Egyptian Journal of Chemistry

In Silico QSAR Studies on N-Aryl-Oxazolidinone-5-carboxamides as Novel HIV-1 Protease Inhibitors

Saloni Mishra[a], Arvind Tripathi[b]

[a]*Department of Chemistry, Feroze Gandhi College, Raebareli, UP 229001, India*

[a]*Data Science Department, Saint Peters' University, 2641 John F. Kennedy Blvd, Jersey City, NJ 07306, US*

[b]*Department of Statistics, UAB, 1720 University Blvd, Birmingham, AL 35294 US*

Abstract

From 2006 to 2009, the number of new cases of human immunodeficiency virus type 1 (HIV-1) in the United States increased. Many have questioned whether this is due to medication response, therapy, viral mutation, infection rates, or a lack of improved knowledge on virus protection. Regardless of the causes, new methods of prevention, such as education against contracting the virus, and therapy to make persons afflicted with the virus suffer from a manageable chronic condition are required. Because protease is required for the life cycle of HIV virions, protease inhibitors have been the most widely used anti-HIV drugs. However, because of the virus's ability to evolve at such an alarming rate, many researchers and physicians remain unconvinced and skeptical that antiretroviral medications will stay effective against it. Therefore, we developed topomer CoMFA (Comparative Molecular Field Analysis) and HQSAR (Hologram Quantitative Structure Activity Relationship) models on a series of N-aryl-oxazolidinone-5-carboxamides which is HIV-1 protease inhibitors. The developed models show good statistics in terms of q^2 and r^2 values. The best predictions obtained with topomer CoMFA model (r^2 = 0.967, q^2 = 0.874) and HQSAR model (r^2 = 0.973, q^2 = 0.902). The HQSAR model was fitted to predict the biological activity using atom, bond, connections and chirality as parameters and fragment size (6-9). The validity of developed models was confirmed by test set and good predictive correlation coefficient for topomer CoMFA (0.9576) and HQSAR (0.913) is found. The contour plot of topomer CoMFA suggests that at phenylsulfonamide moiety (R1 fragment) sterically bulky and electropositive groups are favored by the model for enhancing the activity. In addition, the atomic contribution map of HQSAR suggests that 4-methoxy, dioxolane and amino groups at phenylsulfonamide moiety are preferred and, moreover, 3 and 4 acetyl groups at phenyloxazolidinone moiety are favored.

Keywords: Topomer CoMFA; HQSAR, HIV-1; Protease Inhibitors; N-Aryl-Oxazolidinone-5-carboamides

_____

1. Introduction

HIV-1 protease is a retroviral aspartyl protease that is required for the life cycle of HIV, the retrovirus that causes AIDS [1, 2].HIV virions remain un-infectious in the absence of a functional HIV-1 protease [3, 4].Consequently, altering the active site of HIV protease or inhibiting its activity prevents HIV from replicating and infecting new cells, [5] making HIV-1 protease inhibition the subject of much pharmaceutical research [6]. In fact,

HIV-1 protease inhibitors are the most potent anti-AIDS drugs reported to date and are crucial components of highly active antiretroviral therapy (HAART) [7, 8]. In the past few decades, the continuous efforts of structure based drug design and advanced clinical trials led to the discovery of some FDA approved drugs for HIV-1 protease including, saquinavir [9], indinavir [10], ritonavir [11], amprenavir [12], lopinavir [13], and darunavir (TMC-114) [14], and these all are competitive inhibitors that bind in the active site of the enzyme of HIV-1. The development and clinical introduction of anti-AIDS HIV-1 protease inhibitors is regarded as a significant

_____

achievement of structure-based drug design [15].However, the emergence of HIV-1 mutants that are resistant to current drug regimens is a critical factor in the clinical failure of antiviral therapy [16]. The advent of multidrug-resistant (MDR) protease variants poses a significant challenge to the efficacy of the majority of currently licensed protease inhibitors [17, 18]. Development of next-generation HIV-1 protease inhibitors active against MDR virus has been the focus of intense research efforts in recent years [14, 19]. Consequently, there is a growing need for the discovery of novel classes of protease inhibitors with acceptable pharmacological profiles and decreased susceptibility to drug resistance, with an emphasis on broad spectrum action against MDRvariants [16].Designing inhibitors that interact with the same residues could be one technique for reducing the likelihood of drug resistance of HIV-1 protease that are necessary to recognize the substrate [20, 21]. For this purpose, N-phenyloxazolidinone-5-carboxamides (figure 1) into the (hydroxyethylamino) sulfonamide scaffold as P2 ligands [16] could be regarded a viable option for additional computational research in order to obtain a better drug that can overcome the limitations outlined.Crystal structures of inhibitors incorporating phenyloxazolidinone-based P2 ligands bound to wild-type HIV-1 protease revealed that the orientation of the oxazolidinone moiety allows it to make a complex network of hydrogen bonds with invariant Asp29 residue of the protease [16].

Computer-Aided Drug Design (CADD) has evolved as an efficient method of discovering prospective lead compounds and assisting in the creation of new drugs for a wide range of ailments [22]. CADD can be structure- or ligand-based, with the former exploring and learning the target protein structure to quantify the interaction levels of all compounds under consideration and the latter using chemical similarity criteria and predictive quantitative structure-activity relationship (QSAR) models from the molecules to identify the known actives and inactives [23, 24]. QSAR analysis had been used in several research on a wide variety of compounds. Using 3D QSAR, Mernissi et al. investigated 30 1,3,5-triazine derivatives for their activity against cancer cell lines. Models had been developed using the Comparative Molecular Field Analysis (CoMFA) and the Comparative Molecular

Similarity Indices (CoMSIA) to analyze these molecules [25]. Abdulrahman et. al. did QSAR study on benzene sulphonamide derivatives which had been synthesized and evaluated as anti-oxidants and predicted new compounds that have effectiveness as antioxidants [26]. Adeniji also used multivariate QSAR model to correlate the chemical structures of the 1,2,4-triazole analogues as anti-tubercular agent with their observed activities [27]. Hassan et. al. used 2D-QSAR to study of some new antiproliferative pyrazoles and pyrazolopyridines as potential CDK2 inhibitors [28].

Similarly, many computational studies have been performed on N-aryl-oxazolidinone-5-carboxamides as anti-HIV agent. Earlier, the validated predictive quantitative structure activity relationship (QSAR) modeling was performed, which shows the importance of atom-based descriptors like RTSA indices, Wang–Ford charges and different whole molecular descriptors [29]. Another study on the N-aryl-oxazolidinone-5-carboxamides has also been done using quantum chemical and topological fingerprints descriptors [30]. Using molecular modeling, Mukesh Sharma [31] inferred that the SssOE index, SsNH2 count, SsOH count, SsFE index, and highly occupied molecular orbital energy are significant for N-aryl-oxazolidinone-5-carboxamide derivatives to have anti-HIV protease activity and TNF4, TTF2 are alignment-independent descriptors that boost anti-HIV-1 protease activity. The 3D QSAR tests demonstrated steric descriptors with negative coefficients near the R2 and R3 regions of the ring, indicating that fewer bulky groups impair oxazolidinone-5-carboxamide anti-HIV efficacy.

Through similar approaches, here we are exploring N-aryl-oxazolidinone-5-carboxamides for anti-HIV-1 activity. The aim of the study is twofold: first, to enhance the biological activity by elucidating the structural basis secondly, providing the guidelines in designing the more potent compounds for the anti-HIV-1 activity of N-aryl-oxazolidinone-5-carboxamides. In the present study, we are exploring the descriptors favored by the HQSAR model and looking at the steric and electrostatic field contribution by performing the 3D QSAR study by analyzing the topomer CoMFA model.

_____

Egypt. J. Chem. **66**, No.SI 13 (2023)

_____

2. Experimental

2.1. Data Set

The dataset of thirty-eight HIV-1 protease inhibitors, used for the topomer CoMFA and HQSAR analysis were selected from the literature [16]. The values of biological activity (Ki) in dataset vary from 0.0008nM to 188.8nM. The Ki values were converted into the corresponding pKi (-logKi) values for modeling purposes and used as a dependent variable in both the analysis. The training set (thirty compounds) and test set (eight compounds) were selected in such a way that the structurally diverse molecules possessing wide range of activities included in both the sets and are shown in Table 1. The training set was used to derive the model whereas the test set was used to validate the model. The generation of the molecular structures, analysis, calculations, and visualizations were performed using the Sybyl-X 1.2(Tripos Ltd.) [32].

2.2. Topomer CoMFA

A new methodology, topomer CoMFA, has been introduced by R. Cramer to overcome the challenges of the alignment problem of CoMFA [34]. This method merges CoMFA [33] and topomer technology yielding effective results[34].In topomer CoMFA , with respect to rotation and translation, an invariant three-dimensional (3D) representation of a molecular component that is derived from its two-dimensional (2D) topology by topomer alignment is referred to as a topomer [35]. Structural fragments contain a common feature that is the open valence or common attachment bond. In the present topomer CoMFA analysis, all molecules of training set were divided into two fragments and one common core, as R1 (blue), R2 (green) and common core (black) groups shown in Figure 1.
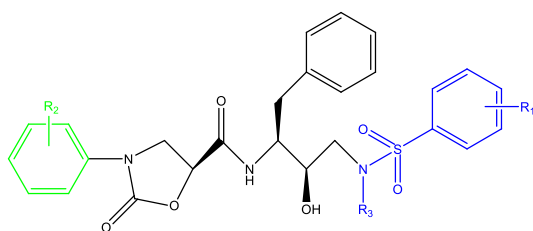

Figure 1- Fragmentation pattern (R1 and R2) for all moleculesof dataset in topomer CoMFA analysis. R1 fragment is represented by the blue color and R2 fragment is denoted by the green color.

Each fragment of a topomer was utilized in conjunction with topomer alignment to create a 3D

representation that was invariant[36]. In addition, each fragment's contribution to the overall activity is also determined by topomer CoMFA. Using the results of a leave-one-out (LOO) cross validation analysis, the optimal number of components (ONC) is used to determine the correlation coefficient r^2. At a regular space grid of 2, steric and electrostatic fields are computed and automatically fixed into a 1000-point cube that is optimally positioned to contain a topomer.When calculating the field value at a lattice point, an atom's steric or electrostatic contribution is multiplied by 0.85n, where n is the total number of acyclic single bonds that can be found between that atom and the open attachment bond [34].

2.3. HQSAR

In this work, we have also explored the two-dimensional molecular features related to the biological activity of a series of the N-Aryl-Oxazolidinone-5-carboxamides using molecular hologram Quantitative Structure Activity Relationship (HQSAR) methodology [37, 38]. This is a powerful technique that is based on fragment-based strategy in the drug designing. This methodology uses the strategy to translate chemical structures into binary bit strings, known as fingerprints. HQSAR uses molecular hologram [39, 40] (an extended form of fingerprint), which encodes more information (e.g., branched, and cyclic fragments, stereochemistry) than the traditional 2D fingerprint.With the fingerprint method, molecular structures are broken down into their smallest possible linear, branching, and overlapping components. A cyclic redundancy check (CRC) algorithm [41] is then used to assign a unique integer to each of these fragments.These integers are then hashed to a bin in an integer array of fixed length range in 50 to 50029. These arrays are known as molecular hologram and the bin occupancies of the molecular holograms are used as the descriptors in statistical analyses [40]. The descriptors (molecular holograms) are expected to encode the chemical and topological information of molecules. As a result, a molecular hologram is presented as a string of integers [39].

3. Result & Discussion

3.1. Topomer CoMFA Analysis

In the present study, 3D-QSAR topomer CoMFA model has been derived for N-aryl-Oxazolidinone-5-

_____

_____

carboxamides and derivatives against HIV-1 activity of training set with good predictivity in terms of q^2 (Leave-one-out cross validation) and r^2 (correlation coefficient) values. The model displayed $q^2 = 0.874$ and $r^2 = 0.967$ with standard error of estimates s=0.33 and optimum number of components was eight for the best LOO cross validation. This model was validated by a test set of eight compounds which were excluded during the model creation of topomer CoMFA giving satisfactory r^2=0.9576. Table 2 provides the statistical summary of the model. The experimental and predicted values of activity of the series of N-aryl-Oxazolidinone-5-carboxamides is given in Table 1 and fragment contribution of R1 and R2 fragments are given in Table 3. The residual values are also given in Table 1. The relationship between the experimental and predicted values of activities of both training and test set were fairly determined as the residual values are less than 0.69 log unit. The graph of predicted versus experimental values of activity for training and test set of topomer CoMFA analysis is given in figure 2.
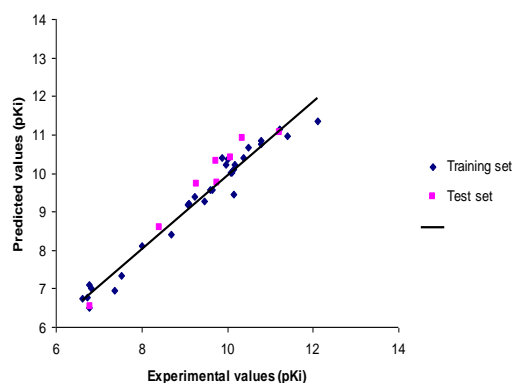

Figure 2- Scatter plot for Topomer CoMFA analysis. Predicted versus experimental values of the training set (blue diamond) and the test set (pink square) compounds

3.2. Contour Plot Analysis

Steric and electrostatic contour plots of topomer CoMFA for R1 and R2 fragments of most active molecule (Compound 5) are shown in figure 3. Contour level, color scheme and estimated volume of contour plot of compound 5 are given in Table 4. In the steric contour plot, green color denotes sterically bulky group is favored and the yellow color indicates sterically bulky group is disfavored by the model. Moreover, in the electrostatics contour plot, red color indicates electronegative-group is favored and blue color indicates electropositive-group is favored by the model. The steric contour plot of R1 fragment for compound 5 (figure 3A) indicates that the methoxy group at phenylsulfonamide moiety and bulky

isobutyl group at R3 substitution are sterically favored. In addition, the electrostatic contour plot of R1 fragment (figure 3B), shows that being an electropositive, methoxy group at phenylsulfonamide moiety occupies blue region and is favored whereas the isobutyl is in open region showing no information that either electropositive or electronegative groups are favored/disfavored.
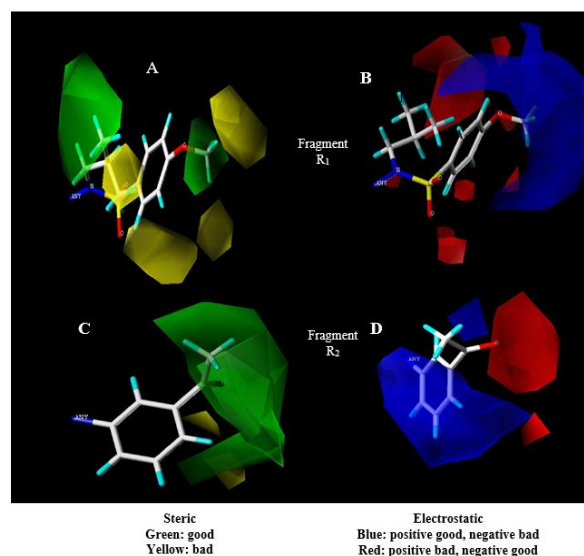

Figure 3- Steric and electrostatics stdev* coefficient contour plots for compound 5 by topomer CoMFA analysis. (A) Steric contour plot for the R1 fragment. (B) Electrostatic contour plot for the R1 fragment. (C) Steric contour plot for the R2 fragment. (D) Electrostatics contour plot for the R2 fragment. Sterically favored/unfavored areas are shown in green/yellow contour, while the blue/red areas show favorable sites for the positively/negatively charged groups.

The steric contour plot of R2 fragment (figure 3C) indicates that acetyl group at position 3 in phenyoxazolidinone moiety sterically favored, as well as electronegative acetyl group occupies red region in the contour plot (Figure 3 D) and is favored. The second highest active compound 6 (pKi=11.3979) has almost same structure as of compound 5 except the position of acetyl group at phenyloxazolidinone moiety, which is at position 4. The contour plot (Figure 4) of compound 6 has the same information for the R1 fragment as of compound 5. The steric contour plot (figure 4C) of R2 fragment of compound 6 indicates sterically favored acetyl group at phenyoxazolidinone moiety, whereas the electrostatic contour plot (Figure4D) shows electronegative acetyl group occupies blue region and is disfavored by the model, hence, diminishing the activity.
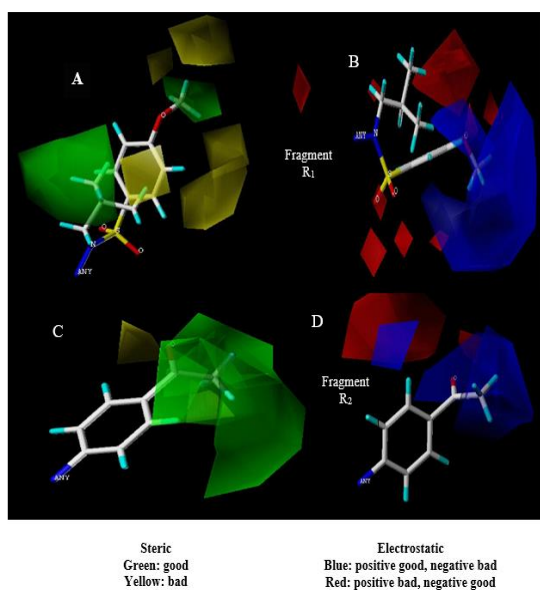
_____

Steric   Electrostatic
Green: good   Blue: positive good, negative bad
Yellow: bad   Red: positive bad, negative good

Figure 4- Steric and electrostatics stdev* coefficient contour plot for compound 6 by topomer CoMFA analysis. (A) Steric contour plot for the R_1 fragment. (B) Electrostatic contour plot for the R_1 fragment. (C) Steric contour plot for the R_2 fragment. (D) Electrostatics contour plot for the R_2 fragment. Sterically favored/unfavored areas are shown in green/yellow contour, while the blue/red areas show favorable sites for the positively/negatively charged groups.

The compound 24 has moderate inhibitory activity (pKi=8.00). The steric contour plot of compound 24 (figure 5 A) of fragment R1 indicates that the sterically bulky group favored at the phenyl sulfonamide moiety by the model whereas the electrostatic contour plot (Figure 5B) indicates that electronegative trifluoromethoxy group occupies blue region and is disfavored by the model and therefore, diminishing the activity.
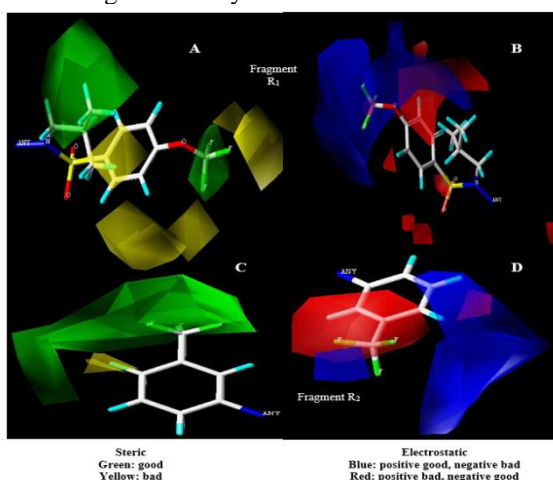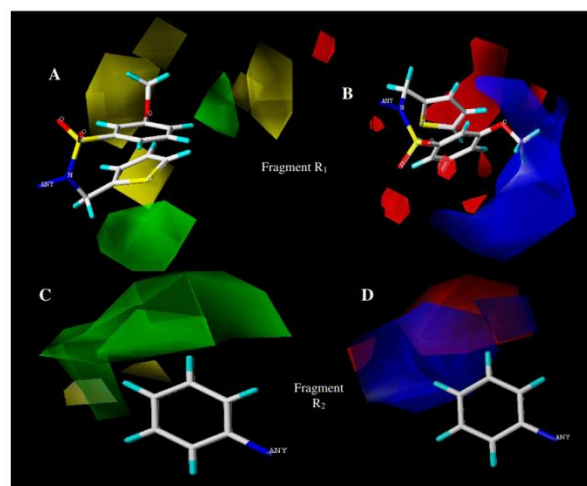

Steric   Electrostatic
Green: good   Blue: positive good, negative bad
Yellow: bad   Red: positive bad, negative good

Figure 5- Steric and electrostatics stdev* coefficient contour plot for compound 24 by Topomer CoMFA analysis. (A) Steric contour plot for the R_1 fragment. (B)

Electrostatic contour plot for the R_1 fragment. (C) Steric contour plot for the R_2 fragment. (D) Electrostatics contour plot for the R_2 fragment. Sterically favored/unfavored areas are shown in green/yellow contour, while the blue/red areas depict favorable sites for the positively/negatively charged groups.

Again, the steric plot of R2 fragment of compound 24 (figure 5C) favored the sterically bulky group and electrostatic contour plot (figure 5D) trifluoromethyl electronegative group occupies red region, favored by the model. The compound 31, which is least active compound (pKi=6.6221) in the series. In steric contour plot of R1 fragment of Compound 31 (figure 6A) methoxy group occupies yellow region and bulky group is not favored by the model at the phenylsulfonamide moiety whereas (2thiophenyl) methyl is in open region where sterically bulky group neither favored nor disfavored. In (figure 6B), (2thiophenyl) methyl group is again in open region whereas methoxy electropositive group occupies blue region and is favored by the model. However, R2 fragment (figure 6C and D) phenyl is in open region in both steric and electrostatic contour plot showing group is neither favored nor disfavored either through sterically or electrostatically.

Steric   Electrostatic
Green: good   Blue: positive good, negative bad
Yellow: bad   Red: positive bad, negative good

Figure 6- Steric and electrostatics stdev* coefficient contour plot for least active compound 31 by topomer CoMFA analysis. (A) Steric contour plot for the R_1 fragment. (B) Electrostatic contour plot for the R_1 fragment. (C) Steric contour plot for the R_2 fragment. (D) Electrostatics contour plot for the R_2 fragment. Sterically favored/unfavored areas are shown in green/yellow contour, while the blue/red areas depict favorable sites for the positively/negatively charged groups.

Through the above discussion, we concluded that the 3 acetyl, 4 acetyl and 3 trifluoromethyl groups of

_____

phenyl ring at oxazolidinone moiety are well suited and enhancing the activity. Moreover, 4 methoxy at phenylsulfonamide moiety is well suited. Isobutyl at R3 substitution is also better than any other group and enhancing the activity. These findings are in the agreement with the structure activity relationship (SAR) study of N-aryl-oxazolidinone-5-carboxamides and derivatives[16].

3.3. HQSAR Analysis

The HQSAR analysis involves varying of some fragment distinctions during the generation of the molecular fragments, and the distinctions used were atoms (A), bonds (B), connections (C), hydrogen atoms (H), chirality (Ch) and donor and acceptor (DA); several combinations of these parameters were considered during the generation of HQSAR models. The HQSAR analyses were performed by screening the 12-default series of hologram length values ranging from 53 to 401 bins. After that, the partial least square (PLS) method was employed to relate the data set compounds patterns of fragment counts to the experimental biological activity of these compounds. The statistical results obtained from PLS analyses using the default fragment size (4-7) and several fragment distinction combinations are given in Table 5.

According to Table 5, the best statistical results among all models using the training set compounds were obtained from model 7 (q^2 = 0.914), which was obtained by using the combination of fragment distinctions: A, B, C and Ch, with optimum number of components 5. This indicates that atom types, bonds, connections, and chirality are essential features of the molecular structures for biological activity. This finding is in agreement with experimental determination as the carbonyl of the 3-acetyl group in Compound number 5 forms van der Waals (VDW) contacts with the residues Gly48, Gly49, and Pro81' of the protease [16]. As the SAR study reported that compounds with the (S)-enantiomer of substituted phenyloxazolidinones at P2 show highly potent inhibitory activities against HIV-1 protease16 and here chirality distinction is essential feature.

The influence of different fragment sizes had been studied on the important statistical parameters in the next step of an HQSAR analysis. Fragment size parameters control the minimum and maximum length of fragments to be included in the hologram fingerprint and can be varied to incorporate larger or smaller fragments during the analyses [17]. The HQSAR results for several fragment sizes, using the best statistical model (model 7, Table 5), are given in

Table 6. The variation of fragment size given in Table 6 provided change in the statistical parameters, in which fragment size 6-9 gives the better result. We get high cross-validated correlation coefficient (q^2 =0.902) and non cross validated correlation coefficient (r^2= 0.973) for this model, with a low cross-validated standard error (SEE = 0.282) which indicates high predictive capability of the HQSAR model.

The selected model was validated by a test set of eight compounds giving satisfactory r^2=0.913. Based on the final HQSAR analysis, the relationshiop between the experimented and predicted activities of both training and test set were fairly determined as the residual values are less than log unit 0.67. The graphical result of both training and test sets of HQSAR analysis is given in figure 8. The experimental and predicted values of activity of the series of N-aryl-Oxazolidinone-5-carboxamides find by HQSAR analysis are also given in Table 1.

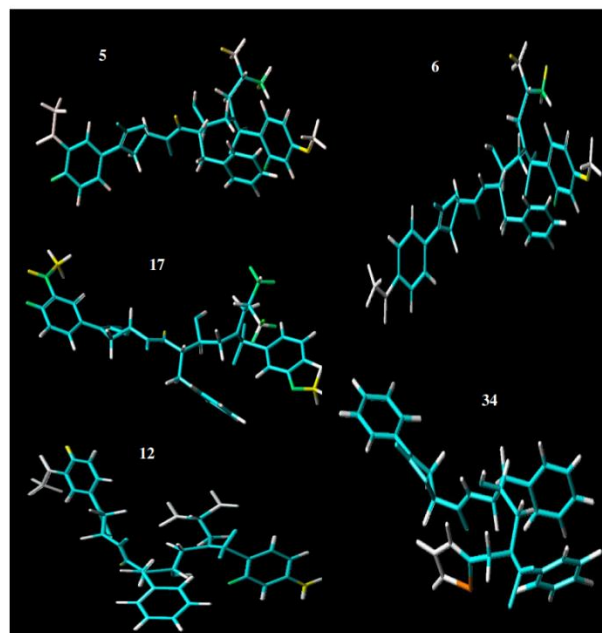Atomic contribution map for few compounds are shown in Figure 7.

Figure 7. Atomic contribution map for few molecules obtained by HQSAR analysis. Color codes red, red orange, and orange show the unfavorable or negative contribution to the activity, the color codes yellow, green blue, and green denote favorable or positive contribution to the activity. The white color shows the moderate contribution to the activity.

The most active compound 5 (pKi =12.0969) and compound 6 (pKi=11.3979) color coded with green, and yellow color indicates positive contribution for inhibitory activity. Methoxy group at phenylsulfonamide moiety with yellow color shows

_____

_____

the positive contribution towards the activity by the model. Isobutyl group at R3 substitution of compound 5 and 6 is in green and yellow color also showing its positive contribution. These contributions are well supported by the above findings of topomer CoMFA analysis.
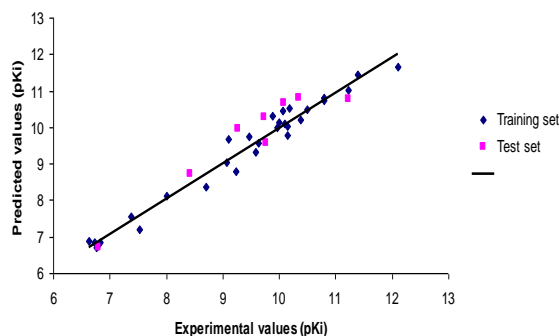

Figure 8- Scatter plot for HQSAR analysis. Predicted versus experimental values of the training set (blue diamond) and the test set (pink square) compounds
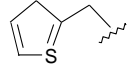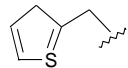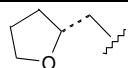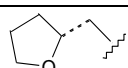
The atomic contribution of compound 12 (pKi=10.4949) shows that the electron donating amino group in position 4 at phenylsulfonamide moiety with yellow color positively contributing towards the activity. Moreover, the atomic contribution of compound 17 (pKi=11.2218) indicates that electron donating dioxolane group at 3, 4 position of phenylsulfonamide moiety with yellow and green color contributing positively towards activity by the model. It also shows that electron withdrawing acetyl group at position 3 of phenyloxazolidinone moiety with green and yellow color contributing positively towards the activity. Whereas the contour plot of compound 34 (pKi=6.7690) is one of the lowest inhibitory activity, shows that (2-thiophenyl) methyl at R3 substitution negatively contributing towards the activity as one fragment is of orange in color.

Table 1-Chemical structure, Biological activities and predicted activities for series of N-Aryl-Oxazolidinone-5-carboxamides
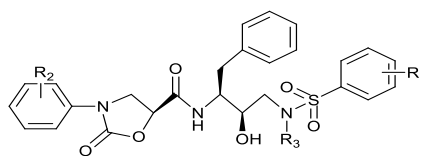
Co. No.	Substitution			Experimental pKi	Predicted pKi (HQSAR)	Residual	Predicted pKi (Topomer CoMFA)	Residual
	R1	R2	R3					
1.	4-OCH$_3$	H		10.0000	10.1352	-0.1352	10.3705	-0.3705
2*.	4-OCH$_3$	3-F		10.0809	10.652	-0.5711	10.4051	-0.3242
3.	4-OCH$_3$	3,4-diF		10.1805	10.5076	-0.3271	10.218	-0.0375
4*.	4-OCH$_3$	3-CF$_3$		11.2218	10.784	0.4378	11.058	0.1638
5.	4-OCH$_3$	3-Ac		12.0969	11.6587	0.4382	11.3438	0.7531
6.	4-OCH$_3$	4-Ac		11.3979	11.4528	-0.0549	10.96	0.4379
7*.	4-OCH$_3$	3-OCH$_3$		10.3468	10.804	-0.4572	10.9115	-0.5647
8*.	4-NH$_2$	H		9.2757	9.951	-0.6753	9.7088	-0.4331
9*.	4-NH$_2$	3-F		9.7696	9.58	0.1896	9.7434	0.0262
10.	4-NH$_2$	3,4-diF		9.6383	9.5583	0.0800	9.5562	0.0821

_____

No.								
11.	4-NH$_2$	3-CF$_3$		10.3768	10.2168	0.1600	10.3963	-0.0195
12.	4-NH$_2$	3-Ac		10.4949	10.4737	0.0212	10.6821	-0.1872
13*.	4-NH$_2$	4-Ac		9.7352	10.268	-0.5328	10.2982	-0.5630
14.	3,4-OCH$_2$O)-	3-F		9.9706	9.9943	-0.0237	10.209	-0.2384
15.	3,4-OCH$_2$O)-	3,4-diF		10.0706	10.4526	-0.3820	10.0218	0.0488
16.	3,4-OCH$_2$O)-	3-CF$_3$		10.7959	10.7539	0.0420	10.8619	-0.0660
17.	3,4-OCH$_2$O)-	3-Ac		11.2218	11.0131	0.2087	11.1477	0.0741
18.	3,4-OCH$_2$O)-	4-Ac		10.7959	10.8115	-0.0156	10.7638	0.0321
19.	3-F, 4-OCH$_3$	3-F		10.1549	9.7657	0.3892	9.4643	0.6906
20.	3-F, 4-OCH$_3$	3,4-diF		9.4647	9.7444	-0.2797	9.2772	0.1875
21.	3-F, 4-OCH$_3$	3-CF$_3$		10.1427	10.0458	0.0969	10.1172	0.0255
22.	3-F, 4-OCH$_3$	3-Ac		9.8761	10.3050	-0.4289	10.4031	-0.5270
23.	3-F, 4-OCH$_3$	4-Ac		10.0969	10.1034	-0.0065	10.0192	0.0777
24.	4-OCF$_3$	3CF$_3$		8.0000	8.1213	-0.1213	8.119	-0.1190
25.	4-OCF$_3$	3-Ac		8.6990	8.3806	0.3184	8.4048	0.2942
26*.	3-OCH$_3$	H		8.4202	8.711	-0.2908	8.5969	-0.1767
27.	3-OCH$_3$	4-Ac		9.0757	9.0274	0.0483	9.1863	-0.1106
28.	4-OCH$_3$	3-F		9.5901	9.3266	0.2635	9.58	0.0101
29.	4-OCH$_3$	3,4-diF		9.2366	8.8002	0.4364	9.3929	-0.1563
30.	4-OCH$_3$	4-Ac		9.0969	9.6643	-0.5674	9.1987	-0.1018
31.	3-OCH$_3$	H		6.6221	6.8726	-0.2505	6.7341	-0.1120
32.	3-OCH$_3$	3-F		6.7240	6.8517	-0.1277	6.7602	-0.0362
33.	3-OCH$_3$	4-Ac		7.5302	7.1894	0.3408	7.3235	0.2067
34.	2,4,5-triF	H		6.7690	6.7130	0.0560	6.4973	0.2717

_____

35*.	2,4,5-triF	3-F		6.7953	6.692	0.1033	6.5319	0.2634
36.	2,4,5-triF	4-Ac		6.7755	6.7342	0.0413	7.0868	-0.3113
37.	3-OCH$_3$	H		7.3768	7.5731	-0.1963	6.9643	0.4125
38.	3-OCH$_3$	3-F		6.8239	6.8484	-0.0245	6.9989	-0.1750

*Compounds marked as * are in test set*

Table 2 Statistical results of topomer CoMFA

ONC	q2	r^2	SEE	r^{2}pred
8	0.874	0.967	0.33	0.9576

ONC=optimum number of components, q^2= Leave-one-out cross validated correlation coefficient, r^2=Non cross validated correlation coefficient, SEE= standard error of estimates, r^{2}pred=predicted correlation coefficient

Table 3 Topomer CoMFA contour plot for the R$_1$ and R$_2$ fragment of compound 5

Contour	R1			R2		
	Contour level	Color	Volume estimates	Contour level	Color	Volume estimates
Steric	-0.029	Yellow	33.3	-0.001	Yellow	1.1
	0.039	Green	36.0	0.014	Green	54.8
Electrostatic	-0.016	Red	22.1	-0.006	Red	20.8
	0.019	Blue	96.5	0.006	Blue	34.6

Table 4- R$_1$ and R$_2$ fragment contribution in Topomer CoMFA Analysis

Co. No.	R$_1$ Fragment	R$_2$ Fragment
1	2.73	0.03
2*	2.73	0.07
3	2.73	-0.12
4*	2.73	0.72
5	2.73	1.01
6	2.73	0.62
7*	2.73	0.57
8*	2.07	0.03
9*	2.07	0.07
10	2.07	-0.12
11	2.07	0.72
12	2.07	1.01
13*	2.07	0.62
14	2.53	0.07
15	2.53	-0.12
16	2.53	0.72

17	2.53	1.01
18	2.53	0.62
19	1.79	0.07
20	1.79	-0.12
21	1.79	0.72
22	1.79	1.01
23	1.79	0.62
24	-0.21	0.72
25	-0.21	1.01
26*	0.95	0.03
27	0.95	0.62
28	1.90	0.07
29	1.90	-0.12
30	0.97	0.62
31	-0.91	0.03
32	-0.92	0.07
33	-0.91	0.62
34	-1.15	0.03
35*	-1.15	0.07
36	-1.15	0.62
37	-0.68	0.03
38	-0.68	0.07

*Compounds marked as * are in test set*

Table 5- HQSAR analysis for various fragment distinction using default fragment size (4-7)

Model	Fragment Distinction	r^2	SEE	q^2	HL	ONC
1.	A/B	0.941	0.412	0.869	307	5
2.	A/B/C	0.930	0.443	0.862	257	4
3.	A/B/C/H	0.935	0.443	0.832	97	6
4.	A/B/C/H/Ch	0.974	0.282	0.894	199	6
5.	A/B/C/H/Ch/DA	0.963	0.334	0.825	199	6
6.	A/B/H	0.943	0.416	0.853	307	6
7.	A/B/C/Ch	0.970	0.293	0.914	353	5
8.	A/B/DA	0.927	0.460	0.817	97	5
9.	A/B/C/DA	0.939	0.421	0.827	199	5
10.	A/B/H/DA	0.936	0.442	0.787	353	6
11.	A/B/C/H/DA	0.942	0.420	0.772	199	6
12.	A/B/H/Ch/DA	0.960	0.348	0.844	199	6
13.	A/B/C/Ch/DA	0.976	0.264	0.871	353	5

_____

Egypt. J. Chem. **66**, No.SI 13 (2023)

_____

r^2= Non cross validated correlation coefficient, SEE= Non cross validated standard error, q^2= leave-one-out cross validated correlation coefficient, HL=Hologram length, ONC= Optimal number of components, Fragment distinction: A= Atom, B=Bond, C= Connections, H=Hydrogen atoms, Ch= Chirality and DA= Donor and Acceptor

Table 6- HQSAR analysis for the influence of various fragment sizes using the best fragment distinction (A, B, C and Ch)

Fragment size	r^2	SEE	q^2	HL	ONC
2-5	0.927	0.459	0.873	353	5
3-6	0.954	0.365	0.900	97	5
4-7	0.970	0.293	0.914	353	5
5-8	0.968	0.300	0.898	257	4
6-9	0.973	0.282	0.902	151	5
7-10	0.978	0.259	0.907	353	6

r^2= Non cross validated correlation coefficient, SEE= Non cross validated standard error, q^2= leave-one-out cross validated correlation coefficient, HL=Hologram length, ONC= Optimal number of components, Fragment distinction: A= Atom, B=Bond, C= Connections, H=Hydrogen atoms, Ch= Chirality and DA= Donor and Acceptor

The above findings are in agreement with SAR study of the compounds [16]. Moreover, the main field (steric and electrostatic) contribution analyzed by topomer CoMFA analysis and the main fragments contribution analyzed by HQSAR analysis are related to the important interactions that determine the preferred binding site and group which enhances the inhibitory activity is methoxy group at the phenylsulfonamide moiety of the compounds 5 /6 forms a hydrogen bond with the backbone nitrogen of the protease residue Asp30' with a distance of 3.4/3.3 Å.16 The 4-acetyl group on the phenyloxazolidinone moiety of the compound 6 does not form any hydrogen bond with the protease but extends beyond the protease into the solvent [16]. The carbonyl of the 3-acetyl group in compound 5 forms van der Waals (VDW) contacts with the residues Gly48, Gly49, and Pro81' of the protease. The methyl group of the acetyl moiety forms VDW contacts with the phenyl ring of Phe53 in the flap region in both the structures [16].

4. Conclusions

In the present study, we obtained the topomer CoMFA and HQSAR models with good statistical values. The robustness of these models was confirmed using the test set prediction. Topomer CoMFA and HQSAR analysis provide great insight for more improvement of potency over existing compounds. The information obtained from HQSAR model shows the importance of atom, bond, connection and chirality parameters.

The overall study indicates that, in topomer CoMFA analysis, sterically bulky and electron donating like 3 methoxy group at phenylsulfonamide moiety is important for inhibitory effect; this was also supported by the HQSAR analysis. The contour plots of topomer CoMFA also suggests that electron withdrawing 3 acetyl, 4 acetyl and 3 trifluoromethyl groups at phenyloxazolidinone moiety are preferable for good inhibitory activity. The HQSAR analysis also prefers the dioxolane and amino groups at phenylsulfonamide moiety. Good statistical results and satisfactory predictive power of developed topomer CoMFA and HQSAR models with contour plots and atomic contribution maps respectively indicate that these models can serve as computational tools for rational design of novel HIV-1 protease inhibitors with enhanced activity and provide good prediction prior to synthesis.

5. Conflicts of interest

There are no conflicts to declare.

_____

6. References

[2] Brik, A., & Wong, C. H. (2003). HIV-1 protease: mechanism and drug discovery. Organic & biomolecular chemistry, 1(1), 5-14.

[3] Kräusslich, H. G., Ingraham, R. H., Skoog, M. T., Wimmer, E., Pallai, P. V., & Carter, C. A. (1989). Activity of purified biosynthetic proteinase of human immunodeficiency virus on natural substrates and synthetic peptides. Proceedings of the National Academy of Sciences, 86(3), 807-811.

[4] Kohl N. E. et al. Proceedings of the National Academy of Sciences U.S.A. 85; 4686–4690; 1988.

[5] Seelmeier, S., Schmidt, H., Turk, V., & Von Der Helm, K. (1988). Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. Proceedings of the National Academy of Sciences, 85(18), 6612-6616.

[6] McPhee, F., Good, A. C., Kuntz, I. D., & Craik, C. S. (1996). Engineering human immunodeficiency virus 1 protease heterodimers as macromolecular inhibitors of viral maturation. Proceedings of the National Academy of Sciences, 93(21), 11477-11481.

[7] Bartlett, J. A., DeMasi, R., Quinn, J., Moxham, C., & Rousseau, F. (2001). Overview of the effectiveness of triple combination therapy in antiretroviral-naive HIV-1 infected adults. Aids, 15(11), 1369-1377.

[8] Gulick, R. M., Mellors, J. W., Havlir, D., Eron, J. J., Meibohm, A., Condra, J. H., ... & Richman, D. D. (2000). 3-year suppression of HIV viremia with indinavir, zidovudine, and lamivudine. Annals of internal medicine, 133(1), 35-39.

[9] Dorsey, B. D., Levin, R. B., McDaniel, S. L., Vacca, J. P., Guare, J. P., Darke, P. L., ... & Schleif, W. A. (1994). L-735,524: the design of a potent and orally bioavailable HIV protease inhibitor. Journal of medicinal chemistry, 37(21), 3443-3451.

[10] Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., ... & Kong, X. P. (1995). ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. Proceedings of the National Academy of Sciences, 92(7), 2484-2488.

[11] Kaldor, S. W., Kalish, V. J., Davies, J. F., Shetty, B. V., Fritz, J. E., Appelt, K., ... & Tatlock, J. H. (1997). Viracept (nelfinavir mesylate, AG1343): a potent, orally bioavailable inhibitor of HIV-1 protease. Journal of medicinal chemistry, 40(24), 3979-3985.

[12] Sham, H. L., Kempf, D. J., Molla, A., Marsh, K. C., Kumar, G. N., Chen, C. M., ... & Norbeck, D. W. (1998). ABT-378, a highly potent inhibitor of the human immunodeficiency virus protease. Antimicrobial agents and chemotherapy, 42(12), 3218-3224.

[13] Robinson, B. S., Riccardi, K. A., Gong, Y. F., Guo, Q., Stock, D. A., Blair, W. S., ... & Lin, P. F. (2000). BMS-232632, a highly potent human immunodeficiency virus protease inhibitor that can be used in combination with other available antiretroviral agents. Antimicrobial agents and chemotherapy, 44(8), 2093-2099.

[14] Wlodawer, A., & Vondrasek, J. (1998). Inhibitors of HIV-1 protease: a major success of structure-assisted drug design. Annual review of biophysics and biomolecular structure, 27(1), 249-284.

[15] Condra, J. H., Schleif, W. A., Blahy, O. M., Gabryelski, L. J., Graham, D. J., Quintero, J., ... & Emini, E. A. (1995). In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. Nature, 374(6522), 569-571.

[16] Ali, A., Reddy, G. K. K., Cao, H., Anjum, S. G., Nalam, M. N., Schiffer, C. A., & Rana, T. M. (2006). Discovery of HIV-1 protease inhibitors with picomolar affinities incorporating N-aryl-oxazolidinone-5-carboxamides as novel P2 ligands. Journal of medicinal chemistry, 49(25), 7342-7356.

[17] Condra, J. H., Schleif, W. A., Blahy, O. M., Gabryelski, L. J., Graham, D. J., Quintero, J., ... & Emini, E. A. (1995). In vivo emergence of HIV-1 variants resistant to multiple protease inhibitors. Nature, 374(6522), 569-571.

[18] Clavel, F., & Hance, A. J. (2004). HIV drug resistance. New England Journal of Medicine, 350(10), 1023-1035.

[19] A. Surleraux, D. L., Tahri, A., Verschueren, W. G., Pille, G. M., De Kock, H. A., Jonckers, T. H., ... & Wigerinck, P. B. (2005). Discovery and selection of TMC114, a next generation HIV-1 protease inhibitor. Journal of medicinal chemistry, 48(6), 1813-1822.B.Miller, W. H., All, F. E., Bondinell, W. E., Callahan, J. F., Calvo, R. R., Eggleston, D. S., ... & Huffman, W. F. (1996). Synthesis of ribofuranosyl glycosides of echiguanines a and b, inhibitors of phosphatidylinositol 4-kinase. Bioorg. Med. Chem. Lett, 6, 2481.C. Ghosh, A. K., Sridhar, P. R., Leshchenko, S., Hussain, A. K., Li, J., Kovalevsky, A. Y., ... & Mitsuya, H. (2006). Structure-based design of novel HIV-1 protease inhibitors to combat drug resistance. Journal of medicinal chemistry, 49(17), 5252-5261.

[20] King, N. M., Prabu-Jeyabalan, M., Nalivaika, E. A., & Schiffer, C. A. (2004). Combating susceptibility to drug resistance: lessons from HIV-1 protease. Chemistry & biology, 11(10), 1333-1338.

[21] Prabu-Jeyabalan, M., Nalivaika, E. A., King, N. M., & Schiffer, C. A. (2003). Viability of a drug-resistant human immunodeficiency virus type 1 protease variant: structural insights for better antiviral therapy. Journal of virology, 77(2), 1306-1315.

[22] Baig, M. H., Ahmad, K., Rabbani, G., Danishuddin, M., & Choi, I. (2018). Computer aided drug design and its application to the development of potential drugs for neurodegenerative disorders. Current neuropharmacology, 16(6), 740-748.

[23] Osakwe, O., & Rizvi, S. A. A. (2016). The significance of discovery screening and structure optimization studies. Social aspects of drug

discovery, development and commercialization, 109-128.

[24] Kalyaanamoorthy, S., & Chen, Y. P. P. (2011). Structure-based drug design to augment hit discovery. Drug discovery today, 16(17-18), 831-839.

[25] Aziz Ajana, M., El Khatabi, K., EL-MERNISSI, R. E. D. A., Khaldan, A., ElMchichi, L., Lakhlifi, T., & Bouachrine, M. (2022). 3D-QSAR, ADMET, and Molecular Docking Studies for Designing New 1, 3, 5-Triazine Derivatives as Anticancer Agents. Egyptian Journal of Chemistry.

[26] Abdulrahman, S. H., Al-khayat, R. Z., & Ali, W. K. (2022). QSAR of antioxidant activity of some novel sulfonamide derivatives. Egyptian Journal of Chemistry, 65(8), 1-6.

[27] ELIJAH ADENIJI, S. H. O. L. A. (2019). Geometrical and topological descriptors for activities modeling of some potent inhibitors against Mycobacterium Tuberculosis: a genetic functional approach. Egyptian Journal of Chemistry, 62(7), 1235-1247.

[28] Hassan, G. S., Georgey, H. H., Mohammed, E. Z., George, R. F., Mahmoud, W. R., & Omar, F. A. (2021). Mechanistic selectivity investigation and 2D-QSAR study of some new antiproliferative pyrazoles and pyrazolopyridines as potential CDK2 inhibitors. European Journal of Medicinal Chemistry, 218, 113389.

[29] Halder, A. K., & Jha, T. (2010). Validated predictive QSAR modeling of N-aryl-oxazolidinone-5-carboxamides for anti-HIV protease activity. Bioorganic & medicinal chemistry letters, 20(20), 6082-6087.

[30] Singh, S., Singh, S., & Shukla, P. (2011). Modeling of novel HIV-1 protease inhibitors incorporating N-Aryl-oxazolidinone-5-carboxamides as P2 ligands using quantum chemical and topological finger print descriptors. Medicinal Chemistry Research, 20(9), 1556-1565.

[31] Sharma, M. C. (2014). Structural requirements of N-aryl-oxazolidinone-5-carboxamide derivatives for anti-HIV protease activity using molecular modelling techniques. Journal of Taibah University for Science, 8(2), 111-123.

[32] SYBYL-X 1.2; Tripos International: South Hanley Rd., St. Louis, Missouri 63144, USA, 1699.

[33] Cramer, R. D., Patterson, D. E., & Bunce, J. D. (1988). Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. Journal of the American Chemical Society, 110(18), 5959-5967.

[34] Cramer, R. D. (2003). Topomer CoMFA: a design methodology for rapid lead optimization. Journal of medicinal chemistry, 46(3), 374-388.

[35] Jilek, R. J., & Cramer, R. D. (2004). Topomers: a validated protocol for their self-consistent generation. Journal of chemical information and computer sciences, 44(4), 1221-1227.

[36] Cramer, R. D., Clark, R. D., Patterson, D. E., & Ferguson, A. M. (1996). Bioisosterism as a molecular diversity descriptor: Steric fields of single "topomeric" conformers. Journal of medicinal chemistry, 39(16), 3060-3069.

[37] Lowis, D. R. (1997). Tripos technical notes. HQSAR: a New, Highly Predictive QSAR Technique, 1.

[38] Salum, L. B., & Andricopulo, A. D. (2009). Fragment-based QSAR: perspectives in drug design. Molecular diversity, 13(3), 277-285.

[39] Flower, D. R. (1998). On the properties of bit string-based measures of chemical similarity. Journal of chemical information and computer sciences, 38(3), 379-386.

[40] Heritage TW, Lowis DR. In Rational Drug Design: Novel Methodology and Practical Applications. New York: Oxford University Press; 1999.

[41] Knuth, D. E. (1973). Sorting and searching.