

Research Article

2D-QSAR Study of Indolylpyrimidines Derivative as Antibacterial against *Pseudomonas aeruginosa* and *Staphylococcus aureus*: A Comparative Approach

Prasanna A. Datar

Department of Pharmaceutical Chemistry, Sinhgad Institute of Pharmacy, Narhe, Pune 411041, India

Correspondence should be addressed to Prasanna A. Datar; d_pras_anna@rediffmail.com

Received 6 February 2014; Revised 26 April 2014; Accepted 26 April 2014; Published 5 June 2014

Academic Editor: Gabriela Mustata Wilson

Copyright © 2014 Prasanna A. Datar. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A set of 15 indolylpyrimidine derivatives with their antibacterial activities in terms of minimum inhibitory concentration against the gram-negative bacteria *Pseudomonas aeruginosa* and gram-positive *Staphylococcus aureus* were selected for 2D quantitative structure activity relationship (QSAR) analysis. QSAR was performed using a combination of various descriptors such as steric, electronic and topological. Stepwise regression method was used to derive the most significant QSAR equation for predicting the inhibitory activity of this class of molecules. The best QSAR model was further validated by a leave one out technique as well as by the random trials. A high correlation between experimental and predicted inhibitory values was observed. A comparative picture of behavior of indolylpyrimidines against both of the microorganisms is discussed.

1. Introduction

Pseudomonas aeruginosa (PA), a gram-negative pathogen, has been known as a major cause of hospital acquired infection and antimicrobial resistance [1, 2]. *Pseudomonas aeruginosa* is responsible for various infectious cases such as nosocomial pneumonia urinary tract infections, surgical wound infections, and bloodstream infections [3]. Structural differences exist between cell walls of gram-positive and gram-negative bacteria. Gram-positive bacteria have more peptidoglycan layers as compared to gram-negative bacteria. Therefore the cell wall of gram-positive bacteria is thicker than the cell wall of gram-negative bacteria.

Gram-negative bacterial cell wall is different from gram-positive bacterial cell wall by having an outer membrane of lipoproteins that covers the peptidoglycan layer. The outer membrane of gram-negative bacteria is made up of phospholipids, lipoproteins, and lipopolysaccharides. The outer membrane is negatively charged and helps prevent the bacteria from being phagocytosed by macrophages. The outer membrane provides protection from effects of antibiotics, digestive enzymes, and heavy metals.

Many chemical entities work as antibacterial agents by inhibiting the DNA synthesis of cell wall by blocking the enzymes such as DNA gyrase and dihydrofolate reductase and even inhibiting enzymes processing the development of peptidoglycan layer [4, 5].

Approach of antibacterial drug is initially a surface phenomenon. The wall of gram-positive and gram-negative bacteria will resist the surface interaction. Therefore there is a difference in antibacterial activity of chemical entity towards gram-positive and gram-negative bacteria. In the present work, we have made an attempt to differentiate the behavior of chemical entities towards gram-positive and gram-negative bacteria using QSAR molecular descriptors which explains surface phenomenon. This would have been possible to express in terms of Hammett parameters and pKa values of the compounds as descriptors, but our approach gives emphasis on calculated surface parameter to express the antibacterial activity. There are many equations in the literature expressing antibacterial activity using Hammett and pKa values but these equations are insignificant in virtual screening of molecular databases for finding significant antibacterial hits.

The objective of the present investigation was to study the usefulness of QSAR in the prediction of the antibacterial activity of indolylpyrimidine derivatives against *Pseudomonas aeruginosa* (PA) and *Staphylococcus aureus* (SA) and understand how multiple linear regression (MLR) equations can explain the structural key points correlating to differential behavior in activity against both gram-positive and gram-negative strains.

The pharmaceutical importance of pyrimidine compounds lies in the fact that they can be effectively used as analgesic, anti-inflammatory, anticonvulsant, insecticidal, herbicidal, antitubercular, anticancer and antidiabetic agents. The indole ring is known to exhibit anti-inflammatory, antimicrobial and antifungal activities [6–10]. The fused ring system of substituted indolylpyrimidines is remarkably effective as antitumor and antibacterial activity [11, 12]. The QSAR method requires data collection from the same laboratory experiment, molecular descriptor selection, QSAR model development and finally model validation. A QSAR study has predictive ability and even provides clues for mechanism of drug receptor interactions [13, 14].

2. Biological Activity

Very few research articles are available from the same laboratory on indolylpyrimidines as antibacterial agents. Indolylpyrimidines tested were obtained from the studies reported by Panda and Chowdhary [15] where *in vitro* antibacterial activity against *Pseudomonas aeruginosa* and *Staphylococcus aureus* was carried out at uniform concentration of 100 $\mu\text{g}/\text{mL}$ and zone of inhibition is reported in distance unit, that is, millimeters (mm). The values are directly proportional since the more the distance in mm for zone of inhibition is, the more the potency would be. The distance in mm is used as dependent variable in QSAR study. Zone of inhibition against *Pseudomonas aeruginosa* (PAantibact) varies from 10 mm to 24 mm while, for *Staphylococcus aureus*, it varies from 13 mm to 31 mm (SAantibact). Figures 1 and 2 give the activity distribution plot for antibacterial activity against *Pseudomonas aeruginosa* and *Staphylococcus aureus*, respectively. The plot shows blue columns representing training set while brown columns represent test set distribution.

3. Data Sets

Due to constraints of data selection based on the same laboratory experimental studies, study with exact numerical values was selected for QSAR study. According to the variation in substituents at various positions in a set of indolylpyrimidine derivatives, molecules were divided into the training set and test set using sphere exclusion method. The training set comprised 15 derivatives (4a, 4c, 4e, 4g, 4i, 5c, 5f, 5g, 5h, 5j, 6a, 6b, 6d, 6i, and 6j) while the and test set comprised 9 derivatives (4d, 4f, 4h, 4j, 5a, 5d, 6c, 6e, 6f, and 6g) of indolylpyrimidines published by Panda and Chowdhary [15], which have been shown to possess antibacterial activity against *Pseudomonas aeruginosa*.

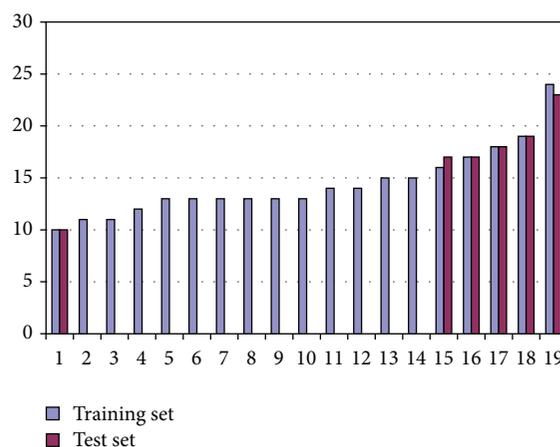


FIGURE 1: Activity distribution plot for antibacterial activity against PA.

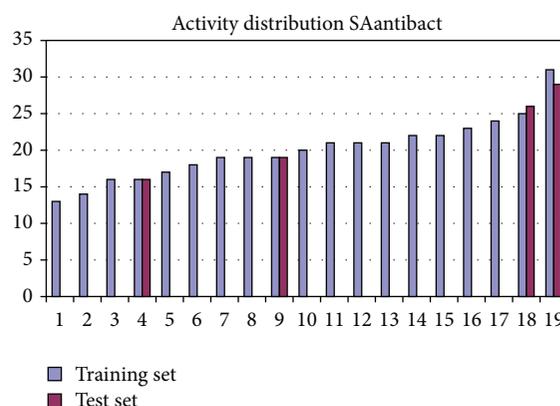


FIGURE 2: Activity distribution plot for antibacterial activity against SA.

The case of QSAR model studied for activity against *Staphylococcus aureus* has a training set comprising 14 molecules (4a, 4c, 4d, 4f, 5a, 5b, 5d, 5g, 5h, 6b, 6c, 6f, 6i, 6j) and test set comprising 6 molecules (4b, 5c, 5f, 6a, 6d, 6h). The test compounds were selected manually considering the activity distribution and structural diversity as compared to the training set. Chemical structures of indolylpyrimidines and corresponding biological data are shown in Table 1.

4. Molecular Modeling

All molecular modeling studies were performed by using Schrodinger software running on windows platform. Ligprep module was used to draw the 3D structures. The 3D structures were further cleaned up and subjected to energy minimization using conjugate gradient method using MMFF force field. The minimization was performed until the RMS gradient value reaches a value smaller than 0.1 kcal/mol $^{\circ}\text{A}$. Optimization was further performed using BFGS method until the RMS gradient attains a value smaller than 0.0001 kcal/mol $^{\circ}\text{A}$. The lowest energy structure of each molecule was subjected to calculate molecular descriptors.

TABLE 1: Structures of indolylpyrimidines and their antibacterial activity used for the QSAR studies.

Compound number	R ₁	R ₂	Zone of inhibition against <i>Pseudomonas aeruginosa</i> in mm (PAantibact)	Zone of inhibition against <i>Staphylococcus aureus</i> in mm (SAantibact)
4a	OH	H	13	16
4b	OH	p-NH ₂	24	26
4c	OH	p-Br	19	20
4d	OH	p-Cl	18	19
4e	OH	o,p-OH	14	14
4f	OH	p-F	17	18
4g	OH	p-CH ₃	13	10
4h	OH	p-OCH ₃	17	21
4i	OH	p-OH	11	16
4j	OH	p-NO ₂	23	22
5a	SH	H	12	13
5b	SH	p-NH ₂	18	19
5c	SH	p-Br	15	16
5d	SH	p-Cl	16	17
5f	SH	p-F	14	19
5g	SH	p-CH ₃	10	21
5h	SH	p-OCH ₃	13	25
5j	SH	p-NO ₂	19	16
6a	NH ₂	H	11	10
6b	NH ₂	p-NH ₂	10	24
6c	NH ₂	p-Br	17	23
6d	NH ₂	p-Cl	15	19
6e	NH ₂	o,p-OH	13	16
6f	NH ₂	p-F	23	22
6g	NH ₂	p-CH ₃	13	14
6h	NH ₂	p-OCH ₃	25	29
6i	NH ₂	p-OH	13	21
6j	NH ₂	p-NO ₂	24	31

5. Descriptor Generation

The numerical descriptors are responsible for encoding information of important features of the molecular structure and can be categorized into different properties such as electronic, geometric, hydrophobic and topological. Various molecular descriptors were calculated such as molecular weight, dipole moment and partition coefficient (ClogP), surface

area descriptors, H-bond donor count, H-bond acceptor count, ionization potential and electron affinity. Among the set of equations, generated descriptors that explain surface phenomenon were retained as shown in Table 2. Biological activity and descriptor values were scaled to unit variance as shown in Table 3.

Pearson's correlation matrix was used to select the suitable descriptors for MLR analysis. Pearson's correlation matrix

TABLE 2: Description of descriptors used in QSAR study.

Descriptor name	Description
mol_MW	Molecular weight of the molecule.
dipole	Computed dipole moment of the molecule.
EA (eV)	PM3 calculated electron affinity.
donorHB	Estimated number of hydrogen bonds that would be donated by the solute to water molecules in an aqueous solution. Values are averages taken over a number of configurations, so they can be noninteger.
PSA	Van der Waals surface area of polar nitrogen and oxygen atoms.
SASA	Total solvent accessible surface area (SASA) in square angstroms using a probe with a 1.4 Å radius.
PISA	π (carbon and attached hydrogen) component of the SASA.
WPSA	Weakly polar component of the SASA (halogens, P and S).
FISA	Hydrophilic component of the SASA (SASA on N, O, and H on heteroatoms).
FOSA	Hydrophobic component of the SASA (saturated carbon and attached hydrogen).

TABLE 3: Descriptors used in 2D QSAR studies.

Compound number	mol_MW	dipole	FOSA	FISA	PISA	WPSA	donorHB	EA (eV)	PSA
4a	287.3	5.2	0	109.3	423.9	0	2	0.631	57.9
4b	302.3	3.4	0	170.7	377.8	0	3.5	0.394	83.9
4c	366.2	6.9	0	109.1	375.9	77.9	2	0.747	57.9
4d	321.7	7.3	0	109.3	376.3	72.5	2	0.784	57.9
4e	319.3	4.3	0	201.0	354.2	0	4	0.428	98.5
4f	305.3	7.4	0	109.3	386.7	47.2	2	0.804	57.9
4g	301.3	5.0	89.4	109.3	368.1	0	2	0.608	57.9
4h	317.3	5.9	93.0	111.3	365.1	0	2	0.693	67.8
4i	303.3	5.1	0	164.7	382.2	0	3	0.605	80.6
4j	332.3	14.3	0	207.5	365.3	0	2	1.852	103.0
5a	303.3	4.2	0	47.1	420.0	81.2	1.8	0.721	35.2
5b	318.3	3.2	0	108.6	373.9	81.2	3.3	0.47	61.2
5c	382.2	5.7	0	46.9	371.9	159.2	1.8	0.836	35.2
5d	337.8	6.1	0	47.2	372.3	153.8	1.8	0.874	35.2
5e	335.3	3.6	0	139.3	350.5	79.9	3.8	0.507	75.8
5f	321.3	6.3	0	47.1	382.7	128.5	1.8	0.894	35.2
5g	317.4	3.9	89.4	47.2	364.2	81.2	1.8	0.696	35.2
5h	333.4	4.9	93.0	47.1	361.2	81.2	1.8	0.721	44.1
5i	319.3	4.0	0	102.6	378.2	81.2	2.8	0.692	57.9
5j	348.3	13.1	0	145.3	361.4	81.2	1.8	1.903	80.4
6a	286.3	4.3	0	113.8	422.3	0	3	0.388	60.7
6b	301.3	1.8	0	175.3	376.1	0	4.5	0.166	86.7
6c	365.2	6.3	0	113.6	374.2	77.9	3	0.511	60.7
6d	320.7	6.7	0	113.8	374.6	72.5	3	0.555	60.7
6e	318.3	5.4	0	205.3	352.7	0	5	0.204	101.3
6f	304.3	6.7	0	113.8	385.0	47.2	3	0.572	60.7
6g	300.3	3.9	89.4	113.9	366.4	0	3	0.369	60.7
6h	317.3	5.9	93.0	111.3	365.1	0	2	0.693	67.8
6i	302.3	5.4	0	169.3	380.5	0	4	0.365	83.4
6j	331.3	13.7	0	212.0	363.7	0	3	1.716	105.8

was performed on all descriptors by using “BuildQSAR” module available in Schrodinger software. Tables 4 and 5 show correlation matrix for the descriptors used in the resulting models for *Pseudomonas aeruginosa* and *Staphylococcus aureus*, respectively. The descriptors correlated above 0.5 were

eliminated from the QSAR study. The correlated descriptor that duplicates the meaning of other descriptors was eliminated. ClogP was found to be correlated to FISA by 0.94 and to FOSA by 0.46, respectively. Correlation matrix shows that ClogP correlates well with FISA and PSA (Tables 4 and 5).

TABLE 4: Pearson’s correlation matrix for descriptors in QSAR model of *Pseudomonas aeruginosa*.

	PAantibact	Dipole	mol_MW	donorHB	FISA	PSA	ClogP
PAantibact	1.0	—	—	—	—	—	—
Dipole	0.803	1.0	—	—	—	—	—
mol_MW	0.321	0.170	1.0	—	—	—	—
donorHB	0.037	0.075	0.186	1.0	—	—	—
FISA	0.090	0.076	0.080	0.581	1.0	—	—
PSA	0.122	0.115	0.053	0.536	0.979	1.0	—
ClogP	-0.013	0.013	0.510	-0.898	-0.93	-0.914	1.0

TABLE 5: Pearson’s correlation matrix for descriptors in QSAR model of *Staphylococcus aureus*.

	SAantibact	Dipole	EA (eV)	PSA	PISA	WPSA	FOSA	ClogP
SAantibact	1.0	—	—	—	—	—	—	—
Dipole	0.200	1.0	—	—	—	—	—	—
EA (eV)	0.175	0.016	1.0	—	—	—	—	—
PSA	0.328	0.810	0.087	1.0	—	—	—	—
PISA	0.618	0.065	0.084	0.038	1.0	—	—	—
WPSA	0.055	0.065	0.001	0.506	0.080	1.0	—	—
FOSA	0.182	0.032	0.001	0.072	0.212	0.001	1.0	—
ClogP	-0.488	-0.159	0.113	-0.933	0.010	0.855	0.468	1.0

Since ClogP is unable to express charge of a molecule, therefore descriptors based on solvent accessible surface area such as FISA and FOSA were used to overcome this limitation [16, 17].

Linear QSAR equations were developed by a stepwise addition of terms. Each descriptor was chosen as input for variable selection method such as stepwise addition method. The method selects the descriptor that contributes to the antibacterial activity of indolylpyrimidine derivatives. To reduce the variation in the biological data, stepwise equations were generated using autoscaling of the dataset.

MLR method only can be used when a relatively small number of molecular descriptors are used. The ratio of descriptor to number of molecules in a QSAR equation generated was kept as 1 : 5. Thus for training set of 15 molecules, 3 descriptors were optimal. However, PCR and PLS would also be used which allow using more descriptors, but too many descriptors may cause difficulties in model interpretability. Besides, using several factors (principal components or latent variables) can make model interpretation tedious. Validation is a crucial aspect of any QSAR analysis [18, 19]. Predictive power of selected MLR equation is validated by the live one out (LOO) technique. The QSAR equations were validated by the calculation of the following statistical parameters: CV $r^2(q^2)$, r^2 , ran_r^2 , and std_error (Tables 6a and 6b).

Randomization was performed by randomly shuffling the dependent parameter and then generating the equation using MLR method for the same set of descriptors and training set molecule. The resulting r^2 is denoted by ran_r^2 . If ran_r^2 is consistent and equivalent to r^2 value of the QSAR model, then the QSAR equation is spurious. Consider

$$\text{CV } r^2 = \frac{1 - \text{Press}}{\text{SSY}}, \quad (1)$$

where,

$$\text{Press} = \sum (Y_{\text{obs}} - Y_{\text{calc}})^2, \quad (2)$$

$$\text{SSY} = \sum (Y_{\text{obs}} - Y_{\text{mean}})^2, \quad (3)$$

where Y_{obs} , Y_{calc} , and Y_{mean} are observed, calculated, and mean values, respectively, “ n ” is the number of compounds, and “ p ” is the number of independent parameters. The value PRESS is indicated by prediction sum of squares. The PRESS value can be used to calculate CV $r^2(q^2)$, called cross-validated r^2 , which represents the prediction ability of the QSAR equation. This is a good way to validate the prediction of a regression equation. The CV $r^2(q^2)$ value ranges from zero to one. To calculate PRESS, each observation is individually omitted once. The remaining $n - 1$ observations are used to calculate a regression equation and value of the omitted observation is estimated. This is done “ n ” times, once for each observation. The difference between the actual Y value (Y_{obs}) and the predicted Y value (Y_{calc}) is called the prediction error. The sum of the squared prediction errors is indicated as PRESS value.

A compound is said to be outlier whose activity cannot be predicted by generated QSAR equation. The structural diversity of these molecules is responsible for their nonpredictability. Equation 1 in Table 6(a) gives outliers, which can be identified by Z score value [20]. Z score value is calculated by the following formula:

$$Z \text{ score} = \frac{(x - x_{\text{mean}})}{s}, \quad (4)$$

where initially the mean is subtracted from every value and then the mean-shifted values are divided by the standard

TABLE 6: Best MLR models for the prediction of antibacterial activity.

(a) Stepwise regression result for antibacterial activity against PA								
Eq. number	Equation	N	r^2	CV r^2 (q^2)	Std_error (s)	Pred. r^2	Ran_ r^2	Z score r^2
1	PAantibact = (0.8455) dipole + (0.0464) mol.MW + (0.0135) FISA - 7.5189	15	0.879	0.709	0.499	0.399	0.264	4.938
1.1	PAantibact = (1.0249) dipole + (0.0538) mol.MW - (0.0555) SASA + 21.7974	15	0.878	0.673	1.504	0.518	0.391	1.876
(b) Stepwise regression result for antibacterial activity against SA								
Eq. number	Equation	N	r^2	CV r^2 (q^2)	Std_error (s)	Pred. r^2	Ran_ r^2	Z score r^2
2	SAantibact = -0.2246 PISA - 0.0595 WPSA + 2.9410 EA (eV) + 107.5483	15	0.874	0.767	0.222	1.469	0.356	3.938
2.1	SAantibact = -(0.1343) PISA + (0.1660) PSA + (0.0494) FOSA + 61.1447	15	0.876	0.743	2.210	1.544	0.288	3.461

deviation (s). Z score is a value that estimates in terms of the number of standard deviations the value above or below the mean of a data set.

6. Results

In QSAR equation development for *Pseudomonas aeruginosa*, the regression equation indicated dipole moment as the most significant in contribution to inhibitory activity (Equation 1, Table 6(a)) since it has the highest correlation with the activity. Molecular weight (mol.MW) is an additional parameter to dipole, which significantly increases the correlation coefficient from 0.7910 to 0.8587. The selection of variable was such that it minimizes the mean squared error of prediction. Similarly, the addition of a third parameter, FISA, also increased the correlation coefficient from 0.8587 to 0.8793. Other regression equations were also obtained just by altering third parameter FISA with donorHB or PSA and so forth. Tables 6(a) and 6(b) give the best equations (1 and 2) selected in comparison to 1.1 and 2.1, respectively, since the latter ones are having high standard error (s) value and therefore are insignificant.

In Table 6(a), Equation 1, all the descriptors are directly proportional to the activity and it is highly correlated to dipole more than molecular weight. Similarly, in case of QSAR equation developed for activity against *Staphylococcus aureus* is shown in Table 6(b), PISA is a prominent parameter. Successive addition of WPSA or PSA and the third parameter EA (eV) enhanced QSAR equation significantly. N is the number of molecules in training set, r^2 is squared correlation coefficient and F -test is a variance-related statistical value that compares two models differing by one or more variable. The QSAR equation is supposed to be good if the F -test is above a threshold value. The statistical quality of the resulting models, as depicted in Table 6(a), is determined by r , standard error (std_error), and randomization test (ran_ r^2) [21, 22].

A data point was considered to be an outlier if its residual value exceeded two times the value of standard error of estimate of the model. It is noteworthy that all these equations were derived using the entire data set of compounds ($n = 28$) and the outliers were identified for both of the QSAR equations. For QSAR equation of *Pseudomonas aeruginosa*, outliers were obtained as 5b, 6h with Z score value of 4.938. While QSAR equation of activity against *Staphylococcus aureus* had outliers as 4e, 4g, 5j, 6e, 6g with Z score value of 3.938. The reason for the molecules being found as outliers was investigated. It was found that the descriptor value for these molecules was away from the range of descriptor values of the training set molecules [23]. Removal of these outliers has improved the statistics of the equations.

The best measure of reliability of a 2D-QSAR model is a high q^2 , not just a high r^2 that could be a result of overfitting to data. More often, a value of $q^2 > 0.5$ is considered acceptable [24–26]. Self-consistency of the derived models was verified using the leave one out (LOO) process and the predictability of each model was assessed using cross-validated r^2 , called q^2 .

Figures 3 and 4 shows plots of predicted versus experimentally observed inhibitory activity for training set and test set of indolylpyrimidines against PA respectively. While Figures 5 and 6 shows plots of predicted versus experimentally observed inhibitory activity for training set and test set of indolylpyrimidines against SA respectively. The plots for QSAR models of PA and SA show a very good fit in the range of $r^2 = 0.87$ to 0.80. It indicates that these models can be successfully applied to predict the antibacterial activity of this class of molecules. Randomization studies show ran_ r^2 of 0.2 and 0.3 for Equation 1 (Table 6(a)) and 2 (Table 6(b)), respectively; thus the equations are not chance correlations. Moreover, it was possible to use the reported QSAR models to predict the activity of analogous molecules for antibacterial activity against PA and SA. The

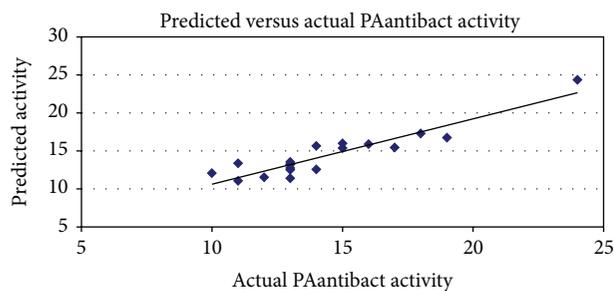


FIGURE 3: Plots of predicted versus experimentally observed inhibitory activity for training set of indolylpyrimidines against PA.

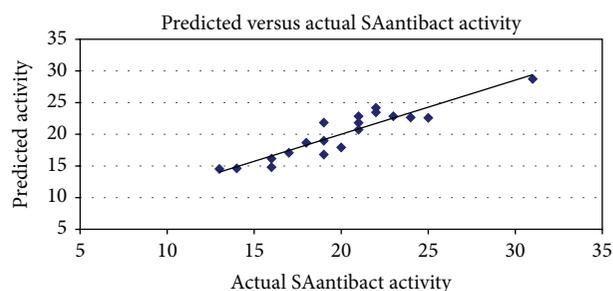


FIGURE 5: Plot of predicted versus experimentally observed inhibitory activity for training set of indolylpyrimidines against SA.

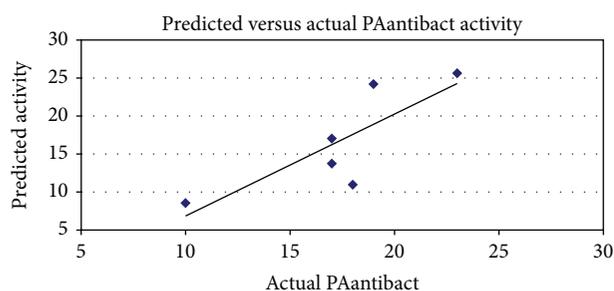


FIGURE 4: Plot of predicted versus experimentally observed inhibitory activity for test set of indolylpyrimidines against PA.

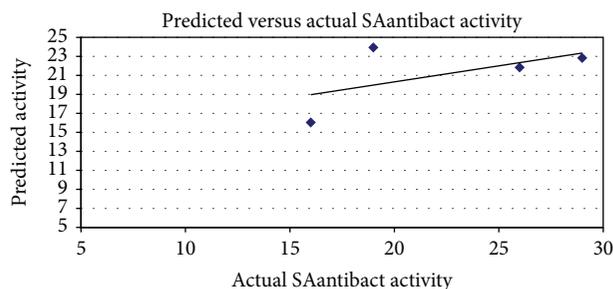


FIGURE 6: Plot of predicted versus experimentally observed inhibitory activity for test set of indolylpyrimidines against SA.

applicability domain of the derived QSAR models is restricted to substituted indolylpyrimidine derivatives.

7. Discussion

It clearly showed that QSAR Equation 1 in Table 6(a) obtained for antibacterial activity against PA includes dipole, mol.MW, and third descriptor FISA. Thus the equation is a combination of size of molecule, polar nature, and hydrophilicity of molecule as shown in Table 6(a). While QSAR equation obtained for antibacterial activity against SA includes PISA, PSA, and third descriptor FOSA. Thus QSAR Equation 2 in Table 6(b) shows combination of hydrophobicity and electron density centers and weakly polar groups (halogens P and S).

Since both of the activities were conducted in the same laboratory and at the same concentration of drug, that is, 100 $\mu\text{g}/\text{mL}$, the comparisons with respect to activity were possible.

Equation 1 in Table 6(a) gives descriptors for PA activity such as dipole, molecular weight, and FISA. In QSAR Equation 1 of Table 6(a), all the descriptors are directly proportional to the activity. Dipole is highly correlated with the activity followed by molecular weight. It is observed that the higher the dipole and the higher the PA activity, the higher the FISA activity and the higher the PA activity and molecular weight to be in the range of 320 to 332. Overall, this equation gives the relationship of polarity of a molecule as the essential characteristics required for the antibacterial activity against PA.

In QSAR Equation 2 which is developed for SA activity, as shown in Table 6(b), PISA is a prominent parameter. Successive addition of WPSA and the third parameter EA (eV) enhanced the model significantly. In interpretation of SA activity, the following descriptors were found to have general observations. PISA is a π (carbon and attached hydrogen) component of the SASA. The lower the PISA value the higher the SA activity. The lower value of PISA is due to presence of electronegative groups the higher SA activity is. Higher value of PISA is due to presence of protons connected to C, N, and O and has more influence at the R_2 position. WPSA is a weakly polar component of the SASA (halogens P and S). Sulfhydryl group and halogens are not essential but impart moderate activity. Among halogens, activity diminishes at R_2 as $\text{Br} > \text{F}$ in the presence of S. Electron acceptor is an atom that has a more positive value of electron affinity and the electron donor atom has less positive electron affinity. Electron affinity follows the trend of electronegativity. Fluorine (F) has a higher electron affinity than oxygen and chlorine most strongly attracts extra electrons. Atoms whose anions are stabler than neutral atoms have a greater electron affinity. Electron withdrawing group at R_2 is having more electron affinity and electron donating groups have moderate electron affinity which seems to be responsible for high to moderate activity, respectively. The activity would be higher for a combination of anion and proton donor at R_1 or R_2 position or vice versa, respectively. Overall, this equation gives the relationship of hydrophobicity of a molecule as essential characteristics required for the antibacterial activity against SA. Compounds 4b, 4j, 6f, 6h, and 6j were active at both the gram-positive and gram-negative bacterial strains.

This can be explained by using both of the QSAR equations. The compounds have high electron affinity and high dipole along with high values of FISA for activity against PA. The molecules have sufficient PISA contribution and low values of WPSA. Any extension to substituted phenyl by bulky groups would be favorable for activity against SA.

Polar surface area (PSA) descriptor represents drug transport properties such as intestinal absorption and blood-brain barrier penetration [27]. It is the sum of the contributions to the molecular (usually van der Waals) surface area of polar atoms such as oxygen, nitrogen, and their attached hydrogens. The polarity and polarizability of a molecule have been well known to be important for description of various physicochemical properties and chemical reactivity of molecule. Molecular polarity accounts for chromatographic retention on a polar stationary phase [28]. Most often, dipole moment is used which reflects only global polarity of molecule. Local polarities can be represented by atomic charges in the localized regions of the molecule.

8. Conclusions

The 2D-QSAR study on a series of indolylpyrimidine compounds showed that the presence of functional groups that balance the dipole and hydrophobicity would lead to increase in the activity of indolylpyrimidines against both PA and SA. The regression equation indicates that presence of polar components weighs more than hydrophobicity in the R₂ position of indolylpyrimidines for activity against *Pseudomonas aeruginosa*, while absence of protonated electronegative bulky groups at R₂ is suitable for antibacterial activity against *Staphylococcus aureus*.

Abbreviations

PA:	<i>Pseudomonas aeruginosa</i>
SA:	<i>Staphylococcus aureus</i>
MLR:	Multiple linear regression
QSAR:	Quantitative structure activity relationship
MMFF:	Merck molecular force field
RMS:	Root mean square
LOO:	Leave one out.

Conflict of Interests

The author declares no competing financial interests regarding the publication of this paper.

Acknowledgments

The author thanks his father, Mr. Anand Datar, and his guide, Dr. E. C. Coutinho, for providing facilities as well as continuous encouragement during this work.

References

- [1] L. Nicoll, "Pseudomonas aeruginosa: infections and treatment," *New England Journal of Medicine*, vol. 332, pp. 616–617, 1995.
- [2] J. Vila and T. Pal, "Update on antibacterial resistance in low-income countries: factors favoring the emergence of resistance," *Open Infectious Diseases Journal*, vol. 4, no. 1, pp. 38–54, 2010.
- [3] C. Van Delden and B. H. Iglewski, "Cell-to-cell signaling and *Pseudomonas aeruginosa* infections," *Emerging Infectious Diseases*, vol. 4, no. 4, pp. 551–560, 1998.
- [4] P. S. Charifson, T. H. Grossman, and P. Mueller, "The use of structure-guided design to discover new anti-microbial agents: focus on antibacterial resistance," *Anti-Infective Agents in Medicinal Chemistry*, vol. 8, no. 1, pp. 73–86, 2009.
- [5] J. Vila, J. Sánchez-Céspedes, and E. Giralt, "Old and new strategies for the discovery of antibacterial agents," *Current Medicinal Chemistry: Anti-Infective Agents*, vol. 4, no. 4, pp. 337–353, 2005.
- [6] H. D. Patel, B. D. Mistry, and K. R. Desai, "Synthesis and antimicrobial activity of pyrazolo [3,4-d] pyrimidines," *Indian Journal of Heterocyclic Chemistry*, vol. 13, no. 2, pp. 179–180, 2003.
- [7] A. Kreutzberger and M. Sellheim, "Antimycotic agents. XIX. [1,2]. 4,6-disubstituted 2-(cyanamino)pyrimidine derivatives," *Journal of Heterocyclic Chemistry*, vol. 22, no. 3, pp. 721–723, 1985.
- [8] V. Lather and P. V. Chowdary, "Synthesis and antimicrobial activity of N1-(arylidine hydrazidomethyl)-indoles, 2-(substituted aryl)-3-(N1-indolyl acetamidyl)-4-oxo-thiazolidines and 5-benzylidene derivatives of thiazolidinones," *Indian Journal of Pharmaceutical Sciences*, vol. 65, no. 6, pp. 576–579, 2003.
- [9] G. S. Gadaginamath, A. S. Shyadligeri, and R. R. Kavali, "Chemoselectivity of indole-dicarboxylates towards hydrazine hydrate: part III-synthesis and antimicrobial activity of novel 4-thiazolidinonylindoles," *Indian Journal of Chemistry B: Organic and Medicinal Chemistry*, vol. 38, no. 2, pp. 156–159, 1999.
- [10] P. Renukadevi and J. S. Biradar, "Synthesis and antimicrobial activity of 3,5-disubstituted-2-[1'-phenyl-5'-thioalkyl-s-triazol-2'-yl] indoles and 3,5-disubstituted-2-[1'-substituted aminomethyl-4'-phenyl-5' (4'h)-thione-s-triazol-3-yl] indoles," *Indian Journal of Heterocyclic Chemistry*, vol. 9, no. 2, pp. 107–112, 1999.
- [11] B. Jiang, C. Yang, W. Xiong, and J. Wang, "Synthesis and cytotoxicity evaluation of novel indolylpyrimidines and indolylpyrazines as potential antitumor agents," *Bioorganic and Medicinal Chemistry*, vol. 9, no. 5, pp. 1149–1154, 2001.
- [12] M. A. A. Radwan and M. El-Sherbiny, "Synthesis and antitumor activity of indolylpyrimidines: marine natural product meridianin D analogues," *Bioorganic and Medicinal Chemistry*, vol. 15, no. 3, pp. 1206–1211, 2007.
- [13] C. Hansch, A. Leo, and D. H. Hoekman, *Exploring QSAR, Hydrophobic, Electronic and Steric Constants*, American Chemical Society, Washington, DC, USA, 1995.
- [14] C. Hansch, A. Leo, and D. H. Hoekman, *Exploring QSAR, Fundamentals and Application in Chemistry and Biology*, American Chemical Society, Washington, DC, USA, 1995.
- [15] S. Panda and P. V. R. Chowdary, "Synthesis of novel indolylpyrimidine antiinflammatory, antioxidant and antibacterial agents," *Indian Journal of Pharmaceutical Sciences*, vol. 70, no. 2, pp. 208–215, 2008.
- [16] D. F. Veber, S. R. Johnson, H. Cheng, B. R. Smith, K. W. Ward, and K. D. Kopple, "Molecular properties that influence the oral bioavailability of drug candidates," *Journal of Medicinal Chemistry*, vol. 45, no. 12, pp. 2615–2623, 2002.
- [17] T. Masuda, T. Jikihara, K. Nakamura, A. Kimura, T. Takagi, and H. Fujiwara, "Introduction of solvent-accessible surface area in

- the calculation of the hydrophobicity parameter log P from an atomistic approach," *Journal of Pharmaceutical Sciences*, vol. 86, no. 1, pp. 57–63, 1997.
- [18] Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. OECD Environment Health and Safety Publications Series on Testing and Assessment No. 69. OECD: Paris, 2007, <http://www.oecd.org/dataoecd/55/35/38130292.pdf>.
- [19] J. G. Topliss and R. P. Edwards, "Chance factors in studies of quantitative structure-activity relationships," *Journal of Medicinal Chemistry*, vol. 22, no. 10, pp. 1238–1244, 1979.
- [20] R. D. Cramer III, D. E. Patterson, and J. D. Bunce, "Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins," *Journal of the American Chemical Society*, vol. 110, no. 18, pp. 5959–5967, 1988.
- [21] G. W. Snedecor and W. G. Cochran, *Statistical Methods*, Oxford and IBH, New Delhi, India, 1967.
- [22] S. Chatterjee, A. S. Hadi, and B. Price, *Regression Analysis by Examples*, Wiley VCH, New York, NY, USA, 2000.
- [23] M. V. Diudea, *QSPR/QSAR Studies for Molecular Descriptors*, Nova Science, Huntington, NY, USA, 2000.
- [24] A. M. Doweyko, "3D-QSAR illusions," *Journal of Computer-Aided Molecular Design*, vol. 18, no. 7–9, pp. 587–596, 2004.
- [25] Y. Marrero Ponce, J. A. Castillo Garit, F. Torrens, V. Romero Zaldivar, and E. A. Castro, "Atom, atom-type, and total linear indices of the "molecular pseudograph's atom adjacency matrix": application to QSPR/QSAR studies of organic compounds," *Molecules*, vol. 9, no. 12, pp. 1100–1123, 2004.
- [26] A. Golbraikh and A. Tropsha, "Beware of q^2 !," *Journal of Molecular Graphics and Modelling*, vol. 20, no. 4, pp. 269–276, 2002.
- [27] N. Strazielle and J. Ghersi-Egea, "Factors affecting delivery of antiviral drugs to the brain," *Reviews in Medical Virology*, vol. 15, no. 2, pp. 105–133, 2005.
- [28] L. Buydens, D. L. Massart, and P. Geerlings, "Prediction of gas chromatographic retention indexes with topological, physicochemical, and quantum chemical parameters," *Analytical Chemistry*, vol. 55, no. 4, pp. 738–744, 1983.



Hindawi
Submit your manuscripts at
<http://www.hindawi.com>

