Journal of Computer-Aided Molecular Design, **13:** 419–434, 1999.
KLUWER/ESCOM

An electron-conformational method of identification of pharmacophore and anti-pharmacophore shielding: Application to rice blast activity

Isaac B. Bersuker[a,b,*], Süleyman Bahceci[a], James E. Boggs[a] & Robert S. Pearlman[b]
[a]*Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry; and* [b]*College of Pharmacy, The University of Texas at Austin, Austin, TX 78712, U.S.A.*

Received 19 June 1998; Accepted 21 October 1998

Key words: drug design, molecular modeling, QSAR

Summary

In extension and improvement of previous results, a novel method is worked out for pharmacophore identification and activity prediction in structure-activity relationships. In this method, as in our previous works, each molecular system (conformation) of the training set is described by a matrix with both electron structural parameters (atomic charges, bond orders, etc.) and interatomic distances as matrix elements. This description includes a rather full geometry of charge and/or reactivity distribution thus providing a much better representation of the molecular properties in their interaction with the target. By multiple comparison of these matrices for the active and inactive compounds of the training set, a relatively small number of matrix elements are revealed that are common for all the active compounds and are not present in the same combination in the inactive ones. In this way a set of electronic and geometry parameters is obtained that characterize the pharmacophore (Pha). A major improvement of this scheme is reached by introducing the anti-pharmacophore shielding (APS) and a proper treatment of the conformational problem. The APS is defined as molecular groups and competing charges outside the basic skeleton (the Pha plus the inert neighbor atoms that do not affect the activity) that hinder the proper docking of the Pha with the bioreceptor thus diminishing (partially or completely) the activity. A simple empirical formula is derived to estimate the relative contribution of APS numerically. Two main issues are most affected by the APS: (1) the procedure of Pha identification is essentially simplified because only a small number of molecular systems with the highest activity and simplest structures (systems without APS) should be tried for this purpose; (2) with the APS known numerically, we can make a quantitative (or semiquantitative) prediction of relative activities. The contributions of different conformations (of the same molecular system) that possess the Pha and different APS is taken into account by means of a Boltzmann distribution at given temperatures. Applied to an example, rice blast activity, this approach proved to be rather robust and efficient. In validation of the method, the screening of 39 new compounds yields approximately 100% (within experimental error) prediction probability of the activity qualitatively (yes, no), and with $r^2 = 0.66$ quantitatively.

Introduction

Identification of the pharmacophore (Pha), the group of atoms in a specific arrangement (state) which is responsible for the bioactivity of a series of molecular systems, is one of the most important problems in molecular modeling for drug design and screening, especially when the structure and properties of the bioreceptor remain unknown. There are many monographs, review articles, and original works devoted to this problem (see e.g. [1–9] and references therein) in which the reader can find a full account of this intensively and extensively developing field. With no intention to review or evaluate the methods employed in the broad number of approaches to the solution of this problem, we only note that in the majority of these works the main emphasis is placed on the use of mathematical statistics, while the proper drug-receptor

*To whom correspondence should be addressed.

interaction is roughly presented by a few parameters like donor properties or hydrogen bonding. In our paper we pay more attention to the chemical aspect of the problem, trying to give a much better description of the molecular systems to represent the biological activity under consideration.

The present work starts from a QSAR approach for biological activity, partly suggested earlier [10]. The main idea of this approach is to represent the molecular system by a set of electronic and structural parameters arranged in a matrix form with electronic atomic characteristics as diagonal elements and off-diagonal elements representing electronic bonding parameters (for chemically bonded pairs of atoms) and interatomic distances (for other pairs). The set of such matrices for a series of trial molecular systems for which the biological activity under consideration is known (while the bioreceptor is unknown), is processed in comparison with their activity or inactivity to reveal a common group of matrix elements that are present in the active compounds and not present in the same combination in the inactive ones. In this way the submatrix of activity representing the Pha is revealed. In this approach the Pha is presented by not just a group of certain atoms, but a set of numbers describing electronic and geometric characteristics which can be the same for different atoms and different for the same atoms in different environment.

Several problems, solved by the simplest version of this method [10–22], demonstrate the efficacy of the main features of this approach as a whole (see e.g., origin of musk odor [11], meat odor [12], garlic activity [13], sandalwood odor [20], inhibitory activity of glycolic acid oxidase [21], etc.). Based on the active fragment revealed by this method for garlic activity, new garlic compounds were synthesized [10,17]. In the most elaborate example of musk odor [10,11] the qualitative prediction of odorant activity obtained by this approach was shown to be more than 90%. Kansy et al. [23] question this number stating that they checked our rules for musk odor by performing the conformational analysis of the same compounds and got a much lower prediction probability. This is because *just conformational data are definitely not sufficient* to verify the results: our method reveals a set of numbers (submatrix of activity) which include both geometry and *electronic structure*, and the rules of activity also emphasize the role of additional (to the Pha) bulky groups which may deter the *steric accessibility* of the Pha. In the previous works [10–22] the latter factor was introduced as a (revealed by inspection) qualitative agent that may reduce the activity to zero. Both the charge distribution and accessibility factors are basically ignored in Reference 23. But the discussion raises the issue of a better account for the accessibility (shielding) problem which is treated in this publication.

In the present paper we report the results of a reconsideration of this approach that led us to a general extension of the idea of Pha identification in drug design. The most important extension of the Pha concept is achieved in this method by introducing the so-called *anti-pharmacophore shielding* (APS) which allows us to reach a quantitative or semi-quantitative level of prediction of activities. The novel idea of introducing the APS as an attribute to the Pha seems to be of major importance to the solution of all such problems. As compared with the earlier version [10–22], our new procedure also accounts for the *multiconformational aspect* of the problem.

The next section describes briefly the main features of the electron-conformational method. Then follows the way of introducing APS, the general computer-implemented scheme of the method as a whole, and its application to an example, rice blast activity (RBA), that also demonstrates the procedures of this method in more detail.

The electron-conformational method of pharmacophore identification

The essence of the electron-conformational (EC) method is as follows. Assume that we have a series of N molecules with known biological activity or inactivity (the training set). The experimental method used for activity measurements falls beyond the discussion in this paper, but we require that the test of activities (inactivities) was carried out by the same 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 one Pha that is responsible for the activity under consideration.

First we evaluate approximately the energies of the conformations of the training set of compounds 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. If there are many conformations with equal or close energies that can be described as a function of a con-

tinuous parameter (e.g., an angle of internal rotation), smaller intervals of the latter can be chosen to consider each such interval as describing a separate conformation. Then for each molecule (conformation) the electron-conformational 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 a molecular system taken as an example (the compound S35 in the rice blast activity considered below). The whole number of independent elements in this matrix is $n(n + 1)/2$, where n equals 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, atomic activities (see below), valence activities, polarizabilities, HOMO (LUMO) participation, and so on. 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 (total, covalent, ionic), or polarizability; (2) if i and j label non-bonded atoms, then $a_{ij} = R_{ij}$ is their interatomic distance. In this way, each matrix 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 each ECMC, both the electronic and topological (geometrical) structure of the molecule are incorporated.

This description of charge distribution and geometry is valid for any molecular system including organometallic and coordination compounds. In the simple version of this approach employed earlier [10] only the case of 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 those mentioned above can be tried in order to choose the best electronic description. For the atomic characteristics a_{ii}, a more elaborate Fukui-Klopman type [24, 25] parameter that includes both charge and orbital-controlled interactions with the target atoms is suggested (Bersuker, to be published), but not yet fully realized.

As a result, each heavily populated conformer of each molecule with s atoms is described by its ECMC, where s varies from one molecule to another. In the previous version of this method, the next step in the design was to compare all these ECMC with the activities of the corresponding molecules, or with the values of activities in a certain interval (when the activities are known quantitatively), and to reveal those matrix elements which, within a given tolerance (see below), are common for all the active compounds and are not present 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 forming the EC submatrix of activity (ECSA). If different electronic structure parameters are to be tried (valence activities, polarizabilities, HOMO (LUMO) energies and so on), the construction of the ECMC and their processing to reveal the ECSA should be repeated for each set of these parameters; then by comparison of the ECSA obtained with different electronic structure parameters, one can decide which of them are the best in separating the active compounds from the inactive ones.

An important question is whether the ECMC 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 matrices are different, and no transposition of the columns or rows (no change in the order of numeration of atoms) is allowed (ECMC matrices 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. But it is important that the way of numeration of atoms of the molecular system (e.g., ‘clockwise’ or ‘anticlockwise’) that determines the arrangement of the matrix elements in the ECMC, once chosen, should be preserved for all conformations and enantiomers.

It is assumed that the resulting ECSA represents the Pha of the activity under consideration or the value of activity within the interval tried. 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 distribution and topology. Provided the tolerances are given, the required values of the electronic and conformational parameters for the activity in question are obtained quantitatively even when the activity is known qualitatively (yes, no). When the activity is known quantitatively, this procedure can be re-

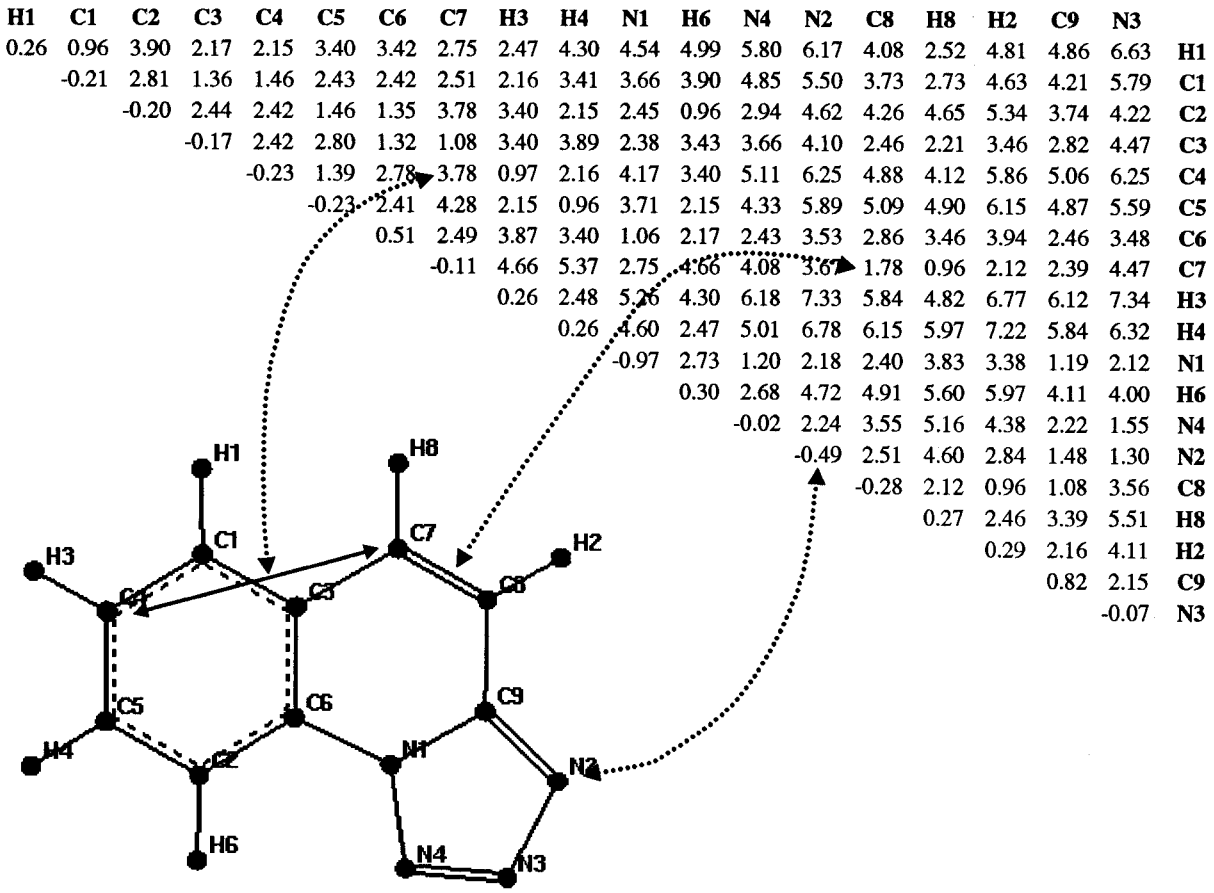


Figure 1. The general appearance of the electron-conformational matrix of congruity (ECMC) illustrated, by way of example, for compound S35 in the rice blast activity problem discussed below. The atomic charge on N2 is −0.49, the C7=C8 bond order is 1.78, and the C4···C7 distance is 3.78 Å.

peated for each interval of activity values, and then the electronic and conformational parameters emerge as a function of the activity. The description of the Pha provided by the ECSA can be used directly in screening 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 3D chemical structure [26].

By introducing the APS and consequent changes, the whole procedure is essentially improved. Based on the assumption of one Pha, the set of systems for Pha identification (*Pha training set*) may be much smaller because it can be chosen from the active and inactive compounds that have the simplest structure with the smallest number of conformations, the reduction of the activity in the remaining compounds being attributed to either APS or lack of Pha. The choice of this Pha training set, as demonstrated below for rice blast activity, seems to be straightforward and should be done by inspection. The resulting Pha should not depend on this choice, which means that by including additional compounds one can test the completeness of the Pha training set. Practically, chemical experience allows one to avoid most of this additional work. A bad choice of the initial Pha training set just requires more labor but results in the same Pha.

This statement is related also to comparisons with similar inactive compounds without APS. Such comparisons are most important in the estimation of the *tolerance limits*, i.e., the minimal interval of parameter values for which the best separation of active and inactive compounds is achieved. The ECSA obtained from the Pha training set is further compared with the ECSAs of the other active compounds. The procedure is carried out in an iterative dialog regime with the computer (see below).

The main point is that the concept of APS simplifies essentially the whole procedure of Pha identi-

fication because it allows us to do it by considering the simplest structures of the training set, and attribute the complete or partial reduction of the activity in the more complicated structures to their APS.

Basic skeleton and anti-pharmacophore shielding

It is obvious that the presence of the Pha, as outlined above, is a necessary but not sufficient condition of activity. Indeed, in the majority of cases studied so far the Pha contains several atoms (mostly three to five–six) which occupy a limited space volume (often three-four atoms in a plane). Since the molecular system is usually (but not necessarily) larger than the Pha, there may be atoms that occupy positions which may restrain (sterically hinder) the direct interaction of the Pha with the bioreceptor (sterical restrictions). Another restriction may occur from the presence of additional highly charged (or otherwise reactive) atoms which, although outside the Pha, may compete in the interaction with bioreceptor (competing charges), thus violating the proper docking of the substrate with the bioreceptor.

To determine the APS, we have to define first the *basic skeleton*. The Pha, defined above, includes the minimal group of atoms (more precisely, certain 3D arranged charges or reactivity sites) required for the activity. The presence of molecular groups, additional to this minimum set of Pha, may either preserve or diminish (destroy) the activity. In specific cases, the analysis of the group of active molecules shows that there is a 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*. The additions to the Pha group from the basic skeleton seemingly do not participate significantly in (but do not hinder either) the interaction with the bioreceptor. The difference between Pha and basic skeleton is that these additions may be different (may have different ECMC elements) for compounds with approximately the same activity. One can imagine that these additions occupy some empty places in the cavity of the receptor and hence they do not affect the ligand-receptor interaction.

Then the APS can be defined as additional (to the basic skeleton) groups of atoms that cause sterical restrictions and/or have competing charges (reactivities) which diminish the activity, partially or completely. If the bioreceptor is known, the effect of both sterical restrictions and competing charges 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, it can be approached based on statistical comparisons of the structural and electronic parameters of the compounds with their relative values of activity. We suggest the following scheme of APS evaluation which, as shown by the example of rice blast activity, allows for at least qualitative or semiquantitative estimations.

First we assume that a parameter exists which quantitatively characterizes the APS value incorporating both the sterical restrictions and competing charges. Denote it by S. Since the interaction between the substrate and bioreceptor is similar to any other intermolecular interaction (at least from the point of view of sterical or other energy barrier factors), we can assume that the activity A depends on S exponentially:

$$A \sim e^{-s} \tag{1}$$

To evaluate the APS parameter S, consider a model Pha which is planar and has one of the atoms highly charged (this is a common situation which we have found in our Pha identification so far [10–22]; see also the Pha of rice blast activity below in this paper). Then, there is a group of atoms around the Pha that do not affect the activity (do not shield the Pha) which together with the Pha form the basic-skeleton of activity as defined above. Note that the skeleton may be larger than the Pha, but the necessary condition of activity is just the presence of Pha.

Consider two groups of atoms outside the skeleton, 1 and 2, on both sides of the Pha plane (Figure 2), which are bonded to the skeleton atoms L_1 and L_2 at the (farthest atoms) distances d_1 and d_2 from the latter, and assume that one of these groups or another (third group) has a highly charged atom with the charge q. Our experience with rice blast activity (see below) shows that single unilateral groups, i.e., groups that occupy just one side of the Pha plane in the absence of steric groups in its other side, are much less efficient in shielding the activity than bilateral groups (situated in both sides of this plane). This observation can be understood if one takes into account that one-sided steric groups do not restrict the Pha docking to the bioreceptor with its other side (still such one-sided sterical groups will affect the kinetics of drug-receptor interaction). If the Pha is chiral and chirality is significant in the ligand-receptor interaction, the activity may depend on the side that is occupied by the sterical

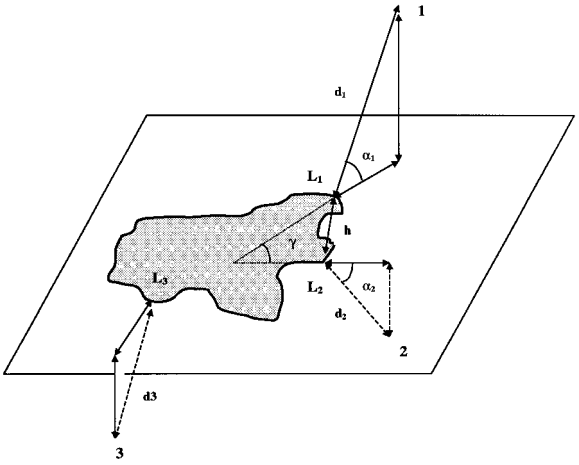


Figure 2. The scheme for estimation of the anti-pharmacophore shielding. The molecular group outlined on the plane is the basic skeleton, while the sterical restricting groups that may include also competing charges are indicated as 1, 2 and 3.

group. In this case the unilateral (one-sided) shielding group does not affect the docking with the receptor only if it does not violate their proper chiral interaction (which may be different for two enantiomers). Therefore the APS problem with chirality may look more complicated. Below in this paper we consider the cases when chirality is not essential in the drug-receptor interaction leaving the chirality APS problem for subsequent publications.

Based on the empirical data and experience accumulated from the RBA problem, considered below, as well as from other problems treated earlier [10–22], we assume that if the additional (steric) groups are in the plane of the Pha forming its extension beyond the skeleton or under a small angle to this plane, their shielding effect is significant. But if two groups contribute to such skeleton-extension shielding, their effect is additive only if their projections on the Pha plan do not coincide and the distance h between the points of connection to the skeleton L_1 and L_2 is not small.

To take into account these features of the steric effects in APS (except chirality), we suggest for the steric part of S the following formula:

$$S_{\text{steric}} = \kappa d_1 \cos\alpha_1 + \kappa' d_2 \cos\alpha_2 \sin\left(\frac{h\pi}{6} + \gamma\right) + \kappa_b [d_1 \sin\alpha_1 d_2 \sin(-\alpha_2)] \quad (2)$$

where the angles $0^\circ < \alpha_1 < 90^\circ$ and $0^\circ < \alpha_2 < 90^\circ$ are the 'outside' dihedral angles between the farthest atoms of the steric groups under consideration and the outside extension of the Pha plane (Figure 2), $0<\gamma<\pi/2$ is the angle between the projections of the two distances d_1 and d_2 on the Pha plane, κ, κ', and κ_b are adjustable coefficients (see below), and it is assumed that $\cos\alpha_1 = 0$ and $\cos\alpha_2 = 0$ for $\alpha > 90^\circ$, $\sin\alpha = 0$ for $\alpha < 0^\circ$, $\sin\alpha = 1$ for $\alpha>\pi/2$, and $d_1 > d_2$ (h is taken in Å). The second term in this equation is nonzero only if $h \neq 0$ and/or $\gamma \neq 0$, and it becomes equal to $\kappa' d_2 \cos\alpha_2$ when $(h\pi/6)+\gamma \geq \pi/2$, i.e. when $h \geq 3$ Å at $\gamma = 0$.

With this formula, a one-sided steric group positioned along the outside extension of the Pha plane contributes to the APS when the angle α is not very large, while its perpendicular projection (large α) is important only when there is a second sterical hindrance from the other side of the Pha plane and the two groups form a two-sided APS; for one-sided steric groups $d_2 = 0$ and the last term is zero. Here it is assumed that the ligand-receptor interaction is achiral. If there are more than two APS groups, they may be handled in the same way. In particular, if a third group at a distance d_3 is present, additional terms of $\kappa'' d_3 \cos\alpha_3$ and two terms $\kappa_b d_1 \sin\alpha_1 d_3 \sin(-\alpha_3)$ and $\kappa_b d_2 \sin\alpha_2 d_3 \sin(-\alpha_3)$, as well as equivalents to the second term, should be added in Equation 2.

Note that the constants κ, κ', and κ'' etc. (κ_i in the general formula Equation 4 below) are different when the distances d_1, d_2, and d_3 are related to different places of the skeleton L_1, L_2, and L_3, respectively. If there are several such places, each of them should have its own weight constant κ_i to account for the different effect of APS at different places L_i.

Passing to competing charges, we note that highly charged (or otherwise reactive) atoms compete with the atoms of the Pha in the interaction with the bioreceptor thus reducing the probability of a proper Pha–bioreceptor docking. This effect depends mostly on the magnitude of the competing charge as compared with the highest charged (reactive) atom of the Pha, although any charge differences, as compared with those of the corresponding (nearest) atoms of the skeleton may affect the activity. Dependent on where the charge is placed with respect to skeleton, its competing effect may be different. Therefore we suggest a choice of several basic locations (three in the example of the rice blast activity) for each of which an independent weight constant κ_{ci} is introduced. As a result we have:

$$S_{charge} = \sum_i \kappa_{ci} |q_i| \quad (3)$$

We can assume now that $S = S_{steric} + S_{charge}$ and

$$S = \sum_{i,j}\left[\kappa_i d_i \cos\alpha_i + \kappa_j d_j \cos\alpha_j \sin\left(\frac{h_{ij}\pi}{6} + \gamma_{ij}\right)\right] + \kappa_b\left[\sum_{i,j} d_i \sin\alpha_i d_j \sin(-\alpha_j)\right] + \sum_i \kappa_{ci}|q_i| \quad (4)$$

The constants κ_i, κ_b and κ_{ci} are very important to adjust the different terms in Equation 4 to the same units (S in Equation 1 must be dimensionless) and to weight their significance in the process of APS.

With the presence of the Pha as a necessary condition of activity and the APS parameter taken in the form of Equation 4, the formula of activity can be given as follows:

$$A = \delta[Pha] A_{\max}\, e^{-S} \quad (5)$$

where δ[Pha] is a kind of Dirac δ-function:

$$\delta[Pha] = \begin{cases} 1, \text{ when Pha is present} \\ 0, \text{ when Pha is absent} \end{cases} \quad (6)$$

and A_{max} is the maximum activity when δ[Pha] = 1 and $S = 0$.

It remains to take into account the conformational problem. Assume that each molecular system n in the training set has m_n conformations, and only some of them have the Pha and hence may be active. According to the Bolztmann distribution, the relative number of molecules in the conformation i is $n_i/N = \exp(-E_i/kT)/N$ where $N = \sum_i n_i$ is the molecular partition function and E_i is the energy of the conformation above the ground state one. With this in mind, the activity of the compound n, including all its conformations, is:

$$A_n = A_{\max}\frac{\sum_{i=1}^{m_n} \delta_{ni}[Pha] e^{-S_{ni}} e^{-E_{ni}/kT}}{\sum_{i=1}^{m_n} e^{-E_{ni}/kT}} \quad (7)$$

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 relative numbers of molecules in the active conformations. These numbers decrease rapidly with the energy of the conformation E_i. Therefore, in most cases we can neglect the contributions of the conformations with higher energies. At room temperatures $kT \sim 0.6$ kcal/mol; hence conformations with energies higher than 1.0–1.5 kcal/mol above the ground state one contribute less than 10%–15% each, and often can be neglected (in view of the approximations made in other calculations and inaccuracies in experimental data), provided they are not extremely large in number. The latter case is expected to be rare because the Pha usually forms a rigid frame of several atomic points that does not involve multi-conformation flexible groups.

If several or all the conformations are active, they still may have different shielding parameters S_i. Thus the conformational factor is significant also in providing different Boltzmann weighted shielding contributions.

The parameters κ_i, κ_b and κ_{ci} can be easily evaluated from a least-squares procedure that gives the lowest value of the sum $\Sigma[A_i - A^i_{exp}]^2$, i.e. the best possible (average) agreement of the activities calculated from Equation 7 with the experimental data A^i_{exp} available for the training set of compounds.

Computer implementation

The overall architecture of the EC method with APS (EC-Pha-APS method) is presented in Figure 3.

Pharmacophore identification module

- Choose the training set of N compounds to include most active, moderately active, and inactive systems, the activity being determined by the same experimental method.
- Construct the 2D molecular sketches and minimize them by molecular mechanics methods to convert to 3D structures. Using existing programs, generate a representative set of conformations for flexible compounds and separate the conformations that are populated significantly at room temperature to be tried.
- Calculate the electronic structure of each molecule (conformation) by semiempirical methods (simple ab initio calculation may also be needed in cases of small limits of tolerance in electronic structure parameters) and generate their matrix representation, the ECMC.
- Separate a group of most active compounds with the smallest number of atoms plus a number of inactive compounds that are most similar to the active ones (the Pha training set), and compare their ECMC to identify the submatrix of activity (ECSA) and the corresponding active fragment.

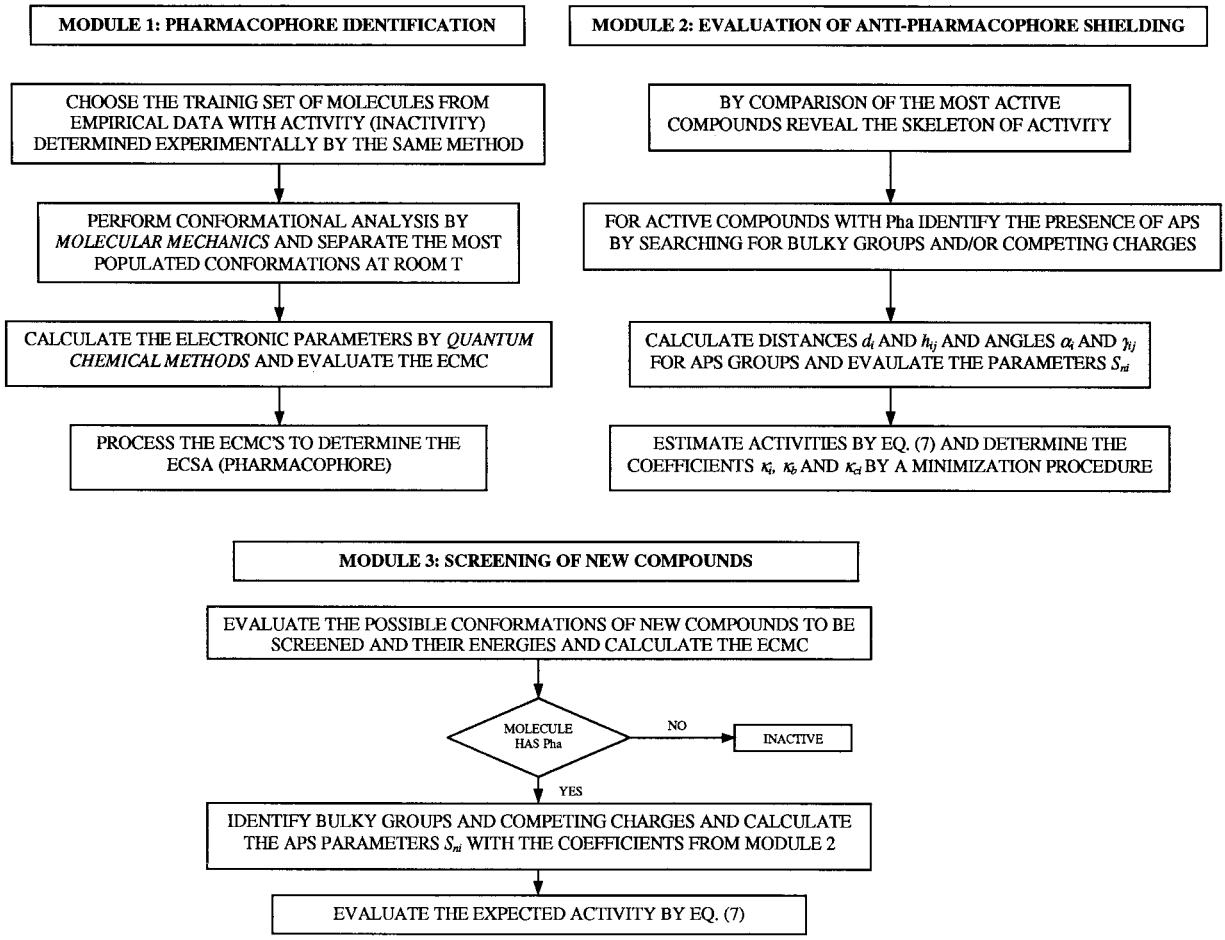


Figure 3. Algorithms and computer implementation of the EC method.

- Check the completeness of the Pha training set by adding compounds; run all the jobs in an iterative regime with the computer and, by means of trial choices that include weakly active and inactive compounds without APS, determine the parameter's tolerance, i.e. the minimal intervals of parameter values for which the best separation of the active compounds from the inactive ones is achieved.

Evaluation of anti-pharmacophore shielding (APS) module

- By comparison (superposition) of all the most active compounds, reveal the basic skeleton of activity.
- By inspection, identify the presence of APS by searching for bulky groups and/or competing charges in the rest of the active and inactive compounds that have the Pha (beyond the Pha training set); calculate the distances d_i and h_{ij} and angles α_i and γ_{ij}.
- Estimate the expected activities A of all compounds using Equation 7: calculate the coefficients κ_i, κ_b, and κ_{ci} using a least-squares minimization procedure that gives the lowest (average) deviations of A from the experimental values A_{exp}, and estimate the APS parameter value S after Equation 4.

Screening of new compounds module

- Evaluate the possible conformations of new compounds to be screened and their energies and calculate the ECMC.
- Predict the presence of biological activity in novel compounds by checking whether they contain the Pha.

- Calculate the APS parameters S with the coefficients from the APS Module.
- Evaluate the expected activity using Equation 7.

In these stages, the most labor-intensive and computer time-consuming ones are the conformational analysis and the processing of the ECMC to determine the ECSA. The general computations can be performed using any of the existing packages for quantum chemical calculations. In the problem of rice blast activity, discussed below, we used SPARTAN [27].

The Pha identification module, on the other hand, uses several codes written by us in the MATLAB environment by Mathworks, Inc. [28]. The traditional methods used for comparison of matrices are based on the algorithms of molecular graph intersection. In an alternative algorithm, partly worked out by us for the EC method, the matrix elements of the 'starting' ECMC (the one for the most active compound) are compared with the ones of all the other ECMC, and in this way, working in an interactive regime with the computer, the submatrix that divides the active and inactive compounds is revealed. This algorithm is much more effective than the traditional one based on the search for intersections of appropriate molecular graphs. We plan to work out a procedure that will identify the APS from the same ECMC that determines the Pha.

Application: Rice blast activity

The example of rice blast activity (RBA) has been chosen among the series of biologically active molecular systems for which there is a significant number of quantitatively tested compounds which can serve for validation of the method and, for simplicity of an example, are relatively not very complicated, but not very simple either. There are many such series available in the literature, and from this point of view the selection of RBA as an example is random.[1]

It comes out that RBA compounds have few conformations, but for our method the conformation problem is not limiting, in principle; as mentioned above, the Pha can be evaluated by trying a smaller set of active (inactive) compounds which can be chosen to have few conformations. The main synthetic work on the RBA data set used in our investigation was carried out more than a couple of decades ago; it led directly to a significant body of active compounds, and ultimately, to a commercial rice blast control agent, tricyclazole [29].

Dreikorn et al. [30] conducted a 'classical' SAR study of the tetrazolo- and triazoloquinolines (TZQs). They found out that the RBA for a portion of the molecules correlates (roughly) with the dipole moment of their component groups in the plane of the molecules. This correlation between the activity and the dipole moment, although interesting, is far from solving the problem. In view of the results of this paper, it is obvious that just the dipole moment can not characterize the molecules sufficiently well to allow for an effective QSAR analysis and to predict the features of RBA. The results of investigation of this series of RBA active and inactive compounds are discussed below as an example of application of the new algorithm; it gives also the details of our method.

The training set of 85 compounds plus 11 conformations (altogether 96 molecular systems) includes the tetrazolo and triazolo quinolines that were considered in the previous QSAR attempts [30]. The skeletons of the training set of molecules are shown in Figure 4, while all the structures are listed in Table 1. The RBA is expressed in relative activity (RELAC) values, as compared to the activity of tetrazolo(1,5)-quinoline (TZQ) equal to 1; zero indicates that no activity against the rice blast disease was found. Note that the experimental data (as in many other cases of experimental determination of biological activity) are rather approximate: the authors [30] grouped them in accordance with their RELAC values between 0 and 1.33 with intervals of $\sim$0.20. The latter (together with additional private information from the authors) can serve as an indication of the estimated error level in the experimental measurements. In spite of such inaccuracies, these data may serve as a basis for the qualitative and semiquantitative analysis given below.

In accordance with the computational scheme, described above, the conformational and electronic parameters of the selected training set have been calculated and the corresponding electron-conformational matrices of congruity (ECMC) were constructed. By way of illustration, the ECMC for compound S35 in Table 2 was given above in Figure 1. In this matrix, as deter-

[1] A reviewer suggested that it might be relevant to consider one of the examples which were treated by the old version of our method without APS and other improvements in order to demonstrate the difference. Unfortunately for the best examples considered earlier [10–22], e.g. musk odor activity [11], there are no quantitative data which could serve for validation of the novel quantitative or semiquantitative prediction. In fact such quantitative data for odorant activity determined as the threshold quantities that manifest odor do exist for limited numbers of compounds, but they are confidential and not published. Without quantitative data there is no way to get the coefficients κ_i in Equations 5 and 7.

Table 1. The training set of compounds under investigation

No.	No. of conformations[b]	Skeleton	R1	R2	R3	R4	R5	R6	R7	R8	RELAC[c]
N1[a]	1	A	–	N	H,H	H,H	H	H	H	H	1.33
N2[a]	1	A	–	N	H,H	H,H	H	H	H	Cl	1.33
N3[a]	1	A	–	N	H,H	H,H	H	H	H	CH_3	1.33
N4[a]	1	A	–	N	H	H	H	H	H	CH_3	1.33
N5	1	E	N	NH	NH	–	–	–	–	–	1.00
N6[a]	1	A	–	CH	H,H	H,H	H	H	H	H	0.91
N7[a]	1	A	–	N	H	CH_3	H	H	H	H	0.80
N8[a]	1	A	CH_3	CH	H,H	H,H	H	H	H	H	0.75
N9[a]	1	A	–	CH	H,H	H,H	H	H	H	CH_3	0.75
N10[a]	1	H	CH	N	N	–	–	–	–	–	0.75
N11[a]	1	F	CH	S	–	–	–	–	–	–	0.75
N12[a]	1	G	N	CH_2	NH	–	–	–	–	–	0.63
N13[a]	1	G	N	CH_2	S	–	–	–	–	–	0.63
N14[a]	1	G	N	S	CH_2	–	–	–	–	–	0.63
N15[a]	1	C	N	CH	N	CH	CH	–	–	–	0.63
N16	1	D	N	$(CH_2)_4$	CH	CH	–	–	–	–	0.50
N17	1	F	N	$(CH_2)_3$	–	–	–	–	–	–	0.50
N19	1	A	–	N	CH_3	H	H	H	H	H	0.42
N20	1	A	–	N	H,H	H,H	H	Cl	H	H	0.42
N21	1	A	–	N	H,H	H,H	CH_3	H	H	H	0.42
N22	1	A	CH_2S	C	H	H	H	H	H	H	0.42
N24	1	B	CH	N	N	–	–	–	–	–	0.42
N25	1	A	–	N	H	H	H	H	CH_3	H	0.41
N26	2	A	CH_2CH_3	C	H	H	H	H	H	H	0.41
N28	1	F	N	$(CH_2)_4$	–	–	–	–	–	–	0.38
N29	1	F	N	S	–	–	–	–	–	–	0.38
N30	1	A	–	N	H,Cl	H,Cl	H	H	H	H	0.33
N31	1	A	–	N	H,Cl	H,Cl	H	H	H	CH_2CH_3	0.33
N32	1	A	–	N	H,H	H,H	H	OCH_3	H	H	0.33
N33	1	A	–	N	H,H	H,H	H	CH_3	H	H	0.33
N34	1	A	OH	C	H	H	H	H	H	H	0.33
N35	1	A	COOH	C	H	H	H	H	H	H	0.33
N36	2	A	$CH_2CH_2CH_2CH_3$	C	H	H	H	H	H	H	0.33
N38	1	L	CH	–	–	–	–	–	–	–	0.19
N39	1	G	CH	S	CH_2	–	–	–	–	–	0.19
N40	1	A	–	N	H,H	H,CH_3	H	H	H	H	0.17
N41	1	A	COOH	C	H	H	H	H	H	H	0.17
N42	3	A	$COOCH_2CH_3$	C	H	H	H	H	H	H	0.17
N43	1	A	NH_2	C	H	H	H	H	H	H	0.17
N44	1	A	i-Propyl	C	H	H	H	H	H	H	0.17
N45	4	A	n-Propyl	C	H	H	H	H	H	H	0.17
N46	1	A	$CH_2OCH_2CH_3$	C	H	H	H	H	H	H	0.17
N47	2	A	$CH_2CH_2OCH_2CH_3$	C	H	H	H	H	H	H	0.17
N48	1	G	CH	NH	–	–	–	–	–	–	0.16
N49	1	A	–	N	H,Ch_3	H,H	H	H	H	H	0.00
N51	1	A	$CONH_2$	C	H	H	H	H	H	H	0.00
N52	1	A	$NHCOCH_3$	C	H	H	H	H	H	H	0.00

Table 1 (continued)

No.	No. of conformations[b]	Skeleton	R1	R2	R3	R4	R5	R6	R7	R8	RELAC[c]
N53	1	A	SH	C	H	H	H	H	H	H	0.00
N54	2	A	CH_2OH	C	H	H	H	H	H	H	0.00
N55	1	A	$CHCH_2$	C	H	H	H	H	H	H	0.00
N56	1	A	Phenyl	C	H	H	H	H	H	H	0.00
N57	1	A	CF_3	C	H,H	H,H	H	H	H	H	0.00
N58	1	A	1-Phenyl	C	H,H	H,H	H	H	H	H	0.00
N59	1	A	1-Pyridyl	C	H,H	H,H	H	H	H	H	0.00
N60	1	A	1(2-Thiophene)	C	H,H	H,H	H	H	H	H	0.00
N61	1	A	1(2-Furanyl)	C	H,H	H,H	H	H	H	H	0.00
N62	1	B	N	CH	CH	–	–	–	–	–	0.00
N63	1	B	N	N	CH	–	–	–	–	–	0.00
N64	1	B	N	CH	N	–	–	–	–	–	0.00
N66	1	C	N	N	N	CH	CH	–	–	–	0.00
N67	1	C	N	CH	N	N	CH	–	–	–	0.00
N68	1	C	N	CH	CH	CH	N	–	–	–	0.00
N69	1	D	N	Naphthyl	N	N	–	–	–	–	0.00
N70	1	G	N	CO	NH	–	–	–	–	–	0.00
N72	1	N	–	–	–	–	–	–	–	–	0.00
N73	1	H	N	N	CH	–	–	–	–	–	0.00
N74	1	H	CH	CH	N	–	–	–	–	–	0.00
N75[a]	1	I	N	CH	CH	CH	CH	–	–	–	0.00
N77[a]	1	I	N	CH	CH	N	CH	–	–	–	0.00
N78[a]	1	I	N	CH	N	CH	CH	–	–	–	0.00
N79[a]	1	I	N	CH	CH	CH	N	–	–	–	0.00
N80[a]	1	J	N	–	–	–	–	–	–	–	0.00
N81[a]	1	K	N	$(CH_2)_4$	–	–	–	–	–	–	0.00
N82[a]	1	K	N	$(CH_2)_3$	–	–	–	–	–	–	0.00
N83[a]	1	K	N	$(CH_2)_5$	–	–	–	–	–	–	0.00
N84[a]	1	I	CH	CH	CH	CH	CH	–	–	–	0.00
N85[a]	1	I	CH	CH	N	CH	CH	–	–	–	0.00
N86[a]	1	I	CH	N	CH	CH	CH	–	–	–	0.00
N87[a]	1	I	CH	CH	CH	N	CH	–	–	–	0.00
N88[a]	1	K	CH	$(CH_2)_3$	–	–	–	–	–	–	0.00
N89[a]	1	K	CH	$(CH_2)_4$	–	–	–	–	–	–	0.00
N90[a]	1	K	CH	$(CH_2)_5$	–	–	–	–	–	–	0.00
N91[a]	1	I	CH	CH	CH	CH	N	–	–	–	0.00
N92	1	D	CH	$(CH)_4$		–	$(CH_2)_4$	–	–	–	0.00
N93	1	F	CH	$(CH_2)_3$	–	–	–	–	–	–	0.00

[a] The molecules are used for the Pha training set.
[b] The conformations which are populated at room temperature.
[c] The rice blast activity is expressed in relative activity (RELAC) values as compared to the activity of the lead tetrazolo(1,5)-quinoline(TZQ) equal to 1; zero indicates that no activity against the rice blast disease was found.

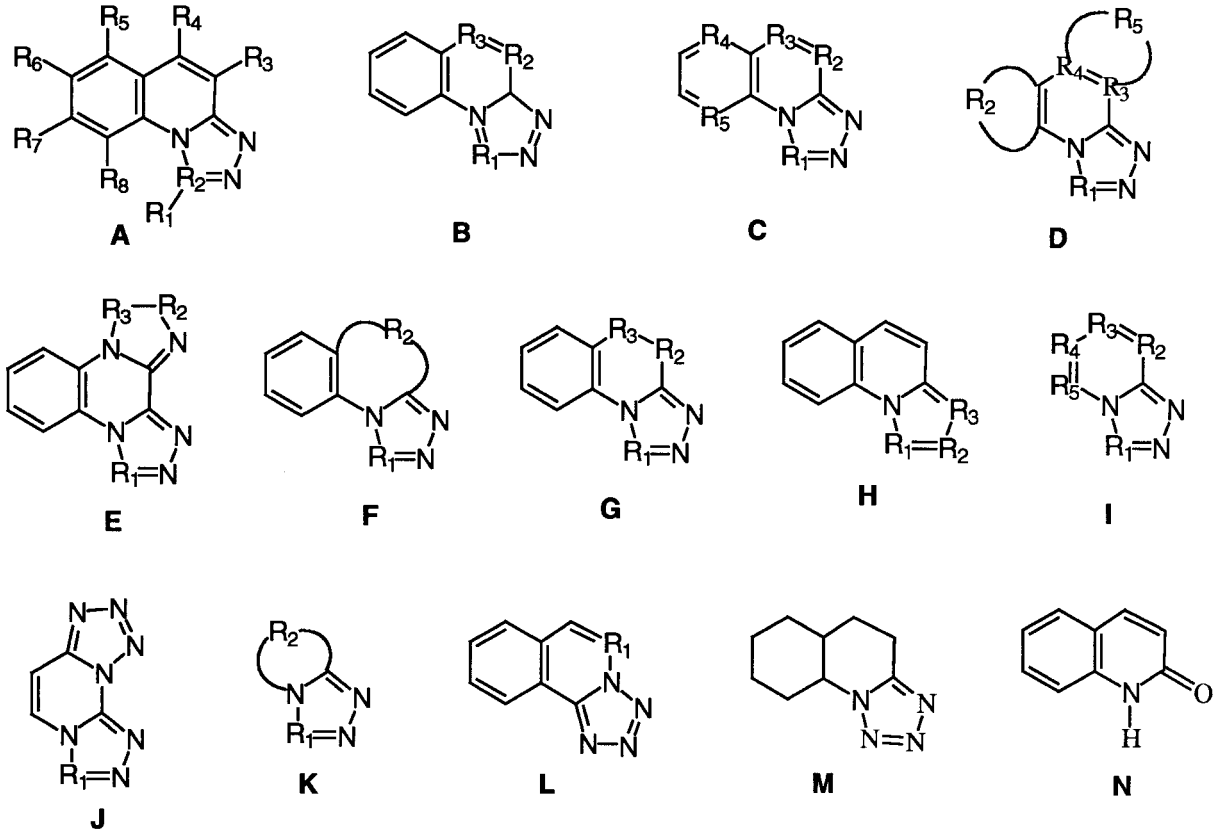


Figure 4. Skeletons of the set of compounds considered in the RBA problem (see Tables 1 and 2).

mined by the EC method, the diagonal elements are the effective charges on atoms, Q_{ii} (in units of electronic charges), while the off-diagonal elements are bond orders W_{ij} for chemically bonded atoms, and optimized distances R_{ij} in Å for chemically nonbonded pairs of atoms.

The next step is to process the massive set of numbers of the ECMCs of a selected part of the training set in a dialogue regime with the computer by means of our programs in order to reveal the ECSA. The main operation here is to take one of the most active molecules as a reference system, and to compare the ECMC of all the active and the inactive compounds of the Pha training set with that of the reference one. The comparison reveals a smaller number of matrix elements that are common for all the active compounds in the Pha training set, the submatrix of activity, which describes the active electron-conformational molecular fragment, the Pha. Figure 5 shows the ECSA and the corresponding atomic arrangement of the four atoms in the Pha for RBA. Our algorithm provides for a double check to make sure that the active fragment is absent in the inactive compounds by comparing ECMC of the reference molecule with those of the inactive ones.

As indicated above, in order to identify the Pha there is no need to compare all the compounds in the training set. Because the Pha is assumed to be present in all the active compounds, it is sufficient to consider only those with the simplest structures, both active and inactive, the latter being chosen (by inspection) with no anti-pharmacophore shielding. This reduces significantly the computational work. In our case the Pha training set contains 14 active and 16 inactive compounds (see Table 1).

The pattern that describes the Pha contains a set of interatomic distances correlated with atomic charges on certain locations given in the ECSA. In particular, when the planarity of the molecule is distorted by a substituent, the activity decreases and ultimately, beyond a threshold, disappears. This threshold, which is in fact the *tolerance limit* for the molecular parameters, is obtained as follows. The ECSA determined from comparison of most active and similar inactive

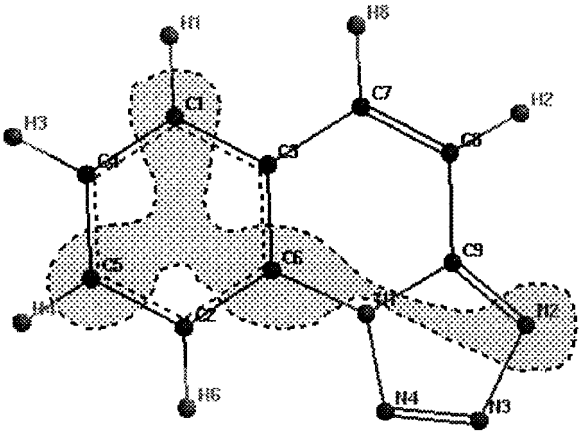
Table 2. Screening of a new set of compounds on RBA

No.[a]	Skeleton	R1	R2	R3	R4	R5	R6	R7	R8	A_{exp}	A_{calc}
S1	A	–	N	H	NH_2	H	H	H	H	0.00	0.18
S2	C	N	N	CH	CH	CH	–	–	–	0.00	0.00
S3	A	–	N	H	H	NO_2	H	H	H	0.00	0.00
S4	M	–	–	–	–	–	–	–	–	0.00	0.00
S5	B	CH	N	CH	–	–	–	–	–	0.00	0.23
S6	E	CH	N	N	–	–	–	–	–	0.00	0.05
S7	I	N	N	CH	CH	CH	–	–	–	0.00	0.00
S8	E	CH	N	CH	–	–	–	–	–	0.00	0.04
S9	A	–	N	H	H	H	NO_2	H	H	0.00	0.00
S10	A	–	N	H	H	H	H	H	NO_2	0.00	0.00
S11	A	–	N	H	H	NH_2	H	H	H	0.00	0.00
S12	A	Cyclopropyl	C	H	H	H	H	H	H	0.17	0.33
S13	A	OH	C	H,H	H,H	H	H	H	H	0.17	0.18
S14	A	Ph	C	H	H	H	H	H	H	0.17	0.16
S15	A	–	N	H	H	H	Cl	H	H	0.17	0.39
S16	A	–	N	H	H	H	CH_3	H	H	0.17	0.45
S17	A	H	C	H	Cl	H	H	H	H	0.33	0.42
S18	G	CH	CH_2	S	–	–	–	–	–	0.33	0.39
S19	D	N	$(CH)_4$	–	–	$(CH_2)_4$	–	–	–	0.38	0.59
S20	A	–	N	H	H	H	H	H	NH_2	0.42	0.45
S21	A	–	N	Cl	H	H	H	H	H	0.42	0.20
S22	A	–	N	CH_3	H	H	H	H	H	0.42	0.25
S23	B	CH	CH	N	–	–	–	–	–	0.42	0.23
S24	A	H	C	H	H	H	OCH_3	H	H	0.42	0.06
S25	A	CF_3	C	H	H	H	H	H	H	0.42	0.10
S26	A	S-Et	C	H	H	H	H	H	H	0.42	0.17
S27	H	N	CH	CH	–	–	–	–	–	0.50	0.56
S28	A	H	C	H	H	H	H	H	H	0.67	0.73
S29	A	–	N	H	CH_3	H	H	H	H	0.75	0.82
S30	A	H	C	H	CH_3	H	H	H	H	0.79	0.45
S31	A	H	C	CH_3	Cl	H	H	H	H	0.92	0.42
S32	A	H	C	H,H	H,H	H	H	H	Cl	0.92	0.71
S33	A	–	N	H	Cl	H	H	H	H	0.92	0.76
S34	A	CH_3	C	H	H	H	H	H	H	1.00	0.71
S35	A	–	N	H	H	H	H	H	H	1.00	1.33
S36	A	–	N	H	H	H	H	H	Cl	1.00	1.33
S37	A	H	C	H	H	H	H	H	F	1.17	0.73
S38	A	Cl	C	H	H	H	H	H	H	1.17	0.75
S39	A	H	C	H	H	H	H	H	CH_3	1.33	0.73

[a] All compounds have one conformation populated at room temperature.

compounds as described above is further compared with the remaining weakly active compounds that are assumed to have the Pha. By gradually changing the limits of tolerance for charges, bond orders and interatomic distances, we finally obtain the tolerance limits indicated in the ECSA in Figure 5, which give the best separation of the active compounds from the similar inactive ones.

With the known Pha, the basic skeleton is determined by superimposing the most active compounds N1, N2, N3, and N4 that have the RELAC 1.33 (Figure 6). Then the APS is defined as additional (to the basic skeleton) groups that may produce anti-



C1	C5	C6	N2	
-0.17±0.10	2.42 ±0.03	2.44 ±0.03	5.48 ±0.34	C1
	-0.16 ±0.10	2.42 ±0.05	5.91 ±0.10	C5
		0.46 ±0.08	3.53 ±0.08	C6
			-0.46 ±0.05	N2

Figure 5. The active fragment (pharmacophore) responsible for rice blast activity contains four directly nonbonded atoms, more precisely, four charges situated at distances shown in the submatrix of activity (ECSA).

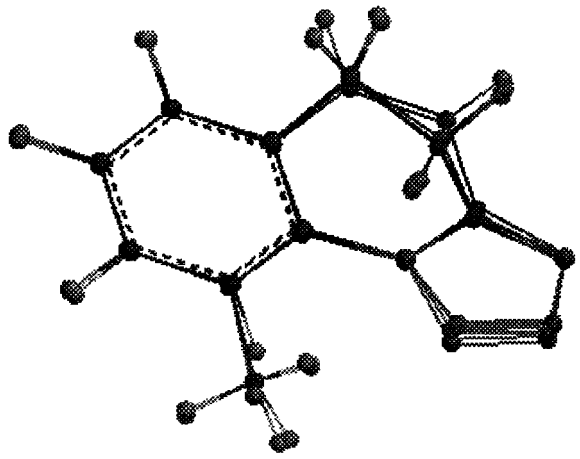


Figure 6. The skeleton of activity – the maximum extension of the pharmacophore that does not influence the RBA.

pharmacophore shielding as described above (Figure 2). Identification of APS is done by searching bulky groups and/or competing charges in the rest of the active and inactive compounds that have Pha. In order to estimate the *S* value, the distances d_i and h_{ij} and angles α_i and γ_{ij} were calculated; the pharmacophore plane for each molecule was determined by using SPARTAN's user-friendly graphic interface. The expected activities *A* are estimated from Equation 7 in which the conformations with energies more than 1 kcal/mol above the ground state one were neglected as contributing less than 15% (which is negligible in view of the inaccuracies in the experimental data and approximations made in the calculations). The coefficients κ_i, κ_b and κ_{ci} are obtained by using a least-squares minimization procedure that gives the lowest deviations of *A* from the experimental values A_{exp}. The results are shown in Figure 7 ($r^2 = 0.88$).

The parameters of the Pha (Figure 5) together with the coefficients of APS in Figure 7 can be used for nonexperimental screening of new compounds. Only 2 out of 42 active compounds (N16 and N38) of the training set with activities 0.50 and 0.19, respectively, have no Pha, while compounds N8 and N53 have calculated activities, respectively, 0.42 and 0.39, not shown experimentally. For a better fit with experimental data another two compounds, N5 and N40, are excluded from the training set: because of their unusual nonfit to the scheme, their experimental data look suspicious.

Nonexperimental screening of new rice blast active compounds – validation of the EC method

For nonexperimental screening of new compounds we determine their low-energy conformations, calculate their ECMC, and check for the presence of the ECSA (Pha). The data for an additional set of 39 compounds listed in Table 2 were kindly provided to us by Dreikorn and Durst (personal communication). For the compounds that have the Pha, the APS parameters are calculated and, using Equation 7, the activities *A* are predicted. Figure 8 shows the predicted RELAC values *A* in comparison with the experimental data. Qualitatively (yes, no), the prediction probability is ∼100% (within the experimental errors). On the other hand, quantitatively the prediction is only $r^2 = 0.66$ which is quite understandable in view of the approximations made and other factors of activity not taken into account in this paper (including lipophilicity).

Concluding remarks

The electron-conformational (EC) approach to the problem of identification of the pharmacophore (Pha)

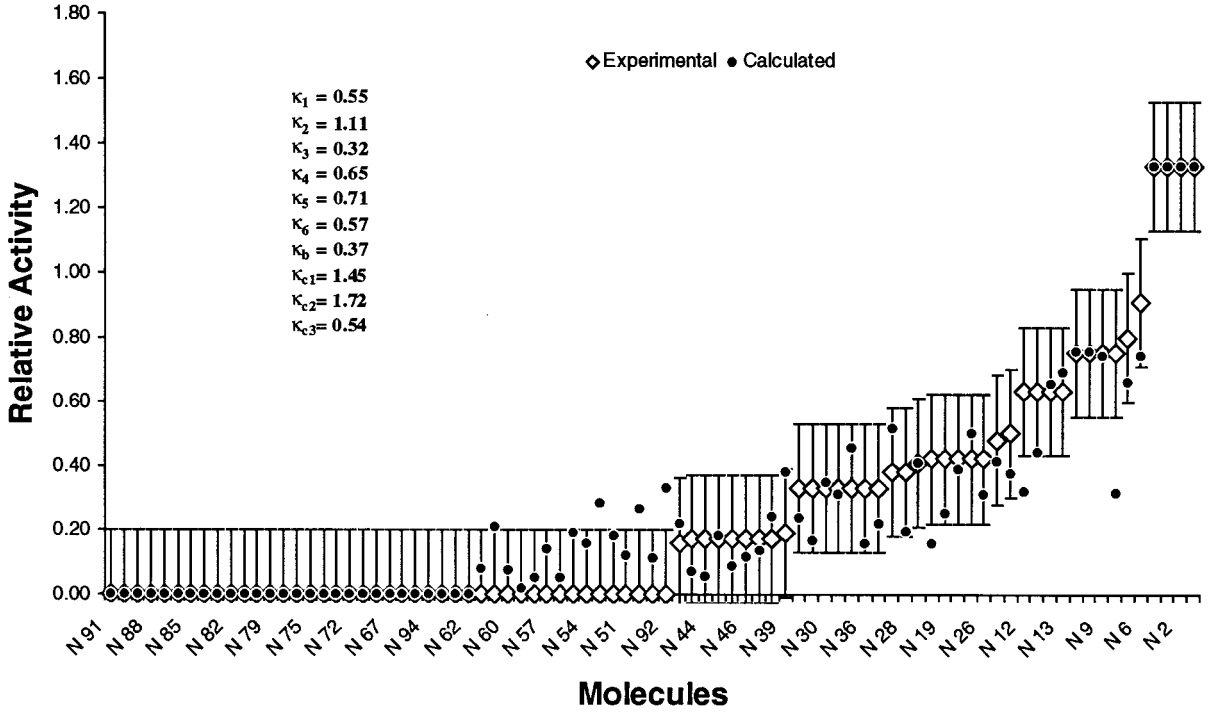


Figure 7. Calculated and experimental RELAC values in the training set. The vertical lines indicate the estimated error in the experimental values.

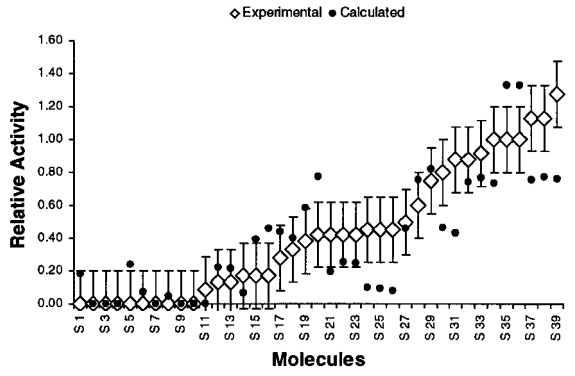


Figure 8. Predicted (by Equation 7) RELAC values of RBA of a series of 39 compounds in comparison with the experimental values.

and anti-pharmacophore shielding (APS) described in this paper may be regarded as an extended QSAR method in which the modeling of the molecular systems of the training set includes their rather full description by means of electronic and geometric parameters arranged in a single matrix, and the idea of APS that diminishes the activity is incorporated. As compared with the previous version [10–22] this paper addresses additionally two major issues: the conformational problem and the APS. The latter seems to be of special importance allowing for a quantitative (or semiquantitative) prediction of the activity that depends on the molecular structure (obviously, the activity quantitatively may depend also on some macroscopic parameters, e.g., lipophilicity).

A distinguishing feature of this method is that it reveals the Pha in terms of charges, bond orders, and interatomic distances, but not specific atoms. This means that the same activity may be manifested by different atoms and atomic groups which, within the given tolerance, have the same topology of charge (reactivity) distribution. In this way the same Pha describes the activity of compounds from quite different classes, sometimes with dramatically different atomic groups. For example, in the case of musk odorant activity [11], in some compounds a bromine atom occupies the same position in space and has approximately the same charge distribution and hence produces the same activity as the t-butyl group. For the RBA considered in this paper we have a similar case: the Pha of compound S18 from the set of new compounds to be screened has four carbon atoms which, within the tolerance limits, have the same charge distribution as the three carbons and one nitrogen in the Pha of many other active compounds.

A novel package of computer programs which implement this method was worked out with several modules taken from existing commercial programs. This package of programs works fast enough: each of the procedures described in Figure 3 is carried out within minutes of CPU time on a workstation.

Applied to the problem of rice blast activity, taken as an example, the EC method is shown to be highly efficient: all the active systems of the 96 relevant conformations of 85 compounds of the training set are described by the same Pha, and the activity is matched also quantitatively when the APS is taken into account, while the nonexperimental screening of another 39 compounds predicts their activity with an almost 100% probability qualitatively (yes, no) and with $r^2 = 0.66$ quantitatively.

Acknowledgements

This work has been supported in part by a grant from the Robert A. Welch Foundation and it is based in part upon work supported by Texas Advanced Research Program under Grant No. 003658–345.

References

1. Marshall, G.R., In Wolff, M.E. (Ed.) Burger's Medicinal Chemistry and Drug Discovery, Vol. 1, John Wiley and Sons, New York, NY, 1995, pp. 573–659.
2. Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications, ESCOM, Leiden, The Netherlands, 1993.
3. Höltje, H.D. and Folkers, G., In Mannhold, R., Kubinyi, H. and Timmerman, H. (Eds.) Molecular Modeling: Basic Principles and Applications, VCH Publishers, New York, NY, 1997, pp. 9–61.
4. Cohen, N.C. (Ed.) Guidebook on Molecular Modeling in Drug Design, Academic Press, New York, NY, 1996.
5. Oprea, T.I. and Waller, C.L., In Lipkowitz, K.B. and Boyd, D.B. (Eds.) Review of Computational Chemistry, Vol. 11, Wiley-VCH, New York, NY, 1997, pp. 127–182.
6. Martin, Y.C., In Martin, Y.C. and Willett, P. (Eds.) Designing Bioactive Molecules, American Chemical Society, Washington, DC, 1998, pp. 121–148.
7. Greco, G., Novellino, E. and Martin, Y.C., In Lipkowitz, K.B. and Boyd, D.B. (Eds.) Review of Computational Chemistry, Vol. 11, Wiley-VCH, New York, NY, 1997, pp. 183–240.
8. Jurs, P.C., Chou, J.T. and Yuan, M., In Olson, E.C. and Christofferson, R.E. (Eds.) Computer-Assisted Drug Design, ACS Symposium Series 112, American Chemical Society, Washington, DC, 1979, pp. 103–129.
9. Doucet, J.P. and Weber, J., Computer-Aided Molecular Design: Theory and Applications, Academic Press, San Diego, CA, 1996, pp. 364–404.
10. Bersuker, I.B. and Dimoglo, A.S., In Lipkowitz, K.B. and Boyd, D.B. (Eds.) Review of Computational Chemistry, Vol. 2, VCH, New York, NY, 1991, pp. 423–460.
11. Bersuker, I.B., Dimoglo, A.S., Gorbachov, M.Yu., Pesaro, M. and Vlad, P.F., New J. Chem., 15 (1991) 307.
12. Bersuker, I.B., Dimoglo, A.S. and Gorbachov, M.Yu., Nahrung-Food, 32 (1988) 461.
13. Bersuker, I.B., Dimoglo, A.S. and Gorbachov, M.Yu., Nahrung-Food, 33 (1989) 405.
14. Bersuker, I.B., Dimoglo, A.S. and Gorbachov, M.Yu., In Hadzi, D. and Jerman-Blazic, B. (Eds.) QSAR in Drug Design and Toxicology, Vol. 10, Elsevier, Amsterdam, 1987, pp. 43–48.
15. Bersuker, I.B., Dimoglo, A.S., Gorbachov, M.Yu., Vlad, P.F. and Koltsa, M.N., In Hadzi, D. and Jerman-Blazic, B. (Eds.) QSAR in Drug Design and Toxicology, Vol. 10, Elsevier, Amsterdam, 1987, pp. 340–342.
16. Bersuker, I.B., Dimoglo, A.S. and Gorbachov, M.Yu., Bioorgan. Khim., 13 (1987) 38.
17. Bersuker, I.B. and Dimoglo, A.S., In First World Congress on the Health Significance of Garlic and Garlic Constituents, Washington, DC, 1990.
18. Dimoglo, A.S., Bersuker, I.B., Popa, D.P. and Kuchkova, K.I., Theor. I Eksp. Khim., 5 (1989) 590.
19. Dimoglo, A.S., Vlad, P.F., Shvets, N.M., Koltsa, M.N., Guzel, Y., Saracoglu, M., Saripinar, E. and Patat, S., New. J. Chem., 19 (1995) 1217.
20. Dimoglo, A.S., Beda, A.A., Shvets, N.M., Gorbachov, M.Yu., Kheifits, L.A. and Aulchenko, I.S., New. J. Chem., 19 (1995) 149.
21. Guzel, Y., J. Mol. Struct. (Theochem), 366 (1996) 131.
22. Guzel, Y., Saripinar, E. and Yildirim, I., J. Mol. Struct. (Theochem), 418 (1997) 83.
23. Kansy, M., Ulmshneider, M. and van de Waterbeemd, H., In Sanz, I., Giraldo, J. and Manaut, F. (Eds.), QSAR and Modelling: Concepts, Computational Tools and Biological Applications, Prous Science Publishers, Barcelona, Spain, 1995, pp. 633–638.
24. Fukui, K., Theory of Orientation and Stereoselection, Springer-Verlag, Berlin, 1975, p. 134.
25. Klopman, G., In Klopman, G. (Ed.) Chemical Reactivity and Reaction Paths, Wiley, New York, NY, 1974, pp. 55–165.
26. Pearlman, R.S., In Kubinyi, H. (Ed.) 3D QSAR in Drug Design: Theory, Methods and Applications, ESCOM, Leiden, 1993, pp. 41–79.
27. SPARTAN, V. 5.0.3, Wavefunction, Inc., Irvine, CA, 1997.
28. MATLAB, V. 5.2.1, Mathworks, Inc., Natick, MA, 1998.
29. Dreikorn, B.A. and Thibault, T.D., Triazolo(4,3a)quinoxalines for the Control of Rice Blast, US Patent 4,008,322, 1977.
30. Dreikorn, B.A and Durst, G.L., In Baker, D.R., Fenyes, J.G. and Basarab, G.S. (Eds.) Synthesis and Chemistry of Agrochemicals IV, ACS Symposium Series 584, American Chemical Society, Washington, DC, 1995, pp. 354–374.