Three-Dimensional Molecular Shape Analysis–Quantitative Structure–Activity Relationship of a Series of Cholecystokinin-A Receptor Antagonists

John S. Tokarski and Anton J. Hopfinger*

Department of Medicinal Chemistry and Pharmacognosy (M/C 781), College of Pharmacy, University of Illinois at Chicago, 833 South Wood Street, Chicago, Illinois 60612-7231

Received May 11, 1994®

The three-dimensional molecular shape analysis–quantitative structure–activity relationship (3D-MSA–QSAR) technique has been applied to develop correlations between the calculated physicochemical properties and the *in vitro* activities of a series of 3-(acylamino)-5-phenyl-2*H*-1,4-benzodiazepine cholecystokinin-A (CCK-A) antagonists. 3D-MSA–QSARs were developed for varying subsets of 53 analogs (*J. Med. Chem.* **1988**, *31*, 2235–2246). An active conformation is hypothesized for these compounds using the loss in biological activity–loss in conformational stability principle. After placing all compounds in the active conformation and performing pairwise molecular shape analysis, it was determined that not any one analog serves as the best shape reference compound. Nonidentical volumes of allowed receptor space are mapped out by different antagonists. A shape reference compound that consists of selected overlapped structures expands the definition of the accessible receptor space. This type of mutant improves the predicted activity of analogs over the value predicted if only one compound is chosen as the reference. Molecular shape, as represented by common overlap steric volume and nonoverlap steric volume, is the major factor contributing to the affinity of this class of compounds. Intramolecular conformational stability, as measured by the difference in energy of the active conformation and the global minimum energy conformation, is also important. It is further concluded from the 3D-MSA–QSAR models that part of the binding pocket for the 3-amido substituent has a preference for lipophilicity. The method used in this study of fragmenting the antagonist into spheres of varying radii and measuring lipophilicity isolates the substructure with highest probability of interacting with the receptor. Two indicator variables marking the presence of an *N*-methyl group and an *o*-fluoro atom on the 5′-phenyl substituent of the benzodiazepine ring structure also contribute significantly to the 3D-MSA–QSAR models. The 3D-MSA–QSAR results have led to the proposal of a 3D pharmacophore model for the benzodiazepine CCK-A antagonists.

Introduction

Cholecystokinin (CCK) is a family of peptides that functions as a digestive hormone in the periphery and as a neuromodulator in the central nervous system.[1,2] Two receptor subtypes of CCK have been characterized.[3] The CCK-A receptor subtype predominates in peripheral target organs such as the pancreas, gallbladder, and ileum but is also found in discrete regions of the central nervous system.[3,4] CCK-B receptors exist in numerous brain regions and show a structural and pharmacological profile similar to that of gastrin receptors.[5,6] Peripheral activities of CCK include stimulation of biliary and pancreatic secretion, gallbladder contraction, and inhibition of gastric emptying time.[7] Centrally, CCK may modulate dopaminergic and opiate-mediated neural transmission.[8] There is evidence that CCK-A receptors mediate CCK-facilitated dopaminergic efflux in rat brain.[9] In addition, satiety, produced by exogenously administered CCK, seems to be mediated by CCK-A receptors.[10] CCK-B agonists have been shown to cause panic attack in human volunteers,[11] and CCK-B antagonists have been found to ameliorate anxiety in animal studies.[12] CCK-A receptor involvement in this response has not been ruled out.

Potential therapeutic uses of CCK receptor antagonists have been proposed in the treatments of appetite disorders, pancreatitis, pancreatic carcinoma, abnormalities of gastric motility, biliary tract disease, irritable bowel syndrome, pain management, and psychiatric disorders.[13]

The geometry of the CCK-A receptor has not been determined, and therefore, little is known about the agonist-antagonist-binding pocket. Information about this receptor, however, may be gleaned from an analysis of a series of competitive inhibitors synthesized at Merck.[14] An inspection of the data set of the phenyl analogs in Table 1 suggests that conformational behavior might play a role in determining inhibition potency. Ortho-substitution could be expected to alter conformational profiles with respect to the flexible torsion angles of the compounds. Thus, we felt that molecular shape analysis (MSA)[15,16] should be performed on these compounds along with the other CCK antagonists of Tables 2 and 3. The results of these modeling studies are developed into three-dimensional quantitative structure–activity relationships (3D-QSARs) and are described herein.

Methods

a. Biological Activity. Evans et al.[14] reported the receptor binding, expressed as IC_{50}, to CCK receptors in rat pancreatic and guinea pig brain tissues and guinea pig gastric glands. This study focused on the pancreatic receptor values so that only CCK-A receptor antagonism was considered. Also, the receptor antagonism activities were expressed in the common QSAR form:

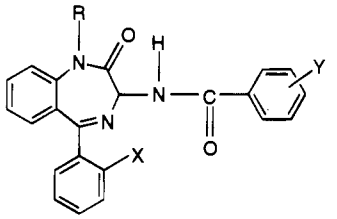
$$\text{activity} = -\log(IC_{50}, \mu M \times 10^{-6}) \quad (1)$$

where IC_{50} is the micromolar concentration required for half-

* To whom all correspondence should be addressed.
® Abstract published in *Advance ACS Abstracts,* September 1, 1994.

0022-2623/94/1837-3639$04.50/0

Table 1. Pancreas Receptor-Binding Affinities for 3-(Benzoylamino)benzodiazepines[a]

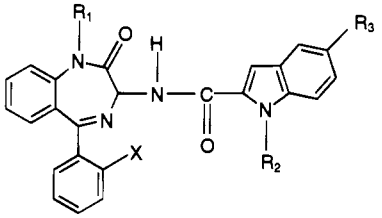


compd	X	R	Y	stereo	−log(IC50) CCK, pancreas
12	F	H	*p*-Cl	*RS*	8.22
61	F	H	*p*-NO_2	*RS*	6.96
62	F	CH_3	*p*-Cl	*RS*	8.64
64	H	CH_3	*p*-Cl	*RS*	8.08
65	H	H	*p*-Cl	*RS*	7.39
66	F	H	H	*RS*	6.82
67	F	H	*o*-Cl	*RS*	5.25
69	H	H	*o*-Cl	*RS*	5.12
70	H	CH_3	*o*-Cl	*S*	5.77
72	F	H	*p*-CF_3	*RS*	7.82
73	F	H	*p*-CH_3	*RS*	7.68
75	H	CH_3	*o*-Cl	*RS*	5.47
76	H	H	*m*-Cl	*RS*	7.09
77	F	H	*p*-OCH_3	*RS*	7.02
78	F	H	*p*-$N(CH_3)_2$	*RS*	6.57
79	F	H	3,4-di-OCH_3	*RS*	5.77
81	H	H	3,4-di-Cl	*RS*	7.54
83	H	CH_3	*p*-Br	*S*	8.60
85	F	CH_3	*p*-Cl	*S*	8.54
87	H	H	*p*-SCH_3	*RS*	6.66
88	H	H	*p*-F	*RS*	6.32
90	F	CH_3	*m*-Br	*S*	8.46
91	H	H	*m*-SCF_3	*RS*	5.82
92	H	H	*p*-CF_3	*RS*	6.62
93	F	CH_3	*p*-Br	*S*	9.10
94	F	CH_3	*p*-*t*-Bu	*S*	7.72
95	F	CH_3	*p*-I	*S*	9.10
96	F	CH_3	*o*-Br	*S*	5.21
97	F	CH_3	*p*-CN	*RS*	7.37
99	H	H	*p*-Ph	*RS*	4.00
101	H	H	*p*-*t*-Bu	*RS*	5.85
102	H	H	3,5-di-Cl	*RS*	6.30
103	H	H	*p*-OH	*RS*	5.75
104	F	CH_3	*m*-I	*S*	9.12
105	H	H	*p*-CN	*RS*	6.14
108	F	CH_3	*o*-I	*S*	6.10
112	H	CH_3	*o*-NH_2, *p*-Cl	*RS*	9.05
114	H	CH_3	*p*-CF_3	*S*	8.89
117	H	CH_3	*p*-Br	*RS*	9.00
118	F	CH_3	*p*-CF_3	*RS*	7.70
119	F	CH_3	*p*-CF_3	*S*	7.96
122	F	CH_3	*p*-*t*-Bu	*RS*	7.00
123	H	CH_3	*o*-I	*S*	5.96

[a] The receptor binding has been previously reported[11] and is expressed here as −log(IC50) calculated by eq 1.

maximal inhibition of binding of [125I]CCK-33 or [125I]CCK-8 to rat pancreas. The set of inhibition potencies, based upon the IC50 measure, for the selected compounds studied is given in Tables 1−3. In general, most of the isomers at the isomeric carbon of the benzodiazepine (BZD) ring structure have not been resolved with respect to inhibition potency. Therefore, this study concentrated on the racemates. However, the activity data of the *S*-isomers, the more active of the two isomers, have also been analyzed. The range in inhibition potency, IC50, for the racemates at the pancreatic receptor is 1−125 000 μM. This is a reasonably large range in inhibition potency concentration from the point of view of constructing statistically significant correlation relationships. Not every analog that the Merck group studied was included in the analysis. Fifty-three compounds were selected on the basis of potency, optical isomerism, and ease of modeling. The *R*-isomers were excluded. Also, the phenyl analogs containing a para *n*-propyl and *n*-pentyl substituent were not considered

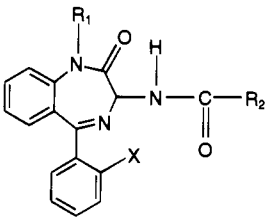
Table 2. Pancreas Receptor-Binding Affinities for 3-[(2-Indolylcarbonyl)amino]benzodiazepines[a]



compd	X	R_1	R_2	R_3	stereo	−log(IC50) CCK, pancreas
2	H	CH_3	H	H	*S*	10.10
9	H	H	H	H	*RS*	8.34
43	F	H	H	H	*RS*	8.68
44	F	CH_3	H	H	*RS*	8.85
45	F	CH_3	CH_3	H	*RS*	8.86
47	F	H	H	5′-F	*RS*	8.30
48	F	H	H	5′-Cl	*RS*	7.00
50	H	CH_3	H	H	*RS*	9.10
51	H	CH_3	CH_3	H	*RS*	8.85
53	F	H	H	5′-Br	*RS*	6.92
54	F	H	H	5′-OH	*RS*	7.15
55	F	H	H	5′-OCH_3	*RS*	6.70
56	F	H	CH_3	H	*RS*	8.33
58	F	CH_3	H	H	*S*	9.22

[a] Binding affinities defined as in Table 1, footnote a.

Table 3. Pancreas Receptor-Binding Affinities for 3-(Acylamino)benzodiazepines[a]

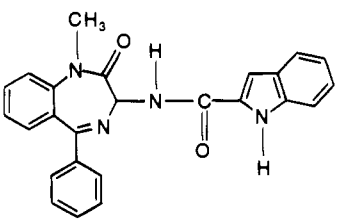


compd	X	R_1	R_2	stereo	−log(IC50) CCK, pancreas
5	H	H	3-indolyl	*RS*	5.96
6	H	H	CH_2-3-indolyl	*RS*	6.00
7	H	H	$(CH_2)_2$-3-indolyl	*RS*	5.32
8	H	H	$(CH_2)_3$-3-indolyl	*RS*	4.70
11	F	H	2-(1-CH_3)-1-indenyl	*RS*	7.66
13	F	H	2-quinazolinyl	*RS*	6.52
15	F	H	(*E*)-2-phenylethenyl	*RS*	7.96
16	F	H	(phenylamino)methyl	*RS*	7.05
17	F	H	2-benzofuranyl	*RS*	8.05
23	F	H	2-pyrrolyl	*RS*	6.96
24	H	H	2-naphthyl	*RS*	7.70
26	H	H	1-naphthyl	*RS*	5.80

[a] Binding affinities defined as in Table 1, footnote a.

because of the high flexibility of these substituent chains. The numbering scheme of the compounds for this entire article is based on that chosen by Evans et al.[14]

b. Building the Molecules. Two distinct crystal structures were observed in the X-ray structural analysis of **2**.[14]



2

These structures, A and B, are very similar and differ

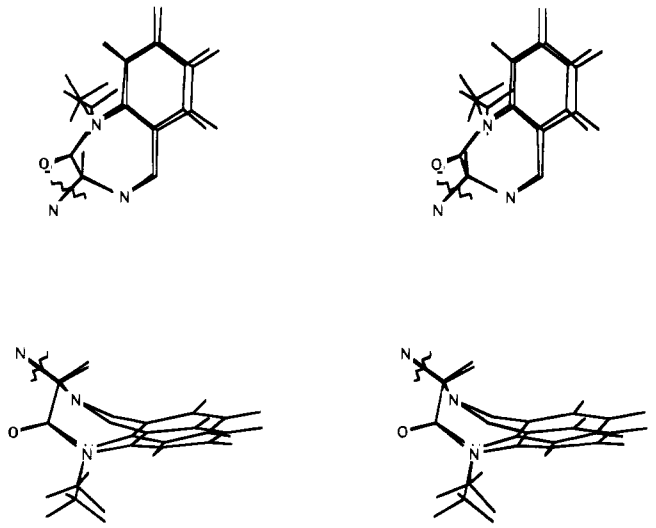
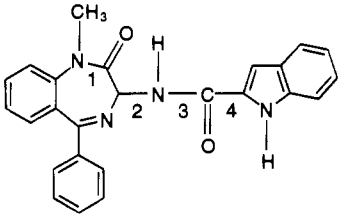


Figure 1. Stereoview comparison of the MNDO-optimized BZD ring structure and the crystal structure. Only the BZD ring structures with the 3-amido nitrogen and the 3-hydrogen are shown. The three atoms, the 3-carbon of the BZD ring, the 3-amido nitrogen, and the 3-hydrogen, of each structure were overlapped for the shape comparison. Two different perspectives are shown. The crystal structure is the one above the overlapped MNDO-optimized structure in each view.

Table 4. Comparison of Selected Bond Lengths from Different Sources for Compound **2**



source	bond lengths (Å)			
	1	2	3	4
crystal	1.359	1.448	1.346	1.476
Mopac	1.425	1.458	1.406	1.488
MMFF	1.352	1.431	1.347	1.363
experimental[a]	1.333	1.472	1.333	1.470

[a] The values listed as experimental are an average based on determinations of many related molecules (see ref 19).

primarily in the orientation of the 3-(2-indolecarboxamide) substituent with respect to the benzodiazepine ring. The A and B X-ray crystal coordinates of **2** produced by the Merck group were obtained, and the structures were built with the Chemlab-II molecular modeling package.[17] Both crystal structures (A and B) were used as starting geometries in separate structure optimizations using MNDO (Mopac version)[18] and the MMFF option in Chemlab-II, an extended version of Allinger's MM2 program.[19] The purpose of these calculations was to evaluate how sensitive the molecular geometry, particularly the BZD ring structure, is to the method of computation. The principal concern is the puckering of the seven-membered ring which can have a significant influence on the location of the 3-amido pharmacophore. Mopac optimized to a BZD ring geometry similar to the crystal structure (Figure 1) as compared to MMFF. The only major difference between the MNDO-optimized structure and the crystal structure was that both the endocyclic and exocyclic amide bond lengths were calculated to be almost 5% larger than in the crystal structure (see Table 4). The lengthening of the endocyclic amide bond does change the puckering of the BZD ring, but only slightly. MMFF predicts a BZD ring structure with a different molecular shape than the crystal structure and also calculates the bond lengths of the two major flexible torsion angles of the 3-amido substituent to be shorter than observed experimentally[20] (see Table 4). Consequently, the MMFF-minimized BZD ring structure was not used in the modeling studies.

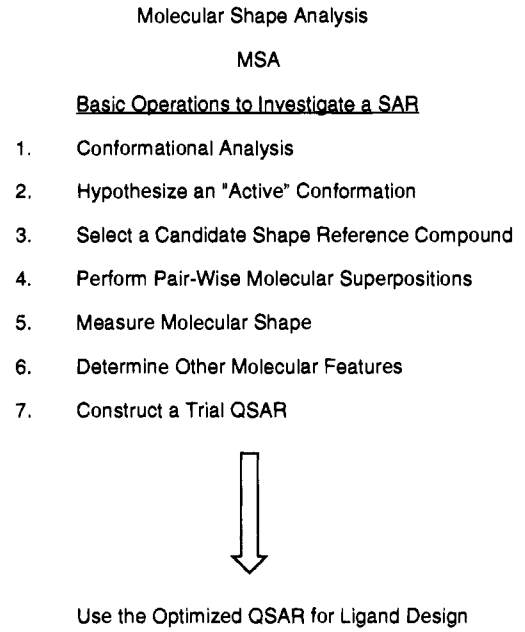


Figure 2. Basic operations of performing MSA and investigating a SAR.

The A and B crystal structures of **2** were both used to construct the analogs to test whether these structural variants would affect the conformational properties of the CCK antagonists. Since the crystal coordinates obtained for **2** were for the *S*-isomer, all other compounds subsequently built were of the same enantiomer. Fragments from the Chemlab-II library were used along with the crystal BZD structure and amide functionality to build the compounds that did not contain a 3-amido indole ring. The starting geometry of the crystal structure of **2** was retained for the indole analogs. The crystal structure of cinnamamide[21] was used to build compound **15**. Also, any necessary substituents on the BZD ring structure were added, or subtracted, to complete the analogs. The geometry of the entire structure of each analog was optimized using MNDO (Mopac version, keywords: MMOK). For all structures, MNDO-calculated atomic charges were employed in the modeling.

c. Molecular Shape Analysis. There are seven operations involved in the current formulation of MSA.[22] The manner in which these operations, as listed in Figure 2, were employed is summarized below.

(1) Conformational Analysis. In this study, the Chemlab-II modeling package option SCAN was used to perform a fixed valence conformational energy scan at 10° increments of ϕ_1 and ϕ_2 (see Figure 3 for definition of ϕ_1 and ϕ_2). The reference conformation is that for which the torsion angles ϕ_1 and $\phi_2 = 0°$. The amide group was kept transplanar as observed in the crystal structures of **2** for all scans. Flexible side chains were also scanned at various resolutions, depending on the number and type of fragments and if their torsion angles were coupled to ϕ_1 and ϕ_2.

For compounds **6**–**8**, **15**, and **16**, see Figure 4, more than two major torsion angles exist. In these cases, higher order conformational scans are required. The reference conformations of these compounds are those in which all rotated torsion angles are set at 0° (see Figure 4 for the definition of the rotated torsion angles). For compounds **6** and **15**, ϕ_1, ϕ_2, and ϕ_3 were scanned at 10° increments. For compounds **7**, **8**, and **16**, ϕ_1, which is common to all analogs, was scanned at 10° increments and all other angles were scanned at 30° increments.

A fixed valence geometry molecular mechanics force field composed of dispersion/steric, electrostatic, and hydrogen-bonding contributions was used to estimate the conformational energy. The nonbonded steric MMFF parameters from Chemlab-II were used to compute the dispersion/steric interactions.

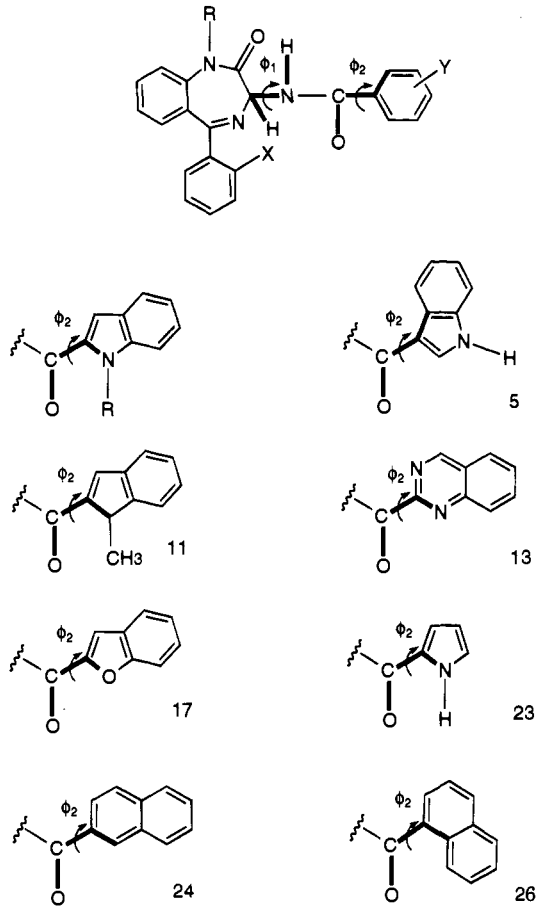


Figure 3. Two rotatable bonds which were scanned for the phenyl, indole, and other analogs (shown with thickened bonds). Since the torsion angle ϕ_1 is identical for all analogs, only unique torsion angles ϕ_2 are shown. The exact ring system is shown for those compounds of Table 3. There are two choices of the fourth atom of ϕ_2 for the compounds of Table 3. Both choices of ϕ_2 were examined in the conformational analysis. The torsion angle, which when placed in the active conformation and resulted in maximal common steric overlap volume and minimal nonoverlap volume with the shape reference compound, is shown with thickened bonds.

The electrostatic interactions were calculated using a Coloumb potential with a molecular dielectric of 3.5 and MNDO-calculated atomic charges. When hydrogen-bonding atoms were present, the hydrogen-bonding potential developed by Hopfinger was used.[23]

Another consideration in performing conformational analysis is the sensitivity of the preferred conformer states to choice in force field. A second set of nonbonded steric parameters, the Z set,[23] a "softer" potential set,[24] was used to assess the role of the force field on conformational behavior. The similarity of the two conformational profiles from the two sets of potentials was used to evalulate conformational sensitivity to force field, i.e., could the relative conformational properties of a series of compounds deferred from one potential force field be compared favorably with the respective properties obtained from employing a different force field? If so, this would support conclusions derived from the conformational analysis. The findings of this comparative force field analysis study appear in the Results section.

The global conformational energy minimum was used to define the relative stability of each conformational state sampled, that is, the relative conformational stability of a compound is defined as the difference in energy between a particular conformation and the global free space intramolecular conformational energy minimum.

(2) Selection of the Shape Reference Compound and Active Conformation. Once conformational analysis is performed, the active, or biologically relevant, conformation is sought. One modeling strategy used in this MSA step is termed the LBA–LCS (loss in biological activity–loss in conformational stability) strategy.[25] This method attempts to identify stable intramolecular conformer states for active analogs that are not stable for inactive analogs. The premise implicitly made is that the loss in activity for the inactive analogs is a result of not being able to adopt the "active" conformation energetically available to the active compounds. An arbitrary cutoff energy difference of 8 kcal/mol was chosen because this limit allows sampling an exhaustive amount of torsion angle space in which an active conformation, if present, should be found. In the study of the test set, conformations within 8 kcal/mol of the global energy minimum for each analog were considered as candidates for the active conformation. A search for an active conformation began with the conformational analysis of a subset of compounds from the entire data set, namely, **50**, **51**, **64**, **75**, and **112**. These compounds were selected to include very active, inactive, and moderately active antagonists. Furthermore, these compounds have the same substitution pattern on the BZD ring, ensuring that a comparison of their conformational energy profiles is meaningful and that differences in activity can be attributed to the substituents on the 3-amido aromatic ring. In addition, the analogs of the subset are all *RS*-racemates. Analysis of the torsion angle space of the ortho- and equivalent non-ortho-substituted phenyl analog was used to study the possible relationship between conformational stability and activity. Compound **51**, an *N*-methylindole analog, was included because the *N*-methyl group occupies the same area in space as does the *o*-phenyl substituent.

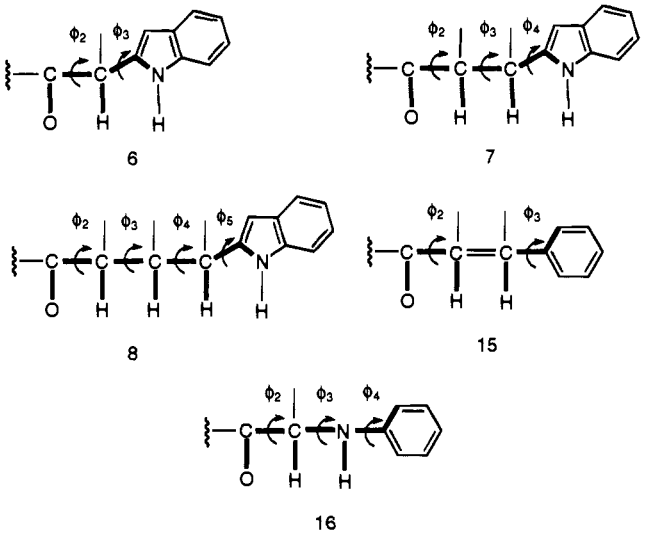


Figure 4. Structures of compounds with more than two rotatable bonds, namely **6**–**8**, **15**, and **16**. The torsion angles which were rotated are shown with thickened bonds. The BZD ring structure and ϕ_1 are not shown since these features are common to all the analogs.

The active conformation was postulated by observing which conformational energy minimum common to the active analogs, compounds **50**, **51**, **64**, and **112**, was not available to the inactive analog, compound **75**. The postulated active conformation was then assigned to all compounds. This conformation was easily examined for those compounds which have only the torsion angles ϕ_1 and ϕ_2, as in the analogs of the subset. If an analog had substituents on the phenyl or indole ring, and if there existed energetically equivalent multiple conformations of these substituents for the identical values of ϕ_1 and ϕ_2, then that overall conformation which maximized the shape similarity between the analog and compound **2**, the most active antagonist of the series, was chosen. The assumption was made at this step that compound **2**, being most active, has the preferred molecular shape for the dataset for the activity. For compounds **6**–**8**, **15**, and **16**, which have more than two torsional angles, the postulated active conformation was that conformation which statistically gave the best fit when its

shape was compared to that of the selected shape reference compound within the defined energy cutoff. In a trial 3D-MSA–QSAR, the active conformations for the more flexible analogs, with compound **2** as the reference compound, were determined using partial least squares (PLS).[26]

Once all compounds were assigned an active conformation, it was possible to test every analog as a candidate shape reference compound. The criterium for selecting the shape reference compound is, as usual, to optimize the statistical significance of the corresponding QSAR.

(3) Molecular Superposition. MSA requires that each compound in the data set be compared to the shape reference compound. Such a comparison necessitates a pairwise molecular superposition. The geometric criterion for pairwise analog molecular superposition was to place the N–C–C atoms of the amide linkage consisting of the amide nitrogen, the amide carbonyl, and the first carbon bonded to the amide carbonyl of each pair of molecules upon one another. The reason that these three atoms were chosen was not only because they are common to all analogs but also because they play a role in determining activity, i.e., they contribute to the spatial pharmacophore. The analogs which have the amide functionality replaced by a methylene linkage,[27] or an *N*-methyl amide group,[14] are much less active than their amide counterparts. All compounds have the BZD ring structure with a 3-amide fragment in common, and it was assumed that the 3-amido substituents interact with the identical region of receptor space. Thus, no other possible superpositions were considered given these constraints.

(4) Quantitative Measures of Molecular Shape. Two descriptors of relative shape similarity were considered. The common overlap steric volume, V_{ov}, between each analog, u, in the data set and the reference compound, v, is defined as,

$$V_{ov} = V_u \cap V_v \tag{2}$$

where V_u and V_v are the volumes of the isolated molecules, respectively. V_{ov} is a measure of how much a molecule shares a common spatial region with the reference compound, under the chosen superposition, and is computed in Chemlab-II.[17]

The other shape similarity descriptor considered in this investigation is the nonoverlap volume, V_{non},[17] defined as,

$$V_{non} = V_{uv} - V_v \tag{3}$$

where molecule v is the reference compound and V_{uv} is the composite volume of the superimposed pair. The nonoverlap volume measures the regions of space not shared by the molecules.

V_{ov} and V_{non} do not contain information regarding the intramolecular stability of each conformation of the pair of compounds from which they are derived. Conformational energetics has been built into the estimation of molecular shape similarity by including the loss in conformational stability of an analog (relative to the intramolecular global minimum energy conformation) corresponding to the geometric measure of molecular shape similarity. For compounds with more than two torsion angles, a trial QSAR was established by maximizing V_{ov} and minimizing V_{non} with the least loss in conformational energy, ΔE. As mentioned earlier, PLS analysis of the shape desciptors and ΔE with the dependent variable, activity, was used to find which conformation of the more flexible compounds **6–8**, **15**, and **16** would be best to use in formulating a 3D-MSA–QSAR for the entire set of analogs.

An additional method, besides systematic conformational searching, was used to explore conformations of compounds **6–8**, **15**, and **16** that were perhaps missed because of the torsion angle increment chosen in the conformational scans. The molecular similarity application in the Quanta molecular modeling package[28] was used to maximize the overlap of two structures, one structure serving as the template, the other as the working structure. The template structure is fixed. The conformation of the working structure is modified by rigid-body fit to the template and the rms difference of the fit minimized with respect to the torsion angles of the working structure. The intramolecular energies of the rigid-fit conformations were evalulated and, if within the specified ΔE, included in the 3D-MSA–QSAR analysis.

(5) Determination of Other Molecular Features. The molecular decomposition–recomposition, MDR, technique[25] was used to determine which nonshape physicochemical properties might be potential molecular descriptors in a 3D-MSA–QSAR. The MDR technique assumes that a molecule can be decomposed into a set of substructure molecules such that each set of individual substructure molecules can be analyzed, in terms of QSARs and/or molecular modeling, using the structure–activity data of the corresponding parent (whole) compounds. Clearly, the MDR assumption is predicated upon selecting specific parent molecular structure–activity data in which the only changes in chemical structure occur in the substructure of interest. The remainder of the parent molecule remains constant for the substructure congeners being analyzed. MDR analysis also implicitly assumes that physicochemical property changes in one substructure of the whole molecule do not modify the properties of any other substructure, that is, the substructures are uncoupled from one another with respect to computation of the molecular descriptors of interest.

In this particular investigation, the QSAR analyses are based upon a molecular substructure that is bonded to the remainder of the parent compound by a saturated single carbon bond. Thus, no resonance/electronic structure complications should occur in the estimation of electronic descriptors, characteristic of the parent compound, by using its substructure representative in a 3D-MSA–QSAR analysis. Also, the size and flexibility of the substituents used in the studies are sufficiently limited to minimize any long-range conformational coupling between a substructure and the remainder of the parent molecule. Hence, the MDR technique could be carried out for this set of CCK antagonists. The 3-amido substituent with the amide nitrogen possessing an *N*-methyl group was used as the substructure to determine possible molecular descriptors. All fragments were assigned the postulated active conformation.

In addition to molecular shape, the descriptors used in trial QSARs are described below.

(a) Lipophilicity Measures. (i) LOGP. The lipophilicities of the substructures were determined using the program CLOGP.[29]

(ii) π-Constants. π-Constants[30] for the substituents in the ortho, meta, and para positions of the phenyl analogs (Table 1) were determined. The 5′-substituent on the indole analogs (Table 2) was treated alternatively as corresponding to the meta or para position of the phenyl analogs, and the *N*-methyl group of the indole ring was categorized as an ortho substituent. The π-constants for these indole substituents were also evalulated as descriptors.

(iii) Scheme for Estimating a Lipophilic Contribution from a Part of the Molecule. A third type of lipophilicity descriptor was constructed by measuring the hydrophobicity of only part of a substructure.[22] This strategy was adopted to test for a receptor-specific lipophilic pocket with which only part of a ligand interacts. The sum of the atomic and/or group π-constants was used to assess the possibility of lipophilic interactions in the CCK-A receptor.

The sum of the π-constants, $\Sigma\pi$, was developed upon the basis of two assumptions. The premises were that (1) the location of a receptor lipophilic pocket can be identified by the atoms in the pharmacophoric substructure of one of the most active compounds and (2) the lipophilic character of the substructure can be estimated by summing up atomic π-constant contributions of atoms near the lipophilic pocket.

Using these two assumptions, the following procedure to estimate the relative local lipophilic effect of a substructure was devised.

1. The substructures of all compounds were aligned as defined above.

2. A sphere was generated about the geometric center of the *p*-I substituent of compound **95** (see Figure 5). The radius was varied in increments of 0.5 Å between 1.5 and 5.5 Å. This

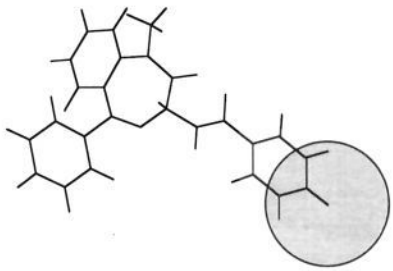


Figure 5. Spherical model used to estimate the lipophilicity term ($\Sigma\pi$ r Å). The atomic contributions are summed when the center of the atom lies within the sphere. The origin of the sphere is the geometric center of the p-I substituent of compound **95** (**95** shown). The radius is varied between 1.5 and 5.5 Å.

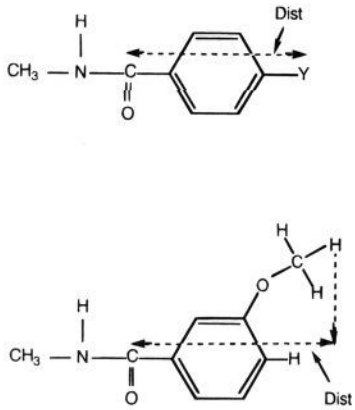


Figure 6. Distance between the carbonyl carbon and the furthermost atom corresponding to the direction of the p-phenyl substituent.

region in space was chosen because compound **95** is very active, and its high activity must come, at least in part, from the p-I atom.

3. Every atom, or group, which has at least its geometric center within the sphere was assigned its π-constant. The sum of the π-constants, over the number of atoms in the sphere, was computed for each analog. The notation for this descriptor is ($\Sigma\pi$ r Å), where r indicates the radius of the lipophilic sphere.

(b) Relative Energy, ΔE. The relative conformational stability of a compound, ΔE, defined as the difference between a particular conformation and the apparent global minimum energy conformation was considered as a descriptor.

(c) C-to-Furthermost Atom Distance, dist. The distance between the carbonyl carbon and the furthermost atom corresponding to the direction of the p-phenyl substituent was considered as a descriptor. For analogs that do not have the equivalent of a para substituent, the distance of the furthermost atom transposed to the axis corresponding to the position of a p-phenyl substituent was calculated as shown in Figure 6. This descriptor may estimate how far a compound might extend "into" a binding site.

(d) Molar Refractivity, MR. The molar refractivities of the substructures were determined using the program CLOGP.[30]

(e) Hammett Constant, σ_x. The Hammett σ-constant[30] was evalulated as a possible molecular descriptor. The σ-constants of the ortho, meta, and para substituent on the phenyl analogs and the 5′-substituent and the N-methyl group on the indole analogs were examined in the trial 3D-MSA–QSARs.

(f) Ionization Potential, IP. The ionization potential, IP, of the substructures, as calculated by MNDO, was considered as a molecular descriptor.

(g) HOMO and LUMO Energies. The *h*ighest *o*ccupied, HOMO, and *l*owest *u*noccupied, LUMO, *m*olecular *o*rbital *e*nergies, *calculated by MNDO, were considered as molecular* descriptors.

(h) Partial Atomic Charges, Q_x. The partial atomic charges were computed using MNDO. Individual charges on the amide nitrogen, the amide proton, the carbonyl carbon, and the ortho, meta, and para carbons of the phenyl ring were used in trial 3D-MSA–QSARs of the phenyl analogs. Sums of atomic charges from various atom combinations were also considered.

(i) Dipole Moment, Ut, x, y, z, and θ. The total dipole moment, Ut, and its x, y, and z components were calculated by the MNDO method. In addition, the angle, θ, between the total dipole and the z-axis (the aromatic ring of the pharmacophore, as it is oriented in this study, resides in the yz-plane) was also calculated. Each of these dipole representations was evalulated as a molecular descriptor.

(j) Indicator Variables, I_F and I_M. Indicator variables were defined for the presence of a methyl group on the endocyclic amide nitrogen and also for the fluorine atom of the 5-phenyl ring of the BZD ring. The variable I_M was given the value of 1 for the presence of methyl and 0 for hydrogen. The variable I_F was given the value of 1 for the presence of fluorine and 0 if there was a hydrogen in its place. The reason *these descriptors were defined was that different activities are* realized for compounds which had identical 3-amido substituents but which differ in the substitution pattern on the BZD ring.

(6) Construction of Trial 3D-MSA–QSARs. A useful aspect to MSA in generating a QSAR is that not only are all combinations of molecular descriptor sets considered in optimizing the statistical significance of the QSAR but also that the QSAR is optimized by cycling through steps 2–6 of the MSA process (see Figure 2).

Multidimensional linear regression analysis[31] was mainly used to optimize the 3D-MSA–QSARs. However, PLS regression was also used in the optimization procedures. PLS was particularly useful in selecting the QSAR conformations of highly flexible compounds. PLS also reinforced the choice of shape reference compounds for use in building trial QSARs.

Trial 3D-MSA–QSARs were generated using the multidimensional linear regression analysis facilities in the SAS software package.[32] All combinations of molecular descriptors were analyzed one at time and then pairs, etc., up to a limit where further significance of the regression equation was not realized. The partial-F test[33] was applied to each regression equation to determine whether the addition of any particular descriptor to the model significantly contributed to the prediction of the activity. In each case, the cross-correlation descriptor matrix was examined to eliminate trial QSARs in which pairs of descriptors had cross-correlation coefficients greater than 0.50. Analogs were considered as outliers when the difference in predicted and observed activities exceeded 25% of the activity range as measured by $-\log(IC_{50})$. Regression analyses were performed minus the outliers and examined to see if significant improvements in the regression equations resulted. 3D-MSA–QSARs were constructed for the entire set of 53 analogs. However, subsets consisting of only the phenyl compounds, only the p-phenyl-substituted analogs, only the indole compounds, the phenyl and indole compounds combined, and only the S-isomers were also examined. Cross-validation ($n = 1$) was also performed to help assess whether the models, based on the compounds included, could be applied to related analog series.[34]

Results

a. Conformational Analysis and Selection of the Active Conformation. A comparison of the conformational profiles of the A and B crystal structures of compound **2** without minimization with the profiles of the completely optimized structures of the two crystals reveals that with geometry optimization the conformational energy profiles of the A and B structures become virtually identical. Hence, the structures derived from *the A crystal were arbitrarily used throughout the* balance of the analyses.

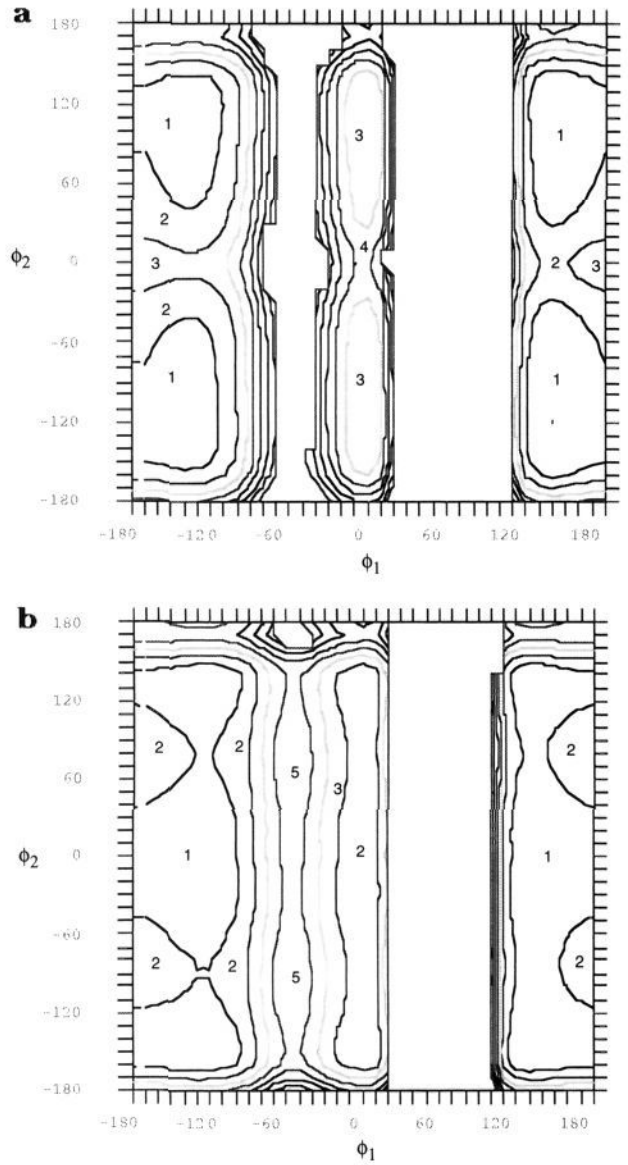


Figure 7. (a) Conformational energy map of the completely optimized crystal A structure of **2** with MNDO charges using MM2 nonbonded potentials. (b) Conformational energy map of the completely optimized crystal A structure of **2** with MNDO charges using the Z set of nonbonded potentials.

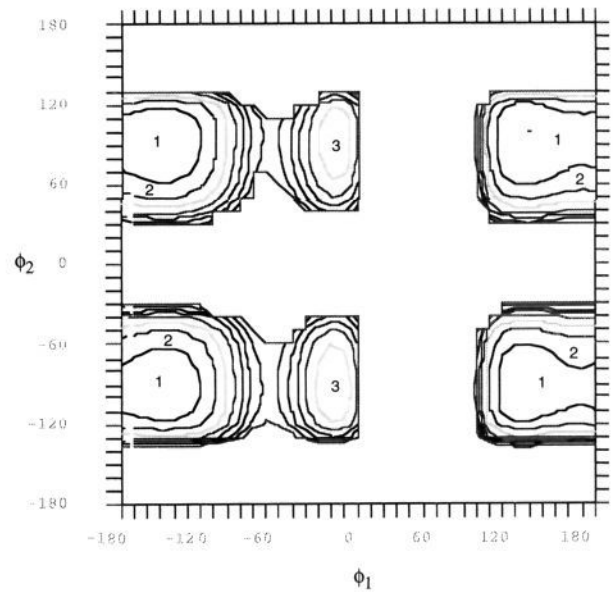


Figure 8. Conformational energy map of compound **75** using MM2 nonbonded potentials.

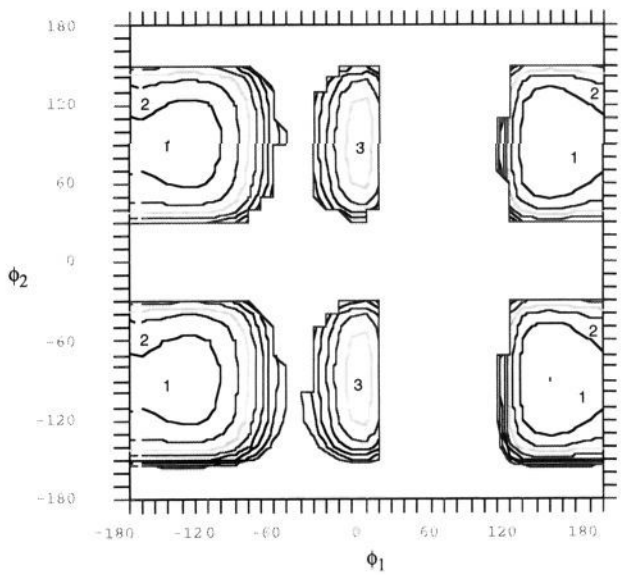


Figure 9. Conformational energy map of compound **64** using MM2 nonbonded potentials.

The conformational profiles of the phenyl and indole analogs of Tables 1 and 2 can be characterized by the conformational energy (ϕ_1, ϕ_2) maps of the test set, *compounds* **50**, **75**, **64**, **112**, *and* **51**, *shown in Figures* 7a–11, respectively. Note that the conformational energy maps of compound **50** are the same as its *S*-isomer, compound **2**, since all analyses were performed with the structures of the *S*-isomer. Each figure shows the map generated using MM2 nonbonded potentials. At first inspection, the energy maps look very similar. All maps have common conformational local energy minima at $\phi_1 \cong 130°, -130°$, $\phi_2 \cong 90°, -90°$. A comparison of the map of the *o*-Cl analog, compound **75**, with those of other corresponding analogs of the test set shows that the *o*-Cl analog displays the greatest reduction in conformational flexibility with respect to ϕ_2. This finding suggested the conformer state $\phi_1 = -130°, \phi_2 = 130°$ as being a possible active conformation candidate. A map (Figure 12) showing the subtraction of conformer states of compound **112** by the equivalent conformer states of compound **75** is an example of one of the subtraction maps created in an effort to elucidate *the active conformation. This comparison map exhibits* the conformer state $\phi_1 = -130°, \phi_2 = 130°$ for compound **112** as being a conformation that is at least 2 kcal/mol more stable than the corresponding conformer state of compound **75**. Similar findings were found for the other active analogs of the subset. The conformational energy map of the *o*-I-phenyl analog, compound **123**, (Figure 13) displays another example of the effect of an *o*-halo substituent on the conformational profile of the phenyl analogs. The upper left quadrant of the energy map shows that the instability of the $\phi_1 = -130°, \phi_2 = 130°$ conformer state is more pronounced than that of the *o*-Cl analog. ΔE, between the apparent global minimum and the conformer state $\phi_1 = -130°, \phi_2 = 130°$, for each compound is given in Table 5, and it appears that this conformation is unstable for the *o*-halo-substituted

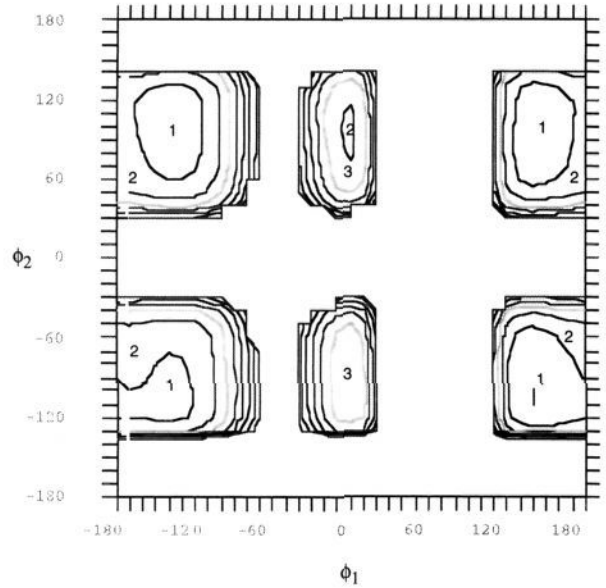


Figure 10. Conformational energy map of compound **112** using MM2 nonbonded potentials.

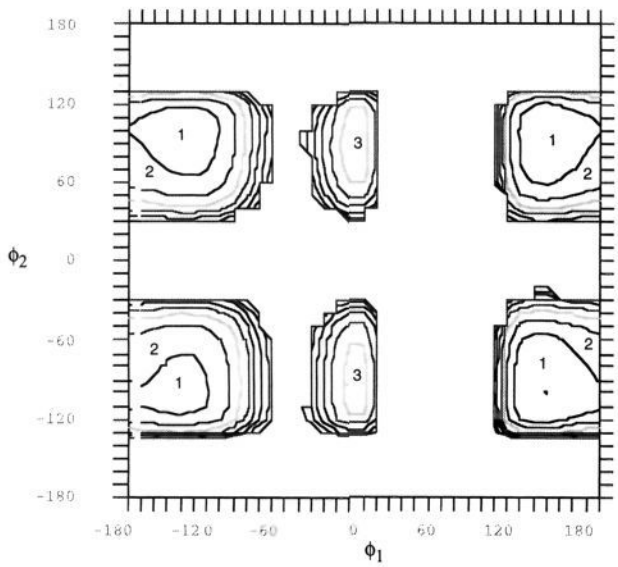


Figure 11. Conformational energy map of compound **51** using MM2 nonbonded potentials.

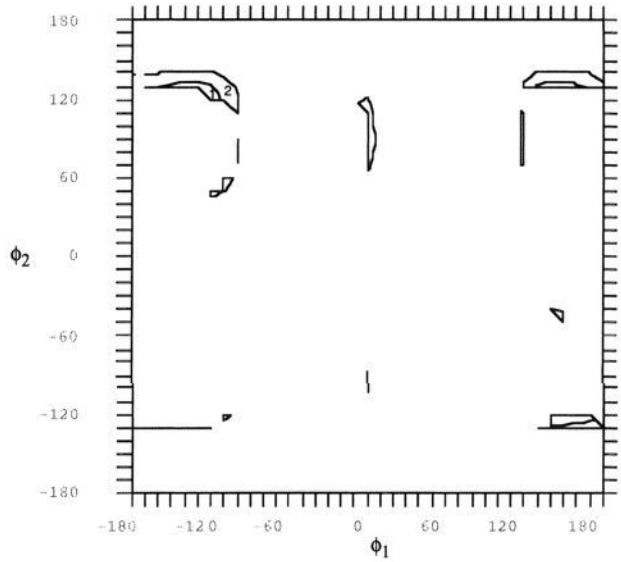


Figure 12. Subtraction energy map showing subtraction of conformer states of compound **112** minus compound **75** using a 2 kcal/mol energy tolerance.

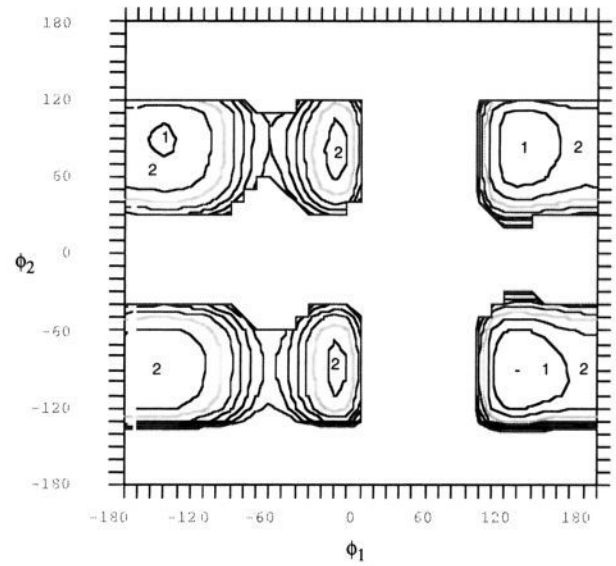


Figure 13. Conformational energy map of compound **108** using MM2 nonbonded potentials.

phenyl analogs but is near the respective apparent global minimum for the other compounds of the data set.

The sensitivity of conformational stability and, ultimately, the fidelity of the corresponding 3D-MSA–QSAR, to force field representation, was also explored. The nonbonded potential set Z[23] which is a "soft" set of potentials, relative to the MM2 potentials, was also used in the conformational analysis. An example of the effect of using the Z set of potentials is given in Figure 7b which shows the energy map for compound **50**. The conformational energy maps obtained with the two different force fields are quite similar. However, there are two noticeable differences for all analogs of the subset: (1) small shifts in the precise location of minima and (2) greater allowed conformational flexibility for maps generated using the Z set of potentials (compare part a of Figure 7 with part b). ΔE values, using the Z set of potentials between the conformer state $\phi_1 = -130°$, $\phi_2 = 130°$ and the apparent global minimum, are given in Table 5. Also given in Table 5 are the ΔE values for the neighboring conformer state $\phi_1 = -130°$, $\phi_2 = 140°$. It is for this conformer state that larger differences in ΔE between the active analogs and

Table 5. Energy Difference, ΔE, between the Apparent Global Minimum and the Conformer State (ϕ_1, ϕ_2) for both the Nonbonded Steric Parameters of the MMFF Set and the Z Set

	ΔE, MMFF	ΔE, Z set	
compd	$\phi_1 = -130°$, $\phi_2 = 130°$	$\phi_1 = -130°$, $\phi_2 = 130°$	$\phi_1 = -130°$, $\phi_2 = 140°$
2	0.55	0.37	0.41
51	0.31	1.16	1.89
64	1.25	0.23	0.24
70	4.22	1.11	7.39
112	0.85	0.05	0.02
108	*a*	3.20	*a*

[a] The energy at these conformations of compound **108** is greater than the cutoff energy of 8 kcal/mol used to create the conformational energy maps.

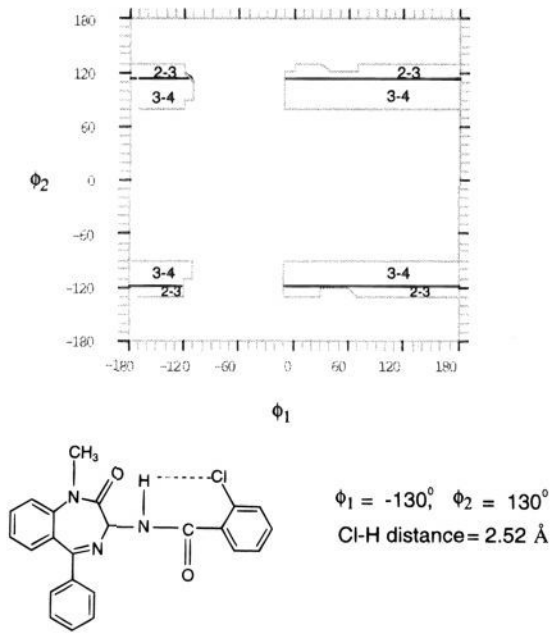


Figure 14. Conformational distance scan showing distances between the *o*-chlorine atom and the amide hydrogen. Areas marked 2–3 indicate areas where the distance between these two atoms is in the range of 2–3 Å. Areas marked 3–4 indicate areas where the distance between these two atoms is in the range of 3–4 Å. The Cl–H distance as shown in figure is 2.52 Å when compound **75** is in the active conformation.

(inactive) compound **75** are seen. If one considers the MM2 potential set as too "hard" and the Z set as too "soft", then the "true" force field may be between the two. Correspondingly, the postulated active conformation may reside in the $\phi_1 = -130°$, $\phi_2 = 130-140°$ region of torsion angle space.

The postulated active conformation was further explored. An attempt to rationalize the difference in ΔE for the analogs of the test set, especially compounds **75** and **112** which have in common an ortho substituent, was undertaken by looking at intramolecular steric interactions. The distance between the ortho substituent and selected atoms with which possible unfavorable interactions might occur was calculated as a function of conformation in the form of conformational distance maps. The atoms which might interact with the ortho substituent include the amide nitrogen and hydrogen and the BZD endocyclic carbonyl oxygen. One of the distance scans of the inactive *o*-Cl analog, compound **75**, indicates that the conformer state ($\phi_1 = -130°$, $\phi_2 = 130°$) is present in that area of space where unfavorable interactions between the *o*-Cl atom and the amide hydrogen may occur. Figure 14 is an example of one of the distance scans and shows the distance of the *o*-Cl atom and the amide hydrogen as a function of ϕ_1 and ϕ_2. The distance between the Cl atom and the amide hydrogen is 2.52 Å when compound **75** is in the postulated active conformation. This distance is sterically unfavorable. The distance scans of the *o*-NH_2, *p*-Cl analog, compound **112**, show that the conformer state ($\phi_1 = -130°$, $\phi_2 = 130°$) is present in that area of space where favorable interactions may occur. The distance between the nearest amino hydrogen and the amide nitrogen is 2.28 Å and the distance between this hydrogen and the BZD ring carbonyl oxygen is 2.39 Å when compound **112** is in the postulated active conformation. Also, the distance between the amino nitrogen and the amide proton is 2.66 Å. Although these distances are longer than a typical hydrogen bond, a significant attractive electrostatic interaction occurs between the atoms. The distance between the nearest amino hydrogen and the amide proton is 2.21 Å and from energy calculations does not present a bad interaction. This analysis explains why compound **112** can adopt the postulated active conformation despite possessing an ortho substituent.

Table 6. Torsion Angles for the Active Conformation of Compounds with More than Two Rotatable Angles

compd	ϕ_1 (deg)	ϕ_2 (deg)	ϕ_3 (deg)	ϕ_4 (deg)	ϕ_5 (deg)
6	−130	−110	−130		
7	−130	−150	−90	0	
8	−130	−120	−30	−60	−150
15	−130	−60	20		
16	−130	0	180	−90	

Intramolecular steric interactions were also examined for compound **51** in the postulated active conformation. The closest distances between the nearest hydrogen of the *N*-methyl group, which occupies the area of space equivalent to the *o*-phenyl substituent, and the amide proton, the endocyclic oxygen, and the amide nitrogen are 2.32, 2.60, and 2.72 Å, respectively. These are acceptable steric distances.

Application of the *LBA–LCS* principle led to postulating that the active conformation of compounds, with two torsional angles, ϕ_1 and ϕ_2, is the comformer state $\phi_1 = -130°$, $\phi_2 = 130°$.

Conformational scans were performed for all the remaining compounds in the data set and the ΔE values calculated. For compounds **6**–**8**, **15**, and **16**, which have more than two torsion angles, the postulated active conformation was determined using a PLS analysis of the independent variables V_{ov}, V_{non}, and ΔE versus activity. For these compounds, the molecular shapes of all the conformer states within ΔE were compared to the most active analog, compound **2**, which had its torsion angles set at the values corresponding to the proposed active conformation. All conformations within a ΔE of 8 kcal/mol were examined, except for compound **8**. Because of the large number of conformations available to this compound, only those conformers within a ΔE of 5 kcal/mol were used. The conformations produced by the molecular similarity application in Quanta were also tested. The conformation for each of the flexible analogs which proved most statistically significant with respect to similarity to compound **2** was chosen as the active conformation for the 3D-MSA–QSAR of the entire dataset. The values of the torsion angles of the proposed active conformations for these five compounds are listed in Table 6.

b. Shape Reference Compound. All compounds listed in Tables 1–3, in the proposed active conformation, were considered as candidates for the shape reference compound. The results of directly regressing common overlap volume, V_o, and nonoverlap volume, V_{non}, against $-\log(IC_{50})$ using different shape reference candidates for the entire set of analogs of Tables 1–3 are reported in Table 7. Some compounds have identical 3-amido fragments, differing only in the substituents on the BZD ring. Only one of these fragments was used to calculate the statistical values. Wherever possible, the compounds which had similar substituents on the

Table 7. Importance of the Shape Reference Compound on the Correlation of the Common Overlap Steric Volume, V_{ov}, and Nonoverlap Volume, V_{non}, with $-\log(IC_{50})$ by R^2 Comparison[a]

	R^2			R^2	
compd	V_{ov}	V_{non}	compd	V_{ov}	V_{non}
5	0.00	0.08	**72**	0.09	0.19
6	0.03	0.04	**73**	0.11	0.17
7	0.00	0.19	**76**	0.00	0.06
8	0.03	0.05	**77**	0.00	0.08
11	0.21	0.34	**78**	0.00	0.08
13	0.14	0.23	**79**	0.00	0.07
15	0.05	0.21	**81**	0.11	0.18
16	0.02	0.15	**87**	0.05	0.16
17	0.13	0.19	**88**	0.02	0.08
23	0.11	0.15	**90**	0.06	0.00
24	0.14	0.23	**91**	0.00	0.06
26	0.00	0.05	**95**	0.10	0.22
47	0.29	0.42	**96**	0.05	0.00
48	0.21	0.41	**99**	0.01	0.05
50	0.34	0.43	**101**	0.01	0.18
51	0.28	0.42	**102**	0.00	0.08
53	0.19	0.40	**103**	0.02	0.09
54	0.28	0.42	**104**	0.06	0.01
55	0.17	0.40	**105**	0.07	0.13
61	0.18	0.19	**108**	0.05	0.00
64	0.13	0.17	**112**	0.12	0.17
66	0.00	0.06	**117**	0.13	0.19
70	0.00	0.05	mutant	0.18	0.38

[a] Terms are significant at the 0.05 level except for those which have an $R^2 < 0.06$.

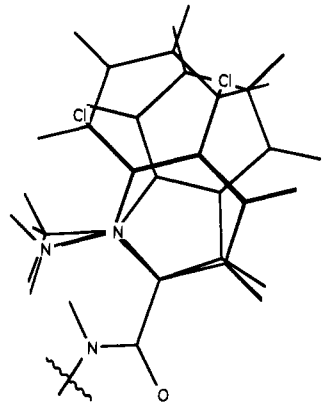


Figure 15. Shape reference compound "mutant" composed of the overlapped structures of **24**, **51**, **81**, and **112**. Only the 3-amido substituent is shown. The amide nitrogen, the amide carbonyl, and the attached aromatic carbon were used to overlap these structures which are in the active conformation.

BZD ring were picked as the representative for that particular fragment type. PLS was also applied to the molecular shape measures and verified the ranking of the compounds in terms of significance as the reference compound. It is clear from Table 7 that selection of the shape reference compound can markedly alter the shape descriptor–inhibition potency correlation. In trial 3D-MSA–QSAR analyses of all the descriptors, a few equally significant regression equations emerge which depend on the shape parameters measured using a select few reference compounds. In other words, no single analog is the best shape reference compound. For this reason, a shape reference "compound" which is made up of the overlapped structures of these select few analogs, namely, **24**, **51**, **81**, and **112** (Figure 15), and is termed a "mutant" was also included in the MSA.

c. 3D-MSA–QSARs. The optimum 3D-MSA–QSAR, in terms of the correlation coefficient and the F-statistic measure, for all the phenyl analogs found in Table 1 is

Table 8. SAR Table for All Phenyl Analogs Found in Table 1

						$-\log(IC_{50})$		
compd	$V_{ov,m}$	$V_{non,m}$	ΔE	I_M	I_F	obs	pred[a]	residual
12	111.72	2.08	0.78	0	1	8.22	7.81	0.41
61	111.67	2.48	0.78	0	1	6.96	7.78	−0.82
62	111.72	2.08	0.78	1	1	8.64	8.77	−0.13
64	111.72	2.08	0.78	1	0	8.08	8.02	0.06
65	111.72	2.08	0.78	0	0	7.39	7.06	0.33
66	96.05	2.07	0.79	0	1	6.82	6.84	−0.02
67	107.73	5.36	4.12	0	1	5.25	5.46	−0.21
69	107.73	5.36	4.12	0	0	5.12	4.72	0.40
72	113.43	12.84	0.78	0	1	7.82	7.33	0.49
73	109.65	6.57	0.78	0	1	7.68	7.44	0.24
75	107.73	5.36	4.12	1	0	5.47	5.67	−0.20
76	111.81	2.09	0.79	0	0	7.09	7.06	0.03
77	104.07	18.92	0.78	0	1	7.02	6.42	0.60
78	111.42	34.73	0.78	0	1	6.57	6.01	0.56
79	112.54	35.02	0.76	0	1	5.77	6.08	−0.31
81	127.22	2.09	0.78	0	0	7.54	8.02	−0.48
87	115.53	20.24	0.87	0	0	6.66	6.26	0.40
88	99.52	2.08	0.77	0	0	6.32	6.32	0.00
91	117.00	29.02	0.70	0	0	5.82	5.97	−0.14
92	113.43	12.84	0.78	0	0	6.62	6.58	0.04
99	112.15	50.72	0.78	0	0	4.00	4.44	−0.44
101	124.13	43.56	0.85	0	0	5.85	5.53	0.32
102	116.27	13.23	0.77	0	0	6.30	6.74	−0.44
103	99.07	5.71	0.74	0	0	5.75	6.11	−0.36
105	106.68	2.53	0.76	0	0	6.14	6.74	−0.60
112	122.70	2.19	0.50	1	0	9.05	8.85	0.20
117	115.33	4.49	0.79	1	0	9.00	8.11	0.89
118	113.43	12.84	0.78	1	1	7.70	8.29	−0.59
122	124.13	43.56	0.85	1	1	7.00	7.23	−0.23

[a] Values of $-\log(IC_{50})$ predicted by eq 4.

$$-\log(IC_{50}) = 0.062\ (\pm 0.014)V_{ov,m} - 0.054\ (\pm 0.007)V_{non,m} - 0.575\ (\pm 0.087)\Delta E + 0.958\ (\pm 0.216)I_M + 0.746\ (\pm 0.181)I_F + 0.732\ (\pm 1.615) \quad (4)$$

$$N = 29, R = 0.94, F = 33.3, \mathrm{SD} = 0.46$$

where $V_{ov,m}$ is the common overlap steric volume with the mutant as the shape reference compound, $V_{non,m}$ is the nonoverlap volume with the mutant as the shape reference compound, and the other descriptors are as defined above. N is the number of compounds, R is the correlation coefficient, SD is the standard deviation of fit, and the statistical ratio F is used to determine the probability of a linear relationship between the dependent variable and the set of independent variables and the constant. The values in parentheses define the 95% confidence limits. Predicted and residual $-\log(IC_{50})$ values and the descriptor values are given in Table 8.

A large number of the phenyl analogs are para-substituted. Thus, an individual 3D-MSA–QSAR was generated for these compounds in order to explore the interaction of the CCK-A receptor with this particular substituent. The optimum regression equation for all phenyl analogs with only a para substituent is

$$-\log(IC_{50}) = -0.032\ (\pm 0.009)V_{non,m} + 0.793\ (\pm 0.344)(\Sigma\pi\ 1.5\ \text{Å})^2 + 0.917\ (\pm 0.343)I_M + 0.714\ (\pm 0.265)I_F + 6.469\ (\pm 0.296) \quad (5)$$

$$N = 20, R = 0.89, F = 14.1, \mathrm{SD} = 0.59$$

where $(\Sigma\pi\ 1.5\ \text{Å})^2$ is the square term of the lipophilic sphere descriptor for a radius of 1.5 Å. The cross-validated R^2 is 0.18 less than the parent R^2. Predicted and residuals $-\log(IC_{50})$ values and the descriptor

Table 9. SAR Table for All Para-Substituted Phenyl Analogs Found in Table 1

compd	$V_{non,m}$	($\Sigma\pi$ 1.5 Å)^2	I_M	I_F	−log(IC$_{50}$) obs	−log(IC$_{50}$) pred[a]	residual
12	2.08	0.88	0	1	8.22	7.82	0.40
61	2.48	0.00	0	1	6.96	7.10	−0.14
62	2.08	0.88	1	1	8.64	8.73	−0.09
64	2.08	0.88	1	0	8.08	8.02	0.06
65	2.08	0.88	0	0	7.39	7.10	0.29
66	2.07	0.05	0	1	6.82	7.16	−0.34
72	12.84	1.21	0	1	7.82	7.73	0.09
73	6.57	0.53	0	1	7.68	7.39	0.29
77	18.92	0.18	0	1	7.02	6.71	0.30
78	34.73	0.29	0	1	6.57	6.30	0.27
87	20.24	0.00	0	0	6.66	5.82	0.84
88	2.08	0.14	0	0	6.32	6.51	−0.19
92	12.84	1.21	0	0	6.62	7.02	−0.40
99	50.72	0.24	0	0	4.00	5.03	−1.03
101	43.56	0.18	0	0	5.85	5.21	0.64
103	5.71	0.19	0	0	5.75	6.44	−0.69
105	2.53	0.12	0	0	6.14	6.48	−0.34
117	4.49	1.19	1	0	9.00	8.18	0.82
118	12.84	1.21	1	1	7.70	8.65	−0.95
122	43.56	0.18	1	1	7.00	6.84	0.16

[a] Values of −log(IC$_{50}$) predicted by eq 5.

values are given in Table 9. As the lipophilicity of the para substituent increases so does the binding affinity. However, there is a trade-off between size and lipophilicity because an increase in V_{non} is detrimental to activity.

The optimal 3D-MSA–QSAR for the indole analogs of Table 2 is

$$-\log(IC_{50}) = 3.825\ (\pm0.253)(\Sigma\pi\ 2.0\ \text{Å}) - 2.413\ (\pm0.153)(\Sigma\pi\ 2.0\ \text{Å})^2 + 0.484\ (\pm0.093)I_M + 6.932\ (\pm0.091) \quad (6)$$

$$N = 12, R = 0.99, F = 160.4, \text{SD} = 0.13$$

Predicted and residual −log(IC$_{50}$) values and the descriptor values are given in Table 10. Shape descriptors are not found in this regression of highly congeneric analogs. Activity can be predicted using only a lipophilic descriptor and an indicator variable. There is a quadratic dependence with the ($\Sigma\pi$ 2.0 Å) descriptor indicating there may be an optimal lipophilicity for these indole analogs. The optimal ($\Sigma\pi$ 2.0 Å) value is calculated to be 0.79 which corresponds to having a hydrogen at the 5′-position for the indole analogs.

The optimum 3D-MSA–QSAR for the combined set of phenyl and indole analogs of Tables 1 and 2 is

$$-\log(IC_{50}) = 0.044\ (\pm0.008)V_{ov,m} - 0.049\ (\pm0.005)V_{non,m} - 0.544\ (\pm0.080)\Delta E + 0.503\ (\pm0.194)(\Sigma\pi\ 1.5\ \text{Å})^2 + 0.853\ (\pm0.174)I_M + 0.494\ (\pm0.137)I_F + 2.566\ (\pm0.911) \quad (7)$$

$$N = 41, R = 0.95, F = 50.5, \text{SD} = 0.43$$

Predicted and residual −log(IC$_{50}$) values and the descriptor values are given in Table 11. Molecular shape plays an important role as the structural diversity of the data set increases. An increase in lipophilicity also contributes to an increase in binding affinity.

A statistically comparable equation is found by replacing the descriptor ΔE by the Hammett σ-constant for the ortho substituent of the phenyl analogs and for the N-substituent of the indole analogs. The N-methyl/hydrogen of the indole analogs was considered as an ortho substituent since the N-methyl substituent occupies the same space as the o-phenyl substituent. The equation is

$$-\log(IC_{50}) = 0.0398\ (\pm0.009)V_{ov,m} - 0.049\ (\pm0.006)V_{non,m} - 1.431\ (\pm0.247)\sigma_o + 0.496\ (\pm0.213)(\Sigma\pi\ 1.5\ \text{Å})^2 + 0.814\ (\pm0.189)I_M + 0.533\ (\pm0.149)I_F + 2.514\ (\pm1.001) \quad (8)$$

$$N = 41, R = 0.94, F = 41.4, \text{SD} = 0.47$$

Table 10. SAR Table for All Indole Analogs Found in Table 2

compd	($\Sigma\pi$ 2.0 Å)	($\Sigma\pi$ 2.0 Å)^2	I_M	−log(IC$_{50}$) obs	−log(IC$_{50}$) pred[a]	residual
9	0.71	0.50	0	8.34	8.43	−0.10
43	0.71	0.50	0	8.68	8.43	0.25
44	0.71	0.50	1	8.85	8.92	−0.07
45	0.71	0.50	1	8.86	8.92	−0.06
47	0.97	0.94	0	8.30	8.37	−0.07
48	1.54	2.37	0	7.15	7.10	0.05
50	0.71	0.50	1	9.10	8.92	0.18
51	0.71	0.50	1	8.85	8.92	−0.07
53	1.58	2.50	0	6.92	6.95	0.03
54	0.05	0.00	0	7.15	7.12	0.03
55	0.05	0.00	0	6.70	6.73	−0.03
56	0.71	0.50	0	8.33	8.43	−0.10

[a] Values of −log(IC$_{50}$) predicted by eq 6.

The descriptor values are included in Table 11. The correlation coefficient between ΔE and σ_o is 0.99. The σ_o values for the amino, methyl, and hydrogen substituents are small and negative, whereas this value is much larger and positive for the halogens. Both ΔE and σ_o are descriptors which correlate the negative influence of an o-halo substituent. Size does not appear to be a factor since the chlorine atom occupies the same volume as a methyl group. Furthermore as halogen size increases, activity does not decrease proportionately. Lipophilicity of the ortho substituent also does not correlate with activity. However, the variety of the ortho substituent is quite limited.

The optimal 3D-MSA–QSAR for all the *RS*-compounds of Tables 1–3 is

$$-\log(IC_{50}) = 0.038\ (\pm0.008)V_{ov,m} - 0.035\ (\pm0.006)V_{non,m} - 0.207\ (\pm0.066)\Delta E + 0.659\ (\pm0.252)(\Sigma\pi\ 1.5\ \text{Å})^2 + 0.764\ (\pm0.223)I_M + 0.609\ (\pm0.167)I_F + 2.671\ (\pm0.937) \quad (9)$$

$$N = 53, R = 0.89, F = 31.2, \text{SD} = 0.59$$

Predicted and residual −log(IC$_{50}$) values and the descriptor values are given in Table 12. The residual of compound **8** falls outside the 25% activity range for this set of analogs. The removal of this antagonist from the analysis produces a small improvement in eq 9. Interestingly, the coefficient of the descriptor ΔE shows a marked decrease when the compounds of Table 3 are added to the phenyl and indole analogs. The value of this coefficient is basically consistent in the 3D-MSA–QSARs of the subset containing the phenyl analogs and the combined subset of phenyl and indole analogs. In addition, the residuals of the o-chloro analogs increase significantly in eq 9 when compared to the residuals for these analogs from eqs 4 and 7. The analogs with more

Table 11. SAR Table for All Phenyl and Indole Analogs Found in Tables 1 and 2

compd	$V_{ov,m}$	$V_{non,m}$	ΔE	$(\Sigma\pi$ 1.5 Å$)^2$	I_M	I_F	σ_o[a]	−log(IC50) obs	−log(IC50) pred[b]	residual
9	120.85	2.17	0.37	0.24	0	0	−0.06	8.34	7.65	0.69
12	111.72	2.08	0.78	0.88	0	1	−0.06	8.22	7.85	0.37
43	120.85	2.17	0.37	0.24	0	1	−0.06	8.68	8.15	0.53
44	120.85	2.17	0.37	0.24	1	1	−0.06	8.85	9.00	−0.15
45	138.87	2.17	1.30	0.24	1	1	−0.12	8.86	9.28	−0.42
47	121.62	5.27	0.37	0.55	0	1	−0.06	8.30	8.19	0.11
48	122.03	17.13	0.37	0.14	0	1	−0.06	7.15	7.41	−0.26
50	120.85	2.17	0.37	0.24	1	0	−0.06	9.10	8.51	0.59
51	138.87	2.17	1.30	0.24	1	0	−0.12	8.85	8.79	0.06
53	108.57	23.20	0.37	0.07	0	1	−0.06	6.92	7.06	−0.14
54	120.60	8.91	0.37	0.03	0	1	−0.06	7.15	7.70	−0.55
55	120.76	27.36	0.37	0.12	0	1	−0.06	6.70	6.85	−0.15
56	138.87	2.17	1.30	0.24	0	1	−0.12	8.33	8.43	−0.10
61	111.67	2.48	0.78	0.00	0	1	−0.06	6.96	7.39	−0.43
62	111.72	2.08	0.78	0.88	1	1	−0.06	8.64	8.71	−0.07
64	111.72	2.08	0.78	0.88	1	0	−0.06	8.08	8.21	−0.13
65	111.72	2.08	0.78	0.88	0	0	−0.06	7.39	7.36	0.03
66	96.05	2.07	0.79	0.05	0	1	−0.06	6.82	6.74	0.07
67	107.73	5.36	4.12	0.05	0	1	1.20	5.25	5.28	−0.03
69	107.73	5.36	4.12	0.05	0	0	1.20	5.12	4.79	0.33
72	113.43	12.84	0.78	1.21	0	1	−0.06	7.82	7.56	0.26
73	109.65	6.57	0.78	0.53	0	1	−0.06	7.68	7.37	0.31
75	107.73	5.36	4.12	0.05	1	0	1.20	5.47	5.64	−0.03
76	111.81	2.09	0.79	0.05	0	0	−0.06	7.09	6.94	0.15
77	104.07	18.92	0.78	0.18	0	1	−0.06	7.02	6.34	0.68
78	111.42	34.73	0.78	0.29	0	1	−0.06	6.57	5.94	0.63
79	112.54	35.02	0.76	0.18	0	1	−0.06	5.77	5.92	−0.15
81	127.22	2.09	0.78	0.88	0	0	−0.06	7.54	8.04	−0.50
87	115.53	20.24	0.87	0.00	0	0	−0.06	6.66	6.14	0.52
88	99.52	2.08	0.77	0.14	0	0	−0.06	6.32	6.46	−0.14
91	117.00	29.02	0.70	0.05	0	0	−0.06	5.82	5.89	−0.07
92	113.43	12.84	0.78	1.21	0	0	−0.06	6.62	7.07	−0.45
99	112.15	50.72	0.78	0.24	0	0	−0.06	4.00	4.67	−0.66
101	124.13	43.56	0.85	0.18	0	0	−0.06	5.85	5.47	0.38
102	116.27	13.23	0.77	0.05	0	0	−0.06	6.30	6.60	−0.30
103	99.07	5.71	0.74	0.19	0	0	−0.06	5.75	6.30	−0.56
105	106.68	2.53	0.76	0.12	0	0	−0.06	6.14	6.74	−0.60
112	122.70	2.19	0.50	0.88	1	0	−0.17	9.05	8.84	0.21
117	115.33	4.49	0.79	1.19	1	0	−0.06	9.00	8.40	0.60
118	113.43	12.84	0.78	1.21	1	1	−0.06	7.70	8.42	−0.72
122	124.13	43.56	0.85	0.18	1	1	−0.06	7.00	6.81	0.19

[a] Hammet σ-constant used in derivation of eq 8. [b] Values of −log(IC50) predicted by eq 7.

than two torsion angles, as a whole, must expend a relatively larger amount of energy to adopt the proposed active conformation than the other antagonists and, therefore, are influencing the significance of ΔE.

A 3D-MSA–QSAR consisting of all the *RS*-compounds minus compounds **8**, **15**, and **16** is found to be the optimal regression equation that maintains the statistical measures of fit found in eq 7. The optimal equation is

$$-\log(IC_{50}) = 0.042\ (\pm 0.006)V_{ov,m} - 0.045\ (\pm 0.005)V_{non,m} - 0.364\ (\pm 0.061)\Delta E + 0.560\ (\pm 0.201)(\Sigma\pi\ 1.5\ \text{Å})^2 + 0.788\ (\pm 0.177)I_M + 0.487\ (\pm 0.138)I_F + 2.483\ (\pm 0.745) \quad (10)$$

$$N = 50,\ R = 0.93,\ F = 49.7,\ \text{SD} = 0.47$$

The regression equation with $V_{ov,50}$, the common overlap steric volume descriptor with compound **50** as the shape reference compound, in place of $V_{ov,m}$ has the following statistics of fit: $N = 50$, $R = 0.93$, $F = 44.0$, SD = 0.49. The correlation coefficient between $V_{ov,m}$ and $V_{ov,50}$ is 0.75. A plot of predicted versus observed values of activity from the data in eq 10 is given in Figure 16.

A closer look as to why the three compounds **8**, **15**, and **16** are problematic was undertaken. When searching for the active conformation of compound **8**, only those conformations within 5 kcal/mol of the apparent global minimum were considered because of the large number of conformations available to this compound. This might be too low of an energy cutoff, and the active conformation for this compound may reside more than 5 kcal/mol above the global minimum.

There is no obvious explanation why the removal of compounds **15** and **16** increases the quality of the regression equation. Perhaps the method of energy calculation was not accurate. A smaller ΔE value would certainly decrease the residuals for these analogs in eq 9.

An analysis of the compound subset consisting of all the *S*-isomers found in Tables 1–3 minus compounds **70** and **96**, which are outliers, yields the following 3D-MSA–QSAR

Table 12. SAR Table for All Analogs Found in Tables 1–3

compd	$V_{ov,m}$	$V_{non,m}$	ΔE	$(\Sigma\pi$ 1.5 Å$)^2$	I_M	I_F	−log(IC50) obs	−log(IC50) pred[a]	residual
5	108.57	14.60	0.62	0.05	0	0	5.96	6.24	−0.28
6	111.94	29.99	2.30	0.05	0	0	6.00	5.52	0.48
7	121.23	39.05	4.20	0.15	0	0	5.32	5.19	0.13
8	106.60	71.10	4.91	0.05	0	0	4.70	3.28	1.42
9	120.85	2.17	0.37	0.24	0	0	8.34	7.32	1.01
11	131.95	13.67	0.66	0.24	0	1	7.66	7.89	−0.23
12	111.72	2.08	0.78	0.88	0	1	8.22	7.92	0.30
13	105.27	21.14	0.34	0.24	0	1	6.52	6.67	−0.15
15	114.20	9.70	2.50	0.24	0	1	7.96	6.97	0.99
16	109.40	17.63	4.20	0.13	0	1	7.05	6.08	0.97
17	117.00	2.13	0.21	0.24	0	1	8.05	7.82	0.23
23	82.14	2.24	0.00	0.00	0	1	6.96	6.36	0.60
24	134.49	2.22	0.86	0.24	0	0	7.70	7.74	−0.04
26	121.60	14.27	5.46	0.05	0	0	5.80	5.75	0.05
43	120.85	2.17	0.37	0.24	0	1	8.68	7.93	0.75
44	120.85	2.17	0.37	0.24	1	1	8.85	8.70	0.16
45	138.87	2.17	1.30	0.24	1	1	8.86	9.20	−0.34
47	121.62	5.27	0.37	0.55	0	1	8.30	8.06	0.24
48	122.03	17.13	0.37	0.14	0	1	7.15	7.38	−0.23
50	120.85	2.17	0.37	0.24	1	0	9.10	8.09	1.01
51	138.87	2.17	1.30	0.24	1	0	8.85	8.56	0.26
53	108.57	23.20	0.37	0.07	0	1	6.92	7.11	−0.19
54	120.60	8.91	0.37	0.03	0	1	7.15	7.55	−0.40
55	120.76	27.36	0.37	0.12	0	1	6.70	6.96	−0.26
56	138.87	2.17	1.30	0.24	0	1	8.33	8.43	0.10
61	111.67	2.48	0.78	0.00	0	1	6.96	7.33	−0.37
62	111.72	2.08	0.78	0.88	0	1	8.64	8.69	−0.05
64	111.72	2.08	0.78	0.88	1	0	8.08	8.08	0.00
65	111.72	2.08	0.78	0.88	0	0	7.39	7.32	0.08
66	96.05	2.07	0.79	0.05	0	1	6.82	6.77	0.05
67	107.73	5.36	4.12	0.05	0	1	5.25	6.42	−1.17
69	107.73	5.36	4.12	0.05	0	0	5.12	5.81	−0.69
72	113.43	12.84	0.78	1.21	0	1	7.82	7.83	−0.01
73	109.65	6.57	0.78	0.53	0	1	7.68	7.45	0.23
75	107.73	5.36	4.12	0.05	1	0	5.47	6.57	−1.10
76	111.81	2.09	0.79	0.05	0	0	7.09	6.77	0.32
77	104.07	18.92	0.78	0.18	0	1	7.02	6.57	0.45
78	111.42	34.73	0.78	0.29	0	1	6.57	6.37	0.20
79	112.54	35.02	0.76	0.18	0	1	5.77	6.33	−0.56
81	127.22	2.09	0.78	0.88	0	0	7.54	7.91	−0.37
87	115.53	20.24	0.87	0.00	0	0	6.66	6.22	0.44
88	99.52	2.08	0.77	0.14	0	0	6.32	6.36	−0.04
91	117.00	29.02	0.70	0.05	0	0	5.82	6.04	−0.21
92	113.43	12.84	0.78	1.21	0	0	6.62	7.22	−0.60
99	112.15	50.72	0.78	0.24	0	0	4.00	5.19	−1.19
101	124.13	43.56	0.85	0.18	0	0	5.85	5.85	0.00
102	116.27	13.23	0.77	0.05	0	0	6.30	6.55	−0.25
103	99.07	5.71	0.74	0.19	0	0	5.75	6.25	−0.50
105	106.68	2.53	0.76	0.12	0	0	6.14	6.60	−0.46
112	122.70	2.19	0.50	0.88	1	0	9.05	8.56	0.49
117	115.33	4.49	0.79	1.19	1	0	9.00	8.33	0.67
118	113.43	12.84	0.78	1.21	1	1	7.70	8.59	−0.89
122	124.13	43.56	0.85	0.18	1	1	7.00	7.22	−0.22

[a] Values of −log(IC50) predicted by eq 9.

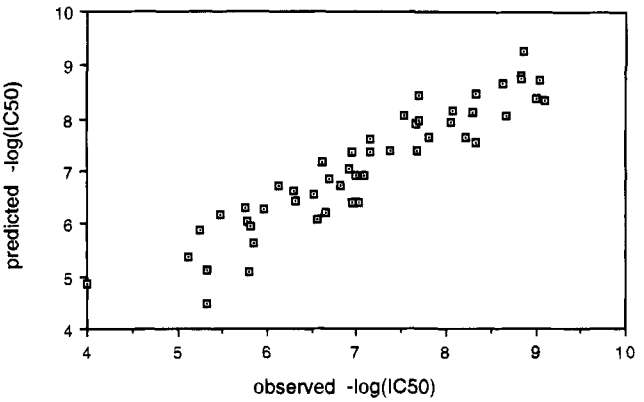


Figure 16. Plot of the observed values of $-\log(IC_{50})$ versus the predicted values of $-\log(IC_{50})$ as predicted by eq 10.

$$-\log(IC_{50}) = 0.135\ (\pm 0.029)V_{ov,m} - 0.053\ (\pm 0.012)V_{non,m} - 0.114\ (\pm 0.029)\Delta E - 6.364\ (\pm 3.286) \quad (11)$$

$$N = 14, R = 0.96, F = 36.3, \text{SD} = 0.39$$

The cross-validated R^2 is 0.20 unit less than the conventional R^2. Predicted and residual $-\log(IC_{50})$ values and the descriptor values are given in Table 13. There is no obvious explanation why compounds **70** and **96**, the *o*-chloro- and *o*-bromophenyl analogs, respectively, should be outliers. One observation is that these compounds along with the *o*-iodophenyl analogs, compounds **108** and **123**, all basically have the same overall steric shape in the active conformation. The activities of these *o*-halophenyl analogs do not correlate to their respective values of ΔE. The ΔE of the *o*-bromophenyl analog, compound **96** is about 2 times and the ΔE of the *o*-iodophenyl analogs, compounds **108** and **123**, is about 3 times the value for the *o*-chlorophenyl analog, yet their activities are all about the same. One possible reason that the ΔE of the *o*-bromo- and *o*-iodophenyl analogs may be too large is that the bromine and iodine atoms cannot be handled by the Mopac version of MNDO. The optimized structure of the *o*-chlorophenyl analog was used to build compounds **96**, **108**, and **123**. The chlorine atom was replaced by the other halogen atoms, and the partial atomic charge of the chlorine atom in compound **70** was also used for the bromine and iodine atoms. The halogen–carbon bond length was modified to the appropriate standard bond length. This approach may have introduced error into the estimation of the ΔE. Additional evidence that there is inconsistency in the ΔE values is that the regression coefficient for this descriptor increases considerably to a value of 0.416 (±0.073), similar to that found in the other 3D-MSA–QSARs, when compounds **108** and **123**, instead of compounds **70** and **96**, are deleted from the set of all the *S*-isomer compounds. The statistics of fit for the equation of all the *S*-isomer compounds, minus compounds **108** and **123**, are $N = 14$, $R = 0.96$, $F = 38.2$, SD = 0.44, which are similar to those of eq **11**.

The cross-validated R^2 values for each model are within 0.1 unit of the conventional R^2 unless otherwise noted. Thus, it may be assumed that these models should be applicable to closely related compounds.[34]

Discussion

In the postulated active conformation, *o*-halophenyl analogs experience unfavorable nonbonded interactions.

Table 13. SAR Table for All *S*-Isomer Analogs Found in Tables 1 and 2

				$-\log(IC_{50})$		
compd	$V_{ov,m}$	$V_{non,m}$	ΔE	obs	pred[a]	residual
2	120.85	2.17	0.37	10.10	9.73	0.37
58	120.85	2.17	0.37	9.22	9.73	−0.51
83	111.72	2.08	0.78	8.60	8.46	0.14
85	115.33	4.49	0.79	8.54	8.82	−0.28
90	114.71	5.17	0.78	8.46	8.70	−0.24
93	115.33	4.49	0.79	9.10	8.82	0.28
94	124.13	43.56	0.85	7.72	7.91	−0.19
95	120.05	10.02	0.78	9.10	9.16	−0.06
97	106.68	2.53	0.76	7.37	7.76	−0.39
104	118.40	11.73	0.78	9.12	8.85	0.27
108	110.72	17.54	13.63	6.10	6.03	0.07
114	113.43	12.84	0.78	8.89	8.12	0.77
119	113.43	12.84	0.78	7.96	8.12	−0.16
123	110.72	17.54	13.63	5.96	6.03	−0.07

[a] Values of $-\log(IC_{50})$ predicted by eq 11.

The N-methylated indole analogs can adopt energetically stable conformations in the region of torsion angle space of the postulated active conformation because rotation of the methyl group allows a minimization of unfavorable steric contacts. The *o*-amino *p*-chloro analog, compound **112**, can also adopt this conformation as this active conformer state is, in fact, near its global minimum. Intramolecular conformational stability, therefore, is critical for binding affinity and is represented by the ΔE descriptor, the difference between the fixed valence geometry intramolecular energy of the active conformation and the global minimum energy conformation. The orientation of the aromatic ring, with respect to the amide hydrogen and carbonyl oxygen, both potential intermolecular hydrogen-bonding sites, determines the overall shape and position of the 3-amido pharmacophore. Figure 17 shows a representative analog of high activity from each of the Tables 1–3 in the active conformation. Ideally if $\phi_1 = -130°$, then ϕ_2, as depicted in the figure, should be 130–140°. Exclusions from this torsion angle space result in a decrease in activity.

A molecular shape descriptor appears in almost every 3D-MSA–QSAR. Mutant, a molecular volume hybrid consisting of the union of structures of compounds **24**, **51**, **81**, and **112**, is the most statistically significant shape reference "compound". This "compound" is essentially carving out the allowed space of the binding site. Volume not included by this space is detrimental to binding affinity. Examples of analogs that demonstrate unfavorable nonoverlap volume with reference to the mutant structure are given in Figure 18. Areas of receptor steric interaction are noted schematically in Figure 19. Compounds with atoms/groups that extend too far above and/or below this plane exhibit a marked decrease in activity. A steric interaction between the receptor and the antagonist may occur 1.3–2.0 Å above and below the plane in the area indicated by Figure 19a. This receptor mapping is derived from compounds **78**, **79**, **87**, **91**, and **99**. The lower activity of these analogs is largely attributed to the amount of nonoverlap volume which occurs above/below the plane of the aromatic ring. If it is accepted that the BZD ring structure acts as an anchor in the binding site, then it can be postulated that if ϕ_2 is rotated to a value of 90° the aromatic ring of even the most active compounds would interact with the nonoverlap volume depicted in Figure 19a. This supports the idea that the binding site prefers the aromatic

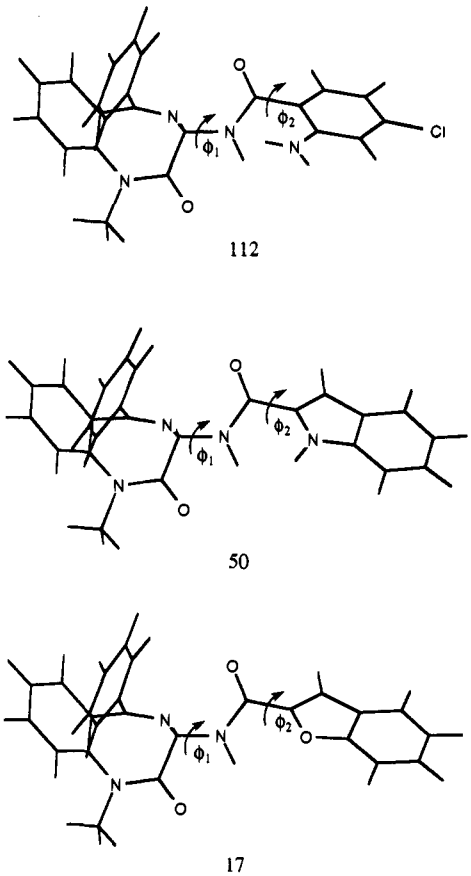


Figure 17. Compounds **112**, **50**, and **17** shown in the active conformation. Ideally $\phi_1 = -130°$, $\phi_2 = 130-140°$.

ring at a particular orientation. Volume extending outward in the plane of the aromatic ring 4.5–5.0 Å away from the illustrated axis may also result in decreased activity and is proposed from examining compounds **5** and **26**. Compound **5** can adopt the active conformation without a significant loss of intramolecular energy. The low activity of this analog is likely attributable to the steric interaction of the nonoverlap volume with the postulated receptor "wall" (Figure 19b). The inactivity of compound **26** may be due to its high ΔE and also to the amount of nonoverlap volume in this same region. The steric/lipophilic area of receptor space 1 Å away from the accessible surface, as depicted in Figure 19b, is carved out by compounds **48**, **53**–**55**. The inactivity of these compounds is most likely due to an unfavorable steric/lipophilic interaction of the 5′-substituent with the receptor.

Although compound **99** is not an outlier in the 3D-MSA–QSARs, the residual for this analog is larger than that for more other compounds. In basically all of the optimal 3D-MSA–QSARs, shape alone is not able to adequately explain the low activity of this analog. The answer may lie in the descriptor dist. The distance between the carbonyl carbon and the furthermost atom corresponding to the direction of the phenyl para substituent is 9.65 Å for compound **99** which is the largest among all analogs. The next closest distance is 8.94 Å belonging to compound **55**. Although dist does not appear in any of the optimal 3D-MSA–QSARs, this descriptor may still be a determinant of activity. All other compounds in the data set may have acceptable dist values, and only compound **99** extends beyond the proposed binding site. Therefore it appears that there

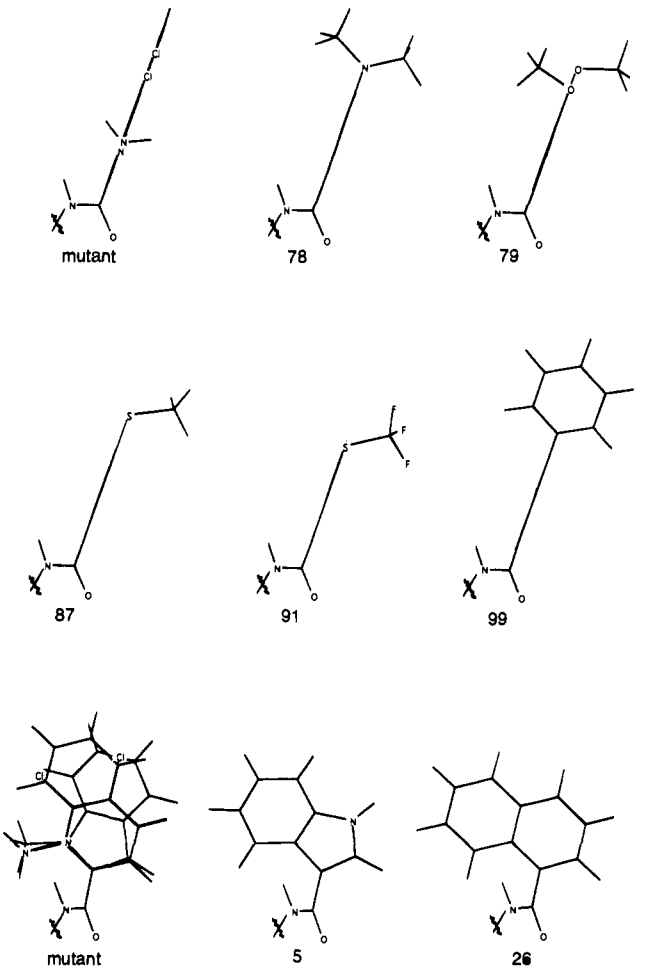


Figure 18. Structures of various analogs that demonstrate nonoverlap volume which may interact with the CCK-A receptor. The mutant is shown as a reference in two different views. The first view is looking perpendicular to the plane of the aromatic ring. The second view is looking above the plane. Compounds **5**, **26**, **78**, **79**, **87**, **91**, and **99** are also shown.

is a limit to the length of the 3-amido substituent of approximately 9.65 Å, as depicted in Figure 19a.

The finding that common overlap steric volume is also an important descriptor for activity suggests an analog must occupy a critical minimum volume to exhibit moderate activity. The most active analog contains an indole ring, further suggesting that this is near the ideal shape and volume. However, the mutant is predicted to be the optimal shape reference compound for common steric overlap volume when all other descriptors are included in the regression equation. Compound **50**, which contains an indole ring and is present in the structure of the mutant, is also a highly significant shape reference compound for common overlap steric volume. This behavior may indicate that in order for an analog to show affinity it should optimally possess a shape similar to an indole ring. However, moderate activity can be realized if the structure extends outward in certain positions of the plane of the aromatic group as is the case for compounds **11**, **17**, **24**, and **81**. The pharmacophoric receptor accessible volume is essentially mapped out by the shape reference "compound", the mutant. This volume is depicted in Figure 19b. Steric volume occuring within the volume of the mutant is acceptable.

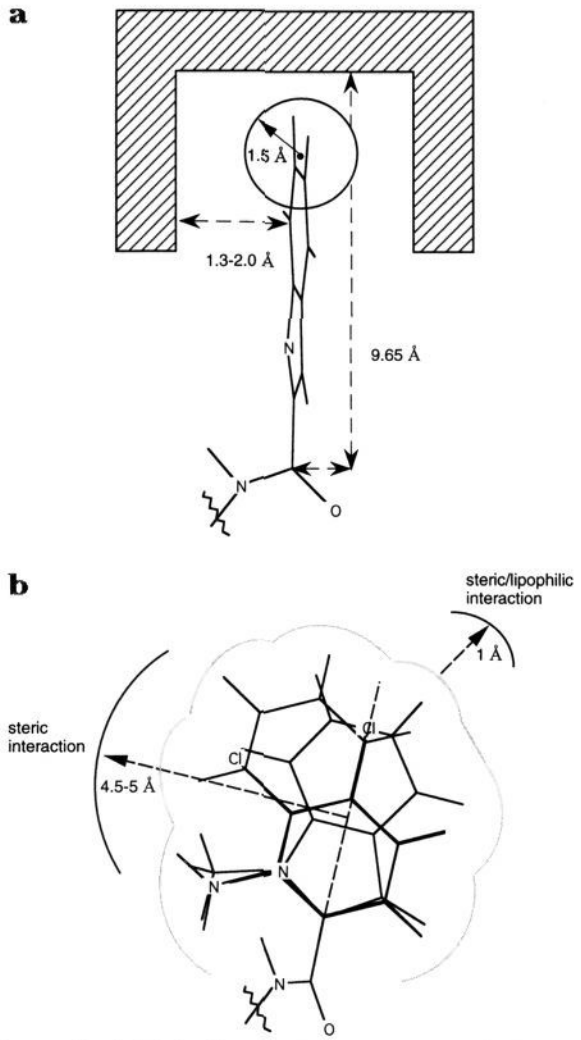


Figure 19. (a) 3-Amido substituent of compound **50** shown in the proposed CCK-A receptor pharmacophoric binding site. The distance to the nonallowed volume above and below the plane of the aromatic ring may occur at 1.3–2.0 Å. Also the maximum distance of 9.65 Å from the carbonyl carbon to the receptor "wall" is depicted. A lipophilic interaction sphere of radius 1.5 Å is also shown. (b) Planar view of the 3-amido substituent of mutant bound in the proposed CCK-A pharmacophoric binding site. The accessible volume is depicted within the gray boundaries. Potential steric/lipophilic interaction sites are indicated.

Lipophilicity is involved in the expression of binding affinity. However, the descriptor LOGP does not appear to be significant. Rather, the sum of the π-constants within a sphere of radius ranging between 1.5 and 3.0 Å appears in many of the trial 3D-MSA–QSARs. The term ($\Sigma\pi$ 1.5 Å), or its square, is a descriptor that is found to add significance to the optimal regression equation. This suggests that the location in space corresponding to a sphere of 1.5 Å about the phenyl *p*-iodine of compound **95** may be near a hydrophobic pocket of the receptor. This interpretation is strengthened with the comparison of the poorly active phenyl *p*-OH analog, compound **103**, and the moderately active *p*-Cl analog, compound **65**. These analogs have identical substitution patterns on the BZD ring structure and only differ in the para substituent. However, there is a 2 orders of magnitude difference in their activities. The lipophilic interaction sphere is depicted in Figure 19a.

An *N*-methyl group on the BZD ring structure is favorable for binding affinity. The indicator variable I_M appears in the optimal 3D-MSA–QSARs containing all the compounds. This may indicate that the methyl group is making a favorable interaction with the receptor. Also the *N*-methyl group has been found to affect the conformation of the BZD ring structure in that it influences the seven-membered ring flip[35] which, in essence, could influence the time-averaged position of the 3-amido substituent.

The presence of a fluorine on the ortho position of the 5-phenyl ring of the BZD ring structure enhances binding as it has a positive coefficient in the 3D-MSA–QSAR. This suggests there is a favorable interaction of the fluorine atom with the receptor. The addition of a fluorine does not always increase binding affinity, however, as there are a few cases contrary to this trend. For example, compare the activities of compounds **114** and **119** which show a reverse trend. In more cases than not, however, the presence of fluorine increases binding affinity.

Overall, a 3D pharmacophore for the 3-amido substituent of the benzodiazepine CCK-A antagonists is derived from the 3D-MSA–QSARs, and an active conformation is hypothesized for these compounds using the loss in biological activity–loss in conformational stability principle. A shape reference compound that consists of selected overlapped structures expands the definition of accessible receptor space. The elimination of one structure from the mutant at a time could define what area of ligand space each compound occupies by examining which analogs become outliers in the resultant 3D-MSA–QSARs.

Acknowledgment. John S. Tokarski is a grateful recipient of an American Foundation for Pharmaceutical Education Fellowship. Resources of the Laboratory of Molecular Modeling and Design at UIC were used over the course of the project. We also appreciate the helpful discussions with Prof. William Dunn III and members of the laboratory regarding this work.

References

(1) Williams, J. A. Cholecystokinin: A Hormone and Neurotransmitter. *Biomed. Res.* **1982**, *3*, 107–121.
(2) Dockray, G. J. The Physiology of Cholecystokinin in Brain and Gut. *Br. Med. Bull.* **1982**, *38*, 253–258.
(3) Dourish, C. T.; Hill, D. R. Classification and Function of CCK Receptors. *Trends Pharmacol. Sci.* **1987**, *8*, 207–208.
(4) Moran, T. H.; Robinson, P. H.; Goldrich, M. S.; McHugh, P. R. Two Brain Cholecystokinin Receptors: Implications for Behavioral Actions. *Brain Res.* **1986**, *362*, 175–179.
(5) Saito, A.; Sankaran, H.; Goldfinde, I. D.; Williams, J. A. Cholecystokinin Receptors in Brain: Characterization and Distribution. *Science* **1980**, *208*, 1155–1156.
(6) Lin, C. W.; Holladay, M. W.; Barrett, R. W.; Wolfram, C. A. W.; Miller, T. R.; Witte, D.; Kerwin, J. F., Jr.; Wagenaar, F.; Nadzan, A. M. Distinct Requirements for Activation at CCK-A and CCK-B/Gastrin Receptors: Studies with a C-Terminal Hydrazide Analogue of Cholecystokinin Tetrapeptide (30–33). *Mol. Pharmacol.* **1989**, *36*, 881–886.
(7) Mutt, V. Cholecystokinin: Isolation, Structure, and Functions. In *Gastrointestinal Hormones*; Glass, G. B. J., Ed.; Raven Press: New York, 1980; pp 169–221.
(8) Ravard, S.; Dourish, C. T. Cholecystokinin and Anxiety. *Trends Pharmacol. Sci.* **1990**, *11*, 271–273.
(9) Vickroy, T. W.; Bianchi, B. R.; Kerwin, J. F.; Kopecka, H.; Nazdan, A. M. Evidence that Type A CCK Receptors Facilitate Dopamine Efflux in Rat Brain. *Eur. J. Pharmacol.* **1988**, *152*, 371–372.

(10) Moran, T. H.; Ameglio, P. J.; Schwartz, G. J.; McHugh, P. R. Blockade of Type A, not Type B CCK Receptors Attenuates Satiety Actions of Exogenous and Endogenous CCK. *Am. J. Physiol.* **1992**, *262*, R46–R50.
(11) de Montigny, C. Cholecystokinin Tetrapeptide Induces Panic-like Attacks in Healthy Volunteers. *Arch. Gen. Psychiatry* **1989**, *46*, 511–517.
(12) Woodruff, G. N.; Hughes, J. Cholecystokinin Antagonists. *Annu. Rev. Pharmacol. Toxicol.* **1991**, *31*, 469–501.
(13) Gertz, B. J. Potential Clinical Applications of a CCK Antagonist. In *Neurol. Neurobiol., 47*; Wang, R. Y., Schoenfeld, R., Eds.; Alan, R. Liss: New York, 1988; pp 327–342.
(14) Evans, B. E.; Rittle, K. E.; Bock, M. G.; Dipardo, R. M.; Freidinger, R. M.; Whitter, W. L.; Lundell, G. F.; Veber, D. F.; Anderson, P. S.; Chang, R. S. L.; Lotti, V. J.; Cerino, D. J.; Chen, T. B.; Kling, P. J.; Kunkel, K. A.; Springer, J. P.; Hirshfield, J. Methods for Drug Discovery: Development of Potent, Selective, Orally Effective Cholecystokinin Antagonists. *J. Med. Chem.* **1988**, *31*, 2235–2246.
(15) Mabilia, M.; Pearlstein, R. A.; Hopfinger, A. J. In *Molecular Graphics and Drug Design*; Burger, A. S. V., Roberts, G. C. K., Tute, M. S., Eds.; Elsevier: Amsterdam, 1986; p 158.
(16) Mabilia, M.; Pearstein, R. A.; Hopfinger, A. J. Molecular shape analysis and energetics-based intermolecular modeling of dihydrofolate reductase inhibitors. *Eur. J. Med. Chem.-Chim. Ther.* **1985**, *20*, 163–174.
(17) Pearlstein, R. A. *Chemlab-II Users Guide, V10.0*; Chemlab Inc.: 1780 Wilson Dr., Lake Forest, IL 60045, 1988.
(18) Stewart, J. J. P.; Seiler, F. K. QCPE #455 (Ver. 4.0), 1987.
(19) Allinger, N. L. Conformational Analysis. 130. MM2. A Hydrocarbon Force Field Utilizing V_1 and V_2 Torsional Terms. *J. Am. Chem. Soc.* **1977**, *99*, 8127–8134.
(20) *Tables of Interatomic Distances and Configuration in Molecules and Ions. Supplement, Special Publication No. 18*; The Chemical Society: London, 1965; pp S3s–S23s. *Handbook of Chemistry and Physics*, 50th ed.; CRC Press: Cleveland, 1969; pp F-154–F-157.
(21) Iwamoto, T.; Kashino, S.; Haisa, M. Structure of Cinnamamide. *Acta Crystallogr.* **1989**, *C45*, 1110–1112.
(22) Burke, B. J. Developments in Molecular Shape Analysis to Establish Spatial Molecular Similarity Among Flexible Molecules. Ph.D. Thesis, Dept. of Medicinal Chemistry and Pharmacognosy, University of Illinois at Chicago, IL, 1992.
(23) Hopfinger, A. J. Conformational Energies and Potential Functions. In *Conformational Properties of Macromolecules*; Academic Press: New York, 1973; Chapter 2.
(24) Crawford, R. J.; Pearlstein, R. A.; Mabilia, M.; Hopfinger, A. J. Characterization and Selection of Atom-Pair Potential Functions for Conformational Analysis. *Tetrahedron Comput. Methodol.* **1988**, *1*, 185–206.
(25) Hopfinger, A. J. A QSAR investigation of dihydrofolate reductase inhibition by baker triazines based upon molecular shape analysis. *J. Am. Chem. Soc.* **1980**, *102*, 7196–7206.
(26) Wold, S.; Albano, C.; Dunn, W. J., III; Edlund, U.; Esbensen, K.; Geladi, P.; Helberg, S.; Johansson, E.; Lindberg, W.; Sjöström, M. Multivariate Data Analysis in Chemistry. In *Chemometrics: Mathematics and Statistics in Chemistry*; Kowalski, B. R., Ed.; D. Reidel Publishing Co.: Dordrecht, Holland, 1987; pp 17–96.
(27) Evans, B. E.; Bock, M. G.; Rittle, K. E.; Dipardo, R. M.; Whitter, W. L.; Veber, D. F.; Anderson, P. S.; Freidinger, R. M. Design of Potent, Orally Effective, Nonpeptidal Antagonists of the Peptide Hormone Cholecystokinin. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 4918–4922.
(28) *Quanta version 3.3*; MSI: 16 New England Executive Park, Burlington, MA 01803.
(29) *Medicinal Chemistry Software, MedChem Software Manuel, Release 3.51*; Ponoma College: Claremont, CA, April 1987.
(30) Hansch, C.; Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*; Wiley-Intercourse: New York, 1979.
(31) Dunn, W. J., III; Wold, S. Relationship Between Chemical Structure and Structure and Biological Activity Modeled by SIMCA Pattern Recognition. *Biorg. Chem.* **1980**, *9*, 505–523.
(32) *SAS Institute, Inc., SAS User's Guide: Basics, SAS Release 5.18*; SAS Institute, Inc.: Cary, NC, 1986.
(33) Kleinbaum, D. G.; Kupper, L. L.; Muller, K. E. Testing Hypotheses for Multiple Regression. In *Applied Regression Analysis and Other Multivariate Methods*; PWS-Kent: Boston, 1988; Chapter 9.
(34) Cramer, R. D., III; Bunce, J. D.; Patterson, D. E. Crossvalidation, Bookstrapping, and Partial Least Squares Compared with Multiple Regression in Conventional QSAR Studies. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18–25.
(35) Hamor, T. A.; Martin, I. L. The Benzodiazepines. In *Progress in Medicinal Chemistry*; Ellis, G. P., West, G. B., Eds.; Elsevier: Amsterdam, 1983; Vol. 20, pp 158–223.