Journal of Biochemistry and Molecular Biology, Vol. 31, No. 5, September 1998, pp. 459–467

JOURNAL OF
Biochemistry
and
Molecular Biology

3D-QSAR of Angiotensin-Converting Enzyme Inhibitors: Functional Group Interaction Energy Descriptors for Quantitative Structure-Activity Relationships Study of ACE Inhibitors

Sanguk Kim, Myung Whan Chi, Chang No Yoon* and Ha-Chin Sung†

Bioanalysis and Biotransformation Research Center, Korea Institute of Science and Technology, Cheongryang, Seoul 130-650, Korea

†Graduate School of Biotechnology, Korea University, Seoul 136-701, Korea

Received 11 May 1998, Accepted 10 June 1998

A new set of functional group interaction energy descriptors relevant to the ACE (Angiotensin-Converting Enzyme) inhibitory peptide, QSAR (Quantitative Structure Activity Relationships), is presented. The functional group interaction energies approximate the charged interactions and distances between functional groups in molecules. The effective energies of the computationally derived geometries are useful parameters for deriving 3D-QSAR models, especially in the absence of experimentally known active site conformation. ACE is a regulatory zinc protease in the renin-angiotensin system. Therapeutic inhibition of this enzyme has proven to be a very effective treatment for the management of hypertension. The nonbond interaction energy values among functional groups of six-feature of ACE inhibitory peptides were used as descriptor terms and analyzed for multivariate correlation with ACE inhibition activity. The functional group interaction energy descriptors used in the regression analysis were obtained by a series of inhibitor structures derived from molecular mechanics and semi-empirical calculations. The descriptors calculated using electrostatic and steric fields from the precisely defined functional group were sufficient to explain the biological activity of inhibitor. Application of the descriptors to the inhibition of ACE indicates that the derived QSAR has good predicting ability and provides insight into the mechanism of enzyme inhibition. The method, functional group interaction energy analysis, is expected to be applicable to predict enzyme inhibitory activity of the rationally designed inhibitors.

* To whom correspondence should be addressed.
Tel: 82-2-958-5068, Fax: 82-2-958-5059
E-mail: cody@kistmail.kist.re.kr

Keywords: ACE inhibitor, Energy descriptor, QSAR.

Introduction

Angiotensin-Converting Enzyme (ACE, EC 3.4.15.1, dipeptidyl carboxy peptidase) is a regulatory zinc protease in the renin-angiotensin system. ACE converts the decapeptide angiotensin I to the vasoconstrictive octapeptide angiotensin II by cleaving the C-terminal dipeptide (Skeggs *et al.*, 1954; Erdos, 1977). 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 three-dimensional structure of ACE is unknown, much information for ACE inhibition has been derived from a traditional structure-activity relationship (SAR) study (Saunders *et al.*, 1987).

The study suggests the following requirements for inhibition of the enzyme: (1) a functional group capable of binding to zinc atom at 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 group for ionic binding to a positively-charged residue (i.e., arginine, lysine) (Wyvratt *et al.*, 1985).

Traditional QSAR applications have long been attempted to correlate biological activity with measurable physicochemical parameters such as Hansch's log P, Hammett's *s*, Taft's *Es*, and MR (Molecular Refractivity) (Hansch *et al.*, 1979). Recently, it has become obvious that in order to develop meaningful quantitative models of structure-activity relationships, the three-dimensional structures of the active compounds must be considered. Hence, current QSAR methods include molecular-shape analysis (Hopfinger, 1980), the hypothetical active site lattice (HASL) (Doweyko, 1988), Crippen's distance

geometry (Ghose *et al.*, 1985) and Voronoi binding site (Boulu *et al.*, 1987), and comparative molecular field analysis (CoMFA) (Cramer *et al.*, 1988; Waller *et al.*, 1993; Raghavan *et al.*, 1995).

The 3D-QSAR methodologies are based on the assumption that the interactions between a ligand (inhibitor) and its receptor (enzyme) are primarily noncovalent and shape-dependent (Sok *et al.*, 1995; Chung, 1996; Ha *et al.*, 1996). Therefore, a QSAR can be derived by analyzing the electrostatic and steric fields surrounding a set of inhibitors.

When a ligand interacts with a receptor at the active site, the potency of each functional group is more important than that of the whole molecule. Thus, within a set of related compounds, it can be expected that only a subset of the components describing the binding energy will account for most of the variance in ligand affinity. In fact, several authors have shown good correlation between particular energy components and biological activities (Blaney *et al.*, 1982; Menziani *et al.*, 1989; Ortiz *et al.*, 1995).

Therefore, the descriptors calculated using electrostatic and steric fields from the precisely defined functional group in a molecule could explain the biological activity of the inhibitor. The advantage of subjecting functional group interaction energies to statistical analysis is that the interaction energy components of little correlation with binding affinity can be reduced, while mechanically important interaction energy components can be identified. The aim of the present work is to develop a more systematic and quantitative approach in the use of structural information in the derivation of QSARs. This can be achieved by using calculated functional group interaction energies directly as descriptors in the QSAR.

In this study, the QSAR analyses were carried out for the 15 dipeptides representing diverse structures and ACE inhibitory activities. In the case of rigid molecules that have a single conformation, QSAR study was carried out using each molecular structure from a database. But, in the case of conformationally flexible polypeptides, the lowest energy conformation of each molecule was used in most studies (Raghavan *et al.*, 1995). Since the ACE inhibitory dipeptides have flexible torsion angles and the receptor complex structures are unknown, the systematic torsion angle search was performed to find out the lowest energy conformation for 3D-QSAR analysis. The descriptors included in the analysis that are correlated with ACE inhibitory activity are the following interaction energy values among six-feature of ACE inhibitors; (1) a negative ionizable group on the C-terminal carboxylate, (2) a hydrophobic region on the R-group of the C-terminal residue, (3) a hydrogen bond donor on the C-terminal amide, (4) a hydrogen bond acceptor on the N-terminal carbonyl, (5) a hydrophobic region on the R-group of the N-terminal residue, and (6) a hydrogen bond donor on the N-terminus, hypothesized by the structures of ACE inhibitory dipeptides (Cheung *et al.*, 1980; Haushin *et al.*, 1990; Pascard *et al.*, 1991). The resulting model was analyzed in order to determine structural features of functional groups that contribute for ACE inhibitory activity and/or may increase biological activity of the inhibitor.

Materials and Methods

Molecular modeling For the analysis of QSAR in ACE inhibitors, a set of 15 dipeptides (Table 1) displaying varying degrees of ACE inhibition was selected from the published data (Cushman *et al.*, 1980). The three-dimensional structures of molecules were constructed using the Biopolymer module of Insight II (Version 2.3.5, San Diego, USA) program. The Biopolymer module produced starting structures with an extended backbone conformation.

A systematic conformational search was then made for each molecule, in which all torsion angles were allowed to rotate in 10, 30, 60-degree increments, starting from the input conformation (Evans *et al.*, 1995). Steepest descents and conjugate gradient energy minimization steps were carried out for the torsionally perturbated structures. The conformers within 1 kcal deviation from the lowest potential energy were selected. Then, each molecule was surrounded by 25 $Å^3$ shell of water molecules and optimized using the cvff DISCOVER forcefield (San diego, USA) with a 10 Å cutoff and a dielectric constant of 1. The energy of solvent box was minimized with periodic boundary conditions to a root mean square (RMS) gradient < 0.01 kcal $mol^{-1}Å^{-1}$. The lowest energy structure calculated from the explicit water environment was subjected to further interaction energy analysis.

Partial atomic charges required for calculation of the electrostatic interaction were computed by a semi-empirical molecular orbital method using the MOPAC program (San Diego, USA). The charges were computed using the MNDO (Modified Neglect of Diatomic Overlap) method.

Functional group interaction energy analysis The nonbond interaction energy values among functional groups of six-feature of ACE inhibitory peptides were used as descriptor terms and analyzed for multivariate correlation with ACE inhibition activity. When a ligand molecule interacts with a receptor, the active site of the receptor takes a pocket shape. Hence, the conformational features of functional groups determine the receptor recognition potency of a ligand molecule.

Although the three-dimensional structure of ACE is unknown, it has been hypothesized that the conformational features of the functional groups were derived from the common physicochemical features on known ACE inhibitors. The 3D arrangement of the dipeptide (Val-Trp) that has the highest degree of ACE inhibition in the training set is represented in Fig. 1.

This structure represents the six-feature of ACE inhibitors that were used in development of anti-hypertensive drug design; a negative ionizable group on the C-terminal carboxylate, a hydrophobic region on the R-group of the C-terminal residue, a hydrogen bond donor on the C-terminal amide, a hydrogen bond acceptor on the N-terminal carbonyl, a hydrophobic region on the

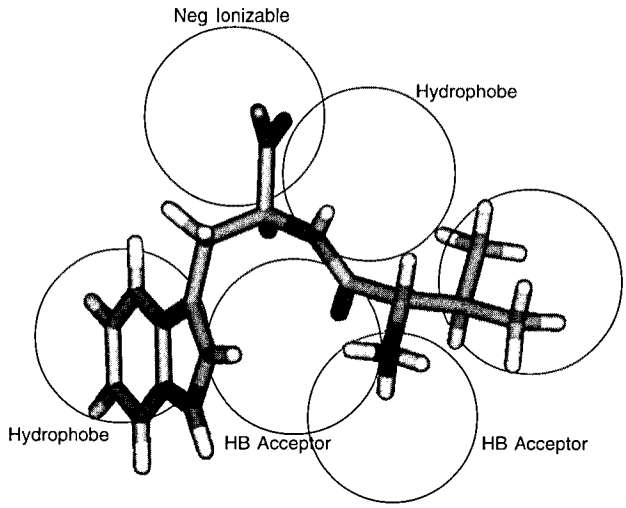

Fig. 1. Structure of ACE inhibitory dipeptide (Val-Trp) and the six-feature hypothesized by the structures of ACE inhibitors.

R-group of the N-terminal residue, and a hydrogen bond donor on the N-terminus. Thus, the atoms of each dipeptide used in the training set were divided into six functional groups (Fig. 2). The interaction energies between atoms of ACE inhibitors were obtained from a series of inhibitor structures obtained by molecular mechanics and semi-empirical calculations. The DISCOVER program was used for the energy calculations.

Nonbond interaction energy between functional groups is composed of a van der Waals and an electrostatic terms. The former can be considered as a size parameter and representing of electron correlation, while the later provides a quantitative measurement of the influence of polarity on the energy and structure. The van der Waals term is represented by the Lennard-Jones potential expression;

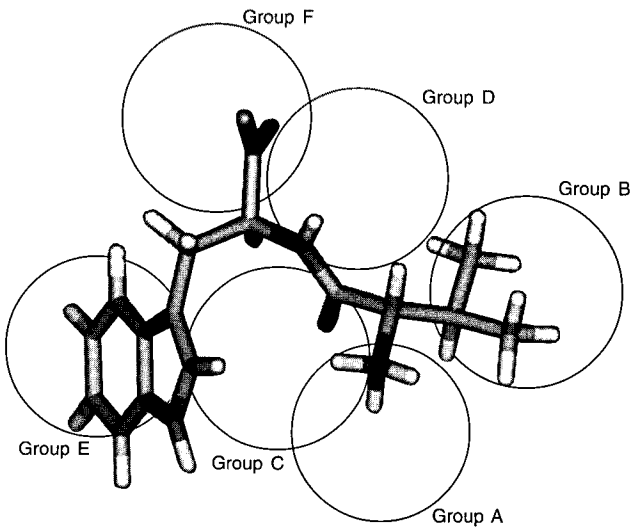

Fig. 2. Structure of ACE inhibitory dipeptide (Val-Trp) and the six functional groups that are selected by the structures of ACE inhibitors.

$$E_{vdw} = \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}$$

A_{ij} and B_{ij} are *Lennard-Jones* potential parameters, and r_{ij} is the distance between the atom i and the atom j.

The Coulombic representation of electrostatic interaction is calculated using the equation:

$$E_{ele} = \frac{q_i q_j}{\varepsilon r_{ij}}$$

where q_i, q_j are the charges of atom i and atom j, and ε is the dielectric constant.

A program was written in FORTRAN77 to get the functional group interaction energy from the output of the DISCOVER program. This program has three steps: It (1) reads pairwise interaction energy values between each atom of a molecule, (2) divides atoms of a molecule into functional groups based on the six-feature of ACE inhibitory activity, and (3) evaluates the total nonbond interaction energy between groups. We obtained the energy descriptors, energy values of 15 group pairwise interactions from 6 functional groups of ACE inhibitors, through the program.

Additional descriptors are the size of each molecule, defined either by the *van der Waals* volume or by the water accessible surface area, and the shape of each molecule, defined by volume per area.

Statistical analysis The method of multiple linear regression has been shown to be suitable in QSAR studies which seek to rationalize the structural features affecting the biological activity. Regression seeks a relationship between a response or dependent variable, which is the vector Y and a descriptor X. The data matrix X is built with columns representing pairs of functional group interaction energy values and rows representing each dipeptide in the training set.

A column of Y variables containing the experimental activities expressed as log ($1/IC_{50}$) was added. Some of the energy descriptors may not contribute to the differences in binding and may therefore add "noise" to the matrix of group interaction energy values (Ortiz *et al.*, 1995). To exclude it from the QSAR, a variable selection procedure is carried out in which the effects of individual variables on predicting ability are evaluated iteratively using all possible combinations of variables and taking into account the magnitude of increment or decrement of R^2. The dimensionality of the data matrix was thus reduced while keeping the amount of information loss to a minimum.

The regression coefficient, R^2, was characterized by the following:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_{exp(i)} - y_{pred(i)})^2}{\sum_{i=1}^{N} (y_{exp(i)} - y_{mean})^2}$$

where $y_{pred(i)}$ corresponds to the activity predicted by the regression model for molecule i, $y_{exp(i)}$ is the experimental activity of molecule i, and y_{mean} is the average experimental activity of a complete set of n compounds.

Table 1. A set of 15 dipeptides have varying degrees of ACE inhibitory activity.

Molecule	IC_{50} (nM)*
Val-Trp	1700
Ile-Tyr	3700
Ala-Trp	10000
Ile-Pro	150000
Ala-Pro	270000
Ala-Val	300000
Val-Pro	420000
Gly-Phe	450000
Ala-Leu	1.6e + 06
Ala-Gly	2.5e + 06
Gly-Glu	5.4e + 06
Gly-Lys	5.4e + 06
Pro-Pro	7.5e + 06
Ala-His	9.0e + 06
Gly-Asp	9.2e + 06

* Data from Cushman *et al.* (1980).

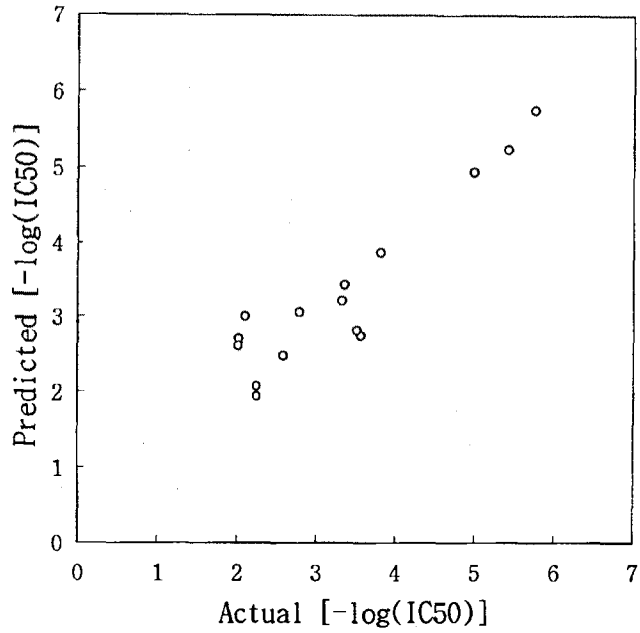

Fig. 3. Predicted vs actual pIC_{50} for the inhibition of ACE using only molecular shape descriptors. The predictive model was derived from the molecular shape descriptor having $R^2 = 0.81$.

Results

Results of the regression analysis of 15 ACE inhibitory dipeptides for correlation of log $(1/IC_{50})$ with the descriptors, such as Volume, Area, Volume/Area, and the number of atoms in molecules are summarized in Table 2. The values of Volume/Area describe the degree of wrinkle surface of a molecule. These molecular size and shape descriptors are only weakly correlated with the experimental activities (regression coefficient $R^2 < 0.66$). The model using all of molecular size and shape descriptors has a regression coefficient $R^2 = 0.81$. The data fitting and the predictive abilities of the model are displayed in Fig. 3. In these cases, models derived from the values of molecular size and shape descriptors alone were insufficient for accurately quantifying and predicting the nature of enzyme inhibitory activity. The molecular size and shape descriptors represent only limited information about the steric nature of the inhibitors to receptor active site.

To explain the interactions between inhibitor and receptor, we have employed other properties of inhibitor molecule, such as hydrophobic and electrostatic interaction potential, as descriptors. The total nonbond interaction energy (Table 3) of each inhibitor was taken from the lowest energy structure calculated in an explicit water environment.

Figure 4 shows the predicting ability based upon total nonbond interaction energy values as descriptor (regression coefficient $R^2 = 0.75$). The total nonbond interaction energy was a better predictor than molecular size or shape descriptor alone in correlating predictive activity with experimental activity. The correlation between total nonbond interaction energies of molecules and their reactivities suggests that more reactive compounds should better interact with receptors than others. But, since there are structural parts of inhibitory molecules that contribute little to interaction with receptors, we can think that these parts serve as noise, i.e., introduced inaccuracies to total nonbond interaction energy. Figure 5 shows the predicting ability based upon four descriptors, total nonbond interaction energies of molecules, and three molecular shape descriptors (regression coefficient $R^2 = 0.83$). Despite that additional descriptors were added in the model, the regression coefficient is not enhanced much, compared to the model which used only three molecular shape descriptors.

The *F*-verification shows a general accuracy of a regression model where we obtain a *F*-value = 18.2. This is not any better than the model using molecular size and shape descriptors only.

The 15 energy values of group pair interaction between six functional groups of ACE inhibitory dipeptides are listed in Table 4. In order to determine the functional group interaction energy descriptors that participate in the best regression model, selection procedures eliminated the columns in the X matrix corresponding to functional group pair interaction energy values that have a little descriptive ability for the regression model. The model using a minimum five energy descriptors of group pair interaction explained above 90% of the variance in inhibition potency (Table 5).

Table 2. List of molecular shape descriptors, such as volume, shape, and atom number, with their R^2 values.

Number	Compound	Activity $-\log(IC_{50})$	Volume (Å^3)	Area (Å^2)	Shape (volume/area)	Number of atom
1	Val-Trp	5.77	917	555.0	1.65	43
2	Ile-Tyr	5.43	907	551.5	1.65	43
3	Ala-Trp	5.00	812	495.8	1.63	37
4	Ile-Pro	3.82	724	451.2	1.60	36
5	Ala-Pro	3.57	594	387.8	1.53	27
6	Ala-Val	3.52	633	410.5	1.54	29
7	Val-Pro	3.38	674	423.3	1.59	33
8	Gly-Phe	3.35	697	443.6	1.57	30
9	Ala-Leu	2.80	682	436.8	1.56	32
10	Ala-Gly	2.60	486	339.8	1.43	20
11	Gly-Glu	2.27	627	421.9	1.49	26
12	Gly-Lys	2.27	684	454.0	1.51	31
13	Pro-Pro	2.12	655	419.6	1.56	31
14	Ala-His	2.04	692	449.4	1.54	30
15	Gly-Asp	2.03	552	373.6	1.47	23
		R^2	0.65	0.57	0.66	0.65

Table 3. Total nonbond interaction energy of ACE inhibitory dipeptides.

Number	Total nonbond interaction energy (kcal)
1	33.07
2	39.92
3	29.06
4	14.35
5	6.05
6	8.57
7	11.19
8	28.49
9	7.85
10	3.73
11	4.24
12	6.85
13	6.46
14	4.13
15	3.21

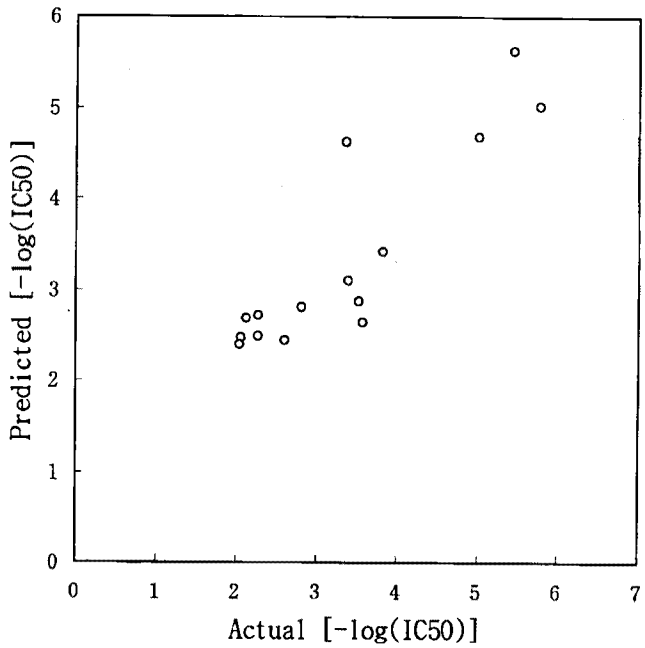

Fig. 4. Predicted vs actual pIC_{50} for the inhibition of ACE using the total nonbond interaction energies as descriptors. The predictive model was derived from a descriptor having $R^2 = 0.75$.

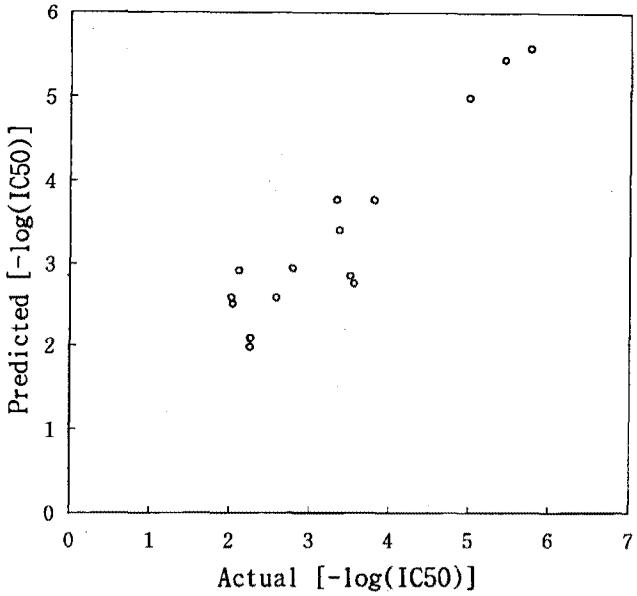

Fig. 5. Predicted vs actual pIC_{50} for the inhibition of ACE using the total nonbond interaction energies of molecules and the molecular shape descriptors. The predictive model was derived from five descriptors having $R^2 = 0.83$.

The regression coefficient R^2 values using the group interaction energy components are comparable to those of models using molecular size and shape descriptors. The predicting ability of those models is significantly improved. Thus, within a set of functional group pairwise interaction energy components, we have obtained the important interaction energy terms that highly correlate with biological activity through the selection procedure.

In order to overcome the difficulty of obtaining a good model when the group interaction energy descriptors are used alone, we considered total nonbond interaction energies and the number of atoms in molecules as additional regressors in the QSAR model. We chose the number of atoms in a molecule as additional descriptors because all three regressors, Volume, Area, and the number of atoms in molecules, have autocorrelation. However, the model including these three regressors above did not improve the results. The numbers of principal components were unchanged from analysis using these additional descriptors. In the final analysis, the first four principal components explain 84% of the variance in the biological data.

The model using the group interaction energy descriptors in combination with the total nonbond interaction energies and the number of atoms in molecules achieved a good 3D-QSAR (Fig. 6).

The experimental ACE inhibitory activity is expressed by the following QSAR equation:

$$\text{Log } (1/IC_{50}) = 1.8E_{AE} + 3.56E_{BC} + 0.51E_{BE} - 6.87E_{BF} - 4.69E_{CF} + 0.04E_{TOE} + 0.06N_{AT} - 2.03$$

$$n = 15\ ,\quad R^2 = 0.95\ ,\quad s = 0.11\ ,\quad F = 40.8$$

where IC_{50} is the concentration inhibiting 50% of the activity of angiotensin converting enzyme, E_{ij} is the interaction energy value between group i and group j which could be taken in the A to F groups, E_{TOE} is the total nonbond interaction energy of the molecule, and N_{AT} is the number of atoms in a molecule. This correlation coefficient R^2 value shows the considerable predicting capacity of the QSAR model shown in the plot of experimental (actual) versus predicted ACE inhibitory activities.

Table 4. The nonbond interaction energy values between six functional groups of ACE inhibitory dipeptides.

Number	Eab	Eac	Ead	Eae	Eaf	Ebc	Ebd	Ebe	Ebf	Ecd	Ece	Ecf	Ede	Edf	Eef
1	2.06	0.14	0.04	−0.47	−0.36	−0.10	0.64	−0.73	−0.44	−0.09	1.33	−0.42	0.69	0.41	0.15
2	2.15	−0.03	0.48	−0.55	−0.40	−0.11	0.77	−0.77	−0.42	−0.09	1.11	−0.40	0.95	0.57	0.09
3	0.37	0.61	0.06	−0.81	−0.64	−0.16	0.71	−0.73	−0.45	−0.33	1.17	−0.57	0.79	1.74	0.02
4	2.91	1.89	−0.38	−0.07	−0.40	−0.10	0.47	2.35	−0.93	0.00	2.14	0.84	−0.06	0.82	0.07
5	0.71	0.84	−0.38	−0.08	−0.36	−0.10	0.17	2.47	−0.34	0.00	2.22	−0.08	−0.08	0.47	0.81
6	0.79	0.03	0.08	−0.13	−0.37	−0.10	0.80	−0.28	−0.37	−0.08	2.17	−0.27	2.48	1.50	−0.07
7	2.78	1.70	−0.38	−0.07	−0.37	−0.10	0.17	2.31	−0.52	0.18	1.73	0.22	−0.08	0.43	0.00
8	0.16	0.10	0.20	−0.45	−0.36	−0.15	0.36	−0.33	−0.28	−0.08	−0.23	−0.40	1.69	0.64	0.04
9	0.87	−0.07	0.07	−0.28	−0.35	−0.10	0.47	−0.33	−0.27	−0.08	1.94	−0.43	1.63	0.97	0.03
10	0.84	1.04	−0.25	0.03	−0.41	−0.11	0.10	−0.14	−0.27	−0.09	2.47	−0.38	0.01	0.17	−0.07
11	0.15	0.96	−0.24	−0.02	−0.35	−0.24	0.60	−0.19	−0.23	−0.10	2.34	−0.40	1.04	0.26	0.05
12	0.15	0.95	−0.24	−0.03	−0.36	−0.24	0.61	−0.16	−0.25	−0.09	2.22	−0.31	0.74	0.11	−0.00
13	0.13	1.00	−0.31	−0.11	−0.27	−0.10	0.11	2.35	−0.55	0.00	2.16	0.57	−0.08	0.75	0.11
14	0.63	−0.12	0.33	−0.14	−0.37	0.31	1.20	−0.29	−0.35	−0.05	2.21	0.22	0.28	0.37	0.24
15	0.16	0.97	−0.24	−0.05	−0.36	−0.25	0.67	−0.23	−0.28	−0.10	1.56	−0.37	0.57	0.71	0.11

Table 5. Descriptors used in training a best predictive model, with the associated R^2 values for each model.

Number of descriptor	R^2 values	C_P values	Functional group pair used in regression analysis
6	0.92108977	4.27822	AB AC BE BF CF DF
6	0.92073392	4.29751	AB AE BC BE BF CF
6	0.91990279	4.34257	AB AE BE BF CE CF
6	0.91837439	4.42544	AB AC AE BE BF CF
6	0.91234414	4.75237	AB AC BC BE CE DF
6	0.91124730	4.81184	AB AC BC BE CE DF
6	0.90915395	4.92533	AB AE BE BF CF DF
5	0.90785403	2.99581	AB AE BE BF CF
5	0.90671724	3.05744	AB AC BC BE DF
5	0.90404285	3.20244	AB AE BE CE CF

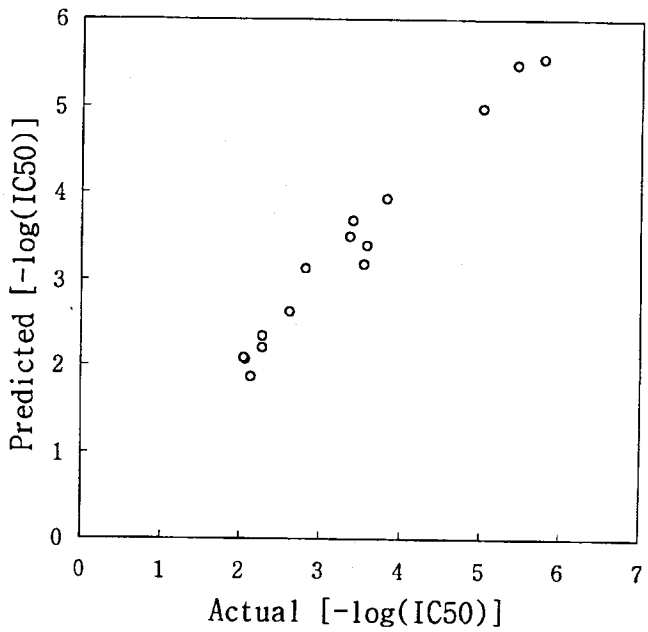

Fig. 6. Predicted vs actual pIC_{50} for the inhibition of ACE using the functional group interaction energy descriptors. The predictive model was derived from functional group interaction energy descriptors having $R^2 = 0.95$.

Thus, only five energy descriptors of group pairwise interaction are required in order to explain the difference in activity. The energy descriptors selected from the model are the group pair interaction energy values between A-E, B-C, B-E, B-F, C-F groups. We can define the selected energy variables of nonbond interaction energy among functional groups as "effective" energies. These effective energies may themselves lack physical meaning, but they may act as statistical descriptors of other physically important interactions.

Discussion

A significant functional group interaction energy contribution to activity may indicate the presence of an electrostatic and a van der Waals interaction with the receptor. An earlier structure-activity relationship study (Cheung *et al.*, 1980) of ACE inhibitors indicated that Pro and aromatic side chains of Tyr and Phe are favored in position-2 or at the C-terminal end of dipeptides, while branched-chain aliphatic amino acids like Ile and Val are most effective at the N-terminus. When the ACE inhibitor interacts with the active site of the enzyme, the potency of each functional group is more important than that of the whole molecule. These nonbond interaction energies between functional groups, consisting of electrostatic and steric forms, are associated with differences in inhibitory potency.

In the absence of an enzyme structure, the interpretation of ligand–receptor interactions is highly intuitive. From an electrostatic consideration, one would expect to see strong electrostatic interactions around the C-terminal carboxyl group as well as the amide bond between P1′ and P2′ residue groups, since these groups are proposed to exhibit strong hydrogen-bonding and charge–charge interactions, respectively. Examination of the "effective" interaction energy components supports this binding hypothesis (Depriest *et al.*, 1993).

The determination of the active site geometry of the inhibitor is the most important factor to explain the variance of biological activity. To determine the active conformation of the molecule, we applied the dihedral angle (ϕ, ψ, ω) increasing method. This creates a larger number of "bad" structures having steric conflict, but it also provides alternative "paths" that may eventually lead to lower energy minima, because it can overcome the energy barrier and have an opportunity for more "downhill" path through conformational space (Evans *et al.*, 1995).

According to the current model of ACE inhibitor binding, the P1 site interacts with a hydrophobic pocket, which recognizes its aromatic ring. We tested the dihedral angle increment of varying intervals to determine the lowest energy conformation. The case of the aromatic group amino acid (Tyr, Trp, Phe) and His had different conformations from 30° to 60° intervals, and the descriptors are substantially different for each conformation. It is apparent that a smaller angle increment would lead to lower energy minima, but the computational burden increases exponentially in proportion to the interval (the torsion angle increment for dipeptide with 10° interval requires more than 10 CPU hours on an SGI R4400 workstation). We, therefore, had to select a more efficient way. We concluded the following regarding torsion angle increment interval; (1) small angle intervals (< 30°) are preferable for determining the lowest energy structure of dipeptides, (2) larger angle intervals (60°) may be useful for dipeptides having Gly, Ala, Pro residues, and

(3) sufficiently small angle intervals (10°) are needed for dipeptides having Tyr, Trp, Phe, and His residues.

These results indicate that a reliable QSAR model can be constructed on the basis of the descriptors from the systematic conformational search, starting from the structure of the molecule constructed using the standard library, atomic charges computed by a semi-empirical method, and the structure optimized with an explicit water condition by molecular mechanics calculation.

It is possible that multiple conformers of ligands could be included in subsequent QSAR analysis and the "best conformation" could be selected on how well it fits the derived model (Nicklaus *et al.*, 1992). This QSAR approach, coupled with a conformational search, proved to be a successful modeling approach. It seems relatively inefficient, but is confirmed by the predictive capacity of the model given by an $R^2 = 0.95$.

The functional group interaction energy values are useful descriptors. In order to improve the predicting ability, additional descriptors, total nonbond interaction energy, and the number of atoms in molecules are included in the model. Additional regressors contribute to the best model, but the explanation for ligand–receptor interaction of those regressors is not clear. Previous reports (Walter *et al.*, 1993; Ortiz *et al.*, 1995) indicated that regressors to describe the molecular interactions must have enthalphic and entropic natures. But the descriptors in this study, the functional group interaction energy descriptors composed of van der Waals and electrostatic energy, represented the enthalphic nature of ligand–receptor interactions only. Although the additional descriptors used in the best regression model improved the predictive R^2 value, they were not entropic descriptors.

It is possible to include the difference in the solvation of molecule and internal degrees of freedom as entropic aspects of binding free energy, and they are supposed to increase the predicting ability of the model. An additional entropic regressor under consideration is the difference in solvation of molecules that can be derived from analysis of vacuum and solvated structures of the inhibitor. The inclusion of this descriptor into the regression equation will be a type of entropic correction method.

Future approaches will include the conformational-search method development of nonpeptide ACE inhibitory compounds and the additional descriptor analysis. Also, we will increase the utility of this QSAR as a pharmacological prescreen for biological activity.

References

Blaney, J. M., Weiner, P. K., Dearag, A., Kollman, P. A., Jorgensen, E. C., Oatley, S. J., Burridge, J. M. and Blake, C. C. F. (1982) Molecular mechanics simulation of protein–ligand interactions: Binding of thyroid hormone analogues to prealbumin. *J. Am. Chem. Soc.* **104**, 6424–6434.

Boulu, L. G. and Crippen, G. M. (1987) Voroni binding site models: Calculation of binding modes and influence of binding data accuracy. *J. Comput. Chem.* **10**, 673–682.

Cheung, H. S., Wang, F. L., Ondetti, M. A., Sabo, E. F. and Cushman, D. W. (1980) Binding of peptide substrates and inhibitors of angiotensin-converting enzyme. *J. Biol. Chem.* **255**, 401–407.

Chung, H. S. (1996) Intramolecular hydrogen bonds in proteinase inhibitor protein, A molecular dynamics simulation study. *J. Biochem. Mol. Biol.* (formerly *Korean Biochem. J.*) **29**, 380–385.

Cramer, R. D., III, Patterson, D. E. and Bunce, J. D. (1988) Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **110**, 5959–5967.

Cushman, D. W., Ondetti, M. A., Cheung, H. S., Autonaccio, M. J., Murthy, V. S. and Rubin, B. (1980) Inhibitions of angiotensin-converting enzyme. *Adv. Exp. Med. Biol.* **130**, 199–225.

DePriest, S. A., Mayer, D., Naylor, C. B. and Marshall, G. R. (1993) 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.* **115**, 5272–5384.

Doweyko, A. M. (1988) The hypothetical active site lattice, an approach to modeling active sites from data on inhibitor molecule. *J. Med. Chem.* **31**, 1396–1406.

Erdos, E. G. (1977) The angiotensin I converting enzyme. *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **36**, 1760–1765.

Evans, J. S., Mathiowetz, A. M., Chan, S. I. and Goddard, W. A., III. (1995) *De novo* prediction of polypeptide conformations using dihedral probability grid Monte Carlo methodology. *Protein Sci.* **4**, 1203–1216.

Ghose, A. K. and Crippen, G. M. (1985) Use of physicochemical parameters in distance geometry and related three-dimensional quantitative structure-activity relationships: A demonstration using *Escherichia coli* dihydrofolate reductase inhibitors. *J. Med. Chem.* **28**, 333–346.

Ha, J. M., Shin, S. Y., Hong, H. N., Suh, D. J., Jang, T. S., Kang, S. W., Kuean, S. J. and Ha, B. J. (1996) Structure-antagonistic activity relationships of an NK-2 tachykinin receptor antagonist, L-659,877 and its analogues. *J. Biochem. Mol. Biol.* (formerly *Korean Biochem. J.*) **29**, 429–435.

Hansch, C. and Leo, A. (1979) *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley-Interscience, New York.

Hausin, R. J. and Codding, P. W. (1990) Crystallographic studies of angiotensin converting enzyme inhibitors and analysis of preferred zinc coordination geometry. *J. Med. Chem.* **33**, 1940–1947.

Hopfinger, A. J. (1980) A QSAR investigation of dihydrofolate reductase inhibition by Baker triazines based upon molecular shape analysis. *J. Am. Chem. Soc.* **102**, 7196–7209.

Menziani, M. C., De Benedetti, P. G., Gago, F. and Richards, W. G. (1989) The binding of benzene sulfonamides to carbonic anhydrase enzyme. A molecular mechanics study and quantitative structure-activity relationships. *J. Med. Chem.* **32**, 951–956.

Meyer, H. (1899) Zur theorie der alkoholnarkose. 1. Mitteilung. *Arch. Exp. Pathol. Pharmakol.* **42**, 109–118.

Nicklaus, M. C., Milne, G. W. and Burke, T. R. (1992) QSAR of conformationally flexible molecules: comparative molecular field analysis of protein-tyrosine kinase inhibitors. *J. Comput.-Aided Mol. Des.* **6**, 487–504.

Ortiz, A. R., Pisabarro, M. T., Gago, F. and Wade, R. C. (1995) Prediction of drug binding affinity by comparative binding energy analysis. *J. Med. Chem.* **38**, 2681–2691.

Pascard, C., Guilhem, J., Vincent, M., Remond, G., Portevin, B. and Laubie, M. (1991) Configuration and preferential solid-state conformations of perindoprilat (S-9780). Comparison with the crystal structures of other ACE inhibitors and conclusions related to structure-activity relationships *J. Med. Chem.* **34**, 663–669.

Petrillo, H. Y., Trippodo, N. C. and DeForrest, J. M. (1989) In *Annual Reports in Medicinal Chemistry*, Robertson, D. W., (Ed.), **25**, pp. 51–60, Academic Press, New York.

Raghavan, K., Buolamwini, J. K., Fesen, M. R., Pommier Y., Kohn, K.W. and Weinstein, J. N. (1995) Three-dimensional quantitative structure-activity relationship (QSAR) of HIV integrase inhibitors: A comparative molecular field analysis (CoMFA) study. *J. Med. Chem.* **38**, 890–897.

Saunders, M. R., Tute, M. S. and Webb, G. A. (1987) A theoretical study of angiotensin-converting enzyme inhibitors. *J. Comput.-Aided Mol. Des.* **1**, 133–142.

Skeggs, L. T., Marsh, W. H., Kahn, J. R. and Shumway, N. P. (1954) Existence of two forms of hypertensin. *J. Exp. Med.* **99**, 275–282.

Sok, D. E. and Kim M. R. (1995) Binding subsites in the active site of Zn^{2+}-glycerophosphocholine cholinephosphodiesterase. *J. Biochem. Mol. Biol.* (formerly *Korean Biochem. J.*) **28**, 94–99.

Walter, C. L. and Marshall, G. R. (1993) Three-dimensional quantitative structure-activity relationship of angiotensin-converting enzyme and thermolysin inhibitors. *J. Med. Chem.* **36**, 2390–2403.

Wyvratt, M. J. and Patchett, A. A. (1985) Recent developments in the design of angiotensin-converting enzyme inhibitors. *Med. Res. Rev.* **5**, 483–531.