

QSAR studies on quinolonyl diketo acid derivatives

S. Dubey*, N. K. Abbas, G. Goutham and P. A. Bhosle

K.L.E.University's College of Pharmacy, Rajajinagar II-Block, Bangalore, India

*Corresponding author: drsonaldubey@gmail.com

ABSTRACT A 3D-quantitative structure activity relationship study on some newly synthesized quinolonyl diketo acid derivatives possessing integrase inhibitory, antiviral and cytotoxicity activity was done. The result presented herein provides QSAR models developed using MLRA, PCA, PCR and PLS analysis. The QSAR models obtained have shown significant correlations in terms of r^2 , q^2 , s^2 and F-values. The models have been validated internally and externally; and shown good predictive ability. Electrostatic descriptors have come out to be the important ones regulating the activity of bifunctional quinolonyl diketo acid derivatives.

Keywords: Quinoline, QSAR, MLRA, PLS, PCA, PCR, integrase inhibition.

Introduction

It has been nearly forty years since the quantitative structure-activity relationship (QSAR) paradigm first found its way into the practice of pharmaceutical chemistry, toxicology, and eventually most facets of chemistry.¹ Its staying power may be attributed to the strength of its initial postulate that activity was a function of structure as described by electronic attributes, hydrophobicity and steric properties as well as the rapid and extensive development in methodologies and computational techniques that have ensued to delineate and refine the many variables and approaches that define the paradigm. The overall goals of quantitative structure-activity relationship retain their original essence and remain focused on the predictive ability of the approach and its receptiveness to mechanistic interpretation.

Rigorous analysis and fine-tuning of independent variables has led to an expansion in development of molecular and atom-based descriptors, as well as descriptors derived from quantum chemical calculations and spectroscopy.² The improvement in high-throughput screening procedures allows for rapid screening of large numbers of compounds under similar test conditions and thus minimizes the risk of combining variable test data from many sources. The formulation of thousands of

equations using quantitative structure-activity relationship methodology attests to a validation of its concepts and its utility in the elucidation of the mechanism of action of drugs at the molecular level and a more complete understanding of physicochemical phenomena such as hydrophobicity. It is now possible not only to develop a model for a system but also to compare models from a biological database and to draw analogies with models from a physical organic database.³

Acquired immunodeficiency syndrome (AIDS) is the consequence of the infection with human immunodeficiency Virus type-1 (HIV-1).⁴ The anti HIV drugs available, either used to inhibit the reverse transcriptase or the protease enzyme activity. Yet complete resistance was not produced. Hence a new series of drugs were synthesized which produced effective integrase enzyme inhibition, thus producing an effective chemotherapeutic agent.

We describe the application of particle swarms for the development of quantitative structure-activity relationship (QSAR) models based on k-nearest neighbor MLRA, PCA, PCR and PLS analysis. Particle swarms is a population-based stochastic search method based on the principles of social interaction. Each individual explores the feature space guided by its previous success and that of its neighbors. Success

is measured using leave-one-out (LOO) cross validation on the resulting model as determined by k-nearest neighbor kernel regression. The technique is shown to

compare favorably to simulated annealing using three classical data sets from the QSAR literature.

Experimental

Some novel bifunctional quinolonyl diketo acid derivatives as HIV-1 integrase inhibitors were selected.^{5,6,7} Then 3-D structures of reasonable conformations were generated from 2D structures.

QSAR studies were performed on 25 novel bifunctional quinolonyl diketoacid derivatives as HIV-1 integrase inhibitors having strand transfer inhibitor activity, 3'-processing inhibitor activity.

Structure of compounds was graphically drawn on the monitor using chemdraw. With the help of CHEM 3D, the different 2D structures are converted to 3D structures. Dragon software provides molecular descriptors but does not perform QSAR analysis. So various descriptors are calcula-

ted with the help of dragon[®]. It has calculated 1664 descriptors each compound. The nature of the descriptors has already been mentioned in the introduction portion.

The various 3D structure files which is saved in MDL. mol files format are provided as input file for dragon. Then DRAGON calculates all the molecular descriptors for that particular 3D compound.

A large number of molecular descriptors were calculated and QSAR was done with the help of **codessa[®]** and **V-Life[®]** software. A set of descriptors found to be irrelevant were discarded from the study. The regression analysis was carried out using heuristic analysis. The PCA, PLS and PCR was done.

Results and discussion

The various novel bifunctional Quinolonyl diketo acid derivatives possessing anti HIV activity were subjected to QSAR analysis. Using Codessa software the QSAR models were developed for 4 sets of activities namely strand transfer inhibitor activity, 3'-processing inhibitor activity, antiviral activity, and cytotoxicity. For each activity we had divided the total compounds into a training set: test set which has number of compounds 24:1, 13:2 for strand transfer inhibitor activity, 3'-processing inhibitor activity. The total of about 1897 different descriptors physicochemical, topological, geometrical, constitutional, 2D and 3D descriptors were calculated by the softwares, Chem 3D, Dragon, Codessa and V-life software. The numbers of descriptors were reduced to final set of 97, by rationally stepwise elimination of lesser significant ones based on their one parameter and

multiparameter r^2 values. In subset only those descriptors were chosen which were having $r^2 \geq 0.1$. The heuristic method was used as final statistical tool for developing QSAR relations using the software Codessa.

Codessa software

Strand transfer inhibitor activity:

The best fit QSAR model found in the case of strand transfer inhibitor activity:
 Activity= -120.98+21.755 *RDF135v -
 38.342*RDF145v+15.71*RDF145m-
 0.936*Ts+71.85*Belm²

r^2	q^2	S^2	F-value	p<
0.9623	0.598	14.53	81.7	0.005

The activity shows positive linear dependence on Constitutional descriptors (RBF); RDF descriptor (RDF135v, RDF145m). Eigen value (BELm², BELv²), WHIM descriptors (L1s,L1m).; RDF

descriptors such as (RDF145e,RDF145v), have a negative effect on activity, and r^2 ranging from 0.9641-0.9542.

3' processing inhibitor

The best fit QSAR model found in the case of 3' processing inhibitor activity:

$$\text{Activity} = 943.16 + 13.822 * \text{RDF155m} - 11.19 * \text{MWCO6} + 76.51 * n_{\text{Cp}}$$

r^2	q^2	S^2	F-value	p<
0.9638	0.8917	274.08	130.8	0.005

The activity shows positive linear dependence on Connectivity index descriptors (X1Av,X0Av). Constitutional descriptors (RBF,RBF) Morse descriptor (Mor26e), RDF descriptor (RDF155m.Topological descriptor (PW2), Walk and path count (MWCO6). Morse descriptor (Mor7e,Mor7u); have a negative effect on activity, and r^2 ranging from 0.9894-0.98

Antiviral activity

The best fit QSAR model found in the case of antiviral activity

$$\text{Activity} = 8.8057 + 0.16712 * G(\text{N..Cl}) - 21.69 * \text{SPAM}$$

r^2	q^2	S^2	F-value	p<
0.9841	0.9650	.0393	217.06	0.005

Cytotoxicity

The best fit QSAR model found in the case of cytotoxicity activity

$$\text{Activity} = -11.56 + 4.523 * \text{Mor7v} - 0.0215 * \text{Mor1e}$$

r^2	q^2	S^2	F-value	p<
0.8571	0.4295	0.3706	13.51	0.025

A graphical representation of the calculated Vs experimental activity for all the three integrase inhibitory activity; and cytotoxic and antiviral activity is shown in figure 1-4.

Principal Component Analysis

As per PCA analysis the Getaway descriptors namely HATS 1v, HATS1p; RDF descriptor namely RDF145e, RDF140e WHIM descriptors namely L1s,L1m,L1s.; RDF descriptors such as RDF145m,topological descriptors like HyDp have a positive effect on *strand transfer activity*. RDF descriptor namely RDF135v, RDF140p,RDF115v; Connectivity index descriptors X1A; Eigen value BEHp2, BELe6 have a negative effect on activity (figure 5).

The connectivity index descriptors X1Av, Morse descriptor Mor7e , Topological descriptor PW2, RDF descriptor namely RDF85m, RDF135u, RDF135m, Walk and path count MWCO6, have positive effect on 3'Processing inhibitory activity. Connectivity index descriptors X4A, X1Av,X0Av; Morse descriptor Mor26e, Mor30m Mor30u; Topological descriptor T12 ; Constitutional descriptors RBN,RBF, have a negative effect on activity(figure 6).

As per the PCA analysis geometric descriptors QYYv, Conectivity index such as X1A, Getaway descriptors HATSp ,HATS2v; Eigen value BEHm2, RDF descriptors such as RDF95v, RDF135e, RDF155v, RDF135u, RDF35m, R5u ; Morse descriptor Mor20p, Mor26m have positive effect on *antiviral activity*. Connectivity index such as Xindex,; Information index as IC4,IC5, ; have a negative effect on activity (figure7).

As per the PCA analysis Morse descriptor Mor21m; geometric descriptors T (O.Cl), G(O.Cl). have negative effect on *cytotoxic activity*. Getaway descriptors Hats7m, WHIM descriptors namely G3e, G3u,G3m, have a have positive effect on activity (figure 8).

V-life software:

Partial least square regression (PLSR) and Principal component regression (PCR) of new chemical entities were generated using V-life molecular design suite software, version 3.0, The result of the study for each activity is given in Table 1 and 2.

TABLE 1: QSAR equations after doing PCR i.e. principal component regression of novel bifunctional quinolonyl diketo acid derivatives using V-LIFE SOFTWARE.

ACTIVITY	QSAR EQUATION	r ²	q ²	F-value	p<
STRAND TRANSFER	Activity=-0.4670+0.5319*T_C_N_2+0.0246*Qudrapole2-0.7936*SsschOcount	0.7328	0.6524	24.68	0.005
3'P	Activity=0.05566-1.1422*SsschO count	0.5796	-0.256	13.38	0.005
ANTIVIRAL	Activity=-3.7251+1.4809*Chlorine count-27.3891*SAMostHydrophilic	0.9117	0.8433	82.58	0.005
CYTOTOXICITY	Activity=4.2502+0.02*-vePotential Surface area	0.8294	0.5110	14.58	0.005

TABLE 2: QSAR equations after doing PLS i.e. partial least square of novel bifunctional quinolonyl diketo acid derivatives using V-LIFE SOFTWARE

ACTIVITY	QSAR EQUATION	r ²	q ²	F-value	p<
STRAND TRANSFER	Activity=0.2167+0.5424*T_C_N_2+0.0361*Qudrapole2-0.1908*XAMostHydrphobicHydrophilic distance	0.727	0.5808	50.73	0.005
3'P	Activity=-1.633-1.2873*SssCHOcount+1.3146chaincount-0.3114*most+ve&-vepotential distance	0.824	0.6388	21.16	0.005
ANTIVIRAL	Activity=-0.3671+1.9156*chlorinecount-0.4734*SssNHE-index	0.944	0.8976	136.5	0.005
CYTOTOXiCITY	Activity=-5.5675+0.0187*-vepotential surfacearea+1.7081*SKmosthydrophobic	0.967	-0.185	90.19	0.005

Conclusion

In our present study, QSAR models have been developed that are based on molecular, sub molecular, physicochemical, 2D and 3D properties obtained from the various softwares. In contrast to using the entire matrix of structural parameters as in PCA analysis an automatic forward inclusion MLR technique was used to arrive at the primary property determinants with minimal number of independent variables or descriptors.

From the result of Codessa Software it can be suggested that the WHIM, Topological, RDF, 2D autocorrelations, Electrostatic, Getaway and 3D Morse descriptors are statistically significant. Correlations ranging between 0.8705-0.9931 were found in this study which is highly significant. These models hold good predictive performance with q² values ranging between 0.7201-0.9624 which was calculated by using LOO method.

Constitutional descriptors and thermodynamic descriptors have minor contribution towards the improvement in relationship. The final mathematical models furnish good correlations, LOO correlations, residuals and overall statistical significance.

Further validation of the QSAR models using external validation (test set of six compounds) has shown good results the descriptors which were coming significant from PCA were similar to those coming significant from heuristics (MLRA), method. i.e. using both the studies almost the same kind of descriptors were showing the positive and negative dependence on the activity. The log transformed values of activity are giving better results because biological data are generally found to be skewed, the log transformation moves the data to a nearly normal distribution.

From the PLS and PCR analysis by V-life software the N count, the positive

potential surface area, the O count, and the Cl count and hydrophilicity are significant.

In both the studies i.e. using Codessa (MLRA & PCA) and V-life software, the

electrostatic descriptors have come out to be the important ones regulating the activity of novel bi-functional quinolonyl diketo acid derivatives.

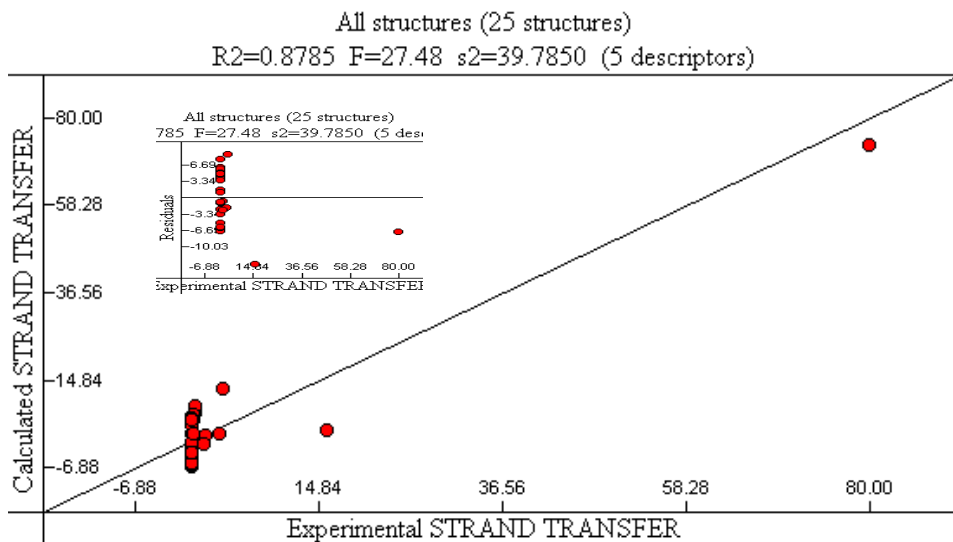


Figure 1: The regression plot of strand transfer inhibitor activity, inset is a residual plot.

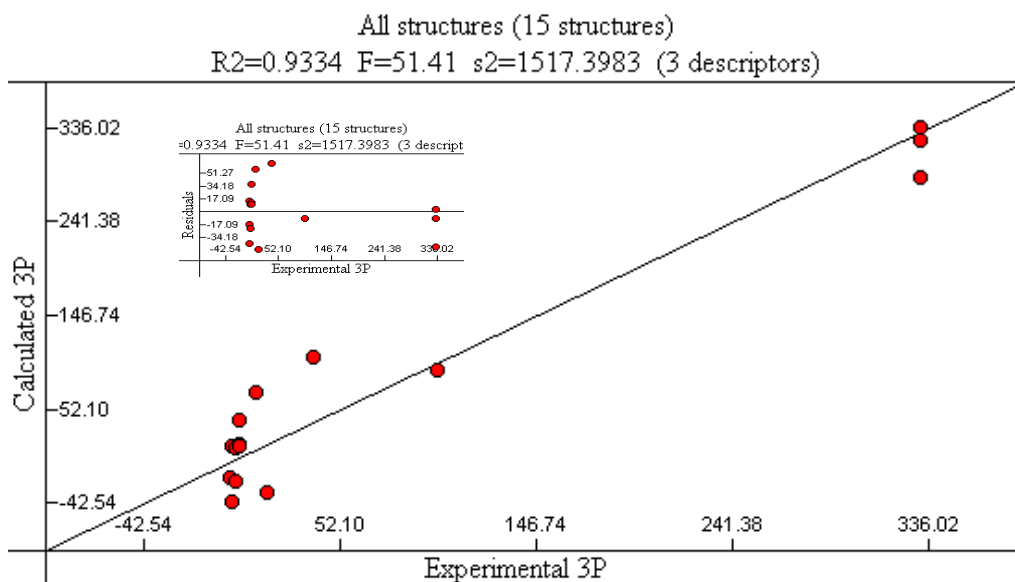


Figure 2: The regression plot of 3'Processing inhibitor activity, inset is a residual plot.

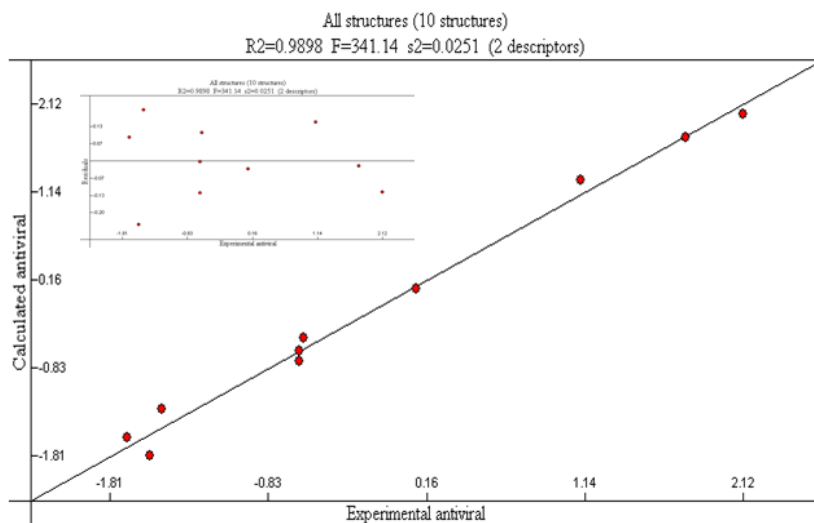


Figure 3: The regression plot of antiviral activity, inset is a residual plot.

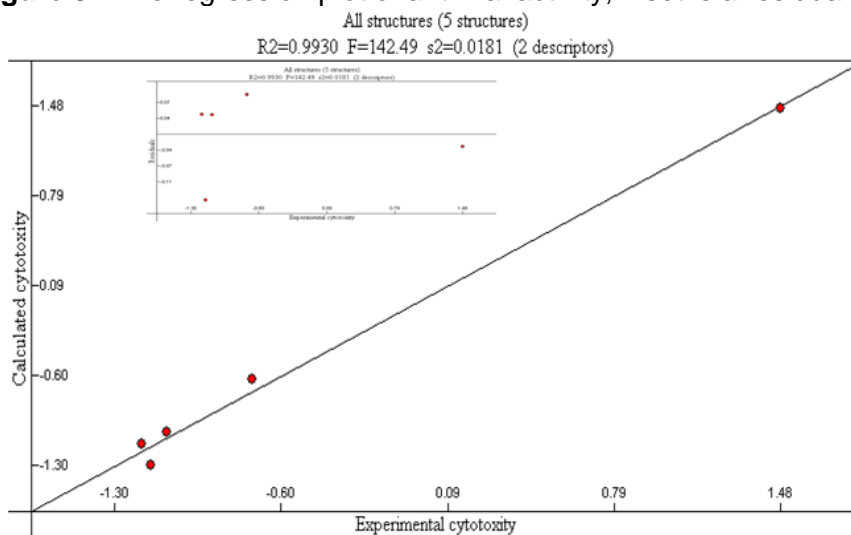


Figure 4: The regression plot of cytotoxicity, inset is a residual plot

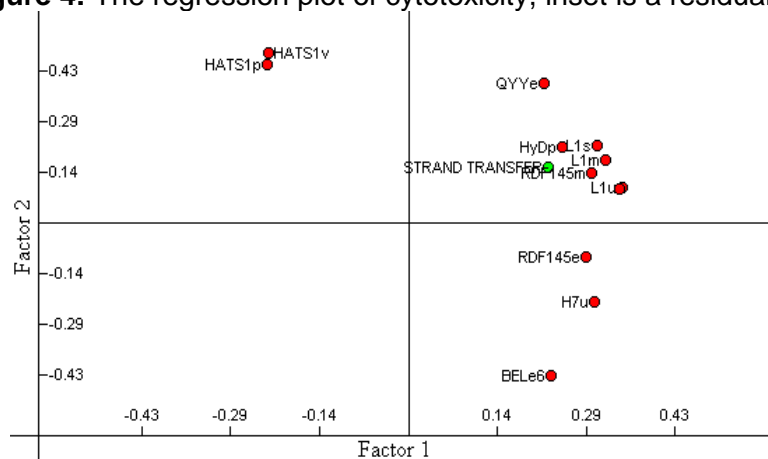


Figure 5 : PCA plot of strand transfer inhibitor activity

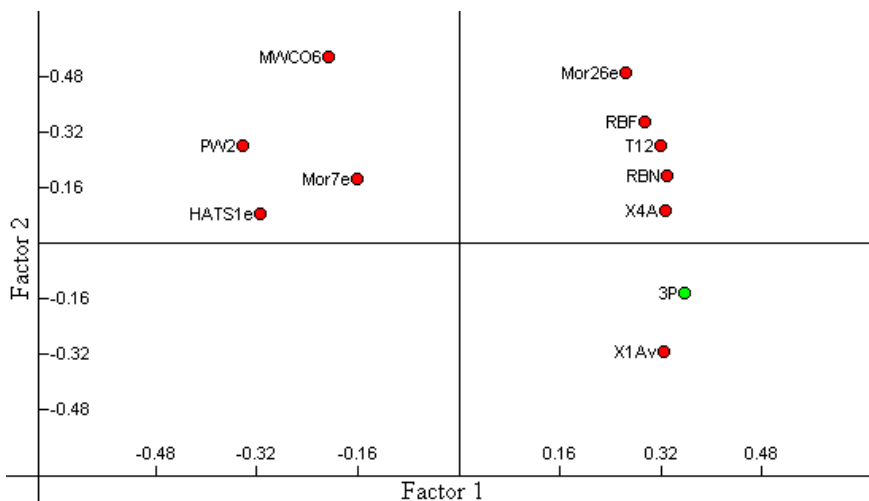


Figure 6 : PCA plot of 3'-processing inhibitor activity

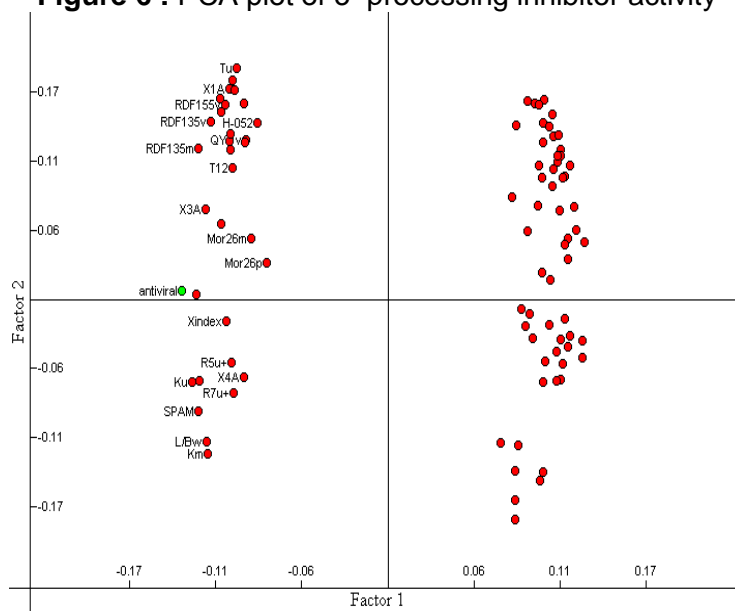


Figure7: PCA plot of antiviral activity.

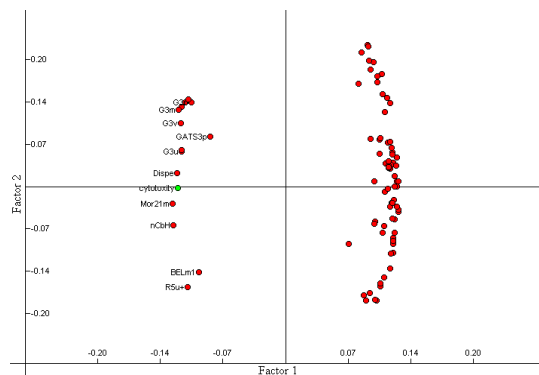


Figure 8: PCA plot of cytotoxicity activity

References

1. Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*, John Wiley & Sons, New York, 1979.
2. Livingstone, D.J. *J. Chem. Inf. Comput. Sci.* 2000, *40*, 195-99.
3. Hansch, C.; Kurup, A.; Garg, R.; Gao, H. *Chem. Rev.* 2001, *101*, 619-24.
4. Mehanna, A. S. Rationale of Design of anti-HIV Drugs. In Burger's Medicinal Chemistry and Drug Discovery, 6th ed., VOL 5, Burger A, Abraham, DJ, Andrako J, Block JH, Borchardt RT, Abraham Ed, A John Wiley and Sons, Inc., Publication: New Jersey, 2003; 457-84.
5. Santo, R. D.; Costi, R; Roux, A.; Artico, M. ; Lavecchia, A. ;Marinelli, L.; Novellino, E. ; Palmisano, L. ;Andreotti, M.; Amici, R.. ; Galluzzo, C. M.; Nencioni, L. Palamara, A.; Pommier, Y.; Marchand, C. Novel Bifunctional Quinolonyl Diketo acid derivatives as HIV-1 Integrase Inhibitors: Design, Synthesis, Biological activities, and Mechanism of action. *J. Med Chem.* 2006,*49*, 1939-194.
6. Bona, R.; Anderotti, M.; Buffa, V. ; Leone, P. ; Galluzzo,C.M. ; Amici, R. ; Paimisano, L. ; Mancini, G. ; Michelini, Z. ; Santo, R. D. ; Costi, R. ; Roux, A. ; Ppmmier, Y.; Marchand, C. ; Vella, S.; and Cara, A. Development of a Human Immunodeficiency Virus Vector –based single cycle Assay for evaluation of anti integrase compounds. *Antimicrobial agents and chemotherapy.*2006. *Vol 50 (10)*.3407-3417.
7. Sato, M.; Motumura, T.; Aramiki H.; Matusda, T; Yamashita. M.; Watanabe, W.; Yamataka, K.; Ikeda, S.; Kodama, E. ; Matsuka, M. ; and Shinkai, H. ; Novel HIV-1 Integrase inhibitors derived from quinolone antibiotics. *J. Med Chem.* 2006, *49*, 1506-1508.