

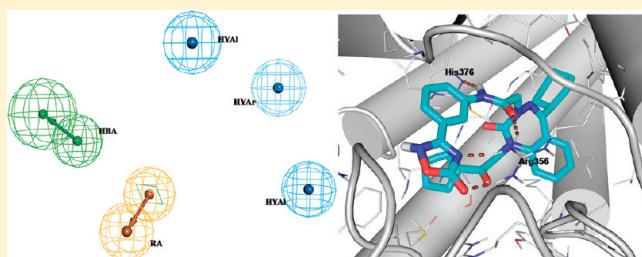
Toward the Identification of a Reliable 3D QSAR Pharmacophore Model for the CCK2 Receptor Antagonism

Amit K. Gupta,^{†,‡} Kanika Varshney,^{†,‡} and Anil K. Saxena^{*,†}

[†]Medicinal and Process Chemistry Division, C.S.I.R.-Central Drug Research Institute, Lucknow 226001, India

Supporting Information

ABSTRACT: The present study describes application of computational approaches to identify a validated and reliable 3D QSAR pharmacophore model for the CCK-2R antagonism through integrated ligand and structure based studies using anthranilic sulfonamide and 1,3,4-benzotriazepine based CCK-2R antagonists. The best hypothesis consisted five features viz. two aliphatic hydrophobic, one aromatic hydrophobic, one H-bond acceptor, and one ring aromatic feature with an excellent correlation for 34 training set ($r^2_{\text{training}} = 0.83$) and 58 test set compounds ($r^2_{\text{test}} = 0.74$). This model was validated through F-test and docking studies at the active site of the plausible CCK-2R where the 99% significance and well corroboration with the pharmacophore model respectively describes the model's reliability. The model also predicts well to other known clinically effective CCK-2R antagonists. Therefore, the developed model may useful in finding new scaffolds that may aid in design and develop new chemical entities (NCEs) as potent CCK-2R antagonists before their synthesis.



INTRODUCTION

In recent decade there has been a revived interest in the cholecystokinin receptor (CCK-R) antagonists as potent therapeutics for gastrointestinal disorders such as acid reflux, gastroesophageal reflux disease, peptic ulcers,¹ GI adenocarcinoma,¹ and tumors.² The CCK-Rs are seven-transmembrane helix spanning receptors that belongs to the superfamily of G-protein-coupled receptors (GPCRs). The CCK-Rs have been classified into two distinct classes, sharing 50% homology³ and designated as CCK-1R and CCK-2R on the basis of their affinity for the peptide agonists CCK and gastrin. Although both CCK and gastrin share an identical carboxyl-terminal pentapeptide sequence they differ in selectivity for the CCK-1R and the CCK-2R.^{4–6} The CCK-1Rs are specific for CCK, whereas CCK-2Rs recognize both CCK and gastrin with high affinity. The gastrointestinal polypeptide hormone gastrin has been known to stimulate gastric acid secretion and gastrointestinal cell growth in peripheral tissues. Hence, efforts have been made to develop agents that inhibit gastrin activity especially on gastric secretion, by acting as CCK-2R antagonists.⁷

The quest for drugs that block the actions of the peptide hormone gastrin mediated by CCK-2R has led to the discovery of a wide range of small-molecule ligands. These include glutamic acid analogues: spiroglumide, itriglumide; 1,4 benzodiazepines analogues: L-365260, YF476; dipeptoids: CI-988; benzobicyclo octane analogues: JB95008, JB93182; 1,5 benzodiazepines analogues: GV191869X; 1-benzazepine-2-one analogues: CP310713; ureidoacetamide analogues: RP73870; quinazoline based analogues: LY-202769; pyrazolidinone and related heterocyclic analogues: LY288513; indol-2-one analogue: AG-041R, and tetronothiodin.⁸ However, the development of many

of these compounds has been limited because of their poor bioavailability led to their low oral potency.

Recently anthranilic sulfonamide and 1,3,4-benzotriazepine based CCK-2R antagonists have been discovered which showed better CCK-2R selectivity with promising pharmacokinetic and *in vivo* activity in inhibiting pentagastrin-stimulated gastric acid secretion in the rat.^{9,10} However, the *in vitro* activity of anthranilic sulfonamide class of CCK-2R antagonists (5 nM) are comparatively far lower than 1,3,4-benzotriazepine CCK-2R antagonists (0.11 nM).

The major bottleneck for structure based drug design of CCK-2R antagonists is the nonavailability of its three-dimensional (3D) structure. Therefore, the ligand-based pharmacophore approach, based on extensive experimental structure–activity relationships (SARs), may be useful in designing of novel compounds with improved biological activity in such cases.¹¹ Earlier a 3D QSAR CoMFA, CoMSIA studies has been reported on the 1,3,4-benzotriazepine derivatives as CCK-2R antagonists,¹² and a pharmacophore model has been reported using diverse CCK-2R antagonists.¹³ However, these studies suffered from a major limitation, as the anthranilic sulfonamide derivatives were not considered for model generation in the former study¹² whereas in the later study anthranilic sulfonamides as well as 1,3,4-benzotriazepine derivates were not used in the model generation.¹³ Hence, both of the published 3D QSAR models may not predict the antagonistic activity of anthranilic sulfonamide as well as 1,3,4-benzotriazepine class of compounds. Considering the importance of these compounds

Received: February 17, 2012

Published: April 24, 2012



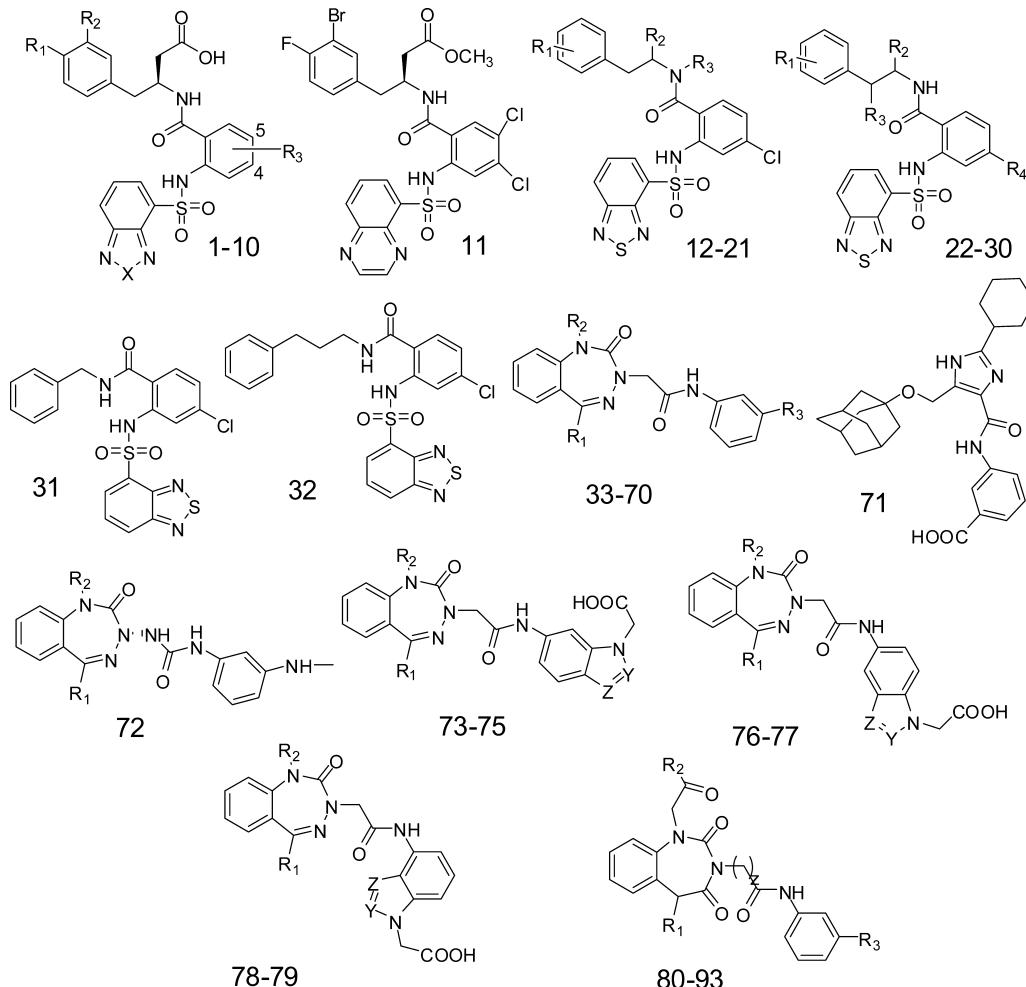


Figure 1. Common structure of the CCK-2R antagonists used in the 3D-QSAR study.

as potent CCK-2R antagonists effective against gastrin induced ulcers and unavailability of the 3D structure of CCK-2R, it appeared of interest to develop a pharmacophore model for determining the essential structural requirements for these classes of molecules. The development and validation of first 3D QSAR pharmacophore model utilizing the anthranilic sulfonamide derivatives and 1,3,4-benzotriazepine class of compounds as CCK-2R antagonists is reported herein.

RESULTS AND DISCUSSION

Biological Data. The chemically diverse anthranilic sulfonamide^{9,14} and 1,3,4-benzotriazepine^{10,15} derivatives as CCK-2R antagonists with K_i values spanning across a range from 0.11 nM to 7943 nM were taken for 3D QSAR pharmacophore model development (Figure 1, Table 1). As the homogeneity of the biological assays is one of the important aspects in QSAR study, therefore the data set was collected from the research groups following the same biological testing protocol.

Pharmacophore Model for Anthranilic Sulfonamide Class of CCK-2R Antagonist. Initially a pharmacophore was developed using only anthranilic sulfonamide derivatives as this class of molecules has shown good CCK-2R selectivity with promising pharmacokinetic activity. Further, there was no report of earlier QSAR studies on this class of CCK-2R antagonist. The data set (compounds 1–32 of Figure 1, Table 1) was rationally divided into a training set of 19 compounds and a

complementary test set of 13 compounds according to the published guidelines.¹⁶ All compounds were built using ISIS Draw 2.5, imported to Accelry's Discovery Studio2.0 (DS 2.0) window,¹⁷ and were optimized using CHARMm force field. A maximum of 255 conformations were generated within an energy threshold of 20.0 kcal/mol above the global energy minimum for all ligands within the Catalyst ConForm module using the "Poling" algorithm.¹⁸ The training set molecules associated with these conformations were submitted to the HypoGen algorithm of DS 2.0. Among the eleven features available for selection, the features viz. H-bond acceptor (HBA), ring aromatic (RA), hydrophobic (HYAl and HYAr), and negative ionizable (NI) feature were selected for pharmacophore generation based on their common 3D fingerprints in all the active molecules of the training set. The rest of the parameters were kept to the default setting as function weight 0.302, and activity uncertainty 3, and interfeature distance 2.97 Å. The HypoGen algorithm of DS 2.0 resulted in the generation of 10 alternative pharmacophores describing the CCK-2R antagonist activity of the training set compounds. The details of the pharmacophore generation procedure have been described in the Experimental Section of this manuscript. The quality of the generated pharmacophore hypotheses was evaluated by considering the cost functions calculated by the HypoGen module during hypothesis generation. The composition (in terms of chemical features that are necessary for activity), ranking score, and statistical parameters associated with the hypotheses are reported in Table 2.

Table 1. Structures of the Training and Test Set Compounds along with Their Biological Activity^a

C.N.	R ₁	R ₂	R ₃	R ₄	X	Z	Y	activity (pKi)
1	Cl	Cl	4-I	-	S	-	-	7.6
2	Cl	Br	4-Br	-	S	-	-	7.9
3	Cl	Br	4-Br	-	CH=CH	-	-	8.0
4	F	Br	4-Cl	-	S	-	-	8.0
5	F	Br	4-Cl	-	CH=CH	-	-	8.0
6	F	Br	4-Br	-	S	-	-	8.2
7	F	Br	4-Br	-	CH=CH	-	-	8.3
8	F	Br	4-I	-	S	-	-	8.3
9	F	Br	4,5-Cl ₂	-	CH=CH	-	-	8.2
10	F	Br	4,5-Cl ₂	-	S	-	-	8.0
11	-	-	-	-	-	-	-	6.6
12	H	H	H	-	-	-	-	5.2
13	3',4'-Cl ₂	H	H	-	-	-	-	5.1
14	H	(S)-COOH	H	-	-	-	-	7.1
15	H	(R)-COOH	H	-	-	-	-	6.8
16	4'-Cl	(S)-COOH	H	-	-	-	-	7.2
17	4'-Cl	(R)-COOH	H	-	-	-	-	6.2
18	3',4'-Cl ₂	(S)-COOH	H	-	-	-	-	7.8
19	3',4'-Cl ₂	(R)-COOH	H	-	-	-	-	6.3
20	3'-Br,4'-Cl	(S)-COOH	H	-	-	-	-	7.4
21	-	-	CH ₃	-	-	-	-	6.5
22	4-Cl	H	Me ₂	Cl	-	-	-	5.1
23	H	H	H	Br	-	-	-	6.2
24	H	H	(S)-Me	Br	-	-	-	6.3
25	H	H	(R)-Me	Br	-	-	-	5.8
26	3,4-Cl ₂	(S)-COOH	(R)-Me	Cl	-	-	-	6.8
27	3,4-Cl ₂	(R)-COOH	(R)-Me	Cl	-	-	-	5.8
28	4-Cl	(S)-COOH	Me ₂	Cl	-	-	-	6.0
29	4-Cl	(S)-COOH	4-Cl-Ph	Cl	-	-	-	5.9
30	3-Br	(R)-Me	(S)-OH	Cl	-	-	-	5.9
31	-	-	-	-	-	-	-	5.6
32	-	-	-	-	-	-	-	6.0
33	c-C ₆ H ₁₁	t-BuCOCH ₂	COOH	-	-	-	-	8.29
34	Me ₂ CHCH ₂	t-BuCOCH ₂	COOH	-	-	-	-	7.05
35	c-C ₅ H ₉	t-BuCOCH ₂	COOH	-	-	-	-	7.62
36	c-C ₇ H ₁₃	t-BuCOCH ₂	COOH	-	-	-	-	8.32
37	1-Ad	t-BuCOCH ₂	COOH	-	-	-	-	7.62
38	c-C ₆ H ₁₁	CH ₃	COOH	-	-	-	-	5.92
39	c-C ₆ H ₁₁	MeCOCH ₂	COOH	-	-	-	-	6.48
40	c-C ₆ H ₁₁	Me ₂ CH(CH ₂) ₂	COOH	-	-	-	-	7.69
41	c-C ₆ H ₁₁	Pyrrolidin-1-yl-COCH ₂	COOH	-	-	-	-	7.79
42	c-C ₆ H ₁₁	EtO(CH ₂) ₂	COOH	-	-	-	-	6.31
43	c-C ₆ H ₁₁	1-AdCOCH ₂	COOH	-	-	-	-	8.25
44	c-C ₆ H ₁₁	2-MeC ₆ H ₄ COCH ₂	COOH	-	-	-	-	8.12
45	c-C ₆ H ₁₁	c-C ₅ H ₉ COCH ₂	COOH	-	-	-	-	9.00
46	c-C ₆ H ₁₁	c-C ₆ H ₁₁ COCH ₂	COOH	-	-	-	-	9.10
47	c-C ₆ H ₁₁	1-Me-c-C ₅ H ₉ COCH ₂	COOH	-	-	-	-	8.92
48	c-C ₆ H ₁₁	t-BuCOCH ₂	CH ₃ COOH	-	-	-	-	8.67
49	c-C ₆ H ₁₁	t-BuCOCH ₂	(CH ₂) ₂ COOH	-	-	-	-	8.73
50	c-C ₆ H ₁₁	t-BuCOCH ₂	SCH ₂ COOH	-	-	-	-	9.19
51	c-C ₆ H ₁₁	t-BuCOCH ₂	N(Me)CH ₂ COOH	-	-	-	-	8.97
52	c-C ₆ H ₁₁	t-BuCOCH ₂	SO ₂ CH ₂ COOH	-	-	-	-	8.37
53	c-C ₆ H ₁₁	t-BuCOCH ₂	OCH ₂ ₂ COOH	-	-	-	-	8.26
54	c-C ₆ H ₁₁	t-BuCOCH ₂	NHCH ₂ COOH	-	-	-	-	8.21
55	c-C ₆ H ₁₁	t-BuCOCH ₂	CONHSO ₂ Me	-	-	-	-	8.41
56	c-C ₆ H ₁₁	t-BuCOCH ₂	1(2) H-tetrazol-5-yl	-	-	-	-	9.00
57	c-C ₆ H ₁₁	t-BuCOCH ₂	1,2,4-oxadiazol-3-yl-5 (2H)-one	-	-	-	-	9.70
58	c-C ₆ H ₁₁	t-BuCOCH ₂	N(Me)-1(2)H-tetrazol-5-yl	-	-	-	-	8.71
59	c-C ₆ H ₁₁	t-BuCOCH ₂	CH ₂ -1(2) H-tetrazol-5-yl	-	-	-	-	8.33
60	c-C ₆ H ₁₁	t-BuCOCH ₂	SCH ₂ -1(2) H-tetrazol-5-yl	-	-	-	-	9.15
61	c-C ₆ H ₁₁	t-BuCOCH ₂	3-C ₆ H ₄ COOH	-	-	-	-	8.80

Table 1. continued

C.N.	R ₁	R ₂	R ₃	R ₄	X	Z	Y	activity (pKi)
62	c-C ₆ H ₁₁	t-BuCOCH ₂	2-thiazol-4-yl-COOH	-	-	-	-	9.18
63	c-C ₆ H ₁₁	t-BuCOCH ₂	4-oxazol-2-yl-COOH	-	-	-	-	9.44
64	c-C ₆ H ₁₁	t-BuCOCH ₂	5-furan-2-yl-COOH	-	-	-	-	9.61
65	c-C ₆ H ₁₁	t-BuCOCH ₂	1-imidazol-4-yl-(1 <i>H</i>)-CH ₂ COOH	-	-	-	-	8.39
66	c-C ₆ H ₁₁	t-BuCOCH ₂	1-pyrrol-2-yl-CH ₂ COOH	-	-	-	-	9.04
67	c-C ₆ H ₁₁	c-C ₅ H ₉ COCH ₂	CH ₂ COOH	-	-	-	-	9.55
68	c-C ₆ H ₁₁	c-C ₅ H ₉ COCH ₂	(CH ₂) ₂ COOH	-	-	-	-	9.47
69	c-C ₆ H ₁₁	c-C ₅ H ₉ COCH ₂	1,2,4-oxadiazol-3-yl-5 (2 <i>H</i>)-one	-	-	-	-	9.97
70	c-C ₆ H ₁₁	c-C ₅ H ₉ COCH ₂	SCH ₂ COOH	-	-	-	-	9.48
71	-	-	-	-	-	-	-	7.71
72	C ₅ H ₅ N	t-BuCOCH ₂	-	-	-	-	-	9.86
73	c-C ₆ H ₁₁	t-BuCOCH ₂	-	-	-	CH	CH	9.53
74	c-C ₆ H ₁₁	t-BuCOCH ₂	-	-	-	N	CH	7.77
75	c-C ₆ H ₁₁	t-BuCOCH ₂	-	-	-	CH	N	9.08
76	c-C ₆ H ₁₁	t-BuCOCH ₂	-	-	-	CH	CH	9.11
77	c-C ₆ H ₁₁	t-BuCOCH ₂	-	-	-	CH	N	8.03
78	c-C ₆ H ₁₁	t-BuCOCH ₂	-	-	-	CH	CH	8.22
79	c-C ₆ H ₁₁	t-BuCOCH ₂	-	-	-	CH	N	8.56
80	Ph	Pyrrolidin-1-yl	H	-	-	1	-	6.00
81	Ph	Pyrrolidin-1-yl	COOH	-	-	1	-	6.42
82	c-C ₇ H ₁₃	Pyrrolidin-1-yl	CH ₃	-	-	1	-	7.63
83	c-C ₇ H ₁₃	t-Bu	CH ₃	-	-	1	-	8.12
84	c-C ₆ H ₁₁	t-Bu	NHCH ₃	-	-	1	-	8.02
85	c-C ₆ H ₁₁	t-Bu	2-methyl-thiazol-4-yl	-	-	1	-	8.02
86	c-C ₇ H ₁₃	t-Bu	CH ₃ COOH	-	-	1	-	8.50
87	c-C ₆ H ₁₁	t-Bu	CH ₃ CH ₂ COOH	-	-	1	-	8.20
88	c-C ₇ H ₁₃	t-Bu	SCH ₂ COOH	-	-	1	-	9.05
89	c-C ₆ H ₁₁	t-Bu	N(Me)-1-((2 <i>H</i>)-tetrazol-5-yl)-amino	-	-	1	-	8.37
90	c-C ₇ H ₁₃	t-Bu	1,2,4-oxadiazol-3-yl-5 (2 <i>H</i>)-one	-	-	1	-	9.37
91	c-C ₆ H ₁₁	t-Bu	1,2,4-oxadiazol-3-yl-5 (2 <i>H</i>)-one	-	-	1	-	8.92
92	c-C ₆ H ₁₁	t-Bu	1,2,4-oxadiazol-3-yl-5 (2 <i>H</i>)-one	-	-	3	-	7.68

^aC.N. compound name.

Table 2. Composition (Features), Cost (Bits), and Statistical Parameters (rmsd and Correlation) Associated with the 10 Best Hypotheses (Pharmacophore Models)

S.N. ^c	hypothesis	features ^a	total cost	Δcost ^b	rmsd	correlation (r)
1	Hypo'-1	HBA, HYAl, HYAr, NI, RA	86.15	27.03	0.655	0.958
2	Hypo'-2	HBA, HYAl, HYAr, NI, RA	87.554	21.7	0.743	0.946
3	Hypo'-3	HBA, HYAl, HYAr, NI, RA	87.901	20.4	0.795	0.938
4	Hypo'-4	HBA, HYAl, HYAr, NI, RA	87.996	18.9	0.799	0.937
5	Hypo'-5	HBA, HYAl, HYAr, NI, RA	88.051	18.3	0.786	0.939
6	Hypo'-6	HBA, HYAl, HYAr, NI, RA	88.110	18.2	0.774	0.942
7	Hypo'-7	HBA, HYAl, HYAr, NI, RA	88.204	18	0.795	0.938
8	Hypo'-8	HBA, HYAl, HYAr, NI, RA	88.209	17.9	0.775	0.942
9	Hypo'-9	HBA, HYAl, HYAr, NI, RA	88.570	17.5	0.831	0.931
10	Hypo'-10	HBA, HYAl, HYAr, NI, RA	88.89	17.4	0.852	0.928

^aHBA: hydrogen bond acceptor, RA: ring aromatic, HYAl: hydrophobic aliphatic, HYAr: hydrophobic aromatic, NI: negative ionizable. ^bΔcost = (null cost - total cost). ^cS.N.: serial number.

The fixed cost (81.89 bits) of the ten top-scored hypotheses was well separated from the null hypothesis cost (113.68 bits). The top ranked pharmacophore model (Hypo'-1) had the best predictive power and statistical significance described by the high correlation coefficient ($r^2_{\text{training}} = 0.92$), low rmsd (0.66), weight cost (1.29), error cost (67.99), and a cost difference of 27.03 satisfying the acceptable range recommended in the cost analysis of the CATALYST procedure.^{16,19} The configuration cost was 16.86 indicating that all generated models have been thoroughly analyzed. The cost difference between total and fixed cost for the best hypothesis was only 4.25 bits indicating

the high probability of the true correlation of the data. It is due to the fact that the lower the cost difference between the total cost and fixed cost, the higher is the probability for the true correlation of the data. Thus Hypo-1 was retained for further analysis as the best pharmacophore model for CCK-2R with five features viz. one H-bond acceptor (HBA), one aliphatic hydrophobic (HYAl), one aromatic hydrophobic (HYAr), one ring aromatic (RA), and one negative ionizable (NI) function at specific geometric locations in 3D space (Figure 2A) and is also statistically the most relevant model. The interfeature distance between all these features has been shown in Figure 2B.

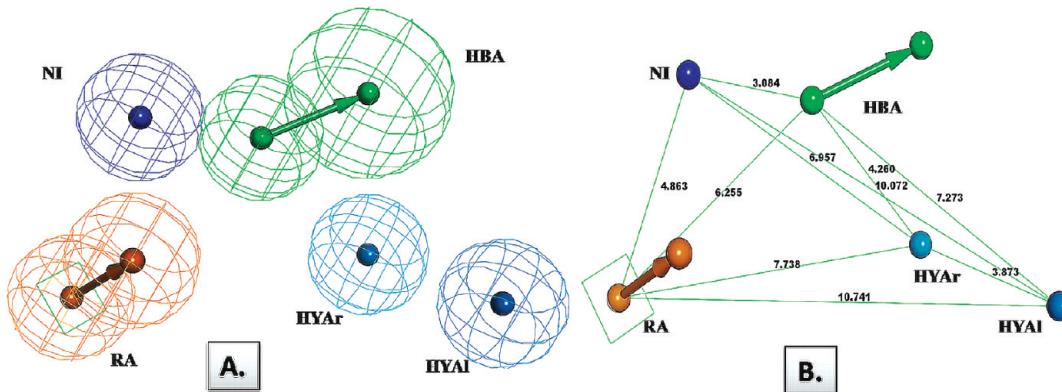


Figure 2. (A) Three dimensional arrangements of pharmacophoric features generated from the training set of 19 substituted antranilic sulfonamides as CCK-2R antagonist. (B) Interfeature distance between different features in the pharmacophore model.

Table 3. Observed and Estimated Activity Values of the Compounds in the Training Set^a

S.N.	compd name	obs. activity (K_i) nM	pred. activity (K_i) nM	error	obs. activity (pKi)	pred. activity (pKi)	error
Training Set							
1	1	25.10	13.97	-1.79	7.61	7.86	1.03
2	2	12.60	25.22	2.00	7.90	7.60	-1.04
3	3	10.00	7.60	-1.32	8.00	8.12	1.02
4	5	10.00	10.71	1.07	8.00	7.97	-1.00
5	6	6.30	10.83	1.72	8.20	7.97	-1.03
6	7	5.00	8.37	1.67	8.30	8.08	-1.02
7	8	5.00	3.58	-1.40	8.30	8.45	1.02
8	11	251.20	849.97	3.38	6.60	6.08	-1.09
9	13	7943.30	2544.66	-3.12	5.10	5.60	1.10
10	16	63.10	136.50	2.16	7.20	6.87	-1.05
11	22	7943.30	4751.22	-1.67	5.10	5.32	1.04
12	23	631.00	1434.91	2.27	6.20	5.84	-1.06
13	24	501.20	931.27	1.86	6.30	6.03	-1.05
14	25	1584.90	1892.22	1.19	5.80	5.72	-1.01
15	26	158.50	162.26	1.02	6.80	6.79	-1.00
16	27	1584.90	326.55	-4.85	5.80	6.49	1.12
17	28	1000.00	498.05	-2.00	6.00	6.31	1.05
18	29	1258.90	817.64	-1.53	5.91	6.09	1.03
19	31	2511.90	2077.14	-1.20	5.60	5.69	1.02
Test Set							
20	4	10.00	7.94	-1.25	7.75	8.10	1.05
21	9	6.30	16.40	2.59	11.59	7.79	-1.49
22	10	10.00	49.40	4.94	13.94	7.31	-1.91
23	12	6309.60	2704.80	-2.33	6.67	5.57	-1.20
24	14	79.40	81.05	1.02	10.02	7.10	-1.41
25	15	158.50	21.37	-7.41	1.59	7.68	4.83
26	17	631.00	331.89	-1.90	7.10	6.48	-1.10
27	18	15.80	335.15	21.21	30.21	6.48	-4.66
28	19	501.20	497.61	-1.00	8.00	6.33	-1.26
29	20	39.80	317.71	7.98	16.98	6.50	-2.61
30	21	316.20	1076.07	3.40	12.40	5.97	-2.08
31	30	1258.90	454.91	-2.77	6.23	6.34	1.02
32	32	1000.00	7514.27	7.51	16.51	5.12	-3.22

^aValues in the error column represent the ratio of estimated activity to experimental activity or its negative inverse if the ratio is less than 1, S.N. = serial number.

The developed pharmacophore model mapped very well to the training set of molecules. The observed K_i along with predicted K_i values for the CCK-2R antagonist activity of the training set compounds are given in Table 3.

A plot between the pKi values of observed activity versus the estimated activities demonstrated a good correlation coefficient

($r^2_{\text{training}} = 0.92$) for training set molecules within the range of uncertainty 3, indicating the high predictive ability of the pharmacophore (Figure 3A).

Pharmacophoric representation of the most potent compound 7 ($K_i = 5$ nM) of the data set showed that the phenyl ring of 3-bromo-4-fluorophenyl group occupies the RA feature

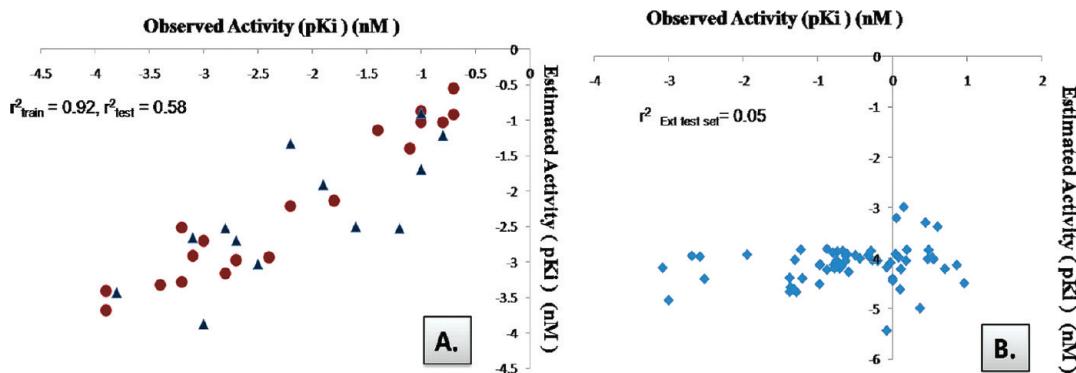


Figure 3. (A) Correlation graph displaying the observed versus estimated pKi (nM) values for the training set of 19 (shown in red circles) and test set of 13 molecules (shown in blue triangles). (B) Correlation graph for observed versus estimated pKi (nM) values of 1,3,4-benzotriazepine class of CCK-2R antagonists.

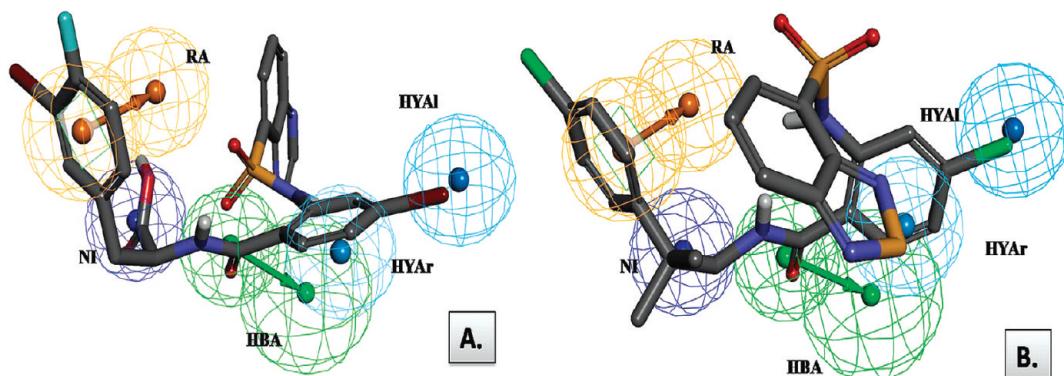


Figure 4. (A) Mapping of the most active compound 7 of the data set onto the pharmacophore. (B) Pharmacophore mapping of the least potent analogue of the data set 22 showing one completely missed feature of the pharmacophore i.e. NI.

of the pharmacophore, while the propanoic acid attached to the same group provides the NI feature (Figure 4A). The ketonic oxygen of the 4-iodobenzamido provides the HBA feature, while the phenyl ring and iodine occupied the HYAr and HYAl features, respectively. The same study carried with the least active compound of the data set 22 ($K_i = 7943.3$ nM) depicts the similar trend for phenyl ring of 4-chlorophenyl group where it occupies the RA feature. Similarly, the phenyl ring and chlorine of 4-chlorophenyl group occupies the HYAr and HYAl features, respectively. The least CCK-2R antagonistic activity of compound 22 may be due to the absence of the negative ionizable group to map the NI feature and the poor mapping of the ketonic oxygen of amide group to the HBA feature due to its altered conformation (Figure 4B).

It was observed from the pharmacophore mapping of the compounds of the training set, that among all active substituted anthranilic sulfonamides (7, 8, 3, 5, 6, 1, 2) provided, the NI and HBA features were provided by the carboxylic group and the ketonic oxygen of amide group, respectively. Hence mapping of these features seemed important for increasing the CCK-2R antagonist potency. The hydrophobic (one HYAl, one HYAr) and one RA feature also seems to be essential for describing the CCK-2R antagonist potency as these features were mapped by all the molecules of the training set. Since the sulfonamide moiety is present at the same position and is common to all the molecules of the training set, hence this feature was not considered for pharmacophore generation by the HypoGen algorithm of DS 2.0.

Test Set Validation. The robustness of the generated model was further assessed by its predictability on the test set containing 13 substituted anthranilic sulfonamide derivatives as CCK-2R antagonists. The observed pKi along with predicted pKi values of the test set compounds are given in Table 3. The overall correlation coefficient value ($r^2_{test} = 0.58$) between the observed and estimated activity of the thirteen test set molecules was indicative of its good predictive quality (Figure 3A).

In the most potent analogue of the test set, compound 9 follows the similar trend as with the highly active analogues of the training set where the similar type functionalities occupy the HYAr, HYAl, RA, and NI features. The least potent member of the test set, compound 12, failed to map NI feature completely, while the HBA feature partially and hence was ranked as least active.

Pharmacophore Mapping of 1,3,4-Benzotriazepine Class of CCK-2R Antagonists. In order to check the predictability of the developed model over other class of CCK-2R antagonists, a series of fifty 1,3,4-benzotriazepine class of CCK-2R antagonists^{10,15} was mapped to this pharmacophore. Interestingly this pharmacophore failed to predict the activity of these compounds ($r^2 = 0.05$) which may be due to the absence of any molecule of this class in the training set (Figure 3B). Therefore it was inferred that in spite of the high predictivity of this pharmacophore model for the substituted anthranilic sulfonamide class of CCK-2R antagonists, it cannot be used as a tool for virtual screening and designing of other classes of chemical entities such as CCK-2R antagonists. Further, the most active molecule considered for the generation of this pharmacophore

Table 4. Composition (Features), Cost (Bits), and Statistical Parameters (rmsd and Correlation) Associated with the 10 Best Hypotheses (Pharmacophore Models)

S.N. ^c	hypothesis	features ^a	total cost	Δcost ^b	rmsd	correlation (r)
1	Hypo-1	HBA, HYAl, HYAl, HYAr, RA	155.85	92.151	1.17	0.909
2	Hypo-2	HBA, HYAl, HYAl, HYAr, HYAr	166.50	82.125	1.40	0.867
3	Hypo-3	HBA, HYAl, HYAl, HYAr, RA	166.92	81.702	1.41	0.865
4	Hypo-4	HBA, HYAl, HYAl, HYAr, RA	167.12	81.507	1.42	0.862
5	Hypo-5	HBA, HYAl, HYAl, HYAr, HYAr	169.99	78.636	1.49	0.848
6	Hypo-6	HBA, HYAl, HYAl, HYAr, RA	171.82	76.804	1.54	0.836
7	Hypo-7	HBA, HBA, HYAl, HYAl, HYAr	173.68	74.94	1.57	0.828
8	Hypo-8	HBA, HYAl, HYAl, HYAr, HYAr	174.56	74.067	1.57	0.829
9	Hypo-9	HBA, HBA, HYAl, HYAl, HYAr	174.70	73.922	1.60	0.823
10	Hypo-10	HBA, HBA, HYAl, HYAr, HYAr	177.03	71.59	1.64	0.812

^aHBA: hydrogen bond acceptor, RA: ring aromatic, HYAl: hydrophobic aliphatic, HYAr: hydrophobic aromatic, NI: negative ionizable. ^bΔcost = (null cost – total cost). ^cS.N.: serial number.

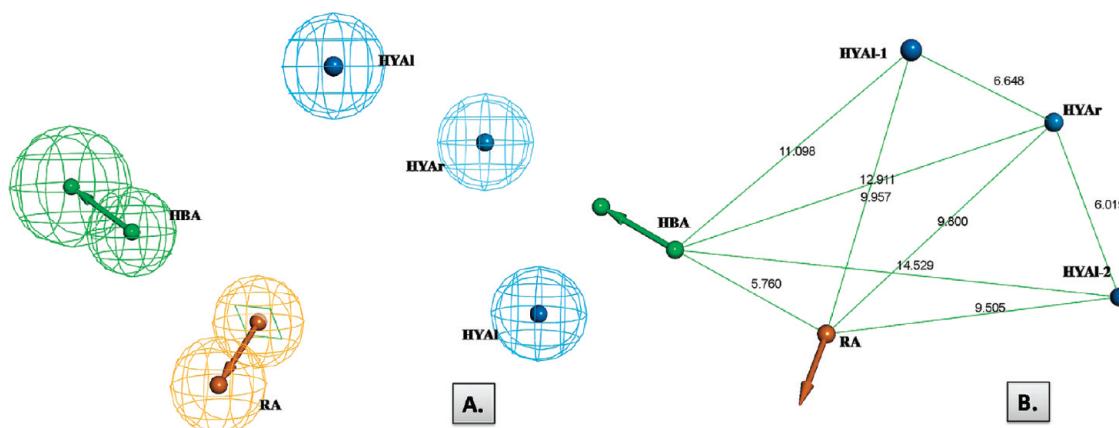


Figure 5. (A)Three dimensional arrangements of pharmacophoric features generated from the training set of 34 anthranilic sulfonamide and 1,3,4-benzotriazepine class as CCK-2R antagonist. **(B)** Interfeature distance between different features in the pharmacophore model.

is 5 nM; therefore, this model even cannot be used for the designing/virtual screening of the highly active sulfonamide class of CCK-2R antagonists with IC₅₀ comparable to the 1,3,4-benzotriazepine class of CCK-2R antagonists (0.11 nM).

In view of the above limitations, the present QSAR study was extended, and the data set comprising the anthranilic sulfonamide derivatives was merged with the data set comprising 1,3,4-benzotriazepine derivatives (compounds 33 to 92 of Figure 1, Table 1) to develop a more versatile and reliable pharmacophore model for CCK-2R antagonists that can even virtual screen/design the sulfonamide class of compounds with CCK-2R antagonist potency comparable with the 1,3,4-benzotriazepine class of CCK-2R antagonists.

Pharmacophore Model for Anthranilic Sulfonamide and 1,3,4-Benzotriazepine Class of CCK-2R Antagonists. The data set comprising anthranilic sulfonamide and 1,3,4-benzotriazepine class of CCK-2R antagonists were divided into a training set of 34 compounds and a complementary test set of 58 compounds according to published guidelines.¹⁶ The structures of test and training set molecules have been given in Figure 1, Table 1(compounds 1 to 92).

Pharmacophore Description. The procedure of pharmacophore generation was similar to the procedure adopted for the above-described pharmacophore. Interestingly, the top 10 alternative pharmacophore hypotheses developed from the training set compounds contained all the features except the NI feature present in the above-described pharmacophore for sulfonamides possibly because of the absence of this feature in some of

the active molecules of the training set. The quality of the generated pharmacophore hypotheses was evaluated by considering the cost functions calculated by the HypoGen module during hypothesis generation. The composition (in terms of chemical features that are necessary for activity), ranking score, and statistical parameters associated with the hypotheses are reported in Table 4.

The fixed cost (131.151 bits) of the ten top-scored hypotheses was well separated from the null hypothesis cost (248.624 bits). The top ranked pharmacophore model (Hypo-1) had the best predictive power and statistical significance described by the high correlation coefficient ($r_{\text{training}} = 0.90$), low rmsd (1.16), weight cost (2.56), error cost (137.619), and a well satisfying cost difference (92.151 bits) recommended for the CATALYST procedure.^{16,19} The configuration cost was 15.66 indicating that all generated models were thoroughly analyzed. The cost difference between total and fixed cost for the best hypothesis was only 24.698 bits indicating the high probability of the true correlation of the data. Thus Hypo-1 was retained for further analysis as the best pharmacophore model for CCK-2R antagonist activity with five features viz. one HBA, two HYAl, one HYAr, and one RA function at specific geometric locations in 3D space (Figure 5A) and it was also the most relevant model statistically. The interfeature distance between all these features has been shown in Figure 5B.

The developed pharmacophore model mapped very well to the training set molecules. The observed K_i along with

Table 5. Observed and Estimated Activity Values of the Compounds in the Training Set^a

S.N.	compd name	obs. activity (K_i) nM	pred. activity (K_i) nM	error	obs. activity (pKi) nM	pred. activity (pKi) nM	error
1	3	10.00	24.19	2.42	8.00	7.62	-1.05
2	6	6.30	6.75	1.07	8.20	8.17	-1.00
3	7	5.00	8.78	1.76	8.30	8.06	-1.03
4	8	5.00	30.28	6.06	8.30	7.52	-1.10
5	10	10.00	7.68	-1.30	8.00	8.11	1.01
6	13	7943.30	581.31	-13.66	5.10	6.24	1.22
7	16	63.10	403.86	6.40	7.20	6.39	-1.13
8	21	316.20	564.01	1.78	6.50	6.25	-1.04
9	22	7943.30	670.19	-11.85	5.10	6.17	1.21
10	23	631.00	936.21	1.48	6.20	6.03	-1.03
11	26	158.50	438.78	2.77	6.80	6.36	-1.07
12	28	1000.00	372.61	-2.68	6.00	6.43	1.07
13	29	1258.90	397.69	-3.17	5.90	6.40	1.08
14	31	2511.90	1095.93	-2.29	5.60	5.96	1.06
15	37	23.99	6.03	-3.98	7.62	8.22	1.08
16	39	331.13	460.55	1.39	6.48	6.34	-1.02
17	41	16.22	25.29	1.56	7.79	7.60	-1.03
18	42	489.78	17.74	-27.60	0.31	7.75	1.23
19	43	5.62	6.06	1.08	8.25	8.22	-1.00
20	45	1.00	4.51	4.51	9.00	8.35	-1.08
21	47	1.20	2.63	2.19	8.92	8.58	-1.04
22	52	4.27	3.85	-1.11	8.37	8.41	1.01
23	56	1.00	4.00	4.00	9.00	8.40	-1.07
24	58	1.95	7.65	3.92	8.71	8.12	-1.07
25	60	0.71	0.36	-0.51	9.15	9.44	1.03
26	64	0.25	0.57	2.26	9.60	9.25	-1.04
27	69	0.11	0.03	-4.40	9.96	10.60	1.06
28	70	0.33	0.28	-1.19	9.48	9.56	1.01
29	75	0.83	0.72	-1.15	9.08	9.14	1.01
30	81	380.19	400.86	1.05	6.42	6.40	-1.00
31	83	7.59	28.05	3.70	8.12	7.55	-1.08
32	86	3.16	13.82	4.37	8.50	7.86	-1.08
33	90	0.43	2.22	5.16	9.37	8.65	-1.08
34	92	20.89	7.33	-2.85	7.68	8.13	1.06

^aValues in the error column represent the ratio of estimated activity to experimental activity, or its negative inverse if the ratio is less than 1, S.N. = serial number.

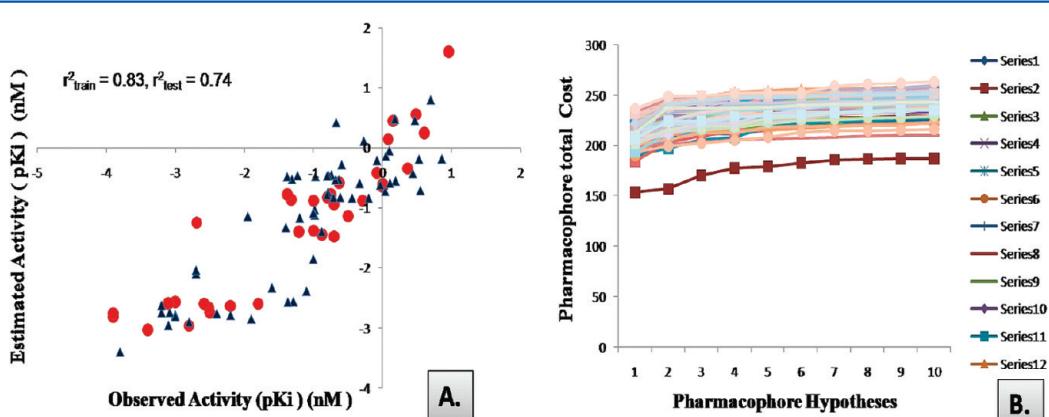


Figure 6. A. Correlation displaying the observed versus estimated pKi (nM) values for the training set of 34 (shown in red circles) and test set of 58 molecules (shown in blue triangles). B. The well differentiated total cost difference between the reported Hypogen run (dark colored lowest line) and the scrambled runs (other colored lines).

predicted K_i values for the CCK-2R antagonist activity of the training set compounds are given in Table 5.

A plot between the pKi values of the observed versus the estimated activities demonstrated a good correlation coefficient

($r^2_{\text{training}} = 0.83$) for training set molecules, indicating the high predictive ability of the pharmacophore (Figure 6A).

Previously a four feature (one HYAl, one HYAr, and two HBD) hypothesis was reported by Chopra et al. for CCK-2R

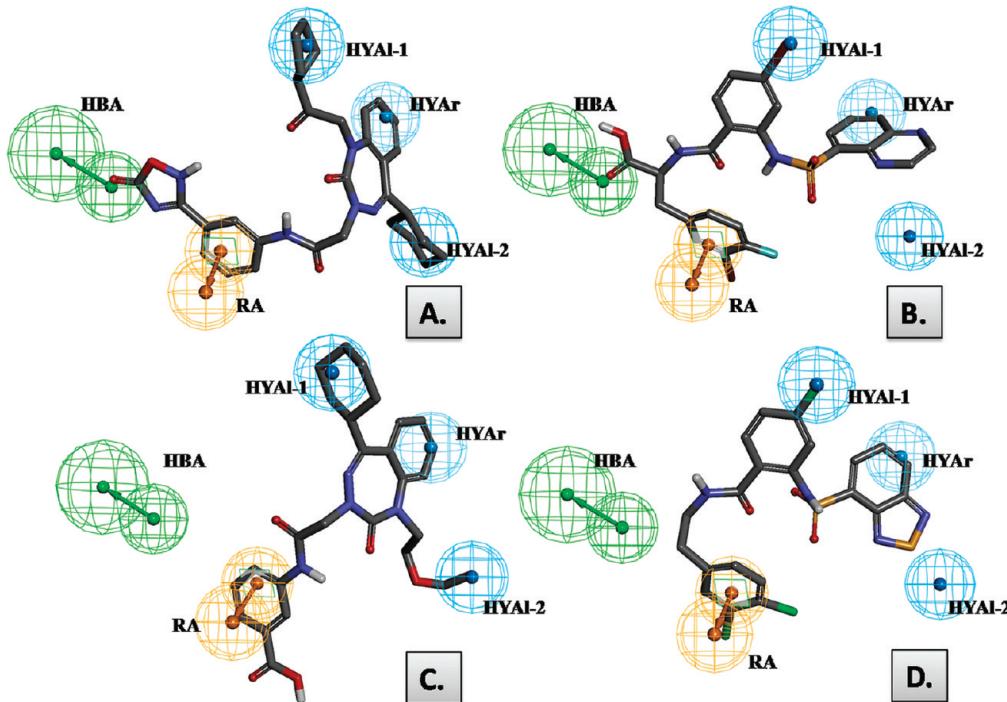


Figure 7. Pharmacophore mapping for (A) the most active compound, **69**, of the benzotriazepine class, (B) the most active compound, **7**, of the anthranilic sulfonamide class, (c) the least potent analogue, **42**, of the benzotriazepine class, and (D) the least potent analogue, **13**, of the benzotriazepine class.

antagonist activity.¹³ Comparison of the this hypothesis with our model describing well differentiated hydrophobic features (two HYAl, one HYAr, and one RA in our model) where one HYAl and HYAr features were corroborated with the earlier studies. However two HBD features were absent in our model which may be due to the absence of such functional groups in training set molecules. Additionally our pharmacophore describes the crucial role of the HBA feature for CCK-2R antagonist activity. This HBA feature corroborated with the earlier CoMFA, CoMSIA studies reported on the 1,3,4-benzotriazepine derivatives as CCK-2R antagonists where the favorable negative charge requirement in terms of 3D contours has been described by the authors.¹²

Pharmacophoric representation of the most potent compound **69** ($K_i = 0.11$ nM) of the data set as well as the 1,3,4-benzotriazepine class showed that the phenyl ring of phenylacetamide occupies the RA function to the molecule, while the carbonyl group of oxadiazole ring provides the HBA functionality. The benzene ring of the benzotriazepine occupied the HYAr functionality, while the cyclohexane and cyclopentane ring attached to the benzotriazepine moiety occupied the HYAl-1 and HYAl-2 features, respectively (Figure 7A). The most potent compound **7** and **10** ($K_i = 5$ nM and 10 nm) of the substituted anthranilic sulfonamide class also mapped well to the developed pharmacophore, where 3-bromo-4-fluoro phenyl ring of both of these molecules mapped the RA feature while the ketonic oxygen of carboxylic group of these molecules mapped the HBA feature. The benzene ring of the quinoxaline moiety mapped the HYAr feature, while the bromo or chloro group attached to the central anthranilic ring mapped the HYAl-1 functionalities (Figure 7B).

The pharmacophore mapping of the least active compound of the substituted anthranilic sulfonamide class of the data set **13** ($K_i = 7943.3$ nM) showed the occupancy of the HYAl-

feature with the halo-phenyl ring of anthranilic functionalities. Similarly, the benzene ring of benzothiazole occupied the HYAr feature; however, the HBA feature was completely missed by this molecule (Figure 7D).

The least active molecule of the benzotriazepine class, **42** ($K_i = 489.78$), has shown maximum deviation from the predicted activity (Figure 7C). Although this molecule followed the pharmacophore mapping trend similar to its other less active series analogues where it also missed the HBA feature however, the comparatively good prediction of this molecule may be due to the mapping of both the HYAl (1 and 2) features, unlike the missing of either of these two features by other less active benzotriazepine analogues.

The pharmacophore mapping of other less active compound of the benzotriazepine data set **81**, **39** ($K_i = 380.19$, and 331.13 nM) showed the mapping of the RA and HYAr feature with the phenyl ring of phenylacetamide and the benzene ring of the benzotriazepine functionality. However the HYAl-1 feature was missed in case **39** and the HYAl-2 feature was missed in case of molecule **81**. It has been observed from the pharmacophore mapping of the compounds of the training set, that among all the active compounds of the benzotriazepines and substituted anthranilic sulfonamides (**69**, **64**, **70**, **90**, **60**, **75**, **7**, **8**, **6**, **3**, **10**) the ketonic oxygen of groups attached to the phenyl ring at the N-3 position in the benzotriazepine series and carboxylic group in the anthranilic sulfonamide series provides the HBA feature. The weak CCK-2R antagonism of compounds (viz. **13**, **22**, **42**) seemed to be due to the absence of HBA function mapping in 3D space for these molecules (Figure 7B). Hence mapping of HBA features seemed to be very crucial for the CCK-2R antagonist potency. The HYAl-2 feature has been missed by all the anthranilic sulfonamide class of CCK-2R antagonists suggesting the nonindispensable nature of this feature for all the class of CCK-2R antagonists. The hydrophobic (HYAl-1, HYAr)

Table 6. Observed and Estimated Activity Values of the Compounds in the Test Set^a

S.N.	compd name	obs. activity (K_i) nM	pred. activity (K_i) nM	error	obs. activity (pKi) nM	pred. activity (pKi) nM	error
1	1	25.10	21.14	-1.19	7.60	7.67	1.01
2	2	12.60	244.54	19.41	7.90	6.61	-1.19
3	4	10.00	71.42	7.14	8.00	7.15	-1.12
4	5	10.00	12.18	1.22	8.00	7.91	-1.01
5	9	6.30	2.93	-2.15	8.20	8.53	1.04
6	11	251.20	580.20	2.31	6.60	6.24	-1.06
7	12	6309.60	2526.23	-2.50	5.20	5.60	1.08
8	14	79.40	713.15	8.98	7.10	6.15	-1.16
9	15	158.50	617.49	3.90	6.80	6.21	-1.10
10	17	631.00	804.62	1.28	6.20	6.09	-1.02
11	18	15.80	14.68	-1.08	7.80	7.83	1.00
12	19	501.20	109.63	-4.57	6.30	6.96	1.10
13	24	501.20	125.29	-4.00	6.30	6.90	1.10
14	25	1584.90	560.77	-2.83	5.80	6.25	1.08
15	27	1584.90	423.49	-3.74	5.80	6.37	1.10
16	27	39.80	215.36	5.41	7.40	6.67	-1.11
17	30	1258.90	909.36	-1.38	5.90	6.04	1.02
18	32	1000.00	613.04	-1.63	6.00	6.21	1.04
19	33	5.13	6.76	1.32	8.29	8.17	-1.01
20	34	89.13	13.85	-6.44	7.05	7.86	1.11
21	35	23.99	3.02	-7.95	7.62	8.52	1.12
22	36	4.79	3.42	-1.40	8.32	8.47	1.02
23	38	1202.26	556.81	-2.16	5.92	6.25	1.06
24	40	20.42	3.34	-6.11	7.69	8.48	1.10
25	44	7.59	25.06	3.30	8.12	7.60	-1.07
26	46	0.79	1.11	1.40	9.10	8.95	-1.02
27	48	2.14	3.89	1.82	8.67	8.41	-1.03
28	49	1.86	0.77	-2.43	8.73	9.12	1.04
29	50	0.65	3.55	5.50	9.19	8.45	-1.09
30	51	1.07	4.15	3.87	8.97	8.38	-1.07
31	53	5.50	2.82	-1.95	8.26	8.55	1.04
32	54	6.17	5.55	-1.11	8.21	8.26	1.01
33	55	3.89	1.90	-2.05	8.41	8.72	1.04
34	57	0.20	0.16	-1.26	9.70	9.80	1.01
35	59	4.68	0.38	-12.41	8.33	9.42	1.13
36	61	1.58	6.90	4.35	8.80	8.16	-1.08
37	62	0.66	0.33	-2.00	9.18	9.48	1.03
38	63	0.36	2.66	7.33	9.44	8.57	-1.10
39	65	4.07	6.67	1.64	8.39	8.18	-1.03
40	66	0.91	5.29	5.80	9.04	8.28	-1.09
41	67	0.28	5.11	18.12	9.55	8.29	-1.15
42	68	0.34	0.35	1.02	9.47	9.46	-1.00
43	71	19.50	366.52	18.80	7.71	6.44	-1.20
44	72	0.14	1.52	11.01	9.86	8.82	-1.12
45	73	0.30	1.53	5.19	9.53	8.81	-1.08
46	74	16.98	2.90	-5.85	7.77	8.54	1.10
47	76	0.78	3.86	4.97	9.11	8.41	-1.08
48	77	9.33	2.99	-3.12	8.03	8.52	1.06
49	78	6.03	2.87	-2.10	8.22	8.54	1.04
50	79	2.75	6.85	2.49	8.56	8.16	-1.05
51	80	1000.00	645.16	-1.55	6.00	6.19	1.03
52	82	23.44	370.98	15.83	7.63	6.43	-1.19
53	84	9.55	10.68	1.12	8.02	7.97	-1.01
54	85	9.55	13.10	1.37	8.02	7.88	-1.02
55	87	6.31	6.09	-1.04	8.20	8.22	1.00
56	88	0.89	1.38	-1.54	9.05	8.86	-1.02
57	89	4.27	3.39	-1.26	8.37	8.47	1.01
58	91	1.20	1.62	1.35	8.92	8.79	-1.01

^aValues in the error column represent the ratio of estimated activity to experimental activity, or its negative inverse if the ratio is less than 1, S.N. = serial number.

and RA feature also seems to be essential for describing the CCK-2R antagonist potency as these features were mapped by all the active molecules of the training set. It is noteworthy to mention that the ketonic oxygen attached with the benzotriazepine nucleus of the 1,3,4-benzotriazepine CCK-2R antagonists provides the same functionality as with the sulfonamide moiety of the anthranilic sulfonamide class of CCK-2R antagonists (Figure 5A and 5B). The inability of the HypoGen algorithm of DS 2.0 to consider this feature for the pharmacophore generation was due to the presence of this functionality at the same position and common to all the active molecules of the training set.

Evaluation of the Hypogen Model. Cat Scramble Validation (Fisher Test). To evaluate the statistical relevance of the model, the Fischer's randomization test²⁰ has been applied. This test involves thorough randomization of the training set to validate and derive the significance of the generated best model. As a consequence, the pharmacophore model corresponding to the Hypo-1 was evaluated for the statistical significance using a randomization trial procedure derived from the Fisher method.¹⁹ These randomized spreadsheets should yield hypotheses with lesser statistical significance than the original model to suggest that the original hypothesis represents a true correlation. Our model was found to be 99% significant in the F-randomization test which substantiates the significance of the model (Figure 6B).

Test Set Validation. The robustness of the generated model was further assessed by its predictability on the test set of 58 substituted anthranilic sulfonamide and benzotriazepine derivatives as CCK-2R antagonists. The observed pKi along with predicted pKi values of these compounds are given in Table 6. The high overall correlation coefficient value ($r_{\text{test}} = 0.86$) between the observed and estimated activity of the 58 test set molecules was indicative of its excellent predictive quality (Figure 6A). The pharmacophore mapping of the test set compounds revealed the consistent mapping trend similar to the training set molecules where the similar type of functionalities mapped to the HYAr, HYAl, RA, and NI features. However the model predicted the activity of the two highly potent compounds **72** and **67** of the test set as moderately active (obs. activity: 0.14; pred. activity: 1.5 and Obs. activity 0.28; pred. activity: 5.17, respectively) which may to be due to improper mapping of the HBA feature by the amino and keto group attached with the phenylacetamide functionality of these molecules respectively.

Pharmacophore Comparison. In order to evaluate the differences between the pharmacophore developed using the compounds of the benzotriazepines and substituted anthranilic sulfonamides classes and the pharmacophore developed using the compounds of only substituted anthranilic sulfonamides class, a pharmacophore comparison study has been carried out using DS2.0 (Figure 8).

This pharmacophore comparison exercise deciphered the indispensable role of HBA, RA, and HYAl-1 feature for CCK-2R antagonist activity and the additive role of HYAl-2 feature. This comparison also explained why the anthranilic sulfonamide class (most active compounds **7**; activity, 5 nM) of CCK-2 antagonists devoid of the HYAl-2 feature were less active than the benzotriazepine derivatives (most active compounds **69**; activity, 0.11 nM) having the HYAl-2 feature. The other difference is of the NI feature present in the pharmacophore of anthranilic sulfonamides and absent in the pharmacophore for the combined

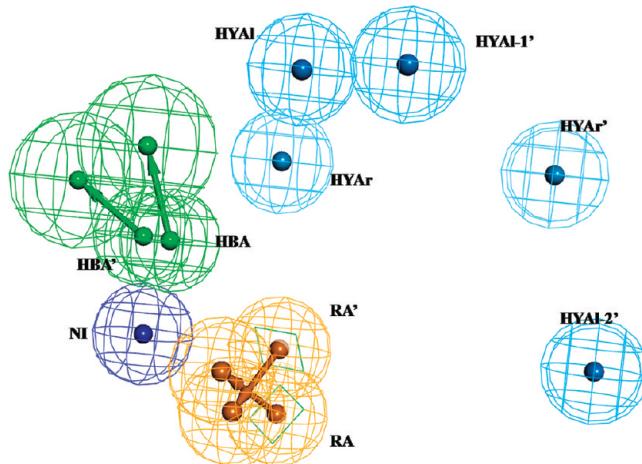


Figure 8. A comparison of pharmacophore developed from anthranilic sulfonamides class (feature are marked as RA, NI, HBA, HYAl, and HYAr) with the pharmacophore developed from benzotriazepines and substituted anthranilic sulfonamides classes together (features are marked as RA', HBA', HYAl-1', HYAl-2', and HYAr').

data set due to the absence of such functionalities in all the active molecules of the training set.

Validation of the Generated Hypothesis for Its Predictability against the Clinically Effective Synthetic CCK-2R Antagonists. The validity of the pharmacophore model developed from benzotriazepines and substituted anthranilic sulfonamides class was further examined against the clinically effective representative of the synthetic CCK-2R antagonists belonging to other diverse classes as an external test set.⁸ The observed and predicted activities of these molecules are listed in Table 7. Interestingly these molecules mapped well to the

Table 7. Observed and Estimated Activity Values of the Clinically Effective CCK-2R Antagonists

S.N.	compound	obs. activity (K_i) nM	pred. activity (K_i) nM	obs. activity (pKi) nM	pred. activity (pKi) nM	error ^a
1	L-736380	0.05	6.48	10.30	9.19	-1.12
2	CP-310713	0.10	0.66	10.00	10.18	1.02
3	YF476	0.11	24.18	9.96	8.62	-1.16
4	RP73870	0.48	2.70	9.32	9.57	1.03
5	JB-93182	1.10	8.85	8.96	9.05	1.01
6	indole-2-one	1.11	15.12	8.95	8.82	-1.02
7	CI988	1.70	2.29	8.77	9.64	1.10
8	CR-2622	20.00	16.28	7.70	8.79	1.14
9	C-2345	700.00	538.11	6.15	7.27	1.18

^aValues in the error column represent the ratio of estimated activity to experimental activity, or its negative inverse if the ratio is less than 1, S.N. = serial number.

features of the best hypotheses, and the overall correlation between observed and predicted pKi values is 0.74 ($r = 0.74$) in spite of their evaluation by different research groups. This pharmacophore mapping exercise further signifies the indispensable role of HBA, RA, and HYAl-1 features for CCK-2R antagonist activity as all the highly active representative CCK-2R antagonists of different classes mapped well to these three features. However, the features HYAl-2 and HYAr may contribute as the additive features for the increment of CCK-2R

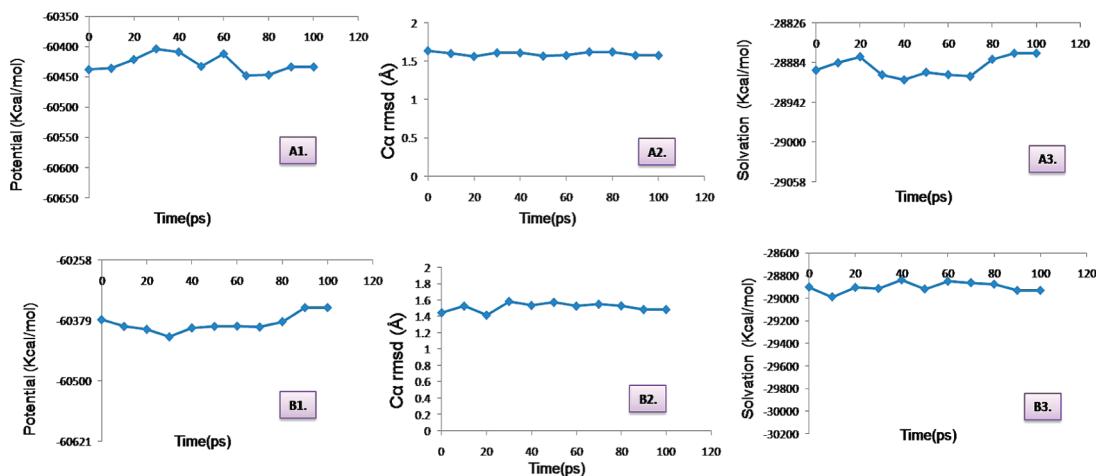


Figure 9. Graphical plot analysis between potential energy, $C\alpha$ backbone rmsd, and solvation energy of 10 sampled structures of ligand-protein complex; A1, A2, and A3 are for compound 7, whereas B1, B2, and B3 are for compound 69 respectively.

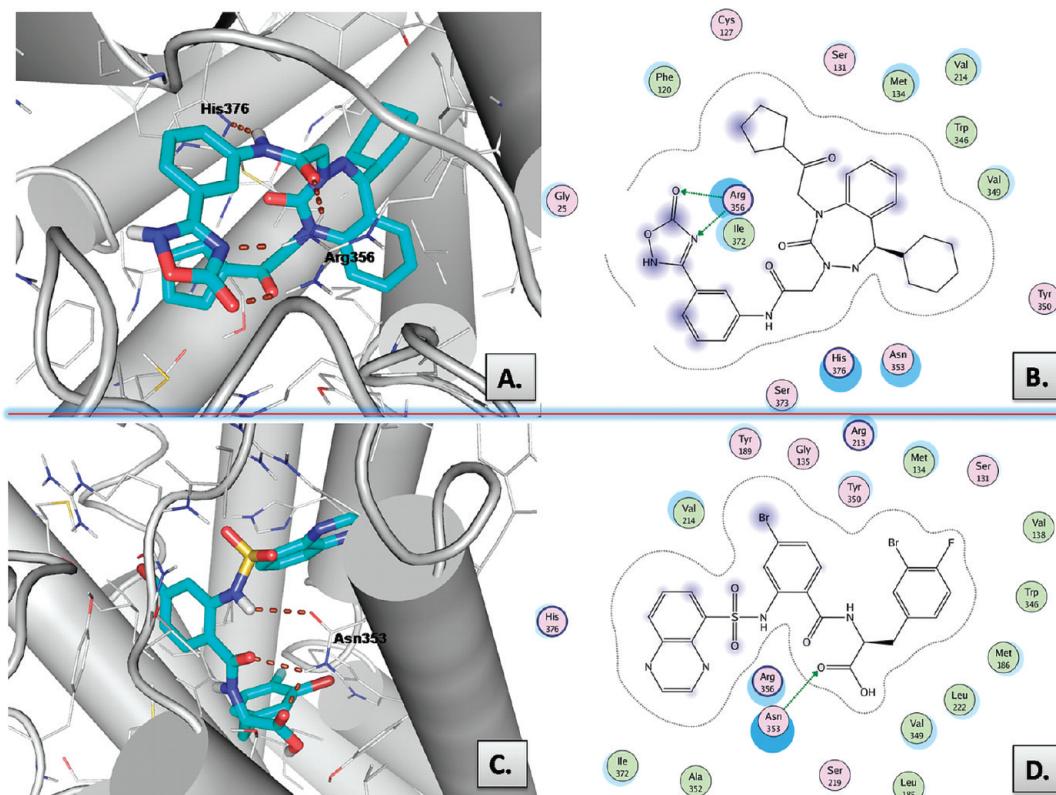


Figure 10. Binding pose analysis of compound 69 and 7 at the binding site of human CCK-2R: (A) 3D binding pose view of compound 69 (cyan color), (B) 2D binding pose view of compound 69, (C) 3D binding pose view of compound 7 (cyan color), and (D) 2D binding pose view of compound 7.

antagonist activity. Since the model discriminated well between active and inactive, it may be considered for virtual screening to prioritize the molecules for the synthesis and evaluation as CCK-2R antagonists.

Molecular Modeling of Human CCK-2R and Docking Studies. In the recent past, we have reported the benefits associated with the integrated ligand based and structure based application of computer-aided drug design (CADD) techniques applied on a different class of inhibitors.²¹ A plausible CCK2R model of rat has been reported several years ago by Jagerschmidt et al.²² This CCK2R model was itself obtained from an AT1a receptor model and was constructed using the

transmembrane (TM) helical positions found in bacteriorhodopsin structure.²³ In order to validate the developed pharmacophore model (Hypo-1), a plausible structural model of the human CCK-2R has been generated using the template of 2.8 Å resolution crystal structure of bovine rhodopsin (PDB ID: 1F88) as described by Foucaud et al.²⁴ Forrest and co-workers have suggested that plausible models of the TM regions of the membrane proteins can be obtained for the template: target sequence homology of $\geq 30\%$.²⁵ The overall sequence identity and similarity between human CCK2R (target) and bovine rhodopsin (template) was found to be 19.6 and 40.4% which further increased to 27.5 and 57.1% if

only the trans membrane (TM) region was considered. The higher sequence similarity (56%) between the TM region of human CCK2R and bovine rhodopsin further justified the selection of bovine rhodopsin (PDB ID: 1F88) as template. The amino acid sequence of the human CCK-2R receptor was retrieved from Swissprot database²⁶ (entry code P32239), and the sequence alignment has been carried out using the Prime²⁷ module of the Schrodinger software package.²⁸ Highly conserved residues in each TM were anchored, and the plausible model for CCK-2R was generated in the prime module of the hSchrodinger software package. Homology models were inspected to ensure that the side chains of the conserved residues were aligned to the template. The outlier residues were examined using the Ramachandran plot, and the loops with outlier residues were refined using the "refine loops" tool with the "extended medium-serial-loop sampling" procedure. The resulting plausible model for CCK-2R was backbone-constrained energy minimized using OPLS 2005 force field using MacroModel.²⁹ Protein was prepared using protein preparation wizard implemented in the Schrödinger software package. The pharmacophore based conformation of highly active molecules of the training set were subjected to the induced fit docking algorithm of the Schrodinger where the amino acid residues of proteins were also allowed to move along with docked ligand. The receptor grid (active site of the protein) was generated using the active site residues namely Asn353, Arg356 (TM6), His 207(EL2), and Tyr189 (TM4) as these residues showed the involvements in CCK-2R antagonism for non-peptide antagonists through earlier site directed mutagenesis and docking studies.²⁴ To ensure the correct binding the initial docked complexes were subjected to low-temperature molecular dynamics simulation (MDS) for 100 ps using MacroModel. The graphical plot analysis of Ca backbone rmsd and potential energy of protein ligand complex of 10 sampled structures for each ligand substantiated the stability of the bound ligand into the active site of CCK-2R (Figure 9).

The binding site analysis of the most active ligands of the training set (compound 69) clearly showed that the carbonyl oxygen of oxadiazole and acetamide group as well as heterocyclic nitrogen of oxadiazole ring were involved in H-bonding with the Arg356 corresponding to the HBA feature of the Hypo-1; additionally the amino hydrogen of acetamide group interacts with His376 through H-bonding. The phenyl ring of acetamide occupies the hydrophobic pocket made by Ile372, His376, Met 134, and Arg356 corresponding to the RA feature of the Hypo-1. Similarly, the benzene ring of the benzotriazepine group, cyclopentane ring and cyclohexane ring attached to the benzotriazepine moiety occupied the hydrophobic pocket made by the residues Val214, Met186 and Val130, Phe120, Arg213 and Asn353, Trp346, Tyr350, and Val349 corresponding to the HYAr, HYAl-1, and HYAl-2 feature of the Hypo-1, respectively (Figure 10A and 10B).

Similar study with the most active ligands of the anthranilic sulfonamide class (compound 7) of the training set showed the carbonyl oxygen of the carboxylic group and amide group interact with Asn353 through H-bonding corresponding to the HBA feature of the Hypo-1. The 3-bromo,4-fluoro phenyl ring well occupied the hydrophobic cleft made by the residues Leu185, Trp346, Tyr350, Val138, and Met134 corresponding to the RA feature of the Hypo-1. Similarly, the benzene ring of the quinoxaline group and halo group attached with the central anthranilic ring was placed into the hydrophobic cavity made by the residues Val349, Ile375, Ile372, and Val214, Tyr189,

Met186, Met134 corresponding to the RA and HYAl-1 features of the Hypo-1 (Figure 10C and 10D).

In summary, the docking studies well corroborate with the Hypo-1 where the importance of hydrophobic functionality at the active site has been described by one RA, one HYAr, and two HYAl features, while H-bond interactions at the binding site has been well described by one H-bond acceptor feature of the pharmacophore.

CONCLUSION

In search of a reliable study for finding essential structural requirements for CCK-2R antagonism a quantitative pharmacophore model has been developed using a training set of 34 substituted anthranilic sulfonamides and benzotriazepine as CCK-2R antagonists. The best hypothesis consisted of five features viz. two aliphatic hydrophobic, one aromatic hydrophobic, one H-bond acceptor, and one ring aromatic feature with an excellent correlation for 34 training set ($r^2_{\text{training}} = 0.83$) and 58 test set compounds ($r^2_{\text{test}} = 0.74$). This model was also found to be 99% significant in the F-randomization test. The docking study with the highly active CCK-2R antagonists of the training set at the active site of the CCK-2R well corroborates with the pharmacophore model. The model also well explained the observed CCK-2R antagonist activity of known clinically effective synthetic CCK-2R antagonists. Therefore, the developed pharmacophore model may not only be useful in finding new anthranilic sulfonamides having better activity than the benzotriazepine class of CCK-2R antagonists but may also aid in design and develop new chemical entities (NCEs) as potent CCK-2R antagonists before their synthesis.

EXPERIMENTAL SECTION

Computational Tools. All calculations were performed on a Linux workstation equipped with four parallel Intel Xeon X5460 processors (2.8 GHZ) with 12 GB total RAM.

Pharmacophore Modeling. HypoGen algorithm allows identification of hypotheses that are common to the "active" molecules in the training set but at the same time not present in the "inactives".¹⁹ This algorithm allows a maximum of five features in pharmacophore generation out of an inbuilt collection of 11, which are hydrogen bond acceptor (HBA), hydrogen bond acceptor lipid (HBAL), hydrogen bond donor (HBD), hydrophobic (HY), hydrophobic aliphatic (HYAl), hydrophobic aromatic (HYAr), positive charged (PC), negative charged (NC), positively ionizable (PI), negatively ionizable (NI), and ring aromatic (RA) features. Among these features available for selection, five features viz. HBA, RA, NI, HYAl, and HYAr present in the training set were selected for pharmacophore generation. The choice and number of features used in the hypothesis construction were HBA (min.1 - max. 5), HBAL (min.1 - max. 5), HYAr (min.1 - max. 5), and HYAl (min.1 - max. 5). A default activity uncertainty value of 3 has been used in the pharmacophore generation.

An uncertainty Δ in the CATALYST paradigm indicates an activity value lying somewhere in the interval from "activity divided by Δ " to "activity multiplied by Δ ". The resultant best ten generated hypotheses were assessed on the basis of the cost relative to the null hypothesis and the correlation coefficients. The total Cost value is the summation of weight, an error, and a configuration cost value. The weight component is a value that increases in a Gaussian form because of the deviation of function weights from the ideal value of 2. Increase of the root-mean-square (rms) divergence between estimated and measured activities for the training set molecules results in a

higher error cost value. The configuration cost is a fixed cost that quantifies the entropy of the hypothesis space. In the standard HypoGen mode, the configuration cost should not exceed a maximum value of 17 (corresponds to a number of 2^{17} pharmacophore models) because high values may lead to chance correlation of the generated hypothesis, since DS 2.0 cannot consider more than 2^{17} models in the optimization phase and so the rest are left out of the process. The HypoGen module performs two additional theoretical cost calculations (represented in bit units) to help the user in assessing the statistical significance of the generated hypothesis. The fixed cost is the lowest possible cost representing a hypothetical, simplest model that fits all data perfectly. It is calculated by adding the minimum achievable error, weight cost, and the constant configuration cost. The null cost represents the maximum cost of a pharmacophore with no features and estimates activity to be the average of the training set molecule's activity data. Its absolute value is equal to the maximum occurring error cost. The statistical significance of the hypothesis increases when the total cost value is close to the fixed cost value, and the difference between null cost and total cost is high. The estimation of the activity of each training set compound is based on the regression analysis between the parameters computed by using the relationship of geometric fit value versus the negative logarithm of activity, and so higher the geometric fit value, the greater is the activity prediction of the compound. The fit function does not depend only on the mapping of the feature but also contains a distance term measuring the distance between the feature on the molecule and the centroid of the hypothesis feature, and both these terms are used in the calculation of geometric fitness.

Sequence Alignment and Model Building. Human CCK-2R sequence was retrieved from Swissprot database²⁶ (entry code P32239). The 2.8 Å resolution crystal structure of bovine rhodopsin (PDB ID: 1F88) was considered as templates for homology modeling of human CCK-2R. Multiple sequence alignment and model building was carried out using the Prime²⁷ module of Schrodinger software package.²⁸

Protein Preparation. Protein was prepared using Protein Preparation Wizard implemented in Schrodinger package using default options: bond orders were assigned, hydrogens were added, metals were treated, and water molecules 5 Å beyond hetero groups were deleted. Hydrogens were then optimized using the exhaustive sampling option, and the protein was minimized to an rmsd limit from the starting structure of 0.3 Å using the Impref module with the OPLS_2005 force field.

Molecular Dynamics Simulations. The MDS was run using settings as OPLS-2005 as force field, water as solvent model, constant dielectric as electrostatic treatment, Polak-Ribiere Conjugate Gradient (PRCG) as minimization method, maximum iterations of 500, molecular dynamics as dynamics method, simulation temperature of 300 K, time step of 1.5 fs, equilibration time of 10 ps, and simulations time of 100 ps. Rest settings were used as the default in the MacroModel.

The key interactions between these inhibitors and the binding site residues of *Ld* AdoHcyase are presented in the two-dimensional (2D) form using the DS, while the 3D-molecular graphics were produced using PyMol program.³⁰

ASSOCIATED CONTENT

Supporting Information

Pharmacophore mapping view for the clinically effective CCK-2R antagonists shown in Table 7. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Phone: +9152226124112/ex 4386. Fax: +9152226123405. E-mail: anilsak@gmail.com.

Author Contributions

[#]These authors have contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are thankful for financial assistance in the form of a fellowship by the Indian Council of Medical Research (ICMR) (A.K.G.), Ministry of Health (M.O.H.) (K.V.), and to Mr. A. S. Kushwaha for technical assistance. The CDRI communication number allotted to this manuscript is 8243.

REFERENCES

- (1) (a) Reubi, J. C.; Schaer, J. C.; Waser, B. Cholecystokinin(CCK)-A and CCK-B/gastrin receptors in human tumors. *Cancer Res.* **1997**, *57*, 1377–1386. (b) Konturek, P. C.; Nikiforuk, A.; Kania, J.; Raithel, M.; Hahn, E. G.; Muehldorfer, S. M. Activation of NFκappaB represents the central event in the neoplastic progression associated with Barrett's esophagus: a possible link to the inflammation and overexpression of COX-2, PPARgamma and growth factors. *Dig. Dis. Sci.* **2004**, *49*, 1075–1083. (c) Herranz, R. Cholecystokinin antagonists: pharmacological and therapeutic potential. *Med. Res. Rev.* **2003**, *23*, 559–605.
- (2) Colucci, R.; Blandizzi, C.; Tanini, M.; Vassalle, C.; Breschi, M. C.; Del Tacca, M. Gastrin promotes human colon cancer cell growth via CCK-2 receptor-mediated cyclooxygenase-2 induction and prostaglandin E2 production. *Br. J. Pharmacol.* **2005**, *144*, 338–348.
- (3) Lee, Y. M.; Beinborn, M.; McBride, E. W.; Lu, M.; Kolakowski, L. F.; Kopin, A. S. The human brain cholecystokinin-B/gastrin receptor. Cloning and characterization. *J. Biol. Chem.* **1993**, *268*, 8164–8169.
- (4) Poirot, S. S.; Dufresne, M.; Vaysse, N.; Fourmy, D. The peripheral cholecystokinin receptors. *Eur. J. Biochem.* **1993**, *215*, 513–529.
- (5) Wank, S. A. Cholecystokinin receptors. *Am. J. Physiol.* **1995**, *269*, 628–646.
- (6) Crawley, J. N.; Corwin, R. L. Biological actions of cholecystokinin. *Peptides* **1994**, *15*, 731–755.
- (7) McDonald, I. M. CCK2 receptor antagonists. *Expert Opin. Ther. Pat.* **2001**, *11*, 445–462.
- (8) Berna, M. J.; Tapia, J. A.; Sancho, V.; Jensen, R. T. Progress in developing cholecystokinin (CCK)/gastrin receptor ligands that have therapeutic potential. *Curr. Opin. Pharmacol.* **2007**, *7*, 583–592.
- (9) Pippel, M.; Boyce, K.; Venkatesan, H.; Phuong, V. K.; Yan, W.; Barrett, T. D.; Lagaud, G.; Li, L.; Morton, M. F.; Prendergast, C.; Wu, X.; Shankley, N. P.; Rabinowitz, M. H. Anthranilic sulfonamide CCK1/CCK2 dual receptor antagonists II: Tuning of receptor selectivity and in vivo efficacy. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6376–6378.
- (10) McDonald, I. M.; Black, J. W.; Buck, I. M.; Dunstone, D. J.; Griffin, E. P.; Harper, E. A.; Hull, R. A. D.; Kalindjian, S. B.; Lilley, E. J.; Linney, I. D.; Pether, M. J.; Roberts, S. P.; Shaxted, M. E.; Spencer, J.; Steel, K. I. M.; Sykes, D. A.; Walker, M. K.; Watt, G. F.; Wright, L.; Xun, W. Optimization of 1,3,4-Benzotriazepine-Based CCK Antagonists to Obtain Potent, Orally Active Inhibitors of Gastrin-Mediated Gastric Acid Secretion. *J. Med. Chem.* **2007**, *50*, 3101–3112.
- (11) Gupta, A. K.; Chakraborty, S.; Srivastava, K.; Puri, S. K.; Saxena, A. K. Pharmacophore modeling of substituted 1,2,4-trioxanes for quantitative prediction of their antimalarial activity. *J. Chem. Inf. Model.* **2010**, *50*, 1510–1520.
- (12) Kaur, K.; Talele, T. T. 3D QSAR studies of 1,3,4-benzotriazepine derivatives as CCK 2 receptor antagonists. *J. Mol. Graphics Modell.* **2008**, *27*, 409–420.

- (13) Chopra, M.; Mishra, A. K. Ligand-Based Molecular Modeling Study on a Chemically Diverse Series of Cholecystokinin-B/Gastrin Receptor Antagonists: Generation of Predictive Mode. *J. Chem. Inf. Model.* **2005**, *45*, 1934–1942.
- (14) Pippel, M.; Allison, B. D.; Phuong, V. K.; Li, L.; Morton, M. F.; Prendergast, C.; Wu, X.; Shankley, N. P.; Rabinowitz, M. H. Anthranilic sulfonamide CCK1/CCK2 dual receptor antagonists I: Discovery of CCKR1 selectivity in a previously CCKR2-selective lead series. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6373–6375.
- (15) John Spencer, J.; John Gaffen, J.; Griffin, E.; Harper, E. A.; Linney, I. D.; McDonald, I. M.; Roberts, S. P.; Shaxted, M. E.; Adatia, T.; Bashall, A. Achiral, selective CCK 2 receptor antagonists based on a 1,3,5-benzotriazepine-2,4-dione template. *Bioorg. Med. Chem.* **2008**, *16*, 2974–2983.
- (16) Kim, S. G.; Yoon, C.; Kim, S.; Cho, Y.; Kang, D. Building a common feature hypothesis for thymidylate synthase inhibition. *Bioorg. Med. Chem.* **2000**, *8*, 11–17.
- (17) Discovery Studio 2.0; Accelrys: 9685 Scranton Road, San Diego, CA 92121.
- (18) Smellie, A.; Teig, S. L.; Poling, T. P. Promoting Conformational Variation. *J. Comput. Chem.* **1995**, *16*, 171–187.
- (19) Gund, P. In *Pharmacophore Perception, Development, and Drug Design*; Gunner, O. F., Ed.; International University Line: California, U.S.A., 2000; pp 3–11.
- (20) Fischer, R. *The principle of experimentation, illustrated by a psycho-physical experiment. the design of experiments*, 8th ed.; Hafner Publishing Co.: New York, 1966; Chapter II.
- (21) Gupta, A. K.; Saxena, S.; Saxena, M. Integrated ligand and structure based studies of flavonoids as fatty acid biosynthesis inhibitors of *Plasmodium falciparum*. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4779–4781.
- (22) Jagerschmidt, A.; Guillaume-Roussellet, N.; Vikland, M. L.; Goudreau, N.; Maigret, B.; Roques, B. P. His381 of the rat CCKB receptor is essential for CCKB versus CCKA receptor antagonist selectivity. *Eur. J. Pharmacol.* **1996**, *296*, 97–106.
- (23) Joseph, M. P.; Maigret, B.; Bonnafous, J. C.; Marie, J.; Scheraga, H. A. A computer modeling postulated mechanism for angiotensin II receptor activation. *J. Protein. Chem.* **1995**, *14*, 381–398.
- (24) Foucaud, M.; Tikhonova, I. G.; Langer, I.; Escrieut, C.; Dufresne, M.; Seva, C.; Maigret, B.; Fourmy, D. Partial Agonism, Neutral Antagonism, and Inverse Agonism at the Human Wild-Type and Constitutively Active Cholecystokinin-2 Receptors. *Mol. Pharmacol.* **2006**, *69*, 680–690.
- (25) Forrest, L. R.; Tang, C. L.; Honig, B. On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophys. J.* **2006**, *91*, 508–517.
- (26) Swiss-prot. <http://www.uniprot.org/uniprot/p32239> (accessed December 12, 2011).
- (27) Prime, version 2.2; Schrödinger, LLC: New York, NY, 2010.
- (28) Schrödinger, version 9.1; Schrodinger, LLC: New York, 2005.
- (29) MacroModel, version 9.8; Schrödinger, LLC: New York, NY, 2010.
- (30) The PyMOL Molecular Graphics System, version 1.3r1; Schrodinger, LLC: Portland, OR, 2010.