Three-Dimensional Quantitative Structure-Activity Relationship of Angiotesin-Converting Enzyme and Thermolysin Inhibitors. II. A Comparison of CoMFA Models Incorporating Molecular Orbital Fields and Desolvation Free Energies Based on Active-Analog and Complementary-Receptor-Field Alignment Rules

Chris L. Waller and Garland R. Marshall*

Center for Molecular Design, Washington University, St. Louis, Missouri 63130-4899

Received March 11, 1993

The utility of comparative molecular field analysis (CoMFA), a three-dimensional Quantitative Structure-Activity Relationship (3-D QSAR) paradigm, as a tool to aid in the development of predictive models has been previously addressed (Depriest, S.D. *et al., J. Am. Chem. Soc.* **1993**, in press). Although predictive correlations were obtained for angiotensin-converting and thermolysin inhibitors, certain inadequacies of the CoMFA technique were noted. Primarily, CoMFA steric and electrostatic fields alone do not fully characterize the zinc–ligand interaction. Previously, this was partially rectified by the inclusion of indicator variables into the QSAR table to designate the class of zinc-binding ligand. Recent advances in molecular modeling technology have allowed us to further address this limitation of the preceding study. Using molecular orbital fields derived from semiempirical calculations as additional descriptors in the QSAR table, predictive correlations were produced based on CoMFA and molecular orbital fields alone—indicator variables no longer being necessary. Arbitrary information concerning the alignment of molecules under study within the active-site introduces ambiguities into the CoMFA study. Crystallographic information detailing the binding mode of several thermolysin enzyme inhibitors has previously been used as a guide for the alignment of additional, noncrystallized, inhibitors. However, this process was complicated by the lack of parameters for zinc in the molecular mechanical force field. Therefore, zinc–ligand interactions were ignored during the standard minimization procedure. The use of field-fit minimization using complementary receptor fields as templates is presented as a possible solution to the problem. Predictive correlations were obtained from analyses based on this method of molecular alignment. The availability of crystallographic data for thermolysin enzyme–inhibitor complexes allowed for an alternate definition of the CoMFA region. Herein, promising results from analyses using actual receptor active-site atom probe atoms are presented.

Introduction

Angiotensin converting enzyme (ACE) is a zinc-containing metallopeptidase which is instrumental in the conversion of angiotensin I to angiotensin II, a potent vasoconstrictor. Therapeutic inhibition of this step in the renin–angiotensin system has proven to be a very effective treatment for the management of hypertension. Although the 3-D structure of ACE is unknown, much information about the structural and electronic characteristics of the molecular binding domain of the enzyme has been determined from traditional structure–activity relationship (SAR) studies.[1-3] There are three requirements which appear, on the basis of the current literature, to be necessary for successful inhibition of the enzyme: (1) a functional group capable of binding to zinc in the active site (i.e. carboxylate, hydroxamate, phosphonate, or sulfhydryl); (2) a carbonyl oxygen capable of accepting a hydrogen bond from some donor residue functional group (i.e. O–H, N–H); and (3) an ionizable C-terminal carboxylate moiety which interacts with a positively charged residue (i.e. arginine, lysine).

Using these three requirements, Mayer *et al.*[4] deduced a hypothetical, active-site geometry from the examination of 28 structurally diverse and conformationally-flexible ACE inhibitors. By using systematic search, the conformational hyperspace available to the amide carbonyl, the carboxyl group, and the Zn-binding functional group of all 28 molecules was explored. A unique active-site model was developed from the analysis of the relative orientations of the active-site groups common to all molecules included in the study. Models of this type allow for the qualitative comparison (evaluation) of potential ACE inhibitors possessing the necessary structural requirements, regardless of the molecular framework, but provide little direct quantitative information regarding affinity.

An early quantitative structure–activity relationship (QSAR) study[3] of ACE inhibitors focused on semiempirically-determined, molecular electronic indices as physical descriptors of the charge–charge interactions between the ligand and postulated binding sites of the receptor. Of 67 total indices (including volume, surface area, and rotatable bond data), the *z* component of the dipole moment (an uncertain quantity for charged species), the square of the electrostatic potential, and the second-order polarization energy were the most descriptive regressors. From this study, it was concluded that the hydrogen-bonding interaction was of primary importance while the overall electrostatic interaction and binding to the carboxylate and zinc make smaller contributions to the biological activity. It is also of interest to note that the overall energy

* To whom correspondence should be addressed.

0022-2623/93/1836-2390$04.00/0

of the highest occupied molecular orbital (HOMO) was determined to be unimportant.

Recently, we reported[6] on the use of comparative molecular field analysis (CoMFA)[7] a 3-D QSAR technique,[8] in the development of predictive models for inhibitors of ACE and thermolysin. A predictive model for the ACE inhibitors was obtained, but several limitations of the current CoMFA methodology were noted. First, the predictive ability of the CoMFA-only model suffered due to a possible deficiency in the ability of the CoMFA fields to discern between different types of Zn-binding ligands. This was believed to result in a systematic trend of overprediction for certain molecular classes (i.e. phosphate–based inhibitors). This trend was partially attenuated by the addition of indicator variables representing the type of Zn-binding functional group present in the molecule. This was rationalized as necessary to account for differences in zinc–ligand interactions which would not be reflected in the electrostatic field calculations used in the CoMFA analysis. While partially successful, this technique also has certain limitations. Most notably, prediction of the affinity of a molecule possessing a novel Zn-binding functional group which was unknown to the model would lack an appropriate indicator variable. Second, while it is generally accepted[9-11] that CoMFA steric (van der Waal's) and electrostatic (Coulombic) fields are capable of adequately representing the enthalpic nature of ligand–receptor interactions, the entropic nature of this interaction is not explicitly considered in a CoMFA analysis, and differences in solvation and internal degrees of freedom could impact the entropic aspects of the free energy of binding.

With respect to the apparent inability of the CoMFA fields alone to distinguish between various Zn-binding functional functional groups, we considered molecular orbital field characteristics as additional regressors in the 3-D QSAR model. Examination of the HOMO's of ACE inhibitors revealed that it is highly dependent on the formal charge pattern of the molecule and is primarily localized at, or near, the ionized Zn-binding functionality. By characterizing both the magnitude (size) and the spatial position (location) of the HOMO with respect to the ligands included in the CoMFA study, it was hoped that the electronic interaction between the ligand and the positively charged zinc in the active site could be fully determined, thus rectifying the systematic overpredictive error noted in previous analyses without the use of the indicator variables. In order to address the lack of entropic characterization, we investigated inclusion of the calculated free energies of solvation (desolvation) in the models.

As before,[6] the analogous series of thermolysin inhibitors was used as a internal calibration test set. Although the X-ray crystal structures of many thermolysin inhibitors bound in the active site are available, certain assumptions were previously made concerning the treatment of the zinc in the molecular mechanics force field. Due to the lack of parameters for zinc in the standard Tripos force field, pertinent bond distances and angles were not considered interesting during the minimization process. Alternate force fields with parameters for zinc require *a priori* knowledge pertaining to coordination geometry and oxidation state of the metal. To help eliminate these ambiguities and to test the concept underlying the available field-fit minimization option, we examined the use of complementary receptor fields as a method of representing the 3-D structure of the binding site as a novel method of ligand alignment.

Methods

A. CoMFA Interaction Energies. All molecular modeling and CoMFA studies were performed on a Silicon Graphics Iris 4D/380 running SYBYL[12] versions 5.5 and 6.0. The steric, electrostatic, and molecular orbital interaction energies for the CoMFA studies were calculated using an sp^3 carbon probe atom with a +1 charge and a distance-dependent dielectric constant at all intersections of regularly-spaced (2 Å) grids of dimensions 26 × 24 × 20 Å for the ACE series and 28 × 28 × 24 Å for the thermolysin series. Additionally, interaction energies for the thermolysin CoMFA studies were determined at coordinates corresponding to the atomic coordinates of thermolysin within a 12-Å sphere of the binding domain centered about the ligand. Therefore, the effective probe corresponded to the respective atom type of the enzyme. The partial charges of the enzyme were determined using either the Kollman all-atom[13] convention or by the method of Gasteiger and Hueckel.[14] Partial atomic charges of the ligand molecules were determined using the PM3 model Hamiltonian[15] within MOPAC 5.0[16] or by the method of Gasteiger and Hueckel. Molecular orbital descriptors for the ligands were taken from the semiempirical calculations above. The cutoff value for all interactions was set to +30 kcal/mol. The electrostatic contributions at lattice intersections yielding maximal (+30 kcal/mol) steric values were ignored. Regression analyses were done using the partial least-squares (PLS) methodology[17] in conjunction with cross-validation.[18]

B. Delphi Desolvation Energies.

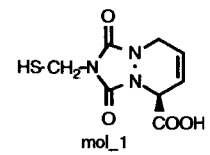
All Delphi free-energy calculations were performed on an IBM Model 550 RISC 6000 workstation. The free energy of desolvation was computed for all of the analogs of the ACE series using the finite difference approximation method[19,20] as employed within Delphi.[21] The molecules in SYBYL mol2 format were read using Insight II version 2.20 and converted to Biosym format by fixing the atomic potentials and accepting the MOPAC/PM3 charges using the Builder module. All molecules were subjected to Delphi calculations using a solute dielectric of 2.0 and a solvent dielectric of 1.0 (vacuum) and 80.0 (aqueous) with a solvent radius of 1.80 Å in either case. Free energies were then calculated as the difference between the total energies (reaction plus grid) of both states.

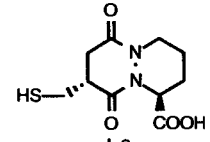
C. "Predictive" r^2 Values. As utilized previously,[6] the "predictive" r^2 will only be based on molecules not included in the training set and is defined as:

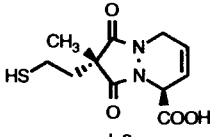
$$\text{"predictive" } r^2 = (\text{SD} - \text{"press"})/\text{SD}$$

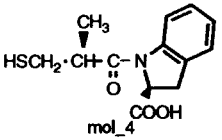
where SD is the sum of the squared deviations between the biological activities of molecules in the test set and the mean activity of the training set molecules and "press" is the sum of the squared deviations between predicted and actual biological activity values for every molecule in the test set. This is analogous to Cramer's definition[18] and likewise can result in a negative value reflecting a complete lack of predictive ability of the training set for the molecules included in the test set.

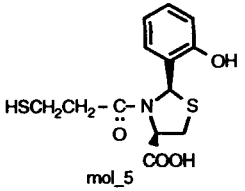
D. CoMFA of ACE Inhibitors. The 68 molecules[3,22-36] (Figure 1 and Table I) used to define the training set previously were also used in this implementation of CoMFA. The original alignment, obtained using the

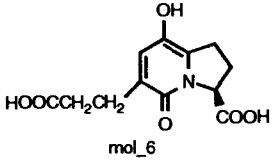

mol_1

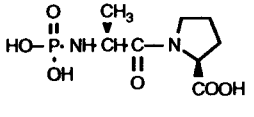

mol_2

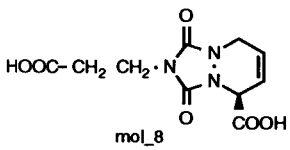

mol_3

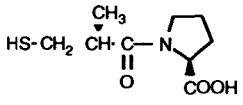

mol_4

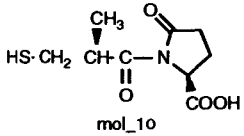

mol_5

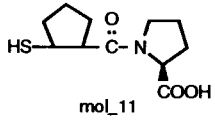

mol_6

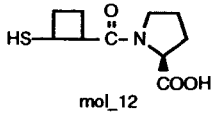

mol_7

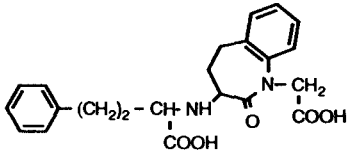

mol_8

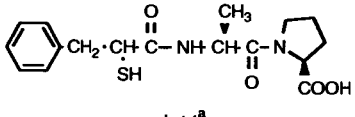

mol_9

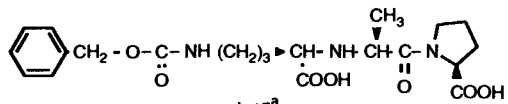

mol_10

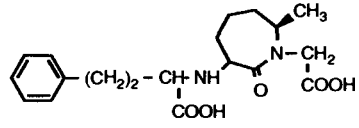

mol_11

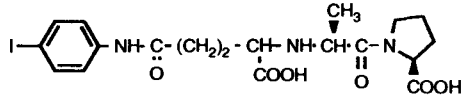

mol_12

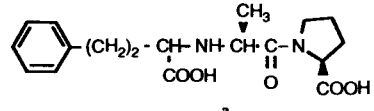

mol_13[a]

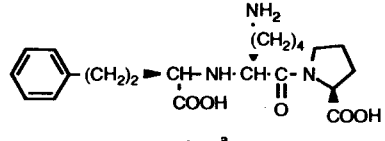

mol_14[a]

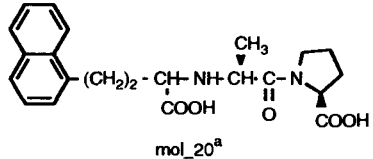

mol_15[a]

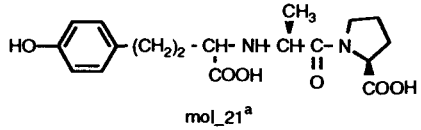

mol_16[a]

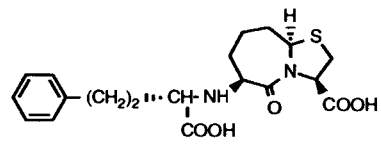

mol_17[a]

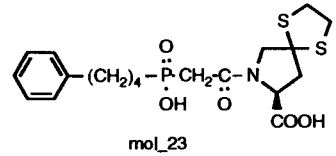

mol_18[a]

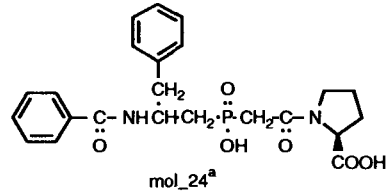

mol_19[a]

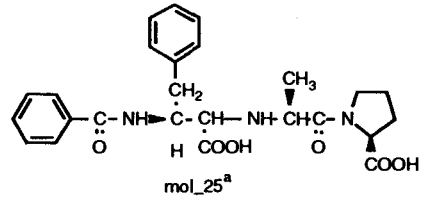

mol_20[a]

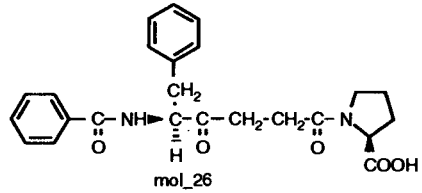

mol_21[a]

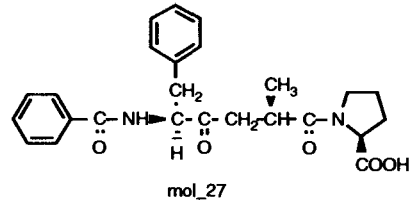

mol_22[a]

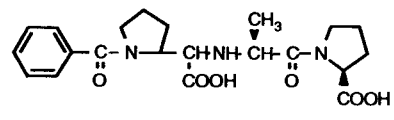

mol_23

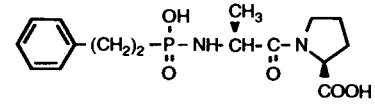

mol_24[a]

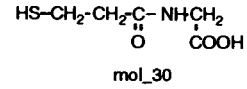

mol_25[a]

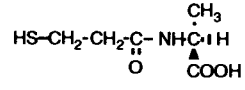

mol_26

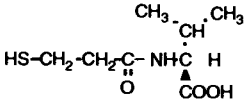

mol_27

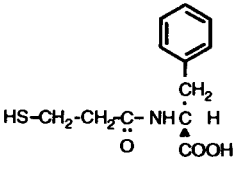

mol_28

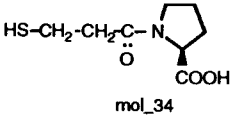

mol_29


mol_30


mol_31


mol_32


mol_33


mol_34

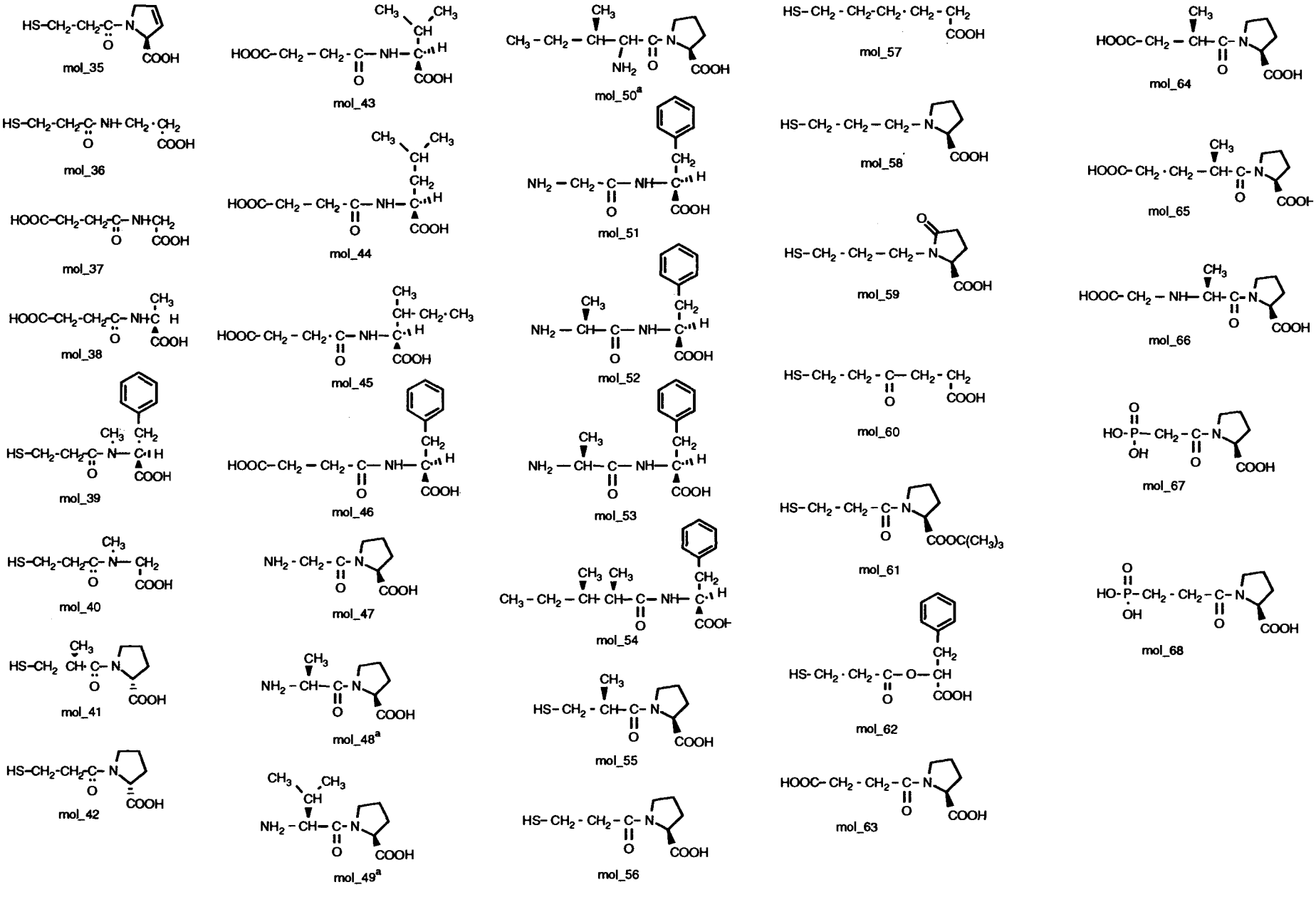


Figure 1. Molecules included in the ACE training set. An a designation means that ambiguous chiral centers were modeled in the S configuration.

Table I. Molecules Included in ACE Training Set

molecule	pIC_{50} (M)	ΔG (kcal)	ref(s)
mol_01	6.15	210.95	22
mol_02	7.15	206.75	23
mol_03	7.00	203.49	23
mol_04	8.22	207.04	24
mol_05	8.43	211.04	5
mol_06	6.34	207.65	5
mol_07	6.11	218.50	25
mol_08	9.00	209.99	5
mol_09	7.64	215.71	26
mol_10	8.05	210.28	27
mol_11	7.19	211.78	5
mol_12	7.31	210.11	5
mol_13	8.77[a]	198.14	5
mol_14	7.30[a]	192.00	28
mol_15	8.54[a]	211.58	5
mol_16	8.52[a]	202.86	5
mol_17	9.64[a]	200.72	5
mol_18	8.92[a]	208.80	29
mol_19	8.92[a]	202.99	29
mol_20	8.96[a]	206.35	5
mol_21	8.55[a]	209.19	5
mol_22	9.22[a]	199.29	5
mol_23	8.40	199.26	5
mol_24	8.00	199.69	5
mol_25	8.11[a]	206.14	24
mol_26	7.92	85.49	5
mol_27	8.52	87.49	5
mol_28	8.54	209.36	5
mol_29	8.15	203.93	5
mol_30	5.55	244.55	3
mol_31	6.07	220.55	3
mol_32	5.80	215.79	3
mol_33	6.37	214.68	3
mol_34	6.70	217.93	3
mol_35	6.19	215.30	3
mol_36	3.31	217.16	3
mol_37	2.70	226.64	3
mol_38	2.87	223.21	3
mol_39	4.51	213.55	3
mol_40	5.52	211.90	3
mol_41	4.96	215.52	3
mol_42	2.74	217.83	3
mol_43	2.96	220.83	3
mol_44	3.21	217.21	3
mol_45	2.98	217.35	3
mol_46	3.26	219.13	3
mol_47	3.35	81.83	3
mol_48	3.64[a]	80.52	3
mol_49	3.38[a]	77.40	3
mol_50	3.89[a]	75.88	3
mol_51	3.22	81.45	3
mol_52	3.72	80.22	3
mol_53	4.28	77.51	3
mol_54	3.03	75.56	3
mol_55	3.62	81.01	3
mol_56	4.77	199.79	3
mol_57	2.96	216.11	3
mol_58	3.62	210.31	3
mol_59	3.19	210.44	3
mol_60	5.62	220.48	3
mol_61	4.41	192.16	3
mol_62	6.15	214.45	3
mol_63	4.48	221.63	5
mol_64	4.99	218.91	5
mol_65	5.31	220.01	5
mol_66	5.62	219.65	5
mol_67	5.08	218.07	5
mol_68	4.32	216.36	5

[a] Activity is for a diastereomeric or racemic mixture.

MULTIFIT algorithm in SYBYL to fit the Mayer geometry,[4] was not altered in order to allow direct comparison with previous results. The initial PLS analysis was performed with full cross-validation (68 groups) and a maximum of 10 components using only the CoMFA steric and electrostatic field values as regressors. Traditionally, this analysis is followed by a non-cross-validated analysis using the optimum number of components (typically that component number which yields the highest cross-validated r^2).

Since it was previously noted that the steric and electrostatic fields alone do not sufficiently characterize the interaction between the zinc–ligand and the receptor,[6] an additional field representing the magnitude and spatial location of the highest occupied molecular orbital (HOMO) of the ligand with respect to the CoMFA region was added to the QSAR table. This was accomplished by calculating the HOMO of each molecule relative to the CoMFA region (lattice) and incorporating these values as a CoMFA steric-only column in the QSAR table. The regression analyses were performed as above with the addition of the HOMO column as a third regressor. Furthermore, desolvation free energies were incorporated into the QSAR table as an additional column of explicit data. These values were used in conjugation with the CoMFA steric and electrostatic columns with and without HOMO columns. The resulting models were tested for their ability to predict the activities of the same test sets versus the CoMFA-only model.

To test the predictive ability of the resulting models, the test set of 20 structurally diverse ACE inhibitors (Figure 2 and Table II) was used from the literature.[3,22-36] As before,[6] individual test sets of carboxylates[37] (Table III), phosphates[38] (Table IV), and thiols[39] (Table V) were used to ascertain the ability of the resultant models to distinguish between different classes of ligands.

E. CoMFA of Thermolysin Inhibitors. Initially, CoMFA studies were performed on the series of 61 inhibitors[240-57] (Table VI) aligned by the MULTIFIT method reported previously. Charges were calculated for the ligands using the PM3 model Hamiltonian in MOPAC 5.0. Since the Tripos force field is not parameterized for zinc, an additional alignment technique was employed in order to eliminate or minimize the ambiguities associated with the original alignment rule. Specifically, the complementary steric and electrostatic fields of the crystal structure for thermolysin were used as template fields on which to field-fit minimize the various ligands. These fields were obtained by probing the molecular-binding domain of the receptor, including all residues within 4 Å of the molecules included in the training set, and effectively reversing the steric and electrostatic field values. All molecules were then minimized with respect to these fields to an energy convergence criterion of 0.001 kcal/mol using the standard Tripos force field.[58] A subsequent standard reminimization was performed to allow all molecules to relax to the nearest local minimum. Limitations on the quantity of heavy atoms which can be handled by the MOPAC program precluded its use in the calculation of partial atomic receptor charges. Therefore, Kollman charges were used on the receptor in the analyses using PM3 ligand charges. This disparity between receptor and ligand charge sets was addressed by using Gasteiger and Hueckel charges on both receptor and ligand in subsequent studies. Additionally, the availability of crystal structure data for the thermolysin molecule allowed for the CoMFA region to be specifically defined in comparative analyses.

As was the case for the ACE series, preliminary analyses were performed with full cross-validation and a maximum of 10 components with combinations of CoMFA steric and electrostatic fields, MOPAC HOMO fields, and Delphi desolvation energies serving as regressors for appropriate

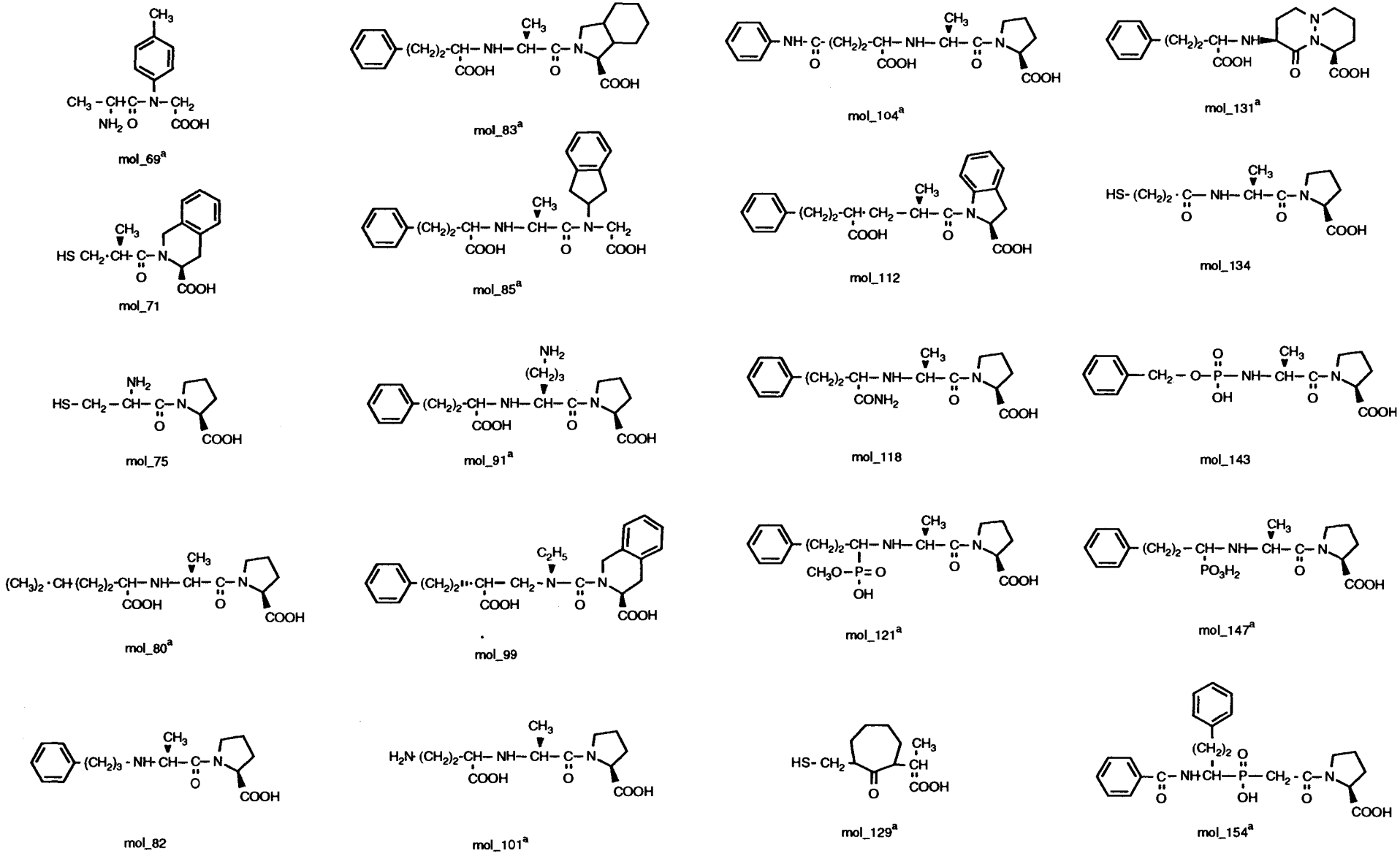
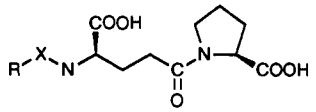
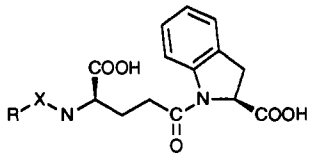
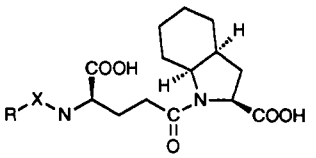


Figure 2. Molecules included in ACE test set of diverse inhibitors. An a designation means that ambiguous chiral centers were modeled in the S configuration.

Table II. ACE Diverse Test Set

molecule	pIC_{50} (M)	ΔG (kcal)	ref
mol_69	5.59	80.25	5
mol_71	7.70	211.32	5
mol_75	7.20	213.14	5
mol_80	8.59	210.56	5
mol_82	5.24	80.78	5
mol_83	8.66	199.87	5
mol_85	7.40	206.07	5
mol_91	8.66	203.71	5
mol_99	8.19	204.04	5
mol_101	6.49	210.37	5
mol_104	8.28	212.67	5
mol_112	8.32	200.35	5
mol_118	4.59	79.29	5
mol_121	6.36	201.77	5
mol_129	7.28	213.87	32
mol_131	7.25	204.57	34
mol_134	5.59	202.71	5
mol_143	7.39	213.44	32
mol_147	7.29	203.90	5
mol_154	7.14	200.34	5

Table III. ACE Carboxylate Test Set

molecule	R	X	pIC_{50} (M)	ΔG (kcal)
				
COO_23A	Z[a]	Lys	7.92	186.43
COO_23E	Z	Pro	5.60	192.81
COO_23J	Z	Ala	6.41	202.40
				
COO_24A	Z	Lys	5.60	183.93
COO_24C	Z	Phe	9.60	189.53
				
COO_25A	Z	Lys	8.46	192.42
COO_25E	Z	Pro	7.82	196.52
COO_26A	H	Lys	7.92	185.63
COO_26C	H	Phe	7.00	193.46
COO_26H	H	D-Phe	6.66	188.58

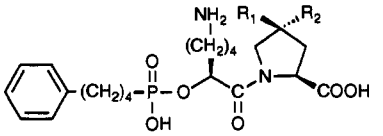
[a] Z = benzyloxycarbonyl.

charge set models. A test data set of 11 diverse thermolysin inhibitors[49,53,57] (Table VII) was used to assess the predictive power of the resulting models.

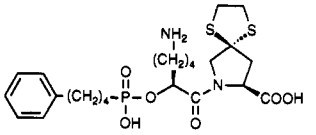
Results

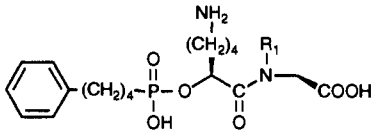
A. CoMFA of ACE Inhibitors. The results of the various CoMFAs of ACE inhibitors are presented in summarized form in Table VIII. The analysis based on CoMFA steric and electrostatic fields yielded results consistent with the previous work[6] (Figure 3), and as expected, produced a similar "predictive" r^2 for the series of 20 diverse inhibitors (0.46) (Figure 4). These results also support the systematic error hypothesis noted previously for the phosphorous-based inhibitors ($r^2_{pred} = 0.42$) (Figure 5) since all were overpredicted. The addition of the HOMO field to the analysis, however, produced a more significant ($r^2_{cross} = 0.722$) model for the training set (Figure 6) and more predictive ($r^2_{pred} = 0.59$) correlation for the same diverse test set (Figure 7). Furthermore, the systematic error would seem to be corrected since the "predictive" r^2 for the phosphate-based inhibitors increased significantly to 0.69 (Figure 8). The combination of CoMFA fields and desolvation energies as regressors yielded generally poorer cross-validated ($r^2_{cross} = 0.646$) results when compared to the CoMFA/HOMO models and predictive results similar to those of the CoMFA-only model for the diverse, phosphate-based, and thiol-based sets of inhibitors (Table IX). The systematic overpredictive trend of the CoMFA-only model for the phosphate-based inhibitors is not significantly altered (Figure 9), while a slight improvement in the predictive ability of the model for the carboxylate test set compared to the CoMFA/HOMO model is noted (Table IX). The most statistically significant model was achieved with the regressor combination of CoMFA steric and electrostatic fields, MOPAC HOMO fields, and Delphi desolvation energies ($r^2_{cross} = 0.729$). This model was superior to all others with respect

Table IV. ACE Phosphate Test Set



molecule	R_1	R_2	pIC_{50} (M)	ΔG (kcal)
SQ29852	H	H	7.44	195.22
SQ29852_2A	H	cyclohexyl	8.49	195.16
SQ9852_2B	cyclohexyl	H	9.31	195.51
SQ29852_2P	H	phenyl	6.47	221.65
SQ29852_2Q	phenyl	H	9.88	200.55
SQ29852_2R	benzyl	H	9.05	195.55
SQ29852_2S	H	phenylthio	9.11	199.06
SQ29852_2T	phenylthio	H	9.28	190.81
SQ29852_2U	H	methylthio	9.28	198.42
SQ29852_2V	methylthio	H	8.72	194.71
SQ29852_2W	methylthio	methylthio	8.59	193.91
SQ29852_2Y	H	OH	9.06	198.74
SQ29852_2Z	OH	H	7.80	216.70
SQ29852_2X			9.20	197.14





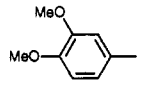
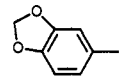
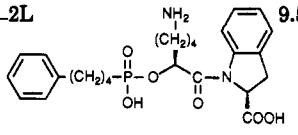
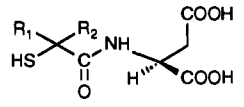
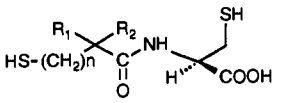
molecule	R_1	pIC_{50} (M)	ΔG (kcal)
SQ29852_2E		8.97	187.08
SQ29852_2G		9.54	192.30
SQ29852_2H		9.94	195.39
SQ29852_2K		9.53	197.72
SQ29852_2L		9.53	193.83

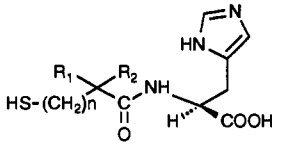
Table V. ACE Thiol Test Set



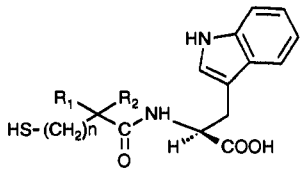
molecule	R_1	R_2	pIC_{50} (M)	ΔG (kcal)
THIOL_2	CH_3	CH_3	2.46	210.11
THIOL_4	H	H	4.11	215.85



molecule	R_1	R_2	n	pIC_{50} (M)	ΔG (kcal)
THIOL_5	H	H	0	4.72	213.57
THIOL_7Aa	CH_3	H	0	5.10	210.02
THIOL_7Bb	CH_3	H	0	3.27	211.21
THIOL_9	CH_3	CH_3	0	2.14	206.44
THIOL_10	H	H	1	4.17	211.60



molecule	R_1	R_2	n	pIC_{50} (M)	ΔG (kcal)
THIOL_12	H	H	0	3.59	220.29
THIOL_14	H	H	1	3.64	220.15



molecule	R_1	R_2	n	pIC_{50} (M)	ΔG (kcal)
THIOL_18	H	H	0	6.12	222.07
THIOL_20A	CH_3	H	0	7.19	215.79
THIOL_20B	H	CH_3	0	6.44	214.92
THIOL_22	CH_3	CH_3	0	4.42	207.77
THIOL_27	H	H	1	5.34	215.88
THIOL_28	H	H	1	<3.00	215.24
THIOL_30A	CH_3	H	1	7.46	214.44
THIOL$_{\beta\beta}$0B	H	CH_3	1	4.48	211.04

to predictive ability for the test sets of diverse, phosphate-based, and thiol-based inhibitors (Table IX).

B. CoMFA of Thermolysin Inhibitors. The results of the CoMFA of thermolysin inhibitors using PM3 ligand charges and the MULTIFIT derived alignment are summarized in Table X. Since these analyses were based on the same alignment rule as previously reported, it was a little surprising to discover that the CoMFA-only analyses yielded poorer cross-validated and conventional statistical results (r^2_{cross} = 0.536 and r^2 = 0.851) (Figure 10). This is further confounded by the fact that the "predictive" r^2 for the test set of 11 diverse inhibitors increased to 0.44 (Figure 11) compared to 0.29 for the original model. The addition of the HOMO field (Figure 12) produced an effect similar in scope to that of the addition of indicator variables in the original analyses of this dataset. Although the statistical significance of the model increased slightly with the inclusion of the HOMO field, the predictive ability of the model was decreased (Figure 13). Desolvation energies, contributing 11% to the overall model including CoMFA steric and electrostatic fields and ΔG values as regressors, did not affect either the statistical significance (Table X) or the predictability (r^2_{pred} = 0.42) of the CoMFA-only model. It is interesting (anomalous) to note that the ΔG-only analysis yielded a "predictive" r^2 value of 0.43 for the 11 diverse inhibitors, yet was statistically insignificant (r^2 = 0.121). The combination of CoMFA fields, MOPAC HOMO fields, and ΔG values did not improve the statistical significance nor the predictive ability of the model with respect to the CoMFA-only model.

The results of the CoMFA of thermolysin inhibitors using Kollman/PM3 charges and complementary receptor field alignment are presented in summarized form in Tables XI and XII. In the initial implementation of complementary receptor field alignment, the results were disappointing. In Table XI, the automatic region defi-

Table VI. Molecules Included in Thermolysin Training Set

molecule	pK_i (M)	ΔG (kcal)	ref(s)
ACE_OHLEU_AGNH2	2.47	19.83	57
BZSAG	6.12	87.87	47
C6PCLTNME	7.28	79.63	48
C6PLTNME	8.82	77.15	48
C6POLTNME	5.84	82.20	48
CBZPHE	3.29	76.52	59
CH3COCH2CO_FAGNH2	2.51	24.26	57
CH3O2S_FAGNH2	0.52	34.40	57
CHO_OHLEU_AGNH2	2.47	20.55	57
CLTZNCRYS	7.47	197.29	41
DAH50	7.96	187.73	2
DAH51	6.22	198.84	2
DAH52	5.55	200.41	2
DAH53	6.66	183.15	2
DAH54	5.77	195.17	2
DAH55	2.42	196.92	2
HOCH2CO_FAGNH2	2.54	23.08	57
NHOHBZMAGNA	6.37	84.64	57
NHOHBZMAGNH2	6.18	88.53	57
NHOHBZMAGOH	6.18	87.00	57
NHOHBZMOET	4.70	78.26	57
NHOHIBMAGNH2	6.32	88.27	57
NHOHLEU	3.72	73.75	57
NHOHMALAGNH2	2.96	91.58	57
OHBZMAGNH2	3.38	91.04	57
P(OPHE)(OME)LEUNH2	0.52	20.09	57
PAAOH	4.06	265.11	57
PHOSPHORAMIDON	7.55	195.59	40
PLEUNH2	4.10	249.09	57
PNHET	0.52	277.34	57
PO3_FAGNH2	5.59	257.44	57
PPHEOH	4.14	260.08	57
P_ILE_AOH	6.44	259.91	57
(R)THIORPHAN	5.64	214.09	50,51
SO2P_FAGNH2	5.16	240.37	57
SO3_FAGNH2	2.37	101.52	57
(S)THIORPHAN	5.74	214.24	50,51
Z-D-APOLA	4.62	202.65	52
Z-D-FPLA	6.32	193.44	52
Z-D-FPOLA	4.52	198.79	52
Z-D-LPOLA	4.38	200.75	52
Z(NH)GLNH2	3.42	20.08	57
Z(NH)GLNHOH	5.57	205.07	57
ZALA	6.07	187.12	54
ZAPOLA	5.74	201.86	52
ZFPLAZNCRYS	10.17	198.44	43
ZFPOLA	7.35	201.26	50
ZG-D-LNHOH	4.32	214.23	55
ZGG-D-LNHOH	3.60	215.48	55
ZGGLNHOH	4.41	215.09	55
ZGGNHOH	3.03	227.00	55
ZGLNH2	1.68	19.05	55
ZGLNHOH	4.89	215.29	55
ZGLNMEOH	2.65	84.70	55
ZGLY	6.39	189.98	54
ZGPCLLZNCRYS	6.74	194.74	56
ZGPLA	7.78	199.11	50
ZGPLLZNCRYS	8.04	196.68	43,53
ZGPOLA	4.89	206.31	52
ZGPOLLZNCRYS	5.05	198.44	44,53
ZLPOLA	6.17	197.53	52

Table VII. Molecules Included in Thermolysin Test Set

molecule	pKi (M)	ΔG (kcal)	ref(s)
ZGPOLF	4.27	194.93	53
ZGPOLG	3.64	207.22	53
ZGPLF	7.12	191.80	53
ZGPLG	6.57	204.09	53
ZGPLNH2	6.12	78.90	53
ZGPOLNH2	3.18	80.88	53
ZLGNH2	2.51	16.78	57
ZFGNH2	3.46	18.73	57
PLFOH	7.72	246.50	57
ZYGNH2	3.66	19.98	57
β PPPhe	2.79	76.74	49

Table VIII. Summary of Results from CoMFA of 68 ACE Inhibitors Based on Alignment Rule Derived from Active-Analog Approach (PM3 Charges on Ligands, Automatic Region Definition)

regressors	r^2_{cross}	r^2	SEE	F-test	cont
CoMFA(S,E)	0.656(7)	0.902(7)	0.711	79.240	78,22
HOMO	0.190(7)	0.411(7)	1.747	5.988	100
ΔG	0.000(1)	0.053(1)	2.111	3.705	100
CoMFA, HOMO	0.722(5)	0.848(5)	0.871	69.435	56,13,31
CoMFA, ΔG	0.646(5)	0.860(5)	0.837	76.217	78,15,7
CoMFA, HOMO, ΔG	0.729(7)	0.869(7)	0.823	57.043	55,8,29,8

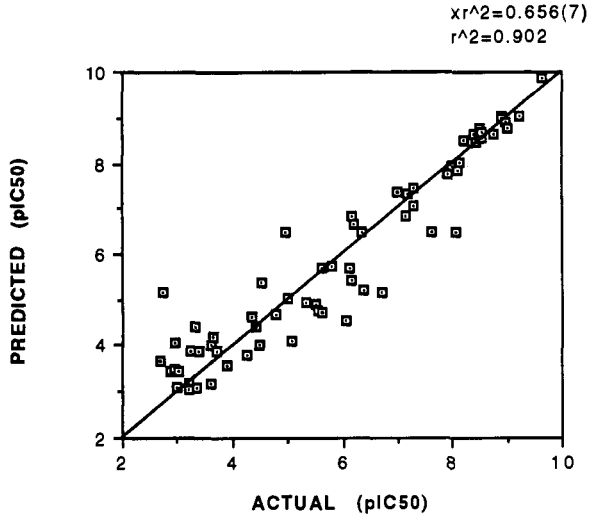


Figure 3. Predicted vs actual pIC_{50} for the CoMFA of 68 ACE inhibitors incorporating only CoMFA (PM3) parameters. The predictive model was derived using seven principal components yielding a cross-validated $r^2 = 0.656$.

nition protocol was used. In the CoMFA-only model, the cross-validated r^2 was 0.454 representing yet another decrease in the significance of the correlation as compared to the original model while the predictive ability with respect to the 11 diverse inhibitors ($r^2_{pred} = 0.21$) was comparable to that of the original model. In the analysis which included the HOMO field, it is interesting to note that the component number (2) which produced the highest cross-validated r^2 was completely comprised of the HOMO field, while a lesser component model (1) exhibited equal contributions from all fields. Predictions for the 11 diverse inhibitors based on this latter model revealed a measure of predictability ($r^2_{pred} = 0.29$) similar to that of the CoMFA-only model presented above.

In the analyses presented in Table XII, the same alignment rule was utilized while the region was defined as spatial positions and point charges of the atoms in the receptor. Once again a decrease in the statistical significance of all model was noted. Conversely to the ACE analyses and also noted in the above thermolysin analyses,

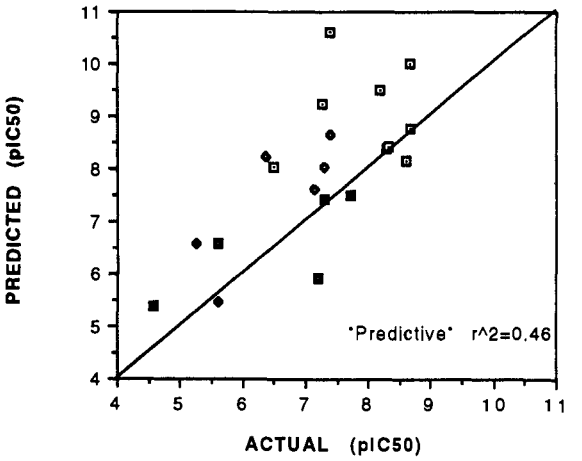


Figure 4. Predicted vs actual pIC_{50} for the test data set of 20 diverse ACE inhibitors incorporating only CoMFA (PM3) parameters. The predictive model was derived using seven principal components yielding a cross-validated $r^2 = 0.656$. "Predictive" $r^2 = 0.46$. Symbols are as follows: ⊡, PredCOOH-(SE); ◆, $PredNH_2$(SE); ■, PredSH(SE); ◆, $PredPO_4$(SE).

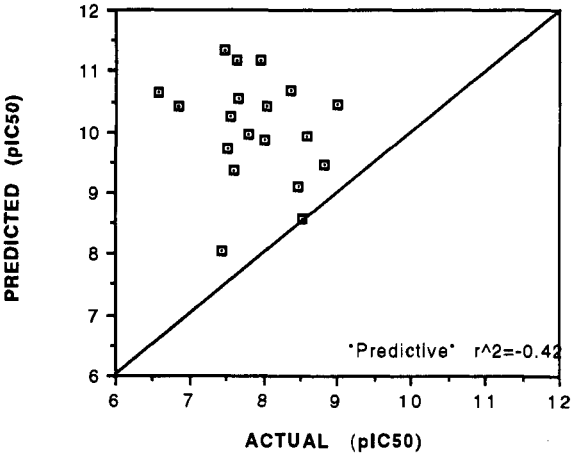


Figure 5. Predicted vs actual pIC_{50} for the test data set of 19 phosphate-based ACE inhibitors incorporating only CoMFA (PM3) parameters. The predictive model was derived using seven principal components yielding a cross-validated $r^2 = 0.656$. "Predictive" $r^2 = -0.42$.

the addition of the HOMO field was detrimental to the overall model while the predictability increased ($r^2_{pred} = 0.40$ for both models) with respect to the automatically-defined region models. Since the charge calculation methods employed in these analyses are very well parameterized for the particular application (PM3 for small molecules and Kollman for macromolecules) and neither is particularly well suited for the other, there is concern over this apparent disparity. In order to address this, a method capable of calculating partial atomic charges for both ligand and receptor was employed in the following analyses.

The results of the CoMFA of thermolysin inhibitors using Gasteiger and Hueckel charges and complementary receptor field alignment are presented in summarized form in Tables XIII and XIV. The initial series of analyses utilized the automatic region definition protocol and yielded more promising results ($r^2_{cross} = 0.518$). As usual, the steric field was dominant in the absence of the HOMO field. Even more promising were the results obtained if the complementary receptor field-fitted molecules were not subjected to reoptimization. But, analyses employing this type of alignment rule must be regarded with caution

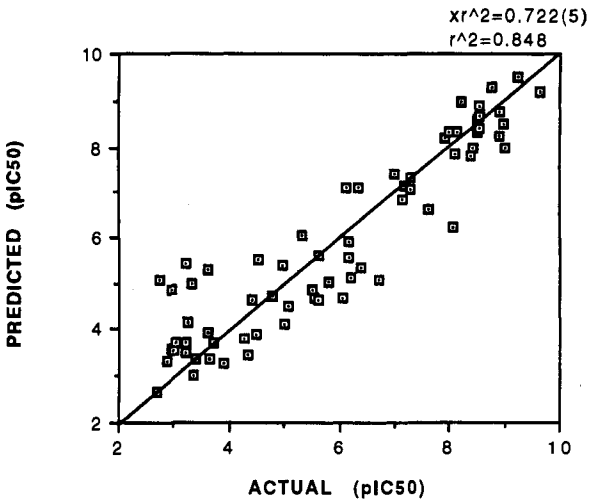


Figure 6. Predicted vs actual pLC_{50} for the CoMFA of 68 ACE inhibitors incorporating CoMFA (PM3) and HOMO (PM3) parameters. The predictive model was derived using five principal components yielding a cross-validated $r^2 = 0.722$.

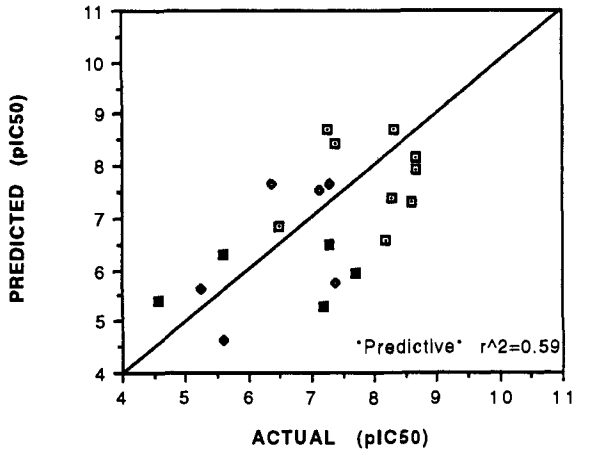


Figure 7. Predicted vs actual pIC_{50} for the test data set of 20 diverse ACE inhibitors incorporating CoMFA (PM3) and HOMO (PM3) parameters. The predictive model was derived using five principal components yielding a cross-validated $r^2 = 0.722$. "Predictive" $r^2 = 0.59$. Symbols are as follows: ⊡, PredCOOH; ◆, PredNH_2; ■, PredSH; ◆, PredPO_4.

since energetically-unfavorable and geometrically-unreasonable conformations are often generated. These results indicate that charge disparity of the two different methods used may have contributed to the apparent lack of statistical significance in the PM3/Kollman analyses. It is possible that a higher level computational technique suited for the calculation of partial atomic charges as well as molecular orbitals for small and moderately-sized molecules could aid in the development of more statistically successful and physically descriptive CoMFA QSAR's.

As reported above in the PM3/Kollman analyses, the statistical significance of the receptor-defined region model decreased compared to analyses using automatically-defined regions and the identical alignment rule. It is of particular interest to note that the predictive ability of the receptor-defined region models was, however, greater than that of automatically-defined region models. This suggests that for statistically-comparable models, the receptor-defined region analysis results would be preferred.

Discussion

Previously,[6] concern was raised over the monodentate versus bidentate binding modes of the ACE and thermolysin inhibitors. With respect to the ACE series of compounds, the Mayer geometry assumes all molecules bind to the zinc in a monodentate manner. The variations in the predictive abilities of the models with respect to phosphate- and carboxylate-based test sets may be attributable to inhomogeneity of binding mode as reflected in the alignment rule. It is clear, however, that the physicochemical nature of the zinc–ligand interaction was not adequately characterized as evidenced by the original CoMFA-only model. Indicator variables proved to be a useful rectification, but introduced additional ambiguities to the QSAR in theory. As presented herein, it was possible to replace the binary indicator variables with actual

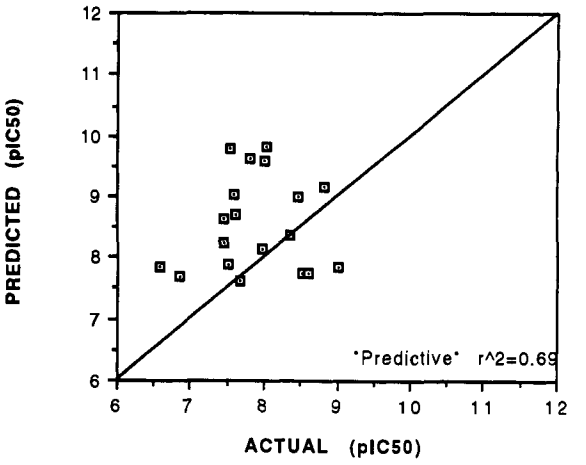


Figure 8. Predicted vs actual pIC_{50} for the test data set of 19 phosphate-based ACE inhibitors incorporating CoMFA (PM3) and HOMO (PM3) parameters. The predictive model was derived using five principal components yielding a cross-validated $r^2 = 0.722$. "Predictive" $r^2 = 0.69$.

Table IX. Predictive r^2 Values for ACE Test Sets[a]

test set	model 1	model 2	model 3	model 4	model 5
20 diverse	0.53	0.46	0.59	0.66	0.66
19 phos-based	−0.40	−0.42	0.69	−0.17	0.71
17 thiol-based	0.20	0.21	0.30	0.16	0.33
10 carb-based	0.60	0.41	−0.38	−0.32	−0.16

[a] Model 1 stands for the CoMFA-only model (Gasteiger and Marsili charges).[6] Model 2 stands for the CoMFA-only model (PM3 charges). Model 3 stands for the CoMFA, HOMO model. Model 4 stands for the CoMFA, ΔG model. Model 5 stands for the CoMFA, HOMO, ΔG model.

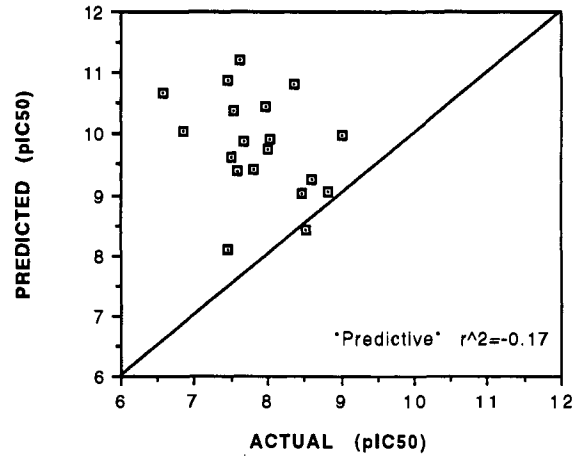


Figure 9. Predicted vs actual pIC_{50} for the test data set of 19 phosphate-based ACE inhibitors incorporating CoMFA (PM3) and ΔG (Delphi) parameters. The predictive model was derived using five principal components yielding a cross-validated $r^2 = 0.646$. "Predictive" $r^2 = -0.17$.

Table X. Summary of Results from CoMFA of 61 Thermolysin Inhibitors Based on Alignment Rule Derived from Crystal Structures (PM3 Charges on Ligands, Automatic Region Definition)

regressors	r^2_{cross}	r^2	SEE	F-test	cont	r^2_{pred}
CoMFA(S,E)	0.536(5)	0.851(5)	0.832	63.056	70,30	0.44
HOMO	−0.064(1)	0.053(1)	2.029	3.327	100	N/A
ΔG	0.052(1)	0.121(1)	1.955	8.116	100	0.43
CoMFA, HOMO	0.596(9)	0.921(9)	0.629	66.482	52,24,24	0.17
CoMFA, ΔG	0.540(5)	0.842(5)	0.857	58.803	57,32,11	0.42
CoMFA, HOMO, ΔG	0.557(6)	0.834(6)	0.888	45.271	43,25,25,7	0.42

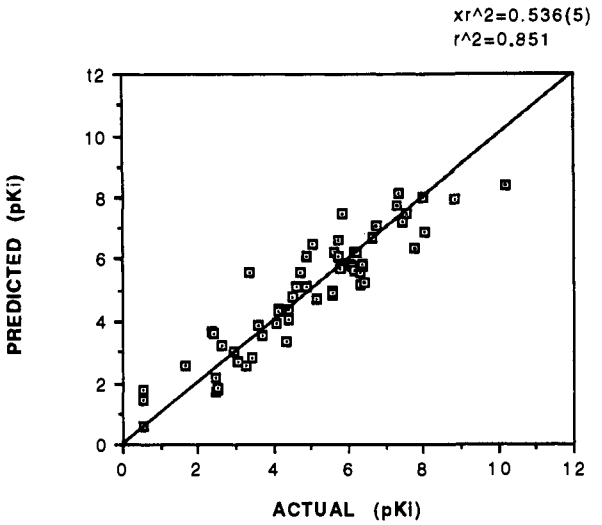


Figure 10. Predicted vs actual pK_i for the CoMFA of 61 thermolysin inhibitors incorporating only CoMFA (PM3) parameters. The predictive model was derived using five principal components yielding a cross-validated $r^2 = 0.536$.

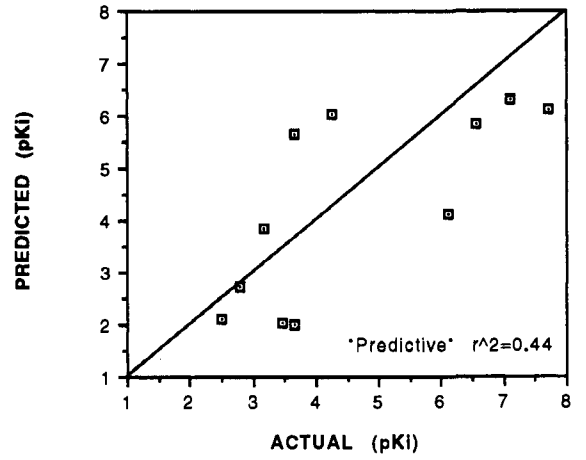


Figure 11. Predicted vs actual pK_i for the test data set of 11 diverse thermolysin inhibitors incorporating only CoMFA (PM3) parameters. The predictive model was derived using five principal components yielding a cross-validated $r^2 = 0.536$. "Predictive" $r^2 = 0.44$.

molecular orbital field indices and produce QSAR models with comparable predictive ability. A better alignment rule for the ACE series based on a more physically-realistic active-site geometry as suggested by Hausin and Codding[59] has been found,[60] and the impact of realignment on the predictability of ACE inhibitors is under investigation. It is entirely possible that CoMFA studies based on a more reliable alignment rule may overcome many of the difficulties encountered, although the results with thermolysin inhibitors would argue otherwise.

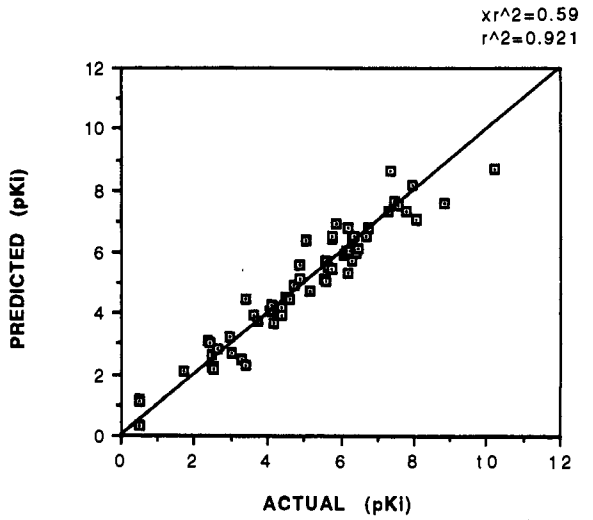


Figure 12. Predicted vs actual pK_i for the CoMFA of 61 thermolysin inhibitors incorporating CoMFA (PM3) and HOMO (PM3) parameters. The predictive model was derived using nine principal components yielding a cross-validated $r^2 = 0.596$.

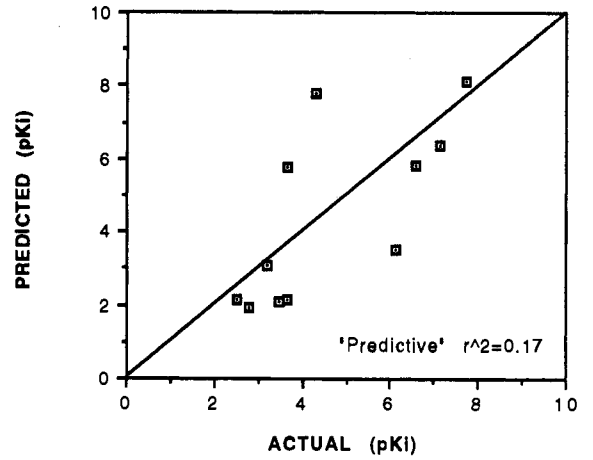


Figure 13. Predicted vs actual pK_i for the test data set of 11 diverse thermolysin inhibitors incorporating CoMFA (PM3) and HOMO (PM3) parameters. The predictive model was derived using nine principal components yielding a cross-validated $r^2 = 0.596$. "Predictive" $r^2 = 0.17$.

Table XI. Summary of Results from CoMFA of 61 Thermolysin Inhibitors Based on Alignment Rule Derived from Field-Fit Minimization to Complementary Receptor Fields (PM3 Charges on Ligand and Kollman Charges on Receptor, Automatic Region Definition)

regressors	r^2_{cross}	r^2	SEE	F-test	cont	r^2_{pred}
CoMFA only	0.454(5)	0.874(5)	0.768	76.014	63,37	0.21
CoMFA, HOMO	0.357(2)	0.674(2)	1.201	59.891	0,0,100	0.29
*CoMFA, HOMO	0.341(1)	0.521(1)	1.444	64.098	30,37,33	

The issue of molecular alignment has been addressed with the analyses of thermolysin inhibitors presented herein. In fitting ligands to the complementary steric and electrostatic fields computed for the known structure of the molecular binding domain of the receptor, ambiguities introduced into the molecular alignment by the zinc are minimized as specific molecular mechanical parameters are no longer necessary to explain the zinc–ligand interaction. The steric and charge properties of the zinc atom are incorporated into the overall complementary receptor fields. The strength of the zinc–ligand interaction is then represented by the orientation of the ligand molecule

Table XII. Summary of Results from CoMFA of 61 Thermolysin Inhibitors Based on Alignment Rule Derived from Field-Fit Minimization to Complementary Receptor Fields (PM3 Charges on Ligand and Kollman Charges on Receptor, Receptor Region Definition)

regressors	r^2_{cross}	r^2	SEE	F-test	cont	r^2_{pred}
CoMFA only	0.371(3)	0.708(3)	1.147	45.975	76,24	0.40
CoMFA, HOMO	0.354(4)	0.718(4)	1.136	35.693	74,24,2	0.40

Table XIII. Summary of Results from CoMFA of 61 Thermolysin Inhibitors Based on Alignment Rule Derived from Field-Fit Minimization to Complementary Receptor Fields (Gasteiger and Huckel Charges on Ligand and Receptor, Automatic Region Definition)

regressors	r^2_{cross}	r^2	SEE	F-test	cont	r^2_{pred}
CoMFA only	0.518(3)	0.747(3)	1.067	56.059	65,35	0.30
Note: Alignment rule-field-fit with no subsequent reoptimization						
CoMFA only	0.546(3)	0.755(3)	1.050	58.552	68,32	

Table XIV. Summary of Results from CoMFA of 61 Thermolysin Inhibitors Based on Alignment Rule Derived from Field-Fit Minimization to Complementary Receptor Fields (Gasteiger and Huckel Charges on Ligand and Receptor, Receptor Region Definition)

regressors	r^2_{cross}	r^2	SEE	F-test	cont	r^2_{pred}
CoMFA only	0.347(1)	0.463(1)	1.527	50.951	53,47	0.34

within the receptor fields during the field-fit minimization routine since the molecule is positioned as such to minimize the steric and electrostatic field-fit energy terms in the overall molecular mechanical force-field equation. This procedure appears to be a logical extension of docking algorithms[61-63] in which rigid molecules (both ligand and receptor) are fit together on the basis of topological complementarity. Experimental support for considering the receptor as a rigid aggregate during the generation of complementary fields comes from the examination of the crystal structures of enzyme–inhibitor complexes where the enzyme presents essentially the same conformation despite the large variations in inhibitor structures.[64] The present implementation seems superior to the above-mentioned docking algorithms in that ligands are fully geometry optimized in the molecular environment of the enzyme active site.

The technique of complementary-receptor-field alignment is, however, not without caveat. As mentioned above, charge disparity between receptor and ligand appears to diminish the statistical significance of the overall model. Additionally, the charge distribution across the receptor active-site upon the binding of a given ligand is likely to be altered. Gasteiger–Hueckel and Kollman charge computation techniques are not sensitive to intermolecular interactions and cannot be used to study this phenomenon. A semiempirical (i.e. MOPAC/PM3) method would appear to be the most viable alternative, but the computational burden precludes routine usage (a single-point, 1SCF, MOPAC energy calculation on a 124 heavy-atom system requires approximately 10 CPU hours on an IBM Model 560 RISC 6000 workstation). At present, suitable alternatives for the computation of partial atomic charges and molecular orbital characteristics are being sought.

Nevertheless, it does appear that this treatment would seem to be a more efficient method for determining the molecular alignment than the previously suggested alternative, namely the use of a molecular mechanics force field including parameters for metals (YETI).[65] The simple fact that YETI requires certain *a priori* knowledge (coordination number of the metal) creates a combinatorial problem (calculating the geometry of every ligand with respect to every possible coordination state of the zinc). It is possible that multiple conformers of ligands could be included in an initial QSAR analysis and the "best conformation" could be selected on how well it fits the derived model.[66] This approach would surely generate a predictive model, yet it seems inefficient. The use of complementary receptor fields which incorporate the metal atom as a point charge and volume contribution to the overall field as field-fit templates appears to be the most direct and relevant solution to former alignment rule ambiguities.

Typically, CoMFA regions are defined as a matrix whose intersections provide coordinates on which to place a given probe atom and whose boundaries extend a given distance beyond those of the ligands in the analysis. An alternative region description (atomic coordinates of the receptor) has been examined herein. At first glance, it would seem that this description most closely replicates that of the pharmacological condition. However, to date this has not yielded very promising results with respect to the statistical significance of the models which were attained. The long-distance receptor ligand interactions are probably as well represented with this technique as with automatically defined regions, but the immediate interactions for some of the ligands in the analysis may be underestimated. Automatically defined regions contain points which fully permeate the entire set of ligands (although many points are eliminated due to commonality within ligands), while all of the points of a receptor-defined region only fully probe ligands which completely fill the volume of the molecular binding domain. This difference would be most readily apparent in the cases of "undersized" ligands where many of the immediate interactions would be underdetermined.

Evidence has been presented which suggests that CoMFA steric and electrostatic fields alone do not sufficiently characterize ligand–receptor interactions with respect to ACE inhibitors. While concluding that the original QSAR model was improved with the addition of MOPAC-derived HOMO field, and better yet with the addition of HOMO fields and desolvation energies, the interaction still remains undercharacterized as evidenced by the low predictive r^2 values for certain test sets of molecules. Additional regressors currently under consideration would include better characterization of the free energy of solvation (desolvation), hydrophobic, and entropic terms. It is possible to include solvation and hydrophobic characteristics[67] of ligands as interaction values on a matrix (CoMFA region), in lieu of total values. The inclusion of simple entropic terms (degrees of freedom, etc.) into the regression equation has been previously addressed and did not yield promising results. Obviously, a better entropic descriptor is desired.

Conclusions

Three potentially useful refinements (HOMO fields as CoMFA fields, field-fit minimization to complementary active-site fields, and active-site-defined CoMFA regions) of the CoMFA technique have been examined. Perhaps, the most innovative of which is the use of MOPAC-derived molecular orbital fields as additional regressors. The success of this application directly confirms the conclusion drawn from the use of indicator variables in the previously

published analyses.[6] Moreover, it represents a potential method by which the biological potency of classes of ACE inhibitors with novel zinc ligands can be predicted, a limitation of the indicator–variable approach.

The most important parameter in a CoMFA study is the relative alignment of molecules within the region when the steric and electrostatic fields are sampled—the alignment rule (anonymous, SYBYL Theory Manual). Variations in molecular superimposition can arise as a direct result of the method chosen for alignment and further as a function of the conformational flexibility of the molecules under study. As previously discussed, it is possible to include multiple conformers in the analysis.[66] Ideally, conformational flexibility should be an implicit component of the alignment rule. For example, an alignment rule based on experimental conformational data of active analogs (as in the thermolysin inhibitor study) allows for a significant reduction of the degrees of freedom of the problem. However, in the absence of direct experimental data regarding the binding mode and active conformation (as in the ACE inhibitor study), an alignment technique which stresses optimal relative intermolecular conformational similarity (such as the alignment rule based on the active–analog approach utilized herein) would seem to be a plausible alternative.

Often, the availability and quality of molecular mechanical parameters dictates the nature of the resulting alignment rules and, therefore, the overall analysis. A possible solution to this problem has been presented in the form of complementary receptor fields. The use of these fields as templates for field-fit minimization allows for the nonparameterized atoms (i.e. Zn) to be incorporated as a van der Waal's radius contribution to overall steric receptor field and as a point-charge contribution to the overall electrostatic receptor field.

Finally, typical CoMFA regions are defined as a lattice of regularly-spaced probe atoms of uniform characteristics (i.e. sp^3 carbon with a charge of +1 at every lattice intersection) extending a given distance away from the ligands included in the training set. The availability of crystallographic data for the thermolysin receptor presented the unique opportunity of using the coordinates, atom type, and partial atomic charge information in the definition of the CoMFA region. In this first implementation of receptor-defined regions, the results were promising.

In general, the results presented herein provide additional support for the CoMFA methodology as a valuable 3-D QSAR tool. The most beneficial aspect of any QSAR model must be the predictive ability of the model with respect to novel compounds. While the original application of this method exhibited a moderate degreee of predictability for ACE inhibitors, we have increased both the statistical significance and the predictability of the model with the inclusion of more physically-descriptive parameters.

Acknowledgment. The authors gratefully acknowledge support for this research from National Institutes of Health grant GM24483.

Supplementary Material Available: The atomic coordinates and partial charges of all molecules (145 pages). Ordering information is given on any current masthead page.

References

(1) Petrillo, W. W.; Trippodo, N. C.; DeForrest, J. M. In *Annual Reports in Medicinal Chemistry*; Robertson, D. W., Ed.; Academic: New York, 1989; Vol. 25, pp 51–60.
(2) Hangauer, D. G. In *Computer-Aided Drug Design: Methods and Applications*; Perun, T. J., Propst, C. L., Ed.; Marcel Dekker: New York, 1989; pp 253–295.
(3) Saunders, M. R.; Tute, M. S.; Webb, G. A. A Theoretical Study of Angiotensin-Converting Enzyme Inhibitors. *J. Comput.-Aided Mol. Des.* **1987**, *1*, 133–142.
(4) Mayer, D.; Naylor, C. B.; Motoc, I.; Marshall, G. R. A Unique Geometry of the Active Site of Angiotensin-Converting Enzyme Consistent with Structure-Activity Studies. *J. Comput.-Aided Mol. Des.* **1987**, *1*, 3–16.
(5) Wyvratt, M. J.; Patchett, A. A. Recent Developments in the Design of Angtiotensin-Converting Enzyme Inhibitors. *Med. Res. Rev.* **1985**, *5*, 483–531.
(6) DePriest, S. A.; Mayer, D.; Naylor, C. B.; Marshall, G. R. 3D-QSAR of Angiotensin-Converting Enzyme and Thermolysin Inhibitors: A comparison of CoMFA Models Based on Deduced and Experimentally Determined Active Site Geometries. *J. Am. Chem. Soc.* **1983**, in press.
(7) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
(8) Marshall, G. R.; Cramer, R. D., III. Three Dimensional Structure-Activity Relationships. *Trends Pharmacol. Sci.* **1988**, *9*, 285–289.
(9) Kim, K. H.; Martin, Y. C. Direct Prediction of Linear Free Energy Substituent Effects from 3D Structures Using a Comparative Molecular Field Analysis (CoMFA) Approach. 1. Electronic Effects of Substituted Benzoic Acids. *J. Org. Chem.* **1991**, *56*, (8), 2723–2729.
(10) Kim, K. H. 3D Quantitative Structure-Activity Relationships: Description of Electronic Effects Directly from 3D Structures Using a GRID-Comaparative Molecular Field Analysis (CoMFA) Approach. *Quant. Struct.-Act. Relat.* **1992**, *11*, 127–134.
(11) Kim, K. H. 3D Quantitative Structure-Activity Relationships: Investigation of Steric Effects with Descriptors from 3D Structures Using a Comparative Molecular Field Analysis (CoMFA) Approach. *Quant. Struct.-Act. Relat.* **1992**, *11*, 453–460.
(12) The program SYBYL is available from Tripos Associates, 1699 S. Hanley Rd., St. Louis, MO.
(13) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. An All Atom Force Field for the Simulation of Proteins and Nucleic Acids. *J. Comput. Chem.* **1986**, *7*, 230.
(14) Purcell, W. P.; Singer, J. A. A Brief Review and Table of Semiempirical Parameters used in the Hueckel Molecular Orbital Method. *J. Chem. Eng. Data* **1967**, *12*, 235–246.
(15) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods. 1. Method. *J. Comput. Chem.* **1989**, *10* (2), 209–220.
(16) MOPAC version 5.0 is available from Quantum Chemistry Program Exchange no. 455.
(17) Wold, S.; Ruhe, A.; Wold, H.; Dunn, W. J. The Covariance Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. *SIAM J. Sci. Stat. Comput.* **1984**, *5* (3), 735–743.
(18) Cramer, R. D.; Bunce, J. D.; Patterson, D. E. Cross-validation, Bootstrapping, and Partial Least Squares Compared with *Multiple Regression in Conventional* QSAR Studies. *Quant. Struct.-Act. Relat.* **1988**, *7*, 18–25.
(19) Klapper, I.; Hagstrom, R.; Fine, R.; Sharp, K.; Honig, B. Focusing on Electric Fields in the Active Site of CuZn Superoxide Dismutase. *Proteins* **1986**, *1*, 47.
(20) Gilson, M.; Honig, B. Calculations of Electrostatic Potentials in an Enzyme Active Site. *Nature* **1987**, *330*, 84.
(21) The products Delphi and Insight II version 2.20 are available from Biosym Technologies, Inc. 9685 Scranton Road, San Diego, CA 92121–2777.
(22) Hassall, C. H.; Kroehn, A.; Moody, C. J.; Thomas, W. A. The Design of a New Group of Angiotensin Converting Enzyme Inhibitors. *FEBS Lett.* **1982**, *147*, 175–179.
(23) Hassall, C. H.; Kroehn, A.; Moody, C. J.; Thomas, W. A. The Design and Synthesis of New Triazolo-, Pyrazolo-, and Pyridazo-pyridazine Derivatives as Inhibitors of Angiotensin-Converting Enzyme. *J. Chem. Soc., Perkin Trans.* **1984**, *1*, 155–164.
(24) Kim, D. H.; Guinosso, C. J.; Buzby, G. C., Jr.; Herbst, D. R.; McCaully, R. J.; Wicks, T. C.; Wendt, R. L. (Mercaptopropanoyl)-indoline-2-carboxylic Acids and Related Compounds as Potent Angiotensin-Converting Enzyme Inhibitors and Antihypertensive Agents. *J. Med. Chem.* **1983**, *26*, 394–403.
(25) Galardy, R. E.; Kontoyiannidou-Ostrem, V.; Kartylewicz, Z. P. Inhibition of Angiotensin-Converting Enzyme by Phosphonic Amides and Phosphonic Acids. *Biochemistry* **1983**, *22*, 1990–1995.
(26) Cushman, D. W.; Cheung, H. S.; Sabo, E. F.; Ondetti, M. A. Design of Potent Competitive Inhibitors of Angiotensin-Converting Enzyme. Carboxyalkanoyl and Mercaptoalkanoyl Amino Acids. *Biochemistry* **1977**, *16*, 5484–5491.

(27) Condon, M. E.; Petrillo, E. W.; Ryondo, D. E.; Reid, J. A.; Neubeck, R.; Puar, M.; Heikes, J. E.; Sabo, E. F.; Losee, K. A.; Cushman, D. W.; Ondetti, M. A. Angiotensin-Converting Enzyme Inhibitors: Importance of the Amide Carbonyl of Mercaptoacyl Amino Acids for Hydrogen Bonding to the Enzyme. *J. Med. Chem.* **1982**, *25*, 250–258.
(28) Ondetti, M. A.; Cushman, D. W. Angiotensin-Converting Enzyme Inhibitors: Biochemical Properties and Biological Actions. *CRC Crit. Rev. Biochem.* **1984**, *16*, 381–411.
(29) Patchett, A. A.; Harris, E.; Tristram, E. W.; Wyvratt, M. J.; Wu, M. T.; Taub, D.; Peterson, E. R.; Ikeler, T. J.; ten Broeke, J.; Payne, L. G.; Ondeyka, D. L.; Thorsett, E. D.; Greenlee, W. J.; Lohr, N. S.; Hoffsommer, R. D.; Joshua, H.; Ruyle, W. V.; Rothrock, J. W.; Aster, S. D.; Maycock, A. L.; Robinson, F. M.; Hirschmann, R.; Sweet, C. S.; Ulm, E. H.; Gross, D. M.; Vassil, T. C.; Stone, C. A. A New Class of Angiotensin-Converting Enzyme Inhibitors. *Nature* **1980**, *288*, 280–283.
(30) Harris, R. B.; Ohlsson, J. T.; Wilson, I. B. Purification of Human Serum Angiotensin-I. Converting Enzyme by Affinity Chromatography. *Arch. Biochem. Biophys.* **1981**, *206*, 105–112.
(31) Ryan, J. W.; Chung, A. Slow Tight Binding Inhibitors of Angiotensin-Converting Enzyme. *Adv. Exp. Med. Biol.* **1983**, *156B*, 1133–1139.
(32) Thorsett, E. D.; Harris, E. E.; Aster, S.; Petersen, E. R.; Taub, D.; Patchett, A. A.; Ulm, E. H.; Vassil, T. C. Dipeptide Mimics: Conformationally Restricted Inhibitors of Angiotensin-Converting Enzyme. *Biochem. Biophys. Res. Commun.* **1983**, *111*, 166–171.
(33) Watthey, J. W. H.; Gavin, T.; Desai, M. Bicyclic Inhibitors of Angiotensin-Converting Enzyme. *J. Med. Chem.* **1984**, *27*, 810–818.
(34) Attwood, M. R.; Francis, R. J.; Hassall, C. H.; Krohn, A.; Lawton, G.; Natoff, I. L.; Nixon, J. S.; Redshaw, S.; Thomas, W. A. New Potent Inhibitors of Angiotensin-Converting Enzyme. *FEBS Lett.* **1984**, *165*, 201–206.
(35) Weller, H. N.; Gordon, E. M.; Rom, M. B.; Pluscec, J. Design of Conformationally Constrained Angiotensin-Converting Enzyme Inhibitors. *Biochem. Biophys. Res. Commun.* **1984**, *125*, 82–89.
(36) Natarajan, S.; Gordon, E. M.; Sabo, J. D.; Godfrey, J. D.; Weller, H. N.; Pluscec, J. Rom, M. B.; Cushman, D. W. *Ketomethyldipeptides*. I. A New Class of Angiotensin-Converting Enzyme Inhibitors. *Biochem. Biophys. Res. Commun.* **1984**, *124*, 141–147.
(37) Sawayama, T.; Tsukamoto, M.; Sasagawa, T.; Nishimura, K.; Yamamoto, R.; Deguchi, T.; Takeyama, K.; Hosoki, K. Angiotensin-Converting Enzyme Inhibitors: Synthesis and Structure-Activity Relationships of Potent Benzyloxycarbonyl Tripeptide Inhibitors. *Chem. Pharm. Bull.* **1989**, *37*, 2417–2422.
(38) Karanewsky, D. S.; Badia, C.; Cushman, D. W.; DeForrest, J. M.; Dejneka, T.; Lee, V. G.; Loots, M. J.; Petrillo, E. W. (Phosphinyloxy)-acyl Amino Acid Inhibitors of Angiotensin Converting Enzyme. 2. Terminal Amino Acid Analogues of (S)-1-[6-Amino-2-[[hydroxy-(4-phenylbutyl)phosphinyl]oxy]-1-oxohexyl]-L-proline. *J. Med. Chem.* **1990**, *33*, 1459–1469.
(39) Oya, M.; Matusmoto, J.; Takasinea, H.; Watanabe, T.; Iwao, J.-I. Thio Compounds. II. Synthesis and Antihypertensive Activity of Mercaptoacyl Amino Acids. *Chem. Pharm. Bull.* **1981**, *29*, 940–947.
(40) Tronrud, D. E.; Monzingo, A. F.; Matthews, B. W. Crystallographic Structural Analysis of Phosphoramidates as Inhibitors and Transition-State Analogs of Thermolysin. *Fur. J. Biochem.* **1986**, *157*, 261–263.
(41) Monzingo, A. F.; Matthews, B. W. Binding of N-Carboxymethyl Dipeptide Inhibitors to Thermolysin Determined by X-Ray Crystallography: A Novel Class of Transition-State Analogues for Zinc Peptidases. *Biochemistry* **1984**, *23*, 5724–5729.
(42) Holden, H. M.; Matthews, B. W. The Binding of L-Valyl-L-Tryptophan to Crystalline Thermolysin Illustrates the Mode of Interaction of a Product of Peptide Hydrolysis. *J. Biol. Chem.* **1988**, *263*, 3256–3260.
(43) Holden, H. M.; Tronrud, D. E.; Monzingo, A. F.; Weaver, L. H.; Matthews, B. W. Slow- and Fast-binding Inhibitors of Thermolysin Display Different Modes of Binding: Crystallographic Analysis of Extended Phosphonamidate Transition-State-Analogs. *Biochemistry* **1987**, *26*, 8542–3553.
(44) Tronrud, D. E.; Holden, H. M.; Matthews, B. W. Structures of Two Thermolysin-Inhibitor Complexes That Differ by a Single Hydrogen Bond. *Science* **1987**, *235*, 571–574.
(45) Holmes, M. A.; Matthews, B. W. Binding of Hydroxamic Acid Inhibitors to Crystalline Thermolysin Suggest a Pentacoordinate Zinc Intermediate in Catalysis. *Biochemistry* **1981**, *20*, 6912–6920.
(46) Holmes, M. A.; Tronrud, D. E.; Matthews, B. W. Structural Analysis of the Inhibition of Thermolysin by an Active-Site-Directed Irreversible Inhibitor. *Biochemistry* **1983**, *22*, 236–240.
(47) Monzingo, A. F.; Matthews, B. W. Structure of a Mercaptan-Thermolysin Complex Illustrates Mode of Inhibition of Zinc Proteases by Substrate-Analogs. *Biochemistry* **1982**, *21*, 3390–3394.
(48) Grobelny, D.; Goli, U. B.; Galardy, R. E. Binding Energetics of Phosphorus-Containing Inhibitors of Thermolysin. *Biochemistry* **1989**, *28*, 4948–4951.
(49) Kester, W. R.; Matthews, B. W. Crystallographic Study of the Binding of Dipeptide Inhibitors to Thermolysin: Implications for the Mechanism of Catalysis. *Biochemistry* **1977**, *16*, 2506–2516.
(50) Benchetrit, T.; Fournie-Zaluski, M. C.; Roques, B. P. Relationship Between the Inhibitory Potencies of Thiorphan and Retrothiorphan Enantiomers on Thermolysin and Neutral Endopeptidase 24.11 and Their Interactions with the Thermolysin Active Site by Computer Modelling. *Biochem. Biophys. Res. Commun.* **1987**, *147*, 1034–1040.
(51) Roderick, S. L.; Zaluski-Fournie, M. C.; Roques, B. P.; Matthews, B. W. Thiorphan and *retro*-Thiorphan Display Equivalent Interactions When Bound to Crystalline Thermolysin. *Biochemistry* **1989**, *28*, 1493–1497.
(52) Bartlett, P. A.; Marlowe, C. K. Possible Role for Water Dissociation in the Slow Binding of Phosphorous-Containing Transition-State-Analog Inhibitors of Thermolysin. *Biochemistry* **1987**, *26*, 8553–8561.
(53) Bartlett, P. A.; Marlowe, C. K. Evaluation of Intrinsic Binding Energy from a Hydrogen Bonding Group in an Enzyme Inhibitor. *Science* **1987**, *235*, 569–571.
(54) Feder, J.; Brougham, L. R.; Wildi, B. S. Inhibition of THermolysin by Dipeptides. *Biochemistry* **1974**, *13*, 1186–1189.
(55) Nishino, N.; Powers, J. C. Peptide Hydroxamic Acids as Inhibitors of Thermolysin. *Biochemistry* **1978**, *17*, 2846–2850.
(56) Morgan, B. P.; Scholtz, J. M.; Ballinger, M. D.; Zipkin, I. D.; Bartlett, P. A. Differential Binding Energy: A Detailed Evaluation of the Influence of Hydrogen-Bonding and Hydrophobic Groups on the Inhibition of Thermolysin by Phosphorus-Containing Inhibitors. *J. Am. Chem. Soc.* **1991**, *113*, 297–307.
(57) Klopman, G.; Bendale, R. D. Computer Automated Structure Evaluation (CASE): A Study of Inhibitors of the Thermolysin Enzyme. *J. Theor. Biol.* **1989**, *136*, 67–77.
(58) Clark, M.; Cramer, R. D., III; Van Opdenbosch, N. Validation of the General Purpose Tripos 5.2 Force Field. *J. Comput. Chem.* **1989**, *10*, 982–1012.
(59) Hausin, R. J.; Codding, P. W. Crystallographic Studies of Angiotensin-Converting Enzyme Inhibitors and Analysis of Preferred Zinc Coordination Geometry. *J. Med. Chem.* **1990**, *33*, 1940–1947.
(60) Waller, C. L.; Shands, R. F. B.; Marshall, G. R.; Dammkoehler, R. A. Work in progress.
(61) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A Geometric Approach to Macromolecule-Ligand Interactions. *J. Mol. Biol.* **1982**, *161*, 269–288.
(62) Shoichet, B. K.; Bodian, D.; Kuntz, I. D. Molecular Docking Using Shape Descriptors. *J. Comput. Chem.* **1992**, *13*, 380–397.
(63) Meng, E. C.; Shoichet, B. L.; Kuntz, I. D. Automated Docking with Grid-Based Energy Evaluation. *J. Comput. Chem.* **1992**, *13*, 505–524.
(64) Appelt, K. Crystal Structures of HIV-1 Protease-Inhibitor Complexes. *J. Comput.-Aided Mol. Des.*, PD3, **1993**, in press.
(65) Vedani, A.; Huhta, D. W. A. New Force Field for Modeling Metalloproteins. *J. Am. Chem. Soc.* **1990**, *112*, 4759–4767.
(66) Nicklaus, M. C.; Milne, G. W.; Burke, T. R. QSAR of Conformationally Flexible Molecules: Comparative Molecular Field Analysis of Protein-Tyrosine Kinase Inhibitors. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 487–504.
(67) Kellogg, G. E.; Semus, S. F.; Abraham, D. J. HINT: A New Method of Empirical Hydrophobic Field Calculation for CoMFA. *J. Comput.-Aided Mol. Des.* **1991**, *5*, 545–552.