

In silico screening for identification of pyrrolidine derivatives dipeptidyl peptidase-IV inhibitors using CoMFA, CoMSIA, HQSAR and docking studies

M. C. Sharma¹ · S. Jain¹ · R. Sharma¹

Received: 23 July 2017 / Accepted: 6 October 2017 / Published online: 23 October 2017
© Springer-Verlag GmbH Germany 2017

Abstract To explore the relationship between the structures of substituted pyrrolidine derivatives and their inhibition of dipeptidyl peptidase IV inhibitors. The QSAR, including CoMFA, CoMSIA and HQSAR, were applied to identify the key structures impacting their inhibitory potencies. The CoMFA, CoMSIA and HQSAR with cross-validated correlation coefficient (q^2) value of 0.727, 0.870 and 0.939 and r^2 value of 0.973, 0.981 and 0.949. Based on the structure–activity relationship revealed by the present study, we have designed a set of novel dipeptidyl peptidase IV inhibitors that showed excellent potencies in the developed models. Thus, our results allowed us to design new derivatives with desired activities.

Keywords Pyrrolidine · CoMFA · CoMSIA · HQSAR · Docking · Dipeptidyl peptidase IV

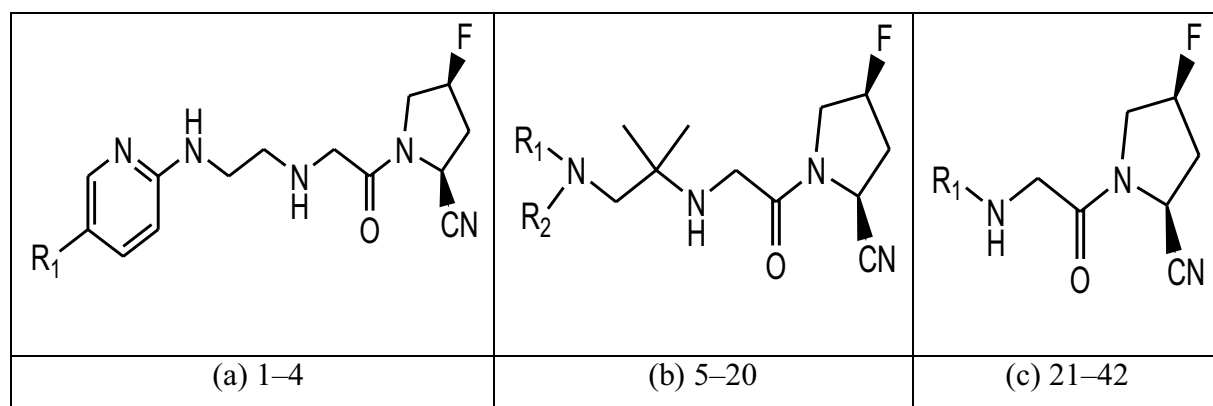
Introduction

Diabetes mellitus is a metabolic disorder, considered as a major public health issue. Million of people worldwide will suffer from type 2 diabetes mellitus (T2DM). T2DM is a chronic disease, developing later in life in most cases, characterized by several defects pancreatic β cell dysfunction, insulin resistance, increase hepatic glucose production, complications, such as stroke, coronary artery disease, hypertension, nephropathy, neuropathy and retinopathy. (Turner et al. 1996; UK Prospective 1998a,

b) Current treatment strategies include reducing insulin resistance, supplementing the insulin deficiency with exogenous insulin, enhancing endogenous insulin secretion, reducing hepatic glucose output, and limiting glucose absorption. Pioglitazone, the widely used antidiabetic agent, has recently been reported to be associated with increased risk of bladder cancer after treatment for 2 years or more, (UK Prospective 1998a, b) Therefore antihyperglycemic agents which are orally available considered to suitable for human being. The GLP-1 in plasma is rapidly degraded by the serine protease dipeptidyl peptidase IV (DPP-4). Thus inhibition of the DPP-4 has appeared as a major target of diabetes research (Holst 1999, 2005; Hui et al. 2005; Deacon et al. 1995). The marvelous interest due to orally intake of active DPP-4 inhibitors which have the potential to control blood glucose level. (Holst 1999) To date, sitagliptin, vildagliptin and saxagliptin have obtained approval as first representatives of this class of novel antidiabetic agents, closely followed by others such as alogliptin. (Hanefeld et al. 2007; Zimmet et al. 2001) Indeed, there is further requirement for more potent and safer DPPIV inhibitor, which have selectivity but does not have specificity and side effect showed by the recently available inhibitors for treatment of T2DM. In the past few decades, a large number of compounds were synthesized and biologically evaluated as DPP-IV inhibitors. (Havale and Pal 2009; Defronzo et al. 2008) Therefore, the SAR and QSAR studies were mainly focused on the presently available series. The QSAR models and docking methodology applied on a series of pyrrolidine to search the interaction between DPP-IV's inhibitors and the receptor by using Hologram quantitative structure activity relationship (HQSAR), comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) methods (Saqib and Siddiqi 2009;

✉ M. C. Sharma
mukeshcsharma@yahoo.com

¹ School of Pharmacy, Devi Ahilya University,
Indore 452 001, India

Table 1 Structures and biological activities (actual, predicted and residual pIC_{50} data) of pyrrolidine analogues used in QSAR study representing both training set and test set

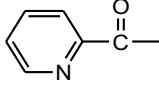
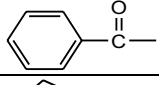
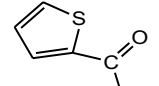
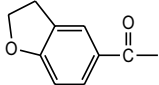
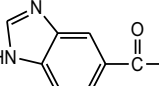
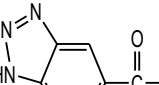
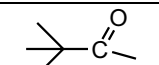
S. N.	Compounds		IC ₅₀	pIC ₅₀	CoMFA		CoMSIA		HQSAR	
					Pred	Res	Pred	Res	Pred	Res
	R ₁	R ₂								
1.	CN	–	1.1	8.959	8.891	0.068	8.606	0.353	8.71	0.249
2.	H	–	2.8	8.553	8.627	– 0.074	8.509	0.044	8.38	0.173
3.	Cl	–	2.7	8.569	8.792	– 0.223	8.539	0.03	8.27	0.299
^a 4.	CONH ₂	–	2.4	8.620	8.68	– 0.060	8.716	– 0.096	7.93	0.69
5.		H	8.2	8.086	8.026	0.060	8.051	0.035	6.89	1.196
6.		H	4.5	8.347	8.062	0.285	8.108	0.239	7.31	1.037
^a 7.		H	5.4	8.268	8.062	0.206	8.026	0.242	7.20	1.068
8.		H	2.9	8.538	8.627	– 0.089	8.283	0.255	6.59	1.948
^a 9.		H	5.7	8.244	8.375	– 0.131	8.22	0.024	7.28	0.964
10.		H	1.5	8.824	8.607	0.217	8.312	0.512	6.73	2.094
^a 11.		H	5.8	8.237	6.43	1.807	8.372	– 0.135	8.13	0.107

Table 1 (continued)

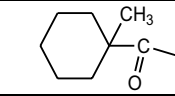
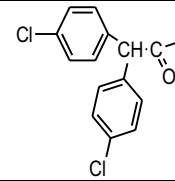
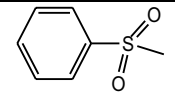
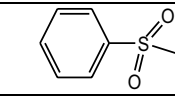
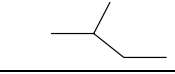
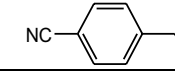
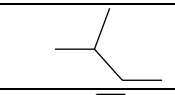
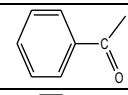
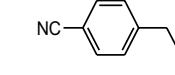
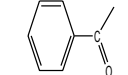
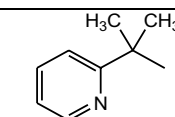
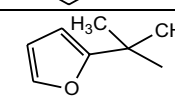
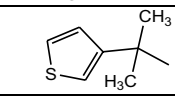
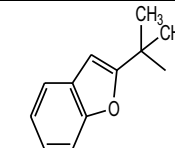
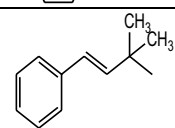
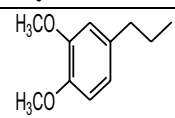
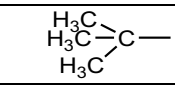
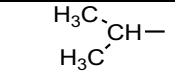
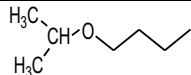
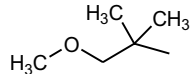
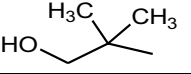
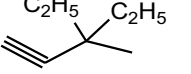
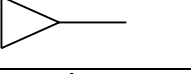



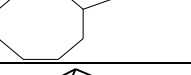




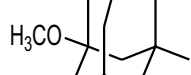
^a 12.		H	13	7.886	7.89	-0.00 4	8.235	-0.34 9	8.69	-0.80 4
^a 13.		H	109	6.963	7.943	-0.98	7.684	-0.72 1	8.316	-1.35 3
14.		H	75	7.125	7.477	-0.35 2	7.975	-0.85	7.765	-0.64
15.		CH ₃	31	7.509	7.403	0.106	7.755	-0.24 6	7.85	-0.34 1
16.		H	316	6.5	6.376	0.124	6.156	0.344	7.79	-1.29
^a 17.		H	39	7.409	8.429	-1.02	8.075	-0.66 6	7.23	0.179
^a 18.	C ₂ H ₅	C ₂ H ₅	252	6.599	6.575	0.024	5.854	0.745	7.82	-0.41 1
19.			7.3	8.137	8.199	-0.06 2	8.475	-0.33 8	7.56	0.577
20.			6.2	8.208	8.468	-0.26	8.306	-0.09 8	8.41	-0.20 2
21.		-	22	7.658	7.516	0.142	7.438	0.22	8.19	-0.53 3
^a 22.		-	8.6	8.066	7.389	0.677	7.549	0.517	7.958	0.108
23.		-	12	7.921	7.8	0.121	7.954	-0.03 3	8.63	-0.70 9
24.		-	88	7.056	7.145	-0.08 9	7.542	-0.48 6	8.049	-0.99 3
25.		-	16	7.796	7.678	0.118	7.459	0.337	8.35	-0.55 4
^a 26.		-	27	7.569	5.127	2.441	8.893	-1.32 4	6.46	1.109
27.		-	2.9	8.538	8.577	-0.03 9	8.45	0.088	7.645	0.893
^a 28.		-	7.8	8.108	7.476	0.632	7.318	0.79	8.37	-0.26 2

Table 1 (continued)

^a 29.		–	8.7	8.060	7.069	0.991	6.612	1.448	8.48	– 0.42
^a 30.		–	13	7.886	6.842	1.044	6.792	1.094	8.22	– 0.33 4
^a 31.		–	4.6	8.337	7.338	0.999	7.181	1.156	8.82	– 0.48 3
32.		–	316	6.5	6.709	– 0.20 9	6.987	– 0.48 7	8.31	– 1.81
33.		–	33	7.481	8.161	– 0.68	8.262	– 0.78 1	8.27	– 0.78 9
34.		–	3.1	8.509	8.284	0.225	8.126	0.383	8.72	– 0.21 1
35.		–	2.6	8.585	8.213	0.372	8.127	0.458	8.55	0.035
36.		–	3.3	8.481	8.563	– 0.08 2	8.14	0.341	8.22	0.261
^a 37.		–	3.3	8.481	8.194	0.287	8.127	0.354	8.36	0.121
^a 38.		–	4.1	8.387	8.203	0.184	8.17	0.217	8.36	0.027
39.		–	3.1	8.509	8.414	0.095	8.5	0.009	8.36	0.149
40.		–	3.1	8.509	8.282	0.227	8.231	0.278	8.30	0.209
41.		–	8.3	8.081	8.074	0.007	8.429	– 0.34 8	8.32	– 0.23 9
42.		–	5.3	8.276	8.278	– 0.00 2	8.484	– 0.20 8	8.38	– 0.10 4

Pred predicted, *Res* residual

^aTest set compounds

Zeng et al. 2007; Peters et al. 2004) for such a purpose, 42 compounds were selected from literature (Fukushima et al. 2008) and divided into a training data base and a test data base. The models resulting from the study provided some beneficial clues in structural modification of these

inhibitors, for further designing inhibitors for the treatment of type 2 diabetes with much improved inhibitory activities against DPP-IV inhibitors.



Fig. 1 The structural alignment of the 42 molecules (a) with their common substructure used for superimposing the compound in the data set (b) of pyrrolidine based series

Table 2 Statistics of CoMFA models of both series on different partial atomic charges

Models	Partial atomic charges	Statistics of pyrrolidine series			
		q^2	r^2	SE	NC
1.	Gasteiger	0.648	0.932	0.244	2
2.	Gast–huck	0.633	0.916	0.27	2
3.	Delre	0.727	0.973	0.15	2
4.	Pullman	0.639	0.921	0.26	2
5.	Formal-charges	0.674	0.926	0.25	2
6.	MMFF94	0.652	0.923	0.259	2

Table 3 Statistics of CoMSIA models of both series on different partial atomic charges

Models	Partial atomic charges	Statistics of pyrrolidine series			
		q^2	r^2	SE	NC
1.	Gasteiger	0.67	0.912	0.22	2
2.	Gast-huck	0.775	0.943	0.22	2
3.	Delre	0.870	0.981	0.22	2
4.	Pullman	0.617	0.972	0.22	2
5.	Formal charges	0.66	0.939	0.22	2
6.	MMFF94	0.67	0.947	0.22	2

Materials and methods

Data set

A set of 42 pyrrolidine derivatives (IC_{50} activity ranges from 1.1 to 316 nM) were collected from the work reported in literature review. (Fukushima et al. 2008) The structures of all the compounds and their biological activities were listed in Table 1 with their IC_{50} values. IC_{50} values

were transformed to corresponding pIC_{50} ($-\log IC_{50}$) values which were used as the dependent variables in the QSAR study. It has been suggested that generated models should be tested on a sufficiently large test set to establish a statistically meaningful and reliable QSAR model; therefore, the molecules were randomly divided into a training set and test set compounds in such a way that both sets cover the structural diversity, chemical prototypes and the complete range of DPP-IV inhibitory activity. Data sets divided into a training set of 26 and test set (labelled with asterisk) of 16 compounds in pyrrolidine based series (Gupta et al. 2011; Gupta and Saxena 2011; SYBYL6.9).

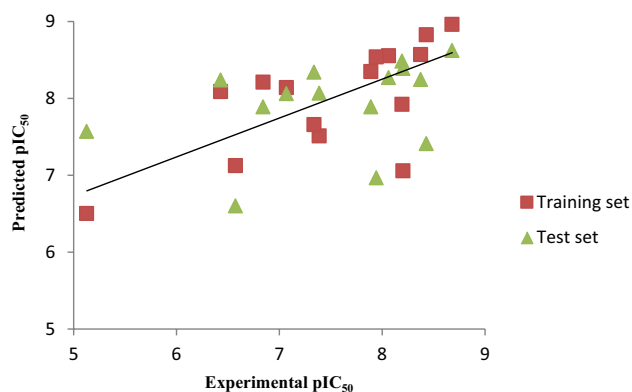
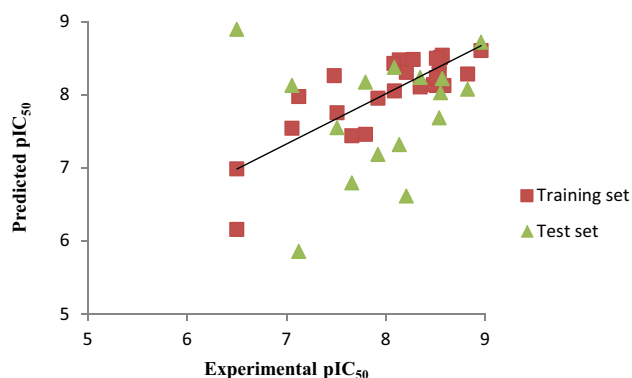
Molecular alignment

The molecular modeling studies were performed using SYBYL X 2.0 (Bush and Nachbar 1993) software running on a core-2 duo Intel processor workstation. The 3D structures of the molecules to be analyzed were aligned on a suitable conformational template, which is assumed to adopt a ‘bioactive conformation’. Hence, in this case the molecular structures of all the compounds were drawn using the most active compound (compound with highest pIC_{50}) as a template and the partial charges were calculated using Gasteiger–Huckel (Viswanadhan et al. 1989) method and geometry optimized using Tripos force field (Cramer et al. 1988) with a distance-dependent dielectric function and energy convergence criterion of 0.001 kcal/mol Å using 1000 iterations. Compound 1 of pyrrolidine based series based series with least IC_{50} value (1.1 nM) was used as the templates. CoMFA and CoMSIA models were constructed based on the structural alignments of both series shown in Fig. 1a and b.

Table 4 Summary of the CoMFA and CoMSIA statistical results for the pyrrolidine molecules of training set best model

Statistical parameters	CoMFA	CoMSIA	Field contribution (%)	CoMFA	CoMSIA
q^2 (cross-validated)	0.727	0.870	Steric	0.634	0.153
r^2	0.973	0.981	Electrostatic	0.366	0.121
F values	41.327	36.138	Hydrophobic	–	0.267
SEE	0.150	0.220	HB donor	–	0.226
N	2	2	HB acceptor	–	0.233

q^2 leave one out cross-validation correlation coefficient, r^2 conventional correlation, *SEE* standard error of estimate, *F* fischer value, *N* optimal number of component, *HB* hydrogen bond

**Fig. 2** Graph of actual vs. predicted pIC_{50} values of all compounds for training and test sets using CoMFA**Fig. 3** Graph of actual vs. predicted pIC_{50} values of both series for training and test sets using CoMSIA

Comparative molecular field analysis (CoMFA) studies

The basic assumption for CoMFA and CoMSIA is that the observed biological properties, i.e. pIC_{50} can be well correlated with the steric, electrostatic and other fields surrounding a set of ligand molecules. (Cramer et al. 1988) In CoMFA analysis, the steric and electrostatic fields were calculated at each lattice with a grid size of 2 \AA using sp^3 hybridised carbon atom with + 1 charge served as a probe atom. The CoMFA fields generated were truncated by the

default energy cutoff of 30 kcal/mol. The Gasteiger–Huckel charge model was determined as the best choice and used in the CoMFA and CoMSIA analyses (Cramer et al. 1988; Klebe et al. 1994).

Comparative molecular similarity index analysis (CoMSIA) studies

The CoMSIA descriptors, namely, steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor, were generated using a sp^3 hybridized carbon atom with + 1 charge, Vanderwaal's radius of 1.4 \AA and hydrophobic and hydrogen bond properties of + 1. CoMSIA similarity indices ($A_{F,K}^q$) between a molecule j and atoms i at a grid point were calculated by using Eq. 1 as follows:

$$A_{F,K}^q(J) = - \sum_{i=1}^n W_{\text{probe},k} W_{ik} e^{-\alpha r_{iq}^2} \quad (1)$$

where q represents the grid point, i is the summation index, over all atoms of the molecule j under computation, W_{ik} is the actual value of the physicochemical property k of atom i , and $W_{\text{probe},k}$ is the value of the probe atom (Pirhadi and Ghasemi 2010; Zhao et al. 2011). Five physicochemical properties steric, electrostatic, hydrophobic, hydrogen bond donor and hydrogen bond acceptor were evaluated. A Gaussian-type distance dependence was used between the grid point q and each atom i in the molecule. The value of the attenuation factor was set to 0.3. (Klebe et al. 1994).

PLS calculations and validations

Partial least square (PLS) regression analysis (Bush and Nachbar 1993) was used to quantify the relationship between DPP-IV inhibitory activity and structural parameters of CoMFA and CoMSIA by using dependent and independent variables, respectively. The optimum number of components (ONC), was the number of components resulting in the highest cross-validated correlation coefficient (r_{cv}^2), which was defined as follows:

$$r_{cv}^2 = 1 - \frac{\sum (Y_{\text{obs}} - Y_{\text{pred}})^2}{\sum (Y_{\text{obs}} - Y_{\text{mean}})^2}$$

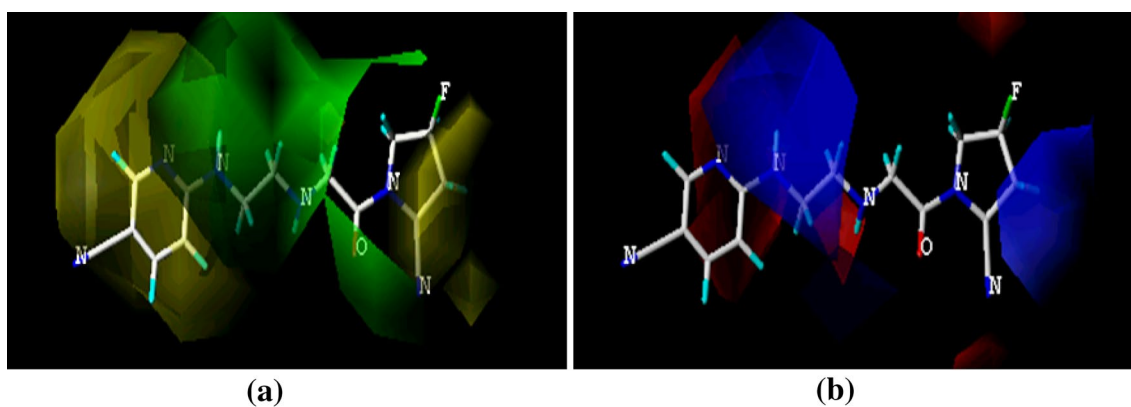


Fig. 4 CoMFA steric and electrostatic contour maps for compound 1 of pyrrolidine based series; **a** steric CoMFA contour of compound 1, **b** electrostatic CoMFA contour of comp. 1

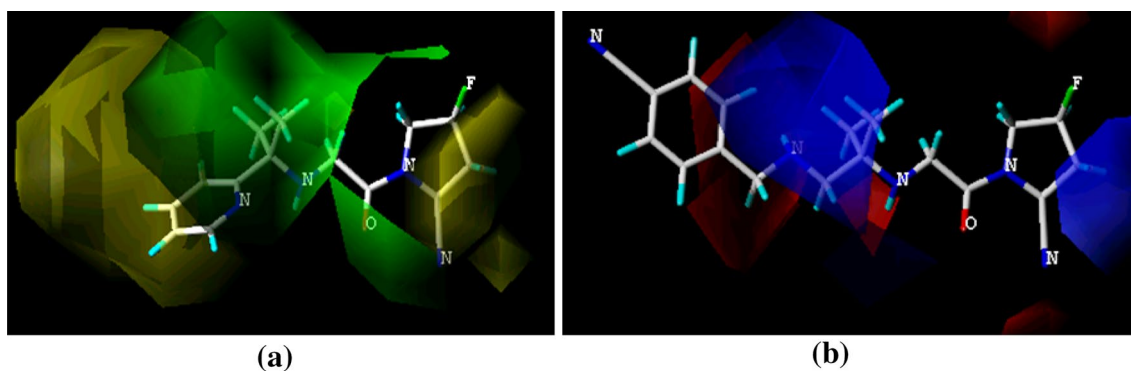


Fig. 5 CoMFA steric and electrostatic contour maps for compound 17 (compound with best dock score) of pyrrolidine based series: **a** steric CoMFA contour of compound 17, **b** electrostatic CoMFA contour of comp. 17

where Y_{Pred} , Y_{obs} and Y_{mean} are predicted, observed and mean values of the target property (pIC_{50}). The predictive r^2 value, based on the test set molecules, is computed by the following formula:

$$r_{\text{pred}}^2 = \text{SD} - \text{PRESS}/\text{SD}$$

where SD is the sum of squared deviation between the biological activities of the test set molecule to the mean activity of the training set molecules while PRESS is the sum of squared deviations between the observed and the predicted activities of the test molecules (Pirhadi and Ghasemi 2010; Zhao et al. 2011; Ai et al. 2011; Ping et al. 2011).

HQSAR studies

HQSAR is a novel technique requires 2D structure which employs specialized fingerprints (molecular holograms) as predicted variable of pharmacological activity. Concerned

sensitivity of the generated model depends on the parameters as hologram length, Fragment size and fragment distinct. The analysis was based on the fragment distinction and fragment size (Sridhara et al. 2011). For the model development, the first and foremost condition is to distinguish molecular fragments based on atoms (A), bonds (B), connections (C), hydrogen atoms (H), chirality (Ch), and donor and acceptor (DA). The statistical parameters obtained after completion of analysis were number of component (NC), cross-validated q^2 , conventional r^2 , Standard error (SE), Best Hologram length (BHL) and predicted pIC_{50} , which provides the useful information on the basis of best model subsequently fingerprint structure saved with color coding.

Molecular docking

The purpose of docking studies was to generate 3D structures of ligands with appropriate binding orientations and conformations. Molecular docking analysis was carried out using the Surflex Dock in SYBYL X 2.0 to explore possible

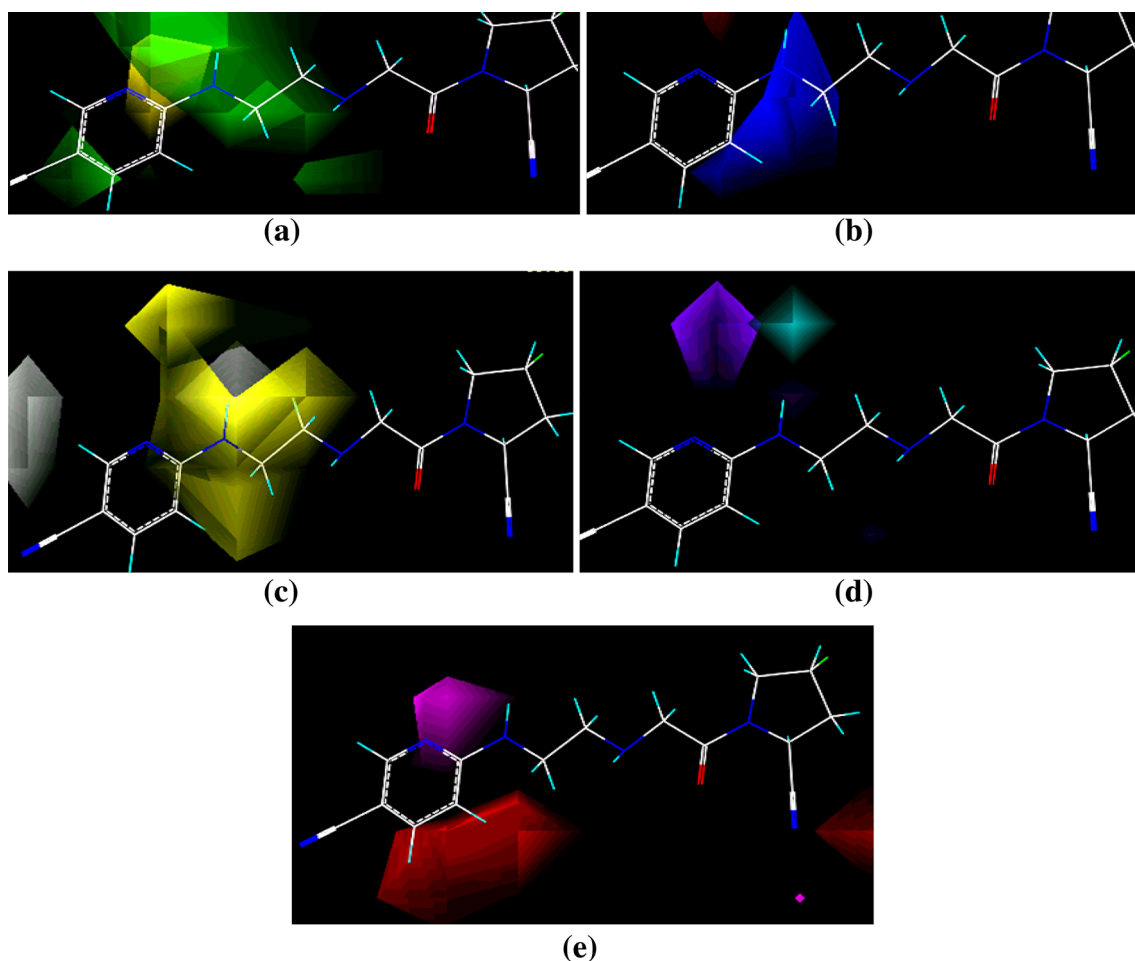


Fig. 6 CoMSIA steric, electrostatic, hydrophobic, H-bond donor, H-bond acceptor contour maps for compound 1 of pyrrolidine based series; **a** steric CoMSIA contour of comp. 1, **b** electrostatic contour

of comp. 1, **c** hydrophobic contour of comp. 1, **d** H-bond donor contour of comp. 1, **e** H-bond acceptor contour of comp. 1

binding conformations and understanding the interactions of dipeptidyl peptidase-IV receptor with various compounds. The crystal structure of dipeptidyl peptidase-IV (DPP-IV) was retrieved from RCSB Protein Data Bank (PDB entry code: 2G5P). The protein structures were utilized in subsequent docking experiments without energy minimization. All ligands and water molecules were removed at first, the polar hydrogen atoms and AMBER7FF99 charges were added. The protomol bloat value was set as 1 and the protomol threshold value as 0.5 when a reasonable binding pocket was obtained.

Results and discussion

CoMFA and CoMSIA techniques were used to derive 3D-QSAR models on a set of 42 chemically diverse pyrrolidine analogues as DPP-IV inhibitors. The best

statistical parameters associated in CoMFA and CoMSIA models are listed Tables 2 and 3. The CoMFA models derived from the 26 pyrrolidine training compounds using both steric and electrostatic fields gave a cross-validated correlation coefficient (q^2) of 0.727 with an optimized component of 2 (Table 4). The best CoMSIA model included steric, electrostatic hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields and gave a q^2 of 0.870 with an optimized component of 2. Contributions of steric, electrostatic hydrophobic, hydrogen bond donor and hydrogen bond acceptor fields were 0.153, 0.121, 0.267, 0.226 and 0.233, respectively. The external predictive ability of the generated CoMFA model of pyrrolidine analogues was evaluated for the test set of 16 molecules, where the obtained predictive r^2 value (r^2_{pred}) of 0.655 supported the high predictive ability of the generated model. Similarly, the CoMFA model, the CoMSIA model of pyrrolidine analogues also showed the lower external predictive ability

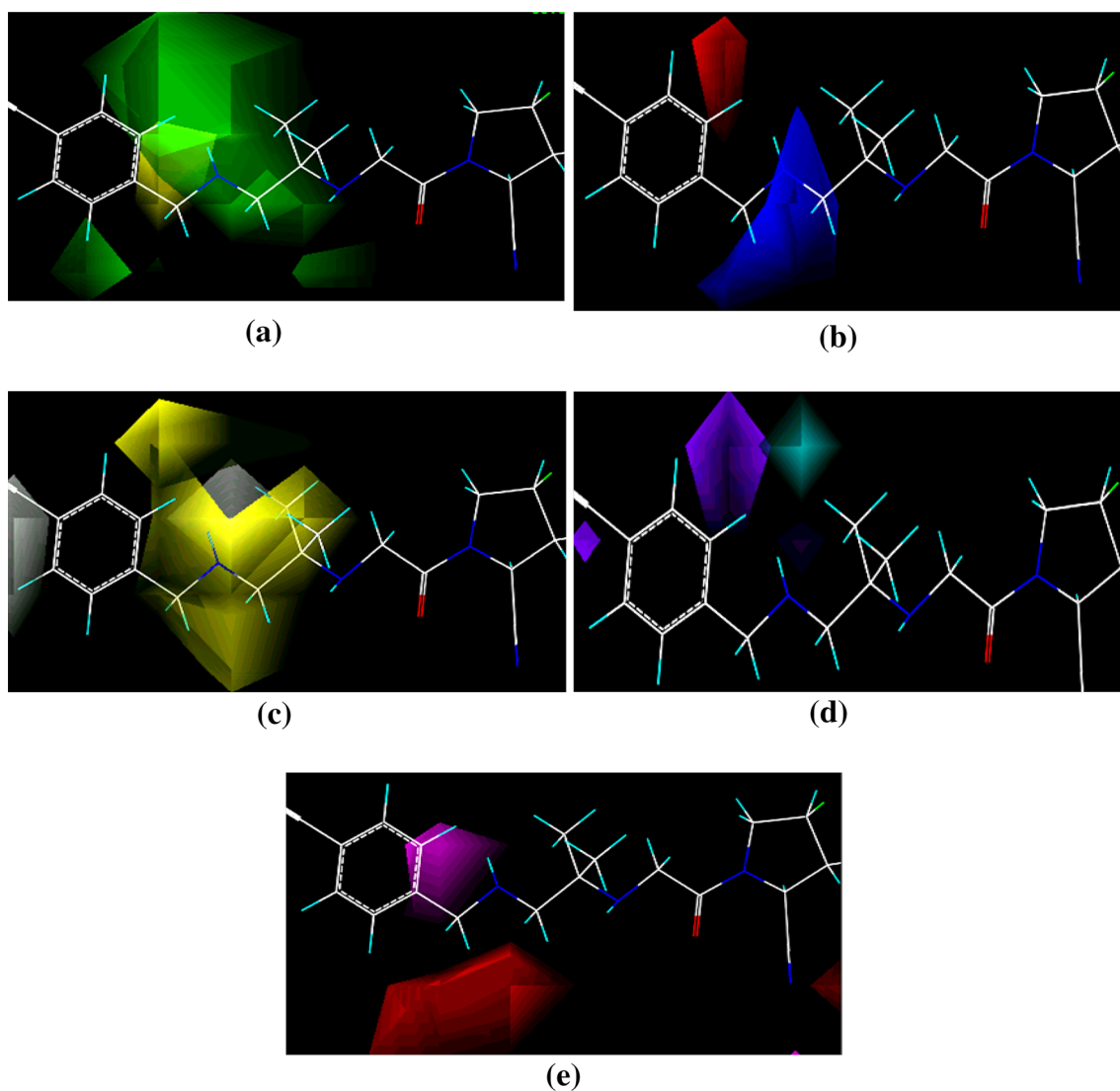


Fig. 7 CoMSIA steric, electrostatic, hydrophobic, H-bond donor, H-bond acceptor contour maps for compound 17. **a** Steric, **b** electrostatic, **c** hydrophobic contour, **d** H-bond donor, **e** H-bond acceptor

Table 5 The determination of statistical parameters for the models of pyrrolidine series based on different fragment distinct with default fragment size (4–7)

Model	Fragment size	Fragment distinct	q^2	r^2	SE	BHL	NC	Ensemble q^2
1.	4–7	A/B	0.859	0.859	0.254	257	6	0.827
2.	4–7	A/B/C	0.920	0.920	0.191	353	6	0.898
3.	4–7	A/C	0.910	0.910	0.202	151	6	0.893
4.	4–7	A/B/C/H	0.939	0.949	0.166	353	6	0.925
5.	4–7	B/C/H	0.903	0.903	0.210	353	6	0.886
6.	4–7	A/B/C/H/Ch	0.938	0.938	0.168	307	6	0.918
7.	4–7	B/C/H/Ch	0.914	0.914	0.198	353	6	0.895
8.	4–7	C/H/Ch	0.903	0.903	0.210	353	6	0.886
9.	4–7	B/C/H/Ch/DA	0.926	0.926	0.184	307	6	0.905
10.	4–7	A/B/C/H/Ch/DA	0.928	0.928	0.181	353	6	0.903
11.	4–7	C/H/Ch/DA	0.928	0.928	0.180	257	6	0.904
12.	4–7	H/Ch/DA	0.780	0.780	0.317	307	6	0.763

NC number of component, SE standard error, BHL best hologram length, A atom, B bond, C connections, H hydrogen, Ch chirality, DA donor and acceptor

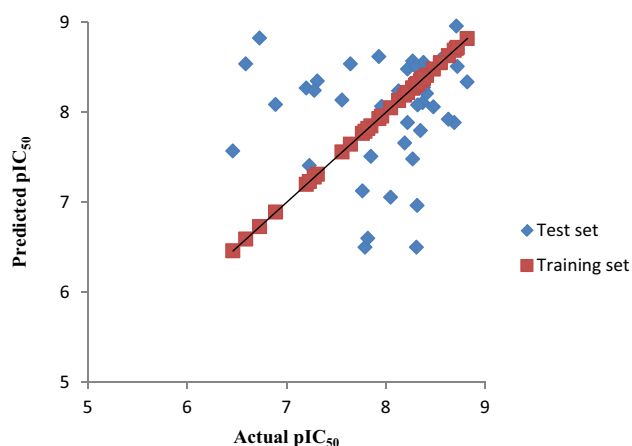


Fig. 8 Graph of actual vs. predicted pIC₅₀ values of both series for training and test sets using HQSAR

(r_{pred}^2) of 0.604 for the external test set. The predicted activity, residual value showed in Table 1 and predicted activity correlation for CoMFA and CoMSIA in Figs. 2 and 3 respectively. The generated contour maps from the above CoMFA analyses are shown in Figs. 4 and 5. As shown in above Fig. 4, an analysis of the CoMFA contour in terms of common steric and electrostatic parameters around the most active molecule of the dataset, compound 1, signified the importance of a sterically favorable region (green contour) near the 2nd position of the substituted cyanopyridine

ring of the molecule, while sterically unfavorable contours (yellow contour) for DDP-IV inhibitory activity were observed near the 2nd and 3rd position of the pyrrolidine ring as well as around the substituted cyanopyrrolidine ring. As shown in Fig. 5 of compound 17, the large green contour was present near the 1st position of substituted cyanophenyl ring demonstrated that bulky group at this position would enhance the activity while the yellow contour was located adjacent to green contour around this ring and the 2nd, 3rd and 4th position of the pyrrolidine ring suggesting steric restriction near these positions. As shown in Fig. 5b, the blue contour near the 2nd position of substituted cyanophenyl ring and near the 3rd position of the pyrrolidine ring demonstrated that electron donating group was required to enhance biological activity at these positions while red contour around the 3rd position of the substituted cyanophenyl ring suggested that electron withdrawing group would be essential to enhance biological activity at this position. CoMSIA, the hydrophobic fields are represented by yellow and white-colored contours (yellow, favored; white, disfavored); the hydrogen bond donor fields are indicated by cyan and purple-colored contoured contours (cyan, favored; purple, disfavored); while the hydrogen bond acceptor fields are denoted by magenta and red contours (magenta, favored; red, disfavored). As shown in above Fig. 6a, 5-cyano group of substituted pyridine ring and 2nd position of this ring marked with green color while 1st position of the ring marked with yellow color. In Fig. 6b, the 2nd and 3rd position of substituted pyridine

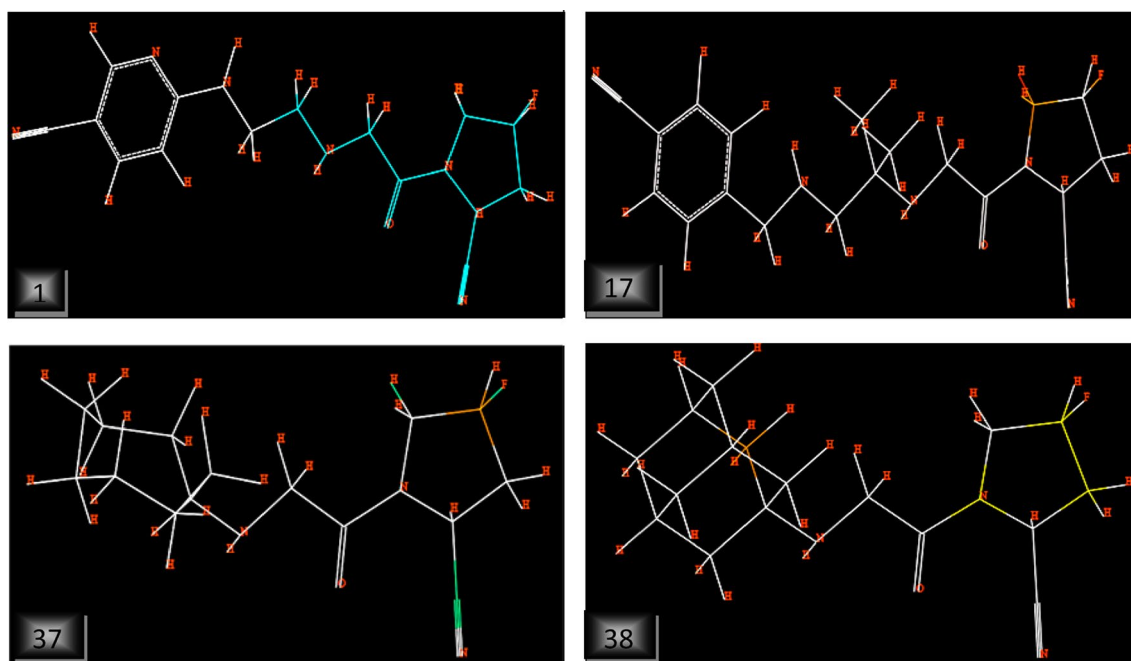


Fig. 9 The compounds (1, 17, 37 and 38) contributing map of pyrrolidine based series which shows the direct relation between the structural fragment and pharmacological activity

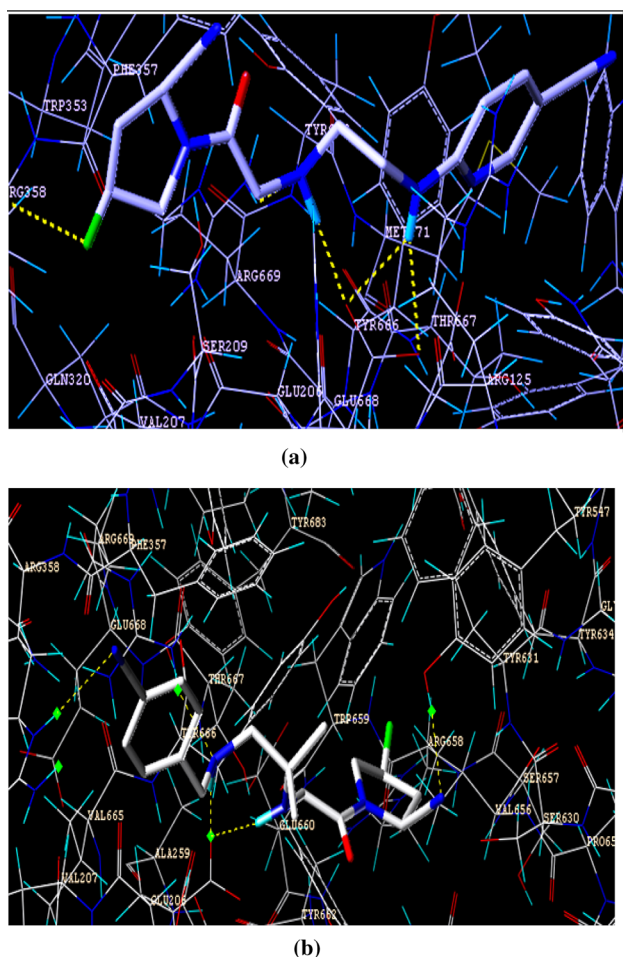


Fig. 10 Stereo-view of the docked conformations of compound 1 (a) and 17 (b) of pyrrolidine based series in the active site of DPP-IV enzyme (2G5P)

ring marked with blue color while red contour was present on the top of this ring. In Fig. 6c, huge yellow contour covered 2nd and 3rd position of the ring. In Fig. 6d, purple contour was present on the top portion of the ring. As shown in above Fig. 7a of compound 17, the 1st, 2nd, 5th and 6th position of substituted cyanopyridine ring marked with green color while yellow contour covered the 1st position of this ring. In electrostatic contour (Fig. 7b), 6th position of this ring marked with red color while blue contour covered the 1st position of this ring. In hydrophobic contour (Fig. 7c), huge yellow contour covered 1st and 5th position of the ring while white contour was present near the 3rd and 4th position of the ring. In H-bond donor contour (Fig. 7d), purple contour covered the 6th position of the ring while cyan contour was present on the top of the amino group. In H-bond acceptor contour (Fig. 7e), magenta contour around the C-2 and C-3 atom of this ring indicated that these atoms act as hydrogen bond acceptor. HQSAR models were predicted on the series of DPP IV

inhibitors and relates with structural fragments (Table 5). The best statistical results among all models using the training set compounds of pyrrolidine based series were obtained for model 1 fragment distinct A/B/C/H ($q^2 = 0.939$, $r^2 = 0.949$) which was resultant using the combination of fragment size (4–7), with six optimum number of components, 353 hologram length and 0.166 as standard error of estimation. The internal consistency plot Fig. 8, available in the forms of q^2 and r^2 , is important and significant, to predict external activity of compounds. The green and yellow color (represents positive contribution), white color (indicates intermediate or moderate contribution), red and orange (negative contribution) and cyan color (represents maximal common structure) suggest the structure fragment requirement for enhancing the binding affinity. As shown in above Fig. 9, in compound 1, the pyrrolidine portion is a common backbone to all compounds into the training set, and therefore, is colored cyan. In compound 17, 5th position of the pyrrolidine ring marked with red colour showed that electropositive group is required at this position to enhance activity. In compound 37, 2nd position of the ring marked with green color showed that bulky group is required to enhance the activity at this position. In compound 38, 1st, 3rd and 4th position of the pyrrolidine ring marked with yellow color showed that no bulky group is required to enhance activity at these positions. The most potent compounds 1 and 17 were selected according to docking scores and performed deeper docking study discussed in Fig. 10. The hydrogen bonds are shown by yellow color broken lines. As shown in Fig. 10, the important amino acid residues which form interaction with compound 1 were Arg-358, Tyr-666 with 2G5P receptor. The interaction of Arg-358 amino acid with 4-fluoro group of pyrrolidine ring is favored for bulkiness. In Fig. 10b, the important amino acid residues which form interaction with compound 17 were Arg-658, Glu-206 and Tyr-667 with 2G5P receptor.

Conclusion

The present work elucidated the successful application of the combination of three different computational approaches (CoMFA, CoMSIA and HQSAR) and molecular docking analysis to identify the essential structural requirements in 3D chemical space for the modulation of DPP-4 inhibitory activity of pyrrolidine derivatives. The CoMFA and CoMSIA contour maps of pyrrolidine analogues showed that the electron donating group at 3rd position of pyrrolidine ring increases the activity while the electron withdrawing groups are favored at 4th and 5th position of pyrrolidine ring. The binding mode of the high active compound at the active

site of DPP-4 Inhibitor (PDB: 2G5P) was explored and hydrogen-bonding interactions were observed between the inhibitor and the target. These results will serve as a useful guideline for designing the novel compounds with desired DPP-IV inhibitory activity.

Acknowledgements The authors are thankful to Head, School of Pharmacy, DAVV Indore for providing necessary facilities. Shikha Jain thanks the All India Council for Technical Education (AICTE), New Delhi, India, for the financial support for this research.

References

- Ai Y, Wang ST, Sun PH, Song FJ (2011) Combined 3D-QSAR modeling and molecular docking studies on pyrrole-indolin-2-ones as aurora kinase inhibitors. *Int J Mol Sci* 12:1605–1624
- Bush BL, Nachbar RB Jr (1993) Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. *J Comput Aided Mol Des* 7:587–619
- Cramer RD, Patterson DE, Bunce JD (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 110(18):5959–5967
- Deacon CF, Johnsen AH, Holst JJ (1995) Degradation of glucagon-like peptide-1 by human plasma *in vitro* yields an N-terminally truncated peptide that is a major endogenous metabolite *in vivo*. *J Clin Endocrinol Metab* 80(3):952–957
- DeFronzo RA, Okerson T, Viswanathan P, Guan X, Holcombe JH, MacConell L (2008) Effects of exenatide versus sitagliptin on postprandial glucose, insulin and glucagon secretion, gastric emptying, and caloric intake: a randomized, cross-over study. *Curr Med Res Opin* 24:2943–2952
- Fukushima H, Hiratate A, Takahashi M, Mikami A, Saito-Hori M, Munetomo E, Kitano K, Chonan S, Saito H, Suzuki A, Takao Y, Yamamoto K (2008) Synthesis and structure-activity relationships of potent 4-fluoro-2-cyanopyrrolidine dipeptidyl peptidase IV inhibitors. *Bioorg Med Chem* 16(7):4093–4106
- Gupta AK, Saxena AK (2011) 3D-QSAR CoMFA and CoMSIA studies on a set of diverse-adrenergic receptor antagonists. *Med Chem Res* 20:1455–1464
- Gupta AK, Bhunia SS, Balaramnavar VM, Saxena AK (2011) Pharmacophore modelling, molecular docking and virtual screening for EGFR (HER 1) tyrosine kinase inhibitors. *SAR QSAR Environ Res* 22(3):239–263
- Hanefeld M, Herman GA, Wu M, Mickel C, Sanchez M, Stein PP (2007) Once-daily sitagliptin, a dipeptidyl peptidase-4 inhibitor, for the treatment of patients with type 2 diabetes. *Curr Med Res Opin* 23:1329–1339
- Havale SH, Pal M (2009) Medicinal chemistry approaches to the inhibition of dipeptidyl peptidase-4 for the treatment of type 2 diabetes. *Bioorg Med Chem* 17:1783–1802
- Holst JJ (1999) Glucagon-like peptide 1 (GLP-1): an intestinal hormone, signaling nutritional abundance, with an unusual therapeutic potential. *Trends Endocrinol Metab* 10(6):229–235
- Holst JJ (2005) Therapy of type 2 diabetes mellitus based on the actions of glucagonlike peptide-1. *Diabetes Metab Res Rev* 18(6):430–441
- Hui H, Zhao X, Perfetti R (2005) Structure and function studies of glucagon-like peptide-1 (GLP-1): the designing of a novel pharmacological agent for the treatment of diabetes. *Diabetes Metab Res Rev* 21(4):313–331
- Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J Med Chem* 37:4130–4146
- Peters JU, Weber S, Kritter S, Weiss P, Wallier A, Boehringer M, Hennig M, Kuhn B, Loeffler BM (2004) Aminomethylpyrimidines as novel DPP-IV inhibitors: a 10(5)-fold activity increase by optimization of aromatic substituents. *Bioorg Med Chem Lett* 14(6):1491–1493
- Ping L, Chen WN, Chen WM (2011) Molecular modeling studies on imidazo [4,5-b]pyridine derivatives as Aurora A kinase inhibitors using 3D-QSAR and docking approaches. *Eur J Med Chem* 46:77–94
- Pirhadi S, Ghasemi JB (2010) 3D-QSAR analysis of human immunodeficiency virus entry-1 inhibitors by CoMFA and CoMSIA. *Eur J Med Chem* 45:4897–4903
- Saqib U, Siddiqi MI (2009) 3D-QSAR studies on triazolopiperazine amide inhibitors of dipeptidylpeptidase-IV as anti-diabetic agents. *SAR QSAR Environ Res* 20:519–535
- Sridhara J, Foroozesh M, Stevens CLK (2011) A QSAR models of cytochrome P450 enzyme 1A2 inhibitors using CoMFA, CoMSIA and HQSAR. *SAR QSAR Environ Res* 22:681–697
- SYBYL6.9, Tripos Inc., St. Louis, MO, USA
- Turner R, Cull C, Holman R (1996) United Kingdom Prospective Diabetes Study 17: a 9-year update of a randomized, controlled trial on the effect of improved metabolic control on complications in non-insulin-dependent diabetes mellitus. *Ann Intern Med* 124:136–145
- UK Prospective Diabetes Study (UKPDS) Group (1998a) Intensive blood-glucose control with sulfonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet* 352:837–853
- UK Prospective Diabetes Study (UKPDS) Group (1998b) Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). *Lancet* 352:854–865
- Viswanadhan VN, Ghose AK, Revankar RG, Robins RK (1989) Atomic physicochemical parameters for three dimensional structure directed quantitative structure activity relationships. 4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally occurring nucleoside antibiotics. *J Chem Inf Comput Sci* 29(3):163–172
- Zeng J, Liu G, Tang Y, Jiang H (2007) 3D-QSAR studies on fluoropyrrolidine amides as dipeptidyl peptidase IV inhibitors by CoMFA and CoMSIA. *J Mol Model* 13:993–1000
- Zhao X, Chen M, Huang B, Ji H, Yuan M (2011) Comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) studies on $\alpha(1A)$ -adrenergic receptor antagonists based on pharmacophore molecular alignment. *Int J Mol Sci* 12:7022–7037
- Zimmet P, Alberti KG, Shaw J (2001) Global and societal implications of the diabetes epidemic. *Nature* 414:782–787