Eur. J. Med. Chem. 34 (1999) 687–699

Original article

The COREPA approach to lead generation: an application to ACE-inhibitors

Verginia Kamenska, Julian Ivanov, Ovanes Mekenyan*

Bourgas University "As. Zlatarov", Department Of Physical Chemistry, 8010 Bourgas, Bulgaria

(Received 17 August 1998; revised 19 January 1999; accepted 21 January 1999)

Abstract – The recently derived algorithm for identifying the COmmon REactivity PAttern (COREPA) of structurally diverse chemicals having similar biological behaviour was employed to recognize the structural requirements for high ACE inhibition. COREPA is based on an assessment of all energetically-reasonable conformations of the compounds under study. It is not dependent upon a predetermined and specified pharmocophore or an alignment of active conformers. The approach describes the common reactivity pattern in terms of global and local stereoelectronic parameter ranges. These ranges are associated with compounds having extreme (highest and lowest) biological activity through a comparison of conformer distributions for specific descriptors. The defined parameter ranges were combined into Boolean expressions providing flexible screening of chemicals by making use of a new chemical rule interpreter allowing a simultaneous search according to all available 2-D and 3-D information. The structural combinations, i.e., COREPAs, derived on a training set of ACE inhibitors were used for screening chemicals in an external test series for identifying those with high ACE inhibition activity. The obtained results showed that the structural requirements derived in the present work can be useful for screening chemicals from databases as potential lead ACE inhibitors. © 1999 Éditions scientifiques et médicales Elsevier SAS

quantitative structure-activity relationships / active analogues / stereoelectronic structure / conformational flexibility / ACE-inhibitors / lead generation / COREPA

1. Introduction

Angiotensin-converting enzyme (ACE) is a Zn-containing metallopeptidase which catalyses the hydrolysis of the terminal dipeptide from the angiotensin I (decapeptide) to produce the octapeptide angiotensin II. The latter appears to be one of the most potent vasoconstrictors. Although the primary amino acid sequence of ACE is known, [1–3] its 3-D structure is still undetermined. Inhibitors of ACE are widely prescribed to control essential hypertension. The structural requirements for high ACE activity within congeneric series of chemicals have been derived from intensive SAR studies [4–6]. The latter, combined with crystallographic data from the analogous enzyme, thermolysin, and its inhibitors, defined the minimal set of the active groups necessary for a chemical to elcite ACE inhibition: a C-terminal carboxyl group for ionic binding to the enzyme; a carbonyl oxygen which hydrogen bonds to some active site residue; some Zn-binding functional groups such as a carboxylate, hydroxamate, phosphonate, or thiolate. This structural information has been used to screen databases of structurally diverse classes of ACE-inhibitors to elicite the common three-dimensional geometry of the pharmacophoric sites consistent with their activity [7–9]. Still, the recognition of the 3-D spatial alocation of the pharmacophore responsible for ACE inhibition potency is a subject of intensive QSAR studies employing a variety of pattern recognition techniques.

A set of 28 inhibitors of angiotensin converting enzyme, selected by Mayer et al. [8], were modelled by the active analogue technique to elucidate the pharmacophore, i.e., the essential 3-D arrangement of functional groups that a molecule must possess to be recognized by the receptor under study. A strategy for systematic sampling of the conformational space has been used, thus, arriving at two plausible alternative active site hypotheses [10, 11], demonstrating that the multiple searches of the conformational space of flexible ligands is an effective way to achieve increasing reliability of the modelling results.

The CoMFA methodology (comparative molecular field analysis) is one of the most applied 3D-QSAR

*Correspondence and reprints

approaches to analyse ACE activity of chemicals. DePriest et al. [12] and Waller et al. [13] applied the CoMFA paradigm. They enlarged the initial set of 28 molecules with structures from additional classes of ACE inhibitors to resolve some uncertainties associated with recent hypotheses for the pharmacophoric site [8]. The CoMFA models of DePriest et al. [7], derived from potential fields, were found to be insufficient for accurately quantifying the enzyme-inhibitor interactions. The model and its predictive ability was improved by introducing a Zink indicator variable explicitly describing the Zink-ligand interaction. The derived 3D-QSAR models have been further used for predicting ACE activity of chemicals not included in the initial training set [12].

Pharmacophore search methods and receptor-site mapping, such as the active analogue technique and CoMFA, face significant challenges, which include the selection of appropriate conformations and obtaining an alignment of these structures. There are a number of good techniques for superimposing molecules [14–17], but developing a robust alignment model is not trivial. Typically, hundreds of alignments are explored to reach an optimal outcome, which, if not carefully evaluated and explained in the context of a presumed mechanism of interaction with the receptor may become susceptible to violation of the criteria of Topliss and Edwards [18] for causality in structure-activity models. Alignment errors can also lead to models that are incorrect or are poorly predictive. Further, the use of the lowest-energy conformers in methods such as CoMFA to assess similarity in pharmocophore search and receptor-mapping alogorithms seems inappropriate because, in complex systems such as biological tissues and fluids, chemicals are likely to exist in a variety of conformational states. In fact the lowest-energy, gas-phase, conformations might be the least likely to interact with the solvent or macromolecules [19], and solvation and binding interactions could more than compensate for energy differences among the conformers of a chemical [20–23].

In an attempt to address the issue of conformational flexibility, Prendergast et al. [24] reported an augmented active analogue technique to identify specific conformers of ligands acting as antagonists to angiotensin-II. All geometrically reasonable conformers were assessed; however, conformational energies were not evaluated and an energy minimization was not performed during the search. In a sense, this methodology can be viewed as an augmented version of the active analogue approach because it accounts for conformational flexibility and it eliminates the necessity of conformer alignment.

Recently, we developed a new pattern recognition approach, named COREPA (COmmon REactive PAttern) employed as a pharamacophore mapping technique. COREPA is a generalization of the active analogue approach and it provides an implicit exploration of a receptor's stereoelectronic shape. As opposed to existing pharmacophore search methods, the COREPA is an attempt to circumvent the problems related to the selection of appropriate conformations and obtaining appropriate template alignments [25, 26]. For each stereoelectronic parameter identified to be associated with the endpoint studied, all energetically-reasonable conformer distributions for the compounds from the training set are superimposed and the parameter ranges common for at least one conformer from all of the compounds identified. The collection of common stereoelectronic parameter ranges defines the common reactivity pattern. Due to the increasing pharmacological interest in potent vasoconstrictor agents, the present work aims to employ the COREPA method for the recognition of stereoelectronic requirements for a high ACE inhibition activity.

2. Materials and methods

2.1. The COREPA algorithm

The methodology to elucidate chemical similarity is based on the assumption that chemicals that elicit similar biological behaviour through a common mechanism of action should also possess similarities in stereoelectronic descriptors. Elucidation of this common reactivity pattern within a set of biologically-similar chemicals requires examination of the conformational flexibility of the compounds to evaluate molecular similarity in the context of the associated variability in specific stereoelectronic parameters.

The principal steps of the algorithm are presented and discussed in detail in Mekenyan et al. [25]. However, in order to follow the forthcoming, they are summarized as follows:

Step 1. Definition of the training set of chemicals. A defined subset of chemicals in the reaction series under investigation are selected as the training set. The training set can include either the most or least active chemicals, as defined by a user-imposed threshold of biological activity. This initial step establishes the extent of biological similarity among the chemicals from which stereoelectronic similarity will be discerned in the subsequent steps of the algorithm.

Step 2. Evaluation of stereoelectronic parameters hypothesized to be associated with biologically similar compounds. A restricted set of parameters, hypothesized to be associated with biological activity, are evaluated based on the normalized sum of dynamic similarity

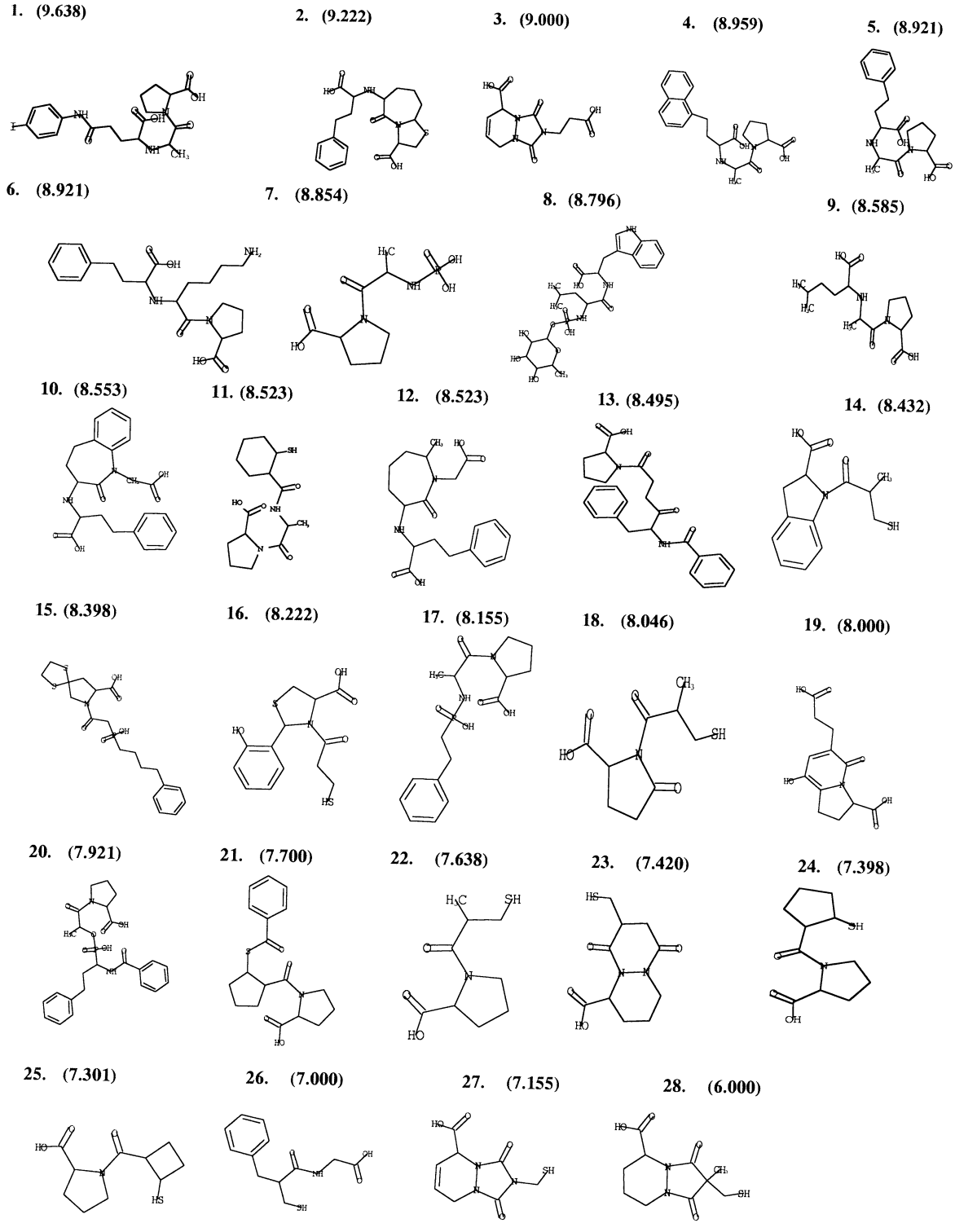

Figure 1. The training set of 28 AR ligands examined in this study (training set A).

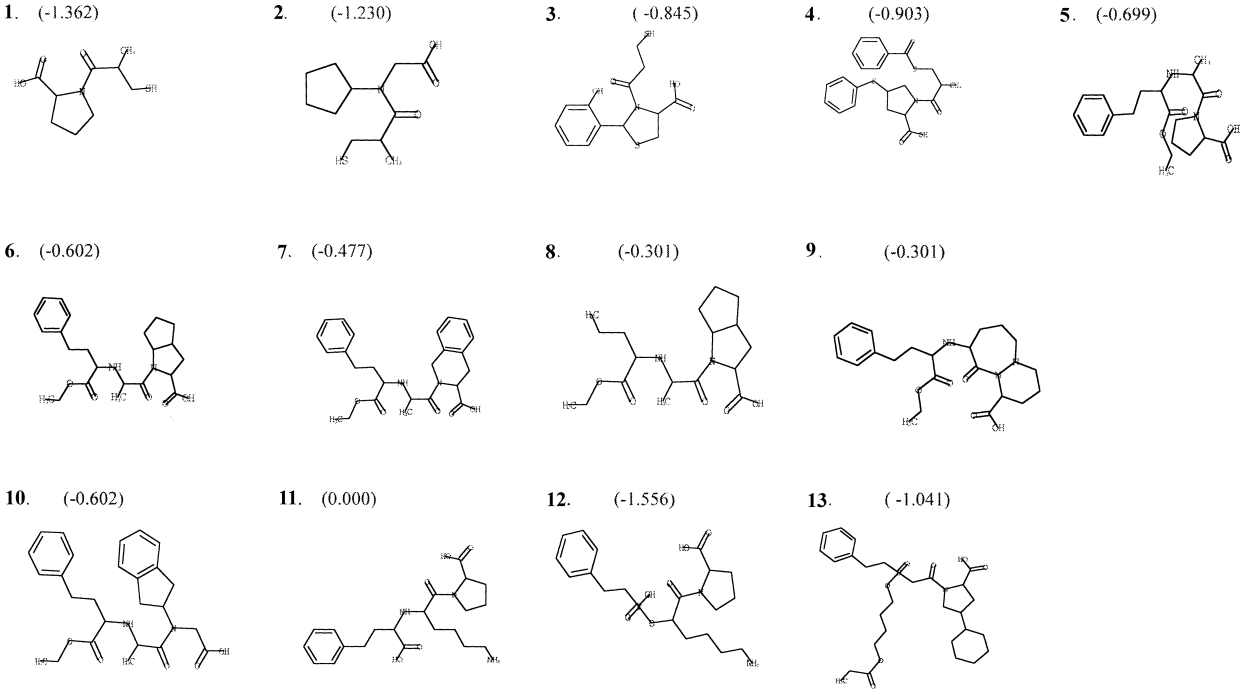

Figure 2. The test set of 13 captopril derivatives (test set B).

indices [27] between each pair of molecules in the training set. The Tanimoto coefficient, TAN, and the related Hodgkin-Richards similarity metric, 3D-TAN, cosine index, COSINE, and an index based on information theory, IF, were used as similarity indices, whereas, the Euclidean metric, DIST, was employed as a dissimilarity measure [27]. The stereoelectronic parameters that provide the maximal measure of similarity among the chemicals in the training set are assumed to be most closely associated with the activity under consideration and are used in the subsequent step of the algorithm.

Step 3. Recognition of the common reactivity pattern. For each stereoelectronic parameter identified in step 2, the conformer distributions of the chemicals from the training set are superimposed and the parameter ranges common for conformers from all of the chemicals identified. The collection of common stereoelectronic parameter ranges defines the common reactivity pattern.

2.2. ACE ligands and binding affinity

A training set of 28 AR ligands examined in this study (training set A) are those included in the learning set of Mayer et al. [8]. The structures are depicted in *figure 1*. The pIC_{50} values (the negative log of the inhibition concentration providing a standard biological response) are listed above the structures in *figure 1*. In the following analyses, sets of known ACE inhibitors such as captopril, pivalopril, rentiapril, zofenopril, enalapril, ramapril, quinapril, perindopril, cilazapril, delapril, lisinopril, SQ 29,852 and fosinopril [28–30] were used as an external validation data set (test set B) to evaluate discrimination abilities of the derived reactivity pattern. Their structures and associated activities (pIC_{50}) are presented in *figure 2*.

2.3. ACE ligand conformations

A primary aspect of the COREPA approach is to evaluate the conformational space for the chemicals under study using a number of conformers that can reasonably be assumed to represent the diversity of relevant stereoelectronic character for the biological process of interest. Sampling of conformational space was performed, aiming to provide up to 50 structurally most distinct conformers for each of the structures included in the training set. The conformational search was performed by making use of the 3DGEN algorithm that initiates from molecular topology and generates all conformers consistent with steric constraints and expert rules [31]. Specific geometric constraints were imposed during structure generation. Thus, torsion resolution (TR) around “saturated” (SP3-SP3) acyclic bonds was chosen

to be 120°, using an initial torsion angle of 60°, with respect to the plane of the preceding three atoms. Up to 7 rotational variables were chosen for the acyclic part of each chemical. For all chemicals, 1.5 Å was set as the distance between non-bonded atoms, while 1.2–2.0 Å was the range imposed for ring closure. After generation of the initial set of conformers, up to 50 of the structurally most diverse conformers were screened for each chemical based on structural dissimilarity of conformers as assessed by the sum of the distances between their non-hydrogen atoms. Each of the generated conformers was submitted to force field optimization based on a simple energy-like function, where only the electrostatic terms are omitted. Subsequently, the conformational degeneracy of the isomers was detected due to molecular symmetry and geometry convergence. Subsequent geometry optimization of the conformers was employed with MOPAC 7 [32], using the AM1 Hamiltonian with the key words "PRECISE" and "NOMM". Finally, these conformers were screened to eliminate those structures with ΔH_f^o values that were 20 kcal/mol or higher than that calculated for the conformer associated with the "absolute" energy minimum (see Results and discussion for an explanation of this threshold). The COREPA analyses were performed on this resulting data set.

2.4. Molecular descriptors

Stereoelectronic parameters were calculated with MOPAC 7 [32], augmented by a new computing module [33], that provides additional reactivity descriptors, using the AM1 all-valence electron semi-empirical Hamiltonian. The electronegativity (EN), dipol moment (μ), volume polarizability (Vol.P), energy of frontier orbitals (E_{HOMO} and E_{LUMO}), and electronic gap ($E_{HOMO\text{-}LUMO}$) were used as global electronic descriptors, whereas the atomic charges (q_i), frontier atomic charges (f_i^{HOMO} and f_i^{LUMO}), and acceptor superdelocalizability indices (S_i^E and S_i^N), as well as atomic self-polarizabilities (π_{ii}), were calculated as local electronic indices (i denotes a specific atom in a molecule). As it was already mentioned, when searching common patterns based upon local parameter distributions, the atomic reactivity indices were not restricted to specific rings in studied derivatives.

The delocalizabilities were calculated by using the following equations:

$$S_i^E = \sum_j^{occ} \sum_a^i C_{ja}\, C_{ja} /(E_{ref} - E_j) \qquad (1)$$

$$S_i^N = \sum_j^{vac} \sum_a^i C_{ja}\, C_{ja} /(E_{ref} - E_j) \qquad (2)$$

where E_j are the molecular orbital (MO) energies; C_j are the corresponding eigenvectors, and a pertains to the atomic orbitals of site i; E_{ref} is a non-zero, fixed, energy reference level, that was set equal to the electronic midgap level for benzene (–4.549 eV), i.e., $E_{ref} = 0.5(E_{HOMO} + E_{LUMO})_{benzene}$. Superdelocalizability indices were used to assess the ability of a reactant to form bonds through charge transfer, i.e., donor superdelocalizability indices are measures of the ability of molecules to donate electron density through orbital transfers, while acceptor superdelocalizability indices are measures of the ability of molecules to accept electron density.

Conformer structures were also assessed based on the steric descriptors GW (sum of geometric distances [34]), L_{max} (the greatest interatomic distance), d_{ij} (steric distance between atoms i and j) and planarity (the normalized sum of torsion angles in a molecule [25]). Finally, Vol.Polar., defined as a sum of atomic self-polarizabilities, and thus, the averaged ability of a compound to change electron density at its atoms during chemical interactions [35, 36], was selected as a physiochemical descriptor. Lower values of Vol. Polar. (Vol.P > 0) reflect higher charge localizations and more polarizable (less lipophilic) molecules [36]. These descriptors were selected because hydrophobicity, steric bulk and size constraints have been reported as important criteria in predicting and interpreting ligand binding interactions.

2.5. Database screening approach

The common reactivity pattern, i.e., the pharmacophore, defined in terms of 2-D structural fragments and associated 3-D stereochemical information (such as parity of stereocentres, distance between pharamacophoric sites, charge and delocalizability requirements) can be further used to screen databases for the presence of chemicals meeting these structural characteristics. A new chemical rule interpreter (CRI) has been developed for this purpose allowing simultaneous search according all available 2-D and 3-D information [37]. With that aim, substructure search techniques were combined with the range requirements for numeric descriptors as introduced in the extended SMILES language [37]. These entities, combining structural fragments with range requirement for numeric descriptors, are further called SD (structure-descriptor) screens. The CRI provides selective classification or screening of the chemicals from an OASIS database file (*.CMP) [33] according to a set of rules, given in terms of SD screens. A particular classification scheme is developed and described in a text file (*.RUL file). A separate rule of the scheme can be used to select from the *.CMP file those chemicals that satisfy it. The

application of the whole classification scheme results in a case-specific assignment of descriptors to the chemicals of the database. The *.RUL file is comprised of three sections, namely, Defines, Rules and Apply sections.

In the Defines section, the pertinent SD screens are described by means of extended SMILES language. It is very often practical to combine multiple SD screens in a single group that is used further as a single composite SD screen. For this purpose, the program supports a dynamic library of SD screens, partitioned in groups. Composite or simple SD screens have user defined labels–screen identifiers. Once defined, screen identifiers can be further used as parts of SMILES strings in order to define new SD screens. Screen identifiers are SD screens on their own since they can serve for SD search or for definition of new screen identifiers. When SD search is performed, each encountered screen identifier is recursively replaced by its actual contents, and the result is positive if at least one match is found.

Explicit SD screens and predefined screen identifiers constitute the basic logical variables of Boolean expressions (BL rules). The BL rules in use are described in the Rules section of the *.RUL file. An SD screen in the context of a BL rule takes logical values “true” or “false” if the SD screen is encountered in the current chemical or not. A BL rule combines a Boolean expression SD screen related with the logical operators “and”, “or” and “not”. The default priority of Boolean operations can be redefined by the use of brackets in unlimited nesting. As in the case of SD screens, separate BL rules can be assigned to different rule identifiers. An explicit BL rule or a rule identifier can be selected, and applied to a *.CMP file to produce in an output *.CMP file the subset of chemicals that match the rule.

In the Apply section, BL rules are employed in conditional *if*, *then*, *else* statements. The condition of a statement is a rule identifier or a Boolean expression of rule identifiers.

Examples are given with some of the structural requirements for high ACE-activity included in Defines and Rules sections of *.RUL file (*table IV)*. Comments are included to provide explanation of the screens.

Defines:

RX: O, S, N, P	Wild-card atom RX holds for any of O, S, N and P atoms.
Z1: O_O{5.6 < DISTANCE < 5.9}	Two O-atoms in a distance range of 5.6–5.9 [Å].
Z4: O{–0.39 < Q < –0.25}_N{–0.39 < Q < –0.25} {6.9 < DISTANCE < 7.4}	O- and N-atoms having charges in the range of –0.39 to –0.25 [a.u.] to be in a distance range of 6.9–7.4 [Å].
Z5: RX{–0.39 < Q < –0.25}_RX{–0.39 < Q < –0.25} {8.7 < DISTANCE < 9.4}	Two wild-card atoms RX (O, S, N and P) having charges in the range of –0.39 to –0.25[a.u.] to be in a distance range of 8.7–9.4 [Å].
Z7: O{0.22 < DONOR_DLC < 0.30}_O{0.22 < DONOR_DLC < 0.30}{8.7 < DISTANCE < 9.4}	Two O-atoms having donor delocalizabilties in the range of 0.22–0.30 $[(a.u.)^2/eV]$ to be in a distance range of 8.7–9.4 [Å].
Z9: RX{0.22 < DONOR_DLC < 0.30}_RX{0.22 < DONOR_DLC < 0.30}{8.7 < DISTANCE < 9.4}	Two wild-card atoms RX (O, S, N and P) having donor delocalizabilties in the range of 0.22–0.30$[(a.u.)^2/eV]$ to be in a distance range of 8.7–9.4 [Å].
Rules:	
r2: “Z7” and “Z8”	Simultaneous fulfillment of the requirements “Z7” and “Z8”.
r4: “Z3” or “Z7”	Fulfillment of any one of requirements “Z3” or “Z7”.
r5: (“Z3” or “Z4”) and (“Z7” or “Z8”)	Simultaneous fulfillment of the structural combination in brackets. The first of these combinations means fulfillment of any one of the requirements “Z3” or “Z4” whereas the second means fulfillment of any one of the requirements “Z7” or “Z8”.
r6: (“Z3” and “Z4”) or (“Z7” and “Z8”)	Fulfillment of any one of the requirements in brackets. The first one means simultaneous fulfillment of the requirements “Z3” and “Z4”, whereas the second one means simultaneous fulfillment of requirements “Z7” and “Z8”.
r15: (“Z5” or “Z6”) or (“Z9” or “Z10”) and not (“Z1” and “Z2”)	Fulfillment of any one of the first two requirements in brackets and at the same time non fulfillment of the requirement defined in the third combination. The first and third of the combinations means fulfillment of any one of the requirements “Z5” or “Z6”, and “Z9” or “Z10”, respectively, whereas the third one means simultaneous fulfillment of the requirements “Z1” and “Z2”.

Table I. ACE-inhibitors, observed binding affinities, number of conformer and parameter ranges for some significant stereoelectronic parameters.

Structure #	pIC_{50} obs.	Conformers	Vol.Polar [(a.u.)^2/eV]	E (HOMO) [eV]	E (LUMO) [eV]	E(HOMO-LUMO) [eV]	μ [D]	ΔH_f^o [kcal/mol]
1	9.638	32	1.511 to 1.546	–9.713 to –8.564	–0.606 to 0.069	8.514 to 9.364	2.113 to 13.257	–213.709 to –194.315
2	9.222	40	1.464 to 1.481	–9.224 to –8.812	–0.055 to 0.43	8.980 to 9.280	0.999 to 6.900	–190.112 to –170.246
3	9.000	46	0.899 to 0.912	–10.943 to –9.827	–0.915 to 0.09	9.532 to 10.132	2.030 to 8.560	–162.691 to –144.868
4	8.959	41	1.607 to 1.631	–8.853 to –8.474	–0.562 to –0.148	8.283 to 8.359	1.093 to 9.888	–177.810 to –158.097
5	8.921	38	1.374 to 1.398	–9.817 to –9.229	0.041 to 0.588	9.506 to 9.907	1.360 to 9.822	–201.710 to –181.716
6	8.921	37	1.648 to 1.664	–9.744 to –9.273	–0.166 to 0.513	9.488 to 9.895	1.296 to 10.996	–215.523 to –195.697
7	8.854	45	0.799 to 0.816	–10.643 to –9.753	–0.236 to 0.708	10.040 to 10.742	1.395 to 7.550	–359.037 to –339.913
8	8.796	33	1.910 to 1.923	–8.837 to –8.147	–0.253 to 0.468	8.566 to 8.633	1.441 to 8.794	–501.299 to –481.409
9	8.585	49	1.242 to 1.259	–10.389 to –9.712	–0.278 to 0.869	9.850 to 10.870	1.375 to 8.565	–245.194 to –225.583
10	8.553	41	1.589 to 1.610	–9.713 to –9.264	–0.587 to 0.045	8.892 to 9.345	2.394 to 9.494	–167.064 to –151.341
11	8.523	44	1.240 to 1.263	–9.426 to –8.643	0.203 to 0.769	8.964 to 9.794	1.141 to 7.011	–184.961 to –164.971
12	8.523	27	1.453 to 1.468	–9.743 to –9.293	–0.147 to 0.547	9.390 to 9.908	2.300 to 7.324	–206.092 to –188.253
13	8.495	32	1.705 to 1.732	–9.723 to –8.994	–0.547 to 0.344	8.957 to 9.694	1.796 to 10.779	–152.300 to –132.998
14	8.432	44	0.985 to 1.014	–9.378 to –8.718	–0.422 to 0.170	8.581 to 9.286	0.531 to 7.336	–92.215 to –73.736
15	8.398	46	1.540 to 1.563	–9.295 to –8.762	–0.529 to 0.039	8.668 to 8.837	3.402 to 9.146	–247.923 to –229.786
16	8.222	45	1.069 to 1.097	–9.338 to –8.629	–0.639 to 0.103	8.419 to 8.981	0.775 to 7.067	–128.738 to –109.576
17	8.155	42	1.278 to 1.294	–9.833 to –9.155	0.069 to 0.512	9.291 to 9.927	2.037 to 8.545	–271.344 to –252.291
18	8.046	50	0.774 to 0.801	–9.763 to –8.738	–0.246 to 0.321	8.818 to 9.798	0.881 to 6.955	–164.372 to –148.330
19	8.000	44	0.936 to 0.948	–9.393 to –8.735	–1.278 to –0.518	8.026 to 8.216	1.850 to 9.097	–237.338 to –219.398
20	7.921	35	1.812 to 1.835	–9.790 to –9.233	–0.516 to 0.125	8.737 to 9.753	1.415 to 7.814	–306.403 to –287.903
21	7.700	42	1.328 to 1.352	–9.671 to –9.028	–0.684 to –0.032	8.856 to 9.263	2.230 to 8.691	–140.147 to –120.234
22	7.638	47	0.778 to 0.799	–9.475 to –8.794	0.049 to 0.831	9.138 to 9.805	1.286 to 7.696	–135.315 to –118.169
23	7.420	49	0.881 to 0.906	–9.629 to –9.054	–0.318 to 0.146	8.940 to 9.500	0.558 to 6.044	–125.260 to –106.604
24	7.398	47	0.891 to 0.920	–9.627 to –8.713	–0.001 to 0.775	9.129 to 9.787	1.995 to 7.923	–139.030 to –119.150
25	7.301	47	0.824 to 0.847	–9.538 to –8.712	–0.010 to 0.844	9.351 to 9.851	0.852 to 7.961	–113.348 to –95.449
26	7.000	48	0.939 to 0.959	–9.526 to –8.771	–0.234 to 0.449	8.863 to 9.745	0.268 to 7.502	–112.093 to –97.395
27	7.155	36	0.773 to 0.791	–9.792 to –9.270	–0.727 to –0.228	8.764 to 9.264	2.582 to 6.031	–59.732 to –47.805
28	6.000	47	0.883 to 0.909	–9.755 to –9.022	–0.394 to 0.195	8.823 to 9.629	1.537 to 6.628	–113.843 to –93.921

3. Results and discussions

3.1. Conformational flexibility and electronic structure

Recently, the 20 kcal/mol threshold for $\Delta\Delta H_f^o$ was assumed to result in an energetically reasonable set of conformations, given the extent to which energy provided during ligand binding could facilitate conformational transformations [22, 23]. In this respect, the range of $\Delta\Delta H_f^o$ values for the conformers of the chemicals under study was selected to be less than 20 kcal/mol *(table I)*.

For a given compound, conformers within the specified range of $\Delta\Delta H_f^o$, exhibited chemically-significant variation in potentially relevant electronic descriptors, as summarized in *table I*. For example, the following parameter ranges are produced by the conformers of compound **7**: 0.944 eV for E_{LUMO} (from –0.236–0.708), 0.89 eV for E_{HOMO}, 0.702 eV for $E_{HOMO\text{-}LUMO}$, 0.017 (a.u.)^2/eV for Vol.P and 6.155 D for μ. Significant parameter variations were also observed for the other compounds. The observation that relatively small energy differences between conformers can be associated with significant variations in electronic structure (i.e., molecular electronic descriptors) highlights the necessity of including all energetically-reasonable conformers when defining common reactivity patterns. Within the employed approximation, these conformers were considered as equally probable because it is difficult to relate their "gas-phase" energies with the conformer preference in complex biological environments.

3.2. Application of the COREPA algorithm

According to the first step of the COREPA algorithm, two learning subsets were selected out of all 28 ACE inhibitors under investigation (training set A). The first one includes the most active chemicals having $pIC_{50} \geq 8.9$, whereas the second consists of the least potent (passive) ligands having $pIC_{50} \leq 7.42$. Further, the dynamic similarity was calculated for each pair of chemicals belonging to the subsets of active and non-active ligands, by using molecular descriptors presented in the preceding section. The normalized (over the pairs of compared chemicals) similarity indices, associated with molecular descriptors are presented in *table II*.

Table II. Normalized similarity measurements between conformer distributions of most active and least active ACE-inhibitors.

Descriptors	Most active ACE-inhibitors					Least active ACE-inhibitors				
	TAN	DIST	IF	3D-TAN	COSINE	TAN	DIST	IF	3D-TAN	COSINE
ACCEPT DLC	0.285	0.472	3.297	0.878	0.935	0.269	0.473	3.144	0.535	0.716
ACCEPT MLK	0.272	0.473	3.282	0.865	0.928	0.279	0.472	3.155	0.408	0.649
BOND ORDER	0.291	0.522	3.319	0.817	0.931	0.289	0.517	3.179	0.747	0.856
DONOR DLC	0.277	0.472	3.289	0.655	0.792	0.277	0.472	3.289	0.655	0.792
DONOR MLK	0.279	0.472	3.292	0.884	0.958	0.279	0.472	3.292	0.884	0.958
POLAR	0.289	0.472	3.302	0.333	0.756	0.289	0.472	3.302	0.333	0.756
POLAR MLK	0.292	0.472	3.305	0.600	0.866	0.292	0.472	3.305	0.600	0.866
POP–HOMO	0.283	0.472	3.293	0.700	0.841	0.283	0.472	3.293	0.700	0.841
POP–HOMO MLK	0.278	0.529	3.292	0.924	0.961	0.278	0.529	3.292	0.924	0.961
POP–LUMO	0.284	0.514	3.296	0.579	0.937	0.284	0.514	3.296	0.579	0.937
POP–LUMO MLK	0.283	0.492	3.297	0.652	0.894	0.283	0.492	3.297	0.652	0.894
Q	0.286	0.478	3.296	0.750	0.949	0.286	0.478	3.296	0.750	0.949
Q MLK	0.287	0.487	3.297	0.207	0.455	0.287	0.487	3.297	0.207	0.455
SPECIAL DISTANCE	0.261	0.475	4.471	0.125	0.223	0.261	0.475	4.471	0.125	0.223
E(HOMO)	0.061	0.969	0.505	0.154	0.272	0.061	0.969	0.505	0.154	0.272
ELECTRONEGATIVITY	0.149	0.958	1.225	0.000	0.000	0.149	0.958	1.225	0.000	0.000
ΔE(HOMO–LUMO)	0.096	0.663	0.673	0.000	0.000	0.096	0.663	0.673	0.000	0.000
GEOM. WIENER	0.012	0.750	0.244	0.000	0.000	0.012	0.750	0.244	0.000	0.000
VOLUME POLARIZAB.	0.000	0.510	0.000	0.000	0.000	0.000	0.510	0.000	0.000	0.000
CALC. HEAT FORM.	0.051	0.719	0.471	0.070	0.134	0.051	0.719	0.471	0.070	0.134
DIPOLE MOMENT	0.257	0.472	1.665	0.056	0.113	0.257	0.472	1.665	0.056	0.113
E(LUMO)	0.104	1.044	0.802	0.077	0.147	0.104	1.044	0.802	0.077	0.147

The data summarised in *table II* suggest that the most active and inactive ACE-ligands are highly similar according to charges (q_i), steric distances (d_{ij}), donor delocalizabilities (S_i^E, S_i^N) frontier charges on HOMO (f_i^{HOMO}) and LUMO (f_i^{LUMO}) and atom polarizabilities (π_i). In this respect, the next search of the common reactivity pattern was based on conformer distributions of chemicals across each of those local molecular descriptors. As such, we have also studied the interatomic distances d_{ij} between non-hydrogen atoms: N-X, O-X and X-X, where X stands for any non-hydrogen atom in the molecules.

To establish common reactivity patterns (step 3), conformer frequency distributions of compounds from each of the two training sets were subsequently examined across all local stereoelectronic descriptors, with results based on q_i and d_{ij}, illustrated in *figure 3* and *figure 4*, respectively. To study the effect of parameter distribution partitioning (i.e., size of parametric windows) on the obtained reactivity pattern, the number of parameter partitions was set at 20 and 40 for each descriptor.

The intersections between conformer distributions were identified separately for chemicals belonging to active and non-active subsets. These intersections are described in terms of parameter ranges occupied by at least one conformer from each chemical belonging to the learning subsets. The collection of those ranges represents the common reactivity pattern necessary for eliciting similar (high or low) biological effect (ACE-inhibition). Stereoelectronic parameters producing the most distinct reactivity patterns for active and passive chemicals and the respective ranges are presented in *table III*.

Even though conformer distributions from the most active and least active sets of compounds overlap, the subset of common partitions that contain conformers from each of the active compounds does not overlap with partitions containing conformers from each of the least active compounds. Thus, the common atomic charge pattern associated with the most active ligands *(figure 3a)* deviates significantly from that of the least active chemicals *(figure 3b)*. For the most active compounds, common partitions corresponding to oxygen and nitrogen atoms having charges from –0.25 to –0.39 [a.u.] *(figure 3a)* and donor delocalizabilities from 0.22–0.30 $[(a.u.)^2/eV]$ were observed. In difference, the charge and donor delocalizability patterns corresponding to O and N atoms for non-active inhibitors were shifted to more positive charge values (–0.26 to –0.36 [a.u.]; *figure 3b*) and lower delocalizabilities. Significant differences have also been found in charge and donor delocalizability patterns, due to the presence of S-atoms in non-active chemicals

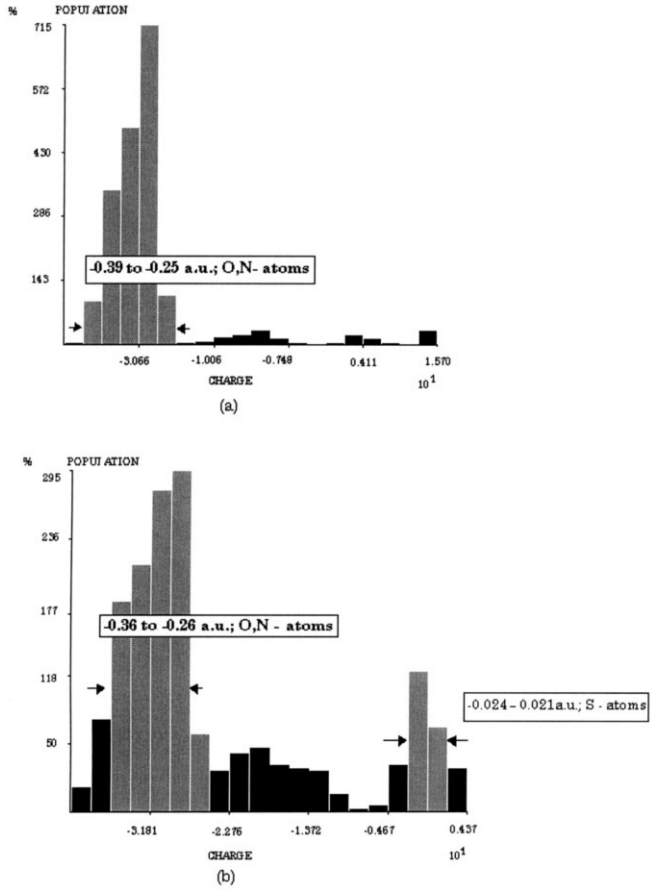

Figure 3. Common reactivity patterns for active **a**. and non-active **b.** ACE-inhibitors across the charges of hetero-atoms (oxygen, nitrogen and sulfur). Grey bars correspond to parameter ranges populated by at least conformers of all chemicals (activities, in *figure 3a*, and nonactives in *figure 3b*), whereas the black bars correspond to parameter ranges populated by conformers of some of the chemicals from active and non-active subsets.

(figure 3b). The maximum common distance range between oxygen atoms for active ligands was found to be from 8.7–9.4 Å (light coloured bars in *figure 4a*), whereas between O and N atoms was from 6.9–7.4 Å (not illustrated). As shown in *figure 4b*, no such common ranges have been established between the electronegative sites in non-active chemicals. As can be seen in *table III*, less significant differences in reactivity patterns of active and non-active chemicals (i.e., small differences in parametric ranges for active and non-active ACE inhibitors) were established for donor delocalizabilities of carbonyl carbons and the remaining C-skeleton.

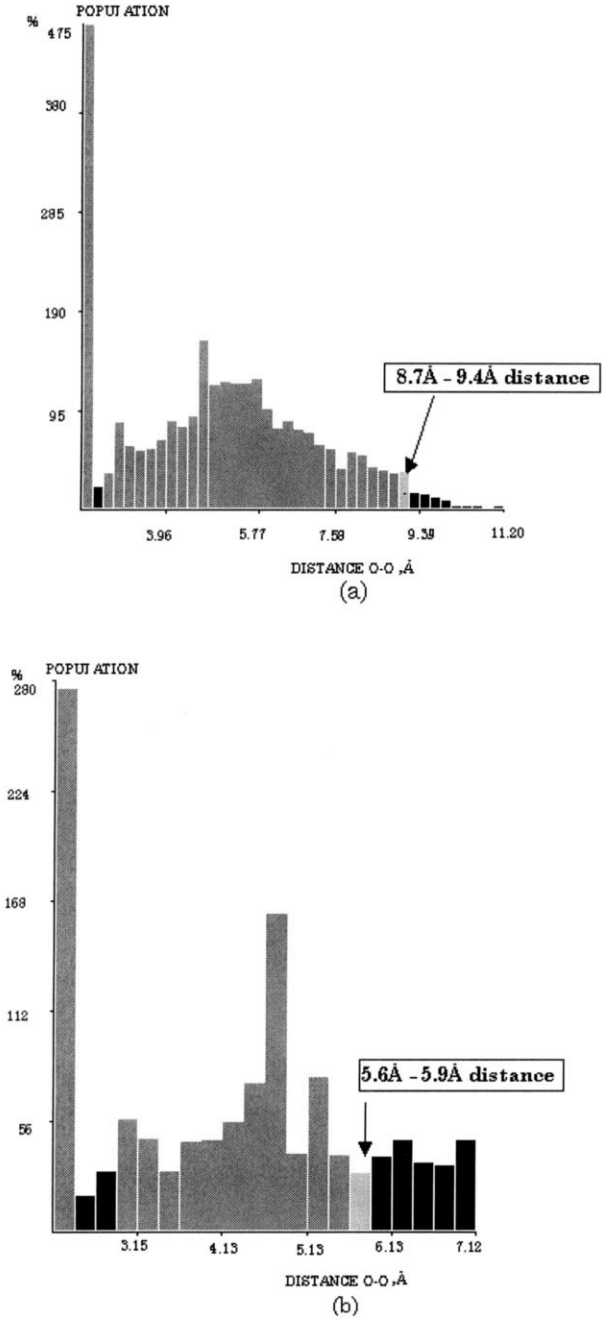

Figure 4. Common reactivity patterns for active **a.** and nonactive **b.** ACE-inhibitors across the distance between oxygen atoms. Grey bars correspond to parameter ranges populated by at least conformers of all chemicals (actives in *figure 4a* and non-actives in *figure 4b*), the black bars correspond to parameter ranges populated by conformers of some of the chemicals from active and non-active subsets. The light grey bar in *figure 4a* are associated with the largest common distance range for active chemicals in the training set.

Table III. Common reactivity patterns for most active and least active ACE-inhibitors from training set A.

		Most active ACE–inhibitors $pIC_{50} \geq 8.9$		Least active ACE-inhibitors $pIC_{50} \leq 7.42$	
Descriptor interatomic distances, [Å]	Atoms/ fragm.	Descriptor ranges		Descriptor ranges	
		Partitioning = 40	Partitioning = 20	Partitioning = 40	Partitioning = 20
	O-O O-N	8.94 to 9.17 6.96 to 7.19	8.74 to 9.39 6.95 to 7.42	5.75 to 5.88 3.71 to 3.80	5.63 to 5.88 3.71 to 3.89
Charge [a.u.]	X	–0.39 to –0.26 (O,N);	–0.39 to –0.25 (O,N);	–0.36 to –0.26 (O,N); –0.013 to 0.010 (S);	–0.36 to –0.26 (O,N); –0.024 to 0.021 (S);
Donor Delocalizability $[(a.u.)^2/eV]$	All	0.111 to 0.126 (C=(O)); 0.149 to 0.212 (C); 0.227 to 0.290 (O,N)	0.111 to 0.126 (C=(O)); 0.142 to 0.22 (C); 0.22 to 0.30 (O,N)	0.113 to 0.121 (C=(O)); 0.146 to 0.163 (C(N)); 0.171 to 0.188 (C(C)); 0.230 to 0.255 (N,OH); 0.263 to 0.280 (O=(C)); 0.381 to 0.410 (S)	0.113 to 0.129 (C=(O)); 0.146 to 0.163 (C-(N)); 0.163 to 0.196 (C-(C)); 0.230 to 0.247 (N,OH); 0.247 to 0.280 (O=(C)); 0.381 to 0.410 (S)

X - for non hydrogens and non-carbon atoms;
ALL - the partitions of the local parameter were analysed for all atoms in the molecule.

3.3. Validation of reactivity pattern

The complete lists of studied structural requirements and the Boolean expressions based on the above defined common parameter ranges are presented in *table IV*. More or less conservative screens were imposed by using flexibility of the Boolean expressions (i.e., the logical combination of the different structural rules by making use of the logical operators “and”, “or” and “not”).

The validation of the method was based on the applications of stereoelectronic screens on chemicals of the internal training set (A) (subset of which was used for deriving reactivity patterns) as well as to an external dataset (B) with well known ACE-inhibitors, used as pharmaceutical agents. The results are presented in *table V*.

The quality of screens is assessed by their ability to: (i) identify conformers of the chemicals from the learning

Table IV. The structural requirements and their Boolean combinations used for selecting highly active ACE inhibitors from test series under investigation.

Expressions of the defined structural requirement used for selecting high ACE inhibition activity from test series under investigation.	Boolean expression of the structural requirements
RX:O, S, N, P	r1: “Z3” and “Z4”
Z1:O_O{5.6 < DISTANCE < 5.9}	r2: “Z7” and “Z8”
Z2:O_N{3.7 < DISTANCE < 3.9}	r3: “Z3” or “Z4”
Z3:O{–0.39 < Q < –0.25}_O{–0.39 < Q < –0.25}{8.7 < DISTANCE < 9.4}	r4: “Z3” or “Z7”
Z4:O{–0.39 < Q < –0.25}_N{–0.39 < Q < –0.25}{6.9 < DISTANCE < 7.4}	r5: (“Z3” or “Z4”) and (“Z7” or “Z8”)
Z5:RX{–0.39 < Q < –0.25}_RX{–0.39 < Q < –0.25}{8.7 < DISTANCE < 9.4}	r6: (“Z3” and “Z4”) or (“Z7” and “Z8”)
Z6:RX{–0.39 < Q < –0.25}_RX{–0.39 < Q < –0.25}{6.9 < DISTANCE < 7.4}	r7: (“Z3” or “Z4”) or (“Z7” or “Z8”)
Z7:O{0.22 < DONOR_DLC < 0.30}_O{0.22 < DONOR_DLC < 0.30}{8.7 < DISTANCE<9.4}	r8: (“Z5” or “Z9”) r9: (“Z5” and “Z9”)
Z8:O{0.22 < DONOR_DLC < 0.30}_N{0.22 < DONOR_DLC < 0.30}{6.9 < DISTANCE < 7.4}	r10: (“Z5” or “Z6”) r11: (“Z5” and “Z6”)
Z9:RX{0.22 < DONOR_DLC < 0.30}_RX{0.22 < DONOR_DLC < 0.30}{8.7 < DISTANCE < 9.4}	r12: (“Z5” or “Z6”) or (“Z9” or “Z10”) r13: (“Z5” and “Z6”) or (“Z9” and “Z10”)
Z10:RX{0.22 < DONOR_DLC < 0.30}_RX{0.22 < DONOR_DLC < 0.30}{6.9 < DISTANCE < 7.4}	r14: (“Z5” or “Z6”) and (“Z9” or “Z10”) r15: (“Z5” or “Z6”) or (“Z9” or “Z10”) and not (“Z1” and “Z2”) r16: (“Z5” or “Z6”) or (“Z9” or “Z10”) and not (“Z1” or “Z2”)

Table V. Screened ACE-inhibitors from test series according to the employed structural requirements, as described by the Boolean expressions listed in *table IV*.

Rules	Training set A			Test set B		
#[a]	Number of selected chemicals	Number of selected conformers	Numbering of selected chemicals[b]	Number of selected chemicals	Number of selected conformers	Numbering of selected chemicals[c]
r1	10	47	**1, 2, 4, 5, 6, 8, 10, 13, 20, 21**	6	42	**5–7, 9–11**
r2	12	75	**1, 2, 4–8, 10, 12, 19–21**	1	9	**12**
r3	14	222	**1–6, 8–10, 12, 13, 19–21**	9	350	**4–7, 9–13**
r4	16	186	**1–10, 12, 13, 17, 19–21**	9	248	**4–7, 9–13**
r5	14	216	**1–6, 8–10, 12, 13, 19–21**	2	60	**12, 13**
r6	12	65	**1, 2, 4, 5, 7, 8, 10, 13, 19–21**	7	51	**5–7, 9–12**
r7	16	250	**1–10, 12, 13, 17, 19–21**	9	374	**4–7, 9–13**
r8	16	222	**1–11, 13, 17, 19–21**	9	317	**4–7, 9–13**
r9	14	188	**1–6, 8–11, 13, 19–21**	2	35	**12, 13**
r10	19	428	**1–6, 8–13, 16, 19–21, 23, 27, 28**	11	596	**3–13**
r11	14	127	**1–6, 8–11, 13, 19–21**	9	134	**4–7, 9–13**
r12	22	495	**1–10, 12, 13, 15–17, 19–21, 23, 27, 28**	11	624	**3–13**
r13	15	156	**1–11, 13, 19–21**	9	162	**4–7, 9–13**
r14	19	416	**1–6, 8–11, 13, 16, 19–21, 23, 27, 28**	2	79	**12, 13**
r15	22	484	**1–13, 15–17, 19–21, 23, 27, 28**	10	619	**3–12**
r16	22	350	**1–13, 15–17, 19–21, 23, 27, 28**	11	265	**3–13**

[b]The numbering of chemicals corresponds to those in *figure 1*
[c]The numbering of chemicals corresponds to those in *figure 2*

subset; (ii) "catch" some chemicals with moderate activity (not used for deriving the pattern); (iii) eliminate all conformers of the chemicals with low activity; and (iv) identify as active, chemicals from external databases, experimentally shown to elicit high activity (i.e., pharmaceutical agents). Beside selection of conformers of most active chemicals from training set A (having $pIC_{50} \geq 8.9$) used for deriving the common reactivity pattern, almost all of the structural rules screened also conformers of ACE ligands with lower activity, i.e., having $pIC_{50} < 8.9$. Still, these are active and medium active inhibitors having $7.7 \leq pIC_{50} \geq 8.9$ (up to chemical **21**, *table I*). Only few of the structural rules, classified as active, the inhibitors having $pIC_{50} < 7.7$. These are the "least restrictive" rules #10, 12, 14–16 with high "concentration" of "or" logical operator, i.e., providing very low screen conservativeness. These rules, screened as active, chemicals having $pIC_{50} \geq 6.00$. It is very difficult to define precisely the pIC_{50} threshold according to which a chemical could be classified as active. Usually, these thresholds are user defined. All chemicals included in training set A are ACE inhibitors and the differences in their activity are only quantitative, not qualitative. In this respect, the obtained results are reasonable: the more restrictive screens, selected as active, the inhibitors with higher pIC_{50} values, whereas, the less restrictive screens select as active all inhibitors under investigation.

The reactivity pattern derived by the most active ACE inhibitors from training set A was used for screening of the chemicals from the test set B. As it was already mentioned, these captopril derivatives are well known ACE inhibitors with different activity, most of them used as pharmaceutical agents. It was hypothesized that the derived reactivity pattern should, screen as active, most of the chemicals from this test set. As can be seen from *table V*, Rules # 3, 4, 7, 8, 11 and 13 selected 70% of the inhibitors from test set B, whereas Rules # 10, 12 and 16 selected 85% of those chemicals. The fact that the last four rules were capable of screening a large percentage of the known ACE-inhibitors from the external dataset is an indication that they could be used for screening of chemicals from databases as potential lead ACE inhibitors. A practical advise is to perform screenings with each of the four individual rule files and then to look for the intersection of predicted subsets of chemicals. The latter are assumed to be the most potent ligands because they meet simultaneously stereoelectronic requirements included in all rule files.

4. Summary and conclusions

The present search for the structural requirements for high ACE-inhibition is based on the COREPA algorithm. This last one is a generalization of the active analogue approach circumventing the need to align conformers of active molecules, while it explicitly addresses variation in conformational flexibility in the context of varying biological activity. The distributions of all energetically reasonable conformers of chemicals from the training subsets across specific molecular descriptors found to be relevant to activity under study are analysed. Thus, instead of template alignment based on single conformer representations of chemicals, their conformer distributions are naturally aligned (ordered) across specific molecular descriptors. The distribution intersections populated by conformers of each of the biologically similar chemicals (actives and non-actives) form the common reactivity pattern. Due to the population character of conformer distributions in the COREPA approach it is difficult to recognize the active conformers of chemicals in classical chemical terms as it is done in other "dynamic" techniques [20–22, 26, 38, 39].

The common reactivity pattern, in terms of charge and distance ranges between electronegative atoms required for high ACE-inhibition, was derived. It was based on the subset of the most active chemicals from the training set A. Then, it was validated by screening all chemicals from the same training set, as well as an external set of active analogues known as pharmaceutical agents (training set B). The various logical combinations of the stereoelectronic rules provided the flexibility of the screening process. The most restrictive screens selected as active the inhibitors with pIC_{50} values higher or equal to the activity threshold used for deriving reactivity patterns. The less restrictive screens in addition, selected as active, inhibitors with moderate activity. Six of the employed screening rules selected as active 70% of the active inhibitors from test set B, whereas four of the rules selected 85% of those chemicals. The large percentage of the screened ACE-inhibitors from the test set by these four stereoelectronic screens is indicative that the associated reactivity pattern could be used for screening of large databases of 3-D structures for the search of potentially active ACE-inhibitors.

Individual screenings were performed by employing each of the stereoelectronic screens on conformers of the chemicals from the test set. Chemicals meeting multiple stereoelectronic requirements, i.e., lying on the intersection of the subsets predicted by single screens, are considered as the most likely candidates for active ligands. Such studies are presently performed by our laboratory for the needs of the Chemicals and Pharmaceutical Research Institute, Sofia, aiming to design original pharmaceutical agents. In large databases, however, chemicals are represented by single conformers because the conformational search for each compound is computationally impractical. In these situations, a less restrictive screening strategy is employed initially that assesses a single conformer per chemical and which is designed to minimize the percentage of false negatives (i.e., compounds incorrectly predicted to be non-active). The "Tweak" technique is also used as a pre-screen alternative, in which the rotatable bonds of the structures are adjusted to produce a conformation which matches as closely as possible a given 3D requirement. Then, in a second stage, more refined screens should be employed on a smaller set of conformationally multiplied chemicals, already passed the pre-screen.

Acknowledgements

This research was financially supported by the Chemicals and Pharmaceutical Research Institute, Sofia (grant # RD-09-173/20.12.1995). The authors also thanks Dr Karabunarliev for developing the Chemical Rule Interpreter software as well as Drs Matey Vitev, Neno Dimov and Kiril Ninov for valuable discussions.

References

[1] Bunning P., Kleemann S.G., Riordan J.F., Biochemistry 29 (1990) 10488–10492.

[2] Chen Y.N.P., Riodan J.F., Biochemistry 29 (1990) 10493–10498.

[3] Soubrier F., Alhenc-Gelas F., Hubert C., Allegrini J., John M., Tregear G., Corvol P., Proc. Natl. Acad. Sci. USA 85 (1988) 9386–9390.

[4] Petrillo W.W., Trippodo N.C., De Forrest J.M., in: Robertson D.W. (Ed.), Annual Reports in Medicinal Chemistry Vol 25, Academic Press, New York, 1989, pp. 51–60.

[5] Hangauer D.G., in: Perun T.J., Propst C.L. (Eds.), Computer-Aided Drug Design: Methods and Applications, Marcel Dekker, New York, 1989, pp. 253–295.

[6] Wyvratt M.J., Patchett A.A., Med. Res. Rev. 5 (1985) 483–531.

[7] Depriest S.A., Shands R.F.B., Dammkeohler R.A., Marshall G.R., in: Silipo C., Vittoria A. (Eds.), QSAR Rational Approaches to the Design of Bioactive Compounds Vol 16, Elsevier Science, Amsterdam, 1991, pp. 405–414.

[8] Mayer D., Naylor C.B., Motoc I., Marshall J.R., J. Comput. -Aided Mol. Des 1 (1987) 3–16.

[9] Andrews P.R., Carson J.M., Caselli A., Spark M.J., Woods R., J. Med. Chem. 28 (1985) 393–399.

[10] Dammkoehler R.A., Karasek S.F., Shands E.B.F., Marshall G.R., J. Comput. -Aided Mol. Des. 3 (1989) 3.

[11] Dammkoehler R.A., Karasek S.F., Shands E.B.F., Marshall G.R., J. Comput. -Aided Mol. Des. 9 (1995) 491–499.

[12] Depriest S.A., Mayer D., Naylor C.B., Marshall G.R., J. Am. Chem. Soc. 13 (1993) 115 5372–5384.

[13] Waller C.L., Shands E.B.F., Dammkoehler R.A., Marshall G.R., J. Comput. -Aided Mol. Des. 11 (1997) 121.

[14] Kearsely S.K., Smith G.M., Tetrahedron Comput. Methodol. 3 (1990) 615–633.

[15] Blaney J.M., Dixon J.S., Perspect. Drug Discov. Des. 1 (1993) 301–319.

[16] Perkins T.D., Deam P.M., J. Comput. -Aided Mol. Des. 7 (1993) 155–172.

[17] Martain Y.C., Bures M.G., Danahar E.A., Delazzar J., Lico I., Pavlik P.A., J. Comput. -Aided Mol. Des. 7 (1993) 83–102.

[18] Topliss J.G., Edwards R.P., J. Med. Chem. 22 (1979) 1238–1244.

[19] Eliel E.L., in: Warr W.A. (Eds.), Chemistry in Three Dimensions in Chemical Structures, Vol. 1, Springer, Berlin, 1993, pp. 1–8.

[20] Mekenyan O.G., Ivanov J.M., Veith G.D., Bradbury S.P., Quant. Struct. -Act. Relat. 13 (1994) 302–307.

[21] Mekenyan O.G., Veith G.D., Call D.J., Ankley G.T., Environ. Health Perspect. 104 (1996) 1302–1309.

[22] Bradbury S.P., Mekenyan O.G., Ankley G.T., Environ. Toxicol. Chem. 15 (1996) 1945–1954.

[23] Wiese T., Brooks S.C., J. Steroid. Biochem. Molec. Biol. 50 (1994) 61–72.

[24] Prendergast K., Adams K., Greenlee W.J., Nachbar R.B., Patchett A.A., Underwood D.J., J. Comput. -Aided Mol. Des. 8 (1994) 491–512.

[25] Mekenyan O.G., Ivanov J.M., Karabunarliev S.H., Bradbury S.P., Ankley G.T., Karcher W., Environ. Sci. Technol. 31 (1997) 3702–3711.

[26] Bradbury S.H., Mekenyan O.G., Ankley G.T., Environ. Toxicol. Chem. 17 (1998) 15–25.

[27] Mekenyan O.G., Ivanov J.M., Karabunarliev S.H., Hansen B., Ankley G.T., Bradbury S.P., in: Chen F., Schüürmann G. (Eds.), Quantitative Structure Activity Relationships in Environmental Sciences VII, SETAC Press, Pensacola FL, USA, 1998, pp. 39.

[28] Cambell W.B., Brooks S.N., Pettinger W.A., Science 184 (1974) 994–996.

[29] Ondetti M.A., Rubin B., Cushman D.W., Science 196 (1977) 441–444.

[30] Natoff I.L., Redshaw S., Drugs Future 12 (1987) 475–483.

[31] Ivanov J.M., Karabunarliev S.H., Mekenyan O.G., J. Chem. Inf. Comput. Sci. 34 (1994) 234–243.

[32] Stewart J.J.P., MOPAC: A general molecular orbital packages; Version 7. 0. Software. Quantum Chemistry Program Exchange no. 455, University of Indiana, Bloomington, IN, USA, (1995).

[33] Mekenyan O.G., Karabunarliev S.H., Ivanov J.M., Dimitrov D.N., Comput. Chem. 18 (1994) 173–187.

[34] Mekenyan O.G., Peitchev D., Bonchev D., Trinajstic N., Bangov I., Arzneim. Forsch. 36 (1986) 176.

[35] Schüürmann G., Quant. Struct. -Act. Relat. 59 (1990) 326–333.

[36] Lewis D.F.V., J. Comput. Chem. 10 (1989) 145–151.

[37] Karabunarliev S.H., Nikolov N., Mekenyan O.G., submitted.

[38] Mekenyan O.G., Schultz T.W., Veith G.D., Kamenska V.B., J. Appl. Toxicol. 16 (1996) 355.

[39] Veith G.D., Mekenyan O.G., Ankley G.T., Call D.J., Environ. Sci. Technol. 29 (1995) 126.