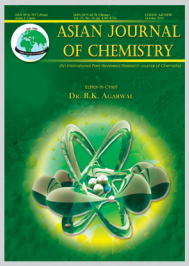
Asian Journal of Chemistry; Vol. 24, No. 1 (2012), 238-248

ASIAN JOURNAL OF CHEMISTRY

www.asianjournalofchemistry.co.in



Studies of Phenylethynyl-pyrrolo[1,2-a]pyrazine as mGluR5 Antagonists Using 3D-QSAR Method

JING LIU[1], YAN LI[1,*], SHUWEI ZHANG[1], CHUNZHI AI[2] and YULIAN YAN[1]

[1]School of Chemical Engineering, Dalian University of Technology, Dalian 116012, Liaoning, P.R. China
[2]Lab of Pharmaceutical Resource Discovery, Dalian Institute of Chemical Physics, Graduate School of the Chinese Academy of Sciences, Dalian 116023, Liaoning, P.R. China

*Corresponding author: Fax: +86 411 84896063; Tel: +86 411 84896062; E-mail: liujing198509@gmail.com

(*Received*: 9 December 2010; *Accepted*: 15 September 2011) AJC-10409

In recent years, interest has been paid in development of compounds with high biological activity for metabotropic glutamate receptors (mGluRs), an interesting therapeutic target in the treatment of cocaine seeking behaviour. In the present work, based on a data set of 84 collected phenylethynyl-pyrrolo[1,2-a]pyrazine mGluR5 antagonists with diverse kinds of structures, a variety of in silico modeling approaches including comparative molecular field analysis (CoMFA), comparative similarity indices analysis (CoMSIA) were carried out to reveal the requisite 3D structural features for activity. Present results show that both the optimal ligand-based CoMFA ($Q^2 = 0.53$, $R^2_{ncv} = 0.92$, $R^2_{pre} = 0.80$, SEE = 0.26, SEP = 0.44) and CoMSIA ($Q^2 = 0.51$, $R^2_{ncv} = 0.85$, $R^2_{pre} = 0.80$, SEE = 0.36, SEP = 0.42) models are reliable with proper predictive capacity. In addition, the analysis about the CoMFA and CoMSIA contour maps shows that: (1) Electropositive groups in Ar substituent are beneficial to enhance the activity; (2) R substituent with HB acceptor also leads to high activity; (3) Bulky R and Ar substituents are not favoured in mGluR5; (4) R substituent with hydrophilic group can improve the biological activity. All these results might provide information for better understanding of the mechanism of antagonism and thus be helpful in design of new potent mGluR5 antagonists.

Key Words: 3D-QSAR, mGluR5, Antagonist, CoMFA, CoMSIA.

INTRODUCTION

In the CNS, L-glutamate is one of the major excitatory neurotransmitters and mediates its actions through activation of both ionotropic and metabotropic receptor families[1]. mGluRs belong to a large subfamily of GPCRs[2]. During the molecular cloning it has revealed that there are eight distinct subtypes of mGluRs called mGluR1-mGluR8, which can be divided into three groups based on the sequence similarities, agonist profiles and preferential signal transduction pathways activated in heterologous systems[2]. Group I mGluRs (mGluR1 and mGluR5) are coupled to the phospholipase C signal transduction pathway, while group II (mGluR2 and mGluR3) and group III (mGluR4, mGluR6, mGluR7 and mGluR8) mGluRs are negatively coupled to adenylate cycase activity[3]. In all of these mGluRs excessive activation of mGluR5 has been implicated in a number of CNS disorders such as the pain, anxiety, depression, drug dependence and mental retardation[4]. Actually, mGluR5 may play a role in the pathophysiology and/or pharmacotherapy of schizophrenia and researches reveal the enhancement of mGluR5 function may produce antipsychotic effects[5]. Some other studies reveal that mGluR5 is also involved in the spatial learning and memory in mice model[6] as well as the therapeutic potential for drug addictions[7]. Therefore, as a potential drug interaction target the mGluR5 is very attractive for research in curing a variety of CNS disorders.

During the last several years in order to develop pharmacotherapies for treatment of the CNS disorders, great efforts have been directed toward the discovery and development of strong mGluR5 antagonists. In 1999, MPEP, the first potent, selective and systemically active mGluR5 receptor antagonist (Fig. 1A) was synthesized[8]. The molecular structure can be divided into three different molecular segments. Part I is a 6-methyl-pyridine ring, part III is a benzene, with part II as a linker between parts I and III. Using MPEP as a template, a number of novel and selective mGluR5 antagonists were synthesized, with most of the structural modifications performed on above three regions. For example, to reveal the structure-activity relationships for part II, a series of pyridinyl-alkynes as the mGluR5 antagonists were sythesized and studied (Fig. 1B)[9]. Some researchers also changed part I region to some other heterocyclic substituents and part III to various aryl rings hoping to

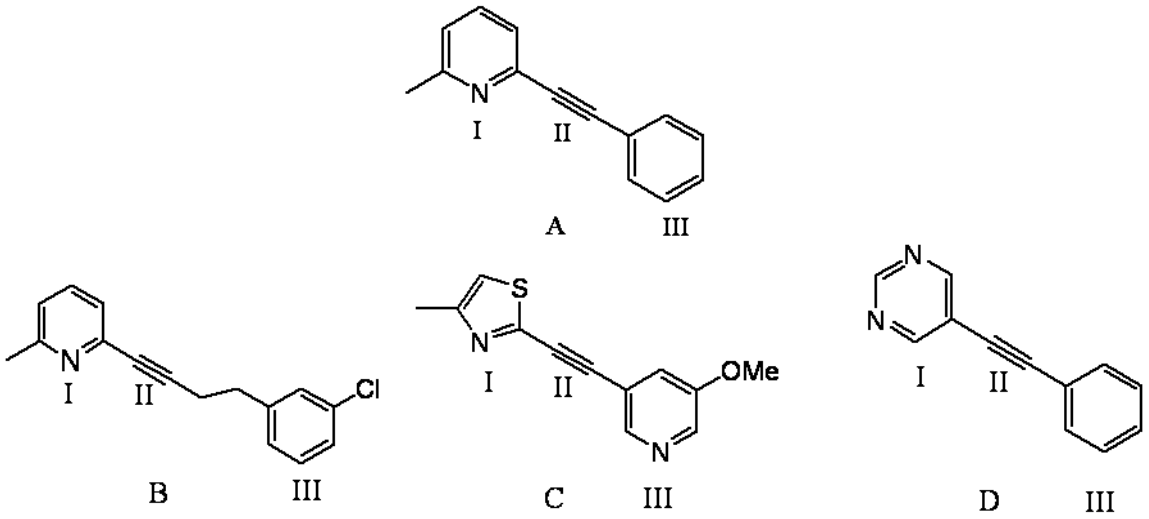


Fig. 1. Molecular structures of the mGluR5 antagonists. (A) Structures of MPEP. (B) A sample of mGluR5 antagonists with structural modifications in part II. (C) and (D) samples of mGluR5 antagonists with changes in parts I and III[8-11]

get more powerful mGluR5 antagonists (Fig. 1C-D)[10-12]. Up to date, mGluR5 as one of the Glu receptors is evidently an increasingly interesting and challenging research area[13].

CADD has been generally accepted and widely applied in the area of modern drug discovery and design for its high efficiency in the design of new compounds and optimization of lead chemicals, thus saving both time and economic costs in the large-scale experimental synthesis and biological tests[14]. QSARs as one of the typical discipline in the area of CADD have been applied widely throughout the world, like in the studies of catechol-*o*-methyltransferase inhibitors[15], to prioritise untested chemicals for more intensive and costly experimental evaluations[16]. QSARs attempt to model complex non-linear relationships between the chemical and physical properties of molecules and their biological activity and modern QSAR techniques employ advanced 2D molecular finger-prints and 3D molecular descriptors coupled with machine learning[17]. However, despite the extensive applications of QSAR study in biological field, up to now, there is no 3D-QSAR study on mGluR5 antagonists.

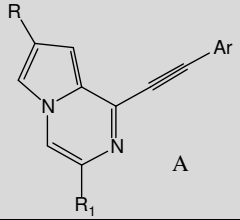
Recently, based on MPEP, a new and up-to-now the largest group of 84 fused phenylethynyl-pyrrolo[1,2-a]pyrazine as selective and potent mGluR5 antagonist were reported[18], containing five similar skeleton types (A-E) (Tables 1-3). Compared with compounds B-D in Fig. 1, these new mGluR5 antagonists are different because they not only have a pyrrolo-[1,2-a]pyrazine substituent in part I, but also possess more systematic substituents in the aryl ring of part III. Therefore, the in silico study on these specific series of compounds may aid in the identification of their requisite structural features for mGluR5 antagonists and thus to explore the possible interaction mechanisms. Thus, in the present work, a combinational computational study on these new fused phenylethynyl-pyrrolo[1,2-a]pyrazine compounds was carried out using the CoMFA, CoMSIA methods. As far as we know, this work provides the first 3D-QSAR analysis for mGluR5 antagonists, which we hope might be helpful for design and screening of novel and potent mGluR5 antagonists.

EXPERIMENTAL

Compounds and activity: A total of 105 phenylethynyl-pyrrolo[1,2-a]pyrazine derivatives prepared as the mGluR5 antagonists[18] were collected in this study, with their *in vitro* inhibitory concentrations (IC_{50}) of the molecules against mGluR5 converted into corresponding pIC_{50} ($-\log IC_{50}$) values and employed as the biological activity (Tables 1 and 2). This data set contains 21 compounds with uncertain pIC_{50} values and thus can not be used in the building of 3D-QSAR models. So we deleted them. In addition, due to with no common sub-structure (Fig. 2, depicted in red), molecules **1** and **2** were unable to be aligned to the others and were deleted from the data set during the molecular alignment process. Finally in a ratio of about 3:1, all 82 compounds with defined pIC_{50} values were divided into a training set of 61 molecules for generating the subsequent QSAR model and a test set of 21 molecules for evaluating the predictive quality of the model. Compounds belonging to the test set were selected in such a way that the structural diversity and wide range of activity in the data set were ensured being included. The pIC_{50} values are considered as dependent variables in the CoMFA and CoMSIA analyses. Energy minimization was performed on SYBYL 6.9 package (Tripos Associates, St. Louis, MO), tripos force field was used and conjugate gradient method with convergence criterion was set as 0.05 kcal/mol in this process. Partial atomic charges were calculated by the Gasteiger-Huckel method[19].

Conformational sampling and alignment: One of the most important adjustable parameters in CoMFA and CoMSIA is the relative alignment of all the compounds to one another so that they have a comparable conformation and a similar orientation of pharmacophoric groups in space[20]. It is preferable to choose an alignment which maintains the bioactive conformation[21]. In the present study, compound **68** had the highest pIC_{50} values ($pIC_{50} = 9.4$) and thus was chosen as the template molecule. And the rest molecules in the data set were aligned to the template by the ALIGN DATABASE command in SYBYL based on an atom-by-atom superimposition

TABLE-1
STRUCTURE OF COMPOUNDS WITH SKELETON TYPE A IN THE DATA SET



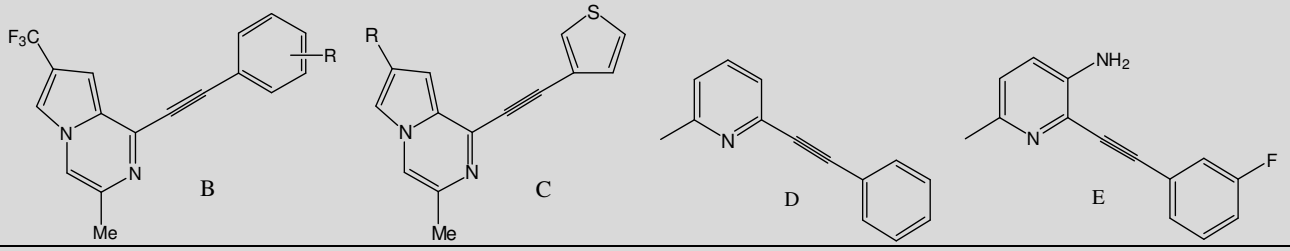
A

No.	R_1	R	Ar	pIC_{50}	No.	R_1	R	Ar	pIC_{50}
3[x,m]	CH_3	H	Ph	8.6	61	Me	CF_3	3-(1-Methyl-1*H*-pyrazol-4-yl) phenyl	<4.9
4[x,m]	CH_3	H	3-Me-Ph	9.4	62	Me	CF_3	3-(3,5-Dimethyl-4-isoxazolyl) phenyl	<4.9
5	CH_3	H	4-F-Ph	7.8	63	Me	CF_3	3-(3-Furanyl)phenyl	5.0
6	CH_3	H	4-Cl-Ph	6.2	64	Me	CF_3	3-(3-Thienyl)phenyl	5.3
7	CH_3	H	4-Br-Ph	<4.9	65[x]	Me	CN	2-F-phenyl	8.5
8	CH_3	H	3-OH-Ph	6.5	66	Me	CN	3-F-phenyl	9.1
9	CH_3	H	5-Pyrimidinyl	5.7	67	Me	CN	3-CF3-phenyl	7.8
10	CH_3	H	4-Isoquinolinyl	<4.9	68	Me	CN	3-Thienyl	9.4
11	H	CF_3	Ph	6.8	69	Me	COOEt	3-Cl-phenyl	7.7
12[x]	CH_3	CF_3	Ph	7.4	70	Me	COOEt	3-CF_3-phenyl	6.7
13[x]	CH_3	CN	Ph	8.9	71	Me	COOEt	3-CN-phenyl	8.5
14	CH_3	(4-Methyl-1-piperazinyl) carbonyl	Ph	7.0	72	Me	COOEt	4-CN-phenyl	<4.8
15	CH_3	1-Pyrrolidinyl carbonyl	Ph	7.1	73	Me	COOEt	4-CF_3-phenyl	<4.8
16[x]	CH_3	4-Morpholinyl carbonyl	Ph	7.8	74	Me	COOEt	4-Cl-phenyl	<4.8
17	CH_3	CH_3	Ph	7.3	75	Me	COOEt	4-F-phenyl	6.4
18	CH_3	COOEt	Ph	8.0	76	Me	COOEt	2-F-phenyl	6.5
19	CH_3	3-Methyl-1,2,4-oxadiazol-5-yl	Ph	8.2	77	Me	COOEt	3-Thienyl	8.1
20	CH_3	CONHMe	Ph	6.9	78[x]	Me	COOEt	3-Pyridyl	5.8
21[x]	CH_3	1-Piperidinyl methyl	Ph	6.8	79	Me	$CONH_2$	2-F-phenyl	6.3
22	CH_3	(4-Methyl-1-piperazinyl) methyl	Ph	6.3	80	Me	$CONH_2$	3-CF_3-phenyl	7.1
23	CH_3	4-Acetyl-1-piperazinyl) methyl	Ph	6.5	81[x]	Me	$CONH_2$	3-Thienyl	7.3
24	CH_3	(3-Oxo-1-piperazinyl) methyl	Ph	6.7	82	Me	CONHMe	2-F-phenyl	5.8
25	Me	Me	4-CF_3	<4.8	83	Me	CONHMe	3-F-phenyl	7.0
26	Me	Me	4-Cl	<5.4	84	Me	CONHMe	3-CF_3-phenyl	6.7
27	Me	Me	4-F	6.4	85	Me	CONHMe	3-Thienyl	7.3
28	Me	Me	4-CN	<4.8	86[x]	Me	1-Pyrrolidinyl carbonyl	2-F-phenyl	6.6
29	Me	Me	3-CF_3	6.8	87	Me	1-Pyrrolidinyl carbonyl	3-F-phenyl	7.7
30[x]	Me	Me	3-Pyridyl	5.7	88	Me	1-Pyrrolidinyl carbonyl	3-CF_3-phenyl	6.8
31	Me	Me	3-Thienyl	7.2	89[x]	Me	1-Pyrrolidinyl carbonyl	3-Thienyl	7.7
52	Me	CF_3	4-Pyridyl	7.1	90	Me	4-Morpholinyl carbonyl	2-F-phenyl	6.9
53	Me	CF_3	2-Pyridyl	7.3	91	Me	4-Morpholinyl carbonyl	3-F-phenyl	7.7
54	Me	CF_3	3-Pyridyl	6.3	92[x]	Me	4-Morpholinyl carbonyl	3-CF_3-phenyl	7.9

55[y]	Me	CF_3	1-Methyl-imidazol-5-yl	<4.9	93	Me	4-Morpholinyl carbonyl	3-Thienyl	8.0
56	Me	CF_3	3-Furyl	7.3	94	Me	(4-Methyl-1-piperazinyl) carbonyl	2-F-phenyl	5.9
57[x]	Me	CF_3	2-Thienyl	7.2	95[x]	Me	(4-Methyl-1-piperazinyl) carbonyl	3-F-phenyl	6.7
58[x,	Me	CF_3	3-Thienyl	8.0	96	Me	(4-Methyl-1-piperazinyl) carbonyl	3-CF_3-phenyl	6.6
59	Me	CF_3	3-(1*H*-Pyrazol-4-yl)phenyl	<4.9	97	Me	(4-Methyl-1-piperazinyl) carbonyl	3-Thienyl	6.8
60[x]	Me	CF_3	3-(4-Isoxazolyl)phenyl	6.1					

[x]Compounds belonged to the test set. [m]Outliers.

TABLE-2
STRUCTURE OF COMPOUNDS WITH SKELETON TYPES B-E IN THE DATA SET



No.	Skeleton	R	pIC_{50}
32	B	3-OMe	7.1
33[x]	B	3-CF_3	6.5
34	B	3-F	7.8
35	B	3-Cl	7.0
36	B	3-Br	7.6
37	B	3-CN	7.9
38	B	3-$NHSO_2CH_3$	<4.9
39	B	3-NHAc	6.1
40	B	3-COOH	<4.9
41	B	3-COO-t-Bu	<4.9
42	B	2-F	6.9
43	B	2-Cl	5.3
44	B	2-CF_3	<4.8
45	B	2,4-bis F	<4.8
46	B	4-F	7.0
47	B	4-Cl	5.7
48	B	4-CF_3	<4.8
49	B	4-CN	<4.8
50	B	4-$NMe_2(CH_3)_2$	<4.8
51	B	4-$NHSO_2CH_3$	<4.8
98	C	(1R,4S)-2-azabicyclo[2.2.1]hept-2-yl carbonyl	6.2
99	C	[(1S,4S)-5-methyl-2,5-diazabicyclo[2.2.1]hept-2-yl] carbonyl	5.4
100[x, m]	C	(2,2-Dimethyl-1-pyrrolidinyl) carbonyl	5.9
101	C	(3,3-Difluoro-1-pyrrolidinyl) carbonyl	6.9
102	C	(2,5-Dimethyl-1-pyrrolidinyl) carbonyl	6.4
103[x,m]	C	[(2R,6S)-2,6-dimethyl-4-morpholinyl] carbonyl	6.5
104	C	(2-Methyl-1-pyrrolidinyl) carbonyl	7.4
105	C	(2-Methyl-1-piperidinyl) carbonyl	7.0
1	D	–	8.0
2	E	–	7.5

[x]Compounds belonged to the test set. [m]Outliers.

TABLE-3
OBSERVED AND CoMFA/CoMSIA PREDICTED mGluR5 INHIBITORY ACTIVITY (pIC_{50} VALUE)

No.	Observed activity	CoMSIA		CoMFA	
		Predicted	Residual	Predicted	Residual
3[x,m]	8.6	7.301	1.299	6.985	1.615
4[x,m]	9.4	6.678	2.722	6.669	2.731
5	7.8	7.033	0.767	7.122	0.678
6	6.2	6.576	-0.376	5.882	0.318
8	6.5	6.936	-0.436	6.822	-0.322
9	5.7	5.668	0.032	5.978	-0.278
11	6.8	6.971	-0.171	6.871	-0.071
12[x]	7.4	6.925	0.475	7.218	0.182
13[x]	8.9	8.901	-0.001	8.824	0.076
14	7.0	6.647	0.353	6.824	0.176
15	7.1	7.245	-0.145	7.501	-0.401
16[x]	7.8	7.774	0.026	7.493	0.307
17	7.3	7.249	0.051	7.021	0.279
18	8.0	7.622	0.378	7.589	0.411
19	8.2	7.854	0.346	8.343	-0.143
20	6.9	6.697	0.203	7.188	-0.288
21[x]	6.8	7.033	-0.233	6.204	0.596
22	6.3	6.41	-0.11	6.306	-0.006
23	6.5	6.823	-0.323	6.455	0.045
24	6.7	6.884	-0.184	6.615	0.085
27	6.4	6.756	-0.356	6.284	0.116
29	6.8	6.765	0.035	6.672	0.128
30[x]	5.7	6.436	-0.736	6.159	-0.459
31	7.2	7.454	-0.254	7.53	-0.33
32	7.1	6.825	0.275	7.044	0.056
33[x]	6.5	6.58	-0.08	6.372	0.128
34	7.8	7.171	0.629	7.459	0.341
35	7.0	7.401	-0.401	7.464	-0.464
36	7.6	7.592	0.008	7.221	0.379
37	7.9	7.717	0.183	7.897	0.003
39	6.1	5.948	0.152	5.934	0.166
42	6.9	6.149	0.751	6.673	0.227
43	5.3	5.92	-0.62	6.046	-0.746
46[x]	7.0	6.683	0.317	7.194	-0.194
47	5.7	6.21	-0.51	5.921	-0.221
52	7.1	7.145	-0.045	7.216	-0.116
53	7.3	7.517	-0.217	7.404	-0.104
54	6.3	6.161	0.139	6.249	0.051
56	7.3	7.333	-0.033	7.194	0.106
57[x]	7.2	6.767	0.433	6.597	0.603
58[x]	8	7.198	0.802	7.311	0.689
60[x]	6.1	5.975	0.125	5.382	0.718
63	5.0	5.062	-0.062	5.187	-0.187
64	5.3	4.822	0.478	5.485	-0.185
65[x]	8.5	8.027	0.473	8.101	0.399
66	9.1	8.907	0.193	8.839	0.261
67	7.8	8.414	-0.614	8.215	-0.415
68	9.4	8.973	0.427	9.153	0.247
69	7.7	7.93	-0.23	7.581	0.119
70	6.7	7.105	-0.405	6.857	-0.157
71	8.5	8.678	-0.178	8.505	-0.005
75	6.4	6.981	-0.581	6.502	-0.102
76	6.5	6.529	-0.029	6.171	0.329
77	8.1	7.845	0.255	8.226	-0.126
78[x]	5.8	6.571	-0.771	5.692	0.108
79	6.3	6.109	0.191	6.434	-0.134
80	7.1	6.565	0.535	6.784	0.316
81[x]	7.3	7.113	0.187	7.481	-0.181
82	5.8	6.133	-0.333	5.987	-0.187
83	7.0	7.088	-0.088	7.021	-0.021
84	6.7	6.568	0.132	6.507	0.193
85	7.3	7.013	0.287	7.407	-0.107
86[x]	6.6	6.25	0.35	7.215	-0.615

87	7.7	7.355	0.345	7.868	-0.168
88	6.8	6.759	0.041	6.823	-0.023
89[x]	7.7	7.325	0.375	7.578	0.122
90	6.9	6.897	0.003	6.779	0.121
91	7.7	7.831	-0.131	7.714	-0.014
92[x]	7.9	7.512	0.388	7.113	0.787
93	8.0	8.029	-0.029	8.174	-0.174
94	5.9	5.688	0.212	5.798	0.102
95[x]	6.7	6.709	-0.009	6.7	0
96	6.6	6.416	0.184	6.597	0.003
97	6.8	6.823	-0.023	6.761	0.039
98	6.2	6.457	-0.257	6.211	-0.011
99	5.4	6.134	-0.734	5.11	0.29
100[x,m]	5.9	6.882	-0.982	7.309	-1.409
101	6.9	6.915	-0.015	7.067	-0.167
102	6.4	6.686	-0.286	6.433	-0.033
103[x,m]	6.5	7.667	-1.167	7.221	-0.721
104	7.4	6.915	0.485	7.341	0.059
105	7.0	6.893	0.107	6.939	0.061

[x]Compounds belonged to the test set. [m]Outliers.

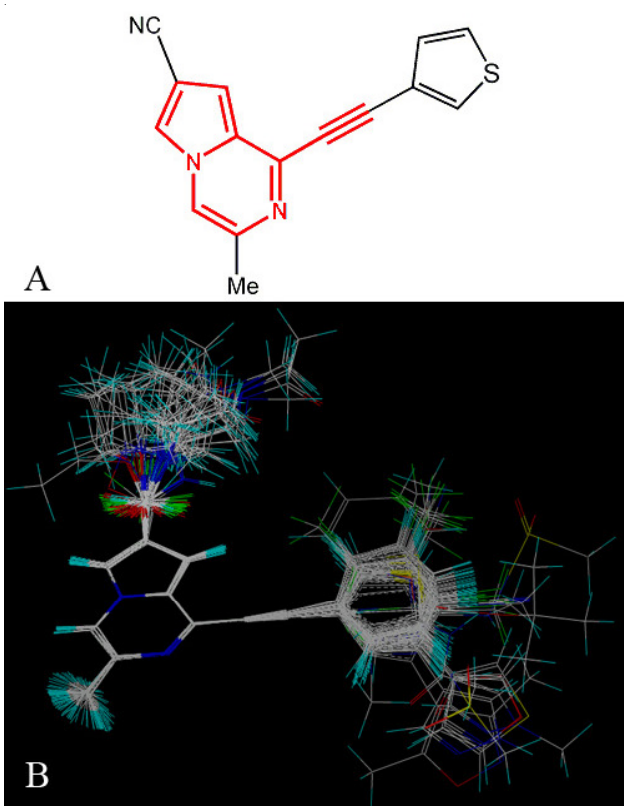


Fig. 2. Molecular alignment of compounds in the whole data set. (A) Common substructure of the molecules is shown in red based on template compound 68. (B) ligand-based alignment of all the compounds. Different colours represent different kind of atoms: white for C, blue for N, red for O, green for F, yellow for S and cyan for H, respectively.

principle. Fig. 2A shows the common substructure depicted in red colour and Fig. 2B shows the resulted ligand-based alignment model.

CoMFA and CoMSIA field calculation: CoMFA[22] and CoMSIA[23] were performed to build the models in order to reveal the relationship between 3D structural features and activity by employing the standard option of SYBYL. In CoMFA method, the superimposed molecules are kept in a 3D grid that steric (vdW interaction) and electrostatic (Coulombic values with a 1/r distance-dependent dielectric function) interactions were calculated using the Tripos force field with a distance-dependent dielectric constant at all interactions in a regularly spaced (2 Å) grid taking a sp^3 carbon atom as steric probe and a + 1 charge as electrostatic probe. The cut off value was 30 kcal/mol (default cutoff) for the Lennard-Jones and Coulomb-type potential and a constant dielectric function. The probe atom was placed at each lattice point and their steric and electrostatic interactions with each atom in the molecule were computed using the CoMFA standard scaling.

Compared with CoMFA method which only calculates the steric and electrostatic interactions, CoMSIA approach not only calculates the steric, electrostatic interactions, but also calculates the hydrophobic, HB donor and HB acceptor interactions. The basic assumption of CoMSIA is that a suitable sampling of the steric, electrostatic, hydrophobic and HB acceptor interactions generated around a set of aligned molecules with a probe atom might provide all the important features for understanding their biological activities and that the changes in binding affinities of the ligands are related to changes in the molecular properties[24]. CoMSIA similarity index descriptors were derived using the same lattice boxes as those used in CoMFA calculations. In CoMSIA, the steric indices are related to the third power of the atomic radii, the electrostatic descriptors are derived from atomic partial charges, the hydrophobic fields are derived from atom-based parameters developed by Viswanadhan and co-workers and the hydrogen bond donor and acceptor indices are obtained from a rule-based method derived from experimental values[25]. A Gaussian function was used to evaluate the mutual distance between the probe atom and each molecule atom. CoMSIA similarity indices (AF) for a molecule j with atom i at a grid point q are calculated by eqn. 1 as follows:

$$A^{q}_{F,k}(j) = -\sum \omega_{probe,k}\,\omega_{ik}\,e^{-\alpha r^{2}_{iq}} \qquad (1)$$

where $\omega_{probe,k}$ is the probe atom with radius 1 Å, charge +1, hydrophobicity +1, hydrogen bond donating +1 and hydrogen bond accepting +1. ω_{ik} is the actual value of the physicochemical property k of atom i. r_{iq} is the mutual distance between the probe atom at grid point q and item i of the test molecule.

α is the attenuation factor, with a default value of 0.3 and an optimal value normally ranging[26] from 0.2-0.4. In this work, five physicochemical properties including steric, electrostatic, hydrophobic, HB donor and acceptor were evaluated using the probe atom.

3D-QSAR models calculation, internal and external validation: The method of partial least squares (PLS) implemented in the QSAR module of SYBYL was used to construct and validate the models. The CoMFA/CoMSIA descriptors served as independent variables and pIC_{50} values as dependent variable in PLS regression analysis. The advantage of this method is that it can reduce the large number of original descriptors to a few principal components (PCs) that are linear combinations of the original descriptors[27]. The optimum number of PCs was determined by the leave-one-out (LOO) cross-validation procedure in which one compound is removed from the data set and its activity is predicted using the model derived from the rest of the data set[28]. Then with this optimal PCs number, a non-cross-validation analysis was carried out and the Pearson coefficient (R^2_{ncv}) and standard error of estimates (SEE) were calculated[27].

During PLS process, to evaluate the reliability of the model generated, several statistical parameters including the Q^2 and above R^2_{ncv} are crucial. As a cross-validated coefficient, Q^2 is used as a statistical index of the predictive power of the model and is calculated by eqn. 2 where the $Y_{predicted}$, $Y_{observed}$ and Y_{mean} are predicted, actual and mean values of the target property, respectively[29].

$$q^2 = 1 - \frac{\sum_Y (Y_{predicted} - Y_{observed})^2}{\sum_Y (Y_{observed} - Y_{mean})^2} \quad (2)$$

When assessing the predictive power of the QSAR model derived using the training set, an independent test set was used and their biological activities were predicted. The predictive R^2 (R^2_{pre}) value is calculated using eqn. 3.

$$R^2_{pre} = \left(\frac{SD - PRESS}{SD} \right) \quad (3)$$

where SD is the sum of squared deviations between the biological activity of the test set and the mean activity of training set molecules and PRESS is the sum of squared deviations between the actual and the predicted activities of the test set molecules[30]. At last, the CoMFA/CoMSIA results were graphically represented by field contour maps, where the coefficients were generated using the field type "Stdev*Coeff".

RESULTS AND DISCUSSION

CoMFA and CoMSIA statistical results: In this study, discarding molecules **1** and **2** who have no common substructures (Fig. 2A), the remaining 82 compounds were aligned according to the ligand-based alignment rules to derive the CoMFA and CoMSIA models. During the modeling, same training set (61 molecules), test set (21 molecules) were used for all the models established. To evaluate the quality of the models we built, several statistical parameters were analyzed, including the cross-validated Q^2, non-cross-validated correlation coefficient R^2_{ncv}, SEE, F-test values, predicted correlation coefficient (R^2_{pre}) and standard error of prediction SEP.

For CoMFA analysis, the steric and electrostatic field descriptors were fitted together in every possible form to build appropriate CoMFA mathematical models. Finally the models using of descriptors steric and electrostatic fields obtained proper reliability (Table-4) and got a result with Q^2 = 0.53, R^2_{ncv} = 0.92, SEE = 0.26, F = 77.49 using eight optimum components. When being tested by the independent test set, the model exhibits a satisfactory predictive ability with R^2_{pre} = 0.80 and SEP = 0.44 for this ligand-based model. In CoMFA model, electrostatic feature is found to make more contribution to the activity (52 %).

For CoMSIA models analysis, five fields including the steric, electrostatic, hydrophobic, HB donor and HB acceptor interactions were calculated using the same data sets as in the CoMFA analysis. All the five parameters were fitted together in every possible form to build appropriate CoMSIA models. Finally, only by using steric, electrostatic, hydrophobic and HB acceptor parameters superior ligand-based models were obtained with the highest Q^2 values (Table-4). The CoMSIA model has a Q^2 value of 0.51 with seven optimum components, an R^2_{ncv} value of 0.85, a SEE value of 0.36 and an F value of 43.43. Furthermore, CoMSIA model indicates that electrostatic feature plays major contribution to the correlation with the mGluR5 antagonist activity. When validated by the independent test set, the model exhibits a satisfactory predictive ability with R^2_{pre} = 0.80 and SEP = 0.42.

Normally, 3D-QSAR studies with a Q^2 greater than 0.4 are considered to be statistically significant[31]. In addition, higher R^2_{ncv} and F values as well as lower SEE values should also be considered as the foundation of a reliable 3D-QSAR model. For our optimal 3D-QSAR models, their Q^2 is larger than 0.5, proving their reliability. But only using the extensively accepted LOO cross-validated Q^2 is insufficient to assess the predictive power of the QSAR models[32]. Therefore, presently we further validated the predictive capacity of above four models by predicting the activity (pIC_{50} value) of the compounds of the external test set. For this purpose, the test set (21 molecules) which representing 33.9 % of the training set were used.

Before the final validation by the test set, an initial inspection of the fitted/predicted activities reveals poor prediction for several inhibitors which were considered as outliers in this work, they are compounds **3**, **4**, **100** and **103**. Several reasons like unmatched structure, different active conformation or more specific molecular mechanisms may result in the existence of outliers. A particular careful examination of the outliers may provide additional information determining their peculiarities; therefore, in this study all outliers are attentively checked and are finally divided as follows:

Compounds **100** and **103** belong to type C skeleton in structure which includes **8** molecules of **98-105**. Having unique and complex R substituents which substructures may greatly affect their binding conformations, the two molecules may have special and different active conformation from others and thus may not fit well in the current binding pocket assumed for most of other molecules with similar skeleton type and substituents.

Compounds **3** and **4**, they all have a higher residual between the experimental and predicted activity (pIC_{50} value residual is larger than 1.2) and thus are treated as outliers. This discrepancy, we speculate, on one hand indicates that these particular molecules may not be typical of the rest of the data and on the other hand suggests the necessity to recruit more plenteous and accurate experimental data with more diversified molecular structures.

After elimination of these outliers, both the CoMFA and CoMSIA models exhibit good prediction (Q^2 is larger than 0.5 and R^2 pre is larger than 0.85), indicating ligand-based alignment rule is good. The observed and CoMFA/CoMSIA predicted mGluR5 receptor inhibitory activities are shown in Table-3.

Fig. 3 depicts the actual *versus* predicted pIC_{50} values plot for both the training (filled black square) and test (filled blue diamond) set molecules of the whole data set based on CoMFA and CoMSIA models. As observed, all the points are rather uniformly distributed around the regression line in the two figures and the predicted activities are almost as accurate as the experimental data, indicating a proper correlation between the predicted and experimental activities of the data set and the reliability of the obtained models.

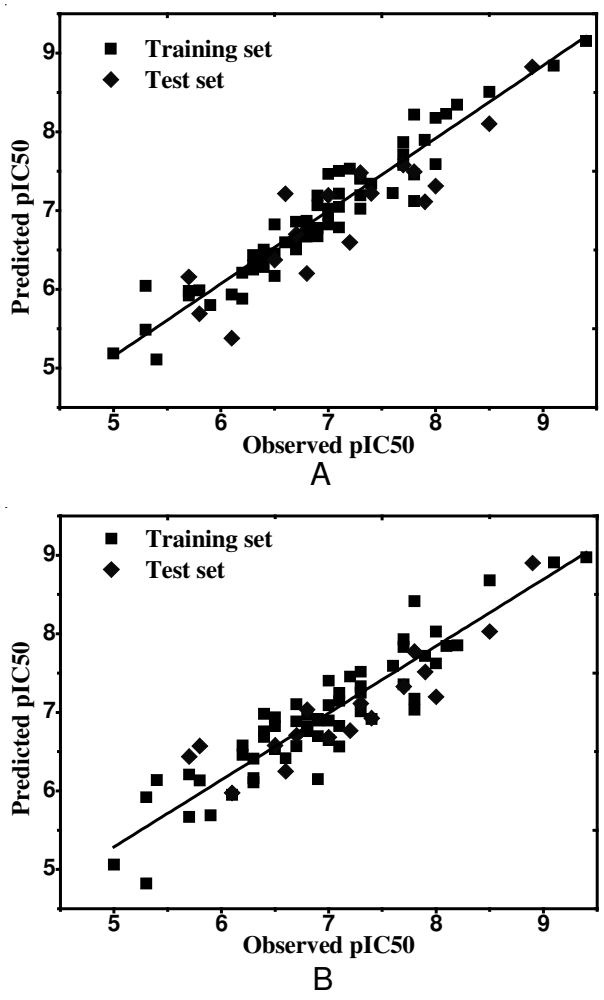


Fig. 3. Ligand-based correlation plots of the predicted *versus* the actual pIC_{50} values using the training set (filled black square) and the test set I (filled blue diamond) based on (A) CoMFA model and (B) CoMSIA model, respectively

3D-QSAR contour maps: In order to view the field effect on the target property, CoMFA and CoMSIA contour maps were generated. The contour maps can identify the important regions where any change in the steric, electrostatic, hydrophobic and HB fields may affect the biological activity and reveal the important features of the ligand-receptor interactions[29]. The visualization of the 3D-QSAR models was performed using the StDev*Coeff mapping option contoured by contribution. The default level of contour by contribution, 80 % for favoured region and 20 % for disfavoured region, was setted during contour analysis. All contour maps obtained from the models are illustrated together with template compound **68** (Fig. 4) which is one of the most active molecules in the whole data set (with pIC_{50} value of 9.4).

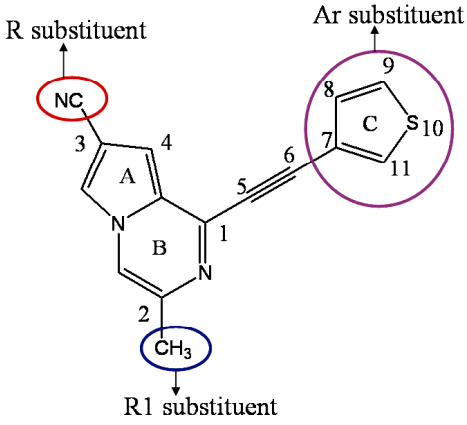


Fig. 4. Structure of compound **68**. Other molecules in the dataset are similar to **68** in structure except the R, R_1 and Ar substituents

CoMFA contour map analysis: The contributions of the steric and electrostatic fields for the CoMFA results were graphically displayed in contour maps in Fig. 5. The blue-red region to the green-yellow region indicates a greater contribution of the electrostatic field towards the inhibitory activity than the steric field. This different contribution is also reflected in Table-4 where the relative contribution for the steric field is 48 % whereas the contribution of the electrostatic field is 52 %. Fig. 5A shows the steric contour map of the optimal CoMFA model, the steric fields defined by the green coloured contours represent regions of favourable steric effect, while yellow coloured contours represent regions of unfavourable steric effect. Three small positive steric (green) regions are found appearing around the ring-C (Ar substituent), explaining why molecules who carry a bulky substituent in ring-C are more active than those compounds without substituent in these particular positions. A good illustration is that compounds **27** and **29** whose R substituent is -F and $-CF_3$, respectively are less active than compounds **3** and **4** with -Ph and -4-F-Ph as Ar substituent, respectively. A green contour is also observed staying above the positon-3 (R substituent) in ring-A, suggesting its favour for sterically bulky groups, which can be examplified by the smaller pIC_{50} values of compounds **82**, **83**, **84** and **85** than chemicals **86**, **87**, **88** and **89**. In contrast, a big negative steric (yellow) region presents itself below the position-3, indicating that a substituent of bulky steric below this position disbenifits the biological activity of the molecules.

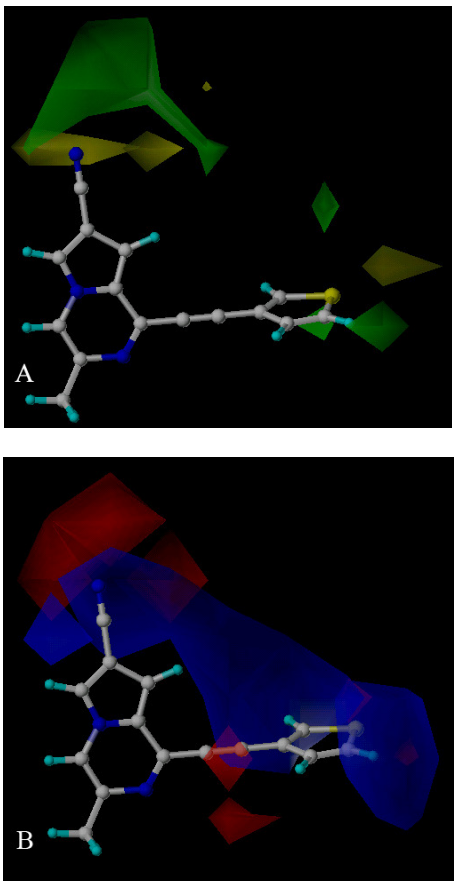


Fig. 5. CoMFA StDev*Coeff contour plots. (A) steric (green/yellow) contour map combined with compound **68**. Green contours indicate regions where bulky groups increase the activity; yellow contours indicate regions where bulky groups decrease the activity. (B) Electrostatic contour map (blue/red) in combination with compound **68**. Blue contours indicate regions where positive charges increase the activity; red contours indicate regions where negative charges increase the activity

TABLE-4
SUMMARY OF CoMFA AND CoMSIA RESULTS*

PLS statistics	Ligand-based model	
	CoMFA	CoMSIA
Q^2	0.53	0.51
R^2_{ncv}	0.92	0.85
SEE	0.26	0.36
F	77.49	43.43
R^2_{pre}	0.80	0.80
SEP	0.44	0.42
OPN	8	7
Contribution		
Steric	0.48	0.14
Electrostatic	0.52	0.37
Hydrophobic	–	0.23
HB acceptor	–	0.26

*Q^2, cross-validated correlation coefficient after the leave-one-out procedure; R^2_{ncv}, non-cross-validated correlation coefficient; SEE, standard error of estimate; F, ratio of R^2_{ncv} explained to unexplained = $R^2_{ncv}/(1-R^2_{ncv})$; R^2_{pre}, predicted correlation coefficient for test set I compounds; SEP, standard error of prediction; OPN, optimal number of principal components.

This conclusion is well proved by the fact that compounds **23**, **24** and **99** with large R substituent in position-3 possess only small pIC_{50} values of 6.5, 6.7 and 5.4, respectively.

Electrostatic contour map based on the PLS analysis of the CoMFA models are shown in Fig. 5B where increasing negative charge is favoured in red regions and increasing positive charge is favoured in blue regions. One big red polyhedron adjacent to the position-3 in ring-A represents the areas where the negative charged group is favoured, exampled by the fact that R substituent in ring-A (position-3) which is -CN bearing negative charges on the N atom increases the activity, like molecules **65** (pIC_{50} = 8.5), **66** (pIC_{50} = 9.1) and **68** (pIC_{50} = 9.4) whose pIC_{50} values are higher than any other molecules in the whole data set. Another one small red polyhedron appearing near the position-9 of ring-C indicates the regions where a negative charge substituent would be favourable and possibly improve the activity. One big blue polyhedron lying above the plane of the ring-A (position-3) indicates substituents with positive charges may improve their activities. There is a big blue contour surrounding the phenyl substituent which shows its favour for groups with positive charges. This conclusion is well illustrated by the poor activity of molecules **5** (pIC_{50} = 7.8), **6** (pIC_{50} = 6.2) and **8** (pIC_{50} = 6.5) with respective fluorine, chlorine and hydroxyl in the ring-C compared with compound **3** (pIC_{50} = 8.6) with no negative atom in ring-C.

CoMSIA contour map analysis: In this study, the optimal CoMSIA model not only calculates the steric and electrostatic fields, but also uses the hydrophobic and HB acceptor fields to correlate with the antagonist activity. All contour maps of the four CoMSIA fields are shown in Fig. 6. The colour scheme used in the CoMSIA steric and electrostatic field contour maps (Fig. 6A and B) is the same as described in the CoMFA contour maps. The steric contour map of CoMSIA model (Fig. 5A) shows similar results as that of the CoMFA one. with only difference in that all the positive steric (green) regions locating around the ring-C (Ar substituent) and positon-3 (R substituent) in ring-A in the CoMSIA model are much larger in size than in the CoMFA one. These results lead to the conclusion that compound with bulky substituent in positon-3 (ring-A) and ring-C may result in an increasing of the activity.

As shown in Table-4, electrostatic field makes the largest contribution to the CoMSIA QSAR models, which suggests that among all descriptors considered, the electrostatic property of the inhibitors may play more important role affecting the binding affinities. Fig. 6B shows the electrostatic contour map of the optimal CoMSIA model, where the blue polyhedron is clearly smaller than in the CoMFA one. One small blue contour is also observed placing near the ring-C (position-8) indicating that the occupancy by electropositive substituent in this region would promote the binding affinities to mGluR5 receptor and the other blue region adjacent to the ring-C (position-10 and 11) also points out the favour of electropositive substituent in this region.

The CoMSIA hydrophobic contour map of affinity for mGluR5 is depicted in Fig. 6C, where yellow and white contours highlight areas that hydrophobic and hydrophilic properties are preferred, respectively. One yellow polyhedron appearing above the ring-A (position-3) indicates that hydrophobic groups (like -OMe, -OEt, -F, -Cl, -Br) are beneficial to enhance the activity, like compounds **76** (pIC_{50} = 6.5) and **77**

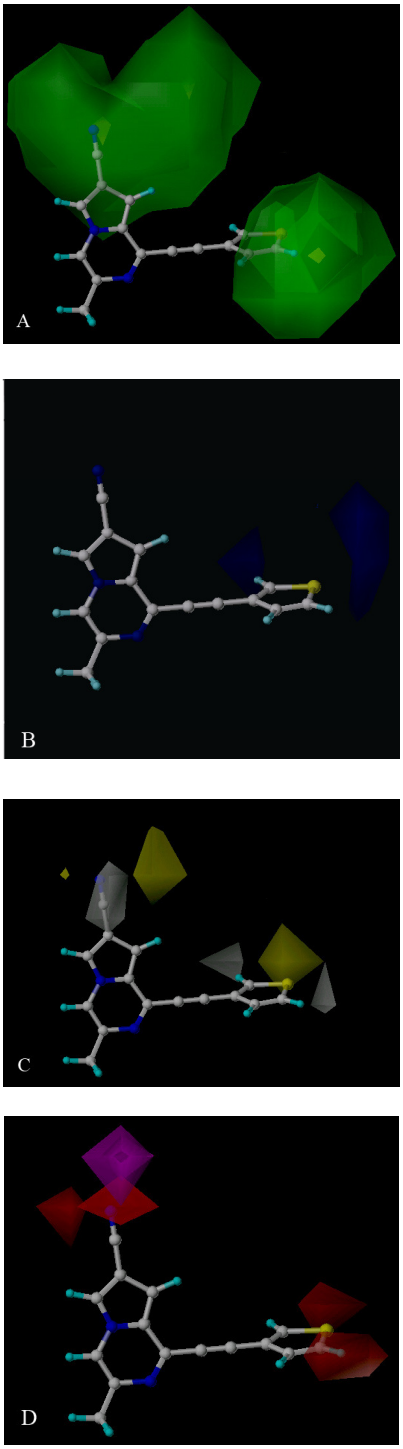


Fig. 6. CoMSIA StDev*Coeff contour plots. (A) steric (green/yellow) contour map in combination with compound **68**. Green contours indicate regions where bulky groups increase the activity; yellow contours indicate regions where bulky groups decrease the activity. (B) Electrostatic contour map (blue/red) in combination with compound **68**. Blue contours indicate regions where positive charges increase the activity; red contours indicate regions where negative charges increase the activity. (C) Hydrophobic contour map (yellow/white) in combination with compound **68**. Yellow contours indicate regions where hydrophobic substituents enhance the activity; white contours indicate regions where hydrophilic substituents enhance the activity. (D) HB acceptor contour map (magenta/red) in combination with compound **68**. Magenta contours indicate regions where HB acceptor on the molecule promotes the affinity; red contours indicate regions where HB acceptor on the molecule demotes the affinity

(pIC_{50} = 8.1) with -COOEt in substituents in position-3 exhibiting higher activities than **79** (pIC_{50} = 6.3) and **81** (pIC_{50} = 7.3) which have -$CONH_2$ substituents in the same position. The other yellow isopleth near the ring-C (position-10) also indicates that hydrophobic groups are beneficial to increase the activity. There is one white contour around the R substituent in position-3 of ring-A revealing the necessity of hydrophilic groups (like hydroxy or cyano) on this position to enhance the activity, like the compounds **65** (pIC_{50} = 8.5), **66** (pIC_{50} = 9.1) and **68** (pIC_{50} = 9.4) with -CN group in this position. Another two white polyhedrons placing near the ring-C (position-9 and 11) also demonstrate the region's favour for hydrophilic groups.

Fig. 6D displays the CoMSIA hydrogen-bond acceptor field. The magenta and red contours indicate regions where HB acceptor group increases and decreases the activity, respectively. A big magenta isopleth adjacent to the position-3 in ring-A supporting the requirement of HB acceptor to improve the activity. This offers an explanation for the higher biological activity of compounds **13** (pIC_{50} = 8.9) and **68** (pIC_{50} = 9.4) which all have the -CN substituent in the position-3 of ring-A than any other molecules in the whole data set. Two small red contours are also found adjacent to the ring-C (position-9 and 10), revealing that HB acceptor group impairs the activity just like the compound group of **65-97**. Molecules with the same R substituent if their Ar substituent without F atoms always show higher activities than with -F atoms, due to the HB acceptor role of the fluorin in ring-C (position-9 and 10).

The detailed contour analysis of ligand-based CoMFA and CoMSIA models enabled us to identify several structural requirements as mentioned in above paragraph for the observed inhibitory activities: (1) Electropositive groups in Ar substituent are beneficial to enhance the activity; (2) R substituent with HB acceptor also leads to high activity; (3) Bulky R and Ar substituents are not favoured in mGluR5; (4) R substituent with hydrophilic group can improve the biological activity.

Conclusion

In this work, depending on the ligand-based alignments a 3D-QSAR study using CoMFA and CoMSIA methods was carried out for the first time to new fused phenylethynyl-pyrrolo[1,2-a]pyrazine derivatives as mGluR5 antagonists. Based on the reasonable Q^2, R^2_{ncv}, R^2_{pre} values of test set, we conclude our ligand-based models are reasonable with proper predictivity (for CoMFA model: Q^2 = 0.53, R^2_{ncv} = 0.92, R^2_{pre} = 0.80 and for CoMSIA model: Q^2 = 0.51, R^2_{ncv} = 0.85, R^2_{pre} = 0.80). In addition, the results from the contour maps have offered insight for the design of mGluR5 antagonists by identifying significant regions for steric, electrostatic, hydrophobic and HB acceptor interactions: (1) electropositive groups in Ar substituent are beneficial to enhance the activity; (2) R substituent with HB acceptor also leads to high activity; (3) Bulky R and Ar substituents are not favoured in mGluR5; (4) R substituent with hydrophilic group can improve the biological activity. All these results should provide information for better understanding of the mechanism of antagonism and thus be helpful to further validate the potential therapeutic applications of the mGluR5 antagonists in future.

ACKNOWLEDGEMENTS

This work is financially supported by the National Natural Science Foundation of China (Grant No. 10801025) and Fund of Northwest A&F University.

REFERENCES

1. J.N. Kew and J.A. Kemp, *Psychopharmacology (Berl.)*, **179**, 4 (2005).
2. N.D.P. Cosford, J. Roppe, L. Tehrani, E.J. Schweiger, T.J. Seiders, A. Chaudary, S. Rao and M.A. Varney, *Bioorg. Med. Chem. Lett*., **13**, 351 (2003).
3. A. Valerio, P. Rizzonelli, M. Paterlini, G. Moretto, T. Knopfel, R. Kuhn, M. Memo and P. Spano, *Neurosci. Res*., **28**, 49 (1997).
4. P.C. Chua, J.Y. Nagasawa, L.S. Bleicher, B. Munoz, E.J. Schweiger, L. Tehrani, J.J. Anderson, M. Cramer, J. Chung, M.D. Green, C.D. King, G. Reyes-Manalo and N.D. Cosford, *Bioorg. Med. Chem. Lett*., **15**, 4589 (2005).
5. A. Krivoy, T. Fischel and A. Weizman, *Eur. Neuropsychopharm*., **18**, 395 (2008).
6. Z. Jia, Y.M. Lu, N. Agopyan and J. Roder, *Physiol. Behav*., **73**, 793 (2001).
7. F.I. Carroll, *Ann. N Y Acad. Sci*., **1141**, 221 (2008).
8. F. Gasparini, K. Lingenhohl, N. Stoehr, P.J. Flor, M. Heinrich, I. Vranesic, M. Biollaz, H. Allgeier, R. Heckendorn, S. Urwyler, M.A. Varney, E.C. Johnson, S.D. Hess, S.P. Rao, A.I. Sacaan, E.M. Santori, G. Velicelebi and R. Kuhn, *Neuropharmacology*, **38**, 1493 (1999).
9. P. Bach, K. Nilsson, T. Svensson, U. Bauer, L.G. Hammerland, A. Peterson, A. Wållberg, K. Österlund, D. Karis, M. Boije and D. Wensbo, *Bioorg. Med. Chem. Lett*., **16**, 4788 (2006).
10. D. Alagille, R.M. Baldwin, B.L. Roth, J.T. Wroblewski, E. Grajkowska and G.D. Tamagnan, *Bioorg. Med. Chem*., **13**, 197 (2005).
11. S. Sharma, A.L. Rodriguez, P.J. Conn and C.W. Lindsley, *Bioorg. Med. Chem. Lett*., **18**, 4098 (2008).
12. A. Wållberg, K. Nilsson, K. Österlund, A. Peterson, S. Elg, P. Raboisson, U. Bauer, L.G. Hammerland and J.P. Mattsson, *Bioorg. Med. Chem. Lett*., **16**, 1142 (2006).
13. H. Bräuner-Osborne, J. Egebjerg, E.O. Nielsen, U. Madsen and P. Krogsgaard-Larsen, *J. Med. Chem*., **43**, 2609 (2000).
14. G.F. Yang and X. Huang, *Curr. Pharm. Des*., **12**, 4601 (2006).
15. C. Ai, Y. Wang, Y. Li and L. Yang, *QSAR Comb. Sci*., **27**, 1183 (2008).
16. E. Papa, F. Battaini and P. Gramatica, *Chemosphere*, **58**, 559 (2005)
17. R. Mueller, A.L. Rodriguez, E.S. Dawson, M. Butkiewicz, T.T. Nguyen, S. Oleszkiewicz, A. Bleckmann, C.D. Weaver, C.W. Lindsley, P.J. Conn and J. Meiler, *ACS Chem. Neurosci*., **1**, 288 (2010).
18. F. Micheli, B. Bertani, A. Bozzoli, L. Crippa, P. Cavanni, R. Di Fabio, D. Donati, P. Marzorati, G. Merlo, A. Paio, L. Perugini and P. Zarantonello, *Bioorg. Med. Chem. Lett*., **18**, 1804 (2008).
19. J. Gasteiger and M. Marsili, *Tetrahedron*, **36**, 3219 (1980).
20. G.B. De Freitas, L.L. da Silva, N.C. Romeiro and C.A. Fraga, *Eur. J. Med. Chem*., **44**, 2482 (2009).
21. N. Dessalew, D.S. Patel and P.V. Bharatam, *J. Mol. Graph. Model*., **25**, 885 (2007).
22. P.L. Jackson, K.R. Scott, W.M. Southerland and Y.Y. Fang, *Bioorg. Med. Chem*., **17**, 133 (2009).
23. V. Murugesan, Y.S. Prabhakar and S.B. Katti, *J. Mol. Graph. Model*., **27**, 735 (2009).
24. Y. Li, Y.H. Wang, L. Yang, S.W. Zhang, C.H. Liu and S.L. Yang, *J. Mol. Struct*., **733**, 111 (2005).
25. E.F. da Cunha, W. Sippl, T. de Castro Ramalho, O.A. Ceva Antunes, R.B. de Alencastro and M.G. Albuquerque, *Eur. J. Med. Chem*., **44**, 4344 (2009).
26. V. Srivastava, S.P. Gupta, M.I. Siddiqi and B.N. Mishra, *Eur. J. Med. Chem*., **45**, 1560 (2010).
27. V.V. Kovalishyn, V. Kholodovych, I.V. Tetko and W.J. Welsh, *J. Mol. Graph. Model*., **26**, 591 (2007).
28 P. Yi, X. Fang and M. Qiu, *Eur. J. Med. Chem*., **43**, 925 (2008).
29. P. Lu, X. Wei and R. Zhang, *Eur. J. Med. Chem*., **45**, 1792 (2010).
30. R. Thaimattam, P. Daga, S.A. Rajjak, R. Banerjee and J. Iqbal, *Bioorg. Med. Chem*., **12**, 6415 (2004).
31. A. Dixit, S.K. Kashaw, S. Gaur and A.K. Saxena, *Bioorg. Med. Chem*., **12**, 3591 (2004).
32. A. Golbraikh and A. Tropsha, *J. Mol. Graph. Model*., **20**, 269 (2002).