

QSAR studies on chalcone derivatives as antibacterial agents against *Bacillus pumilis*

Y. Rajendra Prasad,^{a*} P. Ravi Kumar,^a D. Jesse Smiles,^b and P. Ajay Babu^b

^aPharmaceutical Chemistry Division, AU College of Pharmaceutical Sciences, Andhra University, Visakhapatnam - 530 003, Andhra Pradesh, India

^bProGene Biosciences, Institute of Bioinformatics & Research Centre, 103, Bharat Towers, Dwaraka Nagar, Visakhapatnam - 530 016, India

E-mail: dryejella@yahoo.co.in, ajay_pgb@progenebio.in

Abstract

QSAR analysis on a set of synthesized chalcone derivatives tested for growth inhibitory activity against *Bacillus pumilis* was performed by using multiple regression procedure. The activity contributions of these compounds were determined from regression equation and the validation procedures to analyze the predictive ability of QSAR model were described. The generated model from a 25 molecule training set and 7 molecule validation set using 47 independent variables revealed that an increase in ADME Weight, Kappa2 index and a decrease in HOMO energy as favorable descriptors for *Bacillus pumilis* inhibition.

Keywords: QSAR, chalcone, antibacterial activity, MIC, descriptors

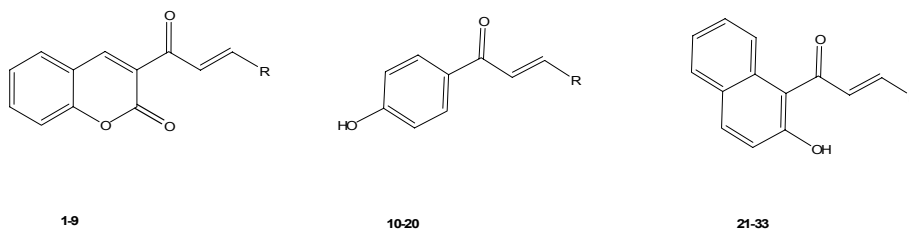
Introduction

Compounds with chalcone as backbone have been reported to exhibit a wide variety of pharmacological effects, including antioncogenic¹, anti-inflammatory², antiulcerative³, analgesic⁴, antiviral⁵, antimalarial⁶ and antibacterial activities⁷. The presence of a reactive α,β -unsaturated keto function in chalcones is found to be responsible for their antimicrobial activity, which may be altered depending on the type and position of the substituent on the aromatic rings. We synthesized a number of chalcones⁸⁻¹⁰ from 3-acetyl coumarin, p-hydroxy acetophenone and 2-hydroxy-1-acetonaphthone with different aromatic as well as heterocyclic aldehydes. The compounds were evaluated for their *in vitro* antibacterial activity against *Bacillus pumilis*.

Structure activity relationship studies delineate the structural requirements for potency of inhibitors. QSAR studies have been investigated on the basis of the fact that the biological activity of the compound is a function of its physicochemical properties. In this study, we report the minimum inhibitory concentration (MIC) and QSAR analysis on a series of chalcone

derivatives tested for their antibacterial activity against *Bacillus pumilis*. The activity data is presented in Table 1.

Table 1. Structures, MIC and logarithmic molar concentration of chalcone derivatives against *Bacillus pumilis*



Compound	R	MIC ^a ($\mu\text{g/ml}$)	C ^b	Observed (log ₁ /C)	Calculated ^c (log ₁ /C)
1	3-pyridyl	900	0.00325	2.49	2.54
2	piperazinyl	900	0.00317	2.50	2.51
3	2-pyrrolyl	900	0.00339	2.47	2.44
4	2,4,5-trimethoxy phenyl	500	0.00136	2.86	2.84
5	3,4-dimethoxy phenyl	600	0.00178	2.75	2.75
6	3-nitro phenyl	600	0.00187	2.73	2.74
7	4-hydroxy-3-methoxy phenyl	700	0.00217	2.66	2.68
8	3-bromo phenyl	500	0.00141	2.85	2.82
9	2-pyridyl	700	0.00252	2.60	2.55
10	anthryl	700	0.00216	2.67	2.65
11	5-bromo furyl	900	0.00307	2.51	2.55
12	2,4,5-trimethoxy phenyl	700	0.00223	2.65	2.65
13	4-hydroxy-3-methoxy phenyl	900	0.00333	2.48	2.49
14	3-bromo-5-chloro-2-hydroxy phenyl	500	0.00141	2.85	2.79
15	2,4-dichloro phenyl	900	0.00307	2.51	2.59
16	3,4,5-trimethoxy phenyl	700	0.00223	2.65	2.69
17	2-pyridyl	900	0.00400	2.4	2.34
18	2-quinolyl	900	0.00327	2.49	2.51
19	4-quinolyl	900	0.00327	2.49	2.52
20	3-pyridyl	900	0.00400	2.4	2.35
21	4-methoxy phenyl	700	0.00230	2.64	2.61
22	3,4,5-trimethoxy phenyl	500	0.00137	2.86	2.86
23	phenyl	900	0.00328	2.48	2.49

24	2-hydroxy phenyl	900	0.00310	2.51	2.55
25	3-bromo phenyl	600	0.00170	2.77	2.77
26	4-dimethyl amino phenyl	700	0.00221	2.66	2.64
27	2,4,5-trimethoxy phenyl	500	0.00137	2.86	2.84
28	4-nitro phenyl	600	0.00188	2.73	2.69
29	4-methyl phenyl	900	0.00312	2.51	2.54
30	9-anthryl	500	0.00134	2.87	2.84
31	2-chloro phenyl	700	0.00227	2.64	2.62
32	2,4-dichloro phenyl	700	0.00204	2.69	2.74
33	3,4-dimethoxy phenyl	700	0.00209	2.68	2.72

^aMinimum Inhibitory Concentration. ^bMIC value expressed in molar concentration. ^cCalculated from Equation 1.

Results and Discussion

Multivariate regression analysis with F stepping (F to enter and F to leave being 4) and cross-validation by leaving-out-one row, to test the predictive power, resulted in ADME Weight, HOMO and Kappa2 index (Table 2) as the most significant descriptors. Equation 1 represents the linear QSAR model from a complete set of 33 chalcone derivatives.

$$\begin{aligned} \log(1/C) = & + 0.003 * \text{ADME Weight} \\ & - 0.061 * \text{HOMO} \\ & + 0.023 * \text{Kappa2 index} \\ & + 1.843 \\ r = 0.969, r^2 = 0.939, q^2 = 0.973, s = 0.037, \\ F = 150.05, n = 33, \text{Predictive Residual Sum of Squares} = 0.04 \end{aligned} \quad (1)$$

The data set was divided into a training set and validation set to cover the whole activity range and the new regression model based on 27 molecule training set was:

$$\begin{aligned} \log(1/C) = & + 0.003 * \text{ADME Weight} \\ & - 0.078 * \text{HOMO} \\ & + 0.028 * \text{Kappa2 index} \\ & + 0.676 \\ r = 0.962, r^2 = 0.926, q^2 = 0.941, s = 0.041, \\ F = 95.91, n = 27, \text{Predictive Residual Sum of Squares} = 0.037 \end{aligned} \quad (2)$$

Inter-correlation between significant descriptors utilized in equation-2 is presented in Table 3, where it is clear that the descriptors are not highly correlated.

Table 2. Calculated values of descriptors used to generate QSAR models

Compound	ADME Weight	HOMO	Kappa2 index
1	277.29	-9.397	7.513
2	284.34	-8.558	7.513
3	265.28	-8.702	6.840
4	366.39	-8.277	10.156
5	336.36	-8.742	9.274
6	321.3	-9.752	8.590
7	322.33	-8.886	8.590
8	355.19	-9.490	7.714
9	277.29	-9.509	7.513
10	324.39	-8.201	8.347
11	293.12	-9.094	6.250
12	314.36	-8.398	9.475
13	270.3	-8.782	7.852
14	353.6	-9.287	7.320
15	293.15	-9.409	7.136
16	314.36	-8.940	9.475
17	225.26	-9.322	6.805
18	275.32	-8.918	7.513
19	275.32	-9.180	7.513
20	225.26	-9.393	6.805
21	304.36	-8.678	8.393
22	364.42	-8.780	10.156
23	274.33	-8.730	7.513
24	290.33	-8.811	7.714
25	353.22	-8.783	7.714
26	317.41	-8.358	8.590
27	364.42	-8.412	10.156
28	319.33	-8.984	8.590
29	288.36	-8.710	7.714
30	374.45	-8.242	9.240
31	308.77	-8.815	7.714
32	343.21	-8.872	7.920
33	334.39	-8.514	9.274

Table 3. Correlation matrix for the three significant descriptors

	ADME Weight	HOMO	Kappa2 index
ADME Weight	1		
HOMO	0.303	1	
Kappa2 index	0.669	0.56188	1

A brief explanation of the descriptors utilized to generate the statistical QSAR model: Molecular weight is an important parameter that signifies the size of the molecule. Molecular Orbital (MO) surfaces represent the various stable electronic distributions of a molecule. According to Frontier Orbital theory, the highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) are crucial in predicting the reactivity of a species. HOMO is the outermost orbital containing the electron and LUMO is the first orbital that does not contain an electron. The Kappa index¹¹ is a molecule shape index based on the assumption that the shape of a molecule is a function of the number of atoms and their bonding relationship.

Equation 2 was used to predict the antibacterial activity for both training and validation sets. Experimental versus predicted values are shown graphically in Figure-1. Equation 2 obtained was of excellent statistical quality that explains 92.6% of variance ($r^2 = 0.926$) of inhibitory activities of chalcone derivatives and has good internal predictivity as shown by q^2 value of 0.941.

The generated QSAR model (Equation 2) indicates that a high value of HOMO energy contributes negatively to the activity. An electron-donating substituent, such as hydroxy, methoxy or methyl group, on the ring increases the energy of the HOMO orbital. A lone pair of electrons on oxygen atom of methoxy group delocalizes into the π space of benzene ring, thereby increasing the energy of HOMO. Similarly, electron-withdrawing substituents, such as halogens, lower the energy of HOMO. An electronegative halogen removes electron density from the σ space of benzene ring, thereby decreasing the energy of HOMO¹². Thus, designing analogs with electron-withdrawing substituents on the ring should improve activity.

On the other hand, a high value of ADME Weight and Kappa 2 index represents a positive contribution to the activity. Molecular weights of chalcone derivatives are in the range of 225.26 to 366.39 daltons respectively. Compounds with 650 daltons molecular weight range tend to be very insoluble and impermeable¹³. Therefore, a marginal increase in molecular weights of the compounds under study should increase antibacterial activity.

Kappa 1 shows the degree of complexity of a bonding pattern. Kappa 2 indicates the degree of linearity of bonding patterns. Kappa 3 indicates the degree of branching at the centre of a molecule, larger for predominantly linear molecules with branching at the ends. Equation 2 suggests that a high value of Kappa 2 index increases the activity. Therefore, designing chalcone derivatives with a high degree of bonding linearity with groups that increase molecular weight

and attachments with electron-withdrawing substituents on the ring increases antibacterial activity against *Bacillus pumilis*. Observed, calculated and predicted activities of chalcone derivatives are presented in Table 4.

Table 4. Observed, calculated and predicted activities of chalcone derivatives using Equation 2

Compound	Observed (log1/C)	Calculated (log1/C) Training set	Predicted (log1/C) Validation set
1	2.49	2.55	
2 ^a	2.5	-	2.50
3	2.47	-	2.43
4 ^a	2.86	-	2.83
5	2.75	2.74	
6 ^a	2.73	2.75	
7	2.66	2.69	
8	2.85	2.82	
9	2.6	2.55	
10 ^a	2.67	-	2.63
11	2.51	2.54	
12	2.65	2.64	
13	2.48	2.48	
14	2.85	2.79	
15	2.51	2.59	
16	2.65	2.69	
17	2.4	2.35	
18	2.49	2.50	
19	2.49	2.52	
20 ^a	2.4	2.35	
21	2.64	2.60	
22	2.86	2.86	
23 ^a	2.48	2.48	
24	2.51	2.55	
25 ^a	2.77	-	2.76
26	2.66	2.63	
27	2.86	2.83	
28	2.73	2.68	
29	2.51	2.53	
30	2.87	-	2.83
31	2.64	2.61	
32	2.69	2.73	
33	2.68	2.72	

^a Validation set.

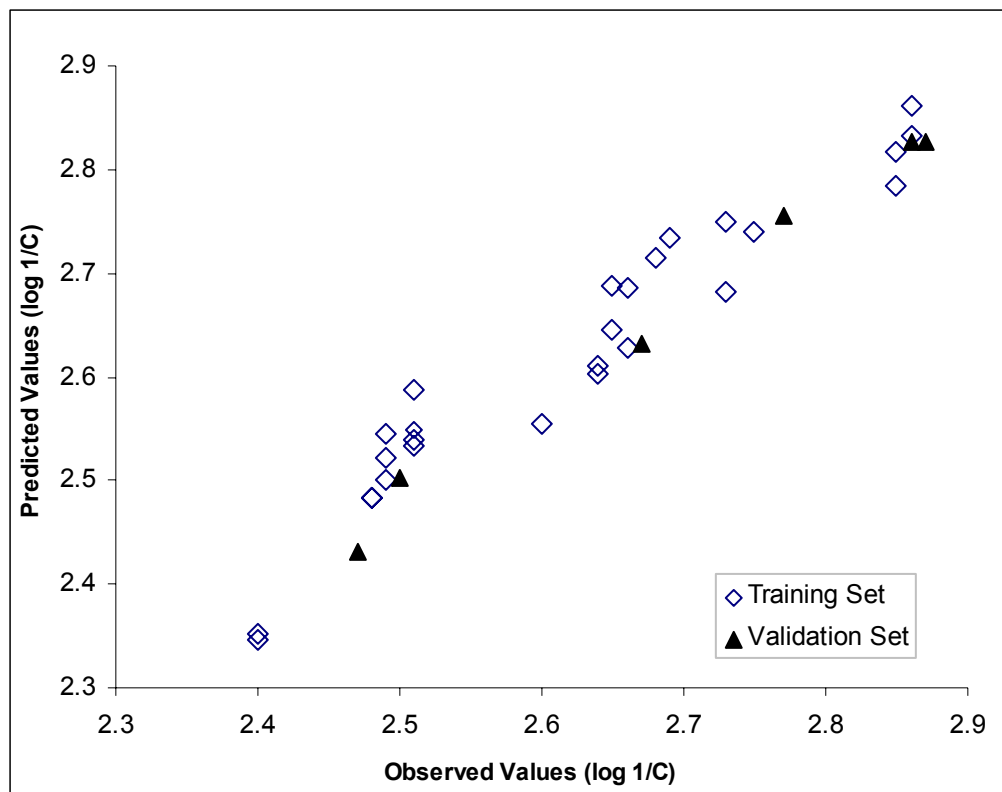


Figure 1. Observed and predicted values ($\log 1/C$) for the training and validation set.

Predictive ability of QSAR model

Equations 3-6 (experimental section) illustrates the predictive ability of QSAR Eq. 2, where, the $R^2_{cv,ext}$ and other parameters are within the limits (Figure 2).

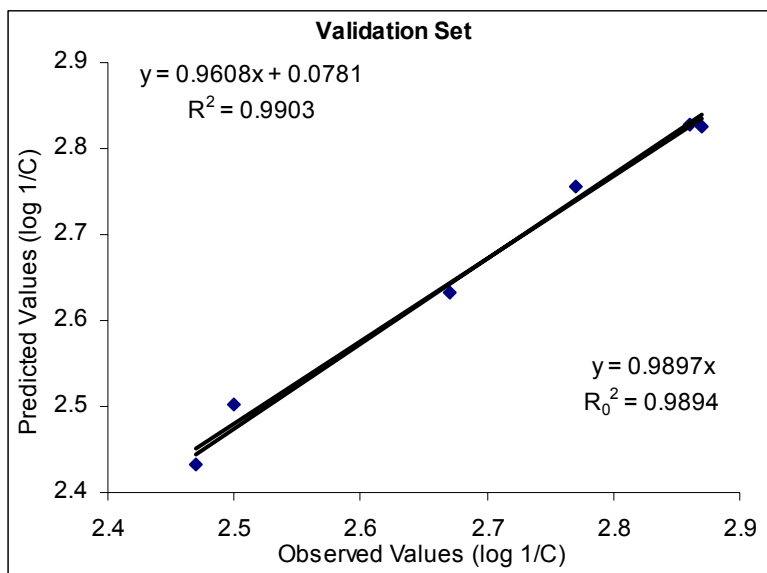
$$R^2_{cv,ext} = 0.841 > 0.5$$

$$R^2 = 0.99 > 0.6$$

$$\frac{(R^2 - R_0^2)}{R^2} = 0.001 < 0.1 \text{ or } \frac{(R^2 - R_0^2)}{R^2} = 0.0004 < 0.1$$

$$k = 1.01 \quad \text{and} \quad k' = 0.989$$

a)



b)

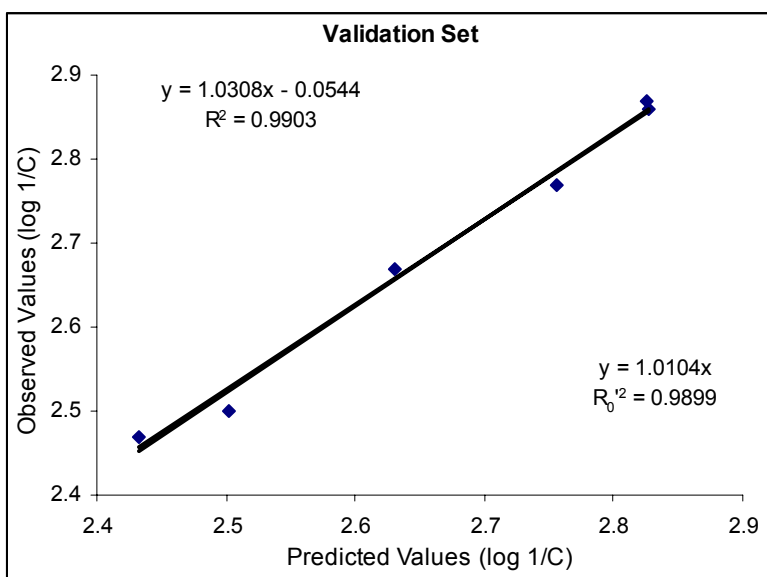


Figure 2. Regression plot between a) predicted vs. observed values (log1/C) and b) observed vs. predicted values (log1/C) for compounds from validation set justifying the predictive ability of QSAR model.

Conclusions

The results of this study suggest that variables like ADME Weight, HOMO and Kappa2 index play an important role in defining inhibitory activity of chalcone derivatives against *Bacillus*

pumilis. The predictive ability (equations 3-6) of QSAR model illustrated the accuracy and robustness of equation 2 on training and validation data sets. The analysis, based on validation procedures, offers a useful technique to synthesize novel analogs for determining antibacterial activity against *Bacillus pumilis*.

Experimental Section

Minimum inhibitory concentration

The synthesized compounds were tested for antibacterial activity against human pathogenic gram positive bacteria, *Bacillus pumilis*. The minimal inhibitory concentrations (MICs) of the compounds assays were carried out as described by Clause¹⁴ with minor modifications. Ampicillin was used as standard antibacterial agent. Solutions of the test compounds and Ampicillin were dissolved in DMSO at concentrations of 5000 µg/ml. The two fold dilution of the compounds and ampicillin were prepared. MIC tests were carried out in Muller Hinton agar medium, pH 7.2, with an inoculum of $(1-2) \times 10^6$ colony-forming units per ml (CFU/ml). The chemical compounds in Muller Hinton Agar serial tube dilutions inoculated with *Bacillus pumilis* were incubated on a rotary shaker at 37°C for 18hr at 150rpm. The minimum inhibitory concentrations of the synthesized compounds were recorded as the lowest concentration of each compound in the tubes with no growth (i.e. no turbidity) of inoculated bacteria. Muller Hinton Agar medium containing DMSO inoculated microorganisms were used as a control. Experiments were repeated at least thrice to find out reproducibility of the assays.

Molecule drawing and software

The inhibitory activity of chalcone derivatives were reported in terms of Minimum Inhibitory Concentration (MIC) against *Bacillus pumilis*. These MIC values were used as log (1/C) (Table 1). The structures of 33 chalcone derivatives were sketched using ISIS Draw 2.3 (www.mdli.com) and the descriptors were calculated using Tsar 3.3 software (www.accelrys.com). Before the calculation of the descriptors, three dimensional structures of all the molecules were generated using Corina 3D, derived charges and the geometries were optimized using the cosmic module of Tsar.

Molecular descriptors

Forty seven molecular descriptors were selected for the study: topological, shape and connectivity indices, 6-membered aromatic rings, 6-membered rings, total lipole and components, ADME molecular weight, h-bond donors, h-bond acceptors, logP and rotatable bond counts. A semi-empirical molecular orbital package Vamp was used to calculate electrostatic properties like heat of formation, HOMO and LUMO.

Multivariate regression analysis

QSAR models were constructed on complete set and training sets, respectively. Validation was done internally using leave-one-out (LOO) technique and externally by predicting the activities of the validation set. The relationship between dependent variable ($\log 1/C$) and independent variables was established by linear multiple regression analysis using Tsar. Significant descriptors were chosen based on the statistical data of analysis. Inter-correlation between these descriptors was checked for independence of the variables. Statistical quality of the generated QSAR equation was judged based on the parameters like correlation coefficient (r), standard error of estimate (s), F-value, cross-validation r^2 (q^2) and predictive residual sum of squares (PRESS).

Predictive ability of QSAR model

Predictive ability of the generated model was estimated externally by predicting the activities of validation set. This criterion may not be sufficient for a QSAR model to be truly predictive¹⁵. An additional condition for high predictive ability of a QSAR model is based on external set cross-validation r^2 , ($R^2_{cv,ext}$) and the regression of observed activities against predicted activities and vice versa for validation set, if the following conditions are satisfied.^{15,16}

$$R^2_{cv,ext} > 0.5 \quad (3)$$

$$R^2 > 0.6 \quad (4)$$

$$(R^2 - R_0^2) / R^2 < 0.1 \text{ or } (R^2 - R_0'^2) / R^2 < 0.1 \quad (5)$$

$$0.85 \leq k \leq 1.15 \text{ or } 0.85 \leq k' \leq 1.15 \quad (6)$$

Calculations relating to $R^2_{cv,ext}$, R_0^2 and the slopes, k and k' are based on regression of observed values against predicted values and vice versa. They were discussed in detail in ref.15.

Acknowledgements

We are thankful to Prof. T. Prabhakar, Department of Pharma Biotechnology, Andhra University, for providing facilities to carry out antibacterial activities.

References

1. Kumar, S. K.; Hager, E.; Pettit, C.; Gurulingappa, H.; Davidson N. E.; Khan, S. R. *J. Med. Chem.* **2003**, *46*, 2813.
2. Hsieh, H. K.; Lee, T. H.; Wang, J. P.; Wang, J. J.; Lin, C. N. *Pharm. Res.* **1998**, *15*, 39.
3. Murakami, S.; Muramatsu, M.; Aihara,; Otomo, S. *Biochem. Pharmacol.* **1991**, *42*, 1447.
4. Viana, G. S.; Bandeira, M. A.; Matos, F. J. *Phytomedicine.* **2003**, *10*, 189.
5. Wu, J. H.; Wang, X. H.; Yi, Y. H.; Lee, K. H. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1813.
6. Liu, M.; Wilairat, P.; Go, M. L. *J. Med. Chem.* **2001**, *44*, 4443.

7. Bekhit, A. A.; Habib, N. S.; El-Din A.; Bekhit, A. *Boll. Chim. Farm.* **2001**, *140*, 297.
8. Rajendra-Prasad, Y.; Ravi-Kumar, P.; Asha-Deepti, Ch.; Venkata-Ramana, M. *Eur. J. Chem.* **2006**, *3*, 236.
9. Rajendra-Prasad, Y.; Ravi-Kumar, P.; Srivani, N.; Srinivasa-Rao, A. *Int. J. Chem. Sci.* **2006**, *4*, 905.
10. Rajendra-Prasad, Y.; Ravi-Kumar, P.; Asha-Deepti, Ch.; Venkata-Ramana, M. *Asian J. Chem.* In Press.
11. Hall, L. H. and Kier, L. B. The molecular connectivity chi indexes and kappa shape indexes in structure-property modeling. *Reviews in Computational Chemistry*: Lipkowitz, K. B.; Boyd, D. B., Eds., 1992; Vol. 9, pp 367.
12. Venkataraman, L.; Park, Y. S.; Whalley, A. C.; Nuckolls, C.; Hybertsen, M. S.; Steigerwald, M. L. Electronics and chemistry: varying single-molecule junction conductance using chemical substituents. *Nano Lett.* **2007**, *7*, 502.
13. Lipinski, C. A. Physicochemical properties and the discovery of orally active drugs: technical and people issues, *Molecular Informatics: Confronting Complexity*, Martin G. Hicks & Carsten Kettner (Eds.), Proceedings of the Beilstein-Institut Workshop, Bozen, Italy, May 13th - 16th, 2002.
14. Clause, G. W. In: *Understanding Microbes: A Laboratory Text Book for Microbiology*, Freeman, New York, USA, 1989; pp 571.
15. Golbraikh, A.; Tropsha, A. *J. Mol. Graph. Mod.* **2002**, *20*, 269.
16. Afantitis, A.; Melagraki, G.; Sarimveis, H.; Koutentis, P.A.; Merkopoulos, J.; Igglessi – Markopoulou, O. *Bioorg. Med. Chem.* **2006**, *14*, 6686.