6.553	6.553	6.553	6.553	6.553	6.56	6.545	6.545	6.545
05	6.48	6.228	6.228	6.228	6.228	6.228	6.233	6.227	6.227	6.227
06	6.59	6.497	6.497	6.497	6.497	6.497	6.49	6.5	6.5	6.5
07	7.46	6.949	6.949	6.949	6.949	6.949	6.959	6.937	6.937	6.937
08	8.22	7.494	7.494	7.494	7.494	7.494	7.461	7.499	7.499	7.499
09	7.89	7.322	7.322	7.322	7.322	7.322	7.321	7.317	7.317	7.317
10	8.52	7.856	7.856	7.856	7.856	7.856	7.87	7.846	7.846	7.846
11	7.54	7.834	7.834	7.834	7.834	7.834	7.799	7.838	7.838	7.838
12	8.3	8.718	8.718	8.718	8.718	8.718	8.724	8.704	8.704	8.704
13	7.72	7.543	7.543	7.543	7.543	7.543	7.553	7.522	7.522	7.522
14	8.52	8.163	8.163	8.163	8.163	8.163	8.18	8.235	8.235	8.235
15	5.29	5.744	5.744	5.744	5.744	5.744	5.754	5.778	5.778	5.778
16	7.72	7.495	7.495	7.495	7.495	7.495	7.535	7.474	7.474	7.474
17	6.98	7.359	7.359	7.359	7.359	7.359	7.353	7.307	7.307	7.307

TABLE 4: Predicted biological activity PA10-PA13 of first set.

SN	IC_{50}	PA10	PA11	PA12	PA13
01	5.82	5.698	5.698	5.698	5.673
02	6.03	6.99	6.99	6.99	7.001
03	6.9	7.852	7.852	7.852	7.788
04	6.29	6.545	6.545	6.545	6.56
05	6.48	6.227	6.227	6.227	6.238
06	6.59	6.5	6.5	6.5	6.486
07	7.46	6.937	6.937	6.937	6.957
08	8.22	7.499	7.499	7.499	7.434
09	7.89	7.317	7.317	7.317	7.315
10	8.52	7.846	7.846	7.846	7.872
11	7.54	7.838	7.838	7.838	7.767
12	8.3	8.704	8.704	8.704	8.715
13	7.72	7.522	7.522	7.522	7.542
14	8.52	8.235	8.235	8.235	8.272
15	5.29	5.778	5.778	5.778	5.802
16	6.98	7.474	7.474	7.474	7.554
17	7.72	7.307	7.307	7.307	7.294

rCV^2=0.349993
r^2=0.71103

PA3=-0.00515022 DH_f-0.0516896 TE-1.01307 εHOMO-0.139844 η-19.772
rCV^2=0.349993
r^2= 0.71103

PA4=-0.00515022 ΔH_f-0.0516896 TE-1.01307 εLUMO+1.8863 η-19.772
rCV^2=0.349993
r^2=0.71103

PA5=-0.00515022 ΔH_f-0.0516896 TE +0.873228 εLUMO-1.8863 χ-19.772
rCV^2=0.349993
r^2=0.71103

PA6=-0.00482084 ΔH_f -0.0530345 TE -0.96664 εHOMO -20.3534
rCV^2=0.463038
r^2=0.710643

PA7=-0.00585388 ΔH_f +0.0254594 Mw-0.959549εHOMO -0.137502 εLUMO -18.291
rCV^2=0.31272
r^2=0.7084

PA8=-0.00585388 ΔH_f+0.0254594 Mw-1.09705 εHOMO -0.275004 ε-18.291
rCV^2=0.31272
r^2=0.7084

PA9=-0.00585388 ΔH_f+0.0254594 Mw-0.822047 εHOMO -0.275004 -18.291
rCV^2=0.31272
r^2=0.7084

PA10=-0.00585388 ΔH_f +0.0254594 Mw -1.09705 εLUMO +1.9191 η-18.291
rCV^2=0.31272
r^2=0.7084

PA11=-0.00585388 ΔH_f +0.0254594 Mw +0.822047εLUMO -

TABLE 5: Values of cross validation (rCV∧2) and correlation (r∧2) coefficients alongwith combinations of descriptors of first set.

SN	rCV^2	r^2	Combinations of descriptors
01	0.349993	0.71103	Δ_f , TE, εHOMO, χ
02	0.349993	0.71103	Δ_f , TE, εHOMO, εLUMO
03	0.349993	0.71103	Δ_f , TE, εHOMO, η
04	0.349993	0.71103	Δ_f , TE, εLUMO, η
05	0.349993	0.71103	Δ_f , TE, εLUMO, χ
06	0.463038	0.710643	Δ_f , TE, εHOMO
07	0.31272	0.7084	Δ_f , Mw, εHOMO, εLUMO
08	0.31272	0.7084	Δ_f , Mw, εHOMO, η
09	0.31272	0.7084	Δ_f , Mw, εHOMO, χ
10	0.31272	0.7084	Δ_f , Mw, εLUMO, η
11	0.31272	0.7084	Δ_f , Mw, εLUMO, χ
12	0.31272	0.7084	Δ_f , Mw, η, χ
13	0.40823	0.706834	Δ_f , Mw, εHOMO

1.9191χ-18.291
rCV^2=0.31272
r∧2=0.7084

PA12=-0.00585388 ΔH_f+0.0254594 Mw+0.822047η-1.09705 χ-18.291
rCV^2=0.31272
r^2=0.7084

PA13=-0.00521352 ΔH_f+0.0267677 Mw-1.00762εHOMO-19.3653
rCV^2=0.40823
r^2=0.706834

The coefficient values in decreasing order of merit and the combination of descriptors are presented in TABLE 5. QSAR model PA1 having combination of descriptors ΔH_f, TE, ε HOMO and χ provides the best model having high degree of predictivity

Second set

The parent compound is a different isostere derivative and is shown in figure-2. The derivatives whose biological activity is reported in terms of IC_{50} are listed in TABLE 6. The compounds of this series are also protease inhibitors, which are peptidic in nature. The peptidic inhibitors discussed in first set are structurally different with the inhibitor of this set hence presented separately. The values of various descriptors of this series are included in TABLE-7. The quantities of descriptors in 10 combinations have been used for MLR analysis and QSAR modeling. Out of them seven QSAR models have been found successful, which are presented as below

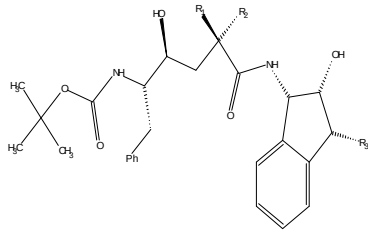

Figure 2 : Isostere

TABLE 6: Second Set: Derivatives of isostere Inhibitor and their observed activity in terms of inhibitory concentrations (IC^{50}).

Comp. no.	Substituents			IC_{50}
	R_1	R2	R3	
1	CH_2Ph	H	H	9.6
2	CH_2Ph	Me	H	8.11
3	CH_2CH_2Ph	H	OH	9.72
4	CH_2-4-CF_3Ph	H	H	9.59
5	(E)CH_2CH=$CHPh$	H	H	9.64
6	$CH_2C_6F_5$	H	H	9.22
7	CH_2-4-CH_3Ph	H	H	9.54
8	CH_2-4-NH_2Ph	H	H	9.51
9	CH_2-4-NO_2Ph	H	H	9.57
10	H	H	H	5.53
11	CH_2-4-$OHPh$	H	H	9.8
12	CH_2CH=CH_2	H	H	7.56
13	CH_2-4-IPh	H	H	9.14
14	$CH_2C(O)Ph$	H	H	8.27
15	CH_2-4-Pyridyl	H	H	9.28
16	CH_2SPh	H	H	9.6
17	CH_2-4-CMe_3Ph	H	H	9.77

Regression equation

PA1 = 0.0129237ΔH_f-0.0594284 TE-1.12447 εHOMO +2.22898 χ-6.38978
rC^2 = 0.543744
r^2 = 0.748351

PA2 = 0.0129237H_f-0.0594284 TE-0.00997822 εHOMO +1.11449 εLUMO-6.38978
rC^2 = 0.543744
r^2 = 0.748351

PA3 = 0.0129237ΔH_f-0.0594284 TE+1.10451 εHOMO+2.22898 η-6.38978
rCV^2 = 0.543744
r^2 = 0.748351

PA4 = 0.0129237ΔH_f-0.0594284 TE+1.10451 0199564 η-

Full Paper

TABLE 7: The values of various descriptors of derivatives of second set.

Comp No.	Heat of Formation (kcal/mol) (Δ_f)	Molecular weight (Mw)	Total energy (TE)	HOMO energy (eV) εHOMO	LUMO energy (eV) εLUMO	Absolute hardness (η)	Electro-negativity (χ)
01	-167.385	544.689	-291.913	-9.526	-0.004	4.761	-4.765
02	-160.352	558.716	-299.065	-9.444	0.144	4.794	-4.65
03	-174.085	558.716	-299.068	-9.439	0.002	4.721	-4.719
04	-325.559	612.688	-346.815	-9.533	-0.804	4.364	-5.169
05	-161.063	584.754	-311.579	-8.986	-0.165	4.41	-4.576
06	-381.389	634.642	-371.501	-9.53	-1.195	4.167	-5.363
07	-176.801	558.716	-299.102	-9.514	-0.04	4.737	-4.777
08	-170.107	559.704	-301.361	-8.789	-0.009	4.39	-4.399
09	-88.427	589.687	-323.59	-9.544	-1.997	3.774	-5.77
10	-192.706	454.565	-248.573	-9.53	0.090	4.81	-4.72
11	-212.313	560.689	-304.135	-9.252	0.024	4.638	-4.614
12	-175.299	494.63	-268.212	-9.502	0.092	4.797	-4.705
13	-147.381	670.585	-300.725	-9.116	-0.579	4.268	-4.847
14	-201.196	572.7	-309.382	-9.368	-0.631	4.368	-5.000
15	-160.591	545.677	-294.064	-9.601	-0.282	4.659	-4.941
16	-158.738	576.749	-301.103	-8.68	-0.017	4.332	-4.349
17	-190.665	600.797	-320.595	-9.425	-0.069	4.678	-4.747

TABLE 8: Predicted biological activity PA1-PA7 of second set.

SN	IC_{50}	PA1	PA2	PA3	PA4	PA5	PA6	PA7
01	9.6	8.886	8.886	8.886	8.886	8.886	8.883	8.685
02	8.11	9.566	9.566	9.566	9.566	9.566	9.563	9.289
03	9.72	9.23	9.23	9.23	9.23	9.23	9.228	9.044
04	9.59	9.212	9.212	9.212	9.212	9.212	9.212	9.218
05	9.64	9.951	9.951	9.951	9.951	9.951	9.953	10.03
06	9.22	9.522	9.522	9.522	9.522	9.522	9.522	9.602
07	9.54	9.151	9.151	9.151	9.151	9.151	9.148	8.934
08	9.51	9.398	9.398	9.398	9.398	9.398	9.403	9.659
09	9.57	9.568	9.568	9.568	9.568	9.568	9.567	9.713
10	5.53	6.088	6.088	6.088	6.088	6.088	6.089	6.27
11	9.8	9.06	9.06	9.06	9.06	9.06	9.06	9.043
12	7.56	7.482	7.482	7.482	7.482	7.482	7.481	7.478
13	9.14	9.022	9.022	9.022	9.022	9.022	9.025	9.231
14	8.27	8.786	8.786	8.786	8.786	8.786	8.786	8.88
15	9.28	8.792	8.792	8.792	8.792	8.792	8.789	8.617
16	9.6	9.52	9.52	9.52	9.52	9.52	9.526	9.839
17	9.77	10.23	10.23	10.23	10.23	10.23	10.21	9.924

6.38978
rCV^2 = 0.543744
r^2 = 0.748351

PA5 = 0.0129237ΔH_f-0.0594284 TE + 1.12447 εLUMO-0.0199564 χ 6.38978
rCV^2 = 0.543744
r^2 = 0.748351

PA6 = 0.0129011ΔH_f -0.0593604 E_T + 1.11166 εLUMO -6.28105
rCV^2 = 0.575431
r^2 = 0.748346

PA7 = 0.0106941ΔH_f-0.050971 TE + 1.424 χ + 2.38164
rCV^2 = 0.487759
r^2 = 0.71775

TABLE 9: Values of cross validation(rCV^2) and correlation(r^2) coefficients alongwith combinations of descriptors of second set.

SN	rCV^2	r^2	Descriptors Used
PA1	0.543744	0.748351	Δ_f , TE, εHOMO, η
PA2	0.543744	0.748351	Δ_f , TE, εHOMO, εLUMO
PA3	0.543744	0.748351	Δ_f , TE, εHOMO, η
PA4	0.543744	0.748351	Δ_f , TE, εLUMO, η
PA5	0.543744	0.748351	Δ_f , TE, εLUMO, χ
PA6	0.575431	0.748346	Δ_f, TE, εLUMO
PA7	0.487759	0.71775	Δ_f , E_T, χ

The coefficient values in decreasing order of merit and combination of descriptors are presented in TABLE 9. A reference to the TABLE indicates that the best results are provided by PA1-PA5.

Third set

Fourteen derivatives of urea isostere whose activity is not reported and method of synthesis could also not be obtained from literature have been studied. These derivatives are listed in TABLE-10. The quantities of descriptors of these derivatives are presented in TABLE 11. With the help of values of various descriptors, 12 MLR equations have been developed using different combination. The predicted activities of these models are included in TABLES 12 and 13.

TABLE 10: Third set: new derivatives of urea isostere inhibitors .

Comp. No.	Substituents			
	R	X	Y	Z
1	2,5-di-(OH)-4-I-C_6H_2.COO	H	C_6H_5	C_2H_5
2	FCH_2-$C_3H_4N_2$.COO	H	C_6H_5	C_2H_5
3	C_2H_5OOC-CH(OH)-CH(OH)-COO	H	C_6H_{11}	CH_3
4	(C_3H_7OOCNH-)(Cl).C_6H_3.COO	H	$(CH_3)_2CH$	C_2H_5
5	(CH_3OOCNH-)(SMe).C_6H_3.COO	H	$(CH_3)_2CH$	CH_3
6	2,5-di-Cl-C_6H_3-C_2N_2S.OOC	H	$(CH_3)_2CH$	H
7	4--F-C_6H_4.COO	H	$(CH_3)_2CH$	C_2H_5
8	CH_3-C_9H_5N.CO	H	$(CH_3)_2CH$	C_2H_5
9	$C_6H_5CS_2$	H	$(CH_3)_2CH$	C_2H_5
10	4-(OH)-3-CH_3-C_6H_3.COO	H	$(CH_3)_2CH$	CH_3
11	2-(OH)-C_6H_4.COO	H	$(CH_3)_2CH$	CH_3
12	CH_3CH(OH)-C_6H_4-CH(OH)-COO	H	$(CH_3)_2CH$	CH_3
13	2,5-di-Cl-C_6H_3-C_2N_2S.OCH_2	H	$(CH_3)_2CH$	C_3H_7
14	(C_6H_5) (CH_3OOC.CHMe.NH)-PO_3	H	C_6H_5	C_2H_5

Regression equation of new derivatives of urea isostere

PA1 = 0.0139659 Mw + 1.45972 εHOMO-1.10652 εLUMO + 14.0227

rCV^2 = 0.623394

r^2 = 0.812674

PA2 = 0.0139659 Mw + 0.353192 εHOMO-2.21305 η + 14.0227

rCV^2 = 0.623394

TABLE 11 : The values of various descriptors of third set.

Comp. No.	Heatof formation (kcal/mole)(Δ_f)	Molecular weight (Mw)	Total energy (Hartree)(TE)	HOMO energy (eV)?εHOMO	LUMO energy (eV) εLUMO	Absolute hardness (η)	Electro-negativity (χ)
01	-129.635	733.558	-348.518	-8.686	-0.735	3.976	-4.711
02	-99.63	599.661	-339.016	-8.655	-0.669	3.993	-4.662
03	-320.463	637.729	-367.07	-8.91	-0.878	4.016	-4.894
04	-118.354	677.196	-372.906	-8.554	-0.863	3.846	-4.708
05	-180.341	600.67	-339.727	-8.689	-0.685	4.002	-4.687
06	-114.069	668.551	-352.017	-8.646	-1.339	3.653	-4.993
07	-152.141	559.636	-316.461	-8.622	-0.815	3.903	-4.719
08	-93.038	590.721	-321.565	-8.755	-1.053	3.851	-4.904
09	-13.827	558.753	-285.484	-8.985	-2.117	3.434	-5.551
10	-157.932	557.645	-312.818	-8.577	-0.585	3.996	-4.581
11	-144.302	543.619	-305.631	-8.619	-0.627	3.996	-4.623
12	-196.355	601.698	-339.232	-8.761	-0.759	4.001	-4.76
13	-48.775	680.648	-350.902	-8.657	-1.205	3.726	-4.931
14	-289.151	677.733	-373.129	-8.699	-0.668	4.016	-4.684

Full Paper

TABLE-12: Predicted biological activity PA1-PA6 of third set.

SN	PA1	PA2	PA3	PA4	PA5	PA6
01	12.401	12.401	12.401	12.401	12.401	12.401
02	10.504	10.504	10.504	10.504	10.504	10.504
03	10.894	10.894	10.894	10.894	10.894	10.894
04	11.949	11.949	11.949	11.949	11.949	11.949
05	10.487	10.487	10.487	10.487	10.487	10.487
06	12.221	12.221	12.221	12.221	12.221	12.221
07	10.155	10.155	10.155	10.155	10.155	10.155
08	10.658	10.658	10.658	10.658	10.658	10.658
09	11.054	11.054	11.054	11.054	11.054	11.054
10	9.938	9.938	9.938	9.938	9.938	9.938
11	9.727	9.727	9.727	9.727	9.727	9.727
12	10.477	10.477	10.477	10.477	10.477	10.477
13	12.225	12.226	12.226	12.226	12.226	12.226
14	11.528	11.528	11.528	11.528	11.528	11.528

r^2 = 0.812674

PA3 = 0.0139659 Mw + 2.56624 εHOMO-2.21305 χ + 14.0227
rCV^2 = 0.623394
r^2 = 0.812674

PA4 = 0.0139659 Mw + 0.353192 εLUMO-2.91943 η + 14.0227
rCV^2 = 0.623394
r^2 = 0.812674

PA5 = 0.0139659 Mw-2.56624 εLUMO + 2.91943 χ + 14.0227
rCV^2 = 0.623394
r^2 = 0.812674

PA6 = 0.0139659 Mw-2.56624 η + 0.353192 ΔH_f + 14.0227
rCV^2 = 0.623394
r^2 = 0.812674

PA7 = 9.51792e-005 ΔH_f + 0.0139981 Mw + 1.45749 εHOMO - 1.09888 εLUMO + 13.9969
rCV^2 = 0.386387
r^2 = 0.812675

PA8 = 9.51792e-005 ΔH_f + 0.0139981 Mw + 0.358616 εHOMO - 2.19776 η + 13.9969
rCV^2 = 0.386387
r^2 = 0.812675

PA9 = 9.51792e-005 ΔH_f + 0.0139981 Mw + 2.55637 εHOMO - 2.19776 χ + 13.9969
rCV^2 = 0.386387
r^2 = 0.812675

PA10 = 9.51792e-005 ΔH_f + 0.0139981 Mw + 0.358616 εLUMO - 2.91499 η + 13.9969
rCV^2 = 0.386387
r^2 = 0.812675

PA11 = 9.51792e-005 ΔH_f + 0.0139981 Mw-2.55637 εLUMO + 2.91499 χ + 13.9969
rCV^2 = 0.386387
r^2 = 0.812675

TABLE 13: Predicted biological activity PA7-PA12 of third set.

SN	PA7	PA8	PA9	PA10	PA11	PA12
01	12.4	12.4	12.4	12.4	12.4	12.4
02	10.502	10.502	10.502	10.502	10.502	10.502
03	10.871	10.871	10.871	10.871	10.871	10.871
04	11.946	11.946	11.946	11.946	11.946	11.946
05	10.477	10.477	10.477	10.477	10.477	10.477
06	12.214	12.214	12.214	12.214	12.214	12.214
07	10.146	10.146	10.146	10.146	10.146	10.146
08	10.654	10.654	10.654	10.654	10.654	10.654
09	11.049	11.049	11.049	11.049	11.049	11.049
10	9.93	9.93	9.93	9.93	9.93	9.93
11	9.719	9.719	9.719	9.719	9.719	9.719
12	10.466	10.465	10.466	10.466	10.465	10.466
13	12.227	12.227	12.227	12.227	12.227	12.227
14	11.511	11.511	11.511	11.511	11.511	11.511

TABLE 14: Values of cross validation (rCV^2) and correlation (r^2) coefficients alongwith combinations of descriptors of third set.

SN	rCV^2	r^2	Combination of descriptors
PA7	0.386387	0.812675	Δ_f, Mw, εHOMO, εLUMO
PA8	0.386387	0.812675	Δ_f, Mw, εHOMO, η
PA9	0.386387	0.812675	Δ_f, Mw, εHOMO, χ
PA10	0.38639	0.81268	ΔMw, εLUMO, η
PA11	0.386387	0.812675	Δ_f, Mw, εLUMO, χ
PA12	0.386387	0.812675	Δ_f, Mw, η, χ
PA1	0.623394	0.812674	Mw, εHOMO, εLUMO
PA2	0.623394	0.812674	Mw, εHOMO, η
PA3	0.623394	0.812674	Mw, εHOMO, χ
PA4	0.623394	0.812674	Mw, εLUMO, η
PA5	0.623394	0.812674	Mw, εUMO, χ
PA6	0.623394	0.812674	Mw, η, χ

PA12 = 9.51792e-005 ΔH_f + 0.0139981 Mw -2.55637 η + 0.358616 χ + 13.9969
rCV^2 = 0.386387
r^2 = 0.812675

The coefficient values in decreasing order of merit and combination of descriptors are presented in TABLE 14; the best among them are models PA7-PA12.

CONCLUSION

The QSAR study of the derivative urea isostere

provides correlation coefficient values above 0.70, and of the other isostere derivative above 0.71. The combination of descriptors providing coefficient value in case of urea isostere derivatives are ΔH_f, εHOMO, χ and in isostere derivatives are ΔH_f, TE, εHOMO, χ.

On the basis of above QSAR models fourteen new derivatives of urea isostere have been proposed as HIV inhibitors. They have correlation coefficient value above 0.80.

REFERENCE

[1] S.Broder, R.Yarchoan, J.M.Collins; Lancet., **2**, 627 (**1985**).

[2] K.J.Lorentsen, C.W.Hendrix, J.M.Collins, D.M. Komhauser, B.G.Petty, R.W.Klecker, C.Flexner, R.H.Eckel, P.S.Lietman; Ann.Intern.Med., **111**, 561 (**1989**).

[3] J.Balzarini, J.Neyts, D.Schols, M.Hosoya, E.Van Damme, W.Peumans, E.De Clarcq; Antiviral.Res., **18**, 191, (**1992**).

[4] M.G.Rosamann; Proc.Natl.Acad.Science.U.S.A, **85**, 4625 (**1988**).

[5] E.De.Clercq; AIDS.Res.Human.Retrov., **8**, 119 (**1992**).

[6] J.F.Mayaux, T.Huet, C.Pernelle, J.Bcaquart, J.C.Gueguen, L.Tahraui, M.Evers, Y.Henin, A.Bousseau, N.Dereu; Abstract of the NIH Conference on Retroviral Integrase, Bethesda, MD, 19-20 (**1995**).

[7] C.A.Stein, Y.C.Cheng; Sci., **261**, 1004 (**1993**).

[8] N.E.Kohl, N.A.Emini, W.A.Schleif, L.J.Davis, J.C.Heimbach, R.A.F.Dixon, E.M.Scolnick, I.S.Sigal; Proc.Natl.Acad.Sci.,U.S.A. **85**,4686 (**1988**).

[9] C.Peng, B.K.Ho, T.WChang, N.T.Chang; J.Virol., **63**, 2550 (**1989**).

[10] A.Wlodawer, J.W.Ericson; Annu.Rev.Biochem., **62**, 543 (**1993**).

[11] K.Appelt; Perspect.Drug.Discovery.Des., **1**, 23 (**1993**).

[12] D.P.Getman, G.A. DeCrescenzo; etal J.Med.Chem., 36, 288 (**1993**). (**B**) M.K.Holloway, J.MWai et al; J.Med.Chem., **38**, 305, (**1995**).

[13] P.P.Singh, F.A.Pasha, H.K.Srivastava ; Indian Journal of Chem., **43B**, 983-991 (**2004**).

[14] P.P.Singh, F.A.Pasha, H.K.Srivastava; QSAR & Comb. Sci., **22(8)**, 843 (**2003**).

[15] P.P.Singh, H.K.Srivastava, F.A.Pasha; Bioorg.Med. Chem., **12(1)**, 171 (**2004**).

[16] H.K.Srivastava, F.A.Pasha, P.P.Singh ; Int.J.Quantum. Chem., **103(3)**, 237 (**2005**).

[17] F.A.Pasha, H.K.Srivastava, P.P.Singh; Mol.Div., **9(1-3)**, 215 (**2005**).

[18] F.A.Pasha, H.K.Srivastava, P.P.Singh; J.Quant.Chem., **104**, 87-100 (**2005**).

[19] F.A.Pasha, H.K.Srivastava, P.P.Singh; Bioorg and Med. Chem., (in press).

[20] R.P.Iczkowski, J.L.Margrave; J.Am.Chem.Soc., **83**, 3547 (**1961**).

[21] R.S.Mulliken; J.Chem.Phys., **2**, 782 (**1934**).

[22] R.G.Parr, R.G.Pearson; J.Am.Chem.Soc., **105**, 7512 (**1983**).

[23] R.G.Parr, R.A.Donnelly, M.Levy, W.E.Palke; J.Chem. Phys., **68** (**1978**).

[24] R.G.Parr, L.V.Szentpaly, S.Liu; J.Am.Chem.Soc., **121**, 1922 (**1999**).