Improved Electron-Conformational Method of Pharmacophore Identification and Bioactivity Prediction. Application to Angiotensin Converting Enzyme Inhibitors

Isaac B. Bersuker,*[,†,‡] Süleyman Bahçeci,[†] and James E. Boggs[†]

Department of Chemistry and Biochemistry, Institute for Theoretical Chemistry, and College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712

Received April 14, 2000

The electron-conformational (EC) method of pharmacophore (Pha) identification and bioactivity prediction, suggested earlier, is given here two major improvements. First, an atomic index of orbital and charge controlled interaction is introduced to better represent the ligand (substrate) in its interaction with the bioreceptor. Second, the multiconformational problem is considered in view of ligand−receptor binding states, resulting in essential simplification of the expression of bioactivity. The details of the improved EC method are demonstrated in application to the problem of angiotensin converting enzyme (ACE) inhibitors. The Pha of the latter is identified by separation of the heavily populated conformations of the chosen 51 compounds (the training set), calculation of the electronic structure, construction of their EC matrixes of congruity, and processing of the latter in comparison with the activities to reveal a common submatrix of all the active only compounds that describes the Pha. The latter contains three oxygen atoms plus a fourth atom X = S, N, O at certain interatomic distances and with restricted electronic parameters (within assumed tolerances), the position of the atom X being more changeable from one active compound to another. For quantitative prediction of the bioactivity, an expression is deduced which takes into account the duly parametrized influence of auxiliary groups (AG) which, being positioned outside the Pha, either diminish the activity (antipharmacophore shielding) or enhance it. It is shown that in case of many conformations of the same compound only one of them, that of the lowest energy which has the Pha, should be parametrized. The 15 parameters chosen to represent the AG in case of ACE inhibitors are weighted by variational (adjustable) coefficients which are determined from a regression treatment of the calculated versus known activities in the training set. Then the formulas with known coefficients are used to validate the method by calculating the bioactivity of other compounds not used in the training set. The prediction of the activity proved to be more than 90% (within experimental error and available compounds) qualitatively (yes, no) and about 60%−70% quantitatively.

INTRODUCTION

Methods of pharmacophore identification and pharmacophore-based bioactivity prediction are in great demand and fast development in modern theoretical and computational chemistry.[1] The pharmacophore (Pha) is usually defined as a group of specific atoms in a given geometric arrangement that are deemed to exert the activity under consideration in a series of compounds. To locate the Pha means to be able to identify the presence or absence of the bioactivity. In this form the Pha serves as a tool for qualitative prediction of activity (yes, no) without any indication of its magnitude quantitatively.

Such a pure qualitative measure of activity may be insufficient for practical use. For instance, if the Pha is present but the activity is small, the compound will be classified experimentally as inactive (comparison of data from different sources shows that the experimental error may be 1−2 orders of magnitude). The heuristic value of the Pha concept would increase significantly if complemented with a *quantitative or a semiquantitative* measure of predicting the activity, at least by orders of magnitude.

In the electron-conformational (EC) method of Pha identification suggested earlier,[2−8] the molecular system is described by means of a matrix, the electron-conformational matrix of congruity (ECMC), in which the matrix elements are parameters of its electronic structure and geometry. Then the Pha is identified by a small number of matrix elements that are approximately the same for all the active compounds of the chosen training set, and not present in the same combination in the inactive compounds (see below). The values of the matrix elements that represent the Pha may vary from one compound to another within some limits (tolerances).

Based on the EC method, we are able to extend and improve the idea of Pha by introducing the following three important extensions:

(1) The Pha should be defined by not just atoms from the periodic table, but by appropriate atom-in-molecule electronic characteristics which may be the same for different atoms and different for the same atoms in different compounds. Although, in general, this statement is quite understandable, the EC method allows for a direct implementation of this definition of Pha.

* To whom correspondence should be addressed. E-mail: bersuker@eeyore.cm.utexas.edu. Telephone: (512) 471-4671. FAX: (512) 471-8696.
† Department of Chemistry and Biochemistry.
‡ College of Pharmacy.

10.1021/ci000453n CCC: $19.00
Published on Web 09/07/2000

(2) Both the electronic characteristics and geometric parameters of Pha vary from one active compound to another within certain limits (tolerances), and the magnitude of the activity may be a function of these variations.

(3) The Pha is a necessary but not sufficient condition of activity. In the presence of Pha the activity of the molecule may be diminished (partially or completely) by antipharmacophore shielding groups (APS) that hinder its proper docking with the bioreceptor, or enhanced by other auxiliary groups (AG) that provide for other properties (e.g., hydrophobicity). Hereafter we call them all AG.

The influence of AG can be parametrized to approximately take into account their role in the activity. Similar to Pha, this parametrization is based on describing AG by the same electronic and geometric parameters known from ECMC. Then by processing these parameters for the training set in comparison with the activities and performing a least-squares minimization procedure, we get the constants that represent the weight of each of these parameters in the activity. Adding also the weight (relative population) of each conformation as a function of its energy and temperature, we obtain a formula for quantitative prediction of the bioactivity. In this way the EC method, by complementing the Pha with parametrized AG, elevates the idea of Pha from a qualitative to a semiquantitative tool for bioactivity prediction.

The quantitative aspect of the EC method is improved also by introducing a special atomic index of electron-donor (proton-acceptor) interaction (an index of orbital and charge controlled interaction, *II*) to better characterize the interaction of the corresponding atom-in-molecule of the ligand with the bioreceptor.

A special novelty introduced to the EC method, as compared with its version in the previous papers, is also the handling of the multiconformation problem. The latter is a significant burden for all the methods that consider ligand−receptor, or substrate−enzyme (S−E) interaction. The solution presented in this paper seems to be of general importance to all these methods.

In this paper we present the improved EC method and its application to the problem of angiotensin converting enzyme (ACE) inhibitors. The Pha identification and bioactivity prediction of ACE inhibitors is one of the most difficult problems in this field, because these inhibitors are relatively large molecules with many low-energy conformations (cf. rice blast activity[2−4]). The complexity of the ACE inhibitors for us is augmented by the fact that the experimental data on ACE inhibitors are usually confidential (so are the data for many other drug systems), and we are thus restrained to use data published long ago (in particular, we could not find any information on experimentally tried molecular systems which proved to be inactive). By solving this rather complex problem in unfavorable conditions (lack of access to all experimental data), we demonstrate the possibilities of our method and claim that more accurate predictions of bioactivity would be possible if experimental data were available.

After describing briefly the main features and procedures of the method as a whole, partly described earlier,[2−4] we introduce the *II* and the handling of the multiconformation problem. Then we demonstrate the details of the method by applying it to an important problem of ACE inhibitors and discuss the results: the Pha obtained, the AG groups, their parametrization, numerical values for the variational (adjustable) constants obtained from the training set, and the validation of the results by predicting the activity of new compounds.

THE METHOD

(a) Conformations and ECMC. The essence of the electron-conformational (EC) method is as follows. Assume we have a series of molecules with known biological activity or inactivity (the training set). We require that the test of activities (inactivities) be carried out by the same experimental method for all compounds of the training set, as well as for new compounds with which the activities predicted by this method are compared. We assume also that there is only one (unique) Pha that is responsible for the activity under consideration.

First we evaluate approximately the energies of the conformations of the molecules in the training set and separate those which are expected to be significantly populated at room temperatures (see below). Second, we calculate the electronic structure of the set of molecules in each of their populated conformations. Then for each molecule (conformation) the EC matrix of congruity (ECMC) is constructed.

The ECMC is a square matrix which is symmetric with respect to the diagonal elements ($a_{ij} = a_{ji}$); the upper half of it is shown in Figure 1 for the compound M21 (captopril) of the training set of ACE inhibitors. The whole number of independent elements in this matrix is $m(m + 1)/2$, where m is the number of atoms in the molecule. For the diagonal elements a_{ii}, an atomic parameter describing its electronic properties in the molecule is chosen (one at a time) from atomic charges (in the old version), index of orbital and charge controlled interaction, or in brief, interaction index (*II*) (see below), valence activities, polarizabilities, etc. The off-diagonal elements a_{ij} are of two kinds: (1) if i and j label two near-neighbor atoms that are chemically bonded, then a_{ij}* may be one of the electronic parameters of the $i-j$ bond, e.g., the bond order, Wiberg index, bond energy, or polarizability; (2) if i and j label nonbonded atoms, then $a_{ij} = R_{ij}$ is their interatomic distance. In this way, each ECMC contains both electronic (a_{ii} and a_{ij}*) and geometric (R_{ij}) characteristics, which are deemed to represent fully enough the properties of the conformation of the molecule under consideration in its interaction with the bioreceptor.

In the simple version of this approach employed earlier[5−8] only one conformation was considered, and only atomic charges and bond orders were used as the electronic parameters a_{ii} and a_{ij}*, respectively. In principle, keeping the same interatomic distances R_{ij} of the given conformation, different combinations of any electronic parameter a_{ii} with any a_{ij}* from both those mentioned above and others can be tried in order to choose the best electronic description.

(b) Index of Electron-Donor (Proton-Acceptor) Interaction. As outlined above, the diagonal elements of the ECMC in the EC method are assumed to characterize the electronic properties of the corresponding atoms. In the applications of this method so far we employed as such the most obvious electronic property of atoms, the atomic charges. However, neither charges nor their derivatives (the electrostatic fields) seem to be the best possible parameters

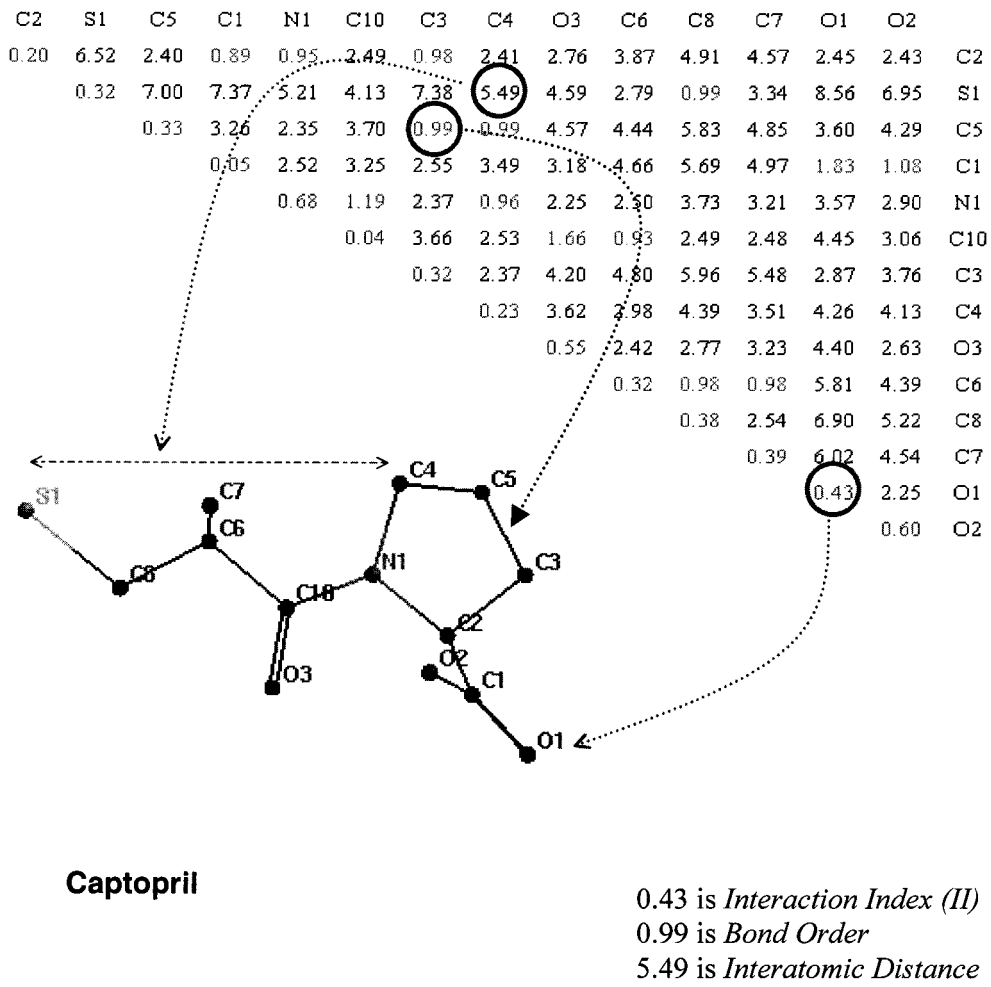


Figure 1. General appearance of the electron-conformational matrix of congruity (ECMC) illustrated, by way of example, for the compound M21 (captopril) in the ACE inhibitor problem.

for representing the atoms of the organic system in its interaction with the environment. Indeed, when neutral biological systems are considered (e.g., drug−receptor interaction) the atomic charges are usually not very large and the interaction is rather of donor−acceptor nature. This notion includes also formation of hydrogen bonds which can be described approximately as due to the interaction of (strong enough) electron-donor atoms with proton donors (poor electron donors). Note that charges, too, reflect the ability of atoms to donate electrons and/or accept protons, but in a rather rough approximation which may not be sufficient for a more refined molecular modeling.

In general, there are methods for calculating of donor−acceptor interactions between molecular systems, but they are rather cumbersome and are not suited for modeling molecular properties in general. The problem is to find a relatively simple parameter that might approximately present the electron donor−acceptor properties of atoms in their interaction with a biological receptor to be used as an index in molecular modeling and QSAR problems. Our idea of such an index is outlined below in this section.

A general expression for the energy of intermolecular interaction is given by the Fukui−Klopman formula:[9]

$$\Delta E_{st} = \frac{q_s q_t}{\epsilon R_{st}} + 2\sum_{n,m} \frac{(c_s^m c_t^n \beta_{st}^{mn})^2}{E_s^m - E_t^n} \qquad (1)$$

where the first term stands for the electrostatic interaction. Here q_s and q_t are the effective atomic charges of the *s* and *t* atoms, respectively, R_{st} is the distance between them, and ϵ is the effective dielectric constant. The second term describes the covalent interactions with c_s^m and c_t^n as the LCAO coefficient of the MO in which the corresponding AO of atoms *s* and *t*, respectively, take part in their molecules, and β_{st}^{mn} and $E_s^m - E_t^n$ are the resonance integral and energy gap between these two AOs, respectively.

This expression includes both electrostatic (the first term) and donor−acceptor (second term) interactions, the latter being determined by perturbation theory based on the MO LCAO calculated electronic structure of the separate molecules. Obviously, eq 1 contains the parameters of both molecular systems. If the electronic structure of these two systems is known, the energy of their interaction can be estimated by using this formula. However, in the majority of cases of molecular modeling (e.g., for QSAR problems) the target system is unknown, and the problem is to get an index which would approximately characterize the given molecule with respect to the interaction with possible targets. For this purpose we modify slightly our goal: we now look for an atomic index of *maximum possible interaction* with the virtual target atom. Such an index would characterize different atoms of the molecular system under consideration with respect to their ability to exert an interaction with other atoms.

Following eq 1, maximum interaction is achieved when $E_{ns} \approx E_{nt}$, i.e., when the donor energy levels (HOMO) of the molecule under consideration are close to (or coincide with) the LUMO of the target, or vice versa. In this case the donor−acceptor term (the second term in eq 1) is much larger than the electrostatic term, which may be neglected.[9] For the assumed $E_{ns} \approx E_{nt}$, perturbation theory for degenerate states should be employed in the deduction of eq 1, which results in the expression[9]

$$\Delta E_{\max} \approx 2c_s^m c_t^n \beta_{st}^{mn} \tag{2}$$

Note that by eliminating the electrostatic term we lower the absolute value of ΔE, while by taking $E_{ns} \approx E_{nt}$ (they are usually different) we increase this value. Thereby, the approximate eq 2 emerges with some compensation of errors.

The coefficients c_s^m and c_t^n determine the electron population of the appropriate atomic orbitals, while the resonance integral depends on the wave functions of both interacting atoms. If the target atom is known, $\Delta E_{\max}$ can be calculated with needed accuracy. In our case the target atoms are unknown. Following our goal of comparing different systems in their interaction with the same target system, we look for the relative contributions of different atoms to $\Delta E_{\max}$ *assuming that the target atom is the same for all of them*. In other words, we look for an atomic index that characterizes the relative ability of different atoms to exert donor−acceptor interactions with a standard target.

In addition to the population of the atomic orbitals that form the HOMO, the relative contribution to β depends on how fast the wave functions of these orbitals decrease in the region of overlap with the wave function of the target atom. Indeed β is strongly dependent on this overlap (approximately $\beta \sim S$, where S is the overlap integral). From general quantum mechanics of atoms, it is well-known that at relatively large distances, *at the outside of the atom*, the wave function is decreasing exponentially and the exponential power is similar to that of a hydrogen-like atom, i.e.

$$\Psi_i \sim \exp(-\sqrt{-2E_i}R) \tag{3}$$

where E_i is the effective orbital energy of the outermost electron of the atom in the given molecular environment and R is the distance from its outer maximum. In fact the function (3) reflects approximately the rate of falling electron density outside the atom as a function of the electron binding.

Based on the formulas 2 and 3, we suggest the following atomic index of relative donor−acceptor reactivity of the atom A of the molecular system under consideration, abbreviated as *interaction index* (*II*):

$$II^{\mathrm{A}} = g^{\mathrm{A}} \exp(-\sqrt{2\alpha^{\mathrm{A}}}R_0) \tag{4}$$

where g^{A} is the electron population of the outer orbital (1s for H, *n*p for the second- and third-row elements) in the atom A, α_i^{A} is the energy of ionization of the electron from this orbital in the valence state (EIVS) of the atom in the molecule, and R_0 is the distance from the maximum density to the point of assumed maximum overlap with the wave function of the target atom.

The EIVS α_i^{A} is a well-known atom-in-molecule parameter that is easily calculated as the orbital ionization energy which depends on the atomic charge and electron configuration.[10] If the electronic structure of the molecule is determined, the g and α values can be obtained directly. The R_0 value is conventional; in fact it determines the scaling of the II^{A} values. We believe that the value of $R_0 = 2$ au is rather appropriate. It means that the maximum overlap with the target atom wave function takes place at ∼4 au, ≈2.1 Å from a second-row atom (∼2 au from the nucleus to the maximum density plus ∼2 au to the maximum overlap), and ∼3 au (≈1.5 Å) from a hydrogen atom. In practical applications, several R_0 values can be tried to find the best description.

With this parametrization the values of *II* range between 0 and 1.0. Calculated *II* values for different atoms in ACE inhibitors are given below. The highest values of *II* characterize the strongest donors of electrons or acceptors of protons, while the lowest values of *II* indicate possible proton donors (electron acceptors). Note that this index is strongly dependent on atomic charges (α_i is charge dependent), and this explains why atomic charges are successful in presenting the electronic properties of atoms in the previous versions of the EC method in QSAR problems. However, the *II* values are deemed to be much more effective.

(c) Pha Identification. With the calculated electronic and geometrical parameters, each heavily populated conformer of each molecule with *m* atoms is described by its ECMC, where *m* varies from one molecule to another. The next step is to process all the ECMC in comparison with the activities of the corresponding molecules and to reveal those matrix elements which, within a given tolerance (see below), are common for all the active compounds and are absent in the same arrangement in the inactive compounds. In this way a smaller number of matrix elements, common to all the active compounds, is separated thus forming the EC submatrix of activity (ECSA).

It is assumed that the resulting ECSA represents the Pha of the activity under consideration. Since the matrix elements describe corresponding electronic and conformational features, the ECSA in principle provides information about the active conformation of the molecule and its active site (the Pha) in terms of charge and interaction-activity distribution in space. The description of the Pha provided by the ECSA can be used directly in qualitative (yes, no) screening of new compounds for the activity under consideration. It can also be employed to formulate a search query which, in turn, could be used to search one or more large databases of three-dimensional chemical structure.[11]

An important question is whether the ECMC can describe chirality, i.e., whether this matrix description allows one to distinguish enantiomers. The answer is yes, enantiomers are described by different ECMC: although the absolute values of the matrix elements of the two enantiomers are the same, their mutual positions in the two matrixes are different, and no transposition of the columns or rows (no change in the order of numeration of atoms) is allowed (ECMC matrixes are not determinants). This means that if the ECSA (the Pha) is present in one of the enantiomers, it may not be necessarily present in the other one or, if present, it will be accompanied by different AG positions (see below) resulting in different activity. It is important how the numeration of atoms in the molecular system is done (e.g., "clockwise" or "anticlockwise" in the Pha plane), because this determines the

arrangement of the matrix elements in the ECMC. Once chosen, this way of numeration must be preserved for all conformations and enantiomers. An example in our previous publication[2] demonstrates in more detail how and when two enantiomers have different activities.

Another important feature of the ECSA is that its parameters are defined within certain limits (tolerances) which are determined as those which best separate the active compounds from the inactive ones. The tolerances and other flexibilities in the Pha (see below) reflect the possible flexibility of the receptor.

By introducing the APS[2−4] and other AG that may influence the activity, the whole Pha identification procedure is improved. Based on the assumption of one Pha, the set of systems for the initial Pha identification can be reduced in size because it can be chosen from some of the active and inactive compounds having the simplest structure with the smallest number of conformations. Then the change of the activity in the remaining compounds may be attributed to either AG (in the presence of Pha) or lack of Pha. Note that as mentioned above, while shielding groups reduce the activity, other AG may enhance the activity, for example, by contributing to the required hydrophobicity. In general, there is no a priori division of the AG groups in "shielding" and "enhancing" the substrate−enzyme (S−E) interaction, and only the final optimization of the parameter values shows the kind of different AG contributions.

The choice of this smaller training set for a preliminary Pha identification should be done by inspection. The resulting Pha obtained from this smaller set is considered as the first approximation to Pha: without comparison with the remaining compounds there is no guarantee that it does not contain extra (with respect to minimum necessary) atoms (the same doubts remain when the training set is not complete in the sense of sufficient diversity). Also these further comparisons, especially with inactive compounds, allow one to determine the tolerances for the ECSA parameters.

(d) Auxiliary Groups and General Formula of Activity. To determine the AG, we have to examine the structures (conformations) of the active compounds by their superposition. The Pha defined above includes the minimal group of atoms (more precisely, certain three-dimensionally arranged charges or reactivity sites) required for the activity. The presence of molecular groups, additional to this minimum set of Pha, may diminish (destroy) the activity and/or provide for hydrophobicity or another property that may enhance the activity. In specific cases, the analysis of the group of active molecules shows that there may be considerable enlargement of the Pha that does not influence the activity. The largest group of atoms that includes the Pha and does not influence the activity is defined as the *basic skeleton*. This notion proved to be important in the rice blast activity problem solved recently.[2−4] In the ACE inhibitor problem under consideration in this paper, the notion of a basic skeleton seems to be less important.

The AG can be defined as atoms or groups of atoms located outside the Pha that cause improper docking with the bioreceptor, or which enhance the activity, for instance, by inducing conformational changes in the bioreceptor. If the latter is known, the effect of AG can be evaluated directly, at least in principle. In our cases of an unknown bioreceptor, this problem is much more complicated. Similar to the idea of Pha, the problem can be approached based on statistical comparisons of their structural and electronic parameters with the relative values of activity. We suggest a general scheme for evaluation of the role of AG which allows for at least qualitative or semiquantitative estimations. The main idea (somewhat similar to that involved in the Pha identification) is to describe each AG by means of structural and electronic parameters and to reveal their role by a minimization procedure as is usually done in QSAR problems. First we take into account that the AG reduce or enhance the S−E binding by an amount E' which reduces (or increases) the activity by a factor of $\exp(-E'/kT)$. We denote $S_{ni} = E_{ni}'/kT$ and introduce the function S in a general way as follows:

$$S_{ni} = \sum_{j=1}^{N} \kappa_j a_{ni}^{(j)} \tag{5}$$

where $a_{ni}^{(j)}$ are AG parameters that describe the jth kind of AG in the ith conformation of nth compound, N being the number of chosen AG parameters, and κ_j are variational (adjustable) parameters. Using this function and taking into account the Boltzmann population of each conformation as a function of its energy and temperature T, we get the following general formula of activity:

$$A_n = A_0 \frac{\sum_{i=1}^{m_n} \delta_{ni}[\mathrm{Pha}] e^{-S_{ni}} e^{-E_{ni}/kT}}{\sum_{i=1}^{m_n} e^{-E_{ni}/kT}} \tag{6}$$

Table 1. The Activities pIC_{50} and Number of Conformations for the 51 Compounds in the Training Set of ACE Inhibitors Shown in Figure 4[a]

molecule	pIC_{50}	no. of conformations	molecule	pIC_{50}	no. of conformations
M01*	9.64	16	M27	6.70	22
M02*	9.22	50	M28	6.37	50
M03*	9.00	28	M29	6.34	39
M04*	8.96	35	M30	6.19	20
M05*	8.92	50	M31	6.15	50
M06*	8.92	50	M32	6.15	3
M07*	8.77	38	M33	6.11	43
M08*	8.55	18	M34	6.07	50
M09*	8.54	14	M35	5.80	49
M10*	8.54	50	M36	5.62	50
M11*	8.52	50	M37	5.62	49
M12*	8.52	50	M38	5.55	41
M13*	8.43	14	M39	5.52	30
M14*	8.40	50	M40	5.31	50
M15*	8.22	35	M41	5.08	19
M16*	8.15	50	M42	4.96	18
M17*	8.11	50	M43	4.77	47
M18*	8.05	19	M44	4.66	13
M19*	8.00	48	M45	4.51	50
M20*	7.92	50	M46	4.41	50
M21*	7.64	22	M47	4.32	34
M22	7.42	9	M48	4.28	50
M23	7.31	25	M49	3.89	50
M24	7.30	37	M50	3.72	32
M25	7.19	17	M51	3.64	11
M26	7.00	9			

[a] The molecules used in the initial Pha training set are marked by *.

where δ is a kind of Dirac δ-function:

$$\delta_{ni}[\text{Pha}] = \begin{cases} 1, & \text{when Pha is present} \\ 0, & \text{when Pha is absent} \end{cases} \quad (7)$$

A_0 is a constant (see below), and m_n is the number of conformations of the nth molecule (Table 1).

In this formula, only those conformations that have the Pha ($\delta_{ni} \neq 0$) contribute to the activity under consideration, and these contributions are weighted in accordance with the contribution of AG and relative numbers of molecules in the active conformations. These numbers decrease rapidly with the energy of the conformation E_{ni} (at $E_{ni} \sim 2.8$ kcal/mol the number of conformations becomes less than a 0.01 part of those in the lowest conformation at $E_{n0} = 0$). In the next section we describe how we handle the multiconformation problem.

To determine the A_0 constant, we choose a reference molecule (l) from the training set for which the activity is known and calculate A_l after eq 6:

$$A_l = A_0 \frac{\sum_{i=1}^{m_l} \delta_{li}[\text{Pha}] e^{-S_{li}} e^{-E_{li}/kT}}{\sum_{i=1}^{m_l} e^{-E_{li}/kT}} \quad (8)$$

Determining A_0 from this equation and substituting it in eq 6, we get

$$A_n = A_l \frac{\sum_{i=1}^{m_l} e^{-E_{li}/kT}}{\sum_{i=1}^{m_n} e^{-E_{ni}/kT}} \frac{\sum_{i=1}^{m_n} \delta_{ni}[\text{Pha}] e^{-S_{ni}} e^{-E_{ni}/kT}}{\sum_{i=1}^{m_l} \delta_{li}[\text{Pha}] e^{-S_{li}} e^{-E_{li}/kT}} \quad (9)$$

Using the experimental data for the activities of the compounds in the training set, we estimate the adjustable constants κ_j by performing a least-squares minimization operation on the function $|A_n^{\text{calc}} - A_n^{\text{exp}}|^2$. With the constants κ_j determined in this way, we can evaluate the expected activity of any molecular system using eq 9. In this formula, only the choice of the $a_{ni}^{(j)}$ parameters remains uncertain. It requires some experience and skill. In a previous paper we have shown how to do it for rice blast activity.[2-4] Before going to the ACE problem, we present the significant progress achieved in handling the multiconformation problem which simplifies essentially eqs 8 and 9.

(e) Handling the Multiconformational Problem. The issue of multiple conformations for each molecular system as a candidate of substrate−enzyme (S−E) interaction is problematic. For instance, in the EC method under consideration in this paper, all significantly populated conformations should be subjected to the electronic structure calculations and construction of ECMC, which then should be processed to reveal the presence (or absence) of the Pha and the AG in order to estimate the activities using eq 9. The estimation of the AG parameters is especially time-consuming because so far we have not yet found an algorithm to extract them automatically from the ECMC, so it must be done individually for each conformation. An additional uncertainty is introduced by the possibility of conformational changes during the interaction with the receptor, which seemed to be uncontrolled, especially when the receptor is unknown.

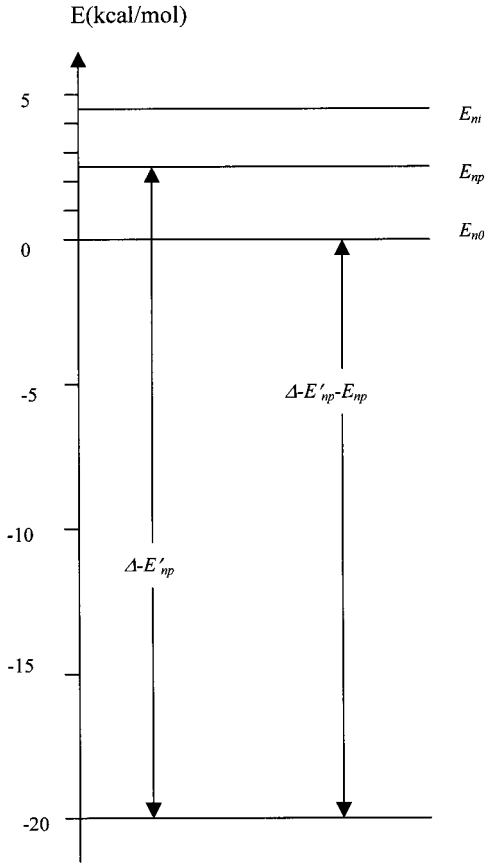


Figure 2. Energy comparison of conformations in the nonbonded and bonded states. Δ is the maximum possible SE bonding in the presence of Pha (p conformation) and E'_{np} is the reduction of bonding energy due to the auxiliary groups. The energy of any conformation after thermal relaxation to the p one and consequent bonding is $E_{ni} - \Delta + E_{np} + E'_{np}$.

We solved this problem by considering some general features of the S−E interaction. First we take into account that there is a preequilibrium (Michaelis−Menten mechanism[12]) in enzyme-catalyzed reactions; i.e., the intermediate enzyme−substrate bound state SE reaches an equilibrium with the reactants E and S,[12] and hence we can assume a Boltzmann distribution of such SE bounded sites with E and S. An important point is that the SE binding energy is of the order of 10−20 kcal/mol, which is much larger than the energy difference $\delta \sim 2-3$ kcal/mol between the significantly populated conformations. Under this condition, if at least one conformation of the ligand has the Pha, *all the other conformations will undergo transformation (relax)* to the latter one in order to bind to the receptor and *to gain energy*.

Denote the conformation with Pha by index p, its energy being E_{np} (read off the ground state conformation with $E_{n0} = 0$). If there are several conformations with Pha, p is the one with the best binding to the receptor, usually that of lowest E_{np}. After binding to the receptor, the energy of this conformation in the bonded state becomes $E_{np} + E_{np}' - \Delta$, where E_{np}' is the contribution of AG, mentioned above, and Δ is the S−E bonding energy. Since $\Delta > \delta$, this bonded state is the lowest in energy (Figure 2); the energies of all the other conformations after thermal relaxation to E_{np} and

bonding become equal to $E_{ni} - [\Delta - (E_{np} + E_{np}')]$. With these energies the Boltzmann distribution in eq 8 yields

$$A_n = A_0 \frac{\sum_i \exp[(-E_{ni} + \Delta - E_{np} - E_{np}')/kT]}{\sum_i \exp(-E_{ni}/kT)} = A_0 \exp[(\Delta - E_{np} - E_{np}')/kT] \quad (10)$$

The value exp Δ is a constant and can be included in A_0. Employing the previous denotations $S_{np} = E_{np}'/kT$, we have

$$A_n = A_0 e^{-E_{np}/kT} e^{-S_{np}} \quad (11)$$

As in the previous section, A_0 can be determined by comparison with a reference compound l for which the activity A_l is known. Then

$$A_n = A_l e^{-(E_{np}-E_{lp})/kT} e^{-(S_{np}-S_{lp})} \quad (12)$$

Thus only the lowest energy conformation with Pha should be parametrized in order to estimate the bioactivity. This is an essential simplification of the formula of activity as compared with eq 9. Because of the exponential factor $\exp(-E_{np}/kT)$, high-energy conformations may be neglected (at $E_{np} \sim 1.5$ kcal/mol at room temperatures $\exp(-E_{np}/kT) \sim 0.08$). In many cases the Pha is present in the ground state conformation (see the example of ACE inhibitors below).

The overall computational scheme of this method is illustrated in Figure 3. The main computational algorithms and programs needed for its realization are based on existing commercial packages for electronic structure calculations and molecular mechanics simulations (e.g., Spartan[13]). The additional programs have been written in the Matlab[14] environment and C/C++.

The electronic structure calculations may be performed by either semiempirical or some low-level ab initio method. The latter is preferable when there are difficulties in fine-tuning of the tolerances of the ECSA parameters. In the examples below, 3-21G(*) ab initio single point calculations were performed.

THE PHARMACOPHORE OF ACE INHIBITORS

Angiotensin converting enzyme (ACE) inhibitors are among the most important (and perhaps most studied) blood pressure regulators and serve also as other heart-related drugs (see refs 15−32 and references therein). There have been numerous studies seeking the origin of ACE inhibitor activity of a variety of compounds.[15−20,33−40] In the majority of these works the main goal was to reveal the Pha. A limited number of chemical compounds that are ACE inhibitors were reported in the literature (see refs 15−40 and references therein). As mentioned above, much data, especially the recent results, are unavailable for public use. We chose the training set of ACE inhibitors for Pha identification and activity prediction as shown in Table 1.[33,34] Because of the large error in experimental determination of the activity, we do not include the weakly active compounds keeping only 6 orders of magnitude of the activity, from 9.64 to 3.64, measured as pIC_{50}. Table 1 contains 51 diverse compounds with structural skeletons shown in Figure 4.

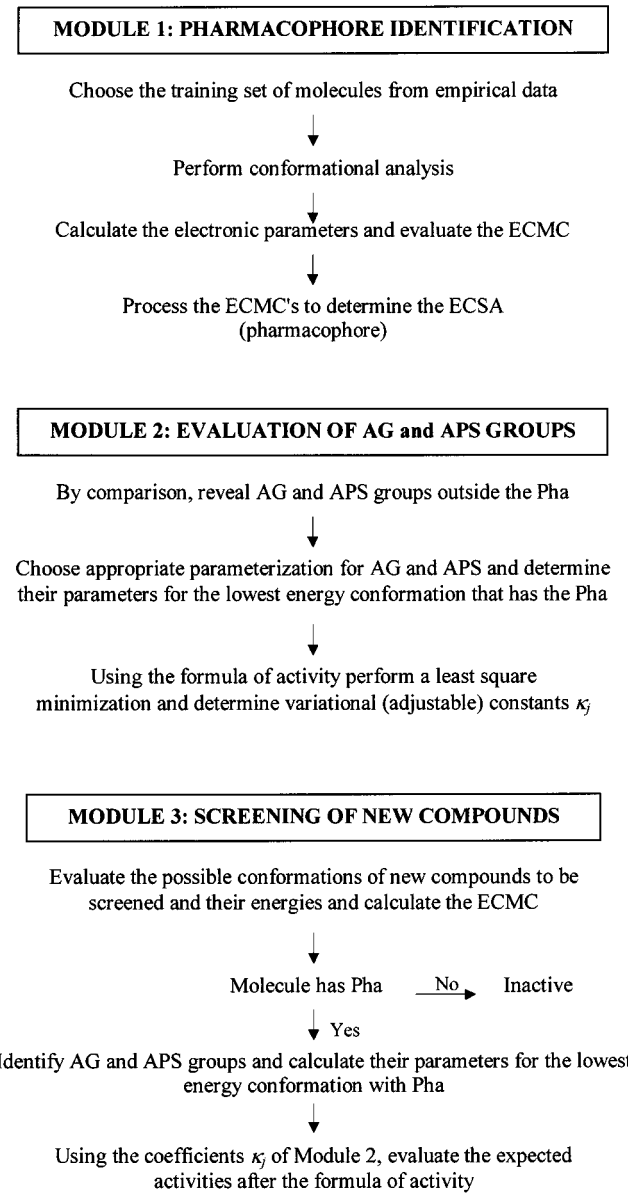
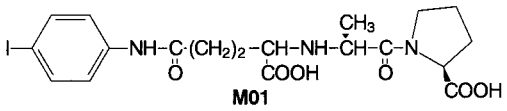


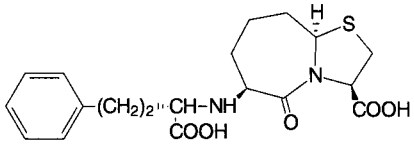
Figure 3. Algorithms and computer implementation of the EC method.

In accordance with the computational scheme, described above, the geometric and electronic parameters of the compounds (conformations) in the training set have been calculated and the corresponding ECMCs were constructed. The conformations were generated using Spartan conformational analysis.[41] The numbers of conformations of each compound are indicated in Table 1. By way of illustration, the ECMC for captopril (compound M21 in Table 1 and Figure 4) is given in Figure 1. In this matrix, as determined by the EC method, the diagonal elements are atomic *II*s, while the off-diagonal elements are bond orders W_{ij} for chemically bonded atoms, and optimized distances R_{ij} in angstroms for chemically nonbonded pairs of atoms.

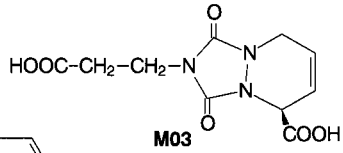
Following the general scheme (Figure 3), the next step is to process the ECMCs of the training set by means of our programs in order to reveal the ECSA. First we consider the structurally simplest but strongly active compounds of the training set that have the smallest number of conformations (the initial training set shown in Table 1). From them, we choose one of the most active and/or well-defined



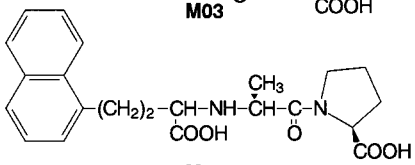
M01



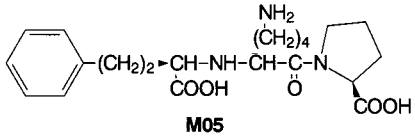
M02



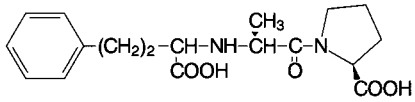
M03



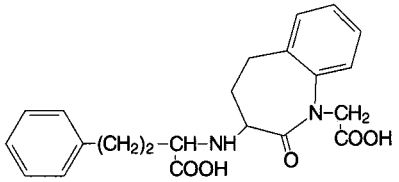
M04



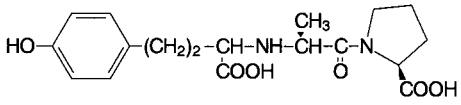
M05



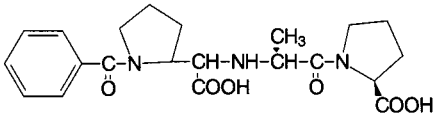
M06



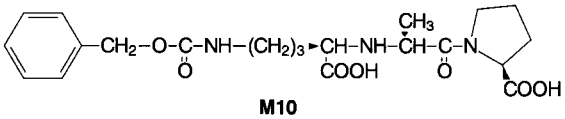
M07



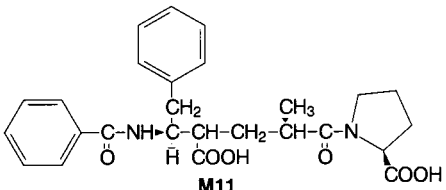
M08



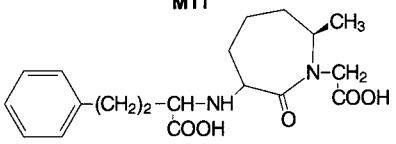
M09



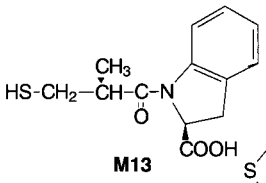
M10



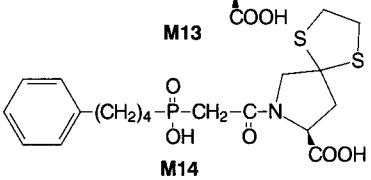
M11



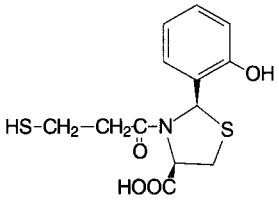
M12



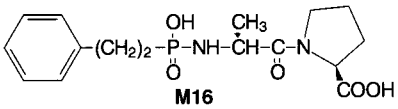
M13



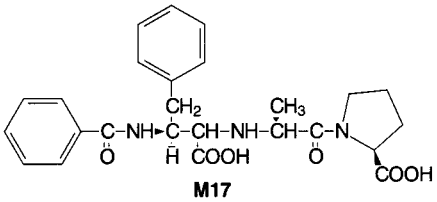
M14



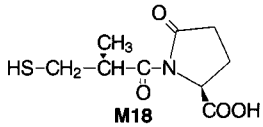
M15



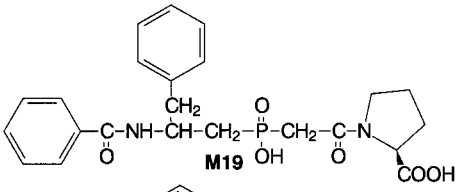
M16



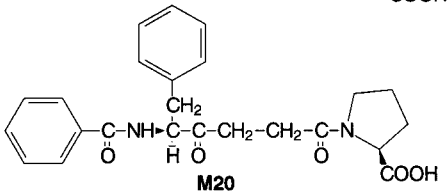
M17



M18



M19



M20

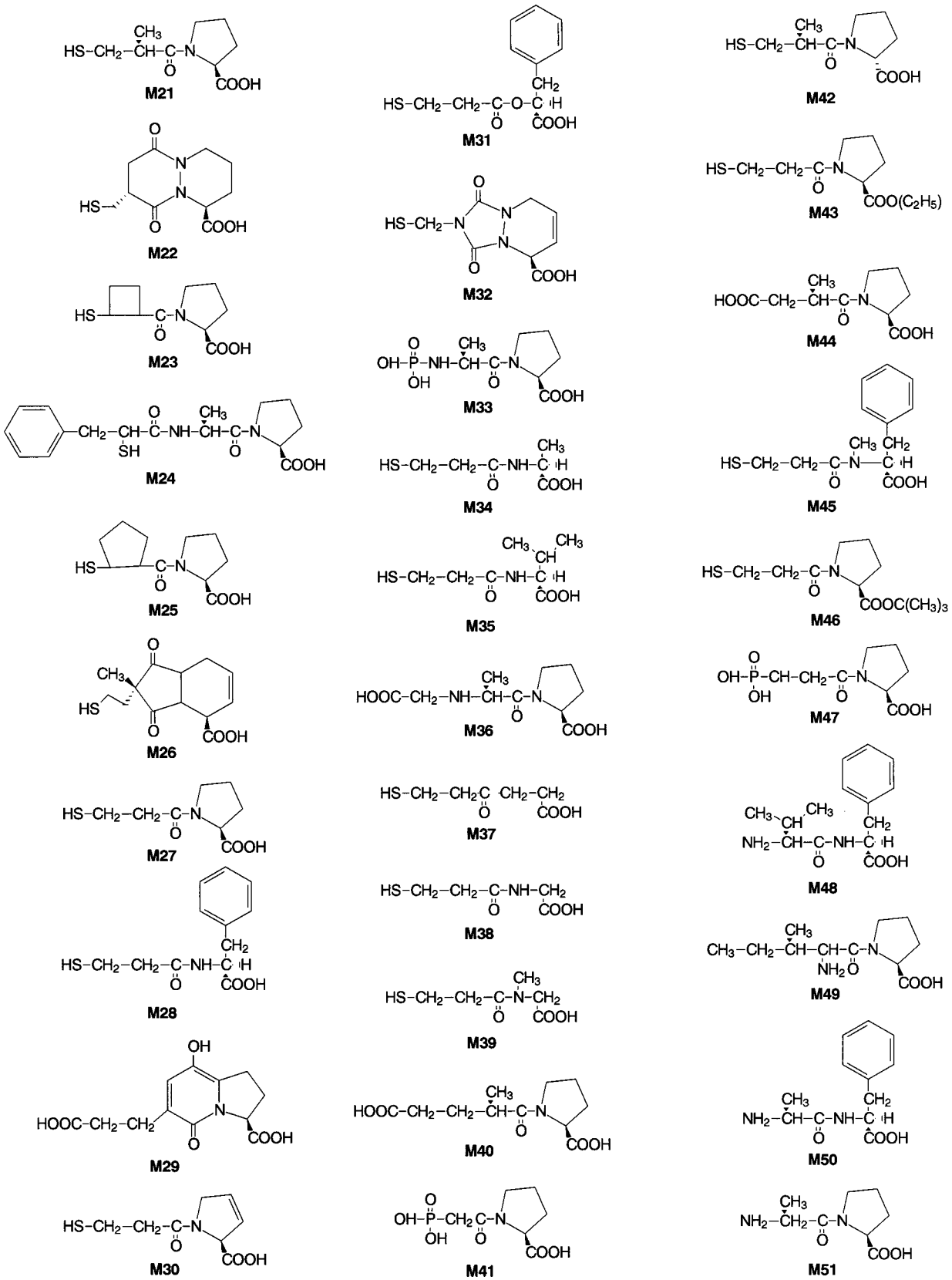


Figure 4. Chemical formulas for the 51 compounds of the training set considered in the ACE inhibitor problem (see Table 1).

(structurally most reliable) compound as a reference (M21, captopril) and compare its ECMC to that of all the others in the initial training set. In this way we determine the preliminary ECSA that may serve as a first approximation to the exact Pha.

The next step is to check the presence of the ECSA obtained from the simplest compounds in the ECMC of each of the lowest energy conformations of the remaining compounds of the training set (in fact according to eq 12 only one conformation with Pha should be found). This

X=S,N,O	O1	O2	O3	
0.38 ± 0.13	6.43±3.00	5.81±2.50	4.66±1.40	X=S,N,O
	0.47±0.04	2.30±0.06	4.47±1.18	O1
		0.54±0.08	3.40±0.83	O2
			0.51±0.05	O3

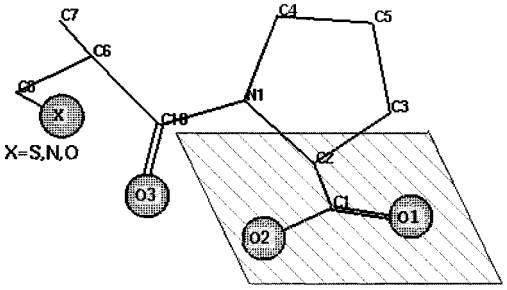


Figure 5. The active fragment (pharmacophore) responsible for ACE inhibitor activity contains three oxygen atoms, O1, O2, O3, and one atom X (X = S, N, O); more precisely, four reactivity points with limited values of *II* situated at distances shown in the submatrix of activity (ECSA).

allows us to reveal the tolerances and a new common ECSA that includes all the active compounds. Unfortunately, inactive compounds of ACE inhibitors could not be found in the literature, so the weakly active compounds were used for such comparisons. Figure 5 shows the ECSA and the corresponding atomic arrangement of the four atoms in the Pha for ACE inhibitors obtained in this way.

As seen from Figure 5 and the ECSA, the Pha has a relatively rigid triangle formed by three oxygen atoms, O1, O2, and O3 and an additional atom X (X = S, N, O) with a more flexible position. Some of the main elements of this Pha were revealed earlier by other authors.[16−20] In particular, the following three features required for ACE inhibitor activity were formulated: (1) a terminal carboxyl group (believed to interact with an arginine residue in the receptor), (2) an amido carbonyl group (with hydrogen bonds to an amide carbonyl in the receptor), and (3) a zinc-binding group. This structural information defines the minimal set of active site groups necessary for ACE inhibition that has been used to analyze databases of diverse structural classes of ACE inhibitors to determine a common three-dimensional geometry for the active site consistent with their activity.[33−40]

It can be seen that all these features are present in the above Pha. However, in addition to the purely qualitative formulation, the EC method gives distances, interaction indexes, and tolerances, all these data being obtained in a general way valid for any compounds and without any assumptions concerning the bioreceptor. Also, most importantly, the previous studies do not reveal flexible positions in the Pha and AG groups that make the prediction of the bioactivity *quantitative or semiquantitative*.

FLEXIBLE POSITIONS AND AG GROUPS. PARAMETRIZATION

The flexible positions of some of the atoms of the Pha are seen directly from the tolerances for the corresponding distances in the ECSA matrix (Figure 5). The atom X has an approximately constant X−O3 distance, but more changeable X−O2 and X−O1 distances. In addition, the superposition of structures reveals several AG hydrophobic groups, as well as shielding groups at X and O1 (Figure 6). There is also an additional oxygen atom (O5) in some of the active compounds (Figure 6).

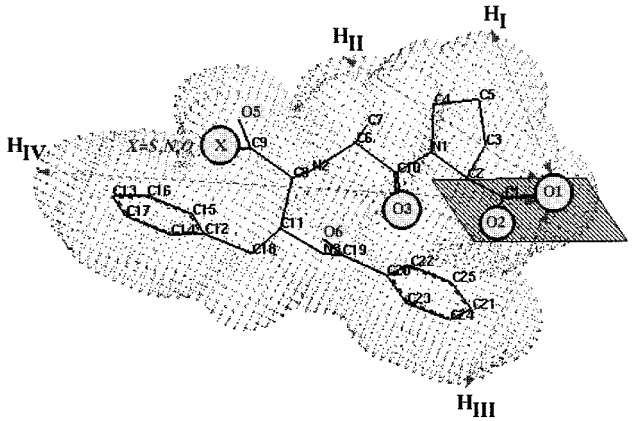


Figure 6. The scheme for estimation of the antipharmacophore shielding (APS) and other auxiliary groups (AG). The four types of hydrophobic groups H_p, p = I, II, III, IV, as well as another AG (O6) and APS (O5) atom are indicated.

After several attempts we chose the following set of parameters to characterize the AG influence after eq 12 (Figure 6).

$$a_{ni}^{(1)} = R_{ni}(O_1-O_3)$$
$$a_{ni}^{(2)} = R_{ni}(O_2-O_3)$$
$$a_{ni}^{(3)} = R_{ni}(X-O_3)$$
$$a_{ni}^{(4)} = R_{ni}(X-O_1)$$
$$a_{ni}^{(5)} = R_{ni}(O_6-O_1)$$
$$a_{ni}^{(6)} = R_{ni}(O_5-X)$$
$$a_{ni}^{(7)} = II(X), \quad X = S, N, O$$
$$a_{ni}^{(8)} = II(O_1)$$
$$a_{ni}^{(9)} = II(O_2)$$
$$a_{ni}^{(10)} = II(O_3)$$
$$a_{ni}^{(11)} = R_{ni}(H_I-O_1)$$
$$a_{ni}^{(12)} = R_{ni}(H_{II}-O_1)$$
$$a_{ni}^{(13)} = \log R_{ni}(H_{III}-O_1)$$
$$a_{ni}^{(14)} = \log R_{ni}(H_{IV}-O_1)$$
$$a_{ni}^{(15)} = \text{surface area} \qquad (13)$$

The first four parameters are just corresponding interatomic distances employed to take into account the influence of their limited flexibility on the activity; $a^{(5)}$ is the distance to the additional oxygen atom, where available; $a^{(6)}$ stands for the APS of a second oxygen (O5) when the X atom is O4 from a carboxyl group; $a^{(7)}$, $a^{(8)}$, $a^{(9)}$, and $a^{(10)}$ are the *II*s of X, O_1,

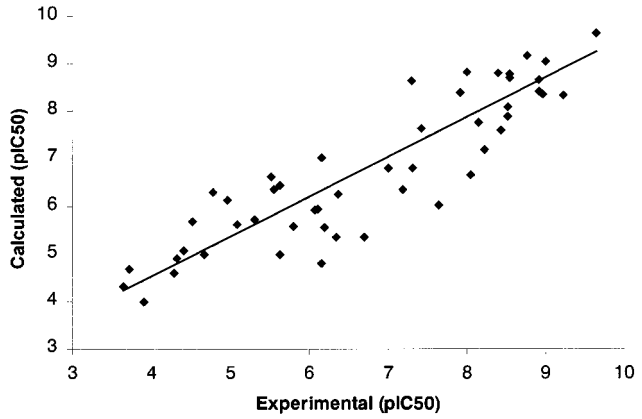


Figure 7. Calculated (optimized) vs experimental pIC_{50} values in the training set of ACE inhibitors ($r = 0.87$).

O_2, and O_3 atoms; the next four parameters take into account four types of hydrophobic groups (defined as H_p, p = I, II, III, IV) that may significantly influence the activity. The groups H_I and H_{II} are those above the [O1O2C1] plane (the numeration O1→O2→C1 is clockwise), the group H_{III} is below the plane, and H_{IV} is not specified with respect to the Pha plane. Since the positions of H_{III} and H_{IV} groups are less specific (this is seen from superposition), we characterize them by log R, where R is the distance from O_1 to the farthest carbon atom of the hydrophobic group plus 1.5 Å to take into account the C−H distance and the van der Waals radius of hydrogen. The last parameter is the surface area of the molecule which seems to influence the activity. Since the constants κ_j are different for the groups above and below the Pha plane, there is no mirror symmetry with respect to this plane, and this property is important to distinguish chirality properties.[2]

This choice of parameters is not unique, but the calculations show that other reasonable choices give similar results with the exception that an increase of the number of the parameters may improve the accuracy of prediction. The estimated values of the 15 parameters for the active p conformation of the compounds in the training set (Table 1) were used to perform a least-squares minimization of the expression $|A_n^{calc} - A_n^{exp}|^2$ as a function of the unknown κ_j constants with A_n^{calc} from eq 12 and A_n^{exp} from the experimental data. The molecule M01 was chosen as a reference compound with known activity A_l. Note that this reference compound is not the same as in the initial training set M21. Distinct from the latter, the reference compound in the parametrization and estimation of the activity should contain a reasonable amount of nonzero AG parameters to be compared with. In the compounds under consideration in Table 1 the Pha is present in many conformations including the ground state one, so the latter were chosen as the p conformations in eq 12.

The calculations yield $\kappa_1 = 17.70$, $\kappa_2 = -9.97$, $\kappa_3 = 2.24$, $\kappa_4 = -0.14$, $\kappa_5 = 0.20$, $\kappa_6 = 1.59$, $\kappa_7 = -7.49$, $\kappa_8 = -14.44$, $\kappa_9 = -11.65$, $\kappa_{10} = -18.87$, $\kappa_{11} = 1.49$, $\kappa_{12} = 0.02$, $\kappa_{13} = -0.85$, $\kappa_{14} = 9.46$, $\kappa_{15} = -1.45$. The calculated A_n versus experimental A_n^{exp} values of the activity are given in Figure 7 with resulting $r = 0.87$. These values of the constants κ_j characterize the relative significance of the parameters a_{ni}, more precisely, their deviation from the corresponding values in the reference compound M01. In particular, it is seen that from the first four distances in (eq 13) the deviations of the O1−O3 distance from that in the reference compound gives the largest contribution in the change of activity, while from the last four hydrophobic groups the deviation in the positions of H_{IV} is most important. From *II* parameters, the most effective ones are *II* of X, O_2, and O_3. Since all AG parameters are taken relative to the reference compound, no judgment about the absolute quality of AG (whether the AG decreases or increases the activity) can be done based on the sign of the κ_j values.

Table 2. The Validation Set of 21 Compounds (Figure 8) Tested as ACE Inhibitors[a]

molecule	exptl pIC_{50}	pred pIC_{50}
T01	8.66	10.08
T02	8.59	7.97
T03	8.32	8.46
T04	8.28	9.30
T05	8.19	8.83
T06	7.92	9.92
T07	7.70	6.88
T08	7.46	8.27
T09	7.40	9.12
T10	7.39	5.83
T11	7.29	6.13
T12	7.28	6.58
T13	7.25	7.42
T14	7.20	5.60
T15	7.19	9.35
T16	6.36	7.96
T17	5.59	4.15
T18	4.72	5.52
T19	4.59	4.28
T20	4.48	5.77
T21	4.17	5.40

[a] The predicted values were calculated after eq 12.

With this set of constants κ_j we can screen any compound for the activity under consideration by estimating its appropriate $a_{ni}^{(j)}$ parameter values and the activity after eq 12.

SCREENING OF NEW COMPOUNDS

For nonexperimental screening of new compounds we used a set of 18 compounds listed in Table 2 (see Figure 8) which were taken from refs 33 and 34. For the active conformations of these compounds that have the Pha (one conformation from each compound), the AG parameters were calculated and, using eq 12, the expected activities A_n were estimated. Figure 9 shows the predicted values A in comparison with the experimental data. Qualitatively (yes or no), the prediction probability is more than 90% (within the experimental errors). On the other hand, quantitatively the prediction is weaker ($r = 0.73$), which is understandable in view of the approximations made and other factors of activity not taken into account in this paper (in particular we have no data for inactive compounds to improve the tolerances in the ECSA parameter values). Although this accuracy is not very high, it is enough to complement the purely qualitative prediction based on Pha (yes or no) with a semiquantitative measure that makes it much more reliable for predictions of activity. Note also that the validation set is taken from different classes of compounds available.

CONCLUSIONS

The improved EC method outlined in this paper is based on a rather full description of the molecular system to

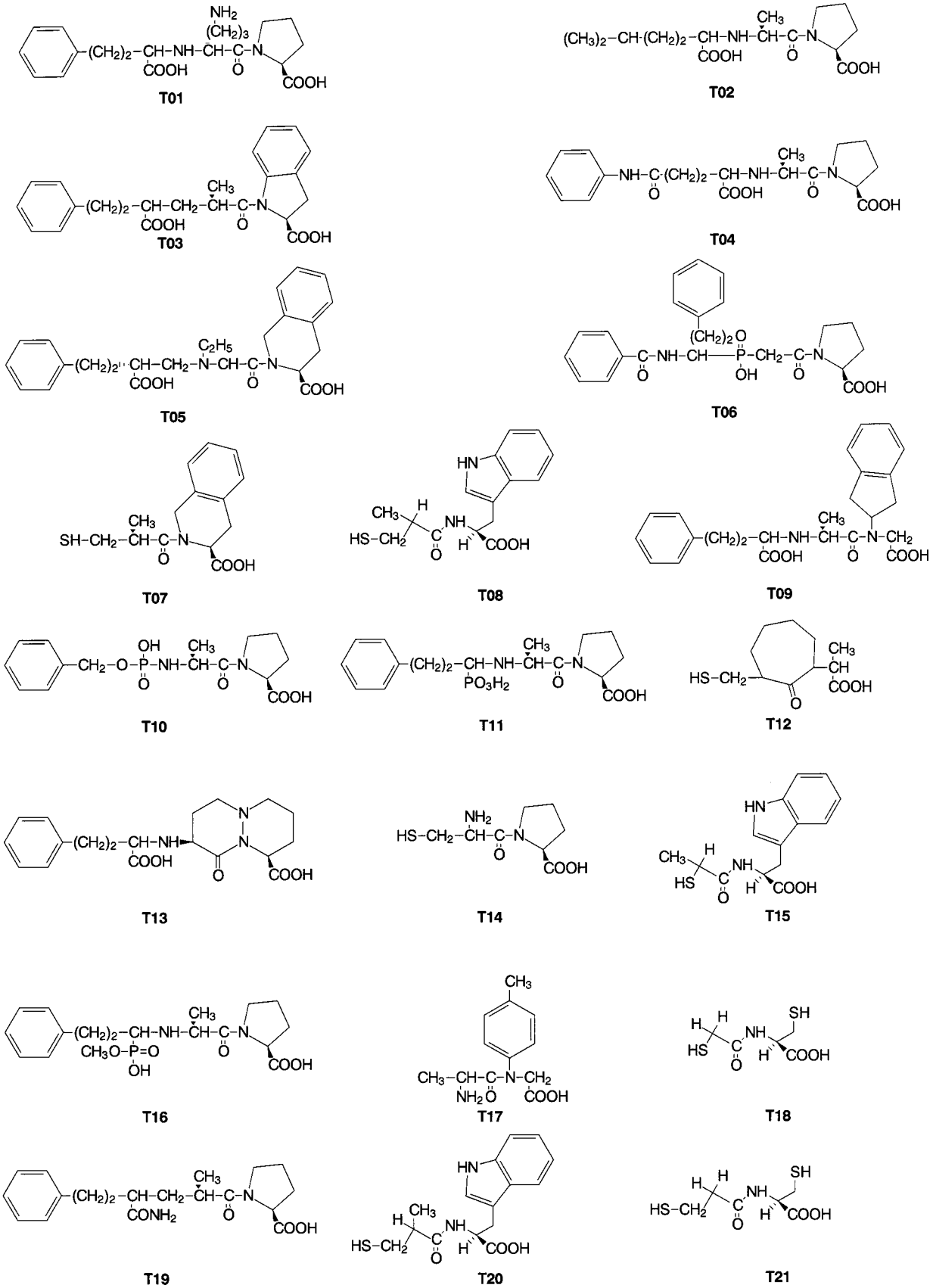


Figure 8. The set of 21 new compounds used for validation in the ACE inhibitor problem (see Table 2).

represent its interaction with the bioreceptor. Distinct from some other methods, it allows for quantitative or semiquantitative prediction of the activity of any kind, provided there are experimental data to form a sufficiently representative training set. The latter may be diverse representing quite different classes of compounds, provided they were tested on the activity under consideration by the same method (in the majority of papers in the literature compounds from the

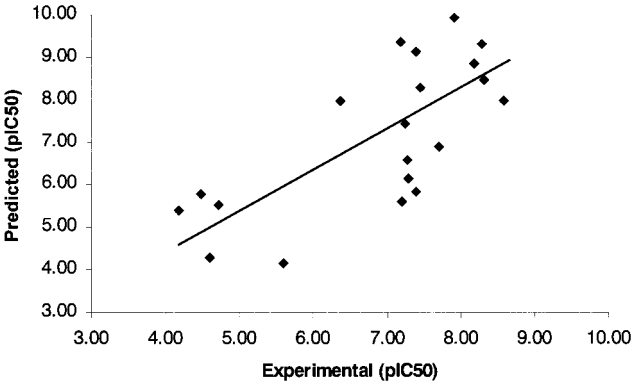


Figure 9. Validation of the EC method with a set of new ACE inhibitors (21 compounds): predicted vs experimental values of pIC_{50} ($r = 0.73$).

same or similar classes only are compared). As compared with the previous versions of this method, two major improvements are introduced: (1) an atomic index of orbital and charge controlled interaction (the interaction index, *II*) that allows for a better presentation of the ligand−receptor interaction and (2) a better handling of the multiconformation problem based on enzyme−substrate kinetics that significantly simplifies the calculations in the EC method. The treatment of the multiconformation problem is of general importance and can be used in other approaches to the problem of ligand−receptor interaction.

In applications of the EC method to specific problems, two parts may be distinguished: Pha identification and parametrization of the influence of AG. While the methodology for the former which allows for more than 90% (within experimental error) prediction of the bioactivity qualitatively (yes, no) is more perfect, the rules of AG parametrization that complement the concept of Pha to give quantitative estimates of the activity are less definitive. Fortunately, variation of the parameter choice within physically grounded limits does not change the results significantly: an acceptable accuracy can be achieved with a variety of reasonable parametrizations.

In this way a quantitative or semiquantitative (by orders of magnitude) prediction of the activity can be achieved. In general, quantitative predictions of bioactivity based on the structural and electronic properties of molecules described approximately by parameters cannot be of very high accuracy because they do not take into account a variety of macro properties (e.g., solubility, transport properties, etc.). Within these limits of accuracy the EC method of Pha identification complemented by AG parametrization seems to take into account the basics of the phenomenon that allow one to approximately predict the bioactivity quantitatively (when the empirical data for the training set are known with high accuracy and include a wide diversity) or semiquantitatively (by orders of magnitude). The results for the ACE inhibitors obtained in this paper, as well as the variety of problems solved earlier by the simpler version of this method[5−8] and partly improved version,[2−4] confirm this statement.

ACKNOWLEDGMENT

The authors are thankful to Dr. K. Lipkowitz for precious comments. This paper is based in part upon work supported by Texas Advanced Research Program under Grant No. 003658-345, and it has been supported in part by a grant from the Robert A. Welch Foundation.

REFERENCES AND NOTES

(1) Guner, O. F., Ed. *Pharmacophore Perception, Development, and Use in Drug Design*; International University Line: La Jolla, 2000.
(2) Bersuker, I. B.; Bahceci, S.; Boggs, E. J. The Electron-Conformational Method of Identification of Pharmacophore and Anti-Pharmacophore Shielding. In *Pharmacophore Perception, Development, and Use in Drug Design*; Guner, O. F., Ed.; International University Line: La Jolla, 2000; pp 457−474.
(3) Bersuker, I. B.; Bahceci, S.; Boggs, E. J.; Pearlman, R. S. An Electron-Conformational Method of Identification of Pharmacophore and Anti-Pharmacophore Shielding: Application to Rice Blast Activity. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 419−434.
(4) Bersuker, I. B.; Bahceci, S.; Boggs, E. J.; Pearlman, R. S. A Novel Electron-Conformational Approach to Molecular Modeling for QSAR by Identification of Pharmacophore and Anti-Pharmacophore Shielding. *SAR QSAR Environ. Res.* **1999**, *10*, 157−173.
(5) Bersuker, I. B.; Dimoglo, A. S. The Electron-Topological Approach to the QSAR Problem. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1991; Vol. 2, pp 423−460.
(6) Bersuker, I. B.; Dimoglo, A. S.; Gorbachov, M. Yu.; Vlad, P. F.; Pesaro, M. Origin of Musk Fragrance Activity: The Electron-Topological Approach. *New J. Chem.* **1991**, *15*, 307−320.
(7) Bersuker, I. B.; Dimoglo, A. S.; Gorbachov, M. Yu.; Greni, A. I.; Vysotskaya, L. E.; Mikhailova, T. V. Study of the Electronic and Structural Properties of the Chemical Compounds in Garlic Aroma. *Nahrung-Food* **1989**, *33*, 405−411. Dimoglo, A. S.; Gorbachov, M. Yu.; Bersuker, I. B.; Greni, A. I.; Vysotskaya, L. E.; Stepanova, O. V.; Lukash, E. Yu. Structural and Electronic Origin of Meat Odour of Organic Hetero-Atomic Compounds. *Nahrung-Food* **1988**, *32*, 461−473. Bersuker, I. B.; Dimoglo, A. S.; Gorbachov, M. Yu. The Electron-Topologic Approach to the QSAR Problem Illustrated by Inhibitor Activity for Thymidine Phosphorylase and α-Chymo-Trypsin. In *QSAR in Drug Design and Toxicology*; Hadzi, D., Jerman-Blazic, B., Eds.; Elsevier: Amsterdam, 1987; Vol. 10, pp 43−48. Bersuker, I. B.; Dimoglo, A. S.; Gorbachov, M. Yu.; Vlad, P. F.; Koltsa, M. N. Structural and Electronic Origin of Odour Properties of Organic Compounds as Revealed by the Electron-Topologic Approach to the QSAR Problem. In *QSAR in Drug Design and Toxciology*; Hadzi, D., Jerman-Blazic, B., Eds.; Elsevier: Amsterdam, 1987; Vol. 10, pp 340−342. Bersuker, I. B.; Dimoglo, A. S.; Gorbachov, M. Yu. The Electron-Topologic Approach to Structure−Activity Relationships. Inhibition of Thymidine Phosphorylase by Uracil Derivatives. *Bioorgan. Khim.* **1987**, *13*, 38−44. Bersuker I. B.; Dimoglo, A. S. Electron-Topology of Garlic's Biologically Active Thioallyl Compounds. In *First World Congress on the Health Significance of Garlic and Garlic Constituents;* Washington, DC, 1990. Dimoglo, A. S.; Bersuker, I. B.; Popa, D. P.; Kuchkova, K. I. Electron-Topological Study of Plant Growth-Regulator Activity in a Series of Analogs of Abscisic Acid. *Theor. Eksp. Khim.* **1989**, *5*, 590.
(8) Guzel, Y.; Saripinar, E.; Yildirim, I. Electron-Topological (ET) Investigation of Structure-Antagonist Activity of a Series of dibenzo-[a,d]cycloalkenimines. *J. Mol. Struct. (THEOCHEM)* **1997**, *418*, 83−91. Guzel, Y. Investigation of the Relationship Between the Inhibitory Activity of Glycolic Acid Oxidase and its Chemical Structure: Electron-Topological Approach. *J. Mol. Struct. (THEOCHEM)* **1996**, *366*, 131−137. Dimoglo, A. S.; Beda, A. A.; Shvets, N. M.; Gorbachov, M. Yu.; Kheifits, L. A.; Aulchenko, I. S. Investigation of the Relationship Between Sandalwood Odor and Chemical Structure: Electron-Topological Approach. *New. J. Chem.* **1995**, *19*, 149−154. Dimoglo, A. S.; Vlad, P. F.; Shvets, N. M.; Koltsa, M. N.; Guzel, Y.; Saracoglu, M.; Saripinar, E.; Patat, S. Electron-Topological Investigations of the Relationship Between Chemical Structure and Ambergris Odor. *New J. Chem.* **1995**, *19*, 1217−1226.
(9) Fukui, K. *Theory of Orientation and Stereoselection*; Springer-Verlag: Berlin, 1975. Klopman, G. The Generalized Perturbation Theory of Chemical Reactivity and its Applications. In *Chemical Reactivity and Reaction Paths*; Klopman, G., Ed.; Wiley: New York, 1974; pp 55−165.
(10) Bersuker, I. B. *Electronic Structure and Properties of Transition Metal Compounds: Introduction to the Theory*; Wiley-Interscience: New York, 1996.
(11) Pearlman, R. S. 3D Molecular Structures: Generation and Use in 3D Searching. In *3D QSAR in Drug Design: Theory, Methods and Application*; Kubinyi, H., Ed.; ESCOM: Leiden, 1993; Vol. 1, pp 41−79.
(12) Schulz, A. R. *Enzyme Kinetics*; Cambridge University Press: 1994.
(13) *SPARTAN*, Version 5.0.3; Wavefunction, Inc., 1997.

(14) *MATLAB*, Version 5.3; Mathworks, Inc., 1999.
(15) Opie, L. H. *Angiotensin Converting Enzyme Inhibitors*; Wiley-Liss: New York, 1992.
(16) Chen, X.; Rusinko, A. III; Tropsha, A.; Young, S. S. Automated Pharmacophore Identification for Large Chemical Data Sets. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 887−896.
(17) Hangauer, D. G. Computer-Aided Design and Evaluation of Angiotensin-Converting Enzyme Inhibitors. In *Computer-Aided Drug Design: Methods and Applications*; Perun, T. J., Propst, C. L., Eds.; Marcel Dekker: New York, 1989; pp 253−295.
(18) Petrillo, E. W.; Trippodo, N. C.; DeForrest, J. M. Antihypertensive Agents. In *Annual Reports in Medicinal Chemistry*; Robertson, D. W., Ed.; Academic Press: New York, 1989; Vol. 25, pp 51−60.
(19) Wyvratt, M. J.; Patchett, A. A. Recent Development in the Design of Angiotensin-Converting Enzyme Inhibitors. *Med. Res. Rev.* **1985**, *5*, 483−531.
(20) Andrews, P. R.; Carson, J. M.; Caselli, A.; Spark, M. J.; Woodsm, R. Conformational Analysis and Active Site Modelling of Angiotensin-Converting Enzyme Inhibitors. *J. Med. Chem.* **1985**, *28*, 393−399.
(21) Stanton, J. L.; Sperbeck, D. M.; Trapani, A. J.; Cote, D., Sakane, Y.; Berry, C. J.; Ghai, R. D. Heterocyclic Lactam Derivatives as Dual Angiotensin Converting Enzyme and Neutral Endopeptidase 24.11 Inhibitors. *J. Med. Chem.* **1993**, *36*, 3829−3833.
(22) Pascard, C.; Guilhem, J.; Vincent M.; Remond, G.; Portevin, B.; Laubie, M. Configuration and Preferential Solid-State Conformations of Perintoprilat (S-9780). Comparison with the Crystal Structures of Other ACE Inhibitors and Conclusions Related to Structure−Activity Relationships. *J. Med. Chem.* **1991**, *34*, 663−669.
(23) Hausian, 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.
(24) Yanagisawa, H.; Ishirara, S.; Ando, A.; Kanazaki, T.; Miyamoto, S.; Koike, H.; Iijima, Y.; Oizumi, K.; Matsushita, Y.; Hata, T. Angiotensin-Converting Enzyme Inhibitors. 2. Perhydroazepin-2-one Derivatives. *J. Med. Chem.* **1988**, *31*, 422−428.
(25) Hassall, C. H.; Krohn, 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.
(26) Kim, D. H.; Guinosso, C. J.; Buzby, G. C.; Herbst, D. R.; McCaully, R. H.; 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.
(27) Gruenfeld, N.; Stanton, J. L.; Yuan, A. M.; Ebetino, F. H.; Browne, L. J.; Gude, C.; Huebner, C. F. Angiotensin Converting Enzyme Inhibitors: 1-Glutarylindoline-2-carboxylic Acid Derivatives. *J. Med. Chem.* **1983**, *26*, 1277−1282.
(28) Galardy, R. E.; Kontoyiannidou-Ostrem, V.; Kortylewicz, Z. P. Inhibition of Angiotensin-Converting Enzyme by Phosphonic Amides and Phosphonic Acids. *Biochemistry* **1983**, *22*, 1990−1995.
(29) Condon, M. E.; Petrillo, E. W.; Ryono, 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.
(30) Patchett, A. A.; Harris, E.; Tristram, E. W.; Wyvratt, M. J.; Wu, M. T.; Taub, D.; Peterson, E. R.; Ikeler, T. J.; Broeke, J. T.; 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. A New Class of Angiotensin-Converting Enzyme Inhibitors. *Nature* **1980**, *288*, 280−283.
(31) Ondetti, M. A.; Rubin, B.; Cushman, D. W. Design of Specific Inhibitors of Angiotensin-Converting Enzyme: New Class of Orally Active Antihypertensive Agents. *Science* **1977**, *196*, 441−444.
(32) 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**, *25*, 5484−5491.
(33) 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.* **1993**, *115*, 5372−5384.
(34) Waller, C. L.; Marshall, G. R. Three-Dimensional Quantitative Structure−Activity Relationship of Angiotensin-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. *J. Med. Chem.* **1993**, *36*, 2390−2403.
(35) Dammkoehler, R. A.; Karasek, S. F.; Shands, E. F. B.; Marshall, G. R. Constrained Search Conformational Hyperspace. *J. Comput.-Aided Mol. Des.* **1989**, *3*, 3−21.
(36) DePriest, S. A.; Shands, E. F. B.; Dammkoehler, R. A.; Marshall, G. R. 3D-QSAR: Further Studies on Inhibitors of Converting Enzyme. In *QSAR: Rational Approaches to the Design of Bioactive Compounds;* Silipo, C., Vittoria, A., Eds.; Elsevier: Amsterdam, The Netherlands, 1991; Vol. 16, pp 405−414.
(37) 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.
(38) Bohacek, R.; De Lombaert, S.; Priestle, J.; Grutter, M. Three-Dimensional Models of ACE and NEP Inhibitors and Their Use in the Design of Potent Dual ACE/NEP Inhibitors. *J. Am. Chem. Soc.* **1996**, *118*, 8231−8249.
(39) Vincent, M.; Schiavi, P. Modelling, Synthesis and Pharmacological Study of Perindopril (S 9490), an Angiotensin I Converting Enzyme Inhibitor. In *Molecular Recognition Mechanism*; Delaage, M., Ed.; VCH: New York, 1991; pp 81−112.
(40) 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.
(41) *Spartan Conformational Analysis* (Osawa: MERCK); SGI Release 5.0.0, 1997.

CI000453N