ASIAN JOURNAL OF CHEMISTRY

3D-QSAR CoMFA/CoMSIA Studies, to Design New VEGFR-2 Inhibitors

S. BATOUCHE and A. GHEID[*]

Centre Universitaire Souk-Ahras, BP. 1553 Route Principale S.A. 41000, Algeria

*Corresponding author: E-mail: hakgheid@yahoo.fr

The vascular endothelial growth factor receptor 2 (VEGFR-2/KDR) is the target of current interest of anticancer drug research. Several classes of ligands have been studied as potential inhibitors against protein tyrosine kinases (PTK). To explore the relationships between the structure of the ligands and their inhibitory activity, 3D-QSAR analysis using CoMFA and CoMSIA techniques were performed on a series of 45 compounds. In this study we used three kinds of tyrosine kinase inhibitors with observable structural diversity in order to construct a universal 3D-QSAR model able to provide more useful information about their inhibitory mechanism. The binding conformation of these compounds was determined by using docking and alignment procedure. Based on their active conformation and their alignments, two highly predictive models were derived, CoMFA model ($q^2 = 0.620$, $r^2 = 0.985$) and CoMSIA model combined hydrophobic, hydrogen-bond donor and hydrogen-bond acceptor ($q^2 = 0.665$, $r^2 = 0.960$), both models were validated by a test set of 8 compounds producing very good predictive r^2 values of 0.840 and 0.810, respectively. The CoMFA and CoMSIA contour maps were used to analyze the contribution of physicochemical properties on the activity of the inhibitors. The best models were used to design a new tyrosine kinase inhibitor with enhanced activity.

Key Words: CoMFA, CoMSIA, Tyrosine kinase, Vascular endothelial growth factor receptor 2 (VEGFR2), Docking, Angiogenesis.

INTRODUCTION

Angiogenesis, the formation of new blood vessel from an existing vascular network[1-3], is an important physiological process that is involved in embryonic development, follicular growth and wound healing[4]. This process is also implicated in the pathogenesis of several diseases including diabetic retinopathy, rheumatoid arthritis, psoriasis, growth and the metastasis of solid tumor[5-8].

Angiogenesis plays a vital role in tumor growth, which cannot grow beyond a certain critical size until they develop new blood vessels to provide oxygen and nutrients.

Tumor angiogenesis is a complex process in which tumor cells in response to hypoxia[9] secrete a number of stimulatory cytokines of which the vascular endothelial growth factor (VEGF) plays a key role[10,11]. The biological effects of VEGF are mediated by two receptors tyrosine kinase, VEGF-R1 (also known as Flt-1) and VEGFR-2 (also known as flk-1/KDR, fetal liver kinase-1/kinase domain-containing receptor)[12,13]. The binding of VEGF to VEGFR-1 and VEGFR-2 receptors at the endothelial cell (EC) surface triggers a number of intracellular signaling events that lead to increase proliferation, migration and survival of the EC[14-17]. Recently, it has been demonstrated that VEGFR-2 is the major mediator of vascular endothelial cell mitogenesis, angiogenesis and microvascular perme-

ability[18-22]. Then the direct inhibition of the VEGFR-2 kinase activity will result in the reduction of tumor angiogenesis and the suppression of tumor growth[23-27]. One such approach for obtaining VEGFR-2 inhibition involves using small molecules to block adenosine triphosphate (ATP) binding to the intracellular kinase domain of the receptor, resulting in diminished VEGF signal transduction. Inhibitors of VEGFR-2 have therefore become a major focus of much research organization. Last decade, several classes of small molecule-based VEGFR-2 kinase inhibitors merged as promising anti-angiogenic agents for possible treatment against a wide variety of cancers[28-34]. Recently, two small molecule drugs that inhibit VEGFR-2 kinase, Sorafenib (BAY-43-9006) and Sunitinib (SU-11248)[35-37] were approved by the FDA for renal and gastrointestinal cancer.

Quantitative structure activity relationship at three dimensional levels (3D-QSAR) is a useful tool for developing new drug with higher activity. In addition 3D-QSAR techniques, such as comparative field analysis (CoMFA)[38] and comparative similarity analysis (CoMSIA)[39] are used in drug design to help understanding about the drug-receptor interaction.

Several applications of 3D-QSAR for the drug design have been published[40-45]. Studies on the VEGFR-2 inhibitors are limited to one kind of inhibitors, one reason is the difficulty

of aligning compounds which have differences in structure. 3D-QSAR studies that covered more different kinds of ligands lead to more useful information about the influence of different physico-chemical and structural parameters on the inhibitory activity.

The aim of this work is to develop robust models using three different kinds of ligands. To address such powerful models, we carried out CoMFA and CoMSIA studies on 45 KDR inhibitors belonging to three types of reported VEGFR-2 inhibitors. These models would offer utility to design a new VEGFR-2 inhibitor with improved efficiency.
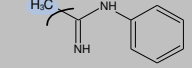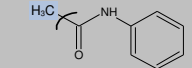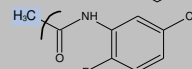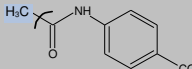
EXPERIMENTAL
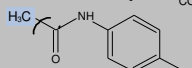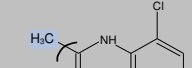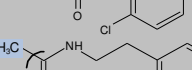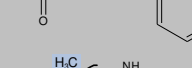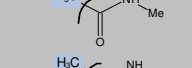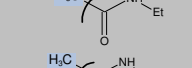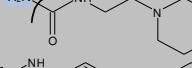
A data set composed of 45 KDR inhibitors and their inhibitory activity which covered nearly four logs units (pIC$_{50}$ = 5.33-9.00), was taken from literature[47-49]. Forty five reversibly ATP-competitive inhibitors of KDR PTK, belonging to three types of KDR inhibitors. They are phenylaminoquinazolines, anilinophthalazine and dianilinopyrimidine. The total set was randomly divided into two subsets: a training set of 37 compounds for generating the 3D-QSAR models and test set of 8 compounds for validating the quality of models. The structure of the KDR inhibitors and their activities are shown in Table-1.

Phenylaminoquinazolines (1-T2)

Anilinophthalazines (9-15)

Dianilinopyrimidine (16-37)
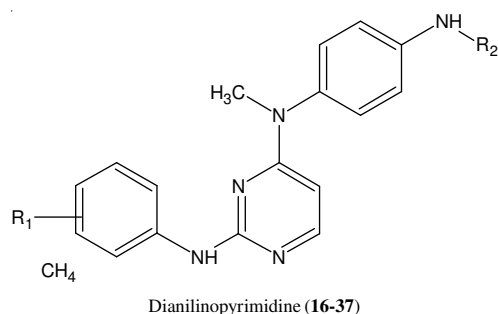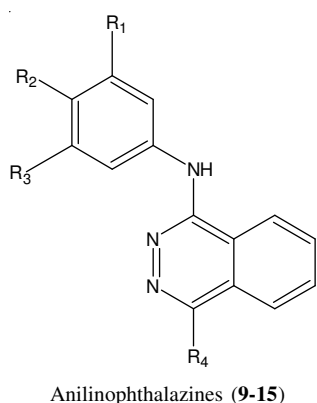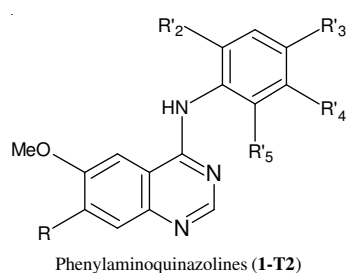
Structure of the three classes of molecules used in the study

TABLE-1
STRUCTURE AND ACTIVITY OF THE 45 INHIBITORS,
T1, T2, T3 … T8 ARE THE TEST SET

Comp.	R'				R	pIC$_{50}$
	1	2	3	4		
1	F	Cl	H	H	4-Pyridyl-N(Me)-(CH$_2$)$_2$O	8.22
2	F	Cl	H	H	MeN(CH$_2$CH$_2$)$_2$CH-O	5.82
3	F	Cl	OH	H	MeN(CH$_2$CH$_2$)$_2$CH-CH$_2$O	8.10
4	F	Me	OH	H	MeN(CH$_2$CH$_2$)$_2$CH-CH$_2$O	8.15
5	F	Br	H	F	MeN(CH$_2$CH$_2$)$_2$CH-CH$_2$O	8.00
T1	F	Br	H	F	HN(CH$_2$CH$_2$)$_2$CH-CH$_2$O	7.82
6	F	Br	H	H	MeN(CH$_2$CH$_2$)$_2$CH-CH$_2$O	7.40
7	F	Cl	H	H	MeO(CH$_2$)$_2$O	8.15
8	F	Br	H	H	4-Me-piperazinyl-(CH$_2$)$_3$O	8.05
T2	F	Cl	H	H	4-Morpholinyl-(CH$_2$)$_3$O	8.05

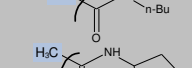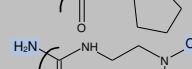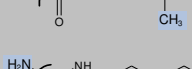|	R$_1$	R$_2$	R$_3$	R$_4$	R$_5$	pIC$_{50}$
9	H	Cl	H			5.77
10	H	Cl	H			5.85
11	H	Cl	H			5.33
T3	H	Cl	H			6.06
12	OMe	H	H			6.62
T4	Me	H	Me			7.40
13	CF$_3$	H	Br			6.42
14	CF$_3$	H	F			6.70
15	H	Cl	H			7.43

	R$_1$	R$_2$	pIC$_{50}$
16	4-MeSO$_2$CH$_2$-		8.40
17	3-MeSO$_2$CH$_2$-		7.74
18	3-MeSO$_2$CH$_2$-		6.75
19	3-MeSO$_2$CH$_2$-		7.02
T5	3-MeSO$_2$CH$_2$-		6.93
20	3-MeSO$_2$CH$_2$-		6.57

No.	R	Structure	Value
21	3-MeSO$_2$CH$_2$-		6.49
T6	3-H$_2$NSO$_2$-		5.30
22	3-MeSO$_2$CH$_2$-		7.46
23	4-MeSO$_2$CH$_2$-		8.70
24	3-H$_2$NSO$_2$-		9.00
25	4-H$_2$NSO$_2$-		9.00
26	3-MeSO$_2$CH$_2$-		8.52
T7	3-MeSO$_2$CH$_2$-		8.22
27	3-H$_2$NSO$_2$-		8.70
28	4-H$_2$NSO$_2$-		8.70
29	3-MeSO$_2$CH$_2$-		7.19
30	3-MeSO$_2$CH$_2$-		7.96
31	3-MeSO$_2$CH$_2$-		7.77
32	4-MeSO$_2$CH$_2$-		7.59
33	3-MeSO$_2$CH$_2$-		7.54
34	3-MeSO$_2$CH$_2$-		7.51
35	3-MeSO$_2$CH$_2$-		6.17
36	4-MeSO$_2$CH$_2$-		5.51
T8	4-MeSO$_2$CH$_2$-		6.00
37	4-MeSO$_2$CH$_2$-		6.82

Fig. 1. Binding mode of inhibitor 38

this is why we chose the active conformation of one molecule of our data base which was established and described by relied on the model of complex KDR/ molecule 16 proposed by Sammond *et al.*[47], Fig. 2.

Fig. 2. VEGFR2 kinase inhibitor 16 bound in the enzyme active site as proposed based on homology modeling

Molecular docking: In order to reproduce this model and to get the active structure of molecule **16**, the 3D structure of the inhibitor **38** was extracted from the binding site of KDR and aligned with the molecule **16**; taking as a point of overlapping, oxygen and nitrogen urea. The result of the superimposition is showed in Fig. 3. The molecule **16** has been docked into the binding site of KDR. After adding all the necessary hydrogen this complex has been minimized using MMFF94 force field with MMFF94 charges and convergence criterion of 0.05 kcal/mol. Indeed, in this complex, molecule **16** adopts

The 3D structure of these compounds was constructed using the molecular modeling package Sybil 7.2, Tripos, Inc., St. Louis. Mo). The geometries of the inhibitors were optimized using MMFF94 force field, using an 8 Å non bond cutoff and an energy convergence gradient value of 0.005 kcal/ mol Å.

VEGFR2 structure: In 3D-QSAR studies, the key step is the alignment of molecules. This alignment can be on a rigid active molecule or more generally on a ligand with known active conformation. This is the case when we have the structure of the receptor in complex with the ligand determined by X-Rays diffraction. In this case the conformation of the ligand is used for the alignment of the other molecules of the study. In present work, we used the X-ray crystal structure of KDR kinase in complex with 4-amino-furo-[2,3-d]-pyrimidine (molecule 38) which was recovered from Protein Data Bank, entry code 1YWN. This three dimensional structure of the ligand in the binding pocket can determine the main protein-ligand interactions, Fig. 1 represents these interactions. But the ligand is structurally very distant from the molecules of our study,
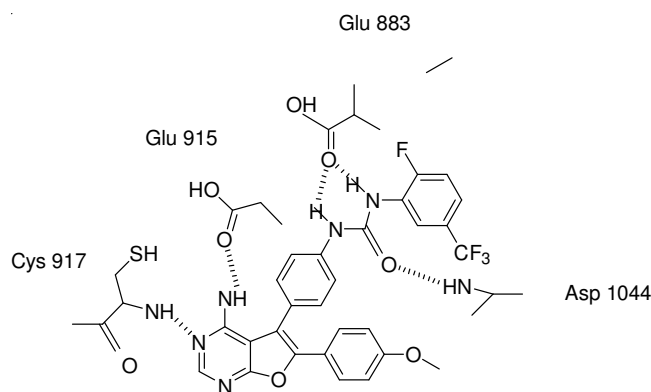

Fig. 3. Fit of molecule **16** and the inhibitor **38**

a conformation similar to the model proposed by Sammond *et al.*[47], the atoms of the molecule **16** interacts with the same residues and the distance between the atoms involved in hydrogen bonds is around 1.7 Å (Fig. 4). Moreover the overlay reflects the existence of common interactions with molecule **16** and ligand **38**, namely those concerning residues Glu 883, Cys 917 this is because the two molecules are positioned in the same region of the pocket.

Fig. 4. Binding mode of molecule **16**

Because of the good superimposition of the molecule **16** with the ligand and its good activity (pIC$_{50}$ = 8.4) we chose molecule **16** as the basis and the template for all lines to present study.
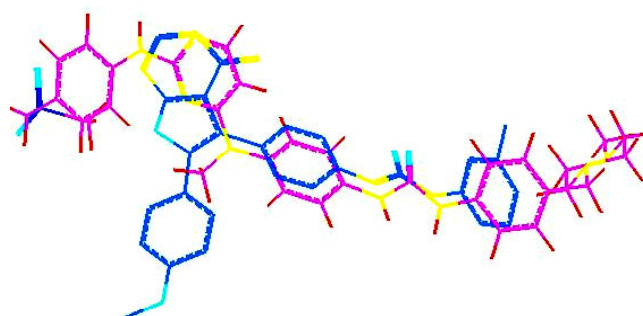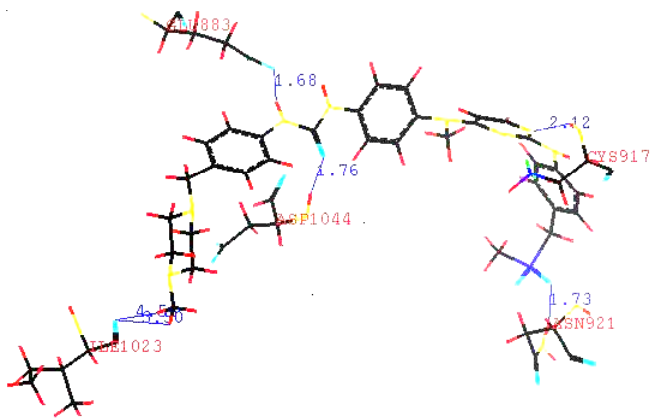
Alignment procedure: A proper alignment of the structure is critical to obtain valid 3D-QSAR models, it's very important that all compounds are aligned in pharmacological active orientation. To obtain a consistent alignment, molecule **16** was selected as the template for the molecular alignment since it exhibits high affinity with KDR (pIC$_{50}$ = 8.4). To work with conformations close to the active conformations, the molecules of the family of phenylamino-quinazoline which have side chains with unknown orientation were first docked into the binding site to avoid steric interactions with residues of the active site. After the minimization of the complex obtained, the molecules of this family were extracted, minimized and aligned on the molecule **16** taking the five points corresponding to atoms and aromatics centers involved in KDR-ligand interactions.

The molecules of the second and third family (anilino-phthalazines and dianilinopyrimidine urea) were aligned on the molecule **16** taking the same points than previous ones (Fig. 5).
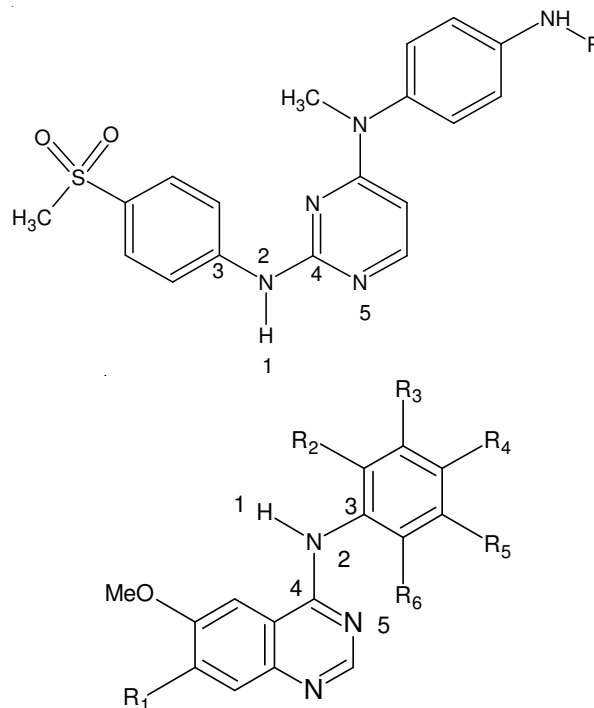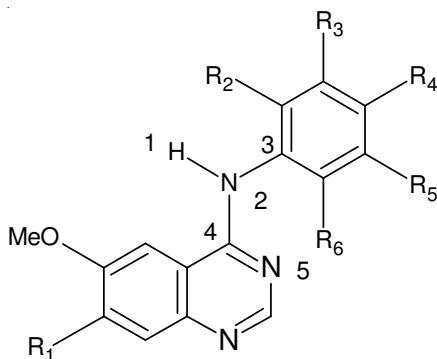
Fig. 5. Common atoms used for the alignment

The result of the alignment of all molecules is shown in Fig. 6.
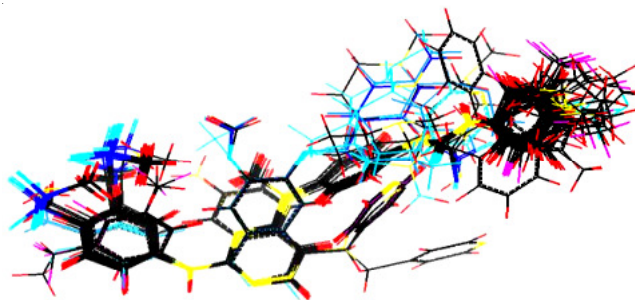
Fig. 6. Alignment of all the compounds

We have also verified that the conformers obtained after the fit are stable by comparing their energy to the energy of the conformers which have the lowest energy obtained after the molecular optimization. For all molecules, the conformations are located in the area of the low energy.

QSAR analysis: To construct good predictive 3D-QSAR models and to evaluate the contributions of steric, electrostatic and hydrophobic effects on the activities of VEGFR2 inhibitors, CoMFA[38] and CoMSIA[39] analyses were performed on the molecules of the training set.

CoMFA: CoMFA (Comparative molecular field analysis) was performed to evaluate the effects of steric and electrostatic fields on the biological activity of the compounds.

After superposition of all the molecules, a three-dimensional cubic lattice with 2 Å grid spacing was generated automatically around the molecules to ensure that the grid extended the molecular dimensions by 4 Å in all directions. Next, the steric field energies were calculated using an sp^3 carbon probe with van der Waals radius of 1.52 Å and a + 1 charge as the electrostatic probe. Energies were truncated to 30 kcal mol^{-1}.

CoMSIA: Comparative molecular similarity index analysis (CoMSIA) introduced by Klebe in 1994 was employed to gain insight into how steric, electrostatic, hydrophobic and hydrogen-bond interactions influence the activity of the inhibitors. Comparative molecular similarity index analysis descriptors were derived by the same lattice box used for the CoMFA calculations. All five CoMSIA similarity index fields (steric, electrostatic, hydrophobic, hydrogen-bond donor and hydrogen-bond acceptor) were evaluated using the sp^3 carbon probe. The CoMSIA similarity indices were calculated between the grid point and each atom of the molecule using a Gaussian function.

Partial least square analysis: Because of the large number of variables generated by the calculation of the fields, it is not possible to correlate these variables with the biological activity according to the method of multi linear regression (MLR). That is the reason that PLS method was performed to correlate the 3D-QSAR fields with the inhibitory activity values. The fields were used as the independent variables and the pIC_{50} values as the dependant variables.

At the beginning of the 3D-QSAR study and in order to reduce the number of variables, a principal component analysis is applied to all the data. The principal components are linear combinations of explanatory variables, the number of principal components was determined by leave-one-out cross-validation (LOO) procedure[50] in which each of the molecules is excluded from the data set and its activity is predicted by the model derived from the remaining compounds in the set. This process is repeated until all molecules are used. The optimum number of components (CN) corresponds to the highest cross validated q^2 and the lowest standard error of prediction (SEP), q^2 value should be greater than 0.5 to indicate significant predictive power.

The model chosen based on the above criteria will be subsequently submitted to a non cross validated analysis in order to define the final form of the model that can be used for predicting the activity.

In this analysis other criteria are determined to decide the quality of the fit; the coefficient of determination r^2, the residual standard deviation s, the F test of Fisher and the probability P. The squared correlation coefficient r^2, which is also used in (multiple) linear regression, is between zero and one and expresses the quality of the PLS analysis. It indicates the proportion of the variation in the dependent variable (here the activity) that is explained by the regression equation and its value should be as close to one as possible. However, r^2 expresses the quality of the data fit rather than the quality of prediction. At the end of the calculations, the final model emerges. It's a graphical representation that shows the regions where the spatial variations of steric and electrostatic properties influence the affinity of ligand for the receptor.

Validation: The series of the molecules of the test set which were not used for the establishment of the model, (training set) were used for the validation of the model. The parameter used to evaluate the predictive capacity of the models CoMFA and CoMSIA is the r^2 predictive[51], this one is calculated on the whole of the molecules of the test set for each model by using the formula.

$$r^2 \text{ predictive} = 1 - \frac{\text{Press}}{\text{SD}}$$

where, SD = sum of the squared deviations between the biological activities of the test set and mean activities of the training molecules and PRESS = sum of squared deviation between predicted and actual activity values for every molecule of the test set.

RESULTS AND DISCUSSION

CoMFA and CoMSIA analyses CoMFA analysis: In this study, we established 2 models. CoMFA I and CoMFA II, The statistical parameters for each model are summarized in Table-2.

TABLE-2 STATISTICAL PARAMETERS OF THE CoMFA MODEL		
PLS analysis	CoMFA I	CoMFA II
Number of compound	37	33[a]
q^{2a}	0.246	0.620
SEP[b]	0.972	0.712
CN[c]	6	5
r^{2d}	0.948	0.985
Sr[e]	0.256	0.142
F[f]	90.355	285.588
Contributions (steric/electrostatic)	0.415/0.585	0.460/0.540

a: Cross-validation correlation coefficient. b: The optimum number of components. c: Non-cross-validation correlation coefficient. d: Standard error of estimate. e: F-ratio.

The Model CoMFA I was constructed with the whole of the molecules of training set produced q^2 value of 0.244 lower than 0.5 indicating that this model is not predictive. After removing molecules 2, 15, 22 and 37 from the training set, we obtained a predictive model, CoMFA II, with higher cross-validated q^2 value of 0.62. For model CoMFA II, the squared correlation coefficient r^2 which indicates the proportion of the biological activity which is explained by the model is 0.985, implies that the model CoMFA II explains 98.5 % of the variation of the biological activity. In this model, the steric descriptors explain approximately 46 % of the variance, while the electrostatic descriptors explain 54 % of the variance. These values suggest a good statistical correlation and satisfactory predictive ability of the CoMFA II model Table-2: statistical parameters of the CoMFA model.

In order to assess the predictive power of the model CoMFA II, the deviation between the experimental and the predicted activities has been calculated for training set as well as test set, the values are presented in Table-3.

CoMSIA analysis: For the CoMSIA analysis, 5 physico-chemical properties were used including, the steric fields (S), electrostatic (E), hydrophobic fields (H), donors (D) and acceptors (A) of hydrogen bonds. The values of these fields were calculated using a probe charged +1. Several models combining the various variables fields were established.

As for the CoMFA analysis, the first CoMSIA analysis carried out on the whole of training set (37 molecules) was not satisfactory. The molecules 2, 15, 22 and 37 which deviated from the model were removed from the data base.

It should be noted that these molecules are identical to the outliers of the CoMFA analysis. Among the CoMSIA models obtained we retained the models corresponding to q^2 > 0.6. The statistical analysis carried out on the various CoMSIA models is summarized in Table-4.

TABLE-3
EXPERIMENTAL, PREDICTED ACTIVITIES AND RESIDUALS
OF THE COMPOUNDS USING CoMFA II AND THE BEST
MODEL FROM CoMSIA

Comp. No.	Experimental pIC$_{50}$	Predicted pIC50		Residuals	
		CoMFA II	CoMSIA (HAD)	CoMFA II	CoMSIA (HAD)
1	8.22	8.13	8.10	0.09	0.12
2	5.82	Outlier	outlier	Outlier	Outlier
3	8.10	8.15	8.04	-0.05	0.06
4	8.15	7.91	8.00	0.24	0.15
5	8.00	8.02	7.93	-0.02	0.07
6	7.40	7.69	7.72	-0.29	-0.32
7	8.15	8.08	8.15	0.06	0.00
8	8.05	8.07	8.02	-0.02	0.03
9	5.77	5.76	5.78	0.01	-0.01
10	5.85	5.89	5.84	-0.037	0.01
11	5.33	5.39	5.19	-0.07	0.14
12	6.62	6.49	6.61	0.12	0.01
13	6.42	6.57	6.69	-0.15	-0.27
14	6.70	6.59	6.70	0.10	0.00
15	7.43	Outlier	outlier	Outlier	Outlier
16	8.40	8.48	8.52	-0.09	-0.12
17	7.74	7.54	7.96	0.21	-0.22
18	6.75	6.63	6.63	0.12	0.12
19	7.02	7.14	7.21	-0.12	-0.19
20	6.57	6.59	6.61	-0.03	-0.04
21	6.49	6.43	6.45	0.06	0.04
22	7.46	Outlier	outlier	Outlier	Outlier
23	8.70	8.51	8.26	0.18	0.44
24	9.00	9.21	8.99	-0.21	0.01
25	9.00	9.07	9.27	-0.07	-0.27
26	8.52	8.46	8.94	0.05	-0.42
27	8.70	8.64	8.61	0.06	0.09
28	8.70	8.52	8.78	0.17	-0.08
29	7.19	7.30	7.10	-0.12	0.09
30	7.96	7.98	7.83	-0.03	0.13
31	7.77	7.80	7.53	-0.04	0.24
32	7.59	7.54	7.27	0.04	0.07
33	7.54	7.71	7.52	-0.17	0.02
34	7.51	7.54	7.22	-0.04	0.29
35	6.17	6.02	6.24	0.155	-0.07
36	5.51	5.64	6.08	-0.14	-0.57
37	6.82	Outlier	outlier	Outlier	Outlier
T1	7.82	7.89	7.77	-0.07	0.05
T2	8.05	7.84	7.88	0.21	0.17
T3	6.06	6.07	6.33	-0.01	-0.27
T4	7.40	6.69	6.67	0.71	0.73
T5	6.93	7.21	6.53	-0.28	0.4
T6	8.30	8.08	8.86	0.22	-0.56
T7	8.22	8.11	8.39	0.11	-0.17
T8	6.00	6.49	5.93	-0.48	0.07

TABLE-5
VALUES OF r^2 PREDICTIVE FOR THE BEST
MODELS CoMFA AND CoMSIA

Model	r^2 Predictive
CoMFA II[*]	0.84
CoMSIA	
H	0.54
HS	0.62
EH	0.67
DH	0.71
AH	0.79
SHA	0.72
EHA	0.78
HSE	0.67
HDA[*]	0.81
EHD	0.73
SHD	0.71
EHDA	0.78
SEHA	0.74
SEHD	0.72
SHDA	0.79
SEHDA	0.77

H: Hydrophobic field. D: Hydrogen bond donor field. A: Hydrogen bond acceptor field. S: Steric field. E: Electrostatic field. *: Best models

the activity. The most predictive CoMSIA model is the one which include the hydrophobic field, the H-bond donor and acceptor (r^2 predictive) yielded of 0.81.

We also carried out a linear regression between the experimental activity and that predicted by the models CoMFA II and CoMSIA (Fig. 7).
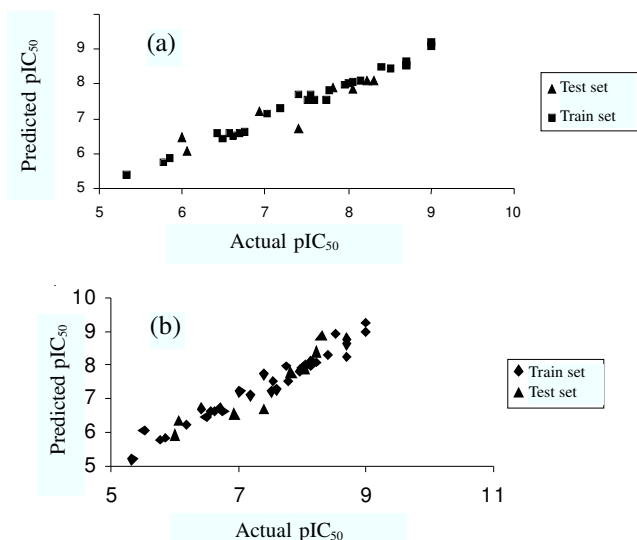

Fig. 7. Correlation between the experimental and predicted activities of (A) CoMFA II and (b) CoMSIA (HAD) model for the training and test set

A good correlation between the predicted activities and the experimental ones was observed.

Graphical interpretation of the results: The CoMFA and CoMSIA steric, electrostatic and H-bonding fields from the final non-cross-validated analysis were plotted as three-dimension coloured contour maps in Figs. 8 and 9. These maps show regions where differences in molecular fields are associated with differences in biological activity.

Validation of the 3D-QSAR models: In order to select the most predictive model and to evaluate the predictive capacity of the models CoMFA and CoMSIA on the whole of compounds of the test set, we calculated r^2 predictive for each model by using the formula (1) described previously

$$r^2 \text{ predictive} = 1 - \text{Press/SD} \qquad (1)$$

By examining Table-5, it is noticed that the model CoMFA II and models which include the hydrophobic field are the most predictive models. Moreover with only the hydrophobic field the model is able to predict in a correct way the values of

TABLE-4
STATISTICAL PARAMETERS OF CoMSIA MODELS

| Fields	Cross validated analysis			Non crossvalidated analysis			Contributions				
	NC	q^2	SEP	r^2	S	F	Steric	Electrostatic	Hydrophobic	H-Bond donor	H-Bond acceptor
Hydrophobic	5	0.648	0.673	0.935	0.289	77.800			1		
Steric-hydrophobic	5	0.657	0.664	0.946	0.264	94.519	0.321		0.679		
Electrostatic-hydrophobic	6	0.616	0.716	0.979	0.167	203.385		0.512	0.488		
Hydrophobic-H-bond donor	5	0.663	0.650	0.952	0.249	106.865			0.670	0.330	
Hydrophobic-H-bond acceptor	6	0.664	0.670	0.961	0.229	101.091			0.521		0.479
Hydrophobic-steric-electrostatic	6	0.665	0.669	0.983	0.153	243.422	0.233	0.336	0.431		
Hydrophobic-H-bond donor-acceptor	6	0.665	0.679	0.960	0.232	103.377			0.416	0.194	0.390
Stéric-hydrophobic-H-bond donor	5	0.660	0.661	0.965	0.212	148.803	0.245		0.483	0.272	
Stéric-hydrophobic-H-bond donor-acceptor	6	0.633	0.700	0.972	0.192	151.967	0.156		0.346	0.179	0.319
Stéric-hydrophobic -H-bond-acceptor	6	0.640	0.693	0.973	0.191	154.585	0.164		0.426		0.390
Electrostatic-hydrophobic-H- bond donor	5	0.662	0.659	0.974	0.148	198.722		0.411	0.375	0.214	
Electrostatic-hydrophobic-H-bond acceptor	6	0.605	0.726	0.979	0.168	201.634		0.348	0.369		0.283
Electrostatic-hydrophobic-H-bond donnor-acceptor	6	0.633	0.700	0.980	0.162	215.437		0.307	0.296	0.167	0.230
Stéric-électrostatic-hydrophobic-H-bond donor	6	0.674	0.660	0.984	0.147	265.073	0.156	0.349	0.303	0.193	
Stéric-électrostatic-hydrophobic-H-bond acceptor	6	0.628	0.705	0.983	0.151	249.547	0.144	0.303	0.311		0.242
Stéric-électrostatic-hydrophobic-H-bond donnor-acceptor	6	0.641	0.693	0.984	0.148	261.519	0.119	0.273	0.254	0.153	0.201
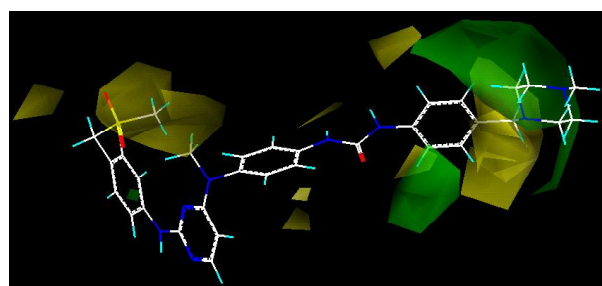
(a)
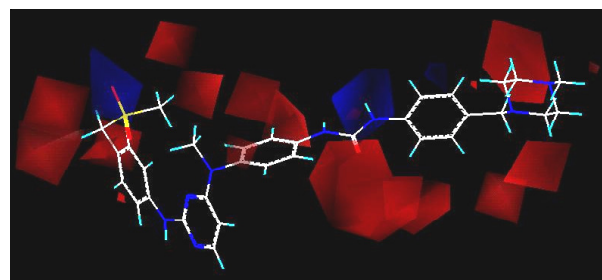
(b)

Fig. 8. CoMFA contour maps in combination with inhibitor 16 (a) the steric fields distribution (b) the electrostatic fields distribution
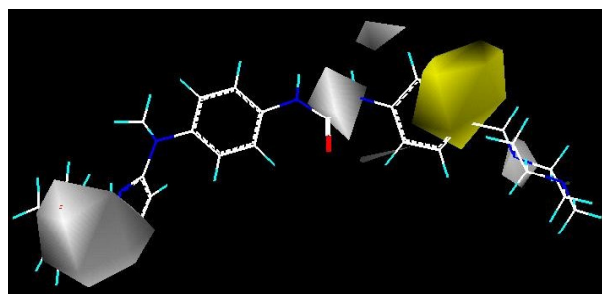
(a)
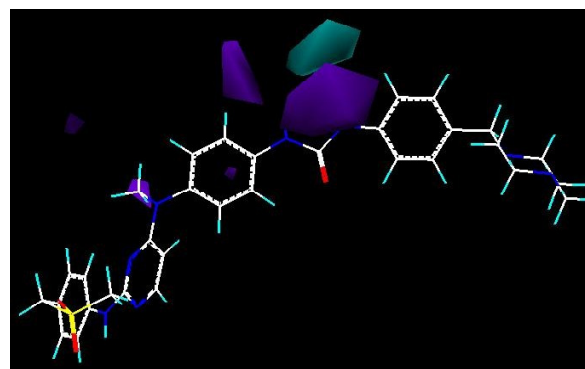
(b)
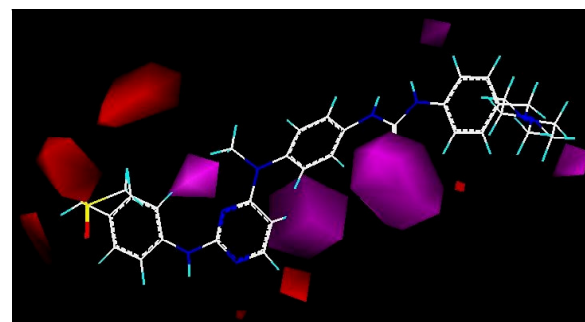
(c)

Fig. 9. CoMSIA (HAD) contour maps in combination with inhibitor 16 (a) the hydrophobic fields distribution (b) the H-bond donor distribution (c) H-bond acceptor fields distribution

CoMFA Contour maps: The CoMFA model is a graphic model, in which one can visualize the space areas where variations of the structural properties *i.e.*, properties steric and electrostatic influence the affinity of the ligands for the receptor, Fig. 8. Green contours represent regions of high steric tolerance

(80 % contribution), while the yellow contours represent regions of low steric bulk tolerance (20 % contribution). The increase in positive charge is favourable in blue regions while increase in negative charge is favoured in red region.

Electrostatic contours of CoMFA (Fig. 8b) show red contours enclosing the urea region indicate that high electron density is expected to increase the activity. Whereas, the blue contours surrounding the sulphuric group suggest that in this region a group with low electron density is expected to increase the activity. The contributions from the steric and electrostatic fields for the present model are 0.46/0.54, indicating that the variations in binding affinity are dominated by electrostatic interactions.

CoMSIA contour maps: As for CoMFA model, CoMSIA model selected, HAD, shown in Fig. 9, is a graphical model which indicates regions in 3D space around the molecules where changes in physicochemical properties are predicted to increase or decrease the biological activity.

The hydrophobic contour map of CoMSIA model, is shown in Fig. 9a. The yellow (contribution level of 80 %) and gray (contribution 20 %), indicate the region where hydrophobic and hydrophilic groups were preferred, respectively. The yellow contour map enclosing the aromatic attached to the urea indicates that a hydrophobic groups would be favourable for binding activity. This region which forms the bottom of the pocket is surrounded and formed by hydrophobic residues, Leu 887, Ile 890, Val 897 and Val 896 which form a hydrophobic binding pocket.

The large gray contour map surrounding the sulfone indicates that the hydrophobic group at this position would decrease the binding affinity. This area is the entrance to the pocket that is generally exposed to the solvent and therefore highly hydrophilic.

The CoMSIA contour maps of the hydrogen-bond donor and hydrogen-bond acceptor fields are shown in Fig. 9(b-c), respectively. Cyan contour (contribution level of 80 %) indicates regions where H-bond donor group increase activity, purple contours (contribution level of 20 %) suggest regions where H-bond donor decreases the activity.

The cyan contour in front of the nitrogen of urea suggests that H-bond donor groups are favoured there. In this region the NH group of the template molecule forms a hydrogen bond with Glu 883 as hydrogen bond donor. The purple contours surrounding the urea moiety indicates that this area is not favoured to H-bond donor group.

Magenta contours (contribution level of 80 %) indicate regions where H-bond acceptor group increases activity, red contours (contribution level of 20 %) represent regions where H-bond acceptor group decreases the activity. A large magenta contour surrounded the oxygen of urea moiety indicates that H-bond acceptor groups are favoured there. This region is formed by residues Asp 1044 and Glu 915. Moreover, the red contours close to the sulfone group indicates that this area is favourable to a substituent of hydrogen-bond donor, this part of the site is lined by residues Arg 1030, Asn 921 and Thr 924 as hydrogen acceptor. The contribution from the hydrophobic, H-bond donor and H-bond acceptor fields for the HAD model are, respectively: 0.416, 0.194 and 0.390 indicating that the variations in the binding affinity are dominated by the hydrophobic interactions.

To verify if CoMFA and CoMSIA models developed are able to predict the inhibitory activity of molecules structurally different from the molecules of our data base, we used a molecule taken from the literature, the Sunitinib known as Sutent[37]. This molecule used in the treatment of gastrointestinal tumors is different from the molecules of present study. The Sunitinib is an inhibitor of tyrosine kinase, it inhibits eight receptor protein-tyrosine kinase; the VEGFR1, VEGFR2, VEGFR3, PDGFRα, PDGFRβ, cKit, Flt3, CSF1-R. The inhibitory activity of Sunitinib for the VEGFR-2 is IC_{50} = 10 nM.

To integrate the Sunitinib in the model and to predict its biological activity, the Sunitinib was first minimized and aligned on the molecule **16** using the nitrogen and oxygen of urea as points of overlapping. Then it was docked in the binding site of KDR. After the complex obtained was minimized, the latter showed that the Sunitinib interacts with the same residues as molecule 16 namely Glu 883 and Asp1044. After extracting the Sunitinib from the hydrophobic pocket, it was minimized and its activity was predicted by each models. Prediction by models CoMFA II and CoMSIA (HAD) are 7.96 and 8.04, respectively. These values are very close to the experimental activity which is 8.0. Both models are good predictive models and could be used to guide the design of new active molecules.

Conclusion

In this study, we have derived 3D-QSAR models using the CoMFA and CoMSIA methods in order to correlate the inhibitory activity of 45 VEGFR-2 PTK inhibitors to their structure. The 45 best binding conformations of these inhibitors, which belong to three different classes of molecules, were determined using docking and alignment. Based on the statistical analyses and the validation on the compounds of the test set, each model shows a good predictive ability. The results demonstrate that the hydrophobic field is the most important parameter in the inhibitory mechanism. The two models chosen according to the r^2 predictive can help to better understanding of inhibitory mechanism. They will also provide a valuable tool for guiding to design new inhibitors with observable structural diversity and enhanced activity.

REFERENCES

1. P.W. Manley, G. Bold, J. Bruggen, G. Fendrich, P. Furet, J. Mestan, C. Schnell, B. Stolz, T. Meyer, B. Meyhack, W. Stark, A. Strauss and J. Wood, *Biochimica et Biophysica Acta*, **1697**, 17 (2004).
2. L. Lu, F. Payvandi, L. Wu, L-H. Zhang, R. J. Hariri, H-W. Man, R.S. Chen, G.W. Muller, C.W. Hughes, D.I. Stirling, P.H. Schafer and J.B. Bartlett, *Microvascular Res.*, **77**, 78 (2009).
3. A.M. Goodwin, *Microvascular Res.*, **74**, 172 (2007).
4. H. Nakamura, Y. Sasaki, M. Uno, T. Yoshikawa, T. Asano, H. Seung Ban, H. Fukazawa, M. Shibuya and Y. Uehara, *Bioorg. Med. Chem. Lett.*, **16**, 5127 (2006).
5. C. Dominguez, L. Smith, Q. Huang, C. Yuan, X. Ouyang, L. Cai, P. Chen, J. Kim, T. Harvey, R. Syed, T. Kim, A. Tasker, L. Wang, M. Zhang, A. Coxon, J. Bready, C. Starnes, D. Chen, Y. Gan, S. Neervannan, G. Kumar, A. Polverino and R. Kendall, *Bioorg. Med. Chem. Lett.*, **17**, 6003 (2007).
6. F.M. Gabhann and A.S. Popel, *Biophys. Chem.*, **128**, 125 (2007).
7. A. Afzal, L.C. Shaw, A.V. Ljubimov, M.E. Boulton, M.S. Segal and M.B. Grant, *Microvascular*, **74**, 131 (2007).

8. H. Nakamura, Y. Sasaki, M. Uno, T. Yoshikawa, T. Asano, H.S. Ban, H. Fukazawa, M. Shibuya and Y. Uehara, *Bioorg. Med. Chem. Lett.*, **16**, 5127 (2006).

9. Q. Tong, L. Zheng, B. Li, D. Wang, C. Huang, G. M Matuschak and D. Li, *Exp. Cell Res.*, **312**, 3559 (2006).

10. E.S. Rennel, M.A. H-Zadeha, E.R. Wheatley, A. Magnussen, Y. Schuler, S.P. Kelly, C. Finucane, D. Ellison, S. Cebe-Suarez, K. Ballmer-Hofer, S. Mather, L. Stewart, D.O. Bates and S.J. Harper, *Eur. J. Cancer*, **44**, 1883 (2008).

11. A. Gangjee, O.A. Namjoshi, J. Yu, M.A. Ihnat, J.E. Thorpe, L.A. Warnke and T. Warnke, *Bioorg. Med. Chem.*, **16**, 5514 (2008).

12. D.I.R. Holmes and I.C. Zachary, *Cellular Signalling*, **20**, 569 (2008).

13. D.F. Hayes, K. Miller and G. Sledge, *Breast*, **16**, S17 (2007).

14. R. Roskoski Jr., *Biochem. Biophys. Res. Commun.*, **375**, 287 (2008).

15. H. Miao, K. Hu, X. Jimenez, E. Navarro, H. Zhang, D. Lu, D.L. Ludwig, P. Balderes and Z. Zhu, *Biochem. Biophys. Res. Commun.*, **345**, 438 (2006).

16. Y. Yamazaki, Y. Nakano, T. Imamura and T. Morita, *Biochem. Biophys. Res. Commun.*, **355**, 693 (2007).

17. Y. Miyazaki, S. Matsunaga, J. Tang, Y. Maeda, M. Nakano, J.R. Philippe, M. Shibahara, W. Liu, H. Sato, L. Wang and R.T. Nolte, *Bioorg. Med. Chem. Lett.*, **15**, 2203 (2005).

18. N.C. Becknell, A.L. Zulli, T.S. Angeles, S. Yang, M.S. Albom, L.D. Aimone, C. Robinson, H. Chang and R.L. Harkins, *Bioorg. Med. Chem. Lett.*, **16**, 5368 (2006).

19. D.I.R. Holmes and I.C. Zachary, *Cellular Signalling*, **20**, 569 (2008).

20. A. Tufro, J. Teichman, N. Banu and G. Villegas, *Biochem. Biophys. Res. Commun.*, **358**, 410 (2007).

21. K. Holmes, O.L. Roberts, A.M. Thomas and M.J. Cross, *Cellular Signalling*, **19**, 2003 (2007).

22. P. Wickens, H. Kluender, J. Dixon, C. Brennan, F. Achebe, A. Bacchiocchi, D. Bankston, D. Bierer, M. Brands, D. Braun, M.S. Brown, C. Chuang, J. Dumas, I. Enyedy, G. Hofilena, Z. Hong, T. Housley, B. Jones, U. Khire, C. Kreiman, E. Kumarasinghe, T. Lowinger, R. Perkins, B. Barton Phillips, R. Schoenleber, W.J. Scott, R. Sheeler, A. Redman, X. Sun, I. Taylor, L. Wang, S. Wilhelm, X. Zhang, M. Zhang, E. Sullivan, C. Carter, M. Miglarese and J. Levy, *Bioorg. Med. Chem. Lett.*, **17**, 4378 (2007).

23. C. Dominguez, L. Smith, Q. Huang, C. Yuan, X. Luoyang, L. Cai, P. Chen, J. Kim, T. Harvey, R. Syed, T.-S. Kim, A. Tasker, L. Wang, M. Zhang, A. Coxon, J. Bready, C. Starnes, D. Chen, Y. Gan, S. Neervannan, G. Kumar, A. Polverino and R. Kendall, *Bioorg. Med. Chem. Lett.*, **17**, 6003 (2007).

24. K. Holmes, L.R. Roberts, A.M. Thomas and M.J. Cross, *Cellular Signalling*, **19**, 2003 (2007).

25. L. Wilmes, M. Pallavicini, L.M. Fleming, J. Gibbs, D. Wang, K. Li, S.C. Partridge, R.G. Henry, D.R. Shalinsky, D. Hu-Lowe, J.W. Park, T.M. McShane, Y. Lu, R.C. Brascha and N.M. Hylton, *Magnet. Resonance Imaging*, **25**, 319 (2007).

26. R. Tripathy, A. Reiboldt, P.A. Messina, M. Iqbal, J. Singh, E.R. Bacon, T.S. Angeles, S.X. Yang, M.S. Albom, C. Robinson, H. Chang, B.A. Ruggeri and J.P. Mallamo, *Bioorg. Med. Chem. Lett.*, **16**, 2158 (2006).

27. R. Tripathy, A. Ghose, J. Singh, E.R. Bacon, T.S. Angeles, S.X. Yang, M.S. Albom, L.D. Aimone, J.L. Herman and J.P. Mallamo, *Bioorg. Med. Chem. Lett.*, **17**, 1793 (2007).

28. A.S. Kiselyov, E.L. Piatnitski, A.V. Samet, V.P. Kisliy and V. Semenov, *Bioorg. Med. Chem. Lett.*, **17**, 1369 (2007).

29. S. Huang, R. Li, P.T. Connolly, G. Xu, M.D. Gaul, S.T. Emanuel, K.R. LaMontagne and L.M. Greenberger, *Bioorg. Med. Chem. Lett.*, **16**, 6063 (2006).

30. Z. Ji, A. Ahmed, D.H. Albert, J. Bouska, P.F. Bousquet, G.A. Cunha, K.B. Glaser, J. Guo, J. Li, P.A. Marcotte, M.D. Moskey, L.J. Pease, K.D. Stewart, M. Yates, S.K. Davidsen and M.R. Michaelides, *Bioorg. Med. Chem. Lett.*, **16**, 4326 (2006).

31. S. Raeppel, S. Claridge, O. Saavedra, F. Gaudette, L. Zhan, M. Mannion, N. Zhou, F. Raeppel, M. Granger, L. Isakovic, R. Déziel, H. Nguyen, N. Beaulieu, C. Beaulieu, I. Dupont, M. Robert, S. Lefebvre, M. Dubay, J. Rahil, I. Wang, H. Ste-Croix, R.A. Macleod, J. Besterman and A. Vaisburg, *Bioorg. Med. Chem. Lett.*, **19**, 1323 (2009).

32. P.A. Harris, A. Boloor, M. Cheung, R. Kumar, R.M. Crosby, R.G. Davis-Ward, A.H. Epperly, K.W. Hinkle, R.N. Hunter III, J.H. Johnson, V.B. Knick, C.P. Laudeman, D.K. Luttrell, R.A. Mook, R.T. Nolte, C.K. Rudolph, J.R. Szewczyk, A.T. Truesdale, J.M. Veal, L. Wang and J.A. Stafford, *J. Med. Chem.*, **51**, 4632 (2008).

33. V. Gracias, Z. Ji, I. Akritopoulou-Zanze, C. Abad-Zapatero, J.R. Huth, D. Song, P.J. Hajduk, E.F. Johnson, K.B. Glaser, P.A. Marcotte, L. Pease, N.B. Soni, K.D. Stewart, S.K. Davidsen, M.R. Michaelides and S.W. Djuric, *Bioorg. Med. Chem. Lett.*, **18**, 2691 (2008).

34. A.S. Kiselyov, D. Milligan and X. Ouyang, *Bioorg. Med. Chem. Lett.*, **17**, 3550 (2007).

35. N. Ferrara, K.J. Hillan and W. Novotny, *Biochem. Biophys. Res. Commun.*, **333**, 328 (2005).

36. C. Eichelberg, R. Heuer, F.K. Chun, K. Hinrichs, M. Zacharias, H. Huland and H. Heinzer, *Eur. Urol.*, **54**, 1373 (2008).

37. R. Roskoski Sunitinib Jr., *Biochem. Biophys. Res. Commun.*, **356**, 323 (2007).

38. H. Richly, B. Schultheis, A. Adamietz, P. Kupsch, M. Grubert, R.A. Hilger, M.E. Ludwig, E. Brendel, O. Christensen and D. Strumberg, *Eur. J. Cancer*, **45**, 579 (2009).

39. R.D. Cramer, *Persp. Drug Discov. Design.*, **1**, 269 (1993).

40. G. Klebe, U. Abraham and T. Mietzner, *J. Med. Chem.*, **24**, 4130 (1994).

41. A.A. San Juan, *Eur. J. Med. Chem.*, **43**, 781 (2008).

42. M. Swale and C.G. Mohan, *J. Mol. Graph. Modell.*, **26**, 1169 (2008).

43. J. Caballero, M. Fernandez, M. Saavedra and F.D. Gonzalez-Nilo, *Bioorg. Med. Chem.*, **16**, 810 (2008).

44. P. Yi, X. Fang and M. Qiu, *Eur. J. Med. Chem.*, **43**, 925 (2008).

45. H. Cao, H. Zhang, X. Zheng and D. Gao, *J. Mol. Graph. Modell.*, **26**, 236 (2007).

46. C. Tintori, M. Magnani, S. Schenone, M. Botta and Docking, *Eur. J. Med. Chem.*, **44**, 990 (2009).

47. D.M. Sammond, K.E. Nailor, J.M. Veal, R.T. Nolte, L.V. Wang, V.B. Knick, C.K. Rudolph, A.T. Truesdale, E.N. Nartey, J.A. Stafford, R. Kumar and M. Cheung, *Bioorg. Med. Chem. Lett.*, **15**, 3519 (2005).

48. G. Bold, K. Altmann, J. Frei, M. Lang, P. W. Manley, P. Traxler, B. Wietfeld, J. Bruggen, E. Buchdunger, R. Cozens, S. Ferrari, P. Furet, F. Hofmann, G. Martiny-Baron, J. Mestan, J. Rsel, M. Sills, D. Stover, F. Acemoglu, E. Boss, R. Emmenegger, L. Lsser, E. Elvira Masso, R. Roth, C. Schlachter, W. Vetterli, D. Wyss and J.M. Wood, *J. Med. Chem.*, **43**, 2310 (2000).

49. L.F. Hennequin, E.S.E. Stokes, A.P. Thomas, C. Johnstone, P.A. Pl, D.J. Ogilvie, M. Dukes, S.R. Wedge, J. Kendrew and J.O. Curwen, *J. Med. Chem.*, **45**, 1300 (2002).

50. B.L. Podlogar and D.M. Fergusson, *Drug Ds. Discov.*, **17**, 4 (2000).

51. D.S. Puntambekar, R. Giridhar and M.R. Yadav, *Eur. J. Med. Chem.*, **41**, 12179 (2006).