

Combined 3D-QSAR and Molecular Docking Analysis of Thienopyrimidine Derivatives as Staphylococcus aureus Inhibitors

MebarkaOuassaf¹, Salah Belaidi^{1*}, Saida Khamouli¹, Houmam Belaidi¹ and Samir Chtita²

¹ University of Biskra, Group of Computational and Medicinal Chemistry, LMCE Laboratory, BP 145 Biskra 07000, Algeria

² Laboratory Physical Chemistry of Materials, Faculty of Sciences Ben M'Sik, Hassan II University of Casablanca, BP7955 Sidi Othmane, Casablanca, Morocco

* Corresponding author: E-mail: prof.belaidi@gmail.com, s.belaidi@univ-biskra.dz.

ABSTRACT:

The discovery of antibacterials is considered one of the greatest medical achievements of all time. In this work, a combination of three computational analyzes: 3D-QSAR, molecular docking and ADME evaluation were applied in thienopyrimidine derivatives as gram-positive bacterium staphylococcus aureus.

The validity of 3D-QSAR model was tested with a set of data which is divided into a training and a test set. The two models constructed (CoMFA and CoMSIA) show good statistical reliability ($q^2=0.758$; $r^2=0.96$; $r^2_{pred}=0.783$) and ($q^2=0.744$; $r^2=0.97$; $r^2_{pred}=0.625$) respectively. In addition, docking methods were applied to understand the structural features responsible for the affinity of the ligands in the binding pockets of *S. aureus* DNA gyrase.

Drug likeness and ADME analysis applied in this series of new proposed compounds, have shown that the five lead molecules would have the potential to be effective drugs and could be used as a starting point for designing compounds against staphylococcus aureus.

Keywords: 3D-QSAR; docking; staphylococcus aureus; thienopyrimidine; admet.

1. INTRODUCTION

The Gram-positive Bacterium *Staphylococcus aureus* is medically important pathogens in infection to deep-seated tissue infection and bacteremia,¹ due to the emergence of bacteria resistant to current therapeutic agents, the exploration of new antibiotics of a diversity of infections.²

The enzymes DNA gyrase B is present in bacteria and absent in humans thereby acting as a potential target in treating the *S. aureus* related diseases.³

Gyrase consists of two heterodimeric subunits, GyrA and GyrB. The inhibitors molecules induce cell death by trapping the gyrase DNA complex, inducing oxidative damage, and preventing DNA replication.⁴

Thienopyrimidines represent important chemical class in drug discovery due to vast range of pharmacological properties including antiallergic,⁵ antiviral,⁶⁻⁷ anti-inflammatory,⁸⁻¹² analgesic,¹³⁻¹⁴ antispasmodic, antibacterial,¹⁴⁻¹⁵ antifungal,¹⁶ antimicrobial,¹⁷⁻²¹ antidiabetic,²² antioxidant,²³ antitumor,²⁴⁻²⁹ antipsychotic³⁰ etc. This useful activity of thienopyrimidine generates our interest in developing a tool for screening novel thienopyrimidine analogs are promise antibacterial agent.³¹

The techniques of QSAR are the most prominent computational means to support chemistry within drug design projects where no three-dimensional structure of the macromolecular target is available, The primary aim of these techniques is to establish a correlation of biological activities of a series of structurally and biologically characterized compounds with the spatial fingerprints of numerous field properties of each molecule, such as steric demand, lipophilicity, and electrostatic interactions.³²

For this study, the modern drug discovery aspects were applied such as 3D-QSAR (three-dimensional quantitative structure-activity relationship), Molecular Docking, ADMET (absorption, distribution, metabolism, excretion, toxicity), etc,

The combination of 3D-QSAR and docking analysis permit the direct visualization and interpretation of molecular modeling results within the active site of gyrase–DNA and some

derivatives were consequently generated, and these compounds were evaluated for their drug likeness and (ADMET) properties.

We believe that the results of this work can offer insight into the structural requirements of *S. aureus* inhibitors, providing some reference to guide the design of novel antimicrobial potency against *staphylococcus aureus*.

2. MATERIALS AND METHODS

2.1 Selection of Dataset

Analogues of thienopyrimidine derivatives reported to have potent and selective inhibitory activity against a gram positive (*S. aureus*), were taken from the literature.³³ The structures of the compounds and corresponding pIC₅₀ values (pIC₅₀=-log IC₅₀), where IC₅₀ is the concentration of compound agreed for 50 that inhibited the visible growth of microorganism after overnight incubation for the whole set of ligands are presented in Table1. For 3D-QSAR study, the 27inhibitors were randomly divided into a training set (20 molecules) and test set (7 molecules).

2.2 Computational Approaches

SYBYL-X2.0 package³⁴ running on windows 7, 64 bits workstation was used to perform 3D-QSAR modeling (CoMFA and CoMSIA). The 2D structures of thienopyrimidine derivatives built using the SKETCH option in SYBYL, by utilizing molecular modeling software package SYBYL-X 2.0 with standard geometric parameters, The Tripos force field was employed to carry out energy minimization of each conformation of the molecule, The Gasteiger–Hückel atomic partial charges by the Powell method with a convergence criterion of 0.01 Kcal/mol Å were estimated during minimization, then subsequently converted into 3D structures.³⁵

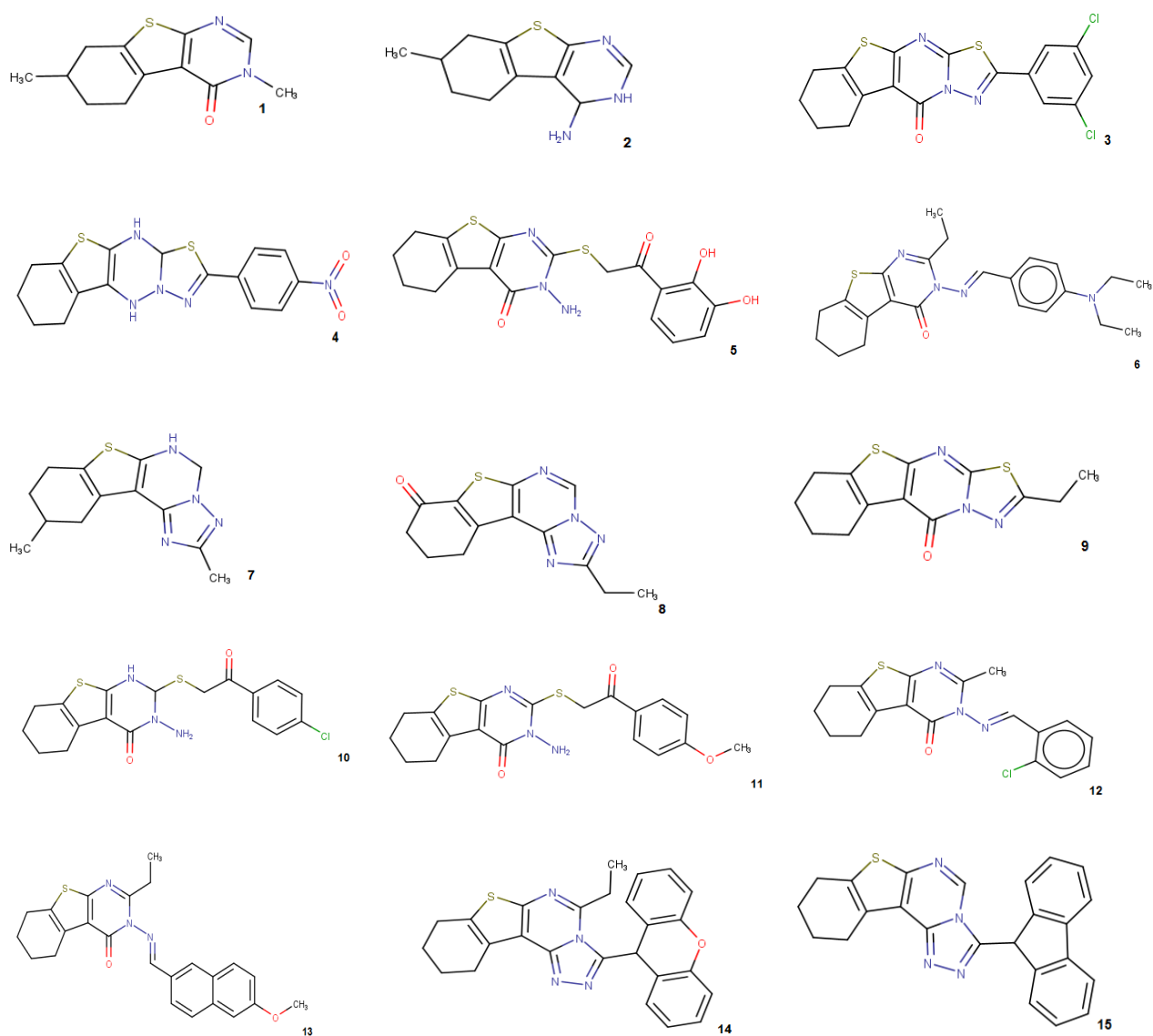
2.3 Molecular Alignment

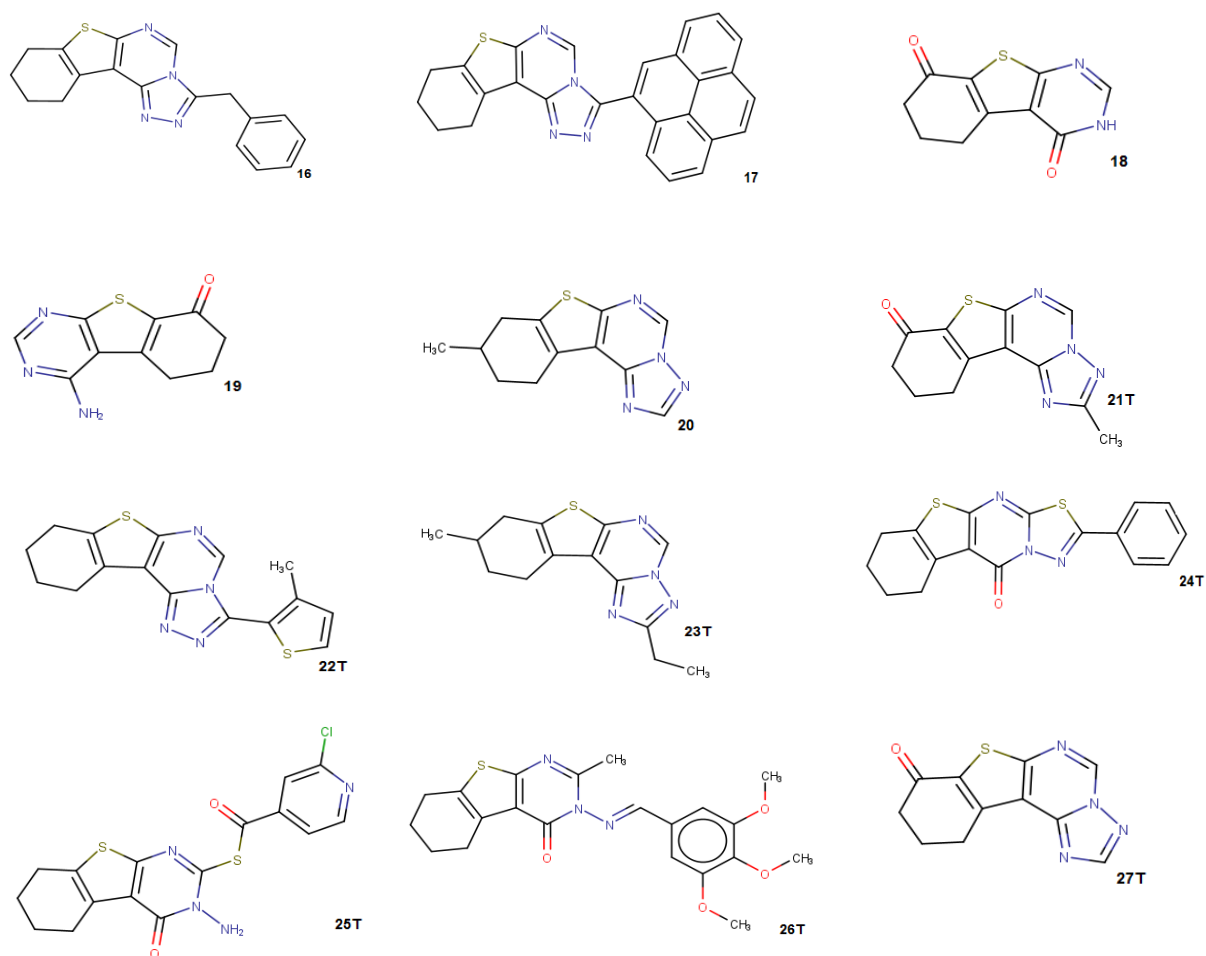
Molecular alignment of compound is a capital step in the construction of 3D-QSAR models.³⁶

88 In the present study ligand-based alignment technique has been chosen in which a
89 template molecule is first isolated over which remaining molecules are aligned, the
90 compound1 was selected as a template and all other molecules were aligned based on the
91 common structure.

92 During the process, all the dataset structures are aligned to the template common
93 substructure using Distill module in SYBYL-X2.0. The superimposed structures of aligned
94 data set are shown in Fig. 2.

95





T=Test set molecule

Fig.1. Structure of thienopyrimidine derivatives

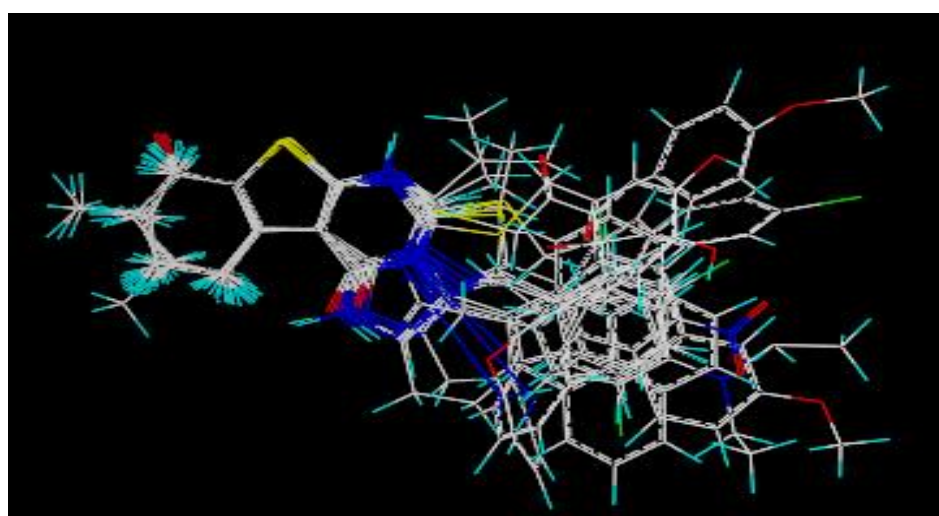


Fig.2. 3D-QSAR structure superposition and alignment of training and test sets

2.4 CoMFA and CoMSIA analysis

The descriptor fields of both methods were calculated in a three-dimensional cubic with one angstrom grid spacing, the frontier of the box extended extra 4 angstrom units from the order of aligned structures in each direction.

For CoMFA method, incorporating steric and electrostatic fields, the probe atom of a charged sp^3 hybridized carbon atom was applied to compute electrostatic and steric fields; the cutoff value was $30 \text{ kcal}\cdot\text{mol}^{-1}$.³⁷

In the case of CoMSIA analysis, five similarity index descriptors consisting of steric (Str), electrostatic (Ele), hydrophobic (Hyd), H-bond donor (HBD), and H-bond acceptor (HBA) fields, A Gaussian function was also applied in calculating the similarity indices, making it accounts for all grid points.³⁸

2.5 Partial Least Square (PLS) Analysis

The PLS statistical method implemented in SYBYL-X 2.0, was used to derive a linear relationship for the 3D-QSAR, and cross-validation was performed using the leave-one-out method.³⁹

In PLS, the independent variables were the CoMFA and CoMSIA descriptors, and pIC_{50} values were used as dependent variables, The ONC was the number of components that led to the highest cross-validated correlated correlation coefficient q^2 (or R^2_{cv})

2.6 Model Validation

The predictive power of CoMFA and CoMSIA models was further validated by using an external test set (inhibitors marked with “T” in Table 2).

To avoid excessive extrapolation upon external prediction, Golbraikh and Tropsha's Criteria followed in developing activity predictors, especially for continuous QSAR, are as follows: (i) correlation coefficient R between the predicted and observed activities; (ii) coefficients of determination predicted versus observed activities R^2_0 and The inhibitors in the test set were given exactly the same pretreatment as the inhibitors in the corresponding training set. The correlation between the experimental and predicted activity for models was

calculated as R_{pred}^2 value observed versus predicted activities $R_0'^2$ for regressions through the origin; and (iii) slopes k and k' of regression lines through the origin.⁴⁰

2.7 Molecular docking

In order to check the reliability of the established 3D-QSAR models, were subjected to docking with DNA gyrase subunit b (PDB ID: 3G7B),⁴ from the Protein Data Bank (RCSB) (<http://www.rcsb.org/pdb>).

Water and co-crystal ligand molecules were eliminated from the structures, molecular docking study was performed using Surflex-dock implemented in SYBYL-X2.0, The ligands and protein preparation steps for the docking protocol were carried out in SYBYL-X 2.0, then results were analyzed using Discovery Studio⁴¹ and MOLCAD module implemented SYBYL-X 2.0.

The MOLCAD program (Molecular Computer Aided Design) was employed to visualize the binding mode between the protein and ligand. MOLCAD calculates and exhibits the surfaces of channels and cavities, as well as the separating surface between protein subunits.⁴²

2.8 Pharmacokinetic Profile

The chemical structure of the compound was submitted in the form of canonical simplified molecular input line entry system (SMILE), to estimate several in silico pharmacokinetic parameters using the Swiss ADME tool⁴³ the pharmacokinetic profile of the compound was evaluated. Gastrointestinal absorption, Blood-Brain Barrier penetration, Skin Permeation, synthetic associability and drug-likeness prediction like Lipinski,⁴⁴ and Veber rules,⁴⁵ interaction of molecules with cytochromes P450 (CYP) and bioavailability score.

3. RESULTS AND DISCUSSION

3.1 3D QSAR studies

CoMFA and CoMSIA 3D-QSAR models were derived using DNA gyrase inhibitors.

The predicted and experimental activity values and their residual values for both the training and test sets of CoMFA and CoMSIA models are given in Table 1.

The results of CoMFA and CoMSIA SYBYL, studies are summarized in Table 2, The q^2 , R^2 , F, and SEE values were computed as defined in PLS analysis showed a q^2 value of 0,758 and R^2 of 0,96 for CoMFA analysis, a non-cross-validated PLS analysis results in a conventional R^2 of 0,944, F = 128, and a standard error of estimation (SEE) of 0,113 for CoMFA analysis.

The steric and electrostatic contributions were 0.576 and 0.246 respectively. These results indicate that steric field contributed highest to the binding affinity.

CoMSIA model was obtained by using the combination of steric, electrostatic, hydrophobic, H-bond donor and H-bond acceptor fields, the statistical results obtained from a combination of these five fields with the four components are (q^2 = 0,744, R^2 = 0,97, F =527, SEE = 0,097).

The corresponding field contributions are 0,116 (steric), 0,201 (electrostatic), 0,253 (hydrophobic), 0,211 (HBD) and 0,169 (HBA), this is suggesting that the hydrophobicity of the molecule influences their inhibitory potential.

Table 1 : Calculated data for the 3D-QSAR model

model	R^2	q^2	F	SEE	ONC	Field contribution					R^2_{pred}
						STR	Ele	Hyd	HBD	HBA	
CoMFA	0,96	0,758	128	0,113	3	0,574	0,426				0,783
CoMSIA	0,97	0,744	527	0,101	4	0,166	0,201	0,253	0,211	0,169	0,625

The higher value of F, greater the probability that the QSAR equation is significant.⁴⁷ The F values for the CoMFA and CoMSIA models were 128 and 527 respectively. The F value stands for the degree of statistical confidence.

Predicted versus experimental final pIC50 values for CoMFA and CoMSIA models and their residues (for the training and test sets) are given in table 2.

The correlation between the predicted and the experimental pIC50 of training and test sets is depicted in Figure 4 for CoMFA and CoMSIA

analysis, illustrate the predicted activities using the CoMFA model are in good agreement with the experimental data, suggesting that the CoMFA model should have a satisfactory predictive ability. Results show that prediction by the CoMSIA model is reasonably accurate.

Finally, the predictability of the proposed models was confirmed using external verification and the R^2_{pred} values were 0,783 and 0,625 for CoMFA and CoMSIA models respectively, the results of these statistics indicated good stability and strong predictive power for the CoMFA and CoMSIA models.

Histogram of residual values obtained from CoMFA and CoMSIA analysis is depicted in Figure 4. They suggest the absence of any outlier compound in the training set whose residual activity is above one.

There is a slight statistical difference between CoMFA and CoMSIA models that indicate the five fields contribute almost as much to the relationship.

Table 2. Experimental and calculated activity (pIC₅₀) for staphylococcus aureus inhibitors of set training and test set for the CoMFA and CoMSIA models.

Compounds	pIC ₅₀ exp. ³³	pIC ₅₀ pred.			
		CoMFA	residue	CoMSIA	residue
1	3,290	3,255	0,035	3,272	0,018
2	3,590	3,420	0,170	3,451	0,139
3	4.000	4,080	-0,080	4,131	-0,131
4	4.000	4,090	-0,090	4,033	-0,033
5	4,220	4,420	-0,200	4,240	-0,020
6	5.000	5,020	-0,020	5,018	-0,018
7	3,890	3,733	0,157	3,725	0,165
8	3,890	3,771	0,119	3,862	0,028
9	4.000	4,065	-0,065	4,012	-0,012
10	4.000	4,035	-0,035	4,005	-0,005
11	4.000	4,014	-0,014	4,029	-0,029
12	5.000	4,715	0,285	4,801	0,199
13	5.000	5,107	-0,107	5,125	-0,125
14	4,190	4,160	0,030	4,176	0,014
15	4,490	4,302	0,188	4,373	0,117
16	4,490	4,242	0,248	4,337	0,153

17	4,490	4,493	-0,003	4,530	-0,040
18	3,290	3,224	0,066	3,390	-0,100
19	3,590	3,515	0,075	3,584	0,006
20	3,590	3,459	0,131	3,608	-0,018
21 T	3,590	3,532	0,058	3,755	-0,165
22 T	4,190	3,964	0,226	4,079	0,111
23 T	3,890	3,685	0,205	3,772	0,118
24 T	4,220	4,331	-0,111	4,345	-0,125
25 T	4,090	4,320	-0,230	4,320	-0,230
26 T	5,000	4,884	-0,156	4,666	0,178
27T	3,890	3,517	0,373	3,751	0,139

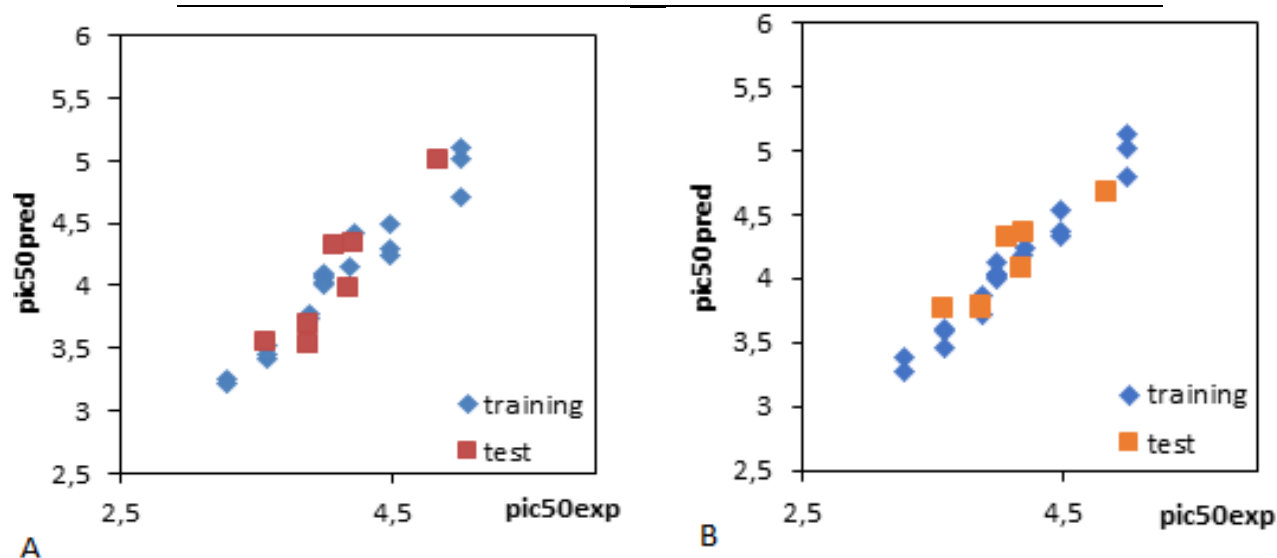


Fig.3. Graph of staphylococcus aureus inhibitors predicted activity of training and test set from a) CoMFA and b) CoMSIA analysis.

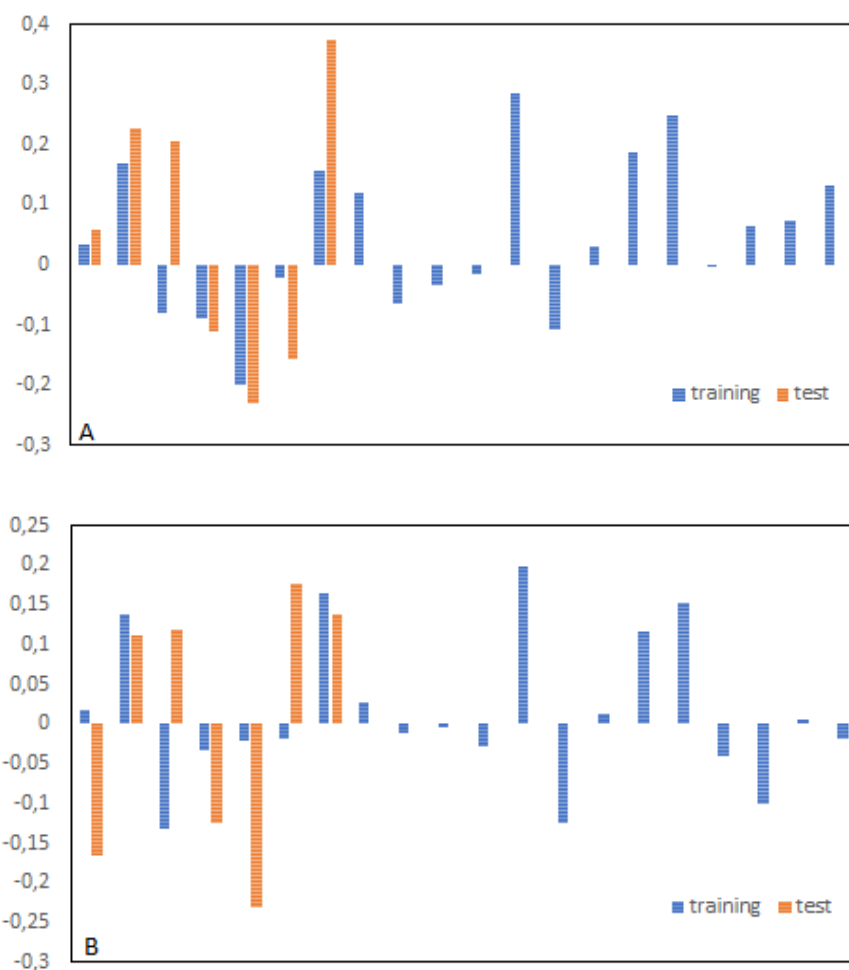


Fig.4. Histogram of residual values from A) COMFA analysis B) COMSIA analysis

3.2 Model validation results

The Table 3 shows statistical parameters associated with CoMFA and CoMSIA models. All the calculated parameters indicated that both models showed a good predictive power. It could be observed that all the Golbraikh–Tropsha criteria criteria:

$$r_{pred}^2 0,6, 0,85 < K < 1,15, 0,85 < K' < 1,1, R^2 \text{ is close to } 1, R'^2 \text{ is close to } 1 \text{ and}$$

$$|R_0^2 - R'^2| < 0,3 \text{ were fulfilled.}$$

Table3. Predictive power results for the external test set; Golbraikh and Tropsha criteria

MODEL	r_{pred}^2	K	K'	R^2	R'^2	$ R_0^2 - R'^2 $
CoMFA	0.783	0.976	1.021	0.941	0.965	0.02
CoMSIA	0.666	0.990	1.007	0.991	0.992	0.001

3.3 CoMFA and CoMSIA contour maps

To visualize the information content of the derived 3D-QSAR model, CoMFA and CoMSIA contour maps were generated to rationalize the regions in 3D space around the molecules where changes in the steric, electrostatic, steric, hydrophobic, H-bond donor, and H-bond acceptor fields were predicted to increase or decrease the activity.

A thorough analysis of the contours obtained determines the vital physicochemical properties responsible in determining the activity and explores the crucial importance of various substituents in their 3D orientation.

The visualization of the results of the CoMFA and CoMSIA models have been performed using the StDev*Coeff mapping option contoured by contribution, the default level of contour with contribution, 80% for favored region and 20% for disfavored region was set during contour analysis.

The CoMFA and CoMSIA steric and electrostatic contour maps were shown in Fig, 5a, b using compound 13 the most active of the series as a reference structure explaining the key structural features required for inhibitory activity.

The steric contour maps of the CoMFA and CoMSIA models are shown in yellow and green colors, the green contours represent regions of high steric tolerance (80% contribution) while the yellow contours represent regions of low steric bulk tolerance (20% contribution).

In the contour map of steric field (Figure 5 a), a large green contour was observed near the naphthalene ring, suggesting the bulky substituent was favored at this region.

Therefore, it is reasonable for the activity order of those compounds, 17(pIC₅₀=4.49)>16(pIC₅₀=4.46)>14(pIC₅₀=4.19)>2(pIC₅₀=2.59)>1(pIC₅₀=2.5), with the corresponding R1 substituent pyrene, 9H-xanthene, Phenyl, Methyl and H respectively.(The figure 9 shows the location of the radicals R1 and R2).

It is clear that the N methyleneamino (the link between ring and thienopyrimidine) is surrounded by most of the yellow areas; the phenomenon demonstrates that bulky groups are unfavorable for increasing the activity.

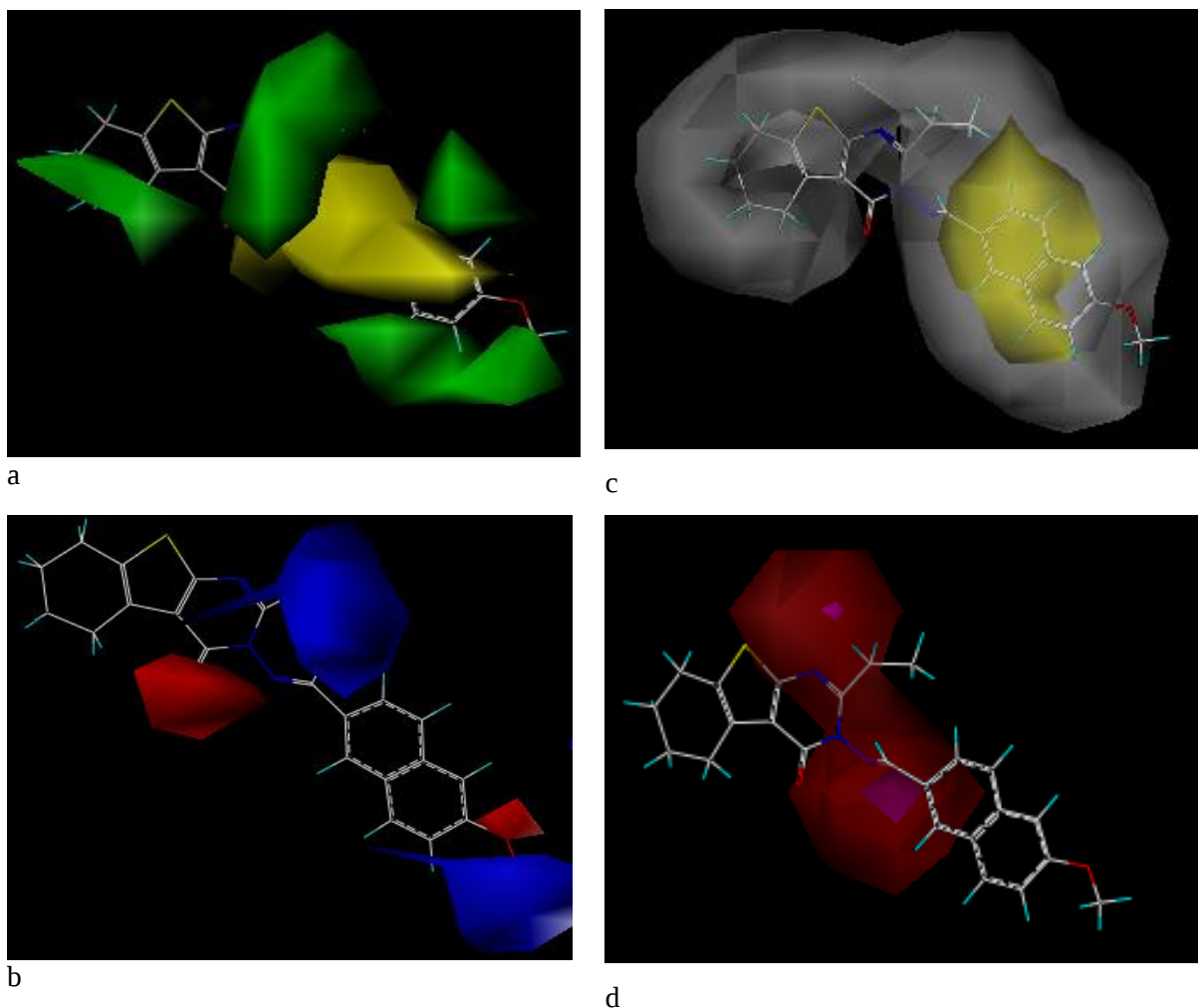


Fig.5.CoMFAandCoMSIA STDEV*COEFF contour maps: a) steric, b) electrostatic, c) Hydrophobic, d) hydrogen-bond acceptor and hydrogen-bond donor fields; based on the most active compound 13.

The electrostatic field (Figure 5b) is indicated by Two blue regions were found near the R_1 and R_2 position, which can explain the fact that the activity of compounds 25 ($R_2 = NH_2$) and 15 ($R_2 = H$) are less potent than the compound 14 ($R_2 = C_2H_5$), therefore, it is reasonable for the activity order of those compounds, $14 (R_1 = C_2H_5) > 12 (R_1 = CH_3) > 11 (R_1 = NH_2)$, because that substitution of electropositive groups at this position would increase the activity and emphasizes that the electronegative environment is undesirable at this position.

The red contour surrounding the oxygen atoms of the methoxy group sheds light on the fact that the activities of compound 13 ($R_2 = O-CH_3$) is higher than that of the compound 16 ($R_2 = H$), and which can explain the fact that the activity of compound 30 which have three group methoxy around the naphthalene ring where any electronegative group at this region would increase the activity.

CoMSIA contribution maps denote those areas within the specified region where the presence of a group with a particular physicochemical property will be favored or disfavored for good inhibitory activity.

CoMSIA calculates both steric and electrostatic fields, as in CoMFA, but additionally uses hydrophobic, HBD and HBA fields, favored and disfavored levels fixed at 80% and 20%, respectively.

The CoMSIA hydrophobic contour map is shown in Fig 5c, represented by yellow (80% contribution) and gray (20% contribution) colored contours.

Yellow colored contours indicated the regions where hydrophobic groups on ligands are favored and gray colored contours represent those areas where hydrophobic groups are unflavored (or favorable for hydrophilic groups on ligands).

The calculated CoMSIA hydrophobic contours (Fig. 5c) display favorable hydrophobic substituents (yellow polyhedral) in proximity of the naphthalene ring. Unfavorable areas (white) are located around the thienopyrimidine and the substituent R1, and in proximity of the R2: methoxy group.

Presence of a big white contour near R1 substituents of thienopyrimidine ring shows the importance of hydrophilic groups on the antibacterial activity in this region.

As shown in Fig.5d, the magenta contours indicate hydrogen bond-accepting groups increase the inhibitory activity, whereas the red contours indicate hydrogen bond-accepting groups decrease the activity, a magenta contour located on the amino between thienopyrimidine and naphthalene around the N atom in the ring pyrimidine suggested that hydrogen bond-accepting groups were favored.

3.4 Docking Analysis

Molecular docking is a computational approach that finds best binding orientation between two biomolecules the ligand and the protein.⁴⁶

The Protein-Ligand interaction plays a vital role in structural based drug design⁴⁷.

In our present study, docking of tested compounds with the primary drug pathway for *S. aureus* was performed, subsequently, the active compound 13 and inactive compound 1 were docked into the ligand-binding pocket of DNA gyrase B protein (code PDB:3G7B), as described in Fig.6.

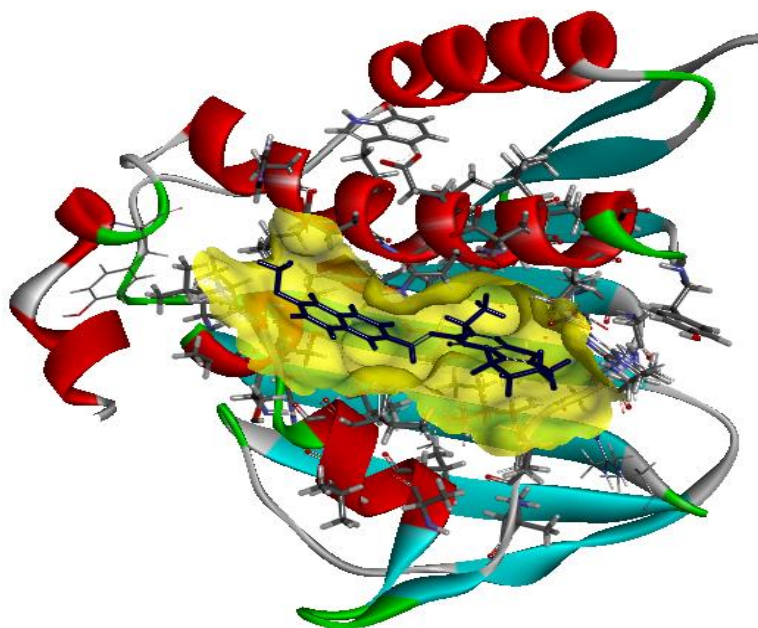


Fig.6 Three-dimensional structure of the receptor proteins DNA gyrase Bin complex with the compound 13.

Docking interactions with two compounds (13 and 1) are shown in the Figs. 7a–c, respectively.

For the low active compound, the Docking results shows carbon-hydrogen bond with Gly85Asp81 and Arg84 residues, pi-alkyl interaction with Ile86 residue.

While compound 13 is stabilized by a number of hydrophobic contacts with the residues Ile 86, Pro87 and Ile 51 residues, as shown in Fig.7, the ligand 13 displayed three hydrogen bond interactions, one of the hydrogen bonds was observed between NH group val130 and O-atom of methoxy group at distance of 2,38 Å.

Another hydrogen bond was observed between of Thr173 and one of the nitrogen atoms of the pyrimidinone ring at a distance of 2,5 Å, the third bond was observed between Asn54 and NH- group.

The key amino acid residues within the docking complex model involved in the interaction between the two compounds (most active, and low active) were Gly 85 and Arg48 corroborating the studies of Berk et al.⁴⁸

The type and the position of interactions were suggested by contour map analysis. This supports the validity of our results.

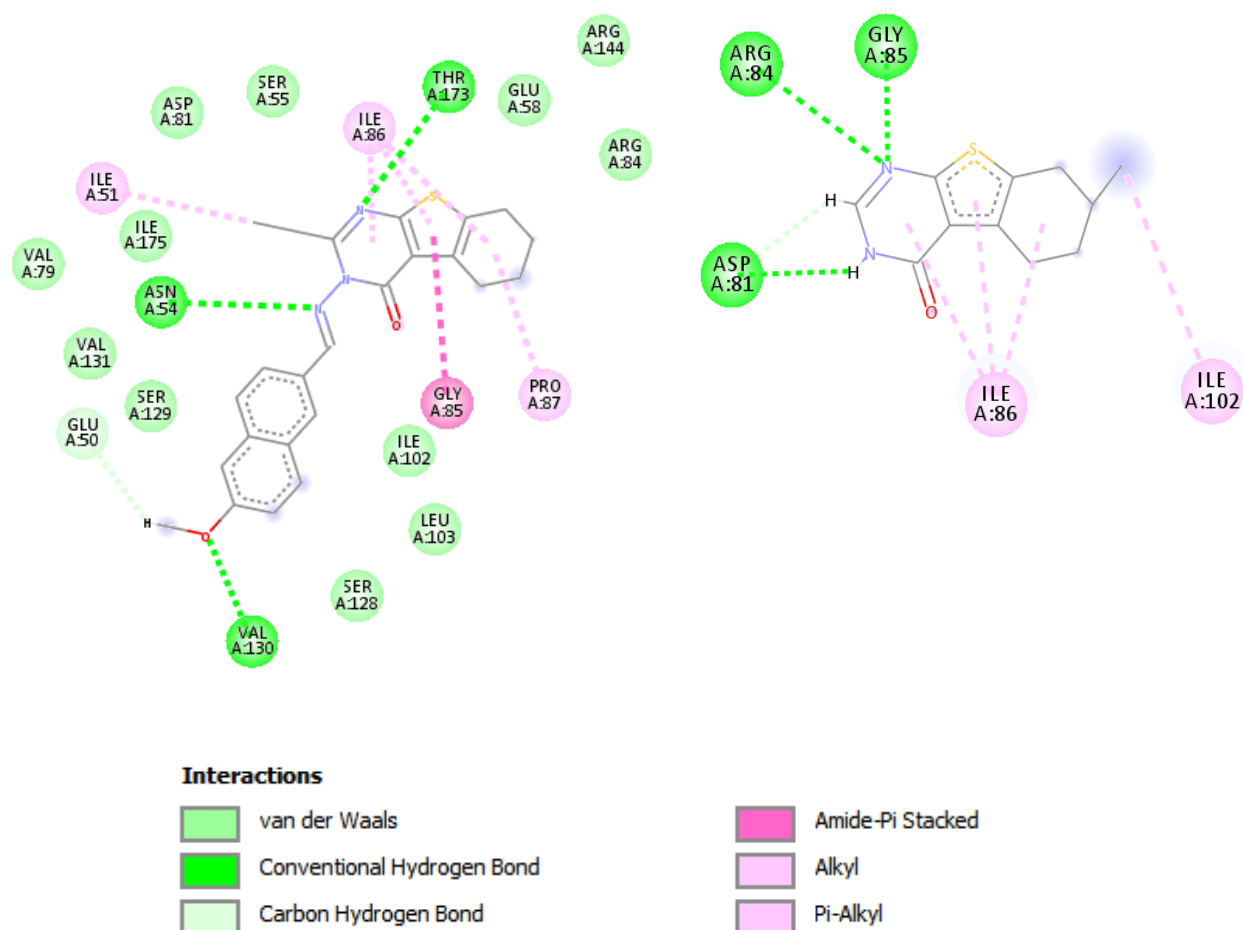


Fig.7.Two-dimensional depiction of the docked conformations of Ligand 13 and ligand 1 with enzyme DNA gyrase protein

To further visualize the binding mode, the molecular computer aided design program (MOLCAD) was conducted, MOLCAD could calculate and display the cavity depth (CD), electrostatic potential (EP), lipophilic potential (LP), and hydrogen bond site (HB) of the binding pocket, which can be used to find the sites that act attractively on ligands by matching opposite colors.

In Figure 8(CD) , the MOLCAD Multi-Channel cavity depth potential surfaces structure of the binding site within the compound13is displayed and the cavity depth color ramp ranged from blue (low depth values = outside of the pocket) to ORANG (high depth values = cavities deep inside the pocket), In Figure 8(CD), the R1 position naphthalene of compound 13 is observed in a blue area, revealing that this position was embedded deep inside the ATP pocket, It can be simply inferred that a bulky group at R1 position maybe favorable, Since the thienopyrimidine site was oriented to a light Yellow/Orange area, which illustrated a minor

group was anchored into a favorable region, this suggests that minor groups may benefit the potency.

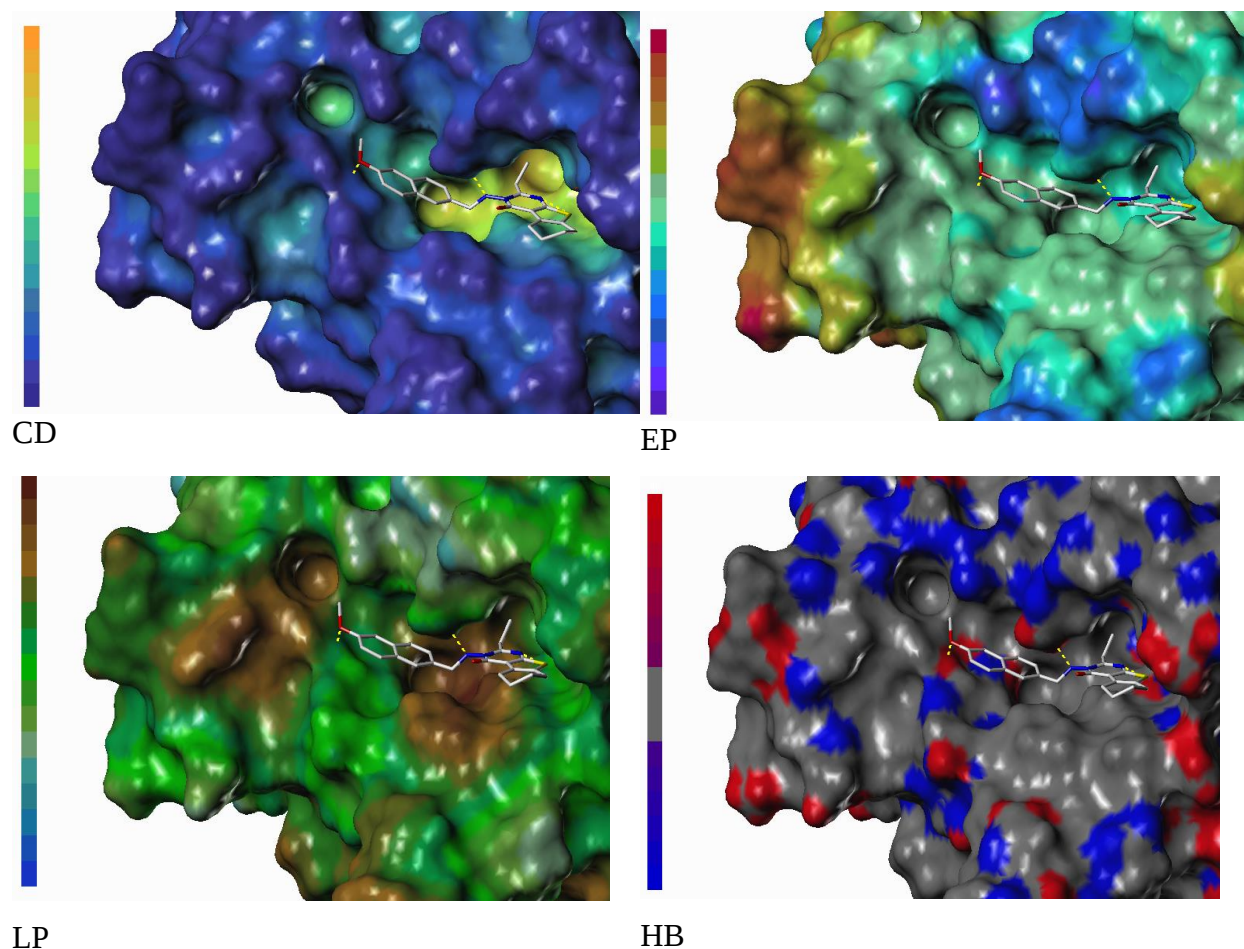


Fig.8. MOLCAD surfaces of the binding site of DNA gyrase protein with molecule 13

In Fig. 8(EP), the MOLCAD electrostatic potential surface of the binding region was demonstrated with the color ramp for EP ranging from red (most positive) to blue (most negative), the position R1 group was found in a blue area, which indicated that electron-donating properties at this site were essential for the potency; the sulfo group was in a yellow area, which suggested that electron-withdrawing properties would be favored; the $-CH_2CH_3$ radical was anchored in a green area which suggested that an electron-donating substituent at this position would be essential for the potency.

These results were well compared with the corresponding CoMFA and CoMSIA electrostatic contour maps.

Figure 8 (HB) displayed the MOLCAD hydrogen bonding sites of the binding surfaces, ligands can be docked to proteins by matching the patterns displayed on the surface,

the color ramp for HB ranges from red (hydrogen donors) to blue (hydrogen acceptors). The nitrogen of thienopyrimidine ring and N methyleneamino of compound 13 was found in the red surface, which suggested that the surface of this site are hydrogen bond donors, and a hydrogen bond acceptor substituent would be favorable; and the naphthalene ring of compound 13 was found in the blue surface, which indicated that the surface of this region are hydrogen bond acceptors, and a hydrogen bond donor substituent be favored.

The observations taken from this hydrogen bonding sites satisfactorily matched to the corresponding CoMSIA hydrogen bond donor contour maps.

Figure 8 (LP) showed the MOLCAD lipophilic potential surface of the binding area, the color ramp for LP ranges from brown (highest lipophilic area of the surface) to blue (highest hydrophilic area). The R1 position was oriented to a brown region, suggesting that a hydrophobic substituent may be favored; the methylene amino was oriented to a blue area, which indicated that a hydrophilic group would be favorable. The observations taken from Fig. 8 satisfactorily matched those of the CoMSIA hydrophobic contour map.

Combined 3D-QSAR and molecular docking analysis is corroborated and these results will help to better interpret the structure-activity relationship of these DNA gyrase inhibitors and provide valuable information into rational drug design.

3.5 New compounds design and activity prediction

Based on the established two sets of 3D-QSAR models and related analysis results, the compound 13 was used as a template to modify its molecular structure, and five new compounds were designed. The structures of the new compounds are shown in Table 4. We substituted R1 and R2 parts with proper groups according to the contour maps. The activities of these designed structures towards *Staphylococcus aureus* antagonist were almost better compared to that of reported thienopyrimidine derivatives.

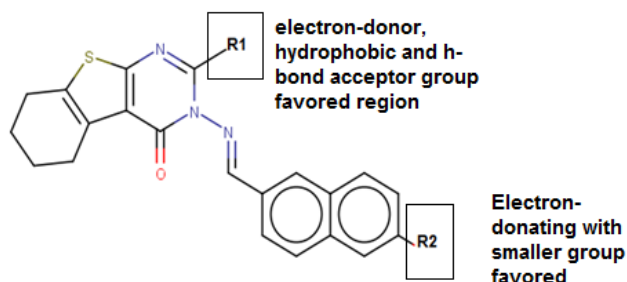


Fig.9. Structure-activity relationship representation.

Table 3. Structures of newly designed molecules and their predicted pIC50 based on CoMFA and CoMSIA 3D-QSARmodels.

	Smile	R1	R2	COMFA	COMSIA
13	<chem>CCc1nc2sc3c(c2c(=O)n1/N=C\c1ccc2c(c1)ccc(c2)OC)CCCC3</chem>	CH ₂ CH ₃	OCH ₃	5.107	5.125
A	<chem>CCCC1nc2sc3c(c2c(=O)n1N=Cc1ccc2cc(OC(C)C)ccc2c1)CCCC3</chem>	-(CH ₂) ₂ CH ₃	OCH(CH ₃) ₂	5.129	5.125
B	<chem>CCOc1ccc2cc(C=Nn3c(CC(C)C)nc4sc5c(c4c3=O)CCCC5)ccc2c1</chem>	CH ₂ CH(CH ₃) ₂	OCH ₂ CH ₃	5.181	5.135
C	<chem>NC1=CC2=CC=C(\C=N\N3C(=O)C4=C(SC5=C4CCCC5)N=C3C3CCCCC3)C=C2C=C1</chem>	C ₆ H ₁₁	NH ₂	5.124	5.126
D	<chem>CCCCC1=NC2=C(C3=C(CCCC3)S2)C(=O)N1N=CC1=CC=CC2=C1C=CC(=C2)N(C)C</chem>	(CH ₂) ₃ CH ₃	N(CH ₃) ₂	5.218	5.154
E	<chem>CCCCCc1nc2sc3c(c2c(=O)n1/N=C/c1ccc2c(c1)cc(cc2)NC)CCCC3</chem>	(CH ₂) ₄ CH ₃	CH ₃ NH	5.137	5.128

3.6 Drug-likeness, bioavailability, synthetic accessibility and alerts for PAINS

Drug likeness may be defined as a complex balance of various molecular properties and structural features, which determine whether a particular molecule is drug or nondrug. Probably, the most widely used filter is Lipinski's Rule-of-five, which proposes that molecules with poor permeation and oral absorption have molecular weight > 500, logP > 5, more than 5 hydrogen-bond donor and more than 10 acceptor groups.⁴⁴ All the molecules exhibited drug likeness characteristics according to Lipinski rules. The other significant properties such as total polar surface area (TPSA) and the number of rotatable bonds and molar refractivity were also calculated. The results are depicted in Table 5. TPSA of a

compound is less than 140 Å² and the number of rotatable bonds is less than 10, as the number of rotatable bonds increases, the molecule becomes more flexible and more adaptable for efficient interaction with a particular binding pocket⁴⁹. Interestingly the compounds E, D and A have 6–7 rotatable bonds and flexible.

So, Lipinski and Veber rules are validated, therefore, theoretically, there would not have a problem with oral bioavailability for all proposed compounds.

Table 3. Prediction of molecular properties descriptors of the new compounds design

Compounds	MW g/mol	logP	H-bond A	H-bond D	log S mol/L	n.rot	REF	TPSA (Å ²)	S.A score	F %	PAINS alert
13	417.52	4.30	4	0	-6.06	4	124.52	84.72	3.92	55	0
A	459.60	4.80	4	0	-6.88	6	138.94	84.70	4.30	55	0
B	458.62	4.93	3	0	-7.30	3	138.94	71.83	4.15	55	0
C	456.60	4.18	3	1	-6.93	3	139.54	101.5	4.32	55	0
D	458.62	4.74	3	0	-6.80	6	141.85	78.73	4.46	55	0
E	458.62	4.78	3	1	-6.99	7	141.75	87.52	4.38	55	0

Drug oral bioavailability is the fractional extent of the drug dosage that finally reaches the therapeutic site of action and is quantitatively symbolized as % F, acceptable probability score is 55%, which indicates that it passed the rule of five. The compounds showed a score of 55%, indicating good bioavailability.

For the discovery of oral administrative drugs, solubility is one of the major descriptors. Highly water solubility was useful for deliver active ingredient in sufficient quantity in small volume of such pharmaceutical dosage. These values are the decimal logarithm of the molar solubility in water (log S). From the results appear in the table 5, it can be said that the compounds tested has poorly water solubility. Low water solubility translates to slow absorption and action.

Activity artifacts in assays present a major problem for biological screening and medicinal chemistry. Such artifacts are often caused by compounds that form aggregates or are reactive under assay conditions. Many pan assay interference compounds (PAINS) have been proposed to cause false-positive assay readouts.⁵⁰ The PAINS violations of proposed compounds are given in table 5. Almost all the compounds showed zero PAINS alert and can be used as lead compounds.

One of the key aspects of CADD (Computer aided design and drafting) activities is help for the selection of most promising molecule which was synthesized and subjected for

biological study is the synthetic accessibility (SA). For given molecule, SA score is the summation of the fragments and corrected by the terms describing size and complexity such as macrocycles, chiral centers, or spiro functions. The SA score ranges from 1 (very easy) to 10 (very difficult).⁵¹ The obtained values were in the range of 3-5 revealed that the compounds here have easy synthesis route.

3.7 ADME evaluation of the new candidates.

The pharmacokinetic studies were performed using online SwissADME tool, the calculated absorption, distribution and metabolism parameters are presented in Table 6 and Table 7 respectively.

Transdermal delivery systems are attractive for both topical and systemic therapeutics. However, the skin barrier, which protects the body from physical and chemical attacks, also hinders the delivery of the required drug dose through the skin to a target organ.⁵²

The results in the table show that all the compounds found (table 6) to be poorly permeable to skin as all compounds have Kp negative values.

Moreover, other parameters used to measure the adsorption and distribution of these drugs is through human intestinal absorption (HIA) or gastrointestinal (GI) adsorption data. These data show that all the compounds are predicted to be well absorbed, except for the compound C, whose absorption is weak. This result is also evident in the BOILED-Egg model. (FIG9)

The Blood–brain partitioning and brain distribution are critical properties for drugs targeting the central nervous system. The Compounds tested are predicted as non-brain penetrant thus, side effects at this level may be diminished.

The BOILED-Egg model is of great support for the users to apprehend the concepts of absorption and distribution, and to figure out what type of chemical modifications must be made to the small molecule to obtain the desired absorption and distribution, in an intuitive and iterative way.⁵³

Table 4. Predicted ADME properties for new inhibitors

Compound	GI absorption	BBB Permeable	log Kp (cm/s)
13	high	yes	-5.85
A	high	no	-4.42
B	high	no	-3.93
C	low	no	-4.54
D	high	no	-4.49
E	high	no	-4.20

Inside circle [yellow] depicts BBB- blood brain barrier, none of the compounds are in this region. The white region that is outer to yellow depicts the human intestinal absorption. Almost all compound lies in this white area. Only molecule C is lying outside [grey area] which indicates poor intestinal absorption.

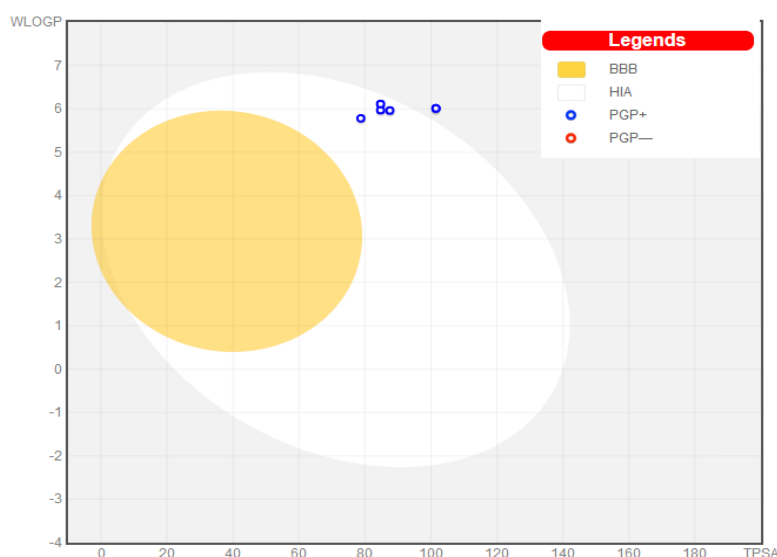


Fig.10. BOILED-Egg model

The study on the potential of compounds to inhibit the cytochrome P450 (CYP) enzymes is important in determining their possible drug interactions and toxicity.⁵⁴ Approximately over 50 % of therapeutic molecules are substrate of five major isoforms

(CYP1A2, CYP2C19, CYP2C9, CYP2D6, and CYP3A4). These enzymes are involved in metabolism of drugs.⁵⁵

Moreover, the compound design presented was found to be substrates of CYP1A2 and CYP2D6.

The compound Dis predicted not to be inhibitors of three of CYP isoenzymes. This fact is very useful, because this compound is expected not to have CYP metabolism interactions with other drugs, and this compound could present a reduced Hepatic toxicity risk. All compounds are found to be substrates of P-Gp, lipophilic substances of low molecular weight tend to be substrates for P-glycoprotein.⁵⁶

Table 5. Metabolism prediction for new inhibitors

Compound	P-Gp Substrata	CYP1A2 Inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 Inhibitor	CYP3A4 Inhibitor
13	yes	yes	yes	yes	no	yes
A	yes	no	yes	yes	no	yes
B	yes	no	yes	yes	no	yes
C	yes	no	yes	yes	no	yes
D	yes	no	yes	yes	no	no
E	yes	no	yes	yes	no	yes

4. CONCLUSION

In this study, CoMFA and CoMSIA 3D-QSAR models were developed for a series of thienopyrimidine derivatives that has antimicrobial potency against *Staphylococcus aureus*; the two models have good statistical results in terms of q^2 and R^2 values.

The good predictive ability of CoMFA and CoMSIA observed for the test set of compounds indicates that these models could be successfully used for predicting the pIC₅₀ values. Moreover, based on the contour's maps of the CoMFA/CoMSIA models, Steric, electrostatic and hydrophobic significant regions were identified to enhance bioactivity as well as H-bond interactions. Docking study was performed to analyze and identify the interactions of possible antimicrobial compounds (The best effective compound being

compound 13 and the weakest compound1) in the active site of DNA gyrase. These results provided crucial clues for designing novel *Staphylococcus aureus* antagonists with high predicted potent activity. A set of 5 novel derivatives were designed by utilizing the structure-activity relationship taken from the present study. In silico analyzes of absorption, distribution, metabolism and excretion were carried out on these new molecules to investigate their activities in compliance with the standard. These five novel lead molecules have better pharmacological properties compared to the study series. The information obtained from this study can further be used for the design of potent inhibitors of *S. aureus* DNA gyrase enzyme.

REFERENCES

1. D. Lowy. *N. Engl. J. Med.***1998**, 339, 520–532.**DOI:**10.1056/NEJM199808203390806.
2. G. Jagadeesan, V. Vijayakuma, M. Palayam, G. Suresh, G. Krishnaswamy, S. Aravindhana, G. H. Peters, *J. Proteomics Bioinform.***2015**, 8, 1–7.
DOI:10.4172/jpb.1000362.
3. M.M. El-Enany, F.S. El-Shafie, *Orient. J. Chem.***1989**, 5, 114-117.
DOI: 10.1016/j.ejmech.2010.08.044.
4. S.M. Ronkin, M. Badia, S. Bellon, A. L. Grillot, C.H. Gross, T.H. Grossman, N. Mani, J.D. Parsons, D. Stamos, M. Trudeau, Y. Wei, P.S. Charifson, *Bioorg. Med. Chem. Lett.***2010**, 20, 2828–2831.**DOI:** 10.1016/0006-2952(75)90415-3.
5. V.P. Litvinov, *Chem Inform*, **2006**, 92, 83-143.**DOI:** 10.1016/S0065-2725(06)92003-0.
6. Kharizomenova, A.N. Grinev, N.V. Samsonova, E.K. Panisheva, N.V. Kaplina, I.S. Nikolaeva, T.V. Punshkina, G.N. Pershin, *Pharm. Chem. J.* **1981**, 15, 645.
7. A.E. Rashad, A.H. Shamroukh, R.E. Abdel-Megeid, A. Mostafa, R. El-Shesheny, A. Kandeil, M.A. Ali, K. Banert, *Eur. J. Med. Chem.***2010**, 45, 5251-5257.
DOI: 10.1016/j.ejmech.2010.08.044.
8. V. Alagarsamy, S. Meena, K.V. Ramseshu, V.R. Solomon, K. Thirumurugan, K. Dhanabal, M. Murugan, *Eur. J. Med. Chem.* **2006**, 41, 1293-1300.
DOI: 10.1016/j.ejmech.2006.06.005.
9. A.B.A. El-Gazzar, H.A.R. Hussein, H.N. Hafez, *Acta. Pharm.***2007**, 57, 395-406.**DOI:** 10.2478/v10007-007-0032-6.

10. N.S. Habib, R. Soliman, K. Ismail, A.M. Hassan, M.T. Sarg, *Bull.Chim. Farm.* **2003**,142, 396-405. DOI: 10.2478/v10007-007-0032-6.
11. M.S.Tolba, M.Ahmed, A.M.Kamal El-Dean, R. Reda Hassanien, M. Farouk, J. *J. Heterocycl. Chem.***2017**, 55, 408-418. DOI: 10.1002/jhet.3056.
12. M.S.El-Shoukrofy, H.A. Abd El Razik, O.M.AboulWafa, A.E.Bayad, I.M. El-Ashmawy,*Bioorg Chem.* **2019**, 85, 541-557.DOI: 10.1016/j.bioorg.2019.02.036.
13. N.S. Shetty, R.S. Lamani, I.A.M. Khazi, *J. Chem. Sci.***2009**. 121, 301-307. DOI:10.1007/s12039-009-0034-7.
14. T. Horiuchi, J. Chiba, K. Uoto, T. Soga, *Bioorg. Med. Chem. Lett.* **2009**, 19, 305-308.DOI: 10.1016/j.bmcl.2008.11.090.
15. M. D. M. H. Bhuiyan, K. M. D. M. Rahman, M. D. K.Hossain, A.Rahim, M. I. Hossain, M. Abu Naser, *Acta Pharm.***2006**, 56, 441–450.
16. C. Sharma, S. Yerande, R.Chavan, A. V. Bhosale, *J. Chem.* **2010**, 7, 655-664. DOI: 10.1155/2010/369141.
17. H. M. Aly, N. M. Saleh, A. Elhady, *Eur.J.Med.Chem.***2011**, 46, 4566-4572. DOI: 10.1016/j.ejmech.2011.07.035.
18. N. Kerru, T. Settypalli, H. Nallapaneni, V.R. Chunduri, *Med. Chem.* **2014**, 4, 623-629. DOI: 10.4172/2161-0444.1000204
19. M. R. Mahmoud, F. S. Abu El-Azm, A. T. Ali, Y. M. Ali, *Synth. Commun.* **2015**, 45, 982 -992.DOI:10.1002/jhet.2824.
20. T. A. Mohamed, M. K.Abou-Elregal, A. S. A.; Youssef, M. M. Hemdan, S. S. Samir, W. S. I. Abou-Elmagd, *Synth. Commun.* **2020**, 50, 399-411. DOI: 10.1080/00397911.2019.1697822.
21. S. Ramamurthy, E. H. Jayachandran, *J. D. Med.* **2015**,7, 38-45. DOI: 10.15254/H.J.D.Med.7.2015.14.
22. J. Deng, Li Peng, G. Zhang, X. Lan, C. Li, F. Chen, Y. Zhou, Z. Lin, L. Chen, R. Dai, H. Xu, L. Yang, X. Zhang, W. Hu, *Eur. J. Med. Chem.* **2011**, 46, 71-76.DOI: 10.1016/j.ejmech.2018.03.041.
23. M.A. El-Sherbeny, M. B.El Ashmawy, H.I. El Subbagh, A.A. El Emam, F.A. Badria, *Eur. J. Med. Chem.***1995**, 30, 445-449. DOI: 10.1016/0223-5234(96)88255-9.
24. Y. Guo, J. Li, J. L. Ma, Z. Yu, H. Wang, W. Zhu, X. Liao, Y. Zhao, *Chin. Chem. Lett.* **2015**, 26, 755 -758. DOI: 10.1007/s11164-015-2088-0 41.

25. S. Kaizhen, M. Junjie, W. Xiao, G. Ping, Z. Yanfang, *Chem. Res. Chin. Univ.* **2014**, *30*, 75-81. DOI: 10.1016/j.cclet.2015.03.026 42.
26. W. Zhu, C. Chen, C. Sun, S. Xu, C. Wu, F. Lei, H. Xia, Q. Tu, P. Zheng, *Eur. J. Med. Chem.* **2015**, *93*, 64-73. DOI: 10.1016/j.ejmech.2015.01.061.
27. T. Becker, A. Sellmer, E. Eichhorn, H. Pongratz, C. Schächtele, F. Totzke, G. Kelter, R. Krumbach, H. Fiebig, F. Böhmer, S. Mhboobi, *Bioorg. Med. Chem. Lett.* **2012**, *20*, 125-136. DOI: 10.1016/j.bmc.2011.11.023.
28. Y. Ni, A. Gopalsamy, D. Cole, Y. Hu, R. Denny, M. Lpek, J. Liu, J. Lee, J.P. Hall, M. Luong, J.B. Telliez, L.L. Lin, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 5952-5965. DOI: 10.1016/j.bmcl.2011.07.069.
29. M. M. Kandeel, H. M. Rafaat, A. E. Kassab, I. G. Shahin, T. M. Abdelghany. *Eur. J. Med. Chem.* **2015**, *90*, 620-632. DOI: 10.1016/j.ejmech.2014.12.009.
30. S. I. Panchamukhi, A. K. Mohammed Iqbal, A. Y. Khan, M. B. Kalashetti, I. M. Khazi. *Pharm. Chem. J.* **2011**, *44*, 694–696. DOI: 10.1007/s11094-011-0545-7.
31. S.G. Li, C. Vilchèze, S. Chakraborty, X. Wang, H. Kim, M. Anisetti, S. Ekins, K. Y. Rhee, W. R. Jacobs, J. S. Freundlich, *Tetrahedron Lett.* **2015**, *56*, 3246–3250. DOI: 10.1016/j.cels.2015.12.005.
32. M.A. Lill. *Drug Discov Today.* **2007**, *12*, 1013–1017. DOI: 10.1016/j.drudis.2007.08.004.
33. M.B. Dewal, A.S. Wani, C. Vidaillac, D. Oupicky, M.J. Rybak, S.M. Firestine, *Eur J Med Chem.* **2012**, *51*, 145-153. DOI: 10.1016/j.ejmech.2012.02.035.
34. SYBYL-X 2.0. St. Louis, MO, USA: Tripos Inc; Available from: <http://www.tripos.com>.
35. W.P. Purcell, J.A. Singer. *J. Chem. Eng. Data.* **1967**, *12*, 235-246. DOI: 10.1021/je60033a020.
36. J. G. Wilkes, I. B. Stoyanova-Slavova, D. A. Buzatu, *J. Comput. Aid. Mol. Des.* **2016**, *30*, 331-345. DOI: 10.1007/s10822-016-9909-0.
37. R.D. Cramer, D.E. Patterson, J.D. Bunce. *J. Am. Chem. Soc.* **1988**, *110*, 5959-5967. DOI: 10.1021/ja00226a005.
38. G. Klebe, U. Abraham, T. Mietzner, *J. Med. Chem.* **1994**, *37*, 4130–4146. DOI: 10.1021/jm00050a010.
39. P. Liu, W. Long. *Int. J. Mol. Sci.* **2009**, *10*, 1978–1998. DOI: 10.3390/ijms1005197
40. A. Golbraikh, A. Tropsha, *J. Mol. Graph. Model.* **2002**, *16*, 357–369. DOI: 10.1016/S1093-3263(01)00123-1.

41. Discovery Studio Visualizer. San Diego: Dassault Systèmes BIOVIA; **2019**.
Available from: <http://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/visualization-download.php>.
42. Y. Ai, S.-T. Wang, P.-H. Sun, F.J. Song. *Int. J. Mol. Sci.***2010**, *11*, 3705–3724.
DOI:10.3390/ijms11103705A.
43. Daina, O. Michielin, V. Zoete. *Sci. Rep.***2017**,*7*, 42717. **DOI:** 10.1038/srep42717.
44. C.A. Lipinski, F. Lombardo, B.W. Dominy, P. J. Feeney, *Adv. Drug Deliv. Rev.*
2001, *46*, 3–26. **DOI:** 10.1016/s0169-409x(00)00129-0.
45. D.F. Veber, S.R. Johnson, H. Y. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, *J. Med. Chem.***2002**, *45*, 2615–2623. **DOI:**10.1021/jm020017n.
46. A. Cherkasov, E.N. Muratov, D. Fourches, A. Varnek, I.I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y.C. Martin, R. Todeschini, V. Consonni, V.E. Kuz'min, R. Cramer, R. Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard, A. Tropsha, *J. Med. Chem.***2014**, *57*, 4977–5010. **DOI:** 10.1021/jm4004285.
47. M. Ouassaf, S. Belaidi, K. Lotfy, I. Daoud, H. Belaidi, *J. Bionanosci.* **2018**, *12*, 1-11. **DOI:**10.1166/jbns.2018.1505.
48. B. Berk, G. Kaynar, M. Ertaş, S.N. Biltekin. *Acta Pharm. Sci.***2017**, *55*, 97-117. **DOI:** 10.23893/1307-2080.APS.0551.
49. M. Ouassaf, S. Belaidi, H. Belaidi, Z. Almi, *J. Fundam. Appl. Sci.* **2018**, *10*, 500-524. **DOI:** 10.4314/jfas. v10i3.33.
50. J.B. Baell, G.A. Holloway, *J. Med. Chem.***2010**, *53*, 2719–2740. **DOI:** 10.1021/jm901137j.
51. P. Ertl, A. Schuffenhauer, *J. Cheminformatics.* **2019**, *82*, 1258-1263. **DOI:** 10.1021/acs.jnatprod.8b01022.
52. M. R. Prausnitz, R. Langer, *Nat. Biotechnol.* **2008**, *26*, 1261–1268.
DOI:10.1038/nbt.1504.
53. A. Daina, V. Zoete, *Chem. Med. Chem.***2016**, *11*, 1117–1121.
DOI:10.1002/cmdc.201600182.
54. P.F. Hollenberg, *Drug Metab. Rev.***2002**, *34*, 17–35. **DOI:**10.1081/dmr-120001387.
55. L. Di. *Expert. Opin. Drug Metab. Toxicol.* **2014**, *10*, 379–393.
DOI:10.1517/17425255.2014.876006.
56. D.J. Begley, *Curr. Pharm. Des.* **2004**, *10*, 1295–1312.
DOI:10.2174/1381612043384844.

618

619